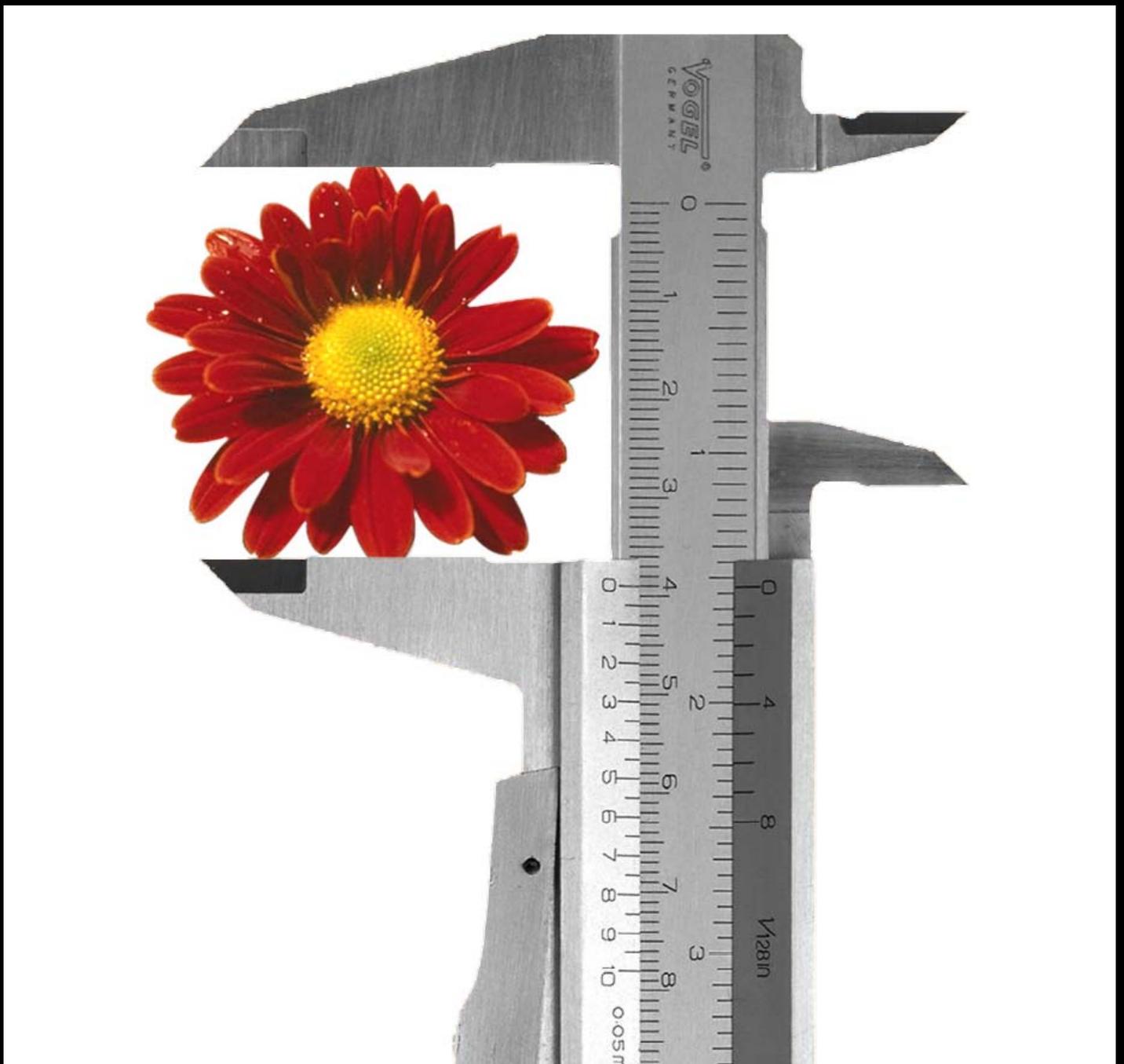


EVALUACIÓN DEL USO DE MEDIDAS DE CALIDAD DE VIDA EN LA PRÁCTICA CLÍNICA APLICACIÓN A PACIENTES CON CATARATAS

José M^a Valderas Martínez

Tesis Doctoral 2005



**EVALUACIÓN DEL USO DE MEDIDAS
DE CALIDAD DE VIDA EN LA PRÁCTICA CLÍNICA.
APLICACIÓN A PACIENTES CON CATARATAS.**

TESIS DOCTORAL

José M^a Valderas Martínez, 2005

Dipòsit legal: B.43967-2005
ISBN: 84-689-4199-9

Departament de Ciències Experimentals i de la Salut
Programa de Doctorat en Ciències de la salut i de la Vida
Universitat Pompeu Fabra (UPF)

EVALUACIÓN DEL USO DE MEDIDAS
DE CALIDAD DE VIDA EN LA PRÁCTICA CLÍNICA.
APLICACIÓN A PACIENTES CON CATARATAS.

Memoria presentada por Jose M^a Valderas Martínez para optar al título de Doctor,
elaborada a partir del trabajo realizado bajo la dirección del Dr. Jordi Alonso Caballero,
en el Institut Municipal d' Investigació Mèdica de Barcelona.
Programa de Doctorat de la Universitat Pompeu Fabra, bienio 1998-2000.

«Così belli gli suoi occhi, se non fosse per la cataratta»

Storia della mia vita. G. Casanova

A Carmina

ÍNDICE

PRESENTACIÓN	1
ACRÓNIMOS	5
RESUMEN	7
INTRODUCCIÓN.....	11
La salud y la calidad de vida relacionada con la salud (CVRS)	11
El uso de los instrumentos de medida de la CVRS en la práctica clínica	13
La interpretación de las medidas de CVRS	16
Las cataratas y el índice de función visual VF-14	18
JUSTIFICACIÓN	23
OBJETIVOS	25
MÉTODOS Y RESULTADOS	27
Síntesis de métodos	27
Artículo 1: Revisión sistemática y meta-análisis	29
Artículo 2: Herramientas de apoyo a la interpretación	47
Artículo 3: Uso sistemático del índice VF-14	59
Síntesis de resultados	73
DISCUSIÓN	79
De la revisión sistemática y meta-análisis	79
De las ayudas a la interpretación	84
Del uso sistemático del índice VF-14	87
CONCLUSIONES	91
BIBLIOGRAFÍA	93
ANEXOS	103
Anexo 1. Glosario	105
Anexo 2 Cuestionario 1 (<i>estudio de evaluación</i>)	113
Anexo 3 Cuestionario 2 (<i>estudio de evaluación</i>)	121
Anexo 4 Formulario clínico (<i>estudio de evaluación</i>)	135

PRESENTACIÓN DE LA TESIS

Esta tesis doctoral se presenta como compendio de publicaciones originales, según la normativa aprobada por la Comissió de Direcció de Doctorat del Departament de Ciències Experimentals i de la Salut de la Universitat Pompeu Fabra. Este trabajo incluye, por este orden, un resumen estructurado del mismo, un capítulo de introducción y objetivos, un capítulo de métodos y resultados en el que se insertan los artículos originales publicados o aceptados para su publicación, un capítulo de discusión y un apartado con conclusiones y recomendaciones. Los tres artículos que se presentan son la producción científica de una línea de investigación centrada en el estudio de los efectos de la incorporación en la práctica clínica de las mediciones de la salud percibida. Por su contribución a la valoración de los resultados de la primera publicación, se incluye también un intercambio epistolar.

Espallargues M, **Valderas JM**, Alonso J. "Provision of feedback on perceived health status to health care professionals: a systematic review of its impact." *Medical Care* 2000;38(2):175-8

Valderas JM, Espallargues M, Alonso J." Provision of feedback on perceived health status to health care professionals". *Medical Care* 2000; 38(8): 878

Valderas JM, Alonso J, Prieto L, Espallargues M, Castells FX. "Content-based interpretation aids for health-related quality of life measures in clinical practice. An example for the visual function index VF-14". *Quality of Life Research* 2004 13: 35–44.

Valderas JM, Rue M, Guyatt G, Alonso J. "The impact of the VF-14 Index, a perceived visual function measure, in the routine management of cataract patients". *Quality of Life Research* 2005 (en prensa).

Presentaciones previas

Algunos resultados parciales pertenecientes a este trabajo se han presentado en los siguientes congresos y reuniones científicas:

- Annual Meeting of the International Society for Quality of Life Research: Baltimore (EEUU), 1998; Barcelona (España), 1999; Vancouver (Canadá), 2000; Amsterdam (Países Bajos), 2001
- International Congress of Psychology. Estocolmo (Suecia), 2000
- Reunión de la Sociedad Española de Epidemiología: Sevilla, 1998; Santiago de Compostela, 1999; Barcelona, 2000

Financiación

Estas investigaciones han contado con la siguiente financiación: Fondo de Investigaciones Sanitarias (FIS Expte. 97/1135), Instituto de Salud Carlos III (Expte. 97/4363, 98/4471 y CM0300118), Consell Interdepartamental de Recerca i Innovació Tecnològica de la Generalitat de Catalunya (1997/SRG 00359, 2001/SGR 00405), Beca Investigación predoctoral “Evaluación controlada del uso de la salud percibida en práctica clínica” Unidad de Investigación en Servicios Sanitarios, Institut Municipal d’ Investigació Mèdica (IMAS-IMIM) 1998.

Premios

El artículo “Content-based approach for facilitating the interpretation of health-related quality of life measures in clinical practice. A proposal for the visual function index VF-14”. Valderas JM, et al. *Quality of Life Research* 2004 13: 35–44 ha recibido el premio a la mejor publicación de Premis Sanitat Baix Lobregat 2004, convocados de forma conjunta por el Col·legi Oficial de Metges de Barcelona (COMB) y la Acèdemia de Ciències Mèdiques de Catalunya i Balears (ACMCB).

Agradecimientos

Este trabajo de tesis doctoral debe su inicio, casi inconsciente, su continuación, laboriosa, en ocasiones extenuante, y su culminación, feliz, a muchas personas. La inmensa mayoría seguirán anónimas en su reconocimiento: los pacientes y médicos participantes en los estudios, el equipo de campo y las encuestadoras, los compañeros de trabajo dentro y fuera del IMIM y muchos más, que no olvido. A otros, que son muchos menos, sí debo poner nombre: Jordi Alonso, mi director de tesis, Marisa y José María, mis padres, y, sobre todo, Carmina. Sé con certeza que sin ellos no hubiera podido.

ACRÓNIMOS

Acrónimos utilizados en el texto y los artículos (*entre paréntesis la correspondencia del acrónimo en inglés cuando éste es el idioma al que corresponde el mismo*)

AV: Agudeza Visual

CI (*Confidence Interval*): Intervalo de Confianza

CTT (*Classical Test Theory*): Teoría Clásica de los Tests

CVRS: Calidad de Vida Relacionada con la Salud

HRQL (*Health Related Quality of Life*): Calidad de Vida Relacionada con la Salud

ICC (*Intraclass Correlation Coefficient*): Coeficiente de Correlación Intraclase

IRT (*Item Response Theory*): Teoría de Respuesta al Ítem

IQR (*Interquartile Range*): Rango Intercuartil

OR (*Odds Ratio*): Razón de Odds

RCT (*Randomized Clinical Trial*): Ensayo Clínico Aleatorizado

SD (*Standard Deviation*): Desviación Estándar

TRI: Teoría de Respuesta al Ítem

VA (*Visual Acuity*): Agudeza Visual

VF14: Índice de Función Visual VF-14

RESUMEN

Antecedentes: En los últimos años, la calidad de vida relacionada con la salud (CVRS) ha experimentado un desarrollo notable como variable de resultado en la investigación clínica. Este impulso no se ha traducido en una difusión análoga en la práctica médica, donde su uso es muy limitado.

Objetivos: Evaluar el impacto en la atención médica del uso sistemático de instrumentos de medida de la CVRS y del suministro de la información a los profesionales de la salud, en general y específicamente para el Índice de Función Visual VF-14, una medida de función visual percibida para pacientes con cataratas. Para la consecución de este último objetivo, se ha desarrollado un método para la construcción de herramientas de ayuda a la interpretación de las medidas de CVRS que se ha aplicado al Índice VF-14.

Métodos: Se realizó una revisión sistemática de ensayos clínicos que evaluaran intervenciones consistentes en el suministro de información (*feed-back*) sobre CVRS a profesionales sanitarios en el período 1966-1997 y se realizó una síntesis cualitativa y cuantitativa (meta-análisis) de los datos obtenidos. Posteriormente se desarrolló un método para la interpretación de las puntuaciones basado en Teoría de Respuesta al Ítem (Análisis de Rasch). Se aplicó el método al Índice VF-14 para la construcción de herramientas específicas, empleando datos pertenecientes a un estudio realizado en 4 hospitales y 6 centros de cirugía de cataratas ambulatoria con 192 pacientes a los que se intervino del primer ojo. A partir de la actividad más difícil que un paciente dado podía realizar sin problemas se desarrolló la herramienta ‘termómetro’; para la construcción de la herramienta de ‘escenarios clínicos’ se identificaron todos los conglomerados significativamente distintos de ítems y luego se estimó la dificultad esperada (respuesta) para realizar una actividad representativa de cada agregado. Para la evaluación del impacto del uso sistemático del Índice VF-14, se diseñó un estudio antes-después en el que se midió la función visual de 833 pacientes con cataratas que acudieron de forma consecutiva a las consultas de 19 oftalmólogos en centros hospitalarios y de atención primaria, públicos y privados de cinco ciudades catalanas. Este estudio antes-después

incluyó a) una sesión educativa y b) la administración posterior de la puntuación VF-14 de todos los pacientes a sus correspondientes oftalmólogos, asistida por las herramientas de ayuda a la interpretación desarrolladas. Se construyeron modelos lineales y logísticos de efectos mixtos para evaluar el impacto en el proceso de la atención sanitaria (correlación entre las valoraciones de la función visual proporcionadas por los oftalmólogos y los pacientes y adecuación de la recomendación efectuada en relación al tratamiento quirúrgico) y su resultado (satisfacción del paciente).

Resultados: En la revisión sistemática se identificaron 21 estudios relevantes de un total de 1186 publicaciones que evaluaran el suministro de información sobre CVRS salud percibida a los profesionales sanitarios. Once de los 20 estudios (55%) que valoraron indicadores de proceso y 4 de los 11 (36%) que evaluaron resultados en los pacientes detectaron al menos una diferencia significativa ($p < 0,05$) a favor del grupo de intervención. La mayor homogeneidad de los estudios centrados en el uso de información sobre la salud mental permitió realizar un meta-análisis. En 11 de ellos se evaluaron cambios en el diagnóstico, con un incremento en el mismo favorable al grupo de intervención (OR combinado = 1,91; IC 95% [1,28;2,83]). En 7 de los 9 estudios que evaluaron el tratamiento no se observaron diferencias entre grupos (OR combinado = 1,15; IC 95% [0,76;1,75]). Las herramientas de ayuda a la interpretación desarrolladas para el VF-14 demostraron una concordancia suficiente entre las respuestas predichas y las observadas. En el estudio de intervención antes-después, el coeficiente de regresión ajustado de la puntuación en el Índice VF14 se incrementó significativamente tras la intervención como predictor de la valoración realizada por el oftalmólogo (coeficiente β : control 0.10 vs intervención 0.35, $p < 0.05$). La intervención no incrementó la adecuación de la decisión médica (OR = 0.90; IC 95% [0.42; 2,69]) ni modificó la satisfacción con la atención sanitaria.

Conclusiones: Los resultados de la revisión sistemática y del meta-análisis indican que la administración de información sobre CVRS y salud percibida a los profesionales de la salud podría modificar algunos componentes del proceso de la atención sanitaria, pero no su resultado. Las características de los estudios evaluados, especialmente la gran variabilidad observada en su diseño, aconseja realizar nuevos estudios que permitan extraer inferencias más robustas. Se ha desarrollado un método para la construcción de herramientas de ayuda a la interpretación de instrumentos de calidad de

vida relacionada con la salud basadas en el contenido y que ha permitido el desarrollo de herramientas específicas para el Índice de Función Visual VF-14, el ‘termómetro’ y los ‘escenarios clínicos’. El método se ha demostrado robusto tanto desde un punto de vista teórico como empírico y debería ser fácilmente generalizable, pero este extremo no ha sido aún evaluado. En un estudio en el que se han aplicado estas herramientas se ha observado que los resultados de la revisión sistemática serían aplicables al Índice de Función Visual VF-14, observándose modificaciones en el proceso de la atención sanitaria, pero no en su resultado. Algunas limitaciones del estudio hacen recomendable acometer estudios complementarios que reproduzcan las observaciones realizadas.

INTRODUCCIÓN

La salud y la calidad de vida relacionada con la salud

Se acepta que el concepto de salud difiere en sociedades distintas e incluso en el seno de una misma si se consideran tiempos distintos. Hace ya más de medio siglo, la Organización Mundial de la Salud definió la salud como el estado de completo bienestar físico, mental y social y no sólo la ausencia de enfermedad o dolencia (WHO-OMS, 1948). Menos ambiciosa, la Real Academia Española propone en su diccionario para la voz 'salud': estado en que el ser orgánico ejerce normalmente todas sus funciones. En cualquiera de los casos queda claro que cuantificar la salud de un individuo se anticipa más complicado que, por ejemplo, cuantificar su altura o su peso. Por complicado que resulte, medir la salud es tarea imprescindible para poder evaluar el resultado de los tratamientos médicos y las intervenciones sanitarias, monitorizar la evolución de pacientes con enfermedades crónicas, identificar a las personas con mayor necesidad de atención y racionalizar la distribución de los recursos, entre otros propósitos. Medir la salud es, en suma, imprescindible para poder tomar decisiones racionales en salud.

La cuantificación de la salud se ha abordado desde dos perspectivas complementarias y a veces mutuamente excluyentes: la poblacional y la individual. La primera permite la comparación de diversos grupos de poblaciones y tiene la ventaja de poder agregar mediciones que carecen de precisión (y utilidad) a nivel individual, como la mortalidad o la morbilidad. Delante de un paciente, podremos decir que está vivo (y no muerto), o si padece una determinada enfermedad (y no está sano), pero no podremos medir con mayor precisión su salud. Las medidas de determinados aspectos individuales pueden permitir una aproximación más refinada. Pero ¿cuáles seleccionar? Las medidas antropométricas, fisiológicas y bioquímicas son el recurso habitual de la medicina clínica. Así, un valor alterado de hemoglobina nos permitirá afirmar que una persona padece anemia. Ahora bien, si entendemos que el concepto de salud incluye dimensiones físicas, psicológicas y sociales, el valor de hemoglobina en sangre nos informará muy poco sobre el estado de salud de esa persona. En la búsqueda de esa información, en las últimas dos décadas se ha producido una auténtica eclosión en la investigación clínica de medidas de la salud centradas en los pacientes, que recogieran

aquellos aspectos que forzosamente no pueden aprehender las mediciones anteriormente citadas. Y sin embargo, por razones que veremos más adelante, este espectacular avance no ha tenido una correspondencia con su uso en la práctica clínica.

Si ardua es la tarea de definir la salud, la conceptualización de la calidad de vida relacionada con la salud (CVRS) es posiblemente aún más ambiciosa. Existe una amplia variedad de definiciones de CVRS en la bibliografía médica. De entre ellas, destaca la formulada por Patrick y Erickson: el valor asignado a la duración de la vida modificada por las deficiencias, los estados funcionales, las percepciones y las oportunidades sociales, que están influidas por la enfermedad, las lesiones, el tratamiento médico o las políticas sanitarias (Patrick DL et al, 1993). Este valor a su vez puede ser asignado por individuos, grupos de individuos o por la sociedad. Lo anterior tiene dos importantes implicaciones. Por un lado, la calidad de vida relacionada con la salud es un concepto de salud que trasciende la vida y su mera duración para incorporar una valoración *subjetiva* de ambas a partir de determinados condicionantes. Por otro lado, quienes asignan el valor ya no son expertos que evalúan la salud de terceros, sino los mismos individuos (o sociedades) cuya salud se pretende cuantificar. Por eso no es infrecuente que estas medidas se describan como '*percibidas*' [por el sujeto]. En el presente trabajo se utilizarán como sinónimos de CVRS términos como salud percibida, función percibida o estado de salud percibido.

La medida de la CVRS se realiza habitualmente mediante cuestionarios estandarizados, administrados bien por un entrevistador o autoadministrados por los propios pacientes. Se han desarrollado numerosos instrumentos de CVRS, fundamentalmente en los países de habla inglesa. Para su utilización en otras culturas y en otros idiomas, se han establecido una serie de normas relativas tanto a su traducción y adaptación como a la evaluación de sus propiedades de medida. Recientemente el Scientific Advisory Committee del Medical Outcomes Trust (organización sin ánimo de lucro que estimula y patrocina el desarrollo, evaluación y distribución de instrumentos estandarizados para la medida de la salud y del resultado de la atención sanitaria) se ha basado en esas normas para la identificación de ocho atributos clave que se deberían tener en cuenta en la evaluación de cualquier instrumento: modelo conceptual, fiabilidad, validez, sensibilidad al cambio, interpretabilidad, administración, formatos alternativos y adaptación transcultural (Scientific Advisory Committee of the Medical Outcomes

Trust, 2002). A la hora de emplear estas medidas es conveniente comprender los conceptos que se están cuantificando y conocer los diferentes instrumentos disponibles al efecto, lo que permitirá seleccionar el más adecuado a los objetivos del usuario. Para facilitar la tarea se han desarrollado varias iniciativas de bibliotecas de instrumentos de CVRS, entre las que destaca una iniciativa nacional a través de la Red de Investigación en Resultados y Servicios de Salud que incluye un catálogo de más de trescientas de tales medidas para su uso en la población española (Red IRYSS, 2004).

El uso de los instrumentos de medida de CVRS en la práctica clínica

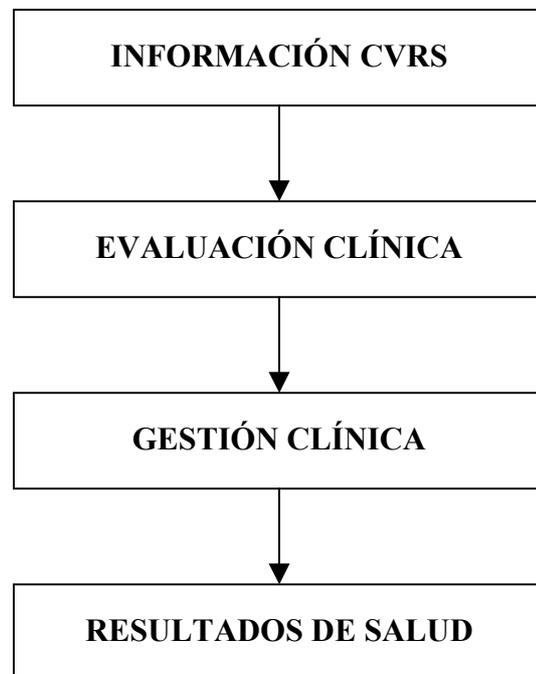
La aparente paradoja a la que se aludía se plantea cuando se considera la enorme importancia que la CVRS ha adquirido en breve plazo como medida de resultado en los programas e intervenciones sanitarias y la escasísima incidencia que ha tenido en el campo de la clínica (Clancy CM et al, 1998; Fletcher A et al, 1992)).

Los usos potenciales de las medidas de salud percibida en la práctica clínica son numerosos, en términos de diagnóstico, tratamiento y seguimiento (gestión clínica): detección de problemas funcionales y emocionales no identificados previamente, monitorización de la progresión de la enfermedad y de la respuesta al tratamiento, mejora de la comunicación entre médico y paciente entre muchos otros (Lohr KN *et al*, 1992; Deyo RA *et al*, 1992). Varios estudios han demostrado que los médicos ignoran con frecuencia algunos problemas percibidos por sus pacientes. Se señala que uno de los mayores beneficios esperados de la aplicación clínica de los instrumentos de CVRS sería precisamente proporcionar información adicional que el médico desconoce (Calkins et al, 1991; Laine et al 1995; Liaw et al, 1996; Mueller-Marquez et al, 1993; Mullan et al,1990). Esta información podría ayudar al médico y al conjunto de los profesionales de la salud a escoger un tratamiento más activo o realizar recomendaciones específicas y por tanto ayudar a los clínicos en el manejo de los pacientes. En última instancia, a mejorar su estado de salud. Atendiendo al modelo de investigación en servicios de sanitarios elaborado por Donabedian y reformulado por Starfield, cabría esperar un efecto sobre el proceso de la atención médica (como las variables relacionadas con la práctica médica y la conducta de lo pacientes) y el

resultado de la atención médica (como las variables relacionadas con la salud de los pacientes y su satisfacción) (Starfield B, 1973).

La medición de la CVRS sistematizaría y facilitaría la obtención de información que probablemente el clínico ya recoge de forma rutinaria y en la que se basa la evaluación clínica general de los pacientes. Es razonable suponer que, al incorporarse de manera sistemática, esta información modificará esa misma evaluación y favorecerá cambios en la gestión clínica de los pacientes y en los resultados de salud, tal como recoge el modelo presentado en la figura 1.

Figura 1. Un modelo para el uso de las medidas de calidad de vida relacionada con la salud (CVRS) en la práctica clínica



Se han propuesto diversos modelos teóricos que sistematizan los usos clínicos de estas medidas (Callaghan CM et al, 1996; Wasson J, Keller A et al, 1992). Callaghan y col. identifican tres estadios en la progresiva implantación de estas medidas en la práctica clínica. En el primer estadio, los resultados de la medición de la salud percibida se transforman en diagnósticos específicos (vinculación a la evaluación). En el segundo, el médico evalúa la necesidad que tiene el paciente de recursos especiales para dirigirlos a las causas específicas que subyacen a la disfunción y formula intenciones (vinculación a

los recursos). Finalmente, ambos, médico y paciente deciden si se debe iniciar un tratamiento nuevo (vinculación a la acción).

Paralelamente se ha identificado algunos problemas que podrían limitar su aplicación en el campo de la práctica clínica (Deyo RA et al, 1989). Así, la longitud de los cuestionarios puede ser excesiva y puede resultar una carga tanto para el paciente como para quien lo administra (especialmente si son profesionales de la salud). Se apuntan problemas logísticos y económicos para recoger los datos y para analizarlos, invocando la necesidad de proporcionar datos de forma inmediata e inteligible para la práctica clínica. El escepticismo reinante entre los profesionales de la salud en cuanto a su significado clínico y a su misma validez tampoco favorece su uso (Fitzpatrick R, et al, 1992).

Cabe señalar aquí que el problema de la introducción de las medidas de calidad de vida en ámbitos distintos de la investigación no sólo afecta a la práctica clínica. Su impacto es escaso en muchos otros campos de la salud pública y a veces incluso en aquellos en los que su papel está mejor establecido, como los ensayos clínicos. Hasta el 2003, de todos los fármacos aprobados para su uso en pacientes oncológicos por la Food and Drug Administration (EEUU) desde 1990, ni una sola de las aprobaciones (71) se basó en la mejora de calidad de vida (la mayoría lo fueron en base a la respuesta tumoral) (Jonson JR et al, 2003).

A pesar de los desarrollos teóricos y metodológicos sobre sus potenciales aplicaciones, el valor de la medición de rutina del estado de salud percibida no ha quedado suficientemente establecido. Los estudios de que se dispone proporcionan información contradictoria sobre su utilidad en el contexto clínico y plantean cuál sería el mejor uso de la salud percibida en la consulta, e incluso si debería tener siquiera lugar.

