



Evaluación de Estudios de Calidad de Vida Relacionada con la Salud

Manuel Madrigal de Torres
Antonio Velandrino Nicolás
María Ruzafa Martínez

1. INTRODUCCIÓN.

1.1. Importancia de la medida de la Calidad de Vida relacionada con la Salud (CVRS) en Atención Primaria.

1.1.1 Facilitar la elección de la estrategia terapéutica más adecuada.

1.1.2 Documentar y evaluar la CVRS de determinadas poblaciones.

2. PASOS A SEGUIR EN LA EVALUACIÓN DE UN ESTUDIO DE CVRS.

2.1. ¿Contiene el instrumento aquellas dimensiones esenciales a medir de acuerdo a la población y enfermedad?.

2.1.1 Dimensiones de la CVRS.

2.1.2 Clasificación de los instrumentos de medida de la CVRS.

2.2. ¿Se especifican las propiedades de medición (psicométricas) del instrumento utilizado para medir la CVRS?.

2.2.1 El modelo conceptual y de medida.

2.2.2 Fiabilidad.

2.2.3 Validez.

2.2.4 Sensibilidad a los cambios.

2.2.5 Interpretabilidad.

2.2.6 Facilidad de administración: carga de respuesta y de administración.

2.2.7 Formas alternativas de administración.

2.3. ¿Ha sido traducido, adaptado y validado al idioma y cultura del país donde lo vamos a emplear (o se ha empleado)?.

2.4. ¿Qué calidad de la evidencia y aplicabilidad clínica proporciona el estudio?.

2.4.1 ¿Son validas estas evidencias?.

2.4.2 ¿Son importantes estos resultados?.

Aplicabilidad de los resultados en la asistencia a mis pacientes.

3. PASOS A SEGUIR A LA HORA DE SELECCIONAR UN INSTRUMENTO PARA MEDIR LA CVRS EN NUESTRA PRÁCTICA CLÍNICA.

4. BIBLIOGRAFÍA.

5. PUNTOS CLAVE.

Escenario Clínico

Varón de 46 años. En una revisión rutinaria se le ha detectado una Hipertensión arterial. Tras la detección, las cifras medias fueron de 156/104 mm Hg. El estudio inicial (exploración, analítica de sangre y orina, ECG, Rx Tórax) fue normal. No fuma, ni toma ninguna medicación. Tras seis meses de tratamiento con medidas higiénico dietéticas las cifras de TA son de 142/98. Presenta una obesidad leve con un IMC de 28. Nos planteamos comenzar con un tratamiento farmacológico.

17

→ 1

INTRODUCCION

En general, curar a un paciente, aumentar su supervivencia, prevenir futuros problemas de salud, o aliviarle de sus molestias son razones suficientes para prescribir un tratamiento. Pero si dicho tratamiento produce un deterioro importante de la Calidad de Vida Relacionada con la Salud (CVRS), el paciente, en ocasiones, puede plantearse si una pequeña ganancia en salud o el beneficio producido por el tratamiento, compensan una posible pérdida en su calidad de vida. Por tanto, sin información acerca del efecto de una medicación sobre la CVRS, ni el médico ni el paciente pueden tomar la decisión apropiada con respecto a un determinado tratamiento. La importancia de la CVRS se ha ido incrementando en los últimos 20 años, llegando a ser un objetivo central de la asistencia sanitaria y una medida de resultados en salud “centrada en el paciente”.

Todas las definiciones de CVRS derivan más o menos de las definiciones de salud y calidad de vida dadas por la Organización Mundial de la Salud (OMS), que ya en 1948 definió la salud como “... un estado de completo bienestar físico, emocional y social y no sólo como la ausencia de enfermedad”¹. Posteriormente, la OMS en 1994 describió la calidad de vida como la “percepción del individuo de su posición en la vida en el contexto de la cultura y sistema de valores en los que vive y en relación con sus objetivos, expectativas, estándares y preocupaciones”². Todo esto implica que el concepto de calidad de vida no puede ser de ningún modo independiente de las normas culturales, patrones de conducta y expectativas de cada persona, teniendo estos aspectos una relevancia especial a la hora de construir y/o adaptar instrumentos para medir la CVRS. Por tanto, la evaluación de Calidad de Vida Relacionada con la Salud en un paciente representa el impacto que una enfermedad y el tratamiento consiguiente tienen sobre la percepción del paciente de su satisfacción y bienestar físico, psíquico, social y espiritual³. Por ello, si se pretende evaluar la CVRS se deberá utilizar un instrumento que evalúe los efectos del estado de salud, y solo de él, sobre la vida normal del individuo.

La Organización Mundial de la Salud en 1995 creó un grupo multicultural de expertos Calidad de Vida (Grupo WHOQOL) que estableció una serie de puntos de consenso, en relación a las medidas de calidad de vida relacionada con la salud, que han sido ampliamente aceptados por diversos grupos de investigadores⁴.

Las medidas de la CVRS deben ser:

1. Subjetivas: Recoger la percepción subjetiva del impacto de la enfermedad sobre el paciente, el cual se encuentra influenciado por sus experiencias, creencias y expectativas.
2. Multidimensionales: Revelar diversos aspectos de la vida del individuo, en los niveles físico, emocional, social, interpersonal, etc.
3. Incluir sentimientos positivos y negativos.
4. Registrar la variabilidad en el tiempo: La edad, la etapa vital que se atraviesa (niñez, adolescencia, adultez, adulto mayor) y el estadio de la enfermedad que se padece, marcan diferencias importantes en los aspectos que se valoran.

1.1. IMPORTANCIA DE LA MEDIDA DE LA CVRS EN ATENCIÓN PRIMARIA

La evaluación de la calidad de vida en atención primaria se considera hoy día imprescindible, ya que es una medida final de resultado en salud que se centra más en la persona que en la enfermedad e incorpora la percepción del paciente como una necesidad en la evaluación de resultados en salud. A pesar de ello, su utilización sigue siendo limitada, probablemente por desconocimiento, sobrecarga asistencial o falta de convicción por parte de los profesionales sobre su beneficio en la atención y mejoría de los pacientes⁵.

Por otra parte, la incorporación de medidas de la CVRS en las consultas de atención primaria resulta de gran utilidad, pues sirve de apoyo en la toma de decisiones clínicas y contribuye a mejorar la relación médico-paciente al posibilitar al enfermo la participación en la elección de la opción más conveniente. Su principal utilidad reside en⁶:

1.1.1. Facilitar la elección de la estrategia terapéutica más adecuada:

- i) Ante alternativas terapéuticas de eficacia similar permite decidir cual elegir en el caso de que existan diferencias en la CVRS que proporcionan. Por ejemplo, en un ensayo clínico aleatorizado que evaluaba la terapia antihipertensiva en hombres comparando captopril vs. enalapril, utilizando varias escalas de CVRS, la conclusión fue que aunque los tratamientos eran similares en términos de eficacia y seguridad sin embargo producían cambios sustanciales y diferentes en la CVRS.⁷
- ii) En problemas de salud moderados y/o crónicos en los que los tratamientos tienen como objetivo controlar la sintomatología hay que asegurarse que su uso no disminuye la CVRS.
- iii) En intervenciones terapéuticas que provocan importantes efectos adversos que pueden afectar de forma negativa a la CVRS resulta una información muy útil para el paciente.
- iv) En enfermedades que carecen de un marcador clínico objetivo y validado que muestre la afectación en el paciente (migraña, epilepsia, ansiedad, etc...) facilita la elección de un tratamiento u otro.
- v) En la aplicación de tratamientos preventivos en enfermedades asintomáticas, al confirmar que su uso no disminuye la CVRS de los pacientes.⁸

1.1.2. Documentar y evaluar la CVRS de determinadas poblaciones:

- i) En grupos poblacionales determinados, con algunas características sociodemográfica y/o patologías de interés. Por ejemplo, la población de adictos a drogas, que obtienen los peores resultados en términos de actividades sociales, sentimientos y calidad de vida propiamente dicha.
- ii) En pacientes cumplidores y no cumplidores del tratamiento. Posibilita el análisis de las variables determinantes de la adherencia al tratamiento y la toma de decisiones.
- iii) En estudios poblacionales permite monitorizar los niveles de salud poblacional más allá de las estadísticas de mortalidad.
- iv) Su asociación con fenómenos de salud relevantes predice de manera independiente la mortalidad, la hospitalización⁹ y la utilización de los servicios de atención primaria¹⁰. Por ejemplo, en un ensayo clínico sobre los efectos de la rehabilitación pulmonar en los resultados fisiológicos y psicosociales de pacientes con EPOC se mostró que la CVRS medida con la Quality of Well-Being Scale predecía de manera independiente la mortalidad a los 6 años¹¹.
- v) Es útil evaluar la mejoría en la CVRS cuando esta medida sea el resultado clínico más importante del tratamiento administrado (cuidados paliativos, antieméticos, etc.)
- vi) Para comprobar si los marcadores fisiopatológicos de enfermedad y/o los indicadores clínicos específicos se correlacionan o no con la CVRS que tiene el paciente. Pues se ha visto que con frecuencia esta asociación es débil o como mucho moderada. Por ejemplo, los indicadores clínicos del asma tienen una correlación inferior a 0,7 con la CVRS percibida por el paciente medida con el Asthma Quality of Life Questionnaire¹².

En definitiva, la medición de la CVRS permite, no sólo la evaluación de los resultados en ensayos clínicos e investigaciones sanitarias, sino también la evaluación de las necesidades reales de una población y, por consiguiente, la planificación adecuada de medidas de intervención sanitaria y de distribución de recursos.

17

→ 2 PASOS A SEGUIR EN LA EVALUACIÓN DE UN ESTUDIO DE CVRS

En la evaluación de un estudio de CVRS se deben considerar los siguientes aspectos:

¿Contiene el instrumento aquellas dimensiones esenciales a medir de acuerdo a la población y enfermedad?

¿Se especifican las propiedades de medición (psicométricas) del instrumento utilizado para medir la CVRS?

¿Ha sido traducido, adaptado y validado al idioma y cultura del país donde se va a emplear (o se ha empleado)?

¿Qué calidad de la evidencia y aplicabilidad clínica proporciona el estudio?

2.1. ¿CONTIENE EL INSTRUMENTO AQUELLAS DIMENSIONES ESENCIALES A MEDIR DE ACUERDO A LA POBLACIÓN Y ENFERMEDAD?

La CVRS constituye un concepto o idea abstracta y compleja, que no puede ser medida directamente y que sólo es medible de forma indirecta mediante su transformación en una variable sintética (“constructo”) compuesta por un conjunto de variables intermedias o dimensiones, cada una de las cuales contribuye a cuantificar algún rasgo o aspecto relevante de los que creemos que engloba este concepto abstracto que es la CVRS. Es decir, las dimensiones que medimos constituyen diferentes aspectos de la CVRS, como pueden ser el funcionamiento físico, el funcionamiento social, la percepción de bienestar, etc.

Para investigar estas dimensiones se diseñan instrumentos de medida que deben reunir una serie de propiedades técnicas que nos garanticen que estamos midiendo aquello que queremos medir y con un error de medición lo más pequeño posible. Los instrumentos de medida de la CVRS suelen ser cuestionarios validados, que con frecuencia contienen varias escalas que se corresponden con las distintas dimensiones que conforman el constructo CVRS. Por tanto, una escala es una serie de preguntas, para las que no hay respuestas correctas o incorrectas, que pretenden cuantificar una determinada característica, en nuestro caso, esta característica es una de las dimensiones del cuestionario (v.g.: percepción del bienestar, funcionamiento social, etc.). La escala puede medir esta característica o dimensión en distintos niveles de medida (nominal, ordinal, intervalo o razón). Generalmente, cuando las respuestas a los diversos ítems vienen en un nivel de medida de intervalo o razón, se pueden sumar obteniendo una puntuación total de la escala. Y sumando las puntuaciones obtenidas en cada escala se obtiene la puntuación total del cuestionario.

En síntesis, estas escalas representan dimensiones o rasgos de un concepto abstracto que no puede ser medido directamente, pero que sí puede ser medido indirecta o parcialmente a través de los ítems o preguntas de cada dimensión. Como no existe ninguna regla teórica que contribuya a la delimitación del número de dimensiones a considerar, éstas se suelen dejar a los conocimientos y experiencia del investigador que diseña el instrumento de medida (cuestionario), y crea lo que se denomina un modelo conceptual.



A la hora de evaluar un cuestionario de CVRS deberemos analizar si se omitieron aspectos importantes de la CVRS. Aunque las áreas ó dimensiones cubiertas por el cuestionario empleado pueden ser relevantes para los pacientes, otros aspectos igualmente importantes pueden haber sido omitidos. Por ejemplo, no es infrecuente que cuestionarios que abordan aspectos importantes de la CVRS de pacientes con una determinada enfermedad, no consideren los efectos indeseables de la medicación empleada o no incluyan una lista de efectos adversos y su frecuencia de aparición. De igual manera, deberíamos tener en cuenta cualquier otro aspecto que consideremos relevante y verificar que haya sido tenido en cuenta por los investigadores.

2.1.1. Dimensiones de la CVRS.

Los aspectos relevantes con respecto a la CVRS pueden variar de unos estudios a otros, pero según el Grupo WHOQoLg en los cuestionarios de CVRS los ítems se agrupan por lo general en las siguientes dimensiones⁴:

1. Funcionamiento físico: grado en que la salud limita las actividades físicas como el cuidado personal, andar, subir escaleras, cargar peso y los esfuerzos moderados e intensos.
2. Funcionamiento emocional: grado en que el sufrimiento psicológico, la falta de bienestar emocional, la ansiedad y la depresión interfieren en las actividades diarias.
3. Funcionamiento social: grado en que los problemas de salud física o emocional interfieren en la vida social habitual, con la familia, amigos y vecinos.
4. Rol funcional: grado en que la salud interfiere en el trabajo y otras actividades diarias, incluyendo el rendimiento y el tipo de actividades comunitarias.
5. Funcionamiento cognitivo: grado en que los problemas cognitivos interfieren en el trabajo o en las actividades diarias.
6. Percepción de salud general y de bienestar: valoración personal de la salud que incluye salud actual, perspectivas y resistencia a enfermar.
7. Síntomas (específicos de cada enfermedad).

Dado que a veces es utópico intentar evaluar todos estos conceptos simultáneamente, en muchas ocasiones los instrumentos de medida de la CVRS solo incluyen ítems para las dimensiones funcionamiento físico, social y emocional. Pero no hay que confundir un instrumento o cuestionario de medida de la CVRS con una simple escala de severidad de los síntomas de una enfermedad o con una lista de síntomas; ni con una escala diagnóstica, o una escala de incapacidad, ansiedad, depresión, fatiga, dolor, etc.

2.1.2. Clasificación de los instrumentos de medida de la CVRS

Existen diferentes criterios a la hora de clasificar los instrumentos de medida de la CVRS, aunque la más aceptada es la propuesta por Guyatt et al¹³, que distingue entre instrumentos genéricos e instrumentos específicos para una enfermedad.

i) Los instrumentos de medición de la CVRS genéricos no hacen referencia, ni están relacionados con algún tipo de enfermedad concreta. Se pueden aplicar tanto a la población general como a grupos específicos de pacientes, dado que suelen incluir un amplio espectro de dimensiones de la CVRS. Según su estructura se clasifican en: a) Perfiles de Salud: incluyen dimensiones genéricas por lo que suelen ser aplicables a pacientes con diferentes estados de salud, independientemente de la edad y del tipo de patología. Permiten efectuar comparaciones de la CVRS global en pacientes con diferentes patologías y b) Medidas de Utilidad o Preferencia: se basan en las preferencias que las personas asignan a diferentes estados de salud y están especialmente diseñadas para realizar estudios de costes y toma de decisiones con repercusiones económicas.

ii) Los instrumentos de medición de la CVRS específicos se centran en aspectos específicos del estado de salud, e incluyen dimensiones de CVRS de una determinada patología, capacidad funcional, o de un determinado grupo de pacientes. Normalmente incluyen preguntas sobre el impacto de los síntomas. La base para este tipo de aproximación es el potencial incremento de la capacidad del instrumento para detectar mejoras o deterioros de la CVRS a lo largo del tiempo (sensibilidad al cambio) al incluir aspectos relacionados con el problema específico.

iii) En ocasiones se puede usar una mezcla de los dos tipos de instrumentos descritos previamente o usar varios específicos. Tenemos así otros instrumentos de medición de la CVRS¹⁴:

- Instrumentos genéricos modificados: la modificación se hace respecto a una población de interés.
- Instrumentos genéricos con suplementos específicos de enfermedad: pretenden medir dimensiones adicionales específicas con respecto a una enfermedad concreta que no están contenidas en un instrumento genérico.
- Instrumentos para aspectos específicos de la CVRS: tratan de explorar con mayor profundidad algunas dimensiones concretas o específicas de la calidad de vida. Estos instrumentos deben ser usados en conjunción con cuestionarios más generales.
- Baterías de cuestionarios: son colecciones de medidas específicas que se puntúan independientemente y se transmiten como puntuaciones individuales.
- Formatos abreviados de cuestionarios más amplios y validados: estos formatos aunque reducen el esfuerzo del que los responde son menos precisos y sensibles.

En líneas generales, los cuestionarios específicos de medición de la CVRS son más adecuados que los genéricos para medir cambios en la evolución de los pacientes, en especial los ocasionados por los tratamientos, dado que generalmente los instrumentos específicos incluyen preguntas sobre los signos y síntomas que el paciente, con una determinada enfermedad, sufre en distintos grados¹⁵.

Cuestionarios Genéricos:

- Permiten comparaciones entre poblaciones
- Muestran diferentes aspectos del estado de salud
- Pueden no detectar diferencias
- Pueden no focalizar en áreas de interés

Cuestionarios Específicos:

- Clínicamente sensibles
- Pueden ser más discriminativos
- No permiten comparación entre poblaciones
- Valor limitado en intervenciones poblacionales

La mayoría de los expertos en evaluación de CVRS recomiendan la utilización de instrumentos específicos para cada enfermedad junto con la administración de algún instrumento genérico.

Los investigadores o diseñadores del instrumento (cuestionario) deberían establecer:

- Que el constructo o concepto y las bases racionales que establecen y describen lo que se intenta medir están claramente especificados.
- Que las dimensiones que componen el instrumento (cuestionario) son importantes para la población para la que está pensado el instrumento y la enfermedad diana. El clínico también debe juzgar si la omisión de algunas dimensiones relativas a la CVRS son importantes o no para los pacientes de su entorno o consulta.
- Que la escala de medida en que se expresan las puntuaciones y las puntuaciones mismas están debidamente justificadas y respaldadas por la evidencia de la que se dispone.
- Que los resultados obtenidos son importantes para la población diana y que ésta ha participado directamente en la elaboración del instrumento (grupos focales, etc.). Si los autores no presentan evidencias directas de que sus resultados son importantes para los pacientes, como mínimo deberían citar trabajos previos donde ello se mencione.
- Alternativamente, los investigadores pueden describir el contenido de sus mediciones con detalle, de tal forma que nos permita en base a nuestra experiencia decidir si lo que está siendo medido es importante o no para nuestros pacientes.
- Verificar que las dimensiones del cuestionario reflejan la influencia de la severidad de la enfermedad, los beneficios esperados, y/o los efectos colaterales de los tratamientos (si son estos los que se están valorando).
- Revisar que las dimensiones exploradas por el instrumento (cuestionario) comprenden aquellas en las cuales se espera que se produzcan cambios.
- Comprobar si se ha usado un instrumento específico en aquellas situaciones en que los cambios globales esperados de CVRS sean pequeños.

Tabla 1 Recomendaciones o lista-guía para la evaluación del Modelo Conceptual y las Dimensiones de la CVRS

2.2. ¿SE ESPECIFICAN LAS PROPIEDADES DE MEDICIÓN (PSICOMÉTRICAS) DEL INSTRUMENTO UTILIZADO PARA MEDIR LA CVRS?

Los instrumentos de medida deben reunir una serie de propiedades técnicas que nos garanticen que estamos midiendo aquello que queremos medir (validez) y con un error de medición lo más pequeño posible (fiabilidad). Pero hay otras características que también deben tomarse en consideración como son su eficiencia, sensibilidad, objetividad, comprensibilidad, simplicidad, etc. Para una correcta valoración de las mismas, el diseñador del cuestionario de CVRS debería describir de forma clara los métodos empleados para recoger los datos, incluyendo:

- a) Tipo de muestreo y tamaño de la muestra; es obvio que es preferible un número grande de sujetos. Según Nunally¹⁶ debe haber cinco sujetos por ítems en la versión inicial del cuestionario. Y si se va a hacer un análisis factorial, la muestra no debe bajar de 150 o 200 sujetos, siendo la recomendación más liberal la de Kline¹⁷ que pone en 100 sujetos el número mínimo aceptable, siempre que se mantenga la norma de que el número de sujetos duplique el número de variables (preguntas del test);
- b) Características de la muestra (sociodemográficas, clínicas, etc.); ésta debe ser semejante a la población en la que se piensa utilizar el instrumento;
- c) Condiciones de aplicación del cuestionario (cómo y dónde fue administrado);
- d) Estadísticos descriptivos del instrumento.

El comité científico asesor del Medical Outcomes Trust¹⁸ publicó en el año 2002 una actualización del conjunto de propiedades o características de los instrumentos de medida de la CVRS y de los criterios o condiciones consideradas como un estándar para juzgar estas propiedades psicométricas, y, así, poder llevar a cabo una valoración de dichos instrumentos. Según este comité científico los atributos o propiedades principales de los instrumentos de medida de calidad de vida y estado de salud son:

- Modelo conceptual y de medida
- Fiabilidad
- Validez
- Sensibilidad a los cambios
- Interpretabilidad
- Carga de respuesta y de administración
- Formas alternativas
- Adaptaciones culturales y de lenguaje

A continuación describiremos las principales propiedades o atributos de los instrumentos de medición de la CVRS y los criterios para evaluar si cumplen dichas propiedades.

2.2.1. El modelo conceptual y de medida

El modelo conceptual, como anteriormente hemos mencionado, es la razón fundamental para la que se ha creado el cuestionario, con la descripción del concepto abstracto o constructo que se quiere medir y la población que se quiere evaluar con dicho instrumento, especificando la relación entre estos conceptos. El modelo de medida queda reflejado en las escalas o dimensiones del cuestionario y en los procedimientos seguidos para crear las distintas puntuaciones que obtendrán los sujetos o grupos.

El diseñador del cuestionario de CVRS debe indicar:

- El constructo o concepto general que intenta medir y sus dimensiones.
- Las bases empíricas y conceptuales para generar el contenido del cuestionario y describir las razones y procedimientos seguidos para la formación de las escalas o subescalas de puntuación, y para seleccionar las preguntas (ítems) finalmente incluidas en el cuestionario.
- En los instrumentos multidimensionales debe justificar las dimensiones incluidas.
- El nivel de medida de la escala, y la adecuación de la distribución de las puntuaciones obtenidas, mediante el cuestionario, a la población o al uso para el que ha sido diseñado. Debe incluir los efectos suelo y techo (porcentaje de individuos en los valores máximos y mínimos de la puntuación).
- Los procedimientos para obtener una puntuación global de la escala o en su caso para obtener las puntuaciones finales de cada una de las dimensiones, a partir de las puntuaciones de cada pregunta.

2.2.2. Fiabilidad

La fiabilidad de una escala se refiere a la capacidad que tiene para medir de forma consistente y precisa la característica que pretende medir. Una balanza es fiable si cada vez que se pesa el mismo objeto se obtiene el mismo resultado. Un cuestionario es fiable si cada vez que se mide a los mismos sujetos se obtiene el mismo resultado. Pero al medir sujetos surge un problema que no es tan evidente cuando se miden objetos: los sujetos son cambiantes; de modo que no es fácil saber si la variabilidad en las mediciones obtenidas se debe a la imprecisión del cuestionario, o a los cambios operados en los sujetos. Las medidas de fiabilidad intentan cuantificar qué cantidad de la variabilidad de las mediciones obtenidas con un cuestionario (puntuaciones observadas) se debe a las puntuaciones verdaderas y qué cantidad se debe a los errores de medida. El análisis de fiabilidad se ocupa de la precisión del instrumento, es decir, de los errores incontrolables, inevitables e impredecibles asociados a todo proceso de medida¹⁹. Por tanto, puede decirse que un instrumento (cuestionario) es fiable si sus mediciones reflejan exactamente los valores verdaderos del constructo o característica que se investiga.

La fiabilidad se evalúa repitiendo el proceso de medición con el fin de analizar la concordancia de las distintas medidas. La fiabilidad es un concepto que admite distintos sentidos o enfoques, que guardan estrecha relación con el procedimiento o método que se haya seguido para la obtención de las puntuaciones o medidas repetidas, y puede entenderse de diversas maneras:

- a) Fiabilidad como constancia o estabilidad en el tiempo, que se refiere a la capacidad del instrumento para arrojar las mismas mediciones cuando se aplica en momentos distintos a los mismos sujetos (método test-retest)
- b) Fiabilidad como equivalencia de los resultados obtenidos por los mismos sujetos en dos instrumentos (cuestionarios) considerados paralelos o intercambiables (método de las formas paralelas o equivalente de la escala)
- c) Fiabilidad como consistencia interna, que recoge el grado de coincidencia o parecido (homogeneidad) existente entre los elementos (preguntas) que componen cada escala (método de la intercorrelación de elementos). En la actualidad, el enfoque de consistencia interna es tal vez al que más recurren los investigadores para estimar la fiabilidad de un instrumento, no sólo porque resulta económico (sólo se requiere una aplicación del cuestionario), sino también porque constituye la mejor forma de valorar una de las más importantes fuentes de error de medición en los cuestionarios: el muestreo o selección de los ítems.

Por ello, el método que se elija para evaluar la fiabilidad de un instrumento depende en parte de la naturaleza del instrumento, y en parte también del aspecto de la fiabilidad que revista mayor interés.

En la fiabilidad entendida como estabilidad o como equivalencia de los resultados, la comparación de los valores obtenidos en las mediciones realizadas con el cuestionario se realiza de manera objetiva mediante el cálculo del coeficiente de fiabilidad (o coeficiente de estabilidad) que corresponde al coeficiente de correlación entre los dos grupos de valores o puntuaciones obtenidas. Los valores posibles para este índice fluctúan de -1 a +1. Si la correlación fuese perfecta, el coeficiente sería 1. Los coeficientes de fiabilidad superiores a 0.70 pueden considerarse satisfactorios en la mayor parte de los casos. Cuanto más alto sea el valor del coeficiente de fiabilidad, tanto más estable será la medición.

En la fiabilidad entendida como consistencia interna podemos presuponer que si varios ítems están midiendo el mismo atributo o dimensión, estos ítems deberían estar intercorrelacionados. Esta correlación se puede medir mediante el coeficiente alfa (o alfa de Cronbach) y la fórmula 20 de Kuder-Richardson (KR-20) que generan un coeficiente de fiabilidad que puede interpretarse de la misma manera que el anteriormente descrito, fluctuando el intervalo normal de valores entre 0 y +1; los valores más elevados reflejan un mayor grado de consistencia interna.

- Debe figurar la fiabilidad estimada y el error estándar para todos los elementos del cuestionario incluyendo la puntuación total y de las escalas o dimensiones cuando proceda.
- Para las medidas con datos de intervalos, incluir información sobre fiabilidad test-retest y coeficientes de correlación intraclase cuando proceda.
- Consistencia Interna: cuando se comparan grupos se recomienda un coeficiente alfa de Cronbach y/o Kuder-Richardson fórmula 20 (KR-20) mínimo de 0.70 cuando se usa a nivel de grupos y de 0.90 cuando se usa a nivel individual.
- Reproducibilidad: cuando se comparan grupos se considera aceptable un coeficiente de fiabilidad test-retest o un coeficiente de correlación intraclase (CCI) mínimo de 0.70 cuando se usa a nivel de grupos y de 0.90 cuando se usa a nivel individual.
- Concordancia interobservadores: es deseable una correlación de 0.80 (o mayor)
- Se debería aportar un gráfico mostrando el error estándar de medida (raíz cuadrada del error de la varianza) sobre el rango de la escala, y adicionalmente la fiabilidad marginal de cada escala. Para escalas nominales y ordinales debería figurar la Kappa simple y ponderada.
- Se debería aportar también información sobre la comparabilidad de los parámetros estimados de los ítems y para la precisión de las medidas en repetidas administraciones.
- Se deben aportar datos relativos a las poblaciones de interés, si los estadísticos pueden diferir sustancialmente entre las diferentes poblaciones en las que el instrumento puede ser usado.