La interpretación de las medidas de CVRS

Uno de los problemas más frecuentemente señalados en relación con la escasa difusión en la clínica de las medidas de salud percibida es el de la interpretación de las puntuaciones, esto es, cómo asignar un significado clínico a una puntuación de CVRS. Este aspecto de los instrumentos es una cuestión de la máxima relevancia. El especial énfasis que se pone en este aspecto podría inducir a pensar que se la dificultad sea intrínseca a este tipo de medida; en realidad, se trata de un problema meramente contextual (Osoba D et al, 1995; Lydick E, 2000).

Cualquier medida requiere, en un principio, términos de comparación para orientar a quienes las usan sobre su significado clínico. Es en el marco de la constelación de conceptos y medidas relacionadas existentes y utilizadas dónde se intenta asignar un significado a los valores de la nueva medida, se contextualiza. El uso continuado favorece que, paulatinamente, no se tengan que buscar las correspondencias y relaciones de los valores de la medida en relación con las otras, si no que se utilizan de forma autónoma. La nueva medida deja de interpretarse en términos de otras para adquirir significado por sí misma. Finalmente, cuando se ha adquirido suficiente experiencia con los resultados de la nueva medida, ésta servirá para contextualizar otras medidas que puedan aparecer. De medida de interpretación incierta habrá pasado a referencia y término de comparación de cuestionarios venideros. Los modelos anteriormente citados, como el de Callaghan tendrían su lugar a partir de este estado de interpretación del instrumento. Una vez conocido el significado de las puntuaciones es cuándo pueden encontrar aplicaciones específicas (vinculación a la evaluación, los recursos o a la acción).

Los métodos de ayuda a la interpretación de las puntuaciones de cuestionarios de CVRS que se emplean durante la primera y la segunda fases descritas se suelen clasificar en categorías complementarias, según se fundamenten en comparaciones con otras medidas relacionadas clínicamente (interpretación basada en el constructo) o en comparaciones con poblaciones o grupos concretos (interpretación basada en normas) (Ware JE et al, 1996; Ferrer et al, 1997), respectivamente. Ésta última es la solución que más adeptos

ha encontrado por su similitud con los resultados de las pruebas diagnósticas empleadas en los análisis clínicos. De hecho, ya se dispone de normas poblacionales para varios instrumentos de CVRS (Ware JE et al, 1993; Crum RM et al, 1993; Klee M et al, 1997). A pesar de su atractivo, esta fórmula no está exenta de inconvenientes. La limitación más importante radica en que el término de comparación escogido es lo que en última instancia determina la gravedad aparente del individuo. Imaginemos una persona enferma con una afectación leve de su calidad de vida. Si se la compara con la población general, la gran mayoría de la cual será sana, lo más probable es que se interprete que está muy afectada. En cambio, si se le compara con pacientes enfermos, como su afectación es leve, probablemente la impresión de gravedad será mucho menor. Por ello se ha profundizado en vías complementarias de interpretación. La exploración del contenido de las preguntas (ítems) y de las actividades descritas en el cuestionario y su relación las puntuaciones de los instrumentos ofrece una tercera opción que se ha dado en llamar basada en el contenido (Ware JE et al, 1996), de escaso desarrollo hasta la fecha, pero que puede permitir una descripción mucho más rica en información. Su mayor atractivo radica en que el contenido de un instrumento es la definición operativa del constructo que se pretende medir. Finalmente, existen distintas aproximaciones estadísticas que han intentado resolver numéricamente el problema de la interpretación, sobre todo en relación con ensayos clínicos aleatorizados. Sus autores han centrado sus esfuerzos en la interpretación de diferencias entre puntuaciones, más que en las puntuaciones en sí mismas (Kazis LE et al, 1989; Juniper et al, 1994; Wyrwich et al, 1999 y 2000; Guyatt GH et al, 1998; Deyo et al, 1995).

Algunos estudios indican que la información sobre salud percibida podría mejorar algunos componentes de la atención médica. Los resultados de algunos estudios indican que la probabilidad de estos efectos beneficiosos se incrementa si los médicos clínicos reciben educación sobre informes sencillos y fáciles de interpretar en relación con instrumentos específicos y en relación con las visitas (Rubenstein et al, 1995). Por ello, si se pretende establecer el uso de estos instrumentos en la práctica clínica habitual o simplemente evaluar su empleo potencial es especialmente relevante la aplicación de herramientas específicas que mejoren la interpretabilidad de las puntuaciones.

Las cataratas y el Índice de Función Visual VF-14.

La catarata es la principal causa de discapacidad visual en el mundo (WHO-OMS, 2000), con una tendencia creciente debido a los cambios en la estructura demográfica de las poblaciones y especialmente por la mayor esperanza de vida (Thylefors B, 1998). Esta opacificación del cristalino produce una disminución gradual de la agudeza visual y alteraciones de la percepción del brillo y el color de los objetos, deslumbramiento (por el sol y otras fuentes de luz) e incluso diplopia (visión doble).

Se trata de una enfermedad asociada principalmente al envejecimiento (*catarata senil*). En nuestro medio, su prevalencia se estima en 15-20% en la población mayor de 64 años y de hasta 50% en los mayores de 74 años (AHCPR Cataract Guideline Management Panel, 1993; Congdon et al, 2004). El diagnóstico se realiza mediante oftalmoscopia directa (disminución del reflejo rojo), con lámpara de hendidura o por simple inspección.

Su tratamiento se basa de forma casi exclusiva en la extracción quirúrgica (*cirugía de cataratas*) y es un procedimiento de efectividad demostrada (Woodcock et al, 2004; Espallargues M et al, 1998; Laidlaw et al, 1998), que en la última década ha vivido un proceso de perfeccionamiento notable. La combinación de cambios en el procedimiento de extracción (facoemulsificación) y el mayor uso de la anestesia local ha favorecido el desarrollo de la cirugía ambulatoria para la extracción de cataratas (Castells X et al, 2001), reduciendo sus contraindicaciones y universalizando su uso. Prueba de ello es un reciente estudio que ha demostrado su efectividad en pacientes de edad avanzada (mayores de 85 años) (Mönestam et al, 2004). Por ello, y a pesar de haberse reducido el coste de cada intervención (Castells X et al, 2000), la cirugía de cataratas es un procedimiento que consume un volumen enorme de recursos y genera listas de espera en la mayoría de los países desarrollados (Taylor HR et al, 2000; Espallargues M et al, 2003).

Se han desarrollado numerosas guías de práctica clínica para la gestión y el tratamiento de las cataratas, pero todas ellas se basan en dos criterios fundamentales para establecer sus recomendaciones: la agudeza visual (considerada un criterio objetivo) y la función visual percibida por el paciente (criterio subjetivo) (AHCPR Cataract Management

Guideline Panel, 1993). Las guías más recientes reconocen el papel preeminente que debería tener la segunda como criterio clave para decidir el tratamiento (American Academy of Ophthalmology, Anterior Segment Panel, 2001; The Alberta Clinical Practice Guidelines Working Group, 2002) y las guías británicas más recientes llegan aún más lejos y especifican que la función visual y no la agudeza visual debería determinar la gestión y el tratamiento de las cataratas (Scottish Intercollegiate Guidelines Network (SIGN), 2001; The Royal College of Ophthalmologists, 2001). En la actualidad existen al menos una decena de instrumentos de medida de resultados percibidos para pacientes con cataratas y resulta sorprendente que en los trabajos anteriores no se recomiende específicamente el uso de ninguno de ellos en la práctica clínica habitual (de Boer et al, 2004). Esta monografía ha revisado las características psicométricas de varios instrumentos específicos relacionados con la visión y concluye que los mejores para la evaluación de función visual en pacientes con cataratas son el Visual Function VF-14 Index (Steinberg EP et al, 1994; Cassard SD et al, 1995) y la Activities of Daily Vision Scale (Mangione CM et al, 1992). Sólo se dispone de una versión adaptada a la población española del primero, el Índice de Función Visual VF-14 (VF14) (Alonso J et al, 1997).

El VF14 es una medida de capacidad visual percibida, específica para pacientes con cataratas y se basa en 14 actividades diarias que pueden verse afectadas. Los pacientes deben describir el grado de dificultad que tiene para realizar cada actividad, por culpa de las cataratas: ‘Ninguna’, ‘Poca’, ‘Bastante’, ‘Mucha’ o ser incluso ‘Incapaz de hacerla’. Se considera que un ítem no es aplicable si la actividad descrita no se realiza ‘por otras razones distintas de la vista’. En su formato estándar, se asigna un valor numérico a cada respuesta (‘Ninguna’ (4), ‘Poca’ (3), ... ‘Incapaz de hacerla’ (0)) y la puntuación total se calcula como promedio de las respuestas aplicables multiplicado por 25. Se obtiene así una puntuación entre 100 (mejor función visual) y 0 (peor función visual). El instrumento original se desarrolló hace una década (Steinberg EP et al, 1994) y su difusión ha sido tal, que se ha incluso se ha ensayado su uso para la evaluación de pacientes con otras patologías oculares distintas de las cataratas: alteraciones retinianas como la degeneración macular senil (Hart PM et al, 1999; Linder M et al, 1999;), el glaucoma (Lee BL et al, 2000) y hasta en candidatos a trasplante corneal (Boisjoly H et al, 1999). La adaptación española ha demostrado ser un instrumento fiable, válido y

sensible al cambio (Alonso J et al; 1997). En la Tabla 1 se presenta los ítems y opciones de respuesta correspondientes al cuestionario VF14.

Se dispone, por tanto de un contrastado instrumento, aplicable a la población española afecta de catarata. En el ámbito de la investigación se ha utilizado como medida de resultado en diversos estudios nacionales e internacionales, entre ellos en un estudio español en que se basa casi exclusivamente una revisión sistemática reciente de la colaboración Cochrane sobre la comparación entre tratamiento ambulatorio *vs* ingreso para la cirugía de la catarata senil (Hammed WW et al, 2004; Castells et al, 2001). También se ha utilizado una adaptación de este instrumento en un estudio cuyo objetivo es el desarrollo de un sistema de priorización de pacientes en listas de espera de cirugía electiva (cataratas, artroplastia de cadera y rodilla) (Espallargues M et al, 2003). A pesar de su extenso uso en investigación, no se dispone hasta la fecha de ningún estudio, nacional o internacional, que valore la aplicación clínica del instrumento VF14 en la evaluación de pacientes de cataratas en la práctica clínica.

Tabla 1. El Índice de Función visual VF-14

ÍNDICE DE FUNCIÓN VISUAL VF-14	
<p>◇ A causa de la vista, ¿cuánta dificultad tiene para ... <i>(incluso llevando gafas)</i>?:</p> <ol style="list-style-type: none">01. Leer letras pequeñas (guía telefónica, nombres de medicamentos, etiquetas de artículos de comida).02. Leer un periódico o un libro.03. Leer letras grandes de un libro o de un periódico, o los números de teléfono.04. Reconocer a personas cuando están cerca.05. Ver escalones, peldaños, o el bordillo de la acera.06. Leer letreros de las calles y tiendas, los números de las casas, o ver los semáforos.07. Hacer trabajos manuales finos, como coser, arreglar un enchufe o clavar un clavo.08. Hacer crucigramas, rellenar un impreso o hacer una quiniela.09. Jugar a las cartas, al dominó o al bingo.10. Participar en actividades como la petanca, buscar setas, cuidar plantas o mirar escaparates.11. Cocinar.12. Ver la TV.13. Conducir de día.14. Conducir de noche.	
<p>↳ Categorías de respuesta para cada pregunta:</p> <ol style="list-style-type: none">0. Incapaz de hacerla1. Mucha dificultad2. Bastante dificultad3. Poca dificultad4. Ninguna dificultadX. No lo hago por otras razones (no por la vista)	

JUSTIFICACIÓN

La información sobre el efecto de la introducción de medidas de calidad de vida en la práctica clínica es insuficiente. Si bien se dispone de estudios que han realizado esta evaluación, sus resultados no son concluyentes y la evidencia es dispersa. Es necesaria una revisión sistemática que sintetice todo este cuerpo de evidencia y proporcione una síntesis cualitativa y cuantitativa y que oriente sobre el papel destinado a este tipo de medidas en la práctica clínica.

La interpretación de las medidas de CVRS es un problema contextual que se ha abordado desde múltiples enfoques, pero que persiste como limitación para su uso clínico. Son necesarias herramientas de apoyo a la interpretación que faciliten su uso en la evaluación y monitorización de los pacientes. Lo anterior es especialmente aplicable a los instrumentos que carezcan de guías explícitas de interpretación, como es el caso del Índice VF14, tanto para su versión adaptada como para el instrumento original.

Finalmente, hay una considerable falta de evaluación del impacto de la medición de la CVRS en el ámbito español. En la evaluación de los servicios sanitarios, las diferencias existentes entre países hacen especialmente recomendable replicar localmente evaluaciones realizadas en otros contextos. La evaluación de los pacientes con cataratas ofrece un contexto y una oportunidad excelentes, por cuanto existe consenso sobre la necesidad de una evaluación de los pacientes basadas en la afectación de las actividades diarias (calidad de vida) y no se dispone de información sobre el impacto de su introducción en la práctica clínica. Particularmente, la decisión sobre la indicación de la cirugía se toma durante la visita médica, lo que justifica la conveniencia de suministrar la información en ese momento.

OBJETIVOS

OBJETIVO GENERAL

Estudiar el impacto en la atención médica del uso sistemático de instrumentos de medición de calidad de vida y del suministro de la información a los profesionales de la salud

OBJETIVOS ESPECÍFICOS

1. Describir el estado actual del conocimiento sobre el impacto en la práctica clínica de intervenciones que incluyan del uso de información sobre salud percibida y su suministro de rutina a los profesionales de la salud.
2. Desarrollar un método general para la obtención de herramientas que faciliten la interpretación de las medidas de salud y favorezcan su uso.
3. Aplicar el método desarrollado al Índice de Función Visual VF-14 para pacientes con cataratas.
4. Evaluar el impacto en la práctica clínica de la introducción del uso sistemático del Índice de Función Visual VF-14.

MÉTODOS Y RESULTADOS

A continuación se describen los tres artículos detallando: *a)* la metodología empleada para alcanzar los objetivos propuestos y *b)* los resultados obtenidos (Tabla 3). Cada artículo va precedido de una traducción de su resumen en español.

Síntesis de métodos

El primero de los artículos es una revisión sistemática de ensayos clínicos que proporciona un resumen cualitativo y cuantitativo (meta-análisis) del impacto del uso de resultados percibidos por los pacientes en la práctica clínica y que abarca el período 1966-1997. Se incluye un intercambio epistolar suscitado en la revista *Medical Care* a raíz de la publicación de los resultados.

El segundo artículo es un trabajo de carácter metodológico. Tal como se ha indicado, una de las barreras para la difusión de las medidas de CVRS en la práctica clínica es la interpretación de las puntuaciones. Se presenta en este trabajo un método para el desarrollo de herramientas de ayuda a la interpretación de medidas de resultados percibidos y que fue elaborado utilizando métodos psicométricos avanzados (Teoría de Respuesta al Ítem). Para su aplicación al índice de Función Visual VF-14 se emplearon las respuestas proporcionadas por los pacientes españoles (Barcelona) incluidos en un estudio de cohortes multicéntrico internacional I-PORT realizado también en EEUU, Manitoba (Canadá), Dinamarca. El objetivo del estudio era la evaluación de los resultados de la cirugía de cataratas y para la muestra de Barcelona se entrevistó individuos a los que se había indicado el tratamiento quirúrgico en diversos hospitales de la red pública (Xarxa d'Hospitals d'Utilització Pública) y privada entre 1993 y 1996. De ellos se seleccionaron todos los pacientes con respuestas válidas para el Índice VF14 (n=192).

El tercer artículo presenta los resultados de un estudio de intervención realizado entre 1999 y 2000 en consultas de oftalmología de consultas externas hospitalarias y de la

atención primaria especializada de Barcelona y en otras ciudades de su área de influencia (Rubí, Sant Boi, Terrassa y Viladecans). El objetivo del estudio fue la evaluación del impacto del uso del instrumento VF-14 en la práctica clínica, específicamente en la evaluación habitual de pacientes afectados de cataratas. En él participaron 19 oftalmólogos y 890 pacientes diagnosticados de cataratas en los que la decisión de operar no hubiera aún tomado en el momento de su inclusión.

Tabla 2. Artículos y objetivos específicos

ARTÍCULO	OBJETIVOS ESPECÍFICOS
<p>"Provision of feedback on perceived health status to health care professionals: a systematic review of its impact."</p> <p>Espallargues M, Valderas JM, Alonso J.</p> <p><i>Medical Care</i> 2000 Feb;38(2):175-86</p>	<p>1. Describir el estado actual del conocimiento sobre el impacto en la práctica clínica de intervenciones que incluyan del uso de información sobre salud percibida y su suministro de rutina a los profesionales de la salud.</p>
<p>"Content-based approach for facilitating the interpretation of health-related quality of life measures in clinical practice. A proposal for the visual function index VF-14".</p> <p>Valderas JM, Alonso J, Prieto L, Espallargues M, Castells FX.</p> <p><i>Quality of Life Research</i> 2004 13: 35–44.</p>	<p>2. Desarrollar un método para la obtención de herramientas que faciliten la interpretación de las medidas de salud y favorezcan su uso.</p> <p>3. Aplicar el método desarrollado al Índice de Función Visual VF-14 para pacientes con cataratas.</p>
<p>"The impact of the VF-14 Index, a perceived visual function measure, in the routine management of cataract patients".</p> <p>Valderas JM, Rue M, Guyatt G, Alonso J.</p> <p><i>Quality of Life Research</i> 2005 (in press).</p>	<p>4. Evaluar el impacto en la práctica clínica del uso sistemático del Índice de Función Visual VF-14.</p>

Artículo 1. "Provision of feedback on perceived health status to health care professionals: a systematic review of its impact." Espallargues M, Valderas JM, Alonso J. Med Care. 2000 Feb;38(2):175-86

Objetivos: Estudio del impacto en el proceso y en el resultado de la atención médica en la práctica clínica del suministro de información sobre estado de salud percibido.

Diseño: Revisión sistemática de ensayos clínicos controlados.

Identificación de los datos: Búsqueda en bases de datos electrónicas (MEDLINE 1966-1997), búsquedas manuales y consultas a expertos en el área.

Análisis de los datos: se estudiaron las diferencias entre grupos control e intervención en proceso de la atención médica (utilización de servicios sanitarios, diagnóstico y tratamiento), resultados en pacientes (estado de salud) y satisfacción del paciente. En un subgrupo de 13 intervenciones que estudiaron el suministro de información relativa a la salud mental de los pacientes, se realizó meta-análisis del impacto en el proceso de atención.

Resultados: Se identificaron 21 estudios que satisfacían los criterios de selección. Once de veinte (55%) de los que estudiaron indicadores de proceso detectaron al menos una diferencia significativa ($p < 0,05$) a favor del grupo de intervención. De los 11 ensayos que evaluaron variable de resultado sólo 4 (36%) detectaron mejorías significativas. Se observó una tendencia similar, con porcentajes menores, en las 8 intervenciones que suministraron información sobre estado de salud general. En 11 intervenciones que evaluaron el uso de información sobre salud mental se detectó un incremento de diagnóstico en el grupo de intervención (OR combinado=1,91; IC95% [1,28;2,83]). En 7 de los 9 estudios que evaluaron el tratamiento no se observó efecto algunos sobre este indicador (OR combinado=1,15; IC95% [0,76;1,75]).

Conclusiones: El suministro de información sobre el estado de salud percibido a los profesionales de la salud parece tener un efecto sobre el proceso de la atención médica, pero no sobre el estado funcional o de salud. Esto es especialmente cierto para la información sobre salud mental. Empero, es necesaria una evaluación más minuciosa de este tipo de intervención.

Este artículo y la carta al director que generó son consultables en:

Espallargues M, Valderas JM, Alonso J

Provision of feedback on perceived health status to health care professionals: a systematic review of its impact.

Medical care. 2000 Feb;38(2):175-86

Goldman RL.

Provision of feedback on perceived health status to health care professionals.

Medical care. 2000 Aug;38(8):877-8

Espallargues M, Valderas JM, Alonso J

Response.

Medical care. 2000 Aug;38(8):878-9

Artículo 2. “Content-based aids for the interpretation of health-related quality of life measures in clinical practice. An example for the visual function index VF-14”. Valderas JM, Alonso J, Prieto L, Espallargues M, Castells FX. *Quality of Life Research* 2004 13: 35–44.

Antecedentes: A pesar de que el desarrollo de instrumentos de calidad de vida está bien establecido, las dificultades que presenta la interpretación de sus resultados puede limitar su utilización en la práctica clínica.

Objetivo: Desarrollar herramientas de ayuda a la interpretación de una medida de salud percibida, el índice VF-14, que sean generalizables.

Diseño: Se utilizó la Teoría de Respuesta al Ítem (Análisis de Rasch), para evaluar las características de los ítems del VF-14. A partir de la actividad más difícil que un paciente podía realizar sin problemas se desarrolló la herramienta ‘termómetro’¹; para la construcción de la herramienta de ‘escenarios clínicos’ se identificaron todos conglomerados significativamente distintos de ítems y luego se estimó la dificultad esperada (respuesta) para realizar una actividad representativa de cada agregado.

Emplazamiento: El estudio se realizó en 4 hospitales y 6 centros de cirugía de cataratas ambulatoria de Barcelona, España.

Pacientes: 192 pacientes a los que se practicó cirugía de cataratas en el primer ojo.

Mediciones: El índice de función visual percibida VF-14 y mediciones clínicas.

Resultados: Los 14 ítems configuraron una única dimensión, si bien tres presentaron desajuste parcial. Para un paciente con una puntuación Rasch de 71, el ‘termómetro’ indicó que la actividad más exigente que podría realizar sin dificultad sería ‘Hacer trabajos manuales finos como coser, arreglar un enchufe o clavar un clavo’. Los ‘escenarios clínicos’ estimaron que ese paciente sería incapaz de ‘Conducir de noche’, que tendría cierta dificultad para ‘Leer letra pequeña’ y que no debería tener ninguna dificultad para ‘Hacer trabajos manuales finos, ...’, ‘Ver la TV’ o ‘Reconocer a personas’. La concordancia entre las respuestas predichas y las observadas varió entre bastante y substancial.

¹ En el texto del artículo se optó por la nomenclatura ‘*ruler*’ (regla) para esta herramienta. Se prefiere aquí la analogía con un ‘termómetro’ (de connotaciones clínicas más evidentes), que preserva la noción de *instrumento que sirve para medir* y que ambas voces comparten en su definición en el Diccionario de la Real Academia.

Conclusiones: Se han desarrollado herramientas para la interpretación simples y basadas en el contenido que deberían facilitar su uso en la práctica clínica. Estas herramientas deben ser fácilmente generalizables a otras medidas de calidad de vida.

[SUBMITTED VERSION]

Content-based interpretation aids for health-related quality of life measures in clinical practice. An example for the visual function index VF-14.

Valderas J.M., MD (1), Alonso J., MD (1,2), Prieto L., PhD (1), Espallargues M., MD (3), Castells F.X., MD (1).

(1) Health Services Research Unit. Institut Municipal d'Investigació Mèdica (IMIM-IMAS),

(2) Universitat Autònoma de Barcelona, and

(3) Catalan Agency for Health Technology Assessment and Research, Barcelona, Spain.

Abstract

Background: In spite of a well-established development of instruments, difficulty in interpreting health related quality of life scores may limit its use in clinical practice.

Objective: To develop generalisable interpretation aids for a measure of perceived functional visual status, the VF-14 index. *Design:* Item Response Theory (Rasch analysis) was used to analyse the performance of VF-14 items. The 'ruler' aid was derived from the most difficult activity (item) a patient is able to do without difficulty; the 'clinical scenarios' aid, identified first all significantly different clusters of items within the index and then estimated the mean expected difficulty (responses) to perform a benchmark item in each cluster. *Setting:* 4 hospital and 6 ambulatory cataract surgery centers in Barcelona, Spain. *Patients:* 198 patients undergoing first eye cataracts surgery. *Measurements:* The self-reported VF-14 index and clinical measures. *Results:* All VF-14 items were found unidimensional with 3 items showing only partial misfit. For a patient with a VF-14 Rasch score of 71 the 'ruler' aid indicated that 'doing fine handwork' would be the most requiring activity he/she would perform without difficulty. The 'clinical scenarios' aid estimated that such a patient would be unable to 'drive at night', would have some difficulty 'reading small print' and no difficulty 'doing fine handwork', 'watching TV' or 'recognizing people'. Concordance between modeled and observed responses was fair to substantial. *Conclusions:* Simple content-based interpretation aids for the VF-14 scores were developed that should facilitate its use in clinical practice. These aids should be easily generalizable to other quality of life instruments.

This work was supported by *Fondo de Investigación Sanitaria* (97/1135). *Instituto de Salud Carlos III* (98/4471) and *Generalitat de Catalunya* (CIRIT 1999 SGR 00240).

Introduction

Despite the increase of the use of health related quality of life (HRQL) instruments in clinical research, its use in clinical practice remains scarce (1-3). It has been emphasized that the difficulty in interpreting these scores might play an important role here (3). Still interpretability is not a problem associated with this specific kind of measurement but a contextual feature (4,5).

Interpretation of HRQL measurements can be considered in one of three consecutive stages demanding specific aids to the interpretation. Shortly after the instrument has been introduced, HRQL scores may be compared with other clinical measures to highlight the connections between clinically related variables, an approach termed construct-based (6,7). Scores can also be interpreted by inspection of the responses to some of the items that form the questionnaire (content-based interpretation) (6). Subsequently, expertise will then allow the clinician to manage the bare values of the HRQL scores and information based on the scores distribution in different populations will provide new opportunities for interpretation, an approach termed norm-based (6) and relying on the psychometric tradition (8). General population norms are available for a number of instruments (9-11), and so are data about the distribution of the scores in well defined groups (12). Conceptual limitations to the use of norms derive from the fact that they result in moving standards, since the individuals serving as comparison frame determine the apparent severity of the individual's status. Finally, the HRQL scores are explicitly used for diagnostic or prognostic purposes and the clinician may then use the measure to enhance interpretation of new coming measures.

Several statistical approaches to the interpretation of HRQL measures have been mainly presented for the interpretation of the results of clinical trials (13-18), their main limitation arising from their focusing on the change of scores rather than on the scores themselves.

Our objective was to synthesize HRQL scores in a way that facilitates their interpretation. Specifically we applied a specific model (Rasch analysis) of Item Response Theory (IRT) methods to an instrument measuring perceived visual function, the Visual Function Index (VF-14), for which no explicit interpretation guidelines existed (19-24).

Methods

The Visual Function Index VF-14

The VF-14 is a measure of perceived functional capacity related to vision based on 14 everyday activities that can be affected by cataracts (25,26) (Table 1). Patients rate their difficulty performing each activity because of their vision. Items are considered not applicable if not performed for reasons other than vision. In its standard form, the score is calculated as the summated responses divided by the number of valid responses and multiplied by 25, in order to render a score between 100 (best function) and 0 (worst function). The Spanish version of the VF-14 has been shown to be reliable, valid and responsive (27).

[TABLE 1 ABOUT HERE]

Patients

One hundred and ninety eight consecutive patients aged 50 or older that were scheduled for cataract surgery were recruited in the context of the International Cataract Surgery Outcomes Study in Barcelona, Spain. Eligibility criteria, data procedures and patients characteristics have been previously presented (28). Patients were interviewed by telephone using a structured questionnaire including information on perceived health status, perceived visual function and satisfaction with vision. Patients underwent a clinical examination pre- and postoperatively and their clinical data forms, including visual acuity (reported here as a decimal value between 1 (best) and 0 (worst)), were completed by the participating ophthalmologists.