Tabla 2. Recomendaciones o lista-guía para la evaluación de la fiabilidad

2.2.3. Validez

Por validez se entiende la capacidad de un instrumento para medir aquello para lo que ha sido diseñado y no otros aspectos distintos de los pretendidos; denota el grado en que el instrumento mide lo que se supone que debe medir. Los problemas de validez se relacionan con la duda de si realmente se estará midiendo el concepto abstracto o constructo que se quiere medir.

Al igual que la fiabilidad, la validez no constituye una característica de todo o nada; no puede afirmarse a ciencia cierta que una herramienta tiene o no validez, pues se trata de una cuestión de grado. En sentido estricto no es el instrumento (cuestionario o escala) lo que se valida, sino las inferencias que se hacen a partir de él sobre determinados aspectos de las personas a las que se les aplica. Aspectos importantes de la validez de un cuestionario o escala son: su capacidad para discriminar entre sujetos de los que se sabe que difieren en la característica medida (validez discriminante); la concordancia existente entre las mediciones obtenidas con el cuestionario y las obtenidas con otras estrategias (por ejemplo, otros cuestionarios) que miden la misma característica (validez concurrente); la articulación de las puntuaciones en dimensiones similares a las propuestas teóricamente (validez de constructo); la concordancia entre las puntuaciones obtenidas con el cuestionario y las valoraciones hechas por expertos (validez de contenido). En el caso de los instrumentos de medida utilizados en el ámbito clínico debe prestarse especial atención a la sensibilidad de la escala para detectar los cambios producidos en la evolución clínica de los sujetos.

Los problemas con los que se enfrenta la validación de un cuestionario son principalmente tres: asegurar que el concepto abstracto o constructo a medir tiene entidad y consistencia; establecer que los ítems o preguntas del cuestionario representan de modo adecuado el constructo a medir, y proporcionar datos sobre la capacidad predictiva del instrumento.

En la bibliografía sobre metodología de la medición con escalas y cuestionarios abundan los términos relacionados con diferentes facetas de la validez para la valoración de un instrumento o cuestionario: validez de contenido, de criterio y de constructo. Estos tres tipos de validez clásicos son sólo aspectos o estrategias de recogida de datos en el proceso de validación.

i) Validez de contenido

La validez de contenido determina si el instrumento contiene ítems representativos de todas las dimensiones que forman la definición del concepto o constructo en estudio. Es decir, describe la idoneidad del muestreo o selección de los ítems para el constructo o concepto que se mide, y hace referencia a cuán representativas son las preguntas del cuestionario respecto del universo de preguntas que podrían hacerse sobre el tema. La validez de contenido de un instrumento se basa necesariamente en el criterio del investigador o en la evaluación por paneles de expertos, pues no existen métodos totalmente objetivos para garantizar la adecuada selección de las preguntas o ítems que deben incluirse en un cuestionario.

A la hora de llevar a cabo la validación de los contenidos han de comprobarse, al menos, dos aspectos: a) La definición del constructo de interés. Esta ha de ser lo más exhaustiva y operativa posible (para que la medición capte adecuadamente todas las dimensiones del mismo), de modo que sea susceptible de someterse a prueba y sea posible derivar indicadores empíricos para su medición. b) La correcta representación del concepto abstracto o constructo en el cuestionario. Se refiere al grado en que los ítems que componen el cuestionario representan todos los aspectos del constructo a medir, es decir, que los ítems sean pertinentes y apropiados en términos del constructo y que midan adecuadamente todas las dimensiones de éste. En lo que se refiere a la evaluación de la pertinencia de los ítems, algunos investigadores calculan índices de concordancia entre observadores y un índice de validez de contenido (IVC) formal para las evaluaciones de los expertos sobre la pertinencia de cada pregunta del cuestionario. Un procedimiento ampliamente utilizado consiste en pedir a los expertos que califiquen los ítems en una escala de cuatro puntos (de 1 = no pertinente a 4 = muy pertinente). El IVC para el instrumento total es el porcentaje del total de ítems a los que los expertos dieron la puntuación de 3 ó 4. Un IVC de 0.80 o más, suele considerarse como indicativo de una validez de contenido elevada.

La validez aparente o facial es una forma especial de la validez de contenido que se refiere al hecho de que el instrumento parezca, de la impresión a las personas evaluadas, de que es adecuado para medir lo que se pretende medir (el constructo). Aunque este parámetro no constituye una prueba aceptable de la validez de un instrumento, puede ser útil que una medida sea aparentemente válida pues en determinadas circunstancias puede ayudar a las personas a motivarse para contestar.

ii) Validez de criterio

La validez de criterio se refiere al grado en que la medición realizada con el instrumento (cuestionario) se correlaciona con otra medida directa del concepto o característica en estudio. Se podría definir también como la concordancia entre los resultados de la medida del instrumento (cuestionario) y los resultados de la medida realizada mediante un criterio o patrón de referencia (gold standard) de lo que se quiere medir. En valoraciones realizadas en el área sanitaria la validez de criterio raramente se realiza debido a la ausencia de medidas con «patrón oro» ampliamente aceptadas.

Una vez que se selecciona un criterio de comparación, es importante considerar los argumentos aportados para elegir esta medida de referencia. La validez de criterio puede evaluarse de manera sencilla y directa correlacionando las puntuaciones del instrumento predictor (cuestionario) con los del criterio de referencia. La magnitud del coeficiente de correlación (coeficiente de validez) constituye una estimación directa de la validez de criterio del instrumento (cuestionario).

En ocasiones se establece una distinción entre dos tipos de validez de criterio que, aunque no suele ser importante, se emplea con cierta frecuencia. La validez predictiva se refiere a la idoneidad de un instrumento para diferenciar entre el desempeño o las conductas de los individuos respecto de un criterio futuro. La validez concurrente denota la capacidad de un instrumento de distinguir a los individuos que difieren en su condición actual respecto de un criterio. La diferencia entre validez predictiva y concurrente radica, por consiguiente, en que el criterio de referencia se pueda medir concurrentemente con el cuestionario, o que haya que medirlo con posterioridad. Hay situaciones en las cuales la validez de criterio no tiene un interés fundamental o no se puede evaluar, por ejemplo, cuando el cuestionario es la única medición del concepto o constructo que se estudia y no existe ningún criterio de referencia que se pueda utilizar.

iii) Validez de constructo

La validez de constructo determina el grado en que una prueba o instrumento (cuestionario) mide un determinado rasgo, característica o constructo (en psicometría un “constructo” es un concepto abstracto no medible directamente que se quiere convertir en variable operativa para que pueda ser medido).

Determinar la validez de constructo de un instrumento constituye una tarea difícil pues hay que dar respuesta a las siguientes preguntas: ¿Qué mide realmente este instrumento de medición? ¿Mide adecuadamente el concepto abstracto que se investiga? No hay que confundir esta validez con la validez de contenido que es la que garantiza que los diversos aspectos o elementos relevantes del concepto abstracto a medir están debidamente representados, mientras que la validez de constructo lo que garantiza es que las medidas o puntuaciones que resultan de la aplicación del instrumento diseñado (cuestionario) pueden ser consideradas como medición de dicho concepto abstracto (“constructo”) y utilizadas como variable para distintos análisis. Las puntuaciones obtenidas al aplicar el instrumento sólo son de interés en la medida en que constituyen un fundamento válido para inferir el grado o cantidad en que el sujeto posee determinadas características.

La validación de constructo puede abordarse de diversas maneras, pero siempre hay que hacer hincapié en el análisis lógico y en la prueba de las relaciones pronosticadas por el investigador que ha diseñado el cuestionario. La validez de convergencia y de discriminación se consideran dos aspectos de la validez de constructo. La validez convergente se refiere al grado en que diferentes formas de medir el mismo constructo (o dimensiones de un constructo) se correlacionan entre sí. La validez discriminante o divergente demuestra que una medida no se correlaciona con otras medidas efectuadas sobre rasgos, dimensiones o constructos diferentes a los que ella mide.

La validación de constructo abarca una serie de técnicas todas dirigidas a valorar: a) Si el constructo teórico que hemos creado parece ser coherente o adecuado; y b) Si las puntuaciones parciales o globales que se obtienen con el cuestionario se corresponden bien con la cuantificación de este constructo. Los análisis para este tipo de validación son empíricos y pueden tener dos enfoques básicos:

- a) Estudios correlacionales sobre todo para tres tipos de relaciones (o, en su caso, de no relaciones):
 - relación con otros modos de medir el mismo constructo (o característica).
 - comprobación de relaciones esperadas (positivas o negativas) con otros constructos (o características).
 - comprobar que no existe relación donde no esperamos que la haya.
- b) Comparaciones entre grupos supuestamente distintos en aquello que estamos midiendo
 - análisis estadísticos de contrastes de medias.

Los diferentes métodos empleados para medir la validez de constructo de un cuestionario se suelen basar en uno de estos dos enfoques básicos, pero principalmente en el uso de las correlaciones. Uno de estos métodos es la técnica de grupos conocidos, según la cual el instrumento se aplica a los grupos que, según se presume, diferirán respecto del constructo a causa de una característica conocida. En este caso, la validez del instrumento sería dudosa si no se registraran diferencias de puntuación entre ambos grupos. En este método, lo más importante es la magnitud de la diferencia y no tanto lo estadísticamente significativo que sea el resultado del test que se ha empleado para la comparación. Otro método para estimar la validez de constructo consiste en examinar las relaciones con base en predicciones teóricas no probadas, lo cual en realidad es una variante de la técnica de grupos conocidos.

El método matricial multirasgo-multimétodo (multitrait-multimethod matrix method, MTMM) desarrollado por Campbell y Fiske²⁰ constituye un avance significativo en el ámbito de la validación de constructos. El procedimiento aplica los conceptos de convergencia (validez convergente) y discriminabilidad (validez discriminante). Para aplicar esta técnica de validación, el investigador debe medir el constructo cuando menos con otros dos instrumentos o estrategias (por ejemplo, cuestionarios) diferentes. Un segundo requisito consiste en medir con los mismos instrumentos otros constructos de los cuales se desea diferenciar el constructo en estudio. Diversos resultados del método de matriz MTMM indican si existe validez de constructo.

En la práctica, la valoración de la validez de constructo se reduce a la comprobación de la consistencia interna de los resultados y de la estructura de relaciones de los diferentes ítems del cuestionario (estructura interna) para comprobar si estas relaciones confirman o contradicen la hipótesis previa de si miden o no el constructo en estudio, así como sus correspondientes dimensiones. Para este análisis se recurre al procedimiento estadístico que se conoce como análisis factorial. Aunque el cálculo es complejo, sus principios resultan bastante sencillos. El análisis factorial es en esencia un método para identificar grupos o conglomerados de ítems relacionados; cada grupo recibe el nombre de factor y debe representar una dimensión del constructo. El procedimiento se emplea, por tanto, para identificar y agrupar diferentes ítems que tienen alguna característica latente o subyacente común. Puede servir también como alternativa para examinar la validez convergente y la discriminante.

En resumen, si un instrumento observa validez de constructo es de esperar una fuerte correlación entre todos los ítems que exploran la misma dimensión del cuestionario. La validez de constructo implica tanto procedimientos lógicos como empíricos.

Para la Validez de Contenido: ¿Contiene el cuestionario los ítems más representativos en cada una de las dimensiones del constructo o concepto que se está midiendo?

- Validez Aparente o Facial del instrumento: ¿El instrumento parece adecuado, a las personas evaluadas, para medir el constructo o concepto de interés?

Para la Validez de Criterio: ¿En qué grado la medición del constructo, realizada con el cuestionario, se relaciona con un criterio de referencia ("gold standard")?

- Validez concurrente: ¿El cuestionario es útil para clasificar individuos cuya diferencia radica en una cualidad o condición actual?

- Validez predictiva: ¿El cuestionario es útil para diferenciarlos respecto a un criterio futuro?

Para la Validez de Constructo: ¿En qué grado el cuestionario mide el constructo, rasgo o dimensión de interés para el investigador? Es decir, ¿las puntuaciones obtenidas apoyan las hipótesis establecidas con relación al constructo evaluado?

- Análisis de validez convergente: ¿Existe una correlación positiva entre dos instrumentos diferentes que miden el mismo constructo?

- Análisis de validez discriminante: ¿Existe una correlación negativa entre dos instrumentos diferentes que miden constructos distintos?

- Método de grupos conocidos: ¿Existen diferencias en la puntuación obtenida entre dos grupos que difieren respecto al constructo evaluado por una característica conocida?

- Método matricial multirasgo-multimétodo: Tras el empleo de dos instrumentos que miden el mismo constructo, ¿establecen correlaciones elevadas? ¿son más bajas al medir conceptos diferentes?

- Análisis Factorial: ¿El cuestionario se puede agrupar en diferentes factores o conglomerados de variables relacionadas que representen las dimensiones del constructo evaluado?

Tabla 3. Recomendaciones o lista-guía para la evaluación de la validez) Adaptado de: Badía Llach X y Lizán Tudela L.²¹

2.2.4. Sensibilidad a los cambios

La sensibilidad de un instrumento se refiere a la medida en que detecta y mide de manera fiable las variaciones por pequeñas que sean de un rasgo, dimensión o constructo. La utilidad de una medida en salud depende de su habilidad para detectar diferencias clínicamente relevantes, por ello, “la sensibilidad a los cambios es definida a menudo como el cambio mínimo considerado como importante por las personas con aquella condición de salud, las personas próximas o importantes para ellos o sus proveedores”²². La sensibilidad a los cambios es una característica importante a valorar en los instrumentos de medida de la salud dado que nos indica el nivel en que sus mediciones discriminan entre los grupos que poseen diferentes cantidades de un atributo o característica (sensitivity), o entre los cambios en ese atributo, característica, o dimensión que han tenido lugar en un sujeto (responsiveness). Por tanto un instrumento de medida de la CVRS debe ser lo suficientemente sensible para detectar cambios relevantes cuando se alteran las condiciones de un paciente, y por otra parte, si dos grupos de pacientes difieren en su CVRS, el instrumento de medida de la CVRS debe ser lo suficientemente sensible para detectar esas diferencias.

La sensibilidad entre grupos se puede evaluar mediante estudios transversales comparando grupos de pacientes en los que se espera encontrar diferencias en la CVRS; sin embargo, la sensibilidad al cambio en un sujeto debe ser evaluada con estudios longitudinales mediante medidas repetidas al mismo sujeto que nos confirmen un cambio en las puntuaciones cuando el estatus del paciente cambia.