The Rasch Model

The Rasch analysis is a probabilistic model of analysis of questionnaires based on IRT. Under the testable assumption of a unidimensional latent (not apparent) variable, person measures (also called *abilities*) and items measures (also called *difficulties*) are simultaneously estimated in interval-scale along the same continuum (19,20). This allows ordering of both persons, after their abilities, and items, considering their intrinsic difficulty. Another feature of the model is the reproducibility of the hierarchic structure of the items in the questionnaire in different patient groups (sample-free measurement) and across test occasions (test-free measurement) (19,20). The terms ‘ability’ and ‘difficulty’ come from the educational field, where IRT was first applied to the measurement of achievement through exercises of varying complexity. Regarding health status measurement, the ability (β) of an individual may be equated to his/her functional status (e. g., visual function). The difficulty (δ) of an item relates to the functional status required for performing the activity described in the item and is inherent to each item.

Item difficulties and person abilities

Person and item parameters derive from modeling the relation of the difference between item difficulty and person ability ($\beta-\delta$) and the probability of a person endorsing a certain response to a given item. If the person ability is higher than the item difficulty ($\beta-\delta>0$) (the functional level required for performing the activity is lower than the actual status of the person), the individual will be expected to perform successfully the activity. The bigger the difference, the higher the probability of successful performance. Conversely, if the person ability is lower than the item difficulty ($\beta-\delta<0$), the person will be expected to fail in performing the activity. When the person ability equates the item difficulty ($\beta-\delta=0$) the outcome is very uncertain and the probability of a successful performance is 0.5.

The Rating Scale Model is a generalization of the Rasch Model for items with more than two response categories (20). The difficulty parameter is here split into single difficulty parameters for each different response category. An average item difficulty and parameters for each response category parameter (called ‘thresholds’ or ‘steps’) are modeled. Category parameters are intuitively expected to be ordered according to the nominal hierarchy of the response categories, but this may not be necessarily the case. When they are not properly ordered they are called ‘reversed’ and indicate imperfect functioning of the rating scale. A different set of response categories should then be considered, or alternatively adjacent categories might be collapsed, if this is clinically meaningful (29).

Under the Rating Scale Model, expected values can be computed that predict the response of the average patient with an ability β_n when confronted to an item with difficulty δ_i (23), using the formula:

$$E_{ni} = \sum_{x=0}^m x \pi_{nxi} = \sum_{x=0}^m x \frac{\sum_{j=0}^x [\beta_n - (\delta_i + \tau_j)]}{\sum_{k=0}^m \sum_{j=0}^k [\beta_n - (\delta_i + \tau_j)]} ;$$

where π_{nxi} is the probability of selecting the response category x out of the $m + 1$ different categories for given person (β_n) and item (δ_i) measures, and τ_j is a category measure (threshold).

Rasch analysis of the VF-14 index was performed with software Bigsteps® 2.73 (30). Based on patients with non extreme raw scores (raw scores other than 100 and 0), the analysis provided with estimates of item difficulties (in logit units) and their standard errors. A linear transformation was used to express the logits into a scale ranging from 100 (best function) to 0 (worst function). One measure was estimated for each person across all items and one difficulty for each item across all persons. Consecutive items in the hierarchy with non overlapping 95% confidence intervals (CI) of their difficulty were considered to pertain to different clusters of items and defined significant gaps (31). Extreme scores have no exact measure in the Rasch model. Bigsteps allows estimation of reasonable measures for each extreme score using a standard Bayesian technique, however (32). Consistency of traditional scores and Rasch person measures was assessed by means of the intraclass correlation coefficient for continuous measures (50).

Goodness of fit

Item fit to the model was assessed by standardized infit and outfit statistics (20). They determine how well each item contributes to define the unidimensional variable. Infit statistics are sensitive to unexpected responses to items near the person's ability level, while outfit statistics are more sensitive to those far away. Both were standardized to approximate a theoretical mean 0 and variance 1 distribution (ZSTD statistics), with misfitting values exceeding the critical value 1.96 (20,33).

The 'ruler' interpretation aid

Based on the item hierarchy revealed by the Rasch analysis (20), a vertical ruler-like figure representing the visual function continuum was constructed, with a scaling ranging from 0 (bottom) to 100 (top). For simplicity, we estimated the item parameters for the response category of 'No difficulty'. A convergent iterative method based on Bolzano's theorem was used to obtain an estimate of the minimal difference between ability and difficulty ($\beta - \delta$) that rendered an expected response of 'No difficulty' in the previous equation (34). Intervals of decreasing range were constructed for this estimation. In each iteration the interval was narrowed selecting as extreme values one extreme and the mean value in the interval of the previous iteration, (the extreme was selected that rendered an expected response different to that of the mean value). The recursive algorithm arrived to an end when the difference between extreme values in the interval was < 0.01 . When adding the difference to any item's difficulty, the minimal ability required for expecting a 'No difficulty' response to this item

was obtained. These values were distributed along the 0-100 ruler. We additionally checked if the empirical data supported this hierarchical distribution of VF-14 items by calculating the number of times that a patient reported more difficulty with a less difficult of two contiguous items (discrepancies).

The 'clinical scenarios' interpretation aid

For a more rich description of the functional capacity of each patient, we estimated the number of statistically different clusters of items in the questionnaire. The most unambiguous and most frequently endorsed item within each cluster was chosen as benchmark. Average expected responses to benchmark items were then calculated for every 5 points scores, from 0 to 100. Finally, expected difficulty answers were tabulated, with conventional non numerical symbols: '-' (none); '+' (some), and '++' (inability). An additional check of the fit, of this approach was conducted by estimating the agreement between observed and expected responses to benchmark items using the weighted κ (kappa) statistic for ordinal measures (35).

Results

Mean age was 70.1 (sd=8,6) and 60.6 % of the patients were women. Median visual acuity was 0.3 for the best (non operative) eye and 0.2 for the worst (operative) eye. About 90% stated to be 'dissatisfied' or 'very dissatisfied' with their vision. VF-14 items endorsed by patient ranged from 4 to 14 items (median= 11).

VF-14 items performance

One hundred and eighty four (92.9%) patients with non extreme scores were susceptible to Rasch analysis. Of the remaining, 11 (5.6%) had maximum (100) and 3 (1.5%) had minimum (0) raw scores. The intermediate response categories of the standard form of the VF-14 showed reversed parameters (thresholds). Thus only the extreme categories were retained in their original form, ('No difficulty' and 'Unable to do the activity'), while the remaining categories (from 'A little difficulty' to 'A great deal of difficulty') were collapsed into a new category ('Some difficulty'). Data presented here result from the analysis of the questionnaire considering this modification. Person abilities were normally distributed (52.86 (sd=18.67)). Item difficulties, standard errors and fit statistics are reported in table 2, with items arranged in decreasing difficulty order.

Eleven of the 14 items completely fit the unidimensional test. The other three items showed partial misfit (item 1, *Driving at night* (INFIT ZSTD=2.0), and item 5, *Seeing steps, stairs* (INFIT ZSTD=-2.4); and item 4, *Recognizing people* (OUTFIT ZSTD=2.1)). None failed the two misfit criteria, thus all items were retained in the questionnaire for the development of interpretation aids.

[TABLE 2 ABOUT HERE]

VF-14 scores provided by Rasch analysis showed a high but not perfect level of agreement with traditional scores (ICC = 0.80). Correlation of the VF-14 Rasch scores and the clinical

measure (visual acuity) was very similar to the one observed with traditional scores previously (27) (Spearman $r = 0.46$ [CI 95% 0.27-0.5]).

The 'ruler' interpretation aid

The item difficulties estimated to perform the activity with 'No difficulties' were used to construct the ruler-like aid presented in Figure 1. The figure indicates the expected performance of the average patient with a given score. As an example, a patient with a Rasch score (person measure or ability) of 71 in the VF-14 is represented by the line. The line divides the ruler into two sections: some activities above the patient's ability level and some below. Items in the lower section are activities that, on average, the patient would be expected to have 'no difficulty' to perform. Activities in the upper section (dotted) would be performed with some degree of difficulty, or not performed at all. Thus we expect a patient with a score of 71 to have no difficulty in '*Recognizing people*' or '*Doing fine work*'. Conversely his/her ability would lay under the level required for having no difficulty in '*Driving at night*', '*Driving during the day*' or '*Reading small print*'.

The patient's ability can be further summarized in terms of the most difficult activity that the patient is expected to perform without difficulty. A VF-14 score of 71 would correspond to a visual function allowing the patient to 'Do fine handwork' without difficulty (any other VF-14 activity placed higher in the hierarchy would be too difficult to be performed without any difficulty).

[FIGURE 1 ABOUT HERE]

There were few discrepancies between the modeled item hierarchy and the observed responses. About two thirds of the patients showed only 2 or less discrepancies and almost 90% 3 or less, (the number of pairs of item being compared varied from patient to patient, due to varying number of items responded).

The 'clinical scenarios' interpretation aid

The Rasch analysis identified five significantly different clusters of visual capacity items within the VF-14 (last column in Table 2). Accordingly, 5 benchmark items, one in each cluster, were selected for defining the visual functions clinical scenarios: '*Driving at night*', '*Reading small print*', '*Doing fine handwork*', '*Watching TV*' and '*Recognizing people*'. The expected response of level of difficulty to each of these items is presented for given VF-14 scores in Figure 2. Thus, to associate the clinical scenarios to scores other than the presented, the first should be approximated to the closest one in the table.

[FIGURE 2 ABOUT HERE]

For the previous patient with a VF-14 score of 71, we will observe that the closest score is 70. For such a score, the patient is expected to be unable of '*Driving at night*', to have some difficulty '*Reading small print*', and to have no difficulty in doing any of the other three activities. The agreement between actual and expected responses to benchmark items ranged from substantial (weighted $\kappa = 0.66$ [SE=0.05], '*Doing fine handwork*') to fair (weighted $\kappa = 0.25$ [SE=0.05], '*Recognizing people*').

Discussion

A method is presented for the development of interpretation aids allowing a meaningful summary of health status measures. It is based on the content of a measuring instrument (content-based) and it relies on two simple and intuitive bases of Item Response Theory (IRT) models (19): 1) each item is an operative definition of the concept being measured, and 2) items describe activities that can be meaningfully ordered according to the increasing ability needed to perform them. By means of one of the IRT models, the Rasch analysis, such a continuum can be built, where items represent marks of increasing functional difficulty on this conceptual ruler. In our approach, the perceived visual function of any person responding to the scale was summarized first in terms of the more difficult activity in the VF-14 index the person was able to pass without any difficulty. We call it the 'ruler' aid, since the items are ordered according to their difficulty. Subsequently, all the questionnaire item clusters were identified, and one item (benchmark) in each cluster was selected. This aid allowed to use only a subgroup of items to characterize a patient's perceived visual function. The patient's average expected responses or abilities to perform benchmark items provided a more clear clinical picture associated with a given score (the 'clinical scenario'). Together with an acceptable goodness of fit, the sequencing of the difficulty of the VF-14 and the clinical scenarios were consistent with clinical experience, which further supports the face validity of our approach. By construction, our approach should be applicable to many other unidimensional health-related quality of life questionnaire.

The analytical model underlying our interpretation aids, IRT, is based on what test takers can do (i.e., it is content-based). It might be argued that a normative framework is implicit in this procedure since previous empirical calibration is needed. Rasch measurement, though, is often defined as 'objective' because it guarantees the invariance of item and person parameters (19). The estimates of ability are freed from the effects of difficulty and the estimates of difficulty are freed from the effects of ability. This is so because each facet of the model is a separate parameter. This property implies that the parameter that characterizes an item does not depend on the ability distribution of the individuals. Conversely, the person parameter does not depend on the set of test items (19). A demonstration of this characteristic is beyond the scope of this paper. Nevertheless, we have empirically verified that while visual function and visual acuity changed dramatically after surgery in our patients, VF-14 item calibrations estimated preoperatively and postoperatively were very similar (data not shown). This provides empirical evidence of the independence of the measures from individual abilities distribution, although further research is necessary with different cataract patient populations, as it has been reported for other health status instruments (36,37).

Several issues in our paper deserve further discussion. First, although the results of Rasch analysis suggest that the VF-14 defines a unidimensional variable of meaningful direction, three items showed partial misfit. These items were retained for practical reasons: two were activities applicable to all but one patient (99.5%) and the third ('*Driving at night*') while not very applicable, offered more discrimination capacity to the questionnaire. In addition, keeping these items facilitates comparability of the results with previous work using the VF14 Index. Nevertheless, it should be assessed if these items are misfitting also in other adapted versions of the VF-14 index (27,38). For an adequate interpretation, any truly misfitting item should be eliminated or replaced by a good performing item to achieve an adequate reliable measure.

Also, we observed imperfect functioning of the original rating scale of the VF-14 index. Thus the original response categories had to be collapsed into clinically meaningful categories. While we are convinced that the feeling of the patients responses has been preserved, we lack empirical evidence of the equivalence of these two response formats. In addition, it should be borne in mind that we choose one particular response category ('No difficulty') as the reference for our ruler-approach. We did so taking into account that this category requires the highest capacity level and it could be the most sensitive in the detection of impairment produced by cataracts. But other options might be valid as well. The extent to which such a decision might influence the clinical interpretation deserves further analysis.

Given the probabilistic nature of Rasch analysis outcomes, some disparity was anticipated when comparing average expected and actual responses. This was specially true since we grouped the VF-14 scores in 5 units for the 'clinical scenarios' approach. Nevertheless, the concordance observed in our study is rather satisfactory and support the validity of this approach.

Finally, it must be noticed that the set of benchmarks items in the 'clinical scenario' aid was selected attending to specific criteria and should not be interpreted as a suggestion of their using as 'short form' of the original VF14.

Although interpretation aids presented here have been developed specifically for the VF-14 index, there is no reason to believe that this method might not be applied to many other functional status instruments. The basic principles upon which they rely (operability and order) apply to most HRQL questionnaires. Whether the test is unidimensional or multidimensional would only imply that as many different aids as separate dimensions are to be developed. Rasch analysis has been successfully applied to well known multidimensional instruments (37). The method presented here should facilitate the use of perceived health measures in clinical practice. But an expansion of these applications is necessary and more evidence of its applicability to other measures will be needed. An empirical evaluation of the effect of routinely providing ophthalmologists with the VF-14 scores of their cataract patients supported by such aids is currently underway.

Appendix

Appendix Table is a 'Response plot', a graphic output elicited by the program Bigsteps®. It estimates 'the average rating that we expect to observe for persons with a particular measure (Rasch score)'. Expected scores for each item are plotted against person measures. The vertical axis lists the items in descending order of difficulty, blinks indicating statistically significant gaps between the clusters of items along the variable. The horizontal axis shows person overall scores with the distribution of the person abilities beneath them. An 'M' marker represents the location of the mean person ability. 'S' markers are placed one standard deviation away from the mean. 'Q' markers are placed two standard deviations away. Between any two colons (:), expected answers can be read inside the quadrant (':') representing the half-score points between response choices). As an example, patients scoring 71 are expected to have a great deal of difficulty (0) to drive at night, some difficulty (1) to read small print and no difficulty (2) to take part in activities or to prepare meals. Approximated minimal ability values necessary for the construction of the 'ruler' approach can be easily obtained from this table. They correspond to the colons (:) between the two top response categories, here 2 ('No difficulty') and 1 ('Some difficulty'). Though comprehensive, this summary is not as user friendly the ruler approach presented here.

This output can be further adapted for the construction of a simpler 'clinical scenarios' interpretation aid. First, according to predefined criteria one representative item can be selected from each cluster. Extreme position and clinical relevance were considered critical for our purposes, but specific needs can differ, like maximizing activities applicable for many people. The expected responses for given person abilities can be read from the table, but only approximately. Unfortunately, lack of precision will not allow extensive use of the method.

References

- (1) Fletcher A, Gore S, Jones D, Fitzpatrick R, Spiegelhalter D, Cox D. Quality of life measures in health care. II: Design, analysis, and interpretation. *BMJ* 1992; 305:1145-1148.
- (2) Clancy CM, Eisenberg JM. Outcomes research: measuring the end results of health care. *Science* 1998; 282(5387):245-246.
- (3) Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals. A systematic review of its impact. *Med Care* 2000; 38(2):175-186.
- (4) Osoba D, Till JE, Pater JL, Young JR. Health-related quality of life: measurement and clinical application. A workshop report. *Can J Oncol* 1995; 5(1):338-343.
- (5) Lydick E. Approaches to the interpretation of quality-of-life scales. *Med Care* 2000; 38(9 Suppl):II180-II183.
- (6) Ware JE, Keller SD. Interpreting General Health Measures. In: Spilker B, editor. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia: Lippincott-Raven Publishers, 1996: 445-460.
- (7) Ferrer M, Alonso J, Morera J, Marrades R, Khalaf A, Aguar MC et al. Chronic obstructive pulmonary disease stage and health-related quality of life. The Quality of Life of Chronic Obstructive Pulmonary Disease Study Group. *Ann Intern Med* 1997; 127(12):1072-1079.
- (8) Nunnally JC, Bernstein IH. *Psychometric theory*. 3 ed. New York: McGraw-Hill, 1994.
- (9) Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 Health Survey Manual and Interpretation Guide*. Boston, MA: The Health Institute, 1993.
- (10) Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA* 1993; 269(18):2386-2391.
- (11) Klee M, Groenvold M, Machin D. Quality of life of Danish women: population-based norms of the EORTC QLQ-C30. *Qual Life Res* 1997; 6(1):27-34.
- (12) Stewart AL, Greenfield S, Hays RD, Wells K, Rogers WH, Berry SD et al. Functional Status and Well-being of Patients With Chronic Conditions. Results from the Medical Outcomes Study. *JAMA* 1989; 262(7):907-913.
- (13) Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; 27(3 supp.):S178-S189.
- (14) Juniper EF, Gordon GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994; 47(1):81-87.
- (15) Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999; 52(9):861-873.
- (16) Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. *BMJ* 1998; 316(7132):690-693.
- (17) Wyrwich KW, Wolinsky FD. Identifying meaningful intra-individual change standards for health-related quality of life measures. *J Eval Clin Pract* 2000; 6(1):39-49.
- (18) Deyo RA, Patrick DL. The significance of treatment effects: the clinical perspective. *Med Care* 1995; 33(4 Suppl):AS286-AS291.
- (19) Wright BD, Stone MH. *Best test design*. Chicago: MESA Press, 1979.
- (20) Wright BD, Masters GN. *Rating Scale Analysis*. Chicago: Mesa Press, 1982.
- (21) Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care* 2000; 38(9 Suppl):II28-II42.
- (22) Coster W, Ludlow L, Mancini M. Using IRT variable maps to enrich understanding of rehabilitation data. *J Outcome Meas* 1999; 3(2):123-133.
- (23) Haley SM, McHorney CA, Ware JE, Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol* 1994; 47(6):671-684.

- (24) Fisher WP, Jr. A research program for accountable and patient-centered health outcome measures. *J Outcome Meas* 1998; 2(3):222-239.
- (25) Steinberg EP, Tielsch JM, Schein OD, Javitt JC, Sharkey PD, Cassard SD et al. The VF-14: an index of functional impairment in patients with cataract. *Arch Ophthalmol* 1994; 112:630-638.
- (26) Cassard SD, Patrick DL, Damiano AM, Legro MW, Tielsch JM, Diener-West M et al. Reproducibility and responsiveness of the VF-14. An index of functional impairment in patients with cataracts. *Arch Ophthalmol* 1995; 113:1508-1513.
- (27) Alonso J, Espallargues M, Folmer-Andersen T, Cassard SD, Dunn E, Bernth-Petersen P et al. International applicability of the VF-14. An Index of Visual Function in Patients with Cataracts. *Ophthalmology* 1997; 104(5):799-807.
- (28) Espallargues M, Alonso J. Effectiveness of cataract surgery in Barcelona, Spain site results of an international study. Barcelona I-PORT investigators. International Patient Outcomes Research Team. *J Clin Epidemiol* 1998; 51(10):843-852.
- (29) Andrich D, de Jong JHAL, Sheridan BE. Diagnostic opportunities with the Rasch model for ordered response categories. In: Rost J, Kangeheine R, editors. *Applications of latent trait and latent class models in the social sciences*. 1997: 59-70.
- (30) Bigsteps: Rasch Analysis for all two-facet models. 1997.
- (31) Dawson TL. Moral and evaluative reasoning across the Life span. *Journal of Applied Measurement* 2000, 1(4): 346-371
- (32) Wright BD, Linacre JM. A user's guide to BIGSTEPS: Rasch-Model Computer Program, version 2.7. Chicago: MESA Press, 1997.
- (33) Linacre JM. Prioritizing misfit indicators. *Rasch Measurement Transactions* 1995, 9(2), 1992: 422-423.
- (34) Apostol TM. *Calculus*. 2 ed. Barcelona: Ed.Reverté, 1973
- (35) Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-174.
- (36) Fisher WP, Jr. Foundations for health status metrology: the stability of MOS SF-36 PF-10 calibrations across samples. *J LA State Med Soc* 1999; 151(11):566-578.
- (37) Prieto L, Alonso J, Lamarca R, Wright BD. Rasch measurement for reducing the items of the Nottingham Health Profile. *J Outcome Meas* 1998; 2 (4):285-301.
- (38) Morales LS, Reise SP, Hays RD. Evaluating the equivalence of health care ratings by whites and Hispanics. *Med Care* 2000; 38(5):517-527.
- (50) Deyo RA, Diehr PD, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Controll Clin Trials* 1991, 12 (Supl): 142-158
- (60) Wright BD, Linacre JM. A user's guide to Bigsteps Rasch Model Computer Program, version 2.7. Chicago: Mesa Press, 1997.

Table 1. Visual Function Index VF-14: items and response options.

VISUAL FUNCTION INDEX VF-14

Do you have any difficulty , even with eyeglasses...

01. reading small print, such as labels on medicine bottles, a telephone book, food labels? (*reading small print*)
02. reading a newspaper or a book? (*reading a newspaper*)
03. reading a large-print book or a large-print newspaper or numbers on a telephone? (*reading large print*)
04. recognizing people when they are closed to you? (*recognizing people*)
05. seeing steps, stairs or curbs? (*seeing steps*)
06. reading traffic signs, street signs? (*reading traffic signs*)
07. doing fine handwork like sewing, knitting, crocheting, carpentry? (*doing fine handwork*)
08. writing checks or filling out forms? (*writing checks*)
09. playing games such as bingo, dominos, card games, mahjong? (*playing games*)
10. taking part in sports like bowling, handball, tennis, golf? (*taking part in activities*)
11. cooking? (*preparing meals*)
12. watching TV? (*watching TV*)
13. driving during the day because of your vision? (*driving during the day*)
14. driving at night because of your vision? (*driving at night*)

Response options:

4. No difficulty
 3. A little difficulty
 2. A moderate amount of difficulty
 1. A great deal of difficulty
 0. Unable to do the activity
 5. Not applicable
-

Table 2. VF-14 Index Rasch rating scale modeling: item parameters estimates and Goodness-of-fit statistics (n=184).

Item	Difficulty*	S. Error	95% CI	Infit	Outfit	Minimal	Cluster
				MNSQ	MNSQ	ability	
<i>Driving at night</i>	84.54	2.75	79.15-89.93	1.58	1.57	97.18	1
<i>Driving during the day</i>	65.10	2.30	60.59-69.61	1.30	1.15	77.74	2
<i>Reading small print</i>	62.71	1.21	60.34-65.08	0.92	0.80	75.35	2
<i>Doing fine handwork</i>	54.77	1.15	52.52-57.02	0.98	0.86	67.41	3
<i>Reading a newspaper</i>	51.90	1.16	49.63-54.17	1.05	0.91	64.54	3
<i>Writing checks</i>	50.16	1.43	47.36-52.96	1.01	0.82	62.80	3
<i>Reading traffic signs</i>	44.43	1.15	42.18-46.68	1.04	0.97	57.07	4
<i>Seeing steps, stairs</i>	41.33	1.15	39.08-43.58	1.12	1.51	53.97	4
<i>Watching TV</i>	40.82	1.16	38.55-43.09	0.73	0.73	53.46	4
<i>Taking part in sports</i>	39.65	1.45	36.81-42.49	1.11	0.95	52.69	4
<i>Playing games</i>	37.69	1.65	34.46-40.92	0.91	0.75	50.33	4
<i>Preparing meals</i>	29.14	1.59	26.02-32.26	0.84	0.77	41.78	5
<i>Reading large print</i>	26.57	1.51	23.61-29.53	1.03	0.65	39.21	5
<i>Recognizing people</i>	23.44	1.55	20.40-26.48	1.09	0.85	36.08	5

* The difficulty indicates the minimal ability needed to perform the item with 'No difficulty'.

CI: Confidence Interval

Figure 1. 'Ruler' approach for interpreting VF-14 Rasch scores. The line represents the ability level of patient with a Rasch score of 71.