Entre las medidas más usadas para medir esta sensibilidad al cambio están la Respuesta Media Estandarizada (razón entre el cambio medio ocurrido y la desviación estándar de este cambio) y el Tamaño del Efecto (razón entre el cambio medio ocurrido y la desviación estándar de la medida inicial). También se han utilizado tests apareados para la comparación de medias. Existe controversia sobre cual es la mejor forma de medir la sensibilidad, pero donde sí hay acuerdo es en que los estudios de sensibilidad que evalúan un nuevo instrumento deberían reportar diversa comparaciones que cubran el rango de situaciones que son más típicas en el área o patología donde se aplicará y utilizará el nuevo instrumento.

2.2.5. Interpretabilidad

Se entiende por interpretabilidad el grado en el que puede asignarse un significado comprensible a las puntuaciones cuantitativas obtenidas con un instrumento de medida (cuestionario). La interpretabilidad de una medida se facilita traduciendo o equiparando una puntuación cuantitativa (o los cambios en las puntuaciones) a una categoría cualitativa u otra medida externa que tenga un significado clínico o más familiar. La interpretabilidad se debe acompañar de una explicación racional de la medida externa cualitativa, los cambios en las puntuaciones, y el modo en el que dichas puntuaciones han de ser interpretadas con relación a la medida externa.

Diversos tipos de información pueden ayudar en la interpretación de las puntuaciones¹⁶:

- Datos comparativos de las distribuciones de puntuaciones derivadas de varios grupos de poblaciones definidas, incluyendo, cuando sea posible, una muestra representativa de la población general.
- Resultados de un amplio grupo de estudios que hayan utilizado el instrumento en cuestión y que aporten algún hallazgo, de manera que permita familiarizarse con el instrumento, ayudando a su interpretación.
- La relación de las puntuaciones en condiciones clínicamente reconocidas, requeridas para tratamientos específicos o intervenciones de efectividad reconocida.
- La relación de las puntuaciones o los cambios en las puntuaciones en situaciones de la vida socialmente reconocidas (tales como el impacto de la pérdida del puesto de trabajo).
- La relación de las puntuaciones o los cambios en las puntuaciones en clasificaciones subjetivas de mínimos cambios clínicamente importantes en personas con la enfermedad, sus familiares o cuidadores.
- La precisión con la que las puntuaciones predicen eventos relevantes conocidos (tales como la muerte o la necesidad de atención institucional).

- Comprobar si se describe razonadamente la selección de los criterios externos o las poblaciones propuestas para la comparación y la interpretación de los datos.
- Para facilitar la interpretación se debe proporcionar la información relacionada con el modo en el que los datos del instrumentos han sido recogidos y presentados.
- Citar los principales “benchmarks” (datos comparativos o normativos) para facilitar la interpretación de las puntuaciones.

Tabla 4. Recomendaciones o lista-guía para la evaluación de la interpretabilidad

En la investigación clínica es importante el concepto de mínimo cambio clínicamente importante, dado que cambios muy pequeños pueden ser estadísticamente significativos sin tener ninguna relevancia clínica (especialmente en tamaños muestrales grandes). El mínimo cambio clínicamente importante se puede definir como la diferencia de puntuación más pequeña, en la dimensión de interés, que los pacientes perciben como beneficiosa, y que obligaría a un cambio en el tratamiento del sujeto en ausencia de efectos indeseables graves o de un coste excesivo¹⁹.

2.2.6. Facilidad de uso: carga de respuesta y de administración

La carga de respuesta se entiende como el tiempo, esfuerzo y otros aspectos que se exigen a quienes se les administra el instrumento de medida (cuestionario). Por otra parte, la carga de administración se define como el grado de trabajo y esfuerzo que deben realizar quienes administran el instrumento (cuestionario), y depende del tiempo precisado para su administración, formación requerida del entrevistador y recursos necesarios para su aplicación y puntuación.

Estas cargas equivalen a lo que otros autores denominan “factibilidad”, que viene a indicar las posibilidades reales y prácticas de que el cuestionario pueda ser respondido, completado y analizado, e incluye dos aspectos: a) la factibilidad para el que responde, es decir, el tiempo y otras necesidades que pueda tener la persona que responde el cuestionario y, b) la factibilidad administrativa, que se define como las necesidades de aquellos que administran el instrumento (cuestionario)²³. La facilidad de uso está relacionada con el nivel de aceptabilidad del instrumento y se refleja en la proporción de rechazos, no respuesta y datos perdidos.

Recomendaciones o lista-guía para la evaluación de la Carga de respuesta:

- Dar información sobre los siguientes aspectos:

- 1) Media y rango de tiempo requerido para completar el instrumento en caso de que sea autoadministrado o bien administrado por un entrevistador (para todos los grupos de la población para los cuales se aplicará el instrumento).
- 2) El nivel de lectura y comprensión requerido para todos los grupos de población para los cuales se aplicará el instrumento.
- 3) Cualquier requisito o cuestión específica que pudiera necesitarse de los encuestados, tales como consultar informes o solicitar copia de la medicación empleada.
- 4) La aceptabilidad del instrumento, por ejemplo indicando el nivel de datos perdidos y el porcentaje de personas que rechazaron responder, así como las razones de ambos hechos.

- Para instrumentos que en apariencia pueden parecer ofensivos o aquellos que aparecen con elevado porcentaje de datos perdidos, proporcionar pruebas de que el instrumento no se aplicó de forma indebida física o emocionalmente al entrevistado (por ejemplo, que no se incluyen preguntas que una minoría significativa de pacientes las considere demasiado comprometidas o desagradables).

- Indicar cuando o bajo qué circunstancias el instrumento no es adecuado para los encuestados.

Recomendaciones o lista-guía para la evaluación de la Carga de administración:

- Proporcionar información sobre cualquier recurso necesario para la administración del instrumento, tales como la necesidad de un hardware o software específicos para gestionar la puntuación o analizar el instrumento:

- Aportar el rango y el tiempo medio requerido por un entrevistador entrenado para administrar el instrumento en una entrevista personal, telefónica o con formato asistido por ordenador, según proceda;

- Indicar el entrenamiento y el nivel educativo o la experiencia profesional requerida por el personal administrativo para la gestión, puntuación u otros usos del instrumento.

- Indicar la disponibilidad de instrucciones para obtener la puntuación.

Tabla 5. Recomendaciones o lista-guía para la evaluación de la Carga de respuesta y de Administración

2.2.7. Formas alternativas de administración

Las formas alternativas de administración son aquellas formas de administración para las que no ha sido originalmente diseñado el instrumento, por ejemplo, si el cuestionario ha sido diseñado para ser autoadministrado, no se debe administrar por un entrevistador si sus características psicométricas (fiabilidad, validez, etc,...) no han sido evaluadas para esta forma de administración.

Las formas alternativas de administración pueden incluir autoadministrado, administrado por entrevistador, valoración por un observador entrenado, entrevista asistida por ordenador, y medidas basadas en el comportamiento. También se consideran formas alternativas de administración aquellos instrumentos originalmente diseñados para ser contestados por informadores directos cuando son administrados a informadores indirectos tales como padres, parejas, cuidadores u otros suplentes.

Los investigadores o diseñadores del instrumento (cuestionario) deberían:

- Poner a disposición la información que pruebe la fiabilidad, validez, sensibilidad e interpretabilidad y carga de respuesta y de administración para cada uno de los modos de administración.
- Proporcionar información sobre la comparabilidad de las formas alternativas, y siempre que sea posible, deberían aplicarse estudios equivalentes ya que las puntuaciones de las formas alternativas pueden compararse entre ellas o con las puntuaciones del instrumento original.

2.3. ¿HA SIDO TRADUCIDO, ADAPTADO Y VALIDADO AL IDIOMA Y CULTURA DEL PAÍS DONDE LO VAMOS A EMPLEAR (O SE HA EMPLEADO)?

La adaptación de un cuestionario de CVRS desarrollado en otro país es muchas veces preferible a la creación de uno nuevo, entre otras razones, porque suele resultar más rápido y menos costoso, pero sobre todo, por el creciente interés en realizar estudios y comparaciones interculturales. Aunque realizar análisis comparativos interculturales de calidad de vida probablemente no siempre tiene sentido dadas las diferentes concepciones culturales sobre lo que es “calidad de vida”, y que ésta suele estar asociada, entre otros aspectos, al grado de desarrollo económico y nivel de vida.

La adaptación de un instrumento de medida de la CVRS a otra cultura es una tarea compleja que va mucho más allá de la mera traducción de un cuestionario. La adaptación transcultural tiene como objetivo el conseguir un instrumento equivalente al desarrollado en el país de origen, preservando el contenido semántico en el uso habitual del lenguaje de la nueva traducción y conservando la equivalencia de significados de los cuestionarios entre las culturas.

Los principales errores o problemas que suelen aparecer cuando se adapta un cuestionario de una cultura a otra, sin seguir unos estándares científicos que garanticen la equivalencia cultural de las dos versiones, se agrupan en:

- Problemas de equivalencia semántica a través de las lenguas: algunos verbos o adjetivos tienen significados o connotaciones diferentes en las dos lenguas.
- Problemas de equivalencia conceptual a través de las culturas: los conceptos y constructos que utilizamos, independientemente de las palabras y frases usadas para expresarlos, no siempre existen en otras culturas.
- Problemas de equivalencia normativa a través de las sociedades: cada sociedad tiene sus propias normas (convenciones sociales) que ejercen gran influencia sobre algunos aspectos de la conducta, e impiden la discusión o aportar información sobre ciertos tópicos.

Para resolver los problemas, fundamentalmente semánticos, que pueden aparecer en el proceso de adaptación cultural de un cuestionario de una lengua o cultura a otra, existen diferentes métodos. El más utilizado suele ser el de traducción / retrotraducción que consiste en un proceso iterativo donde cada ciclo incluye cuatro pasos²⁴:

- 1.- Una persona bilingüe traduce el cuestionario original al idioma diana.
- 2.- Una segunda persona bilingüe sin conocer el cuestionario en el idioma original, traduce de nuevo (retrotraduce) al idioma original la versión traducida al idioma diana.
- 3.- Se comparan el cuestionario original y la versión retrotraducida.
- 4.- En el caso de que existan diferencias importantes entre ambos cuestionarios, se vuelve a elaborar una versión en el idioma diana efectuando las modificaciones pertinentes con la finalidad de eliminar las discrepancias.

Este proceso debe repetirse hasta que ambos cuestionarios sean idénticos o contengan las mínimas discrepancias. Una vez consensuada una primera versión del cuestionario, éste se administra en forma de entrevista semiestructurada a un grupo reducido de pacientes (prueba piloto). Mediante este estudio se determinarán las posibles dificultades de comprensión de los ítems, la calidad de la adaptación cultural, su comprensión, relevancia y aceptabilidad.

No hay que olvidar que la fiabilidad y validez de un cuestionario o escala no son traducibles, por lo tanto una vez finalizada la adaptación transcultural, el resto del proceso de validación del cuestionario traducido es el mismo que cuando se construye un nuevo instrumento de medida de la CVRS.

- ¿Existe una equivalencia semántica y conceptual del constructo y las dimensiones que mide el cuestionario, en las lenguas y grupos culturales?
- ¿Se han empleado traductores profesionales cualificados?.
- ¿Se han combinado los diseños de traducción y retrotraducción?
- ¿Se ha efectuado un estudio piloto de la versión consensuada y adaptada?
- ¿Se ha aplicado el cuestionario a una muestra amplia y representativa de la población objetivo?.
- ¿Se ha comprobado al fiabilidad y validez del nuevo instrumento de medida (cuestionario)?
- ¿Se ha establecido algún tipo de equiparación entre las puntuaciones de la versión original y la traducida que permita las comparaciones?.

Tabla 6. Recomendaciones o lista-guía para evaluar una adaptación transcultural de un cuestionario de CVRS.

2.4. ¿QUÉ CALIDAD DE EVIDENCIA Y APLICABILIDAD CLÍNICA PROPORCIONA EL ESTUDIO?

Para abordar el análisis de un estudio o artículo científico se requiere disponer de una estrategia adecuada. En primer lugar, nos fijaremos en el título y en el resumen (abstract), y si es de nuestro interés pasaremos a leer el apartado metodología, donde analizaremos de forma genérica los aspectos claves a considerar en la lectura crítica de todo artículo científico. Si estos aspectos no están en consonancia con los criterios de calidad científica mencionados anteriormente desecharemos el artículo, pues las inferencias y extrapolaciones serían incorrectas. Una vez superada esta fase, evaluaremos los resultados, valorando si están o no bien interpretados, obteniendo nuestras propias conclusiones en base a las recomendaciones dadas en los apartados anteriores. Por último, hemos de valorar si los resultados pueden ser útiles para nuestra práctica clínica. Para todo ello, hemos de responder a tres cuestiones:

¿Son válidos los resultados?

¿Son importantes estos resultados?

¿Ayudarán los resultados en la asistencia de mis pacientes?

2.4.1. ¿Son validas estas evidencias?

La calidad de la evidencia científica que se desprenda de los estudios consultados dependerá del diseño del estudio. Una de las preguntas a las que deberemos responder es si el tipo de diseño en el que se estudia la CVRS es el adecuado y está bien realizado. Los diseños pueden priorizarse con arreglo a la calidad de la evidencia que proporcionan, pero sin olvidar que el tipo de diseño no es el único apartado a tener en cuenta, pues cada tipo de diseño puede llevarse a cabo con diferentes niveles de calidad. En general, un diseño prospectivo es superior a otro retrospectivo, y el paradigma de la investigación lo constituye el estudio de intervención aleatorizado. La adecuada evaluación crítica de los distintos diseños de investigación no es el objetivo de este capítulo, pues ya han sido explicados en capítulos anteriores, a los que remitimos al lector.

En función del objetivo principal de una investigación unos diseños son más idóneos que otros. Así, para las decisiones terapéuticas, el mejor conocimiento proviene de los ensayos clínicos aleatorizados, y para valorar una prueba diagnóstica o para validar un cuestionario, el estudio más adecuado es el transversal (estudios de corte), que es un diseño de observación en el que se selecciona una muestra de la población para conocer la frecuencia de una exposición y/o un efecto en un momento dado.

Para evaluar críticamente la evidencia acerca de la calidad y validez de los hallazgos que proporciona un estudio de CVRS, al margen del tipo de diseño del estudio, deberemos responder a las siguientes preguntas:

i) Preguntas principales:

- ¿Midieron los investigadores aspectos de la CVRS que los pacientes consideran importantes? ¿El tipo de muestreo, el tamaño de la muestra fueron adecuados? ¿Se administró correctamente?
- El instrumento (cuestionario) utilizado para medir la CVRS, ¿Tiene unas características psicométricas aceptables?: ¿Es fiable? ¿Es válido? ¿Es sensible a los cambios? ¿Es interpretable? ¿Es fácil de usar?
- ¿Fue aceptable la tasa de respuesta?

ii) Preguntas secundarias:

- ¿Se omitieron aspectos importantes de la CVRS?
- Si se incluyó una evaluación económica o de utilidades, ¿se utilizó una metodología adecuada?
- ¿Hay razones para pensar que algunas de dichas características puedan variar en determinados grupos o poblaciones? ¿Existe una versión apropiada para el grupo de edad de interés?

2.4.2. ¿Son importantes estos resultados?

Una vez que hemos considerado que el estudio es válido y sus resultados representan una estimación fiable del efecto de la intervención, pasaremos a considerar la magnitud del efecto observado sobre la CVRS (¿Qué beneficio se obtiene?) y la precisión de dicho efecto (¿Cuál es el grado de certeza con el que puedo afirmar que el efecto real es el observado?) Para responder a estas preguntas nos ayudará:

- a.- Conocer el rango de valores posibles del instrumento de medida (cuestionario).
- b.- Conocer la dirección en que deben interpretarse los valores o puntuaciones obtenidas.
- c.- Conocer los valores correspondientes a la población general o a poblaciones conocidas, que nos puedan servir de referencia.
- d.- Conocer cuál es el mínimo cambio clínicamente importante. Esta información, que no siempre es proporcionada por los autores, se obtiene a partir de relacionar los cambios en la puntuación de CVRS con otros parámetros o características generalmente clínicas.

Sea cual fuere la estrategia utilizada, los investigadores deberían poner a disposición del investigador, elementos que le permitan interpretar la magnitud del cambio observado. No hay que olvidar que la relevancia de los resultados se basa en la magnitud del efecto acompañada del correspondiente intervalo de confianza, y no en la significación estadística, ya que ésta se ve influida por la magnitud y, también, por el tamaño de la muestra.

2.4.3. Aplicabilidad de los resultados en la asistencia a mis pacientes

Por último, quedaría pendiente realizar la valoración sobre el grado de aplicabilidad de los resultados del estudio de CVRS seleccionado a un paciente de nuestra consulta. Para ello deberemos contestar las siguientes preguntas:

1. ¿Pueden aplicarse los resultados a mis pacientes? Considerar si los pacientes del estudio son tan diferentes a los nuestros como para dudar de su utilidad.
¿El diseño del estudio es compatible con las situaciones reales de la práctica clínica en la que nos desenvolvemos? Considerar si el contexto en el que desarrollamos nuestra práctica es muy diferente del contexto en el que se ha realizado el estudio.
2. ¿Se consideran todos los resultados clínicamente importantes?
3. ¿Los beneficios compensan los inconvenientes y costes?

Para responder a estas preguntas deberemos considerar la relevancia clínica de los resultados y las posibles diferencias entre las características de los pacientes incluidos en el estudio científico que evaluamos y los pacientes que acuden habitualmente a nuestra consulta. Otra cuestión importante es la posible interpretación y aplicación de las diferencias observadas entre varios subgrupos de los pacientes participantes en un estudio. En ocasiones es difícil interpretar los datos de CVRS porque:

- El concepto de CVRS no está bien definido.
- Las estrategias para interpretar las puntuaciones no están bien descritas, ni se basan en “puntuaciones de referencia” procedentes de la población general o de otras poblaciones que nos sirvan de comparación.
- No tienen significación biológica. Básicamente son números abstractos que, en ocasiones, no se pueden interpretar o comparar en función de otras características, generalmente clínicas.
- No se especifica el rango teórico y el rango observado de las puntuaciones, y cuando éstas son estandarizadas, no se indica cuáles son las poblaciones de referencia para su cálculo.
- Las unidades de CVRS no están definidas, y sólo pueden interpretarse en función de otras medidas más familiares.

La decisión final de optar o no por un determinado tratamiento dependerá de los valores y preferencias del paciente. En ocasiones, tomar una decisión final resultará fácil como usualmente ocurre ante la propuesta de una intervención sencilla, de bajo riesgo y eficaz. En otras, la decisión será más compleja como puede ocurrir ante la propuesta de una intervención con importantes efectos adversos, eventuales secuelas y beneficios inciertos.

- ¿Hay diferencias cualitativas en la eficacia de un tratamiento entre algunos subgrupos de pacientes?
- ¿Realmente tienen sentido las diferencias desde el punto de vista biológico y clínico?
- ¿Se trata de alguna diferencia importante, tanto desde el punto de vista clínico (beneficioso para algunos, pero inútil o perjudicial para otros) como estadístico?
- ¿Se trata de una diferencia expuesta con anterioridad al inicio del estudio y confirmada por otros estudios independientes?
- ¿Son aplicables estos resultados a un paciente de nuestra consulta?
- ¿La diferencia entre nuestro paciente y los pacientes del estudio es demasiado significativa para que sus resultados puedan serle de utilidad?
- ¿Cuál sería realmente el grado de beneficio potencial del tratamiento aplicado a nuestro paciente?
- ¿Satisfacen el tratamiento y sus consecuencias los valores y preferencias de nuestro paciente?
- ¿Se ajustan el tratamiento y sus consecuencias a esta valoración?

Tabla 7. Recomendaciones o lista-guía para la evaluación de la Aplicabilidad de estudio de CVRS

17

→ 3

PASOS A SEGUIR A LA HORA DE SELECCIONAR UN INSTRUMENTO PARA MEDIR LA CVRS EN NUESTRA PRÁCTICA CLÍNICA

1º. El objetivo o la finalidad que nos guía a estudiar la CVRS

En primer lugar debemos formularnos una cuestión clínica clara a partir de los problemas de un paciente y plantearnos cuales son nuestros objetivos, para qué o por qué necesitamos conocer o estudiar la CVRS, sobre todo diferenciar si pretendemos estudiar la evolución de los pacientes de forma individualizada o en grupo, o si queremos evaluar las diferencias entre dos terapias alternativas, o simplemente queremos describir nuestra población de pacientes. El objetivo del estudio nos ayudará a elegir mejor nuestro instrumento de evaluación dentro de los existentes. En general se puede hacer la siguiente recomendación a este punto: Si queremos estudiar las diferencias entre dos terapias alternativas, o si queremos estudiar la evolución de los pacientes de forma individualizada o en pequeños grupos necesitaremos aplicar un instrumento genérico que disponga de normas poblacionales que nos permitirán ajustar las diferencias atribuibles a la edad y al sexo y otro específico de la patología a estudio con gran sensibilidad a los cambios y sencillo de aplicar²².

2º. Consultar la literatura científica para hallar los artículos clínicos más relevantes

Para ello se dispone de muchas y variadas herramientas como ya se ha explicado en capítulos anteriores y a ellos nos remitimos. Una vez planteada la pregunta clínica bien construida y localizada la evidencia, el siguiente paso es la lectura crítica de la misma.

3º. La pertinencia del instrumento de medida.

Es una consideración prioritaria. Se trata de conocer el grado en que el contenido de una medida concreta se corresponde con los objetivos del investigador que la quiere usar. Esto es especialmente importante debido a los distintos conceptos, definiciones y dimensiones que se pueden incluir en una medida del estado de salud. Por tanto, deberemos seleccionar sólo aquellos instrumentos que contengan las dimensiones esenciales que más se ajusten a nuestros objetivos, acordes a la población y enfermedad diana, y por tanto que mejor vayan a medir aquello que queremos medir.

4º. Elección del instrumento de evaluación de la CVRS

Realizaremos una selección sólo entre aquellos instrumentos que cumplan unos mínimos de calidad con respecto a los atributos o propiedades psicométricas que deben cumplir los instrumentos de evaluación de la CVRS.

5º. Eficiencia y viabilidad del instrumento de medida de la CVRS

A la hora de decidir qué instrumento vamos a elegir para medir la CVRS de nuestros pacientes no deberíamos fijarnos sólo en las propiedades psicométricas del instrumento y que éste haya sido correctamente diseñado, pues puede ocurrir que un instrumento rigurosamente desarrollado nos resulte inservible si su aplicación en la consulta médica de Atención Primaria resulta difícil, compleja y costosa, suponiendo una carga para el paciente y para los profesionales sanitarios.

Por tanto, el siguiente paso sería comprobar si el instrumento reúne determinados aspectos relacionados con la viabilidad de su aplicación en la consulta médica. Para ello deberíamos considerar los siguientes puntos, que se basan en lo visto anteriormente:

Instrumentos con fiabilidad y validez comparables pueden diferir en cuanto a la eficiencia. El número de ítems que forma parte de un instrumento constituye uno de los aspectos de la eficiencia, pues los instrumentos largos tienden a ser más fiables que los cortos. Sin embargo, existe un punto de utilidad decreciente. Es probable que la mayoría de los médicos estén dispuestos a sacrificar un margen modesto de fiabilidad a cambio de la oportunidad de reducir sustancialmente la carga de respuestas de los sujetos y los costes (el esfuerzo) de la recolección de datos. Sin embargo, a factores iguales, es preferible la selección del instrumento más eficiente posible.

No deben pasarse por alto otras cualidades que han de considerarse al valorar un instrumento cuantitativo. La mayor parte de los siguientes criterios son en realidad aspectos vinculados con la fiabilidad y la validez:

- **Comprensión**
Tanto el sujeto como el investigador deben ser capaces de comprender los procedimientos necesarios para garantizar mediciones precisas y válidas.
- **Rapidez**
Para la mayor parte de los instrumentos, el investigador debe asegurarse de que contará con el tiempo necesario para obtener mediciones completas sin apresurar el proceso.
- **Sencillez**
En igualdad de condiciones, siempre es preferible un instrumento sencillo a uno complejo, pues con las mediciones complicadas se corre un riesgo mayor de error.

17

→ 4 **BIBLIOGRAFIA**

1. World Health Organization. Constitution. En: WHO. Basic Document Geneve WHO1948.
2. World Health Organization. Quality of life Assessment. An annotated bibliography. Geneva: WHO (MNH/PSF/94.1), 1994.
3. Patrick D, Erickson P. Health Policy, Quality of Life: Health Care Evaluation and Resource Allocation. Oxford University Press; New York, 1993.
4. WHOQOL GROUP. The World Health Organization Quality of life assessment (WHOQOL). Position Paper from the World Health Organization. Soc. Sci. Med. 1995; 41(10): 1403-9.
5. Piñol Jane A, Sanz Carrillo C. Importancia de la calidad de vida en atención primaria. Gastroenterol Hepatol 2004; 27(Supl 3): 49-55.
6. Lizán Tudela L. Viabilidad de las viñetas COOP/WONCA y otros instrumentos para la valoración de la calidad de vida relacionada con la salud en atención primaria. Aten Primaria 1998; 22: 135-9.
7. Testa AM, Anderson RB, Nackley JF, et al. Quality of life and antihypertensive therapy in men: a comparison of captopril with enalapril. N Engl J Med 1993; 328: 907-13.
8. Soto Alvarez J. Incorporación de estudios de calidad de vida relacionada con la salud en los ensayos clínicos: bases y recomendaciones prácticas. Uso de una lista-guía para su correcto diseño y/o evaluación. An Med Interna 2003; 20(12): 633-44.
9. Siu SI, Reuben DB, Ouslander JB, Osterweil D. Using multidimensional health measures in older persons to identify the risk of hospitalization and skilled nursing placement. Qual Life Res 1993; 2: 253-61.
10. Connelly JE, Philbrick JT, Smith GR, Kaiser DL, Wymer A. Health perceptions of primary care patients and the influence on health care utilization. Med Care 1989; 27:S99-S109.
11. Ries AL, Kaplan RM, Limbreg TM, Prewitt LM. Effects of pulmonary rehabilitation of physiologic and psychosocial outcomes in patients with chronic obstructive pulmonary disease. An Intern Med 1995; 122: 823-32.
12. Alonso J, Ferrer M. Enfermedad pulmonar obstructiva crónica y salud percibida. Med Clin (Barc) 1998; 111:580-2.

- 
13. Guyatt GH, Veldhuyzen Van Zanten SJ, Feeny DH, Patrick DL. Measuring quality of life in clinical trials: a taxonomy and review. *CMAJ* 1989;140:1441-8.
 14. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;27:S217-32.
 15. Laupacis A, Muirhead N, Keown P, Wong C. A diseasespecific questionnaire for Assessing Quality of Life in Patients on Hemodialysis. *Nephron* 1992;60:302-6.
 16. Nunnally JC. *Psychometric Theory*. 2nd ed., McGraw-Hill, New York; 1978.
 17. Kline P. *An easy guide to Factor Analysis*. Sage, Newbury Park; 1994: 74-9.
 18. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality of life instruments: Attributes and review criteria. *Quality of life Research* 2002; 11:193-205.
 19. Pardo Merino A, Ruiz Díaz MA. *Análisis de datos con SPSS 13 Base*. McGraw-Hill, Madrid; 2005.
 20. Campbell D T, Fiske D W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 1959;56:81-105.
 21. Badía Llach X y Lizán Tudela L. Estudios de calidad de vida. En: Martín Zurro. *Compendio de Atención Primaria.. Doyma*, Madrid; 2001: 255.
 22. Rebollo P, Fernández-Vega F, Ortega F. La evaluación de la calidad de vida relacionada con la salud en enfermos con hipertensión arterial. *Hipertensión* 2001; 18(9): 429-39.
 23. Lohr KN, Aaronson NK, Alonso J, Burnam MA, Patrick DL, Perrin EB, et al. Evaluating quality-of-life and health status instruments: development of scientific review criteria. *Clin Therap* 1996;18:979-92.
 24. Behling O, Kenneth S L. *Translating questionnaires and other research instruments. Problems and solutions*. Sage, Iowa 2000: 19-20.

17

→ **5 PUNTOS CLAVE**

La Calidad de Vida Relacionada con la Salud (CVRS) es una medida de resultado centrada en el paciente y representa el impacto que una enfermedad y el tratamiento consiguiente tienen sobre la percepción del paciente de su satisfacción y bienestar físico, psíquico, social y espiritual.