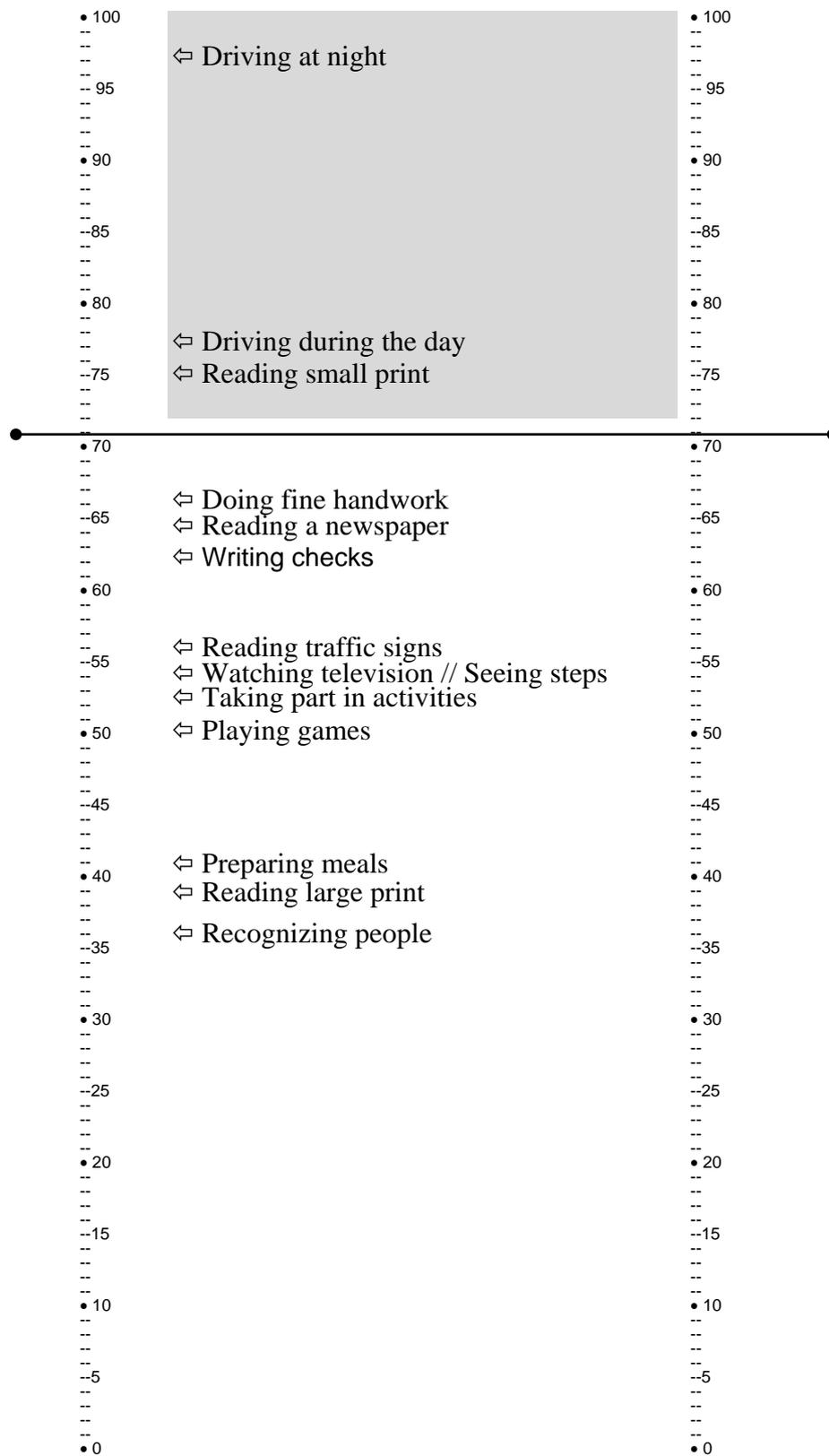


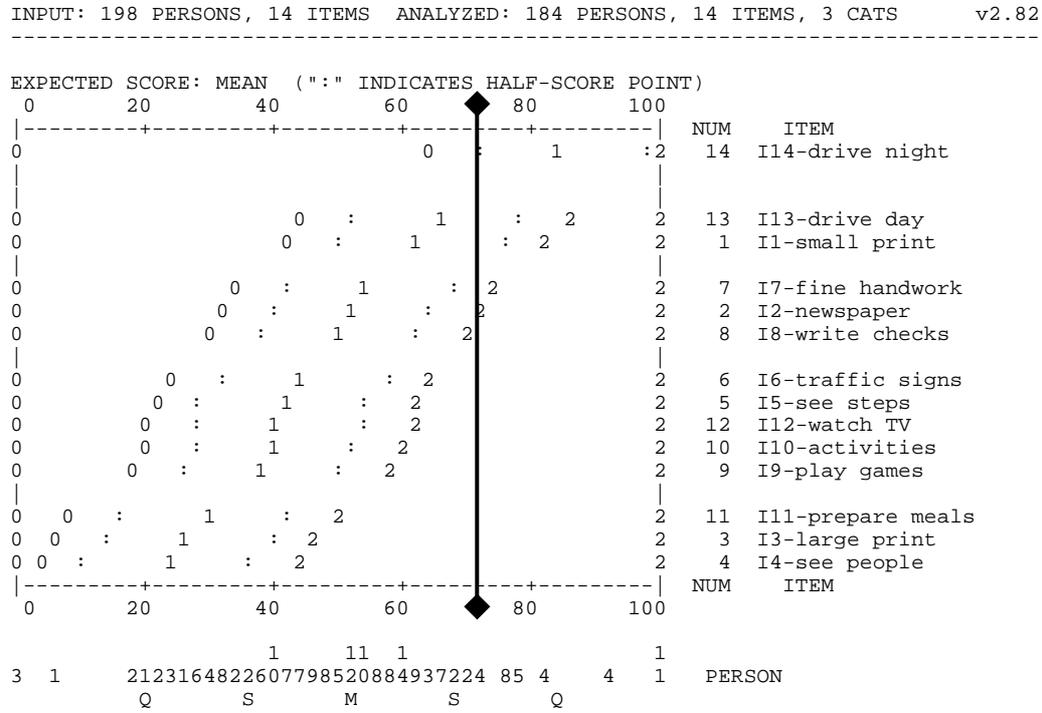
Figure 2. 'Clinical scenario' approach to interpretation of the VF-14, based on the expected performance for selected (benchmark) questionnaire items.*

VF-14 Score	Driving at night	Reading small print	Doing fine handwork	Watching Tv	Recognizing people
100	-	-	-	-	-
95	+	-	-	-	-
90	+	-	-	-	-
85	+	-	-	-	-
80	+	-	-	-	-
75	+	+	-	-	-
70	++	+	-	-	-
65	++	+	+	-	-
60	++	+	+	-	-
55	++	+	+	-	-
50	++	++	+	+	-
45	++	++	+	+	-
40	++	++	++	+	-
35	++	++	++	+	+
30	++	++	++	+	+
25	++	++	++	++	+
20	++	++	++	++	+
15	++	++	++	++	+
10	++	++	++	++	++
5	++	++	++	++	++
0	++	++	++	++	++

Expected answer:
- No Difficulty
+ Some difficulty
++ Unable to do it

* (Footnote) The clinical scenario for a patient with a VF-14 Rasch score of 71 has been highlighted.

Appendix Figure: Expected mean vf-14 score for given patient visual abilities.



Artículo 3. “The Impact of the Vf-14 Index, A Perceived Visual Function Measure, In The Routine Management Of Cataract Patients”. Valderas JM, Rue M, Guyatt G, Alonso J. Quality of Life Research 2005 (in press).

Antecedentes: La evaluación del impacto de la administración de información sobre medidas centradas en los pacientes a los médicos clínicos ha arrojado resultados contradictorios.

Métodos: En 1999-2000, se evaluó con el Índice de Función Visual VF-14 la función visual de 833 pacientes consecutivos con cataratas que acudieron de forma consecutiva a las consultas de 19 oftalmólogos en centros hospitalarios y de atención primaria, públicos y privados de España. Este estudio antes-después incluyó a) una sesión educativa y b) la administración posterior de la puntuación VF-14 de todos los pacientes. Se construyeron modelos lineales y logísticos de efectos mixtos para evaluar el impacto en el proceso de atención (correlación entre las valoraciones de la función visual proporcionadas por los oftalmólogos y los pacientes y la adecuación de la recomendación sobre el tratamiento quirúrgico) y su resultado (satisfacción).

Resultados: El coeficiente de regresión ajustado de la puntuación VF14 se incrementó significativamente tras la intervención como predictor de la valoración realizada por el oftalmólogo (coeficiente (β): control 0.10 vs intervención 0.35, $p < 0.05$). La intervención no incrementó la adecuación de la decisión médica (OR= 0.90; 95%CI [0.42; 2,69]) ni modificó la satisfacción con la atención sanitaria.

Conclusiones: La formación y la administración de rutina de las puntuaciones VF-14 de los pacientes incrementa el acuerdo entre pacientes y médicos sobre la capacidad funcional. La falta de efecto beneficioso en la gestión clínica y el resultado apunta la necesidad de una intervención más intensa para modificar la práctica médica.

Original research

The impact of the VF-14 index, a perceived visual function measure, in the routine management of cataract patients

J.M. Valderas¹, M. Rue¹, G. Guyatt² & J. Alonso^{1,3}, for the Systematic Use of Quality of Life Measures in the Clinical Practice Working Group*

¹Health Services Research Unit, Institut Municipal d'Investigació Mèdica, Barcelona, Spain (E-mail: JALONSO@IMIM.ES); ²Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada; ³Universitat Autònoma de Barcelona, Barcelona, Spain

Accepted in revised form 31 January 2005

Abstract

Background: Evidence about the impact of routine feedback of patient-reported outcomes is contradictory, and there is limited information regarding its use in the routine management of cataract patients. **Methods:** The VF-14 Index was used to assess the visual function of 833 consecutive cataract patients, attending 19 ophthalmologists from public and private hospitals and primary care practices in Spain, in 1999–2000. In this 'before and after' trial, the intervention included (1) an educational session, and (2) the provision of the VF-14 scores of all subsequent patients to the ophthalmologist. Mixed effects linear and logistic models were constructed to assess the effect on the process (correlation between patients' and physicians' assessments of visual function, appropriateness of surgery recommendation) and the outcome of care (satisfaction). **Results:** The adjusted regression coefficient for the VF-14 score significantly increased after the intervention as a predictor of the ophthalmologist's assessment of visual function (β coefficient: control 0.10 vs. intervention 0.35, $p < 0.05$). The intervention did not increase the probability of an appropriate medical decision (OR = 0.90; 95% CI: 0.42; 2.69) and it did not change patient satisfaction with care. **Conclusions:** Routine provision of education and feedback on the patient's VF-14 Index score significantly increases agreement between patients' and physicians' assessments of functional capacity. The lack of a beneficial effect on management or outcome suggests the need for a more intense intervention to change medical practice.

Key words: Cataracts, Disease management, Functional status assessment

Abbreviations: CI – confidence Interval; HRQL – health related quality of life; OVF – ophthalmologist's assessment of the patient's visual function; OR – odds ratio; SD – standard deviation; VF-14: Visual Function Index VF-14

Cataracts are the leading cause of visual impairment worldwide [1–5]. Cataract surgery is an

effective procedure [6–8], performed with increasing frequency, and which absorbs available resources and generates waiting lists [9]. Typically, practice guidelines for cataract management and surgery have strongly relied on two criteria for their recommendations: visual acuity (considered

*The Systematic Use of Quality of Life Measures in the Clinical Practice Working Group are: Alonso J, Castells X, Espallargues M, Prieto L, Valderas JM.

	Journal : QURE	Dispatch : 22-2-2005	Pages : 11
	CMS No. : DO00011745	<input type="checkbox"/> LE	<input type="checkbox"/> TYPESET
	MS Code : QURE 312R1	<input checked="" type="checkbox"/> CP	<input checked="" type="checkbox"/> DISK

41 as the objective criterion) and patient reported
42 visual function (subjective criterion) [10]. The most
43 recent practice guidelines acknowledge the pri-
44 macy of visual function [11–14], stating explicitly
45 that visual function and not visual acuity should
46 determine management [13–14]. Despite the
47 availability of valid health related quality of life
48 (HRQL) instruments, including the Visual Func-
49 tion Index VF-14 (VF-14) [15] and Activity of
50 Daily Vision Scale [16], guideline panels have not
51 made specific recommendations concerning mea-
52 surement and systematic assessment of visual
53 function in clinical practice.

54 Available evidence from systematic reviews
55 suggests that the routine use of HRQL measures
56 in clinical practice may, under some circum-
57 stances, improve the process of care, but not
58 outcome, although methodological limitations
59 may be responsible for failure to detect beneficial
60 effects [17–19]. The likelihood of beneficial effects
61 may increase if clinicians receive education in the
62 use of simple and easy to interpret reports of
63 disease specific instruments. Studies will be more
64 likely to detect beneficial effects if similar health
65 professionals constitute intervention and control
66 groups, and there is adequate control of con-
67 tamination [17–19].

68 Limited interpretability of HRQL scores poses an
69 important barrier to their introduction in clinical
70 practice [17, 20]. Investigators have recently devel-
71 oped novel methods to enhance interpretability of
72 these measures [21]. We took advantage of these
73 developments to study the effect of systematically
74 providing physicians with standardized visual
75 function assessments of their cataract patients.

We anticipated an increase in patient-physician
agreement on visual function, and designated
agreement as the primary endpoint of the trial. In
our conceptual model (Figure 1), increased agree-
ment may facilitate changes in the process of care
(specifically, the appropriateness of the medical
decision) and, eventually, changes in patient sat-
isfaction with care. We, therefore, also examined
the apparent effect of the intervention on the
appropriateness of the medical decision (recom-
mendation of surgery), and patient satisfaction
with care (secondary endpoints).

Materials and methods

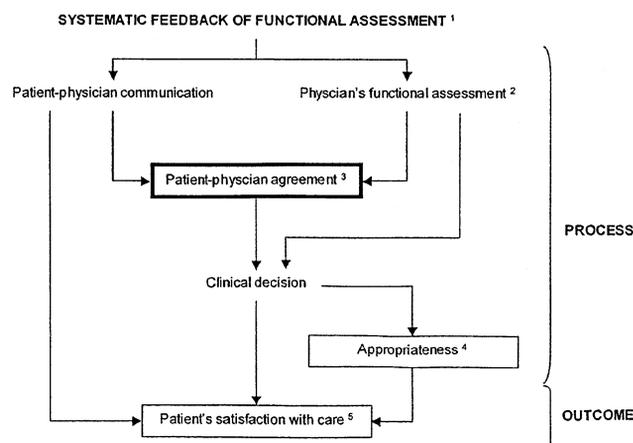
Study design

In this before/after study [22], we enrolled all
consecutive cataract patients attending the partic-
ipating ophthalmology practices who consented to
interviews by trained interviewers before the
medical visit. During the control (*before*) period,
patients received standard care and the ophthal-
mologists did not receive any additional informa-
tion on visual function. Subsequently, the
clinicians participated in an educational session
and then, during the intervention (*after*) period,
received feedback on the VF-14 scores of all their
subsequent patients.

Setting and period

A convenience sample of six ophthalmology
practices in five cities (Barcelona, Terrassa, Rubí,

Figure 1. Conceptual model of the impact of the VF-14 feed-back intervention study



Applying the model to the impact of feedback of the VF-14 Index study: 1. Individual VF-14 scores; 2. Ophthalmologist's rating of visual disability; 3. Agreement on visual disability (primary endpoint); 4. Appropriateness of surgery recommendations (secondary endpoint); 5. Patient's satisfaction with care (secondary endpoint)

105	Viladecans, Sant Boi) in Catalonia, Spain, partic-	little difficulty', 'Some difficulty', 'A moderate	151
106	ipated in the study. Three practices were hospital	amount of difficulty', 'A great deal of difficulty').	152
107	based (one of which was private and the other two	A Spanish version of the VF-14 has proved	153
108	public) and the other three were public, primary	reliable, valid and responsive [24].	154
109	care based settings. We contacted 2 additional	We applied the Rasch model, a mathematical	155
110	private, primary care based practices, but they	model derived from Item Response Theory, to the	156
111	declined to participate, and a third practice was	VF-14 to obtain: (1) scores ranging from 100 (best	157
112	not able to provide the minimum number of	visual function) to 0 (worst visual function), and (2)	158
113	patients required during the control period (10	to develop two interpretation aids: a 'ruler aid'	159
114	patients), and was therefore excluded. Practices	(derived from the most difficult activity (item) a	160
115	refusing to participate did not specify their reasons	patient is able to do without difficulty), and a	161
116	for declining.	'clinical scenarios aid', that provided estimates of	162
117	The control period lasted from May 1999 to	the mean expected difficulty in performing selected	163
118	September 1999, and the intervention period lasted	benchmark items [21]. Patients completed the	164
119	from October 1999 to January 2000. The Clinical	VF-14 after arriving at the clinic, while waiting to	165
120	Research Ethical Committee of the Municipal	see the ophthalmologist. Computer assisted	166
121	Institute of Health Care (CEIC-IMAS) approved	administration by trained interviewers proceeded	167
122	the study.	in convenient privacy and provided immediate	168
123	<i>Participants: physicians and patients</i>	calculation of Rasch scores [25].	169
124	We were able to enrol all the ophthalmologists	<i>Intervention</i>	170
125	(N = 19) working in the selected practices in the	The ophthalmologists attended a 25-min, com-	171
126	study. Ophthalmologists offered all of their con-	puter-assisted multimedia educational session that	172
127	secutive cataract patients the opportunity to par-	described the content, scoring and interpretation of	173
128	ticipate.	the VF-14. A member of the working group (JMV,	174
129	We included patients who were 50 years of age	LP, JA) conducted these sessions for small groups	175
130	and older, who had a diagnosis of cataracts con-	or for a single clinician at the clinicians' clinic. The	176
131	firmed by the participating ophthalmologist, and	ophthalmologists received an interpretation man-	177
132	for whom the ophthalmologist had not yet made a	ual, two quick reference cards summarizing the	178
133	management decision (whether to recommend sur-	manual, and a booklet with relevant articles on the	179
134	gery or not). The only exclusion criterion was having	validity of the VF-14 [15, 23, 24]. The information	180
135	undergone cataract surgery. No patient participated	included neither a threshold for VF-14 suggesting	181
136	in both the control and intervention periods. These	severe visual functional disability, nor any specific	182
137	patients might have attended the public clinics as a	recommendations regarding surgery, nor any spe-	183
138	result of a referral from either primary care or hos-	cific instructions for discussion with patients.	184
139	pital medical services, or might be already under	Beginning the day after the educational inter-	185
140	observation by the ophthalmologists for cataracts,	vention, the clinician received immediate feedback	186
141	or another ophthalmologic condition. Patients	for each patient, by means of a sheet that included	187
142	attending private practices might also have de-	the patient's VF-14 score supported with a ruler-	188
143	manded care on their own initiative.	like graph (Figure 2). We have provided a com-	189
144	<i>VF-14 Index</i>	prehensive description of this reporting method	190
145	The Visual Function Index VF-14 is a measure of	elsewhere [21].	191
146	perceived visual function based on 14 everyday	<i>Measurements</i>	192
147	activities that may be affected by cataracts (e.g.,	The ophthalmologists rated patients' visual func-	193
148	reading a newspaper, watching TV, or taking part	tion using two 10 cm long self-administered visual	194
149	in activities) [23]. Respondents rate their difficulty	analogue scales. They rated the patient's 'Difficulty	195
150	in performing these activities ('No difficulty', 'A	helping himself/herself due to cataracts' and their	196

Name: <u>C. F. S.</u>	VF-14	ID: <u>3 9 6 2</u>
Score: <u>71</u>		

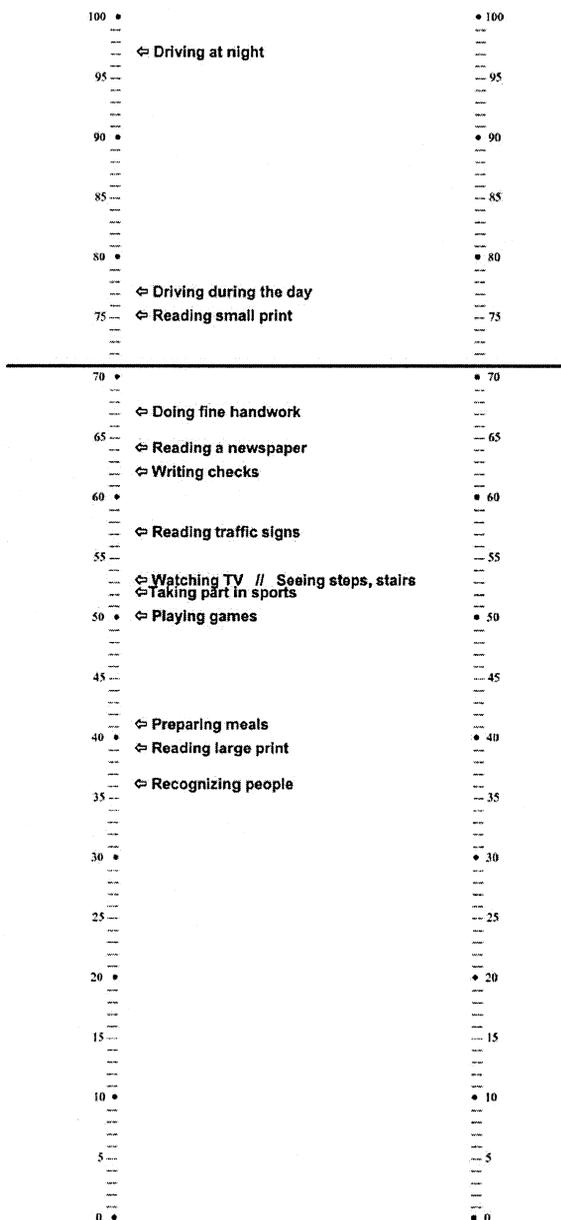


Figure 2. Feedback form (VF-14 score).

197 'Difficulty in performing everyday activities due to
198 the cataracts'. Given the high correlation observed

between these two variables (Spearman's ρ 0.95,
 $p < 0.001$), we constructed, through linear com-

199
200

201 bination, a composite variable 'Ophthalmologist
202 assessment of the patient's visual function' (OVF)
203 ranging from 0 (minimal OVF) to 100 (maximal
204 OVF).

205 Ophthalmologists provided information on
206 whether they had recommended cataract surgery
207 or not, a decision that is typically made during a
208 single visit. Ophthalmologists also specified the
209 optimal timing of surgery (within the next
210 4 months, or later).

211 The ophthalmologists collected clinical infor-
212 mation with standardized forms (cataract type,
213 best corrected visual acuity using Snellen charts
214 (decimal fractions), *fundus* examination, and ocu-
215 lar and medical comorbidities) [10–14].

216 Patients provided, in a telephone interview
217 2 weeks after the visit to the ophthalmologist, so-
218 cio-demographic and clinical information, as well
219 as their level of satisfaction with care. We assessed
220 satisfaction with care using a modified short ver-
221 sion of the Group Health Association America
222 Satisfaction Battery GHAA [26], including items
223 in the following original dimensions: Technical
224 Quality, Communication, Interpersonal Care, and
225 Outcomes. Summary scores ranged from 0 (mini-
226 mal satisfaction) to 100 (maximal satisfaction).
227 The VF-14 Index was re-administered during this
228 telephone interview to test the reliability of this
229 administration alternative.

230 We used conservative criteria for appropriateness
231 of surgery recommendations, based on expert con-
232 sultation (see Appendix 1 for definition of approp-
233 riateness criteria) [27]. Surgery recommendations
234 were classified as appropriate, not appropriate, or
235 of uncertain appropriateness. We excluded the lat-
236 ter from the appropriateness analysis.

237 *Statistical analysis*

238 Bivariate analyses for the characteristics of
239 patients in both periods included the Rao and
240 Scott corrected Pearson statistic for categorical
241 variables [28], and the adjusted Wald test for
242 continuous variables [29]. We performed all
243 bivariate analyses for differences between periods
244 in all outcomes at the ophthalmologist level, and
245 we assessed differences between periods with the
246 Wilcoxon Ranks Test for repeated measures.

247 Given the hierarchical nature of our data
248 (patients nested within ophthalmologists,

ophthalmologists nested within practices) we used 249
multivariate linear or logistic mixed effects models, 250
depending on whether the dependent variable was 251
continuous or dichotomous [30]. We assessed the 252
marginal significance of each fixed effect when all 253
other fixed effects (covariates) were present in the 254
model with conditional *t*-tests based on the 255
restricted maximum likelihood estimate of the 256
variance and the significance of random effects 257
with likelihood ratio tests. Plots of the residuals vs. 258
the predicted values and by quantile–quantile 259
(*q–q*) plots of the random effects allowed assess- 260
ment of normality assumptions of both the resid- 261
uals and the random effects [30]. 262

263 We fitted a regression model, instead of a con-
264 cordance model, in order to describe patient-phy-
265 sician agreement on patients' perceived visual
266 function, because of the different measurement
267 scales of the VF-14 Index and OVF. We included
268 VF-14 and their interaction (period *
269 VF-14) as independent variables in this model,
270 with 'OVF' as the dependent variable. Statistical
271 significance of the coefficient for the interaction
272 term was interpreted as a differential agreement
273 between patients (VF-14 scores) and ophthalmol-
274 ogists ('OVF') across periods (intervention vs.
275 control) [31].

276 Further analysis included a reliability assess-
277 ment of telephone administration of the VF-14
278 Index with the Intraclass Correlation Coefficient
279 [32], as well as internal consistency evaluation of
280 the modified version of the GHAA with
281 Cronbach's α [33].

282 We used the R statistical package [34] for mixed
283 effects modelling, and STATA [35] for all other
284 analyses.

Results 285

286 In total, we recruited 833 patients (control 379,
287 intervention 454); no one refused to participate.
288 We were able to collect re-interview data for
289 94.5% of control and 93.0% of intervention
290 patients. Patients in the control and intervention
291 groups had similar socio-demographic and clinical
292 profiles (Table 1). VF-14 Index scores at recruit-
293 ment were higher for control than for intervention
294 patients (control 62.9 vs. intervention 66.7,
295 $p = 0.032$), but not 2 weeks thereafter (control



Table 1. Socio-demographic and clinical characteristics of patients in control and intervention periods.

	Control (n=379)	Intervention (n=454)
Age	71.7 (8.53)	72.7 (8.24)
Sex (female) [n (%)]:	225 (58.9)	247 (54.3)
Education [n (%)]:		
None	85 (24.1)	82 (19.3)
Primary	200 (56.7)	268 (63.9)
Secondary	44 (12.4)	49 (11.6)
University	24 (6.8)	25 (5.9)
Employment status [n (%)]:		
Paid	20 (5.7)	28 (6.6)
Unemployed	2 (0.6)	2 (0.5)
Retired	229 (65.1)	301 (71.2)
Disabled	22 (6.5)	13 (3.1)
Housework	71 (20.2)	75 (17.1)
Other	8 (2.3)	4 (0.9)
Caregiver of others [n (%)]	60 (17.0)	51 (12.0)
Living alone [n (%)]	74 (21.0)	82 (19.3)
Private insurance [n (%)]	75 (21.2)	83 (19.6)
Visual acuity:		
worst eye	0.32 (0.26)	0.32 (0.24)
best eye	0.53 (0.27)	0.52 (0.25)
both eyes	0.52 (0.28)	0.53 (0.23)
VF-14 Index:		
Recruitment*	62.9 (18.7)	66.7 (17.2)
Reinterview	64.3 (19.4)	65.3 (17.5)
Perceived Health Status [n (%)]:		
Excellent	19 (5.4)	18 (4.2)
Very good	43 (12.1)	54 (12.7)
Good	148 (41.7)	207 (48.8)
Fair	108 (30.4)	111 (26.2)
Poor	37 (10.4)	34 (8.0)

All variables presented as mean and standard deviation unless stated otherwise ([n (%)]).

* $p < 0.05$, all other comparisons not significant at $p < 0.05$ level.

296 64.3 vs. intervention 65.3, $p = 0.423$). Ophthalmologists mean age was 37.6 (standard deviation
297 (SD) 8.7) and 31.6% of them were female. Ten
298 ophthalmologists attended group educational sessions and nine attended individual sessions.

301 Administration of the VF-14 immediately before
302 the visit took a median of 10 min (Interquartile
303 range: 7–11 min). The test–retest Intraclass Correlation Coefficient for the VF-14 was 0.84 (95%
304 CI: 0.82, 0.86). Correlations for the VF-14 score and visual acuity were similar in both periods
305 (Spearman $r = 0.60$ for visual acuity in the best eye, 0.56 for visual acuity in the worst eye), and so
306 was the correlation between OVF and desired timing for surgery (Spearman $r = 0.47$; 95% CI:
307 0.41, 0.52). The modified GHAA had high internal consistency (Cronbach's $\alpha = 0.95$).

The Spearman correlation of patient's perceived (VF-14) and ophthalmologists assessed (OVF) visual function increased significantly after the intervention (control: 0.29, intervention: 0.45, $p < 0.05$) (Table 2). In the linear mixed effects model the coefficient for the interaction term was statistically significant (Table 3, *Agreement*), increasing the adjusted coefficient for VF-14 score from 0.10 in the control period to 0.35 (=0.10 + 0.25) in the intervention period. In other words, the contribution of the VF-14 score to the adjusted ophthalmologist's assessed visual function was increased by the factor 3.5.