La incorporación de medidas de la CVRS en las consultas de atención primaria resulta de gran utilidad, pues sirve de apoyo en la toma de decisiones clínicas y contribuye a mejorar la relación médico-paciente al posibilitar al enfermo la participación en la elección de la opción más conveniente.

Los instrumentos de medida de la CVRS incluyen ítems para las dimensiones funcionamiento físico, social y emocional. Pueden ser a) Genéricos: se pueden aplicar tanto a la población general como a grupos específicos de pacientes, dado que suelen incluir un amplio espectro de dimensiones de la CVRS, o b) Específicos: se centran en aspectos específicos del estado de salud, e incluyen dimensiones de CVRS de una determinada patología.

Los instrumentos de medida deben reunir una serie de propiedades técnicas o psicométricas que nos garanticen que estamos midiendo aquello que queremos medir (validez) y con un error de medición lo más pequeño posible (fiabilidad). Pero hay otras características que también deben tomarse en consideración como son sensibilidad al cambio, interpretabilidad, carga de respuesta y de administración y las diferentes formas alternativas.

Si queremos estudiar las diferencias entre dos terapias alternativas, o si queremos estudiar la evolución de los pacientes de forma individualizada o en pequeños grupos necesitaremos aplicar un instrumento genérico que disponga de normas poblacionales que nos permitirán ajustar las diferencias atribuibles a la edad y al sexo y otro específico de la patología a estudio con gran sensibilidad a los cambios y sencillo de aplicar.

RESOLUCIÓN DE LA SITUACIÓN CLÍNICA

Varón de 46 años. En una revisión rutinaria se le ha detectado una Hipertensión arterial. Tras la detección, las cifras medias fueron de 156/104 mm Hg. El estudio inicial (exploración, analítica de sangre y orina, ECG, Rx Tórax) fue normal. No fuma, ni toma ninguna medicación. Tras seis meses de tratamiento con medidas higiénico dietéticas las cifras de TA son de 142/98. Presenta una obesidad leve con un IMC de 28. Nos planteamos comenzar con un tratamiento farmacológico.

El objetivo del médico en el momento de prescribir un tratamiento antihipertensivo debe ser no sólo normalizar la presión arterial sino también restaurar la CVRS del paciente hipertenso, minimizando al mismo tiempo los efectos secundarios del tratamiento a través de la selección del medicamento más apropiado a las características del paciente concreto y teniendo siempre en cuenta el coste del tratamiento de una manera amplia (no sólo el coste directo, sino también la medida del coste-utilidad).

Ante este paciente, antes de comparar el cambio en la CVRS que producen las distintas terapias hipertensivas de igual eficacia y seguridad, se nos plantea la cuestión de analizar la influencia que podrían tener diversas variables sociodemográficas y clínicas en su calidad de vida, así como de evaluar el efecto que tendría la optimización o intensificación de la pauta terapéutica antihipertensiva en la CVRS. Cuando evaluamos la CVRS en los hipertensos necesariamente tenemos que valorar varios parámetros: a) si el diagnóstico de HTA supone un cambio en la vida del paciente y cómo lo percibe; b) el tratamiento higiénico dietético; c) los efectos secundarios de la medicación antihipertensiva y d) la falta de cumplimiento prescrito.

Por ello es fundamental poder monitorizar la repercusión que la enfermedad en la CVRS del paciente, su evolución al recibir un tratamiento antihipertensivo, así como analizar el coste efectividad y coste oportunidad entre diversos tratamientos y valorar la percepción del propio paciente sobre su estado de salud o enfermedad.

PASOS A SEGUIR EN MBE

1. FORMULAR UNA PREGUNTA CLARA A PARTIR DEL PROBLEMA CLÍNICO A ANALIZAR

En este caso una pregunta a responder sería:

¿En pacientes con HTA la optimización o intensificación del tratamiento influye en la CVRS?

2. BÚSQUEDA SISTEMÁTICA DE LAS MEJORES PRUEBAS DISPONIBLES EN LA BIBLIOGRAFÍA PARA IDENTIFICAR LOS TRABAJOS RELEVANTES

Para realizar una búsqueda sistemática de los artículos científicos que nos pueden interesar, seguimos las indicaciones estudiadas en el capítulo "Búsqueda en la Colaboración Cochrane, en publicaciones secundarias, en bases de datos y en Internet", introduciendo los descriptores o palabras clave: Hypertension and Quality of Life

Así se encontraron 150 citas, de las que tras de la lectura del abstract, y en algunos casos del texto completo, se seleccionaron los siguientes artículos:

- Roca-Cusachs A. Badía X. Dalfó A. Gascón G. Abellán J. Lahoz R. Valera C. Velasco O., en nombre del Grupo MINICHAL. Relación entre variables clínicas y terapéuticas y calidad de vida relacionada con la salud en pacientes con hipertensión arterial. Estudio MINICHAL. Med Clin (Barc) 2003; 121(1):12-7.

- Badía X. Roca-Cusachs A. Dalfó A. Gascón G. Abellán J. Lahoz R. Valera C. Velasco O., on behalf of the MINICHAL Group. Validation of the Short Form of the Spanish Hypertension Quality of Life Questionnaire (MINICHAL). Clin Ther 2002; 24 (12): 2137-54.

3. VALORACIÓN CRÍTICA DE LAS EVIDENCIAS CIENTÍFICAS ENCONTRADAS

El primer trabajo seleccionado trata de un estudio observacional, prospectivo y multicéntrico en pacientes hipertensos que iniciaban o modificaban tratamiento antihipertensivo. Con el objetivo de evaluar, mediante un cuestionario validado, la influencia sobre la CVRS de variables sociodemográficas, clínicas y del efecto de la optimización del tratamiento antihipertensivo. Se evaluó a un total de 736 pacientes hipertensos que debían realizar una visita basal y dos visitas de seguimiento, una al mes y otra a los 6 meses de su inclusión en el estudio tras el inicio o la modificación del tratamiento antihipertensivo, en la que cada médico efectuaba de forma abierta el escalamiento terapéutico que considerara oportuno. En cada una de las visitas se registraron variables clínicas, acontecimientos adversos y clínicos, se administró el cuestionario sobre la CVRS, MINICHAL y la versión española del Índice de Bienestar Psicológico (IBP) y se planteó una pregunta que evaluaba el cambio en el estado de salud general de los pacientes respecto a la visita basal (Escala de Cambio en el Estado de Salud).

Pasos a seguir en la evaluación de un estudio de CVRS

1er paso: ¿Contiene el instrumento aquellas dimensiones esenciales a medir de acuerdo a la población y enfermedad?

El cuestionario MINICHAL es una versión reducida de otro cuestionario específico y validado de calidad de vida para la hipertensión arterial, el CHAL, y está compuesto por 16 ítems agrupados en dos dimensiones: a) estado de ánimo y b) manifestaciones somáticas.

Aunque otros cuestionarios de CVRS específicos de HTA, no adaptados ni validados en España, incluyen hasta 7 dimensiones consideradas relevantes (función cognitiva, funcionamiento sexual, bienestar psicológico, alteraciones del sueño, funcionamiento social, efectos adversos y percepción de la salud general), podemos considerar que las dos dimensiones del cuestionario MINICHAL reflejan la influencia de la severidad de la enfermedad y el modelo conceptual subyacente, dado que incluye preguntas sobre dificultades para conciliar el sueño, para continuar con relaciones sociales habituales, etc.

2º paso: ¿Se especifican las propiedades de medición (psicométricas) del instrumento utilizado para medir la CVRS?

- Evaluación de la fiabilidad:

La fiabilidad como estabilidad (en el sentido de constancia de una medición) fue medida por el método test-retest dando un coeficiente de correlación intraclase CCI de 0.82 para la dimensión estado de ánimo y de 0.75 para la dimensión manifestaciones sintomáticas.

La fiabilidad como consistencia interna (grado en que los ítems se correlacionan entre ellos mismos) fue medida mediante el coeficiente alfa de Cronbach obteniéndose un resultado de 0.87 para la dimensión estado de ánimo y de 0.75 para la dimensión manifestaciones sintomáticas.

Los coeficientes de fiabilidad superiores a 0.70 pueden considerarse satisfactorios en la mayor parte de los casos. Cuanto más alto sea el valor del coeficiente de fiabilidad, tanto más estable será la medición.

- Evaluación de la validez:

- Validez de contenido

Tiene validez de contenido y una validez aparente o facial pues el instrumento contiene ítems representativos de todas las dimensiones que forman la definición del concepto o constructo en estudio. Además fue analizada también mediante un análisis factorial exploratorio (por ejes y con rotación varimax) que definió las dimensiones del cuestionario y permitió aceptar la adecuación de los factores obtenidos.

- Validez de constructo

Se analizó la relación entre las puntuaciones del MINICHAL y las características sociodemográficas y clínicas de los pacientes, observándose que la HTA influye más en el estado de ánimo que en las manifestaciones somáticas. Las puntuaciones obtenidas apoyan las hipótesis establecidas con relación al constructo evaluado.

- Para el análisis de la validez convergente se estudiaron, mediante el coeficiente de correlación de Spearman, las relaciones entre el estado de ánimo y manifestaciones somáticas, y las puntuaciones de las dimensiones del Índice de Bienestar Psicológico (IBP). Las dos dimensiones presentaron correlaciones estadísticamente significativas ($p < 0.01$) con todas las dimensiones del IBP.

- Validez de criterio

No se utilizó ningún criterio externo (gold standard) para medir esta validez en ninguno de sus aspectos (concurrente o predictiva).

- Sensibilidad al cambio

Mediante la prueba de t-student de comparación de medias para datos apareados y el cálculo del tamaño del efecto se evaluaron los cambios en las puntuaciones de cada dimensión entre la visita basal y a los 6 meses. El tamaño del efecto de los cambios aumentaba en ambas dimensiones a medida que los cambios percibidos por los pacientes eran más importantes.

- Interpretabilidad:

Los 16 ítems del cuestionario hacen referencias a los “últimos 7 días” con cuatro posibles opciones de respuesta (“No, en absoluto”, “Sí, algo”, “Sí, bastante” y “Sí, mucho”). El rango de puntuación oscila en la dimensión Estado de Ánimo de 0 (mejor nivel de salud) a 30 (peor nivel de salud). Y de 0 (mejor nivel de salud) a 18 (peor nivel de salud) en la dimensión Manifestaciones Somáticas.

- Carga de respuesta y de administración:

Es un cuestionario autoadministrado y por tanto con baja carga de administración, cuya media de cumplimentación en condiciones de práctica clínica habitual es de 7.2 minutos.

3er paso: ¿Ha sido traducido, adaptado y validado al idioma y cultura del país donde se va a emplear (o se ha empleado)?

Tanto el cuestionario CHAL como el MINICHAL son instrumentos originales creados y validados en nuestro medio. Por lo que en este caso no entramos a valorar este paso.

4. APLICABILIDAD DE LOS RESULTADOS DE NUESTRA VALORACIÓN A NUESTRA PRÁCTICA CLÍNICA

Dado el diseño abierto del estudio, sin aleatorización ni grupo control, no es posible diferenciar si el beneficio observado en CVRS es independiente o no del fármaco que se utiliza. Por otra parte, dado que el estudio se realizó principalmente en centros de Atención Primaria cuyos pacientes eran mayoritariamente hipertensos con afectación orgánica de grado I y II los resultados no pueden ser directamente extrapolables a los pacientes con afectación orgánica de grado III.

En conclusión, los resultados del trabajo apuntan a que la intensificación del tratamiento antihipertensivo no sólo puede incidir favorablemente en el pronóstico vital sino que además tiene un impacto positivo en la CVRS, y que en los pacientes con ciertas características (nivel de educación más bajo, sexo femenino, con hipertensión arterial más grave o mayor repercusión orgánica o con mayor comorbilidad) la intervención terapéutica debe estar especialmente dirigida no sólo a corregir los factores modificables, sino a mejorar una CVRS ya de por sí alterada en este tipo de pacientes.

RECURSOS COMENTADOS

Páginas Web sobre calidad de vida, medición de la salud y aspectos relacionados:

Agency for Healthcare Research and Quality: Web del U.S. Health Service que ofrece información relacionada con la investigación en calidad de vida. Ofrece también información en español.

<http://www.ahrpr.gov/>

European Clearing Houses on Health Outcomes: Grupo que reúne a instituciones de diversos países europeos relacionadas con la investigación de resultados en salud.

<http://www.leeds.ac.uk/nuffield/infoservices/ECHHO/home.html>

Euroqol Group: Se ofrece información el cuestionario Eurqol (medida genérica de CVRS) y referencias bibliográficas relacionadas con el EQ-5D..

<http://www.euroqol.org>

Evaluación Clínica y Económica de Medicamentos: Web española que proporciona recursos en evaluación clínica y económica de medicamentos, y también en CVRS.

<http://www.farmacoeconomia.com>

International Society for Quality of Life Research: Web que ofrece información acerca de esta sociedad y sus congresos internacionales ISOQOL sobre calidad de vida.

<http://www.isoqol.org>

MAPI Research Institute: Web del instituto francés de investigación en calidad de vida. ofrece adaptaciones transculturales de cuestionarios genéricos y específicos clasificado por tipo de medidas, países e idiomas.

<http://www.mapi-research-inst.com>

Medical Outcomes Trust: Web ofrece información sobre la utilización de instrumentos estandarizados de medida de la salud y del resultado de las intervenciones sanitarias.

<Http://www.outcomes-trust.org>

Quality of Life Instruments Database: Base de datos que recoge información sobre cuestionarios genéricos y específicos de medida de la salud y calidad de vida.

<http://www.qolid.org>

Red de investigación en Resultados de Salud y Servicios Sanitarios (Red IRYSS): Red cooperativa de grupos de investigación entre cuyas líneas está la investigación de resultados percibidos y calidad de vida.

<Http://www.rediryss.net/pub/units/rediryss/html/es/dir293/index.html>

Webs de revistas relacionadas con la calidad de vida y la CVRS

Listado de revistas científicas que tratan temas de medida de la calidad de vida relacionada con la salud.