In this model, and in all others as well, there were statistically significant practice and ophthalmologist effects, but no statistically significant effect for any explanatory variable at the

	Journal : QURE	Dispatch : 22-2-2005	Pages : 11
	CMS No. : DO00011745	<input type="checkbox"/> LE	<input type="checkbox"/> TYPESET
	MS Code : QURE 312R1	<input checked="" type="checkbox"/> CP	<input checked="" type="checkbox"/> DISK

Table 2. Effect of the VF-14 feed-back intervention at the ophthalmologist level (n=19)

	Control		Intervention		p value ^a
	Median	Interquartile Range	Median	Interquartile Range	
Correlation VF-14 score-attributed visual function (Spearman coefficient)	0.29	0.17-0.47	0.45	0.34-0.72	0.036
Appropriate medical decision (%) ^b	0.50	0.13-1.0	0.50	0.19-1.0	0.722
Satisfaction with care (score) ^b	61.00	56.25-75.25	59.75	56.54-64.75	0.184

^a Wilcoxon Ranks Test for repeated measures^b Estimates are median values of variables aggregated at the ophthalmologist level**Table 3.** Mixed effects models for all primary and secondary outcomes

Model	Agreement ^a		Appropriateness ^b		Satisfaction ^c	
	β	95% CI	β	95% CI	β	95% CI
<i>Patient level</i>						
Constant	40.08	(22.94; 57.22)	0.51	(-1.42; 4.77)	54.81	(40.48; 69.14)
Period (intervention = 1)	-18.72	(-27.37; -10.07)	-0.10	(-0.84; 0.64)	-1.88	(-4.45; 0.69)
VF-14 score (0-100)	0.10	(-0.12; 0.32)	-	-	0.07	(-0.03; 0.17)
Worst eye visual acuity (0-1)	9.18	(2.06; 16.30)	-	-	6.71	(-1.05; 14.47)
Interaction VF-14 ^a period	0.25	(0.12; 0.38)	-	-	-	-
Random effects	SD	95% CI	SD	95% CI	SD	95% CI:
<i>Practice level</i>						
Constant	11.30	(4.15; 30.78)	1.16	(0.37; 3.67)	4.13	(1.85; 9.25)
VF-14 score	0.22	(0.09; 0.53)	-	-	-	-
<i>Ophthalmologist level</i>						
Constant	9.47	(1.96; 45.79)	1.20	(0.65; 2.21)	2.39	(0.92; 6.15)
VF-14 score	0.12	(0.02; 0.84)	-	-	-	-
Residual	15.68	(14.89; 16.51)	0.89	(0.78; 1.01)	14.99	(14.11; 15.91)

^aAgreement model adjusting also for sex, age, visual acuity in the best eye, cataract type, medical and ocular comorbidity. Dependent variable: OVF; ^bAppropriateness model adjusting also for sex and age. Dependent variable: appropriateness of surgery recommendation (reference category: non-appropriate); ^c Satisfaction model adjusting also for sex, age, education, cataract type, medical and ocular comorbidity and being recommended surgery. Dependent variable: modified GHAA score.

330 ophthalmologist or the practice level (Table 3).
 331 The study did not have enough statistical power to
 332 detect these effects. The participating ophthalmol-
 333 ogists recommended surgery within four months
 334 for slightly fewer patients in the intervention per-
 335 iod (19%) than in the control period (21%)
 336 ($p = 0.039$, Wilcoxon Ranks Test for repeated
 337 measures). Appropriateness of surgery recom-
 338 mendations was not different in the two periods,
 339 both in the bivariate analysis (Table 2), and in the
 340 logistic mixed effects model (odds ratio for period
 341 ($OR_{\text{period}} = \exp(\beta_{\text{period}}) = 0.91$; 95% CI: 0.43,
 342 1.91) (Table 3). The statistical power for the mul-
 343 tivariate analysis was small (β (Type II
 344 error) = 0.45 for observing $OR \geq 2$); this might
 345 be attributed to a high proportion of all surgery

recommendations being classified as uncertain
 (69.8%). In the sensitivity analysis estimates of
 effect on appropriateness showed high variability
 both in size and direction, without any clear trend.

Neither the bivariate nor the linear mixed effects
 analyses demonstrated statistically significant dif-
 ferences between intervention and control groups
 with regard to patients' satisfaction with care
 (Tables 2 and 3, *Satisfaction*).

Our conceptual model predicted that appropri-
 ateness of surgery recommendation would
 improve for physicians, increasing their agreement
 with patients on visual disability. Fourteen
 ophthalmologists (73.7%) demonstrated the
 hypothesized increase in agreement on visual
 disability ($\rho_{\text{intervention}} > \rho_{\text{control}}$), but a decrease in

	Journal : QURE	Dispatch : 22-2-2005	Pages : 11
	CMS No. : DO00011745	<input type="checkbox"/> LE	<input type="checkbox"/> TYPESET
	MS Code : QURE 312R1	<input checked="" type="checkbox"/> CP	<input checked="" type="checkbox"/> DISK

362 the proportion of inappropriate medical decision
 363 ($P_{\text{intervention}} < P_{\text{control}}$) was observed in physicians
 364 both with and without increased agreement
 365 (42.9% vs. 40.0%, $p = 0.9$).

366 Discussion

367 This “before and after” study, with an interven-
 368 tion consisting of education and feedback, showed
 369 significantly increased agreement between patients’
 370 and physicians’ assessments of visual disability
 371 after the intervention. The study demonstrated no
 372 change in other variables related to outcome or the
 373 process of patient care.

374 Effect on outcomes variables

375 The observed improvement in agreement between
 376 the ophthalmologist’s assessment of functional
 377 disability and the patient’s reported visual dis-
 378 ability was substantial in the bivariate analysis
 379 (from a *weak* correlation in the control period to a
 380 *moderate* one after the intervention) [36]. The
 381 regression model also reflected this substantially
 382 improved agreement (the contribution of the
 383 VF-14 score to the adjusted ophthalmologist’s
 384 assessed visual function was increased by a factor
 385 of 3.5).

386 While this improvement in the process of health
 387 care is consistent with previous research [17–20],
 388 investigators have not previously addressed the
 389 issue in cataract patients. Perceived functional
 390 disability is a key measurement informing cataract
 391 management. Its systematic assessment would
 392 provide ophthalmologists with a valid and reliable
 393 measure to complement unstructured patient
 394 interviews, and thus facilitate effective patient-
 395 centred management.

396 We chose a relatively non-intensive intervention
 397 to isolate the effects of feedback from the eventual
 398 effects of additional components in a more com-
 399 plex intervention (e.g. including the provision of
 400 specific guidelines). Previous research is consistent
 401 with changes in practice resulting from more
 402 intensive interventions [17–20]. For instance, the
 403 low intensity of our intervention may have limited
 404 the observed effect on secondary outcomes. Our
 405 very conservative definition of appropriateness

might have also limited the chances of detecting
 any effect on this variable, aside from the fact that
 we lack previous information regarding the valid-
 ity of our appropriateness criteria (see the
 Appendix).

The lack of any observed effect on patient sat-
 isfaction is consistent with previous research [17].
 An increase in patient satisfaction with care at this
 early point of cataract management, before the
 surgery, might be improbable. The satisfaction
 ratings suggest, however, that completing the
 VF-14 and receiving feedback did not cause dis-
 satisfaction. Thus, the results provide some evi-
 dence that patients find routine feedback
 acceptable.

Our results established a link between the doc-
 tor’s assessment of visual disability and desired
 timing for surgery, but showed no association
 between agreement and appropriateness of the
 medical decision. The results are inconsistent with
 the expected cascade of effects and therefore
 challenge our conceptual model.

Finally, a variety of factors (such as professional
 experience, individual patient characteristics or
 available resources) may also influence practice,
 and may have a greater impact on appropriateness
 than this study’s experimental intervention [37].

Strengths and limitations

Our study had some advantages over previous
 assessments. Sample size was large and there was
 sufficient variability in physicians and practice
 settings (size, founding and orientation (primary
 care based vs. hospital based)). Also, we followed
 specific recommendations from previous system-
 atic reviews. Careful statistical analysis reflected
 the hierarchical nature of the data, a key aspect
 seldom considered in feedback assessment [38].
 Previous, well-designed assessments included
 enhanced complex interventions intended to max-
 imize the chances of detecting an impact. This
 approach, however, provides no possibility of
 measuring the effect of individual components of
 the intervention. The present study, on the con-
 trary, was specifically designed to measure the
 effects of feedback.

The main limitation of the study arises from the
 absence of a control group of different physicians
 and patients not receiving the intervention. Cluster

454 randomization of physicians or practices was
 455 considered inappropriate for this study due to the
 456 small sample of practices and their different char-
 457 acteristics. Randomization of ophthalmologists
 458 within practices or patients attending the same
 459 professional would have created an unacceptable
 460 risk of contamination [17].

461 Methodological concerns regarding a “before
 462 and after” study include the possibility that
 463 other factors operating in the study context were
 464 actually responsible for the effects of the inter-
 465 vention [22]. We were not able to identify any
 466 salient change in the environment, such as the
 467 publication of practice guidelines or the approval
 468 of regulatory initiatives concerning cataract
 469 management immediately before or during the
 470 study period. When consulted, none of the par-
 471 ticipating ophthalmologists was aware of such an
 472 intervention. Still there is always the chance that
 473 observed changes were due to some unidentified
 474 confounding variable. For instance, we believe
 475 that factors other than the intervention were
 476 responsible for the decrease in surgery indication
 477 rates, though we are unable to specify what these
 478 factors are.

479 Statistically significant differences in VF-14
 480 scores were observed for patients in the control
 481 and intervention periods. Scores elicited from
 482 patients by telephone at a 2-week follow-up did
 483 not show differences between control and inter-
 484 vention groups, and still reliability for both
 485 administrations of the VF-14 was high [0.84]. This
 486 strengthens our interpretation that the observed
 487 differences are not clinically important.

488 Finally, we included a small number of prac-
 489 tices, all of them from the same regional area.
 490 Although very representative of cataract manage-
 491 ment in Catalonia, Spain, generalizability of our
 492 results might thus be limited.

493 *Conclusion*

494 The results presented here demonstrate a limited
 495 beneficial effect of systematic feedback of VF-14
 496 scores. A cluster randomized study, including a
 497 more intensive intervention addressing the provi-
 498 sion of specific recommendations for cataract
 499 management to health professionals, would pro-
 500 vide stronger inferences for clinical practice. If our
 501 results are confirmed, routine assessment of visual

disability with the VF-14 Index may improve 502
 management of cataract patients. 503

Aknowledgements 504

This study was supported by *Fondo de Investiga- 505*
ción Sanitaria (97/1135), Red de Investigación 506
Cooperativa IRYSS (G03/202). Additional sup- 507
port was provided by Instituto de Salud Carlos III 508
(98/4471), Generalitat de Catalunya (CIRIT 1999 509
SGR 00240). 510

We would like to thank all enrolled ophthal- 511
 mologists (Drs Casado, Cervantes, González, 512
 Folch, Fossas, Hasanein, Juvenet, Llevat, Manza- 513
 no, Mondéjar, Montermoso, Musavi, Raitieri, 514
 Rilo, Salvador, Sedó, Sellarés, Torras, Vendrell, 515
 Vergés) for their participation. We also thank 516
 Dave McFarlane, Rosalia Santed and Gemma 517
 Vilagut for providing technical support, and 518
 Natalia Molina, Cristina Pipó, Carolina Rebato 519
 and Silvia Sarri for their help in data collection. 520

Appendix 1 521
Definition of appropriate medical decision regarding 522
cataract surgery. 523

Available cataract surgery appropriateness criteria [27] were 524
 not considered applicable due to changes in surgical procedures 525
 since their development. We reviewed the literature, consulted 526
 experts and reached a consensus for the definition of specific 527
 criteria. The medical decision was finally considered as 528
 ‘appropriate’, ‘not appropriate’ or ‘uncertain’ by comparing the 529
 recommendation made by the ophthalmologist (whether surgery 530
 had to be performed within the next 4 months) with the 531
 patient’s visual acuity in the worst eye (Snellen decimal fraction 532
 measurements) and his/her VF-14 Index score (see Appendix 533
 Table). Appropriateness of the medical decision was assessed 534
 only for patients without ocular or medical comorbidity. 535

Mean waiting time for cataract surgery in Catalonia is 536
 now 5.3 months (in only one health region out of eight does 537
 it fall below 4 months [39]). This is consistent with evidence 538
 of mean waiting times in European countries lasting from 539
 7 weeks to 8 months [40], and it was the rationale for 540
 selecting the 4 month time frame to indicate immediate vs. 541
 delayed surgery recommendation. 542

We selected very conservative criteria. Immediate surgery 543
 was only considered appropriate if there was a very substantial 544
 impairment in terms of both visual acuity (less than a third of 545
 what is considered to be not impaired visual acuity and well 546
 below the tentative cut-off of 0.5 suggested by the practice 547
 guidelines) and visual function (complete perceived inability to 548
 perform some activities (like ‘driving at night’) and moderate 549

550 amount of difficulty in doing some relevant others (like 'doing
551 fine handwork'). Conversely immediate surgery was only
552 considered inappropriate if both visual acuity and visual func-
553 tion were preserved (visual acuity of more than 70% and almost
554 perfect VF-14 score).

555 Since a recommendation was classified as 'uncertain'
556 depending on a patient's visual acuity (VA) and visual function,
557 but regardless of actual medical advice, patients receiving
558 uncertain recommendations were excluded from the analysis
559 assessing the effect of the intervention on appropriateness.

560 References

- 561 1. Vision 2020. The global initiative. Fact Sheet 213. WHO.
562 2000.
- 563 2. Reidy A, Minassian DC, Vafidis G, et al. Prevalence of
564 serious eye disease and visual impairment in a North
565 London population: Population based, cross sectional
566 study. *BrMedJ* 1998; 316: 1643-1646.
- 567 3. Klaver C, Wolfs R, Vingerling JR, et al. Age-specific
568 prevalence and causes of blindness and visual impairment
569 in an older population. The Rotterdam Study. *Arch Oph-
570 thalmol* 1998; 116: 653-658.
- 571 4. Rodriguez J, Sanchez R, Muñoz B, et al. Causes of blind-
572 ness and visual impairment in a population-based sample of
573 U.S: Hispanics. *Ophthalmology* 2002; 109: 737-743.
- 574 5. VanNewkirk MR, Weih L, McCarty CA, et al. Causes of
575 incident visual impairment: The Blue mountains eye study.
576 *Arch Ophtalmol* 2002; 120: 613-619)
- 577 6. Woodcock M, Shah S, Smith RJ. Recent advances in cus-
578 tomising cataract surgery. *BrMedJ*. 2004; 328: 92-6.
- 579 7. Laidlaw DA, Harrad RA, Hopper CD, et al. Randomised
580 trial of effectiveness of second eye cataract surgery. *Lancet*.
581 1998 Sep 19; 352(9132): 925-929.
- 582 8. Espallargues M, Alonso J. Effectiveness of cataract surgery
583 in Barcelona, Spain site results of an international study. *J
584 Clin Epidemiol* 1998 Oct; 51(10): 843-852).
- 585 9. Taylor HR. Cataract: How much surgery do we have to do?
586 *Br J Ophthalmol* 200; 84: 1-2.
- 587 10. Cataract Management Guideline Panel. Cataract in Adults.
588 Management of functional impairment. Clinical Practice
589 Guideline, Number 4. Rockville MD: US Department of
590 Health and Human Services, Public Health Service, Agency
591 for Health Care Policy and Research. AHCPR pub no 93-
592 0542. 1993.
- 593 11. The Alberta Clinical Practice Guidelines Working Group.
594 Guideline for Surgical & Non-Surgical Management of
595 Cataract in the Otherwise Healthy Adult Eye. Edmonton,
596 AB Canada: The Alberta Clinical Practice Guidelines
597 Program. 2002.
- 598 12. American Academy of Ophthalmology, Anterior Segment
599 Panel. Cataract in the adult eye. San Francisco (CA):
600 American Academy of Ophthalmology (AAO), 2001; 62
601 p.
- 602 13. Scottish Intercollegiate Guidelines Network (SIGN). Day
603 case cataract surgery. A national clinical guideline. Edin-

- burgh (Scotland): Scottish Intercollegiate Guidelines Net-
work (SIGN); 2001 Aug. 20 p. (SIGN publication; no. 53)
14. The Royal College of Ophthalmologists. Cataract Surgery
Guidelines. [on-line] 2001. (http://www.rcophth.ac.uk/publications/guidelines/cataract_surgery.html)
15. Steinberg EP, Tielsch JM, Schein O, et al. The VF-14. An
index of functional impairment in patients with cataract.
Arch Ophthalmol 1994; 112: 630-638.
16. Mangione CM, Phillips RS, Seddon JM, et al. Develop-
ment of the activities of daily vision scale. A measure of
visual functional status. *Med Care* 1992; 30: 1111-1126.
17. Espallargues M, Valderas JM, Alonso J. Provision of
feedback on perceived health status to health care profes-
sionals. A systematic review of its impact. *Med Care* 2000;
38(2): 175-186.
18. Greenhalgh J, Meadows K. The effectiveness of the use of
patient-based measures of health in routine practice in
improving the process and outcomes of patient care: A lit-
erature review. *J Eval Clin Pract* 1999 Nov; 5(4): 401-416.
19. Gilbody SM, House AO, Sheldon T. Routine administration
of Health Related Quality of Life (HRQoL) and needs
assessment instruments to improve psychological outcome—a
systematic review. *Psychol Med* 2002 Nov; 32(8): 1345-1356.
20. Lydick E. Approaches to the interpretation of quality of life
scales. *Med Care* 2000; 38(Suppl 9): I1180-I1183.
21. Valderas JM, Alonso J, Prieto L, et al. "Content-based
approach for facilitating the interpretation of health-related
quality of life measures in clinical practice. A proposal for the
visual function index VF-14". *Qual Life Res* 2004; 13: 35-44.
22. Cook, TD, Campbell DT 1979. Quasiexperimentation:
Design and Analysis Issues for Field Settings. Boston:
Houghton Mifflin Co.
23. Cassard SD, Patrick DL, Damiano AM, et al. Reproduc-
ibility and responsiveness of the VF-14. An index of func-
tional impairment in patients with cataracts. *Arch
Ophthalmol* 1995; 113: 1508-1513.
24. Alonso J, Espallargues M, Folmer-Andersen T, et al.
International applicability of the VF-14. An index of visual
function in patients with cataracts. *Ophthalmology* 1997;
104(5): 799-807.
25. Prieto L, Comas M, Vilagut G (1999) Turbo Pascal
Measurement Program Rasch Measurement Transactions
12: 4, p. 672.
26. Davies AR, Ware JE. GHAA's Consumer Satisfaction
Survey and User's Manual. 2nd ed. Washington, DC:
Group Health Association America Inc., 1991.
27. Lee PP, Kamberg CJ, Hilborne LH, et al. 1993. Cataract
Surgery: A literature review and ratings of appropriateness
and cruciality. RAND, Santa Monica CA.
28. Rao JNK, Scott AJ. On chi-squared tests for multiway
contingency tables with cell proportions estimated from
survey data. *Ann Stat* 1984; 12: 46-60.
29. Eltinge JL, Sribney WM. Estimates of linear combinations
and hypothesis tests for survey data. *Stata Technical Bul-
letin* 1996; 31: 31-42. Reprinted in *Stata Technical Bulletin
Reprints*, Vol. 6, pp. 246-259.
30. Pinheiro JC, Bates DM. Mixed effects models in S and
S_PLUS., Statistics and Computing Series. New York:
Springer-Verlag, 2000.

604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662



Journal : **QURE**
CMS No. : **DO00011745**
MS Code : **QURE 312R1**

Dispatch : **22-2-2005**

Pages : **11**

LE TYPESET
 CP DISK

- 663 31. Kleinbaum DG, Kupper LL Muller KE. Applied Regression Analysis and Other Multivariable Methods. Duxbury Press: Belmont, CA, 1988. 685
- 664 32. Prieto L, Lamarca R, Casado A. Assessment of the reliability of clinical findings: The intraclass correlation coefficient. Med Clin (Barc) 1998 Feb 7; 110(4): 142-145. 686
- 665 33. Nunnally JC, Bernstein IH. Psychometric Theory. 3rd ed. New York: McGraw-Hill, 1994. 687
- 666 34. R Development Core Team (2003). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, URL <http://www.R-project.org>. 688
- 667 35. StataCorp. 2003. Stata Statistical Software: Release 8.0. College Station, TX: StataCorp LP. 689
- 668 36. Swinscow TDV. Statistics at Square One. London: British Medical Association, 1996. 690
- 669 37. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: A systematic review of 102 trials of interventions to improve professional practice. Can Med Assoc J 1995; 153: 1423-1431. 691
- 670 38. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: A randomized controlled trial. J Clin Oncol 2004 Feb 15; 22(4): 714-724. 692
- 671 39. Temps mitjà de resolució i llistes d'espera per Regions Sanitàries. Generalitat de Catalunya. Departament de Salut. Servei Català de la Salut. Available at http://www.10gencat.net/pls/catsalut/cawpk800.Recerca_Regions 693
- 672 40. Working Party of the Standing Committee of the Hospitals of the European Union (HOPE) on management of waiting lists. Measuring and comparing waiting lists: A study in four European countries. Third report. Brussels, April 18, 2004. Available at: http://www.hope.be/07publi/fr_tot.htm; accessed December 24th 2004. 694
- 673 695
- 674 696
- 675 *Address for correspondence:* J. Alonso, Head, Health Services Research Unit, Institut Municipal d' Investigació Mèdica. Dr 697
- 676 Aiguader, 80, 08003 Barcelona, Spain 698
- 677 Phone: +34-932-257-553; 699
- 678 Fax: +34-932-213-237 700
- 679 Email: jalonso@imim.es 701
- 680 702
- 681
- 682
- 683
- 684

Appendix Table

	VA > 0.7 AND VF-14 > 97.18	(VA > 0.3 AND VA ≤ 0.7) OR (VF-14 > 60.34 AND VF-14 ≤ 97.18)	VA ≤ 0.3 AND VF-14 ≤ 60.34
Surgery to be performed within 4 months	Not appropriate	Uncertain	Appropriate
Surgery NOT to be performed within 4 months	Appropriate		Not appropriate

	Journal : QURE	Dispatch : 22-2-2005	Pages : 11
	CMS No. : DO00011745	<input type="checkbox"/> LE	<input type="checkbox"/> TYPESET
	MS Code : QURE 312R1	<input checked="" type="checkbox"/> CP	<input checked="" type="checkbox"/> DISK



Dr. Jordi Alonso
Health Services Research Unit
Institut Municipal d'Investigacio Medica (IMIM)
c/ Dr. Aiguader, 80
E-08003 Barcelona
Spain

January 31, 2005

Subject: Your Submission QURE312R1

Dear Dr. Alonso,

We are pleased to inform you that your revised manuscript, "The Impact of the Vf-14 Index, A Perceived Visual Function Measure, In The Routine Management Of Cataract Patients", has been accepted for publication in *Quality of Life Research*. I appreciate your efforts to respond to both my comments and those of our reviewers, and I believe the paper is far clearer on both the benefits of using the Vf-14 but also the limitations so far of its application.

You will be contacted about proofs and offprints in due course by our Manufacturing Department. Any queries concerning your manuscript should now be addressed to the Editorial Department at: editdept@springer-sbm.com

Please remember to quote the manuscript number, QURE312R1, whenever inquiring about your manuscript.

With best regards,

Kathleen N Lohr, PhD,

Síntesis de resultados

En la revisión sistemática se ha observado una gran variabilidad en los métodos de los ensayos clínicos identificados (de manera especialmente destacable, por sus implicaciones, en cuanto a las intervenciones evaluadas, las variables de resultado evaluadas y al contexto sanitario en el que se han estudiado) (Tablas 3a y 3b). Los estudios que incorporaban información sobre salud mental demostraron mayor homogeneidad, lo que permitió la realización de un meta-análisis que demostró cambios en una variable de proceso (diagnóstico), pero no otras de proceso (tratamiento) ni de resultado.

Tabla 3a. Características de los estudios identificados en la revisión sistemática.

Característica	n (%)
País	
Estados Unidos	18 (86)
Gran Bretaña	3 (14)
Unidad de aleatorización	
Pacientes	10 (48)
Médicos	2 (9)
Centros	9 (43)
VARIABLES DE ESTUDIO*	
Proceso	20 (95)
Utilización de servicios sanitarios	13 (62)
Diagnóstico	13 (62)
Tratamiento	13 (62)
Resultado	11 (52)
Satisfacción	7 (33)
Utilidad percibida por profesionales clínicos	6 (29)
Año de publicación	
1966-1979	2 (9)
1980-1989	7 (33)
1990-1997	12 (57)
Puntuación de calidad, media y rango**	5.4 (2.5-7)

* Cada estudio puede haber evaluado el efecto en más de una variable, por lo que las categorías no son mutuamente excluyentes y la suma de los porcentajes correspondientes a cada categoría excede el 100%.

** Rango de valores posible de 0 (mínimo) a 8 (máximo)

Tabla 3b. Características de las intervenciones identificadas en la revisión sistemática.

Característica	n (%)
Ámbito	
Atención Primaria	15 (71)
Consultas Externas Hospitalarias	6 (29)
Pacientes*	
Sólo pacientes nuevos	5 (24)
Pacientes conocidos y nuevos	4 (19)
Sólo pacientes conocidos	2 (9)
No especificado	10 (48)
Tipo de intervención	
Formación + suministro de información	14 (67)
Suministro de información únicamente	7 (33)
Instrumentos CVRS**	
Genéricos	8 (38)
Específicos de Salud Mental	13 (62)
Otros	2 (9)

* Todos los estudios incluyeron exclusivamente pacientes adultos

** Cada estudio puede haber empleado más de un instrumento, por lo que las categorías no son mutuamente excluyentes y la suma de los porcentajes correspondientes a cada categoría excede el 100%.

En el segundo estudio, se ha utilizado técnicas psicométricas de Teoría de Respuesta al Ítem para desarrollar dos herramientas de ayuda a la interpretación de instrumentos de CVRS basadas en el contenido: el ‘termómetro’ y los ‘escenarios clínicos’. Para un individuo con una puntuación dada, la primera resume su capacidad funcional en términos de la actividad más difícil que el paciente es capaz de realizar sin dificultad. Mediante la herramienta de ‘escenarios clínicos’, es posible predecir la dificultad para realizar actividades seleccionadas incluidas en el propio cuestionario. Las figuras 2 y 3 reproducen las herramientas construidas para la versión española del Índice VF14.

Figura 2. Herramientas para la interpretación: Termómetro

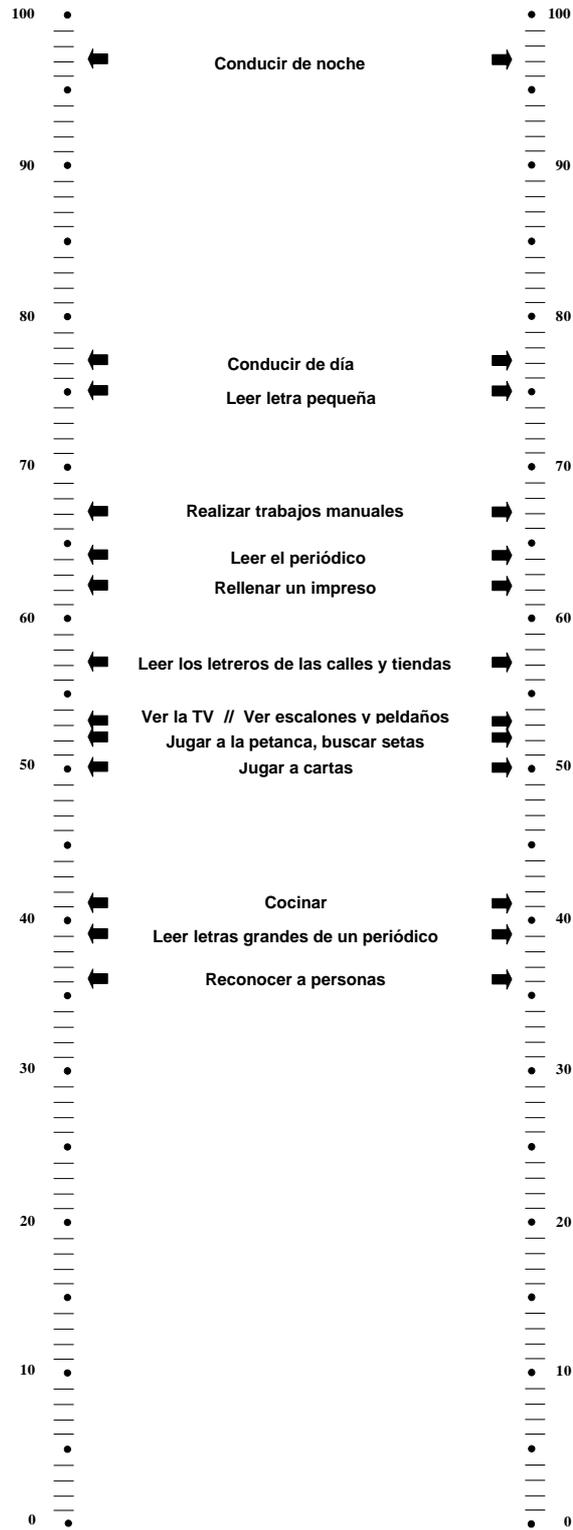


Figura 3. Herramientas para la interpretación: Escenarios Clínicos

Dificultad media para llevar a cabo cinco de las actividades más representativas del VF-14, según la puntuación obtenida en el instrumento.

Puntuación VF-14	Conducir de noche	Ver letra pequeña	Trabajos manuales	Ver TV	Reconocer gente
100	-	-	-	-	-
95	+	-	-	-	-
90	+	-	-	-	-
85	+	-	-	-	-
80	+	-	-	-	-
75	+	+	-	-	-
70	++	+	-	-	-
65	++	+	+	-	-
60	++	+	+	-	-
55	++	+	+	-	-
50	++	++	+	+	-
45	++	++	+	+	-
40	++	++	++	+	-
35	++	++	++	+	+
30	++	++	++	+	+
25	++	++	++	++	+
20	++	++	++	++	+
15	++	++	++	++	+
10	++	++	++	++	++
5	++	++	++	++	++
0	++	++	++	++	++

Grado de dificultad: '-' *Ninguna dificultad*; '+' *Alguna dificultad*; '++' *Incapaz de realizar la actividad*

Finalmente, en el tercer estudio, un ensayo antes-después con 19 médicos y 835 pacientes, se han utilizado las ayudas descritas como apoyo al suministro sistemático de información sobre función visual percibida (VF14) en la atención a pacientes con cataratas y se ha evaluado su impacto en la atención sanitaria. Se ha observado la modificación del proceso de atención sanitaria (mayor correlación médico-paciente en la valoración de la función visual), pero no se han detectado efectos sobre la adecuación de la indicación de la cirugía ni en la satisfacción del paciente con la atención sanitaria (Tabla 4).

Tabla 4. Modelos multinivel para el efecto del suministro del VF14 en la práctica clínica sobre variables de proceso y de resultado seleccionadas.

.Modelo	Acuerdo *		Adecuación ⁺		Satisfacción [‡]	
	β	95% CI	β	95% CI	β	95% CI
Efectos fijos						
<i>Nivel: paciente</i>						
Constante	40.08	[22.94; 57.22]	0.51	[-1.42; 4.77]	54.81	[40.48; 69.14]
Período (intervención=1)	-18.72	[-27.37; -10.07]	-0.10	[-0.84; 0.64]	-1.88	[-4.45; 0.69]
Puntuación VF-14 (0-100)	0.10	[-0.12; 0.32]	-	-	0.07	[-0.03; 0.17]
AV peor ojo (0-1)	9.18	[2.06; 16.30]	-	-	6.71	[-1.05; 14.47]
Interacción VF-14*período	0.25	[0.12; 0.38]	-	-	-	-
Efectos aleatorios						
<i>Nivel: centro</i>						
Constante	11.30	[4.15; 30.78]	1.16	[0.37; 3.67]	4.13	[1.85; 9.25]
Puntuación VF-14	0.22	[0.09; 0.53]	-	-	-	-
<i>Nivel: oftalmólogo</i>						
Constante	9.47	[1.96; 45.79]	1.20	[0.65; 2.21]	2.39	[0.92; 6.15]
Puntuación VF-14	0.12	[0.02; 0.84]	-	-	-	-
<i>Residual</i>	15.68	[14.89; 16.51]	0.89	[0.78; 1.01]	14.99	[14.11; 15.91]

DISCUSIÓN

Sobre la revisión sistemática

Aunque se ha propuesto que la evaluación de la calidad de vida debería ser un elemento fundamental de la atención clínica (Kazis LE et al, 1990), la efectividad de la inclusión formal de este tipo de información (medición y suministro de puntuaciones al personal sanitario) sigue siendo incierta. En la presente revisión sistemática de ensayos clínicos aleatorizados, la administración de información sobre calidad de vida a los profesionales sanitarios, recogida mediante medidas estandarizadas, modifica sólo algunos de los aspectos del proceso de atención: se ha observado un incremento en la frecuencia de diagnósticos y en el uso de servicios sanitarios y de derivaciones en aproximadamente la mitad de los estudios. Pero no se ha demostrado en otros aspectos del proceso ni tampoco en el resultado de la atención, como el estado de salud de los pacientes. Esto último parece estar en contradicción con las opiniones de los clínicos sobre la utilidad potencial de la información sobre calidad de vida recogidas en los estudios evaluados (Rubenstein LV et al, 1989; Kazis LE et al, 1990; White P et al, 1990; Wasson J, Hays R et al, 1992), así como en otros estudios observacionales (Nelson E et al 1987; Schor EL et al, 1995; Meyer KB, 1994).

De acuerdo con la revisión, el efecto de las intervenciones fue mayor en los estudios en los que se había suministrado información relativa a la salud mental (Callahan CM et al, 1996; Johnstone A et al, 1996; German PS et al, 1987; Linn L et al 1980; Yager J et al, 1981; Rand EH et al, 1988; reifler DR et al, 1996; Dowrick C et al, 1995; Hooper EW et al 1984; Zung WWK et al, 1983; Shapiro S et al, 1987; Mazonson PD et al, 1996; Mathias SD et al, 1994; Brody DS et al, 1990; Magruder-Habib K et al, 1990; Moore JT et al, 1987; Gold I et al, 1989). Se sabe que un gran número de pacientes con problemas mentales no están identificados como tales (esto es, no tienen ningún diagnóstico ni se les prescribe medicación alguna) y que, sin embargo, serían susceptibles de recibir intervenciones médicas efectivas (Callahan CM et al, 1996; Johnstone A et al, 1975). Este problema de clasificación podría

deberse en parte a que los pacientes expresan quejas que el clínico no es capaz de interpretar (Feightner JM et al, 1990). Por ello, un abordaje complementario mediante cuestionarios estandarizados de CVRS, al proporcionar información más interpretable para los profesionales de la salud (por ejemplo, una probabilidad dada de la presencia de un trastorno mental) incrementaría la detección de estos pacientes. De hecho, todos los estudios evaluados relacionados con la evaluación de la salud mental usaron instrumentos específicos para el cribado de trastornos mentales. Cuando se usa una prueba diagnóstica con el propósito de realizar un cribado la interpretación de sus resultados suele ser mucho más clara de antemano, por cuanto, por un lado, se inscribe en un modelo de decisión clínica secuencial y estructurado: se aplica la prueba de cribado y en función del resultado (positivo o negativo) se tomarán una serie de decisiones clínicas (Pettitti DB, 2000). Por el otro lado, tal vez más relevante aún, al transformarse la variable de resultado en una variable dicotómica, la información se resume al máximo y maximiza sus posibilidades de interpretación. En definitiva, la interpretación se centra en la alta probabilidad de ser caso, lo que tiene implicaciones clínicas muy claras, mientras que la baja calidad de vida o la discapacidad no la tienen.

Los resultados indican que el uso con el propósito de cribado de medidas de calidad de vida, especialmente las que son específicas para salud mental, pueden tener un efecto beneficioso en la práctica clínica, pero es probable que lo anterior también sea aplicable a pacientes con procesos agudos o reagudizados (por ejemplo, pacientes no conocidos previamente o que acuden a una primera visita, pacientes ancianos frágiles, o aquellas personas que padecen enfermedades que afectan especialmente el estado funcional, como la artritis, las cataratas o la insuficiencia cardíaca congestiva (Rubenstein LV et al, 1989) así como pacientes vulnerables que sigan un tratamiento rehabilitador (como los enfermos con afecciones cardíacas o respiratorias o las derivadas de accidentes vasculares cerebrales).

En contra de lo esperado, en la mayoría de las intervenciones evaluadas no se observó un efecto en el tratamiento, valorado como número de prescripciones o de cambios introducidos el mismo. La mayoría de los médicos a quienes se les preguntó de forma explícita, indicó que no había tomado ninguna decisión ligada a la gestión clínica de sus pacientes atendiendo a la información proporcionada por los

cuestionarios (Rubenstein LV et al, 1989; White P et al, 1995; Wasson J et al, 1995; Goldsmith G et al, 1989). Aunque la mayoría de las intervenciones incluyeron refuerzos educacionales, es posible que las intervenciones realizadas en los estudios evaluados fueron insuficientes para superar la falta de experiencia previa con estas medidas y promover la necesaria familiaridad de los médicos con el tipo de información suministrada. En última instancia, estaríamos de nuevo ante un problema de interpretación de los resultados de las medidas de CVRS, un paso clave en la modificación del comportamiento de los médicos. Adicionalmente, en más del 90% de los casos las intervenciones se realizaron en pacientes conocidos (que habían sido previamente valorados por el médico) y ello puede haber dejado un margen aún más estrecho a la información adicional suministrada, atenuando en consecuencia su posible efecto sobre la salud percibida de los pacientes. Siguiendo el modelo conceptual de Callahan y cols., los profesionales sanitarios no han podido traducir las medidas de función o de estado de salud en intervenciones específicas, lo que interfiere con la vinculación de la medida a los recursos (y por ende a la acción) (Callahan CM et al, 1996).

Los mismos autores de los estudios evaluados han propuesto algunas aproximaciones generales para mejorar la integración de las medidas de estado de salud en la práctica clínica (Rubenstein LV et al, 1989; Kazis LE et al, 1990; Street LR et al, 1994). Así se ha recomendado que se la información se debería proporcionar al clínico mediante informes muy simples y sencillos de interpretar y que se debería disponer de esa información en tiempo real durante la visita. También se ha propuesto que la comunicación medico-paciente se debería estimular activamente proporcionando los resultados también a los pacientes. La mayoría también propone la realización de intervenciones educacionales antes de la implementación de la nueva medida. No existen datos empíricos que apoyen estas recomendaciones y serán necesarios estudios rigurosos y bien diseñados que evalúen estrategias innovadoras.

Es importante destacar la heterogeneidad de determinadas características de los estudios revisados. Si bien la mayoría de los estudios se realizaron en centros de atención primaria, hubo una variabilidad substancial en el tipo de médicos participantes (especialidad clínica, experiencia profesional) y en los pacientes

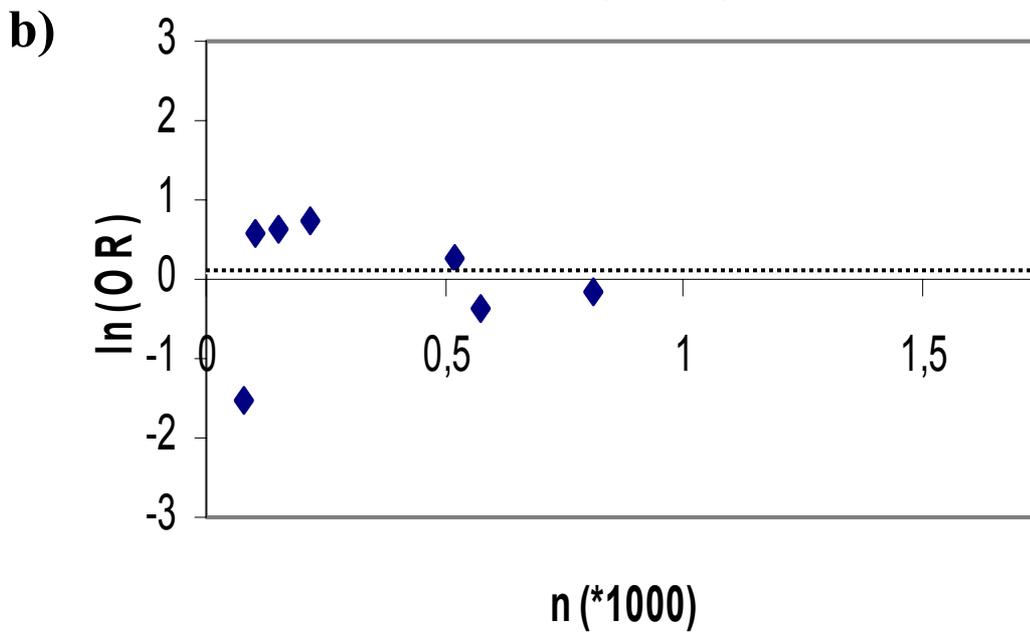
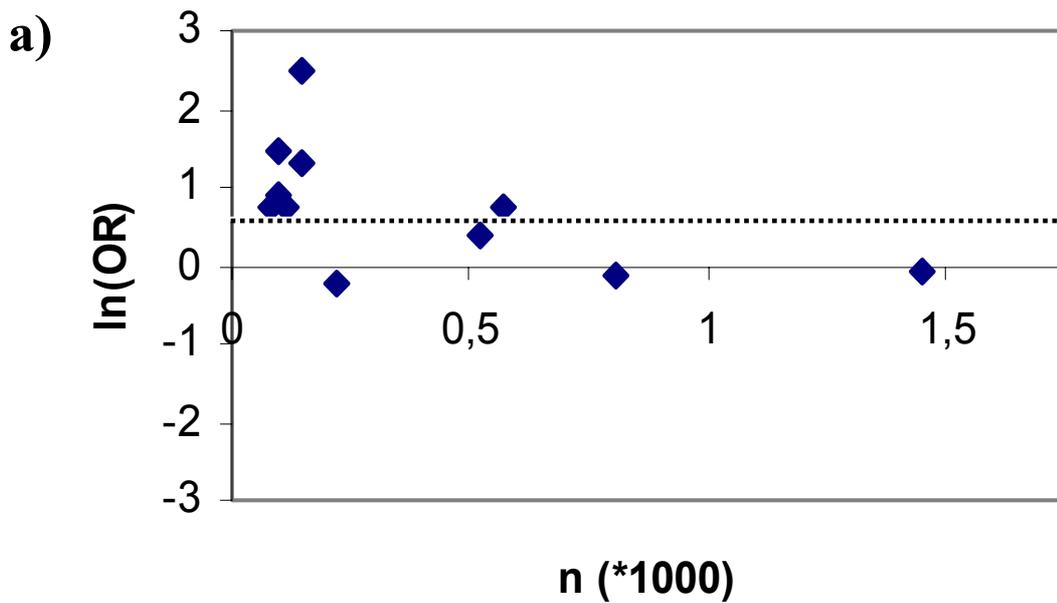
estudiados (patología, grado de conocimiento por parte del médico). Pero la mayor variabilidad se dio en las intervenciones evaluadas, en cuanto a la forma y el momento en que se suministró la información sobre salud percibida. En el 75% de los casos se realizó una intervención educativa, pero el momento de su realización, su periodicidad y su intensidad variaron, así como los instrumentos de calidad de vida utilizados y el momento en que se les suministró la información a los médicos. Estudios diferentes seleccionaron grupos control e intervención distintos, así como diferentes unidades de randomización y de análisis (pacientes, médicos, consultas). Todas estas diferencias limitaron la comparabilidad de los estudios y forzaron a seleccionar un subgrupo de estudios de mayor homogeneidad en el que fuera posible resumir cuantitativamente los efectos observados (meta-análisis).

El análisis que se presenta incluyó sólo los estudios relativos al suministro de información sobre salud mental, porque resultaron más homogéneos. Los resultados son consistentes con el análisis cualitativo general y demuestran un efecto sólo en algunos componentes del proceso asistencial. Aun considerando las limitaciones inherentes a este tipo de técnica, no se puede descartar la existencia de un sesgo de publicación, tal como podría reflejar el gráfico de embudo (*funnel plot*) de los resultados de cada estudio individual que indica tendencia a la no publicación de estudios con resultados negativos en cuanto al diagnóstico (Figura 4.a), pero no al tratamiento (Figura 4.b) (Pettitti DB, 2000; Egger M et al, 2001; Sterne JAC et al, 2001). El análisis cualitativo empleado es muy sensible a la detección de resultados positivos, lo que hace muy probable que los resultados observados sobreestimen el efecto real de las intervenciones evaluadas.

La generabilidad de esta revisión sistemática también es limitada en lo que se refiere a su aplicación a nuestro medio. El perfil medio de las evaluaciones incluidas en la revisión es el de un estudio realizado en la última década (57%) en EEUU (86%), en un servicio de atención primaria (71%) con pacientes adultos y con instrumentos de medida de salud mental (62%). Las significativas diferencias entre el sistema sanitario de EEUU y el español, justifica la precaución con estos resultados deben interpretarse en nuestro entorno (Navarro V, 1989).

Todo lo anterior apunta la necesidad estudios que evalúen el impacto de la medida de la CVRS en nuestro medio y que superen las limitaciones indicadas.

Figura 4. Revisión sistemática: diagramas de embudo de a) diagnóstico y b) tratamiento. Las líneas de puntos representan los valores calculados para el estimador combinado.



Sobre el método general para la obtención de herramientas para la interpretación y su aplicación al Índice de Función Visual VF-14.

Se ha presentado un método para el desarrollo de herramientas de ayuda a la interpretación que proporcionen un resumen inteligible de las medidas de estado de salud. El método se basa en el contenido del instrumento con que se mide y se fundamenta en dos nociones simples e intuitivas de los modelos de Teoría de Respuesta al Ítem: 1) cada ítem es una definición operativa del concepto que se pretende medir, y 2) los ítems describen actividades que se pueden ordenar de forma lógica según la habilidad (estado funcional) creciente necesaria para realizar esas tareas (Wright BD, 1979).

Aplicando los resultados que genera el análisis de Rasch, uno de los modelos de la TRI, se puede construir un *continuum* del concepto medido, en el que los ítems representan hitos de dificultad funcional creciente, como las marcas de un termómetro. En esta aproximación, la función visual percibida del paciente se resume en primer lugar en términos de la actividad más difícil del VF14 que es capaz de realizar sin problemas. En esta herramienta que se ha denominado el ‘termómetro’, los ítems se presentan ordenados según su dificultad. Posteriormente, se han identificado los conglomerados de ítems y se ha seleccionado un ítem de referencia por cada conglomerado. Esta herramienta permite utilizar un subgrupo de ítems relevantes para caracterizar la función visual percibida por el paciente. Las respuestas esperadas de los pacientes (su habilidad para realizar las actividades descritas en los ítems de referencia) proporcionan una descripción clínica más rica de las características funcionales asociadas a una puntuación (su ‘escenario clínico’). La ordenación de los ítems según su dificultad y los escenarios clínicos son consistentes con la experiencia clínica y han proporcionado una bondad de ajuste aceptable, lo que apoya la validez aparente de nuestra propuesta. Por su construcción, este procedimiento debería ser aplicable a cualquier otras medida de CVRS unidimensional. Para aplicarlas a medidas multidimensionales, sería necesario desarrollar tantas herramientas como dimensiones tenga el instrumento.

El modelo analítico que subyace a la construcción de estas herramientas (TRI), se basa en lo que las personas a quienes se administra el instrumento realmente pueden hacer. Se justifica así su caracterización como interpretación basada en el contenido, definido

precisamente por las actividades descritas en los ítems. La conveniencia de una calibración previa podría utilizarse para argumentar un marco normativo implícito similar al de la interpretación basada en normas. La medición Rasch se ha descrito con frecuencia como ‘objetiva’ porque garantiza la invariancia de los parámetros de los ítems y de las personas. Las estimaciones de la habilidad están libres de los efectos de la dificultad y las estimaciones de la dificultad están libres de los efectos de la habilidad (Wright BD, 1979). Y esto es así, porque cada faceta del modelo se calcula de forma independiente. Esta propiedad implica que el parámetro que caracteriza a un ítem no depende de la distribución de la habilidad de los sujetos. De la misma forma, el parámetro de una persona no depende de la selección de ítems que se le presenta. Empíricamente se ha podido verificar la estabilidad de las calibraciones de los ítems del VF14 obtenidas en el segundo estudio (Valderas et al, 2004) al evaluar su concordancia con las obtenidas con las respuestas de los pacientes del tercer estudio incluido en este trabajo (Valderas et al, 2005). Así, se ha calculado un valor de fiabilidad (concordancia) medido mediante el Coeficiente de Correlación Intraclase de 92.9 IC95% [0.79; 0.98]. Esta evidencia empírica de la independencia de las mediciones de la distribución de la habilidad de los individuos, debería completarse con investigaciones en otras poblaciones de pacientes con cataratas, tal como ya se realizó con otros instrumentos (Fisher WP Jr et al, 1999; Prieto L et al, 1998).

Aunque los resultados del análisis de Rasch sugieren que el VF14 define una variable unidimensional con sentido lógico, tres ítems presentaron desajuste parcial. Se optó por mantener la integridad del cuestionario atendiendo a razones prácticas: dos de las actividades eran aplicables a todos los pacientes menos uno (99.5%) y la tercera (*‘Conducir de noche’*) aunque de menor aplicabilidad, ofrecía mayor capacidad de discriminación al instrumento. Otros estudios similares han encontrado resultados contradictorios y así algunos autores han confirmado la necesaria integridad del cuestionario (Friedman DS et al, 2002), mientras que otros han eliminado algunos ítems e incluso han incorporado otros (Veloza et al, 2000). Es preciso profundizar en esta línea de investigación y establecer si también en otras versiones del VF14 se observa un funcionamiento imperfecto (Alonso J et al, 1997; Fisher WP Jr et al, 1999). Para conseguir un instrumento fiable, cualquier ítem cuyo ajuste se confirme inapropiado debería eliminarse o ser sustituido por otro con mejor comportamiento.

También se ha observado un funcionamiento imperfecto de la escala de respuesta original. Por ello algunas categorías se colapsaron en tres categorías con sentido clínico. Estamos convencidos de que de esta manera se preserva el sentido de las respuestas de los pacientes, pero carecemos de evidencia experimental de la equivalencia de estos dos formatos de respuesta. Otros investigadores también han planteado combinar determinadas categorías, atendiendo a la frecuencia de respuesta de cada opción y con proponiendo combinaciones distintas (p. ej., manteniendo las opciones de respuesta ‘Ninguna dificultad’ y ‘Poca dificultad’ y colapsando las demás) (Veloza et al, 2000). Además, hay que considerar que se ha seleccionado una categoría de respuesta (‘Ninguna dificultad’) como referencia para la construcción de la herramienta ‘termómetro’. Se hizo así porque esta categoría es la que requiere mayor capacidad y debe ser la más sensible para la detección de la discapacidad debida a las cataratas. Pero otras opiniones también pueden ser válidas. La medida en que esta decisión puede haber influido en la interpretación clínica merece más análisis y otro tanto es aplicable a la elección de las actividades de referencia. Los autores de un estudio reciente en el que debían seleccionar 3 ítems del VF14 para su administración a pacientes han coincidido en la selección de dos de ellos (leer letra pequeña y ver la TV) de forma independiente (Mönestam et al, 2004).

Finalmente, la naturaleza probabilística del método empleado en las construcción de las herramientas permitía anticipar un cierto grado de disparidad en los datos predichos y observados tanto en las relaciones entre ítems (‘termómetro’) como en las respuestas a los ítems representativos (‘escenarios clínicos’). Esto es especialmente cierto para las herramienta ‘escenarios clínicos’ por dos motivos. Por un lado, el cálculo de las respuestas predichas es una estimación no sesgada basada en el promedio de las respuestas de los pacientes con una misma puntuación, pero no incorpora una medida de la dispersión de esas respuestas. La tabla incorpora un grado de error adicional al reducir a 21 categorías (0, 5, ..., 100) la información de una variable continua (puntuación VF14). La concordancia observada puede considerarse por tanto satisfactoria y apoya la validez del método.

Aunque las ayudas a la interpretación que se presentan aquí se han desarrollado específicamente para el VF14, no hay ninguna razón para pensar que el método general no se pueda aplicar a muchos otros instrumentos de medida de CVRS. Los principios

básicos en los que se apoya (operabilidad y orden) son aplicables a la mayoría de los cuestionarios de CVRS. La mayor limitación radica en que, para cada instrumento, será necesaria una herramienta distinta para cada dimensión. El análisis de Rasch ya se ha aplicado con éxito a instrumentos multidimensionales bien establecidos (Prieto L et al, 1998), pero será necesaria una expansión de estas aplicaciones. Además, por su construcción, el método también se puede aplicar a la interpretación de los resultados de cuestionarios adaptativos computerizados basados en bancos de ítems. Esto es una línea de desarrollo aún incipiente en la medición del estado funcional y de la calidad de vida relacionada con la salud y en la que se esperan notables desarrollos en breve (Holman R et al, 2003). En cualquiera de estas aplicaciones, los métodos que aquí se presentan deberían facilitar el uso de las medidas de salud percibida en distintos campos y especialmente en la práctica clínica.

Sobre el uso sistemático del VF-14 en la práctica clínica

Se ha observado un incremento sustancial (Skinsow TDV, 1996) en la correlación entre la valoración realizada por el oftalmólogo y la percepción del paciente en el análisis divariado (débil en el período de control (antes) y moderada en el de intervención (después)). Este efecto ha sido aún más notable en el modelo de regresión, en el que la contribución ajustada de la puntuación VF14 a la valoración del oftalmólogo de la función visual del paciente se incrementó en un factor 3.5. La mejora en esta variable de proceso de la atención (correlación médico-paciente) es consistente con los hallazgos de la revisión sistemática, así como con otros estudios posteriores (Greenhalgh J et al, 1999; Gilbody SM et al, 2002), pero es la primera vez que se observa en pacientes con cataratas. Todas las guías de práctica clínica coinciden en el papel clave de la función visual percibida en la gestión clínica de las cataratas y su incorporación en la evaluación general de los pacientes proporcionaría a los oftalmólogos una medida válida y fiable que mejoraría la calidad del proceso de la atención y permitiría la concreción de un modelo de gestión efectivo y centrado en los pacientes.

La intervención estudiada ha sido simple con el propósito deliberado de aislar los efectos de la administración sistemática de la información de los eventuales efectos de

otros componentes de una intervención más compleja (que incluyera, por ejemplo, recomendaciones expresas de gestión clínica). La elección de este tipo de intervención podría parecer en contradicción con las conclusiones de la revisión sistemática, que apuntan la conveniencia de intervenciones intensivas y complejas (con diversos componentes de refuerzo). Revisiones sistemáticas posteriores han observado también un mayor impacto en la práctica clínica de intervenciones complejas, pero el problema radica en la imposibilidad de estudiar los efectos de los distintos componentes por separado, excepto evidentemente con diseños más complejos, de tipo factorial (McAlister et al, 2003). El hecho de que la intervención sea poco intensa, y aunque la falta de efecto sobre la satisfacción de los pacientes es consistente con los resultados de la revisión sistemática, puede haber influido en que no se hayan observado cambios en otras variables de resultado. Factores como la experiencia profesional, las características individuales de los pacientes o la disponibilidad de recursos tienen un gran impacto sobre la práctica clínica (Oxman AD et al, 1995) y por ende sobre la adecuación, efectos posiblemente mucho más importantes que la intervención diseñada.