Disease Management & Health Outcomes

<http://www.adis.com/page.asp?objectID=43>

Quality of Life Research

<http://www.kluweronline.com/issn/0962-9343/contents>

Pharmacoeconomics

<http://www.adis.com/page.asp?objectID=52>

Expert Reviews of Pharmacoeconomics & Outcomes Research

<http://www.future-drugs.com>

Clinical Therapeutics

<http://www.clinical-therapeutics.com>

Atención Primaria

http://db.doyma.es/cgi-bin/wdbcgi.exe/doyma/mrevista.salta_a_ultimo?pid=27

Gaceta Sanitaria

http://db.doyma.es/cgi-bin/wdbcgi.exe/doyma/mrevista.salta_a_ultimo?pid=138

Medicina Clínica

http://db.doyma.es/cgi-bin/wdbcgi.exe/doyma/mrevista.salta_a_ultimo?pid=2



Acotado Automático de Términos: Brevemente, consiste en el enlace automático de una palabra o frase (que no es un término MeSH) al término MeSH que mejor lo describe. Una descripción más completa en el apartado correspondiente. Sinónimos (Sin): Automatic Term Mapping, ATM, Mapeo automático de términos.

Aleatorización: En un estudio en el que se comparan dos o más modalidades de tratamiento, es conveniente que los grupos formados sean parecidos en todas las características pronósticas excepto en el tratamiento recibido, de modo que cualquier diferencia en el curso clínico que se registre podrá ser atribuida a los diferentes tratamientos administrados (y sólo a ellos). La asignación aleatoria consiste en distribuir a cada participante de uno de los grupos de tratamiento por un método disciplinado de azar, de modo que cada sujeto tenga exactamente las mismas probabilidades de formar parte de uno u otro grupo de tratamiento. Sin: Randomización.

Alertas: Método usado para la difusión de información mediante mensajes electrónicos a ordenadores, agendas electrónicas e incluso teléfonos móviles. Representa una forma de actualización rápida y eficaz.

Análisis coste/beneficio: Medición de los costes y beneficios de una intervención en términos comensurables, generalmente monetarios. Permite determinar en primer lugar si una intervención concreta ofrece un beneficio global neto de bienestar, y en segundo lugar cómo se compara la ganancia de bienestar obtenida con esta intervención, con la obtenida con intervenciones alternativas. Al valorar todos los costes y beneficios en las mismas unidades, el análisis coste-beneficio compara diversas intervenciones con el criterio del beneficio neto (véase). Análisis económico en el que se consideran los costes de la asistencia sanitaria y la pérdida de ganancias netas debida a muerte o incapacidad, traduciendo todos los costes y beneficios a unidades monetarias.

Análisis coste/efectividad: Modalidad de estudio económico en la que las consecuencias de diferentes intervenciones pueden ser diferentes pero pueden ser medidas en unidades naturales idénticas, Mide los beneficios sanitarios en términos de unidades naturales, como años de vida salvados o mejorías del estado funcional. Dado que los costes no se expresan en las mismas unidades que los beneficios, su razón (ratio coste-efectividad) constituye una herramienta que permite valorar la eficiencia relativa. Análisis económico que compara los costes médicos incrementales y los resultados sobre la salud de programas sanitarios alternativos. Al contrario que el análisis de coste beneficio, los efectos sobre la salud se expresan en diversas unidades, como vida-años ganados, días libres de síntomas, casos evitados, etc., en lugar de en unidades monetarias

Análisis coste-utilidad: Adaptación del análisis coste-efectividad que mide el efecto de una intervención a la vez sobre los aspectos cuantitativos y cualitativos de la salud, mediante el uso de una variable denominada AVAC -Años de Vida Ajustados por Calidad- o QALYs . También evalúa la eficiencia relativa con una razón incremental, que en este caso se denomina ratio coste-utilidad.



Análisis de decisiones (AD): Es la aplicación de métodos explícitos y cuantitativos que cuantifican e integran los efectos del tratamiento, el pronóstico y los valores del paciente para analizar una decisión en condiciones de incertidumbre.

Análisis de minimización de costes: Modalidad de análisis coste-efectividad donde los resultados o beneficios de las alternativas se considera que son equivalentes o intercambiables para el mismo tipo de pacientes. A igualdad de resultados, la atención de los investigadores se centra en la diferencia de costes asociados.

Análisis de sensibilidad: En economía de la salud, técnica por la que se repiten las comparaciones entre recursos necesarios y resultados, pero con sucesivas modificaciones de las suposiciones en las que se basan los cálculos. De este modo, el análisis de sensibilidad pone a prueba la robustez de las conclusiones mediante variaciones de los ingredientes sobre los que existe incertidumbre.

Análisis de regresión: En un conjunto de datos sobre la variable dependiente y sobre una o más variables independientes, x_1, x_2, \dots, x_i , consiste en determinar el modelo matemático más ajustado que describa y como una función de las x o para predecir y a partir de las x . Término general para los métodos de análisis que se ocupan de estimar los parámetros de alguna posible relación entre una variable respuesta y una o más variables explicativas

Análisis de la varianza (ANOVA): Conjunto de técnicas estadísticas para conocer el modo en que el valor medio de una variable es afectado por diferentes tipos de clasificaciones de los datos. Con el análisis de la varianza se pueden ajustar las estimaciones del efecto de un tratamiento según otros factores como sexo, edad, gravedad, etc.

Análisis estratificado: Análisis de los resultados de un ensayo clínico o un estudio observacional, con el fin de examinar si hay diferencias de respuestas asociadas a alguna característica (por ej., edad o sexo).

Análisis factorial: Procedimiento estadístico empleado para la validación de constructo que permite identificar grupos o conglomerados de variables relacionadas.

Análisis multivariante: Cualquier método analítico que permite el estudio simultáneo de dos o más variables dependientes. Término genérico para los métodos que están diseñados para determinar las contribuciones relativas de diferentes causas a un solo suceso o resultado

Análisis por intención de tratar: En un ensayo clínico, análisis de los datos según el tratamiento asignado inicialmente por distribución aleatoria, independientemente de que el paciente no recibiera después este tratamiento.

Árbol de decisiones: Expresión gráfica de un análisis de decisiones.

Artículo científico: informe escrito que comunica por primera vez los resultados de una investigación. Los artículos científicos publicados en miles de revistas científicas componen la literatura primaria de la ciencia

Artículo de revisión: artículos de síntesis que resumen el conocimiento de un tema determinado y durante un periodo de tiempo definido. Junto con los libros, los artículos de revisión constituyen la literatura científica secundaria.

Asiento: Conjunto de información que, en una base de datos, se refiere a un mismo elemento. En Bases de Datos de bibliografía, es el conjunto de información que se refiere a una misma publicación. Sin: Registro, Record.

ATM: Ver Acotado Automático de Términos.

Atributos: Propiedades o características del instrumento que merecen consideraciones separadas e independientes en la evaluación.

Auditoria: Examen o estudio que evalúa e informa sobre la medida en que una condición, proceso o desempeño se ajusta a estándares o criterios predeterminados.

Automatic explosion: Ver Explode.

Automatic Term Mapping: Ver Acotado Automático de Términos.

AVAC: Índice que combina la supervivencia con la calidad de vida. La medida de la calidad de vida no está estandarizada y suele variar de un estudio a otro, según la enfermedad, el tratamiento evaluado y las preferencias de los autores. Para calcular un AVAC se multiplica la duración del estado de salud (en años) por un factor que representa la calidad (“utilidad”) de este estado de salud. Sin: QALY (Quality Adjusted Life Years).

Azar: Se denomina así al proceso que ocurre debido a la casualidad, de forma imprevisible, sin planeamiento ni influencia consciente.

Base de datos: Conjunto de datos homogéneos y ordenados, que se presentan normalmente, de forma legible, por ordenador y que se refieren a una materia determinada. Colección estructurada de datos que está organizada de tal manera que se puede acceder a ella con una amplia variedad de programas de aplicación.

Beneficio neto: Diferencia entre los costes y los beneficios de una intervención.

Bioestadística: Aplicación de la teoría y los métodos de la estadística a las ciencias biológicas. En sentido estricto, rama de la ciencia que aplica los métodos estadísticos a los problemas biológicos, aunque ahora se utiliza la mayoría de las veces para referirse a la estadística aplicada a la medicina y las ciencias de la salud.

Blinding: Ver Enmascaramiento.

Buscadores: Motores de búsqueda de información en Internet que identifican la información contenida en bases de datos formadas a partir de publicaciones previamente indexada. Ver Indización.

Calidad de Vida Relacionada con la Salud (CVRS): Impacto que una enfermedad y el tratamiento consiguiente tienen sobre la percepción del paciente de su satisfacción y bienestar físico, psíquico, social y espiritual.

Campo: Cada una de las unidades de información en que se divide un asiento. Sin: Field.

Carga de administración: Se define como el grado de trabajo y esfuerzo que deben realizar quienes administran un instrumento de medida (cuestionario), y depende del tiempo precisado para su administración, formación requerida del entrevistador y recursos necesarios para su aplicación y puntuación.

Carga de respuestas: Se entiende como el tiempo, esfuerzo y otros aspectos que se exigen a quienes se les administran un instrumento de medida (cuestionario).

CAT: Herramienta de lectura crítica consistente en un resumen breve de las pruebas científicas más relevantes en respuesta a una pregunta clínica concreta.

Causalidad: Indica que las asociaciones encontradas entre las exposiciones y los efectos son realmente de tipo causal y no consecuencia de sesgos ni del efecto de factores de confusión.

Cociente de probabilidad: Véase Razón de probabilidad.

Coefficiente de correlación: Índice que cuantifica la relación lineal entre un par de variables. Medida de asociación que indica el grado en el que dos variables continuas x e y poseen una relación lineal ($y = a + bx$). Su valor puede situarse entre -1 y $+1$. Cuando su valor es 0 no existe ningún grado de correlación. El coeficiente de este tipo que más se utiliza es el de Pearson

Coefficiente de Cronbach: Fórmula utilizada para la comprobación de la consistencia interna de un test que presupone que si varios ítems están midiendo el mismo atributo o dimensión, estos ítems deberían estar intercorrelacionados. Esta correlación la mediremos mediante el coeficiente alfa de Cronbach.

Coefficiente kappa: índice corregido por el azar de la concordancia entre, por ejemplo, juicios o diagnósticos que hacen 2 ó más observadores. El coeficiente adopta desde el valor 1 (cuando hay una concordancia perfecta) hasta el valor 0 cuando la concordancia observada es igual a la debida al azar.

Confusión: Situación en la que la estimación del resultado estudiado está sesgada por algún factor relacionado con la exposición y el resultado.

Consentimiento informado: Procedimiento formal para aplicar el principio de autonomía. Debe reunir por lo menos tres elementos: voluntariedad, información y comprensión. La voluntariedad implica que los sujetos puedan decidir libremente si quieren participar en un estudio, sin que haya persuasión, manipulación ni coerción. La información debe ser comprensible y debe incluir el objetivo del estudio y su procedimiento, los beneficios y riesgos potenciales, las posibles incomodidades derivadas de la participación y la posibilidad de retirarse del estudio en cualquier momento sin que ello de lugar a perjuicio, en este contexto es básica la comprensión de la información, que no sólo depende de esta, sino también de la competencia de quien la recibe, es decir de su capacidad para comprender.

Consistencia interna: Grado en que los ítems de una dimensión se correlacionan entre ellos mismos. Ver: Coeficiente de Cronbach.

Constructo: En psicometría, concepto abstracto no medible directamente que se quiere convertir en variable operativa medible. Generalmente, se utiliza para un concepto que existe en teoría, pero que no se puede observar directamente. Esencialmente, es una variable latente.

Correlación ítem-total:(Correlación total de las preguntas): Verifica la homogeneidad. Proceso mediante el cual comprobamos si cada ítem mide lo mismo que los demás y por lo tanto es sumable en una puntuación total que supuestamente mide el rasgo y que es la que después interpretamos y utilizamos.

Coste: Es el resultado del producto de la cantidad total del recurso consumido y el valor monetario de la unidad de dicho recurso.

Coste oportunidad: En economía de la salud, valor de las oportunidades perdidas por usar recursos en un servicio. Generalmente es cuantificado considerando el beneficio que se obtendría invirtiendo los mismos recursos en la mejor alternativa posible.

Criterios: Condiciones utilizadas como un estándar para juzgar los atributos.

Criterios de selección: Características sociodemográficas, clínicas, etc. que determina que los pacientes puedan ser o no incluidos en un estudio. Nos van a servir para definir la población de estudio.

Curva de Gauss: En matemáticas, la campana o curva de Gauss es la representación gráfica de la ecuación matemática que corresponde a una Distribución normal. Tiene forma de campana y debe su nombre al matemático alemán Carl Friedrich Gauss.

Curvas ROC: Representación gráfica completa de la capacidad de una prueba para discriminar lo normal de lo que no lo es.

Delimitador de campo: Ver Etiqueta.

DeCS - Descriptores en Ciencias de la Salud: Tesauro que puede considerarse, hasta cierto punto, como una versión en español del MeSH - Medical Subject Headings de la U.S. National Library of Medicine. Fue desarrollado a partir de él con el objetivo de permitir el uso de terminología común para búsqueda en tres idiomas (español, inglés y portugués), proporcionando un medio consistente y único para la recuperación de la información independientemente del idioma. Además de los términos médicos originales del MeSH contiene áreas específicas de Salud Pública y Homeopatía.

Descriptor: Término del lenguaje documental, carente de variabilidad, sinonimia y polisemia, escogido para representar el contenido de un documento, que puede ser utilizado en el análisis y recuperación de la información. Relacionado con Palabra clave.

Descriptor mayor: Ver Mayor Topic.

Descriptores en Ciencias de la Salud: Ver DeCS

Descuento: Procedimiento para calcular los costes o beneficios que ocurren en diferentes fechas a su valor en el momento presente, llamado valor actual.

Dimensiones: Aspectos relevantes que engloba el constructo o concepto que se mide a través del cuestionario.

Diseño: Conjunto de estrategias y procedimientos que se seguirán para dar respuestas al problema y comprobar la hipótesis. Engloba al selección del tipo de estudio, la población a estudiar, la definición de las variables y escalas de medición y los métodos de recogida, procesado y análisis de los datos.

Docencia virtual (“on-line” o “e-learning”): Definida como la utilización de las nuevas tecnologías multimedia y de internet para mejorar la calidad del aprendizaje, facilitando el acceso de recursos y servicios; así como los intercambios y la colaboración a distancia.