En relación al modelo conceptual presentado, se observó la relación esperada entre la función visual valorada por el oftalmólogo y el momento adecuado para la cirugía, pero no se pudo establecer una asociación entre el acuerdo médico-paciente y la adecuación de la decisión. Esta discontinuidad, más que una refutación empírica del modelo, probablemente señala la subordinación entre ambas. Las variables ligadas al estilo de práctica médica han demostrado ser los determinantes más importantes de las decisiones médicas y su impacto no debería ser menor en la adecuación (Grytten J et al, 2003; Casparie AF, 1996). En este contexto, el efecto del acuerdo sobre la decisión es probablemente de mucha menor magnitud y, portanto, posiblemente despreciable.

La elección de la adecuación de la indicación de cirugía como variable de resultado deriva de la hipótesis de que una mayor cantidad de información sobre el paciente mejorará la calidad de las decisiones médicas y los resultados asociados. Un estudio reciente ha confirmado peores resultados de la atención en los pacientes con indicación inadecuada precisamente en pacientes con cataratas (Tobacman et al, 2003). En este estudio se utilizaron los criterios propuestos siguiendo el método RAND para la clasificación de la adecuación de la cirugía (Lee PP et al, 1993), que no se han considerado aplicables en nuestro estudio, después de considerar el distinto contexto,

los cambios sustanciales en las técnicas quirúrgicas y el tiempo transcurrido desde la realización del estudio en que se basan (Fitch K et al, 2001).

Este estudio presenta algunas ventajas respecto a estudios previos: se incluyó un número elevado de pacientes y la muestra de centros ofreció una notable variabilidad (tamaño, financiación y nivel de atención). Se siguieron las recomendaciones de los estudios previos y de la revisión sistemática. El análisis estadístico se realizó atendiendo a la naturaleza jerárquica de los datos, un aspecto al que no se había prestado la debida atención en estas evaluaciones, con la excepción de un estudio aparecido este mismo año (Velikova G et al, 2004). El estudio se diseñó de tal forma que maximizara la posibilidad de atribuir al suministro de información sobre CVRS los posibles efectos y que se excluyera la posibilidad de contaminación, esto es la replicación (parcial o total) de la intervención en el grupo control. Este efecto ha sido mensurable en otros estudios (Detmar SB et al, 2002).

En cuanto a las limitaciones de este estudio, la inclusión de un grupo control de médicos y pacientes distintos a los evaluados y que no hubieran recibido la intervención habría incrementado la posibilidad de invocar una relación causal entre la intervención y los efectos observados. Se descartó la aleatorización por conglomerados de los oftalmólogos o de las consultas por cuanto se contaba con un número pequeño de ambos (si bien el número de oftalmólogos fue mayor que en muchos de los estudios incluidos en la revisión sistemática, algunos de los cuales sólo incluyeron uno o dos médicos) y debido a la gran variabilidad en las características de los centros. La aleatorización de los oftalmólogos estratificada por centro o de los pacientes estratificada por médico se descartó porque suponía un riesgo inaceptable de contaminación.

Dos son las limitaciones inherentes al diseño antes-después y que podrían haber actuado en este estudio: la regresión a la media y la posibilidad de que otras intervenciones o acontecimientos hayan operado en el mismo contexto de estudio y sean responsables del efecto observado. No se ha podido identificar ninguna intervención de estas características, como podría ser, por ejemplo la publicación o difusión de una guía de práctica clínica nacional o internacional o una iniciativa regulatoria por parte de la administración sanitaria que fuera aprobada poco antes o durante el período de estudio.

Un análisis de medidas repetidas de las variables de resultado tampoco fue capaz de detectar ninguna tendencia significativa en el período control. No obstante, no es posible descartar la presencia de factores exógenos que hayan afectado al efecto observado.

El ámbito geográfico del estudio fue restringido. Si bien se incluyeron cinco ciudades de tres regiones sanitarias distintas, todas ellas pertenecen a una misma comunidad autónoma (Cataluña), están muy próximas entre sí y comparten ciertas características en cuanto a su población (urbana, envejecida), la organización de los servicios sanitarios (todos los centros pertenecían a la misma entidad proveedora (Institut Català de la Salut), con la excepción de una clínica privada y de un centro de atención primaria vinculado a otro proveedor) y posiblemente los estilos de práctica médica. Todo ello puede limitar la generabilidad de las observaciones.

Los resultados presentados apuntan un beneficio limitado del suministro de rutina de información sobre función visual mediante el índice VF14 en la atención a los pacientes con cataratas. Un ensayo aleatorizado por conglomerados que incluya una intervención más intensa, en el que, por ejemplo se proporcionen recomendaciones específicas para la gestión clínica de los pacientes con cataratas permitiría inferencias más robustas. Si los resultados se confirman, la evaluación sistemática de la disfunción visual mediante el índice VF14 supondría una mejora cualitativa de la evaluación general de estos pacientes.

A la luz de los resultados, es necesaria más investigación y evaluación antes de que se pueda recomendar de forma generalizada el uso de las medidas de salud percibida, especialmente para otras medidas, tanto genéricas como específicas. Se necesita también más investigación para ayudar a los profesionales sanitarios a identificar o elegir aquellos pacientes que se podrían beneficiar más de estas intervenciones. Es probable que estudios más grandes que se concentren en pacientes nuevos o en grupos específicos de pacientes vulnerables con enfermedad activa que incluyan una educación intensiva de los médicos y que eviten la contaminación proporcionen una idea más definida de este asunto tan importante.

CONCLUSIONES

1. La revisión sistemática de la evidencia y su meta-análisis indican que la administración de información sobre la salud percibida a los profesionales de la salud podría modificar algunos componentes del proceso de la atención sanitaria, pero no su resultado.
2. Determinadas características de los estudios evaluados, especialmente la gran variabilidad observada en las intervenciones estudiadas y en otros aspectos clave (medio, instrumentos), aconseja realizar nuevos estudios que permitan realizar inferencias más robustas.
3. Mediante la aplicación de técnicas psicométricas avanzadas (análisis de Rasch), se ha desarrollado un método para la construcción de herramientas de ayuda a la interpretación de instrumentos de calidad de vida relacionada con la salud basadas en el contenido y que ha permitido el desarrollo de herramientas específicas para el Índice de Función Visual VF-14, el 'termómetro' y los 'escenarios clínicos'.
4. El método propuesto se ha demostrado robusto desde un punto de vista teórico y empírico y debería ser fácilmente generalizable, pero este extremo no ha sido aún evaluado.
5. En consonancia con los resultados de la revisión sistemática y el meta-análisis, en el estudio de intervención se han observado modificaciones en algunos componentes del proceso de la atención sanitaria (incremento del acuerdo entre médico y paciente), pero no en otros (adecuación de la indicación de cirugía). Tampoco se han observado cambios en la satisfacción de los pacientes.
6. Algunas limitaciones del estudio de intervención (diseño, intensidad de la intervención, validez externa) hacen recomendable la realización de estudios complementarios que reproduzcan las observaciones realizadas.

BIBLIOGRAFÍA

AHCPR Cataract Management Guideline Panel. Cataract in Adults. Management of functional impairment. Clinical Practice Guideline, Number 4. Rockville MD: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. AHCPR pub no 93-0542. 1993

American Academy of Ophthalmology, Anterior Segment Panel. Cataract in the adult eye. San Francisco (CA): American Academy of Ophthalmology (AAO); 2001. 62 p.

Alonso J, Espallargues M, Folmer-Andersen T, Cassard SD, Dunn E, Bernth-Petersen P et al. International applicability of the VF-14. An Index of Visual Function in Patients with Cataracts. *Ophthalmology* 1997; 104(5):799-807.

Apostol TM. *Calculus*. 2 ed. Barcelona: Ed.Reverté, 1973.

Badia X, Salamero M, Alonso J. *La Medida de la salud : guía de escalas de medición en español* . (3a ed). Barcelona : Edimac, 2002

de Boer MR., Moll AC, de Vet HCW, Terwee CB, VölkerDieben HJM, van Rens GHMB. Psychometric properties of vision-related quality of life questionnaires: a systematic review *Ophthal. Physiol. Opt.* 2004 24: 257-273

Boisjoly H, Gresset J, Fontaine N, Charest M, Brunette I, LeFrancois M, Deschenes J, Bazin R, Laughrea PA, Dube I. The VF-14 index of functional visual impairment in candidates for a corneal graft. *Am J Ophthalmol.* 1999 Jul;128(1):38-44.

Brody DS, Lerman CE, Wolfson HG, et al. Improvement in physicians' counseling of patients with mental health problems. *Arch Intern Med* 1990;150:993-998.

Brylawsky B (Project Leader), et al. The NLM PubMed Project. Available at: <http://www.ncbi.nlm.nih.gov/PubMed/overview.html>.

Calkins DR, Rubenstein LV, Cleary PD, et al. Failure of physicians to recognize functional disability in ambulatory patients. *Ann Intern Med* 1991;114:451-454.

Calkins DR, Rubenstein LV, Cleary PD, et al. Functional disability screening of ambulatory patients: a randomized controlled trial in a hospitalbased group practice. *J Gen Intern Med* 1994;9:590-592.

- Callahan CM, Dittus RS, Tierney WM. Primary care physicians' medical decision making for late-life depression. *J Gen Intern Med* 1996;11:218-219.
- Callahan CM, Hendrie HC, Tierney WM. The recognition and treatment of late-life depression: a view from primary care. *Int J Psychiatry Med* 1996;26:155-171.
- Casparie AF. The ambiguous relationship between practice variation and appropriateness of care: an agenda for further research. *Health Policy* 1996; 35(3) 247-265
- Cassard SD, Patrick DL, Damiano AM, Legro MW, Tielsch JM, Diener-West M et al. Reproducibility and responsiveness of the VF-14. An index of functional impairment in patients with cataracts. *Arch Ophthalmol* 1995; 113:1508-1513.
- Castells X, Alonso J, Castilla M, Comas M. Efficacy and cost of ambulatory cataract surgery: a systematic review. *Med-Clin-(Barc)* 2000;114 Suppl 2:40-7.
- Castells X, Alonso J, Castilla M, Ribo C, Cots F, Anto J. Outcomes and costs of outpatient and inpatient cataract surgery: a randomised clinical trial. *Journal of Clinical Epidemiology* 2001;54(1):23-9. 21104082.
- Clancy CM, Eisenberg JM. Outcomes research: measuring the end results of health care. *Science* 1998; 282(5387):245-246.
- Congdon N, Vingerling JR, Klein BE, West S, Friedman DS, Kempen J, O'Colmain B, Wu SY, Taylor HR. Prevalence of cataract and pseudophakia/aphakia among adults in the United States. *Arch Ophthalmol*. 2004 Apr;122(4):487-94.
- Detmar SB, Muller MJ, Schornagel JH, Wever LDV, Aaronson NK. Health-related quality of life assessments and patient physician communication. *JAMA* 2002; 288(23): 3027-3034.
- Deyo RA, Patrick DL. Barriers to the use of health status measures in clinical investigation, patient care, and policy research. *Med Care* 1989;27:S254-S267.
- Deyo RA, Carter WB. Strategies for improving and expanding the application of health status measures in clinical settings: A researcher-developer viewpoint. *Med Care* 1992;30(suppl):MS176-MS186.
- Deyo RA, Patrick DL. The significance of treatment effects: the clinical perspective *Med Care* 1995; 33(4 Suppl):AS286-AS291.
- Dowrick C. Does testing for depression influence diagnosis or management by general practitioners? *Fam Pract* 1995;12:461-465.

Dowrick C, Buchan I. Twelve month outcome of depression in general practice: Does detection or disclosure make a difference? *BMJ* 1995;311:1274-1276.

Egger M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care: meta-analysis in context*, 2nd ed. London: BMJ Books, 2001.

Espallargues M. *La atención sanitaria de las cataratas seniles: recomendaciones para una atención apropiada*. Barcelona: Agència d'Avaluació de Tecnologia Mèdica. Servicio Catalán de Salud. Departamento de Sanidad y Seguridad Social (DSSS) de Cataluña. Agosto 1998.

Espallargues M, Alonso J. Effectiveness of cataract surgery in Barcelona, Spain site results of an international study. Barcelona I-PORIT investigators. International Patient Outcomes Research Team. *J Clin Epidemiol* 1998; 51(10):843-852.

Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals. A systematic review of its impact. *Med Care* 2000; 38(2): 175-186.

Espallargues M, Sampietro-Colom L, Castells X, Pinto JL., Rodríguez E, Comas M. *Elaboración de un sistema de priorización de pacientes en lista de espera para cirugía de catarata artroplastia de cadera y rodilla*. Barcelona: Agència d'Avaluació de Tecnologia i Recerca Mèdiques. CatSalut. Departament de Sanitat i Seguretat Social. Generalitat de Catalunya. Enero de 2003.

Feightner JW, Worrall G. Early detection of depression by primary care physicians. *Can Med Assoc J* 1990;142:1215-1220.

Fitch K, Bernstein SJ, Aguilar MD, Burnand B, La Calle JR, Lazaro P, et al. *The RAND/UCLA Appropriateness Method User's Manual*. Santa Monica. Rand, 2001.

Fitzpatrick R, Fletcher A, Gore S, et al. Quality of life measures in health care, I: Applications and issues in assessment. *BMJ* 1992;305:1074-1077.

Fletcher A, Gore S, Jones D, et al. Quality of life measures in health care, II: Design, analysis, and interpretation. *BMJ* 1992;305:1145-1148.

Friedman DS, Tielsch JM, Vitale S, Bass EB, Schein OD, Steinberg EP. VF-14 item specific responses in patients undergoing first eye cataract surgery: can the length of the VF14 be reduced? *Br J Ophthalmol*. 2002 Aug;86(8):885-91

German PS, Shapiro S, Skinner EA, et al. Detection and management of mental health problems of older patients by primary care providers. *JAMA* 1987;257:489-493.

- Gilbody SM, House AO, Sheldon T. Routine administration of Health Related Quality of Life (HRQoL) and needs assessment instruments to improve psychological outcome-a systematic review. *Psychol Med*. 2002 Nov;32(8):1345-56.
- Gold I, Baraff LJ. Psychiatric screening in the emergency department: its effect on physician behavior. *Ann Emerg Med* 1989;18:875-880.
- Goldsmith G, Brodwick M. Assessing the functional status of older patients with chronic illness. *Fam Med* 1989;21:38-41.
- Greenhalgh J, Meadows K. The effectiveness of the use of patientbased measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *J Eval Clin Pract*. 1999 Nov;5(4):401-16.
- Grytten J, Sørensen R. Practice variation and physician-specific effects. *J Health Econ* 22 (2003) 403–418
- Guyatt GH, Sackett DL, Cook DJ. User's guides to the medical literature, II: How to use an article about therapy or prevention. A: Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993;270:2598-2601.
- Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. *BMJ* 1998; 316(7132):690-693.
- Hamed WW, Fedorowicz Z. Day care versus in-patient surgery for age-related cataract. *Cochrane Database Syst Rev*. 2004;(1):CD004242.
- Hart PM, Chakravarthy U, Stevenson MR, Jamison JQ. A vision specific functional index for use in patients with age related macular degeneration. *Br J Ophthalmol*. 1999 Oct;83(10):1115-20.
- Hase HD, Luger JA. Screening for psychosocial problems in primary care. *J Fam Pract* 1988;26:297-302.
- Hemingway H, Crook AM, Feder G, Banerjee S, Dawson JR, Magee P et al. Underuse of coronary revascularization procedures in patients considered appropriate candidates for revascularization. *N Engl J Med* 2001; 344(9):645-654
- Hoeper EW, Nycz GR, Kessler LG, et al. The usefulness of screening for mental illness. *Lancet* 1984;1:33-35.

Holman R, Lindeboom R, Glas CAW, Vermeulen M, de Haan RJ. Constructing an item bank using item response theory: the AMC Linear Disability Score project. *Health Services and Outcomes Research Methodology* 2003, 4:19-33

Johnson JR, Williams G, Pazdur R. End Points and United States Food and Drug Administration Approval of Oncology Drugs. *J Clin Oncol* 2003; 21:1404-1411

Johnstone A, Goldberg D. Psychiatric screening in general practice: A controlled trial. *Lancet* 1976;1:605-608.

Juniper EF, Gordon GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994;47(1):81-87.

Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;27(Suppl 3):S178-S189.

Kazis LE, Callahan LF, Meenan RF, et al. Health status reports in the care of patients with rheumatoid arthritis. *J Clin Epidemiol* 1990;43:1243-1253.

Laine C, Davidoff F, Lewis CE, et al. Important elements of outpatient care: A comparison of patients' and physicians' opinions. *Ann Intern Med* 1996;125:640-645.

Laidlaw DA, Harrad RA, Hopper CD, et al. Randomised trial of effectiveness of second eye cataract surgery. *Lancet* 1998; 359(9132): 925-929.

Lee BL, Wilson MR. Health-related quality of life in patients with cataract and glaucoma. *J Glaucoma*. 2000 Feb;9(1):87-94.

Lee PP, Kamberg CJ, Hilborne LH, et al. 1993. Cataract Surgery: a literature review and ratings of appropriateness and cruciality. RAND, Santa Monica CA.

Linder M, Chang TS, Scott IU, Hay D, Chambers K, Sibley LM, Weis E. Validity of the visual function index (VF-14) in patients with retinal disease. *Arch Ophthalmol*. 1999 Dec;117(12):1611-6.

Liaw ST, Young D, Farish S. Improving patient-doctor concordance: An intervention study in general practice. *Fam Pract* 1996;13:427-431.

Linn LS, Yager J. The effect of screening, sensitization and feedback on notation of depression. *J Med Educ* 1980;55:942-949.

Lohr KN. Applications of health status assessment measures in clinical practice: Overview of the third conference on advances in health status assessment. *Med Care* 1992;30:MS4 MS14.

Magruder-Habib K, Zung WWK, Feussner JR. Improving physicians' recognition and treatment of depression in general medical care: results from a randomized clinical trial. *Med Care* 1990;28:239-250.

Mangione CM, Phillips RS, Seddon JM, Lawrence MG, Cook EF, Dailey R, Goldman L. Development of the 'Activities of Daily Vision Scale'. A measure of visual functional status. *Med Care*. 1992 Dec;30(12):1111-26

Mathias SD, Fifer SK, Mazonson PD, et al. Necessary but not sufficient: the effect of screening and feedback on outcomes of primary care patients with untreated anxiety. *J Gen Intern Med* 1994;9:606-615.

Mazonson PD, Mathias SD, Fifer SK, et al. The mental health patient profile: Does it change primary care physicians' practice patterns? *J Am Board Fam Pract* 1996;9:336-345.

McDowell I. Screening for psychosocial problems among primary care patients: a pilot study. *Can Med Assoc J* 1987;137:1095-1100.

McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials: a systematic review. *JAMA*. 2003 May 21;289(19):2545-53.

Meyer KB, Espindle DM, DeGiacomo JM, et al. Monitoring dialysis patients' health status. *Am J Kidney Dis* 1994;24:267-279.

Moore JT, Silimperi DR, Bobula JA. Recognition of depression by family medicine residents: the impact of screening. *J Fam Pract* 1978;7:509-513.

Morales LS, Reise SP, Hays RD. Evaluating the equivalence of health care ratings by whites and Hispanics. *Med Care* 2000; 38(5):517-527.

Mönestan E, Wachmesiter L. Impact of Cataract Surgery on the Visual Ability of the Very Old. *Am J Ophthalmol* 2004;137:145-155

Mueller-Marquez BJ. Is there congruence in hospice nurses' view of their patients' quality of life and their hospice patients' view of their own quality of life? *Am J Hosp Palliat Care* 1993;Sept/Oct:5-6.

Mullan PB, Stross JK. Sensitivity to patients' psychosocial concerns: Relationships among ratings by primary care and traditional internal medicine house officers and patient self assessments. *Soc Sci Med* 1990;31:1337-1345.

Navarro V. Why some countries have national health insurance, others have national health services, and the U.S. has neither. *Soc Sci Med*. 1989;28(9):887-98.

Nelson E, Wasson J, Kirk J, et al. Assessment of function in routine clinical practice: Description of the COOP chart method and preliminary findings. *J Chronic Dis* 1987;40(suppl 1):55S-63S.

Nelson EC, Landgraf JM, Hays RD, et al. The functional status of patients: how can it be measured in physicians' offices? *Med Care* 1990;28:1111-1126.

Nelson EC, Landgraf JM, Hays RD, et al. The COOP function charts: A system to measure patient function in physicians' offices. In: Wonca CC, ed. *Functional status measurement in primary care*. New York: Springer-Verlag; 1995, p. 97-131.

Nowers M. Screening for psychiatric illness in the elderly. *Practitioner* 1989;233:1600-1603.

Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *Can Med Assoc J* 1995; 153: 1423-1431

Patrick LD, Erickson P. *Health status and health policy : quality of life in health care evaluation and resource allocation*. New York : Oxford University Press, 1993

Pettitti, Diana B. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis; Methods for Quantitative Synthesis in Medicine*, 2nd Edition, Oxford University Press, 2000.

Rand EH, Badger LW, Coggins DR. Toward a resolution of contradictions: Utility of feedback from the GHQ. *Gen Hosp Psychiatry* 1988;10:189-196.

Reifler DR, Kessler HS, Bernhard EJ, et al. Impact of screening for mental health concerns on health service utilization and functional status in primary care patients. *Arch Intern Med* 1996;156:2593-2599.

Rubenstein LV, Calkins DR, Young RT, et al. Improving patient function: A randomized trial of functional disability screening. *Ann Intern Med* 1989;111:836-842.

Rubenstein LV, McCoy JM, Cope DW, et al. Improving patient quality of life with feedback to physicians about functional status. *J Gen Intern Med* 1995;10:607-614.

Sackett DL, Richardson WS, Rosenberg W, et al. *Evidence-based medicine: How to practice and teach EBM*. London: Churchill Livingstone; 1997.

Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: Attributes and review Criteria. *Quality of Life Research* 11: 193-205, 2002.

Schor EL, Lerner DJ, Malspeis S. Physicians' assessment of functional health status and well-being: The patient's perspective. *Arch Intern Med* 1995;155:3093-14.

Scottish Intercollegiate Guidelines Network (SIGN). Day case cataract surgery. A national clinical guideline. Edinburgh (Scotland): Scottish Intercollegiate Guidelines Network (SIGN); 2001 Aug. 20 p. (SIGN publication; no. 53)

Shapiro S, German PS, Skinner EA, et al. An experiment to change detection and management of mental morbidity in primary care. *Med Care* 1987;25:327-339.

Shiber A, Maoz B, Antonovsky A, et al. Detection of emotional problems in the primary care clinic. *Fam Pract* 1990;7:195-200.

Swinscow TDV. *Statistics at Square One*. London: British Medical Association; 1996.

Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care: The PRIME-MD 1000 study. *JAMA* 1994;272:1749-1756.

Starfield B. Health services research: a working model. *N Engl J Med* 1973; 289:132-136.

Steinberg EP, Tielsch JM, Schein OD, Javitt JC, Sharkey PD, Cassard SD et al. The VF-14: an index of functional impairment in patients with cataract *Arch Ophthalmol* 1994; 112:630-638.

Sterne JAC, Egger M, Smith GD. Investigating and dealing with publication and other biases in meta-analysis *BMJ* 2001;323:101-5

Street RL Jr., Gold WR, McDowell T. Using health status surveys in medical consultations. *Med Care* 1994;32:732-744.

Taylor HR. Cataract: how much surgery do we have to do? *Br J Ophthalmol* 2000;84:12

Tobacman JK, Zimmerman B, Lee P, Hilborne L, Kolder H, Brook RH. Visual Acuity Following Cataract Surgeries In Relation to Preoperative Appropriateness Ratings. *Medical Decision Making* 2003; 23(2):122-130

The Alberta Clinical Practice Guidelines Working Group. Guideline for Surgical & Non-Surgical Management of Cataract in the Otherwise Healthy Adult Eye. Edmonton, AB Canada: The Alberta Clinical Practice Guidelines Program. 2002.

The Royal College of Ophthalmologists. Cataract Surgery Guidelines. [on-line] 2001. (http://www.rcophth.ac.uk/publications/guidelines/cataract_surgery.html)

Thylefors B. A global initiative for the elimination of avoidable blindness. *American Journal of Ophthalmology* 1998;125(1):90-3. 9437318.

Valderas JM, Alonso J, Prieto L, Espallargues M, Castells FX. Content-based interpretation aids for health-related quality of life measures in clinical practice. An example for the visual function index VF-14. *Quality of Life Research* 2004 13: 35–44.

Valderas JM, Rue M, Guyatt G, Alonso J. The impact of the VF-14 Index, a perceived visual function measure, in the routine management of cataract patients. *Quality of Life Research* 2005 (en prensa).

Veloza CA, Lai JS, Mallinson T, Hauselman E. Maintaining instrument quality while reducing items : application of Rasch analysis to a self-report of visual function. *J Outcome Meas* 2000;4(3):667-680.

Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, Selby PJ. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol.* 2004 Feb 15;22(4):714-24.

Wagner AK, Ehrenberg BL, Tran TA, et al. Patient-based health status measurement in clinical practice: a study of its impact on epilepsy patients' care. *Qual Life Res* 1997;6:329-341.

Wasson J, Hays R, Rubenstein L, et al. The short-term effect of patient health status assessment in a health maintenance organization. *Qual Life Res* 1992;1:99-106.

Wasson J, Keller A, Rubenstein L, et al. Benefits and obstacles of health status assessment in ambulatory settings: The clinician's point of view. *Med Care* 1992;30(suppl 5):MS42 MS49.

White P, Atherton A, Hewett G, et al. Using information from asthma patients: a trial of information feedback in primary care. *BMJ* 1995;311:1065-1069.

WHO-OMS. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 1922 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

WHO-OMS. Vision 2020. The global initiative. Fact Sheet 213. 2000.

Woodcock M, Shah S, Smith RJ. Recent advances in customizing cataract surgery. *BMJ* 2004; 328:92-96.

Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999;52(9):861-873.

Wyrwich KW, Wolinsky FD. Identifying meaningful intraindividual change standards for health-related quality of life measures. *J Eval Clin Pract* 2000;6(1):39-49.

Yager J, Linn LS. Physician-patient agreement about depression: notation in medical records. *Gen Hosp Psychiatry* 1981;3:271-276.

Zung WWK, Magill M, Moore JT, et al. Recognition and treatment of depression in a family medicine practice. *J Clin Psychiatry* 1983;44:3-6.

ANEXOS

ANEXO 1. GLOSARIO

Adecuación: grado en que el beneficio esperado de una intervención sanitaria excede las posibles consecuencias negativas con el suficiente margen como para considerar que merece la pena realizarla; no incorpora los costes asociados

Agudeza Visual (AV, VA): integridad y claridad de la visión de cada ojo, habitualmente cuantificada mediante los optotipos de Snellen y cuyo valor normal se sitúa en 1; un valor de 0 indica falta completa de la capacidad para la visión de ese ojo.

Análisis de Rasch: modelo teórico enmarcado en la *Teoría de Respuesta al Ítem*, y que debe su nombre a uno de sus impulsores, el matemático danés Georg Rasch (1901-1980).

Análisis de sensibilidad: procedimiento por el que se evalúan los efectos de las variaciones en los parámetros de un modelo sobre la estabilidad del mismo o sobre sus resultados

Antes/después: estudio epidemiológico cuasiexperimental en el que los grupos de comparación se construyen en relación al momento en que se produce una determinada intervención o acontecimiento.

Atención Primaria: estructura para la provisión de servicios sanitarios integrados con énfasis en las características de longitudinalidad y continuidad de la atención y que opera en el contexto de la familia y de la comunidad.

Calidad de vida relacionada con la salud (CVRS, HRQL): el valor asignado a la duración de la vida modificada por las deficiencias, los estados funcionales, las percepciones y las oportunidades sociales, que están influidas por la enfermedad, las lesiones, el tratamiento médico o las políticas sanitarias y que puede ser asignado por individuos, grupos de individuos o por la sociedad. En este trabajo se han considerado términos sinónimos salud percibida, función percibida, estado funcional o estado de salud, todos ellos términos afines en el contexto de las medidas de resultados en salud centradas en los pacientes.

Cataratas: opacificación total o parcial del *crystalino* que puede ser secundaria muy diversa patología, pero que se asocia principalmente al envejecimiento (catarata senil) y que produce una disminución de la función visual.

Cirugía de cataratas: intervención quirúrgica para la corrección de las *cataratas*; el método más extendido actualmente es la *faccoemulsificación* y colocación de una lente intraocular

Coefficiente de Correlación Intraclase (CCI, ICC): medida de la *concordancia* para variables continuas

Concordancia: grado de acuerdo observado para los valores de una misma variable proporcionados por distintos evaluadores

Contaminación: reproducción parcial (o total) de la intervención en pacientes asignados al grupo control; se facilita en estudios en los que los pacientes de ambos grupos comparten el mismo profesional sanitario o acuden al mismo centro sanitario si no existe cegamiento

Consultas externas: servicios sanitarios adscritos a un hospital en los que se proporciona atención que no requiere ingreso.

Crystalino: cuerpo transparente lenticular biconvexo e intraocular cuya función es la permitir la convergencia de los rayos de luz sobre la retina para la formación de imágenes; su opacificación se conoce como *catarata*

Cuestionario: catálogo de preguntas; los instrumentos de medida más habituales de CVRS son precisamente cuestionarios estandarizados, que pueden ser específicos (de una determinada patología) o genéricos.

Diagnóstico: identificación de la naturaleza de una enfermedad o afección que se puede apoyar en la exploración física y en pruebas complementarias (p. ej, de laboratorio o de imagen).

Dificultad: en la *Teoría de Respuesta al Ítem*, característica intrínseca de un *ítem* que se relaciona con las respuestas proporcionadas por los individuos.

Ensayo clínico aleatorizado (ECA, RCT): Estudio que evalúa de acuerdo a criterios predefinidos los efectos de un procedimiento, técnica o medicamento en al menos dos grupos de pacientes (control e intervención) y en el que la asignación de los pacientes a cada grupo se hace de forma aleatoria.

Escala analógica visual: escala de respuesta de un instrumento de medida consistente en una línea de longitud variable con extremos interpretables por el individuo y en la que el valor de la respuesta lo debe señalar el propio individuo sobre la línea.

Facoemulsificación: técnica quirúrgica consistente en la fragmentación por ultrasonidos del *crystalino* y posterior aspiración de los fragmentos.

Gestión clínica: conjunto de acciones integradas orientadas al diagnóstico, tratamiento y seguimiento de los pacientes en el curso de la atención médica.

Guía de práctica clínica: conjunto de recomendaciones basadas en el conocimiento científico médico y diseñado para orientar la gestión clínica de los pacientes.

Habilidad: en la *Teoría de Respuesta al Ítem*, grado en que el individuo posee la variable latente o el constructo que se pretende medir.

Índice de Función Visual VF-14 (VF14): instrumento para la medición de la *función visual* en pacientes con *cataratas* originalmente desarrollado por Steinberg y cols. en EEUU y adaptado para su uso en población española por Alonso y cols.; consta de catorce actividades de la vida diaria especialmente susceptibles de verse afectadas por las *cataratas*.

Interpretabilidad: facilidad para asignar un significado comprensible a las puntuaciones del instrumento.

Ítem: cada una de las preguntas de un instrumento de medida y, por extensión de las actividades descritas en ellas.

Meta-análisis: habitualmente en el contexto de una *revisión sistemática*, es un método para la obtención de una síntesis cuantitativa de los resultados observados en estudios independientes.

Práctica clínica: ejercicio profesional de la medicina comúnmente establecido en contacto con pacientes individuales

Privado: aplicado a un servicio sanitario, de titularidad distinta a la pública, habitualmente una sociedad o institución con ánimo de lucro.

Proceso de la atención sanitaria: en el modelo de investigación en servicios sanitarios de Donabedian, conjunto de relaciones que se establecen entre el individuo y los servicios sanitarios en relación con la atención médica.

Público: aplicado a un servicio sanitario, aquel cuya titularidad o control recae en la administración pública (local, autonómica o nacional).

Revisión sistemática: revisión rigurosa basada en la identificación, evaluación y síntesis la evidencia proporcionada por estudios individuales siguiendo una metodología específica preestablecida con el objeto de minimizar los *sesgos* y maximizar su replicabilidad; como se basa principalmente en estudios publicados, uno de sus limitaciones potenciales es el *sesgo de publicación*.

Resultado de la atención sanitaria: en el modelo de investigación en servicios sanitarios de Donabedian, producto de la atención sanitaria, y que considera medidas diversas como la mortalidad, la morbilidad o la función; distintos autores clasifican la satisfacción del paciente con la atención recibida como variable de proceso, de resultado o como categoría independiente.

Sesgo: error sistemático por el que se produce una desviación en los resultados o las inferencias de las verdaderas relaciones entre las variables de estudio; el sesgo de publicación es un tipo de sesgo de selección (diferencias sistemáticas entre los individuos que participan en un estudio y los que no) en el que los resultados de los estudios se relacionan con la probabilidad de su publicación.

Teoría de respuesta al ítem (IRT, TRI): modelo teórico para las medidas centradas en los pacientes que asume que la característica que se pretende cuantificar es una variable latente que se relaciona probabilísticamente con las respuestas observadas en un test o instrumento. La estimación de los parámetros del modelo proporciona una estimación

de la variable latente en cada individuo de forma independiente de los ítems escogidos y de la muestra de individuos.

Teoría clásica de los Tests (TCT, CTT): también Teoría Clásica de la Fiabilidad, es el modelo teórico más sencillo para las medidas centradas en los pacientes y basa la estimación de la precisión de un instrumento (fiabilidad) en la correlación entre la puntuación observada y la verdadera. Asume la equivalencia en las respuestas a los ítems del mismo instrumento y por ello la estimación de la fiabilidad es común para todos los individuos.

Tratamiento: conjunto de procedimientos (y su gestión) orientados a combatir una enfermedad

Utilización de servicios sanitarios: en sentido amplio, toda relación entre un servicio sanitario y el individuo, ya sea por iniciativa del propio individuo o mediada por la relación de agencia de los profesionales sanitarios (como la solicitud de pruebas complementarias o la programación de visitas de seguimiento).

Unidimensionalidad: aplicada a un instrumento de medida de CVRS, característica por la cual ningún subconjunto de *ítems* define (teórica o empíricamente) un concepto o constructo distinto del resto; los instrumentos multidimensionales son, por oposición, aquellos que miden más de una dimensión

**ANEXO 2. CUESTIONARIO RECLUTAMIENTO
(PRIMERA ADMINISTRACIÓN ÍNDICE VF14)**

EVALUACIÓN DEL USO SISTEMÁTICO DE LA FUNCIÓN VISUAL PERCIBIDA EN LA ATENCIÓN AL PACIENTE DE CATARATAS

Cuestionario para la entrevista de reclutamiento

ENCUESTADOR	_____	CENTRO	_____	OFTALMÓLOGO	_____
NOMBRE Y APELLIDOS DEL PACIENTE.....					
.....					
Sexo	1. Hombre	2. Mujer	NUMID: _____		

Hola, me llamo [NOMBRE DEL ENCUESTADOR], le llamo desde el Instituto Municipal de Investigación Médica. Como ya le dijo su médico de la vista, el Dr. [NOMBRE DEL MÉDICO RESPONSABLE], estamos haciendo un estudio sobre las cataratas. Quisiera hacerle algunas preguntas sobre su vista. La entrevista dura, aproximadamente quince minutos y quiero recordarle que todo lo que nos diga es confidencial. ¿Quiere hacerme alguna pregunta?

[SI CONTESTA "NO", CONTINUAR. SI CONTESTA "SÍ", RESPONDER A LAS PREGUNTAS.]

¿Le parece bien que empecemos?

[SI CONTESTA "SÍ" ANOTAR LA HORA DE INICIO Y EMPEZAR LA ENTREVISTA. SI CONTESTA "NO" PREGUNTAR CUÁNDO SERÍA UN BUEN MOMENTO Y CONCERTAR CITA.]

Hora de inicio de la entrevista: : ..

A continuación le preguntaré cuánta dificultad tiene Ud. para hacer algunas actividades de vida diaria, a causa de su vista, incluso cuando lleva gafas. Si Ud. no hace alguna de las actividades por razones distintas a sus problemas de la vista, por favor indíquemelo.

A causa de la vista e incluso llevando gafas, ¿cuánta dificultad tiene usted para realizar las siguientes tareas?

2. A causa de la vista ¿cuánta dificultad tiene para **leer letras pequeñas como la guía telefónica, los nombres de los medicamentos o las etiquetas de artículos de comida?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

3. A causa de la vista ¿cuánta dificultad tiene para **leer un periódico o un libro?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

4. A causa de la vista ¿cuánta dificultad tiene para **leer letras grandes de un libro o de un periódico o los números del teléfono?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

5. A causa de la vista ¿cuánta dificultad tiene para **reconocer personas cuando están cerca de usted?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

6. A causa de la vista ¿cuánta dificultad tiene para **ver escalones, peldaños o el bordillo de la acera?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

7. A causa de la vista ¿cuánta dificultad tiene para **leer letreros de las calles y tiendas, los números de las casas o ver los semáforos?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

8. A causa de la vista ¿cuánta dificultad tiene para **hacer trabajos manuales finos como coser, arreglar un enchufe o clavar un clavo?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

9. A causa de la vista ¿cuánta dificultad tiene para **hacer crucigramas, rellenar un impreso, o hacer una quiniela?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

10. A causa de la vista ¿cuánta dificultad tiene para **jugar a las cartas, al dominó o al bingo?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

11. A causa de la vista ¿cuánta dificultad tiene para **participar en actividades como la petanca, buscar setas, cuidar plantas, mirar escaparates?**, incluso llevando gafas.
- 0 Ninguna
 - 1 Poca
 - 2 Bastante
 - 3 Mucha
 - 4 Incapaz de hacerla
 - 5 No lo hago por otras razones distintas de la vista
12. A causa de la vista ¿cuánta dificultad tiene para **cocinar?**, incluso llevando gafas.
- 0 Ninguna
 - 1 Poca
 - 2 Bastante
 - 3 Mucha
 - 4 Incapaz de hacerla
 - 5 No lo hago por otras razones distintas de la vista
13. A causa de la vista ¿cuánta dificultad tiene para **ver la televisión?**, incluso llevando gafas.
- 0 Ninguna
 - 1 Poca
 - 2 Bastante
 - 3 Mucha
 - 4 Incapaz de hacerla
 - 5 No lo hago por otras razones distintas de la vista
14. A causa de la vista ¿cuánta dificultad tiene para **conducir de día?**, incluso llevando gafas.
- 0 Ninguna
 - 1 Poca
 - 2 Bastante
 - 3 Mucha
 - 4 Incapaz de hacerla
 - 5 No lo hago por otras razones distintas de la vista
15. A causa de la vista ¿cuánta dificultad tiene para **conducir de noche?**, incluso llevando gafas.
- 0 Ninguna
 - 1 Poca
 - 2 Bastante
 - 3 Mucha
 - 4 Incapaz de hacerla
 - 5 No lo hago por otras razones distintas de la vista

Ahora quisiera saber si durante el **último mes** ha tenido alguna de las siguientes molestias

16.1 Durante el último mes ¿ha tenido **visión doble**?

1 Sí → 16.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

17.1 Durante el último mes ¿ha **visto resplandores, halos o anillos alrededor de las luces**?

1 Sí → 17.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

18.1 Durante el último mes ¿ha **tenido visión borrosa**?

1 Sí → 18.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

19.1 Durante el último mes ¿ha **notado que ve los colores diferentes que antes**?

1 Sí → 19.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

20.1 Durante el último mes ¿ha **observado un empeoramiento de la visión**?

1 Sí → 20.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

Hora de finalización de la entrevista: ____:____

Resultado final de la entrevista:

- .1 Entrevista. completa
- 2 Entrevista incompleta
- 3 Falla cognitiva
- 4 Mal reclutado
- 5 Rechaza entrevista
- 6 No localizado
- 7 Otros.....

Grado de comprensión de la entrevista

- 1. Excelente
- 2. Bueno
- 3. Regular
- 4. Malo

OBSERVACIONES

**ANEXO 3. CUESTIONARIO SEGUIMIENTO
(SEGUNDA ADMINISTRACIÓN ÍNDICE VF14)**

EVALUACIÓN DEL USO SISTEMÁTICO DE LA FUNCIÓN VISUAL PERCIBIDA EN LA ATENCIÓN AL PACIENTE DE CATARATAS

Cuestionario para la entrevista telefónica

ENCUESTADOR ____ CENTRO ____ OFTALMÓLOGO ____

NOMBRE Y APELLIDOS DEL PACIENTE

.....

Sexo 1. Hombre 2. Mujer NUMID ____

Hola, me llamo [NOMBRE DEL ENCUESTADOR], le llamo desde el Instituto Municipal de Investigación Médica. Como ya le dijo su médico de la vista, el Dr. [NOMBRE DEL MÉDICO RESPONSABLE], estamos haciendo un estudio sobre las cataratas. Quisiera hacerle algunas preguntas sobre su vista. La entrevista dura, aproximadamente quince minutos y quiero recordarle que todo lo que nos diga es confidencial. ¿Quiere hacerme alguna pregunta?

[SI CONTESTA "NO", CONTINUAR. SI CONTESTA "SÍ", RESPONDER A LAS PREGUNTAS.]

¿Le parece bien que empecemos?

[SI CONTESTA "SÍ" ANOTAR LA HORA DE INICIO Y EMPEZAR LA ENTREVISTA. SI CONTESTA "NO" PREGUNTAR CUÁNDO SERÍA UN BUEN MOMENTO Y CONCERTAR CITA]

Hora de inicio de la entrevista: _:_

A continuación le preguntaré cuánta dificultad tiene Ud. para hacer algunas actividades de vida diaria, a causa de su vista, incluso cuando lleva gafas. Si Ud. no hace alguna de las actividades por razones distintas a sus problemas de la vista, por favor indíquemelo.

A causa de la vista e incluso llevando gafas, ¿cuánta dificultad tiene usted para realizar las siguientes tareas?

2. A causa de la vista ¿cuánta dificultad tiene para **leer letras pequeñas como la guía telefónica, los nombres de los medicamentos o las etiquetas de artículos de comida?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

3. A causa de la vista ¿cuánta dificultad tiene para **leer un periódico o un libro?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

4. A causa de la vista ¿cuánta dificultad tiene para **leer letras grandes de un libro o de un periódico o los números del teléfono?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

5. A causa de la vista ¿cuánta dificultad tiene para **reconocer personas cuando están cerca de usted?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

6. A causa de la vista ¿cuánta dificultad tiene para **ver escalones, peldaños o el bordillo de la acera?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

7. A causa de la vista ¿cuánta dificultad tiene para **leer letreros de las calles y tiendas, los números de las casas o ver los semáforos?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

8. A causa de la vista ¿cuánta dificultad tiene para **hacer trabajos manuales finos como coser, arreglar un enchufe o clavar un clavo?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

9. A causa de la vista ¿cuánta dificultad tiene para **hacer crucigramas, rellenar un impreso, o hacer una quiniela?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

10. A causa de la vista ¿cuánta dificultad tiene para **jugar a las cartas, al dominó o al bingo?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

11. A causa de la vista ¿cuánta dificultad tiene para **participar en actividades como la petanca, buscar setas, cuidar plantas, mirar escaparates?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

12. A causa de la vista ¿cuánta dificultad tiene para **cocinar?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

13. A causa de la vista ¿cuánta dificultad tiene para **ver la televisión?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

14. A causa de la vista ¿cuánta dificultad tiene para **conducir de día?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

15. A causa de la vista ¿cuánta dificultad tiene para **conducir de noche?**, incluso llevando gafas.

- 0 Ninguna
- 1 Poca
- 2 Bastante
- 3 Mucha
- 4 Incapaz de hacerla
- 5 No lo hago por otras razones distintas de la vista

Ahora quisiera saber si durante el **último mes** ha tenido alguna de las siguientes molestias

16.2 Durante el último mes ¿ha tenido **visión doble**?
1 Sí → 16.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

17.1 Durante el último mes ¿ha **visto resplandores, halos o anillos alrededor de las luces**?

1 Sí → 17.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

18.1 Durante el último mes ¿ha **tenido visión borrosa**?

1 Sí → 18.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

19.1 Durante el último mes ¿ha **notado que ve los colores diferentes que antes**?

1 Sí → 19.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

20.1 Durante el último mes ¿ha **observado un empeoramiento de la visión**?

1 Sí → 20.2 ¿Cuánta molestia le ha producido?

- 1 Ninguna
- 2 Poca
- 3 Bastante
- 4 Mucha

2 No

Recuerde que sus respuestas no influirán en su tratamiento médico de ninguna forma, ya que no serán vistas por su médico ni por el personal sanitario. Sus respuestas son importantes para mejorar los servicios médicos. Pensando en la asistencia sanitaria que recibe en relación con su problema de cataratas, "cómo calificaría cada uno de los siguientes aspectos: excelente, muy buena, buena, regular o mala?

21. la **calidad general** de la atención médica que ha recibido en (CENTRO) por sus problemas de la vista?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

21. la **amabilidad** y cortesía que le muestran su **médico** de la vista?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

22. la **amabilidad** y cortesía que le muestra el resto del **personal**?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

23. el **interés personal** de su médico de la vista por sus problemas de la vista?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

24. la **atención** que le presta el médico de la vista **a lo que Ud. dice?**

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

25. el **respeto** que muestra por Ud. el médico de la vista?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

26. la **confianza** que Ud. tiene en su médico de la vista?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

27. la **duración** de las visitas con su médico de la vista?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

28. las **pruebas** que le ha hecho su médico de la vista antes de decidir si tendría que operarse o no?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

29. las **explicaciones** de su médico de la vista acerca de las pruebas que le han hecho?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

30.1 ¿**Ha podido preguntar** cosas sobre las cataratas y la operación?

- 1 Sí → 30.2 Cómo han sido las **respuestas** de su médico de la vista a sus

preguntas sobre las cataratas y la operación?

- 0 Mala
- 1 Regular
- 2 Buena
- 3 Muy buena
- 4 Excelente
- 5 NS/NC

2 No

31.1. ¿Han contado con **su opinión** para decidir si operar o no sus cataratas?

1 Sí → 31.2 ¿**Cuánto**?

- 1 Mucho
- 2 Bastante
- 3 Poco
- 4 Nada

2 No

32. En **general**, Ud. diría que su **salud** es...

- 1 "Excelente"
- 2 "Muy buena"
- 3 "Buena"
- 4 "Regular"
- 5 "Mala"

A continuación le haré, unas preguntas para saber si tiene alguna de las siguientes enfermedades:

- | | | |
|--|------|------|
| 33. Artrosis o reumatismo | 1 Sí | 2 No |
| 34. Bronquitis crónica, asma o enfisema | 1 Sí | 2 No |
| 35. Hipertensión o presión alta | 1 Sí | 2 No |
| 36. Diabetes o azúcar en la sangre | 1 Sí | 2 No |
| 37. Enfermedades del corazón , como la angina de pecho o el infarto | 1 Sí | 2 No |
| 38. Embolias o trombosis cerebral | 1 Sí | 2 No |
| 39. Sordera | 1 Sí | 2 No |
| 40. Depresión o trastornos nerviosos | 1 Sí | 2 No |
| 41. Cáncer | 1 Sí | 2 No |

NOTA: LA NUMERACIÓN CONTINÚA CON 44. NO EXISTEN LAS PREGUNTAS 42 NI 43

Ahora le haré unas preguntas sobre su capacidad para realizar las actividades habituales de la vida diaria . Es decir, ¿hasta qué punto puede valerse por sí mismo/a?

44.1 ¿Necesita ayuda para **sentarse** en una silla o para **echarse y levantarse** de la cama?

1. Sí → 44.2 ¿Cuánto tiempo hace que necesita esta ayuda?
1. Días..... __ __
2. Meses... __ __
3. NS/NC
2. No
3. NS/NC

45.1 ¿Necesita ayuda para prepararse la ropa o **vestirse**?

1. Sí → 45.2. ¿Cuánto tiempo hace que necesita esta ayuda?
1. Días..... __ __
2. Meses... __ __
3. NS/NC
2. No
3. NS/NC

46.1 ¿Necesita ayuda para ir al **lavabo** o para utilizarlo?

1. Sí → 46.2. ¿Cuánto tiempo hace que necesita esta ayuda?
1. Días..... __ __
2. Meses... __ __
3. NS/NC
2. No
3. NS/NC

47.1 ¿Necesita ayuda para entrar o salir de la **bañera** o ducha?

1. Sí → 47.2 ¿Cuánto tiempo hace que necesita esta ayuda?
1. Días..... __ __
2. Meses... __ __
3. NS/NC
2. No
3. NS/NC

48.1 ¿Necesita ayuda para **comer**, por ejemplo, cortar la carne o pelar la fruta?

1. Sí → 48.2 ¿Cuánto tiempo hace que necesita esta ayuda?
1. Días..... __ __
2. Meses... __ __
3. NS/NC
2. No
3. NS/NC

Para terminar le preguntaré algunos datos personales.

49. ¿En qué **fecha nació** Ud.? _ _ / _ _ / 19 _ _

50. Actualmente, ¿cuál es su **estado civil**?

1. soltero/a
2. casado/a
3. divorciado/a
4. separado/a
5. viudo/a

51. ¿Qué **estudios** ha **completado** de los que a continuación le indicamos?

1. No ha aprendido a leer ni escribir
2. Enseñanza primaria
3. Bachillerato elemental/comercio
4. Bachillerato superior/BUP/COU/FP
5. Estudios universitarios

52. ¿Cuál es su situación laboral actualmente?

1. activo
2. en paro
3. jubilado
4. estudiante
5. ama de casa
6. invalidez
7. otros.....

SI CONTESTA "ACTIVO" PASAR A **53**; SI NO, SALTAR A **56**.

→ 53. ¿Cuánta **dificultad** tiene usted **en su trabajo** a causa de la vista?

1. Mucha
2. Bastante
3. Poca
4. Ninguna

→ 54.1 ¿Ha tenido que **cambiar de trabajo** por culpa de su vista?

1. Sí →54.2 ¿Cuál era su trabajo antes de verse obligado a cambiarlo?
.....
2. No

→ 55. ¿**Debería hacerlo** si su vista empeorara?

1. Sí
2. No

56. ¿Cuál ha sido su **ocupación** durante el último año trabajado?
.....

57. ¿Tiene Ud. personas a su cargo o que dependan de su ayuda para valerse?

1. Sí

2. No

58. ¿Vive Ud. **solo**?

1. Sí
2. No

59. ¿Cuál de los siguientes **seguros sanitarios** tiene usted (PUEDE ESCOGER MÁS DE UNA OPCIÓN DE RESPUESTA)?

1. seguridad social
2. mutua privada
3. mutua obligatoria
4. otros:

.....

60. ¿Le ha dicho su oftalmólogo/médico de la vista si se va a operar de cataratas?

1. Sí



61. ¿Sabe ya cuándo se operará?

1. Sí → 63. / /

2. No

2. No

Hora de finalización de la entrevista: ____:____

Fecha de la entrevista: __/__/__

Resultado final de la entrevista:

- .1 Entrevista. completa
- 2 Entrevista parcialmente completa
- 3 Falla cognitiva
- 4 Mal reclutado
- 5 Rechaza entrevista
- 6 No localizado
- 7 Otros.....

Grado de comprensión de la entrevista

- 1. Excelente
- 2. Bueno
- 3. Regular
- 4. Malo

OBSERVACIONES

ANEXO 4. FORMULARIO DATOS CLÍNICOS

Evaluación del uso sistemático de la función visual percibida en la atención al paciente de cataratas



FORMULARIO CLÍNICO

Dr. _____

DATOS DEL PACIENTE (NUMID __/__/__)	FECHA __/__/99
Nombre _____	
Dirección _____	
Teléfono(s) _____	

¿Es la primera evaluación de cataratas del paciente en su consulta?

1. Sí 2. No

Señale si el paciente padece alguna ENFERMEDAD OCULAR además de la catarata (1):

- | | |
|-------------------------------|-------------------------------------|
| 1. Ambliopía | 5. Desprendimiento de retina previo |
| 2. Degeneración macular senil | 6. Otras (especificar) |
| 3. Retinopatía diabética | |
| 4. Glaucoma o PIO elevada | |

Señale si el paciente padece alguna OTRA ENFERMEDAD CONOCIDA (2):

- | | |
|--------------------------------|--|
| 1. Diabetes Mellitus | 5. Enfermedad pulmonar obstructiva crónica |
| 2. Hipertensión arterial | 6. Otras (especificar) |
| 3. Cardiopatía isquémica | |
| 4. Accidente vascular cerebral | |

AGUDEZA VISUAL CORREGIDA (notación decimal)

- | | | |
|--|--|---|
| OD 0, (3.1)
(cuenta dedos am.) | OI 0, (3.2)
(cuenta dedos am.) | ambos 0, (3.3)
(cuenta dedos am.) |
|--|--|---|

PRESIÓN INTRAOCULAR **OD** /mmHg (4.1) **OI** /mmHg (4.2)

FONDO DE OJO

CRISTALINO

¿Se ha examinado el fondo de ojo?

- | | |
|--|--|
| <p>OD (5.1) OI (5.2)</p> <p>1.....normal 1.</p> <p>2..... catarata nuclear 2.</p> <p>3. catarata cortical 3.</p> <p>4. catarata subcapsular posterior 4.</p> <p>5. catarata hipermadura 5.</p> <p>6. catarata congénita 6.</p> <p>7. catarata secundaria 7.</p> | <p>OD (6.1) OI (6.2)</p> <p>1.Sí: normal..... 1</p> <p>2.Sí: anormal..... 2.</p> <p>3.Sí: no valorable..... 3.</p> <p>4.No se ha examinado.....4.</p> |
|--|--|

POR FAVOR, DÉ LA VUELTA AL FORMULARIO

PUNTUACIÓN VF14: _____

Marque en la línea el punto que mejor describa **la dificultad que tiene el paciente para valerse por sí mismo**, a causa de las cataratas (7):

ninguna dificultad |-----| máxima dificultad

Marque en la línea el punto que mejor describa **la dificultad que tiene el paciente para realizar las actividades diarias**, a causa de las cataratas(8):

ninguna dificultad |-----| máxima dificultad

Marque en la línea el punto que mejor describa **la necesidad de la cirugía** en este paciente (9):

ninguna necesidad |-----| máxima necesidad

Marque en la línea el punto que mejor describa **la utilidad de la cirugía** en este paciente (10):

ninguna utilidad |-----| máxima utilidad

Idealmente y sin las restricciones que imponen las listas de espera, Ud **indicaría la cirugía** para este paciente (11):

1. A corto plazo (menos de cuatro meses)
2. A medio plazo (entre cuatro meses y un año)
3. A largo plazo (entre uno y dos años)
4. A muy largo plazo (más de dos años)

¿Ha realizado indicación de cirugía en esta visita? (12)

1. Sí

➔ **¿Se ha incluido al paciente en lista de espera?** (13)

1. Sí 2. No

➔ **Señale qué ojo(s) se va(n) a intervenir** (14): 1.OD 2.OI

2. No

¡MUCHAS GRACIAS!