Dosis Diaria Definida (DDD): Dosis de mantenimiento promedio en la principal indicación de un fármaco en pacientes adultos. Se usa para cuantificar el consumo de medicamentos, sobre todo cuando se compara el consumo de diversos fármacos de un mismo grupo, o cuando se realizan comparaciones internacionales. No se ve afectada por diferencias de precios o de tamaño de las presentaciones. El consumo se suele expresar en número de DDD por 1.000 habitantes y por día, y esta medida da una idea del número promedio de personas tratadas diariamente con un fármaco o grupo de fármacos en una región o país.

Efectividad: Grado en que una determinada intervención origina un resultado beneficioso en las condiciones de la práctica clínica habitual, sobre una población determinada.

Efecto indeseado: Ver Reacción Adversa.

Eficacia: Grado en que una determinada intervención origina un resultado beneficioso en condiciones ideales de aplicación.

Eficiencia: Relación entre los efectos o resultados finales que se alcanzan con una intervención y coste que supone la misma (recursos humanos, materiales y de tiempo).

Encabezamientos de Materias Médicas: Ver MeSH.

Enmascaramiento: Procedimiento por el que se asegura que los sujetos participantes en un ensayo clínico, los observadores o ambos, no conocen el tratamiento u otra característica que pudiera sesgar los resultados. Se denomina simple ciego cuando el sujeto, pero no el observador ignora cuál de los posibles tratamientos recibe; doble ciego es cuando sujeto y observador ignoran el tratamiento recibido; y triple ciego es cuando el sujeto participante, el investigador observador y el investigador que analiza los datos ignoran el tratamiento recibido.

Ensayo clínico: “Toda investigación efectuada en seres humanos para determinar o confirmar los efectos clínicos, farmacológicos y/o demás efectos farmacodinámicos, y/o de detectar las reacciones adversas, y/o de estudiar la absorción, distribución, metabolismo y excreción de uno o varios medicamentos en investigación con el fin de determinar su seguridad y/o su eficacia” (Real Decreto de 16 de febrero de 2004).

Ensayo clínico controlado: Ensayo clínico en el que los sujetos son asignados a dos grupos: uno (grupo experimental) recibe el tratamiento que se está probando y el otro (grupo de comparación o control) recibe un tratamiento alternativo. En el ensayo clínico controlado aleatorizado la asignación a un tratamiento experimental o control se realiza al azar de forma que ni el sujeto ni el investigador responsable de su selección o tratamiento puedan influir en su asignación.

Entry Term: Relacionado con Palabra clave (véase).

Epidemiología: Estudio de la distribución y los determinantes de los estados o acontecimientos relacionados con la salud de poblaciones. Ciencia que investiga las causas y la historia de las enfermedades y el desarrollo de éstas en relación con determinadas intervenciones o exposiciones.

Epistemología: Doctrina de los fundamentos y métodos del método científico. La función de la estadística en el modelo empírico analítico.

Equivalencia: capacidad del instrumento de proporcionar los mismos valores en las mismas condiciones clínicas cuando las mediciones las hacen distintas personas.

Error aleatorio: Las diferencias de medidas entre la realidad y nuestra muestra y son debidas al azar.

Escala tipo Likert: La respuesta del individuo se da mediante una escala numérica, a la cual se le atribuyen las propiedades de las escalas de intervalo, permitiendo por tanto la realización de análisis estadísticos más complejos. Escalas que se utilizan con frecuencia en estudios de actitudes en los que las puntuaciones crudas se basan en respuestas alternativas graduadas a cada una de una serie de preguntas.

Especificidad: Referida a una prueba diagnóstica, es la proporción de personas verdaderamente no enfermas que son catalogadas como tales por medio de la prueba.

Estabilidad: Ver Reproducibilidad.

Estabilidad del test: Véase Método test-retest.

Estadística: Recopilación y organización de hechos o datos numéricos, sobre todo cuando hay muchos. Conjunto de métodos que ayudan a la toma de decisiones frente a la incertidumbre. La estadística no ofrece pruebas ni certidumbre, sino probabilidades.

Estadístico: Medida estadística (media, mediana, etc...) calculada sobre los datos de una muestra.

Estadístico kappa: Medida del acuerdo, corregida por el azar, aplicada a datos binarios u ordinales. Se calcula como la proporción de acuerdo, aparte del que ya sería de esperar por azar, que ha sido observado entre dos repeticiones del mismo instrumento.

Estudios ambispectivos: Estudios que se inician como los retrospectivos (Ver: Estudios retrospectivos) para el análisis de un fenómeno ya ocurrido y, posteriormente, continuamos con un seguimiento prospectivo hasta la aparición de un efecto a largo plazo (Ver: Estudios prospectivos).

Estudios analíticos: Estudios diseñados para examinar asociaciones, cuyo objeto final suele ser identificar o medir los efectos de factores de riesgo o de intervenciones específicas sobre la salud.

Estudios de caso-control: Estudios en los que la selección de la población a estudiar se hace en función de la presencia o ausencia de una enfermedad o efecto, eligiéndose un grupo de individuos que tienen la enfermedad o efecto (casos), y otro que no los presenta (controles).

Estudios de cohortes: Tipo de estudios en los que los sujetos del estudio se seleccionan a partir presencia o ausencia de la exposición que se trata de analizar y de la que se sospecha es responsable del efecto observado.

Estudio descriptivo: Estudio cuyo objetivo es constatar hechos, observar lo que ocurre en nuestra población. Permiten generar hipótesis.

Estudio experimental: Estudio en el que el investigador asigna y controla el factor de estudio.

Estudios de incidencia: Tipo de estudios en los que seguimos a una población determinada a lo largo del tiempo para analizar la aparición o no de un fenómeno de interés. La unidad de análisis es el individuo.

Estudios longitudinales: Estudios en los que la recogida de los datos de interés hace referencia a dos momentos separados en el tiempo (se analizan los sujetos en varios momentos a lo largo del tiempo).

Estudios observacionales: Estudios en los que el investigador no participa directamente en la intervención o exposición (factor de estudio), se limita a observar (registrar) lo que ocurre en la realidad.

Estudios pragmáticos: Estudios cuyo objetivo principal es la evaluación de la adecuación de la práctica asistencial en condiciones normales.

Estudios pre y post-intervención: Estudios en los que se registran la situación basal de la muestra y los cambios que experimenta tras su exposición a una determinada intervención.

Estudios de prevalencia: Estudios en los que se describe una serie de variables en una población determinada y en un momento determinado. La unidad de análisis es el individuo.

Estudios prospectivos: Tipo de estudios en los que los hechos a estudiar no han tenido lugar aun y los sujetos son seguidos hasta obtener el resultado.

Estudio retrospectivo: Estudio en el cual los hechos a estudiar ya han tenido lugar al iniciar el estudio.

Estudio transversal: Tipo de estudio en el que se estudia la exposición y el desenlace al mismo tiempo.

Etiqueta: Abreviatura del nombre de un campo que puede utilizarse en su lugar para dirigir la búsqueda de información. Sin: Tag, Delimitador de campo.

Evaluación crítica: Véase Lectura crítica.

Evaluación económica: Análisis comparativo de opciones alternativas en términos de sus costes y consecuencias. Los estudios de evaluación económica pueden ser parciales, cuando sólo se describen los costes, los resultados o ambos a la vez pero de una única opción o que comparan dos o más alternativas pero sólo en función de sus costes o de sus consecuencias sin valorar ambos a la vez, o completos, cuando se analizan comparativamente dos o más alternativas mediante la valoración de los costes y de las consecuencias de cada una de ellas.

Evaluación de tecnologías sanitarias: Proceso de análisis e investigación dirigido a estimar el valor y contribución relativos de cada tecnología sanitaria (véase) en la mejora de la salud individual y colectiva, teniendo en cuenta su impacto económico, social e, incluso, implicaciones éticas.

Explode: Obtención de las referencias que contienen el descriptor especificado o cualquiera de los descriptores de su mismo campo semántico pero de significado más restringido. Sin: Automatic explosion, Exploding, Explosion, Explosión, Explosión automática.

Exploding: Ver Explode.

Exploración complementaria: Conjunto de pruebas diagnosticas realizadas con la finalidad de confirmar una hipótesis diagnostica o hacer el seguimiento de la patologia de un paciente.

Explosión: Ver Explode.

Explosión automática: Ver Explode.

Factibilidad: Indica las posibilidades reales y práctica de que un instrumento de medida pueda ser contestado, completado y analizado. Está relacionad con el nivel de aceptabilidad del instrumento y se refleja en la proporción de rechazos, no respuesta y datos perdidos.



Factor de confusión: Variable que está simultánea e independientemente asociada al factor de riesgo estudiado y a la enfermedad estudiada, que puede modificar el resultado del estudio. Tal variable debe ser identificada y se debe evitar su influencia antes de la recogida de datos (mediante el emparejamiento o la restricción) o bien en la fase de análisis, mediante el análisis estratificado y la regresión múltiple.

Factor de impacto: Indicador bibliométrico, elaborado por el ISI (véase) y publicado en el Journal Citation Reports, que mide el uso de las revistas. Suele calcularse dividiendo el número de veces que son citados los artículos publicados en una revista durante un plazo de tiempo (habitualmente dos años), entre el número de artículos publicados por esa misma revista y en ese mismo plazo de tiempo. Puede modificarse para calcular el impacto de autores, instituciones... Sin: Impact Factor.

Factor de riesgo: Factor relacionado con el riesgo de que ocurra un suceso (p. ej. enfermar).

Falso negativo: Referido a una prueba diagnóstica, número de enfermos en que el resultado sería negativo.

Falso positivo: Referido a una prueba diagnóstica, número de sanos en que el resultado sería positivo.

Farmacovigilancia: Programas que proporcionan de forma continuada una evaluación sobre la seguridad de los fármacos. El factor clave en este sistema son los profesionales sanitarios que participan en el programa de notificación espontánea de reacciones adversas.

Fiabilidad: Grado de consistencia con que mide el atributo para el que está diseñado. La consistencia implica la ausencia de variabilidad o desacuerdo en repetidas observaciones o mediciones. Grado en que las mismas mediciones de los individuos que se obtienen en diferentes condiciones dan resultados similares.

Fiabilidad inter-observador (inter-ratio): Mide la variación que ocurre entre diferentes observadores al medir a un mismo individuo con un mismo instrumento .

Field: Ver Campo.

Filtros metodológicos: Son estrategias de búsqueda que nos ayudan a encontrar, en bases de datos como Medline, diferentes tipos de estudio sobre tratamiento, diagnosis, etiología o pronóstico. En Pubmed se denominan "Clinical Queries". Para aplicar uno de estos filtros sólo tenemos que combinar nuestra búsqueda con AND al filtro que nos interesa.

Forest/ Forest Plot: Tipo de gráfico ampliamente utilizado para representar los resultados de los diversos estudios seleccionados para una revisión sistemática. El Forest Plot ayuda a decidir si los resultados de los estudios individuales son suficientemente similares para poder ser combinados en una sola estimación.

Generabilidad: Véase Validez externa.

Gold Standard: Es el criterio que define quién tiene "realmente" la enfermedad (o condición) en estudio.

Grado o nivel de significación: En las pruebas de significación estadística, es el valor de p, el cual, en sentido estricto, en un ensayo clínico debe ser preespecificado en la fase de diseño, el nivel aceptado con mayor frecuencia es de 0.05. Ver: P (probabilidad), Significación clínica y Significación estadística.

Grupo de comparación: distintos grupos que intentamos comparar en un estudio analítico, ya sean estudios caso-control, estudios de cohortes o ensayos clínicos.

Grupo control: En un ensayo clínico, el grupo que recibe placebo o bien un tratamiento activo ya conocido, aceptado y establecido. Sin: Grupo de referencia.

Grupo intervención: En un ensayo clínico, el grupo que recibe el tratamiento que se desea evaluar.

Guía farmacoterapéutica: Documento que contiene una relación limitada de medicamentos recomendados para la prescripción en un ámbito determinado, seleccionada a partir de la oferta farmacéutica en función de unos criterios previamente establecidos, con la participación y el consenso de los profesionales a los que va destinada.

Guía de practica clínica: Conjunto de recomendaciones diseñadas para ayudar a los profesionales sanitarios y a los usuarios a seleccionar las opciones diagnósticas y/o terapéuticas más adecuada en el abordaje de una condición clínica específica.

Hipótesis: Suposición, fundada en observaciones o reflexiones, que puede conducir a predicciones refutables. También, conjetura elaborada de forma que pueda ser comprobada y refutada.

History: Archivo de todas las consultas bibliográficas realizadas a lo largo de una sesión en PubMed. También designa al subgrupo de registros que provienen de artículos sobre historia de la medicina.

Holística: Relativa al "holismo" que es la doctrina que propugna la concepción de cada realidad como un todo distinto de la suma de las partes que lo componen.

Impact Factor: Ver Factor de impacto.

Incremento absoluto del riesgo (IAR): Cuando el tratamiento experimental aumenta la probabilidad de un acontecimiento adverso, diferencia absoluta entre las tasas de acontecimientos adversos.

Incremento relativo del riesgo (IRR): Cuando el tratamiento experimental aumenta la probabilidad de un acontecimiento adverso, aumento proporcional de la tasas de acontecimientos adversos en el grupo experimental relativa a la tasa en el grupo de referencia.

Indicador de adhesión a la Guía Farmacoterapéutica: Indicador genérico de prescripción que indica el porcentaje de Dosis Diaria Definida (DDD, véase) de medicamentos prescritos incluidos en la Guía Farmacoterapéutica respecto del total de DDD de medicamentos prescritos.



Índice Bibliográfico Español en Ciencias de la Salud (IBECS): Base de Datos de la Biblioteca Nacional de Ciencias de la Salud (BNCS) del Instituto de Salud Carlos III (ISCIII), contiene referencias de artículos de revistas científico sanitarias editadas en España e incluye contenidos de las diferentes ramas de las ciencias de la salud tales como Medicina (incluyendo Salud Pública, Epidemiología y Administración Sanitaria), Farmacia, Veterinaria, Psicología, Odontología y Enfermería.

Índice Médico Español (IME): Base de datos referencial bibliográfica, creada y gestionada por el Instituto de Historia de la Ciencia y Documentación (IHCD), centro mixto de la Universidad de Valencia y del Consejo Superior de Investigaciones Científicas (CSIC). Cubre 321 publicaciones periódicas editadas en España especializadas en Biomedicina (Administración sanitaria, Farmacia clínica, Medicina experimental, Microbiología, Psiquiatría, Salud pública).

Indización: Proceso por el que el documentalista asigna unos descriptores que describen el contenido conceptual de documento con el fin de facilitar su recuperación

Index: Listado de todos los términos (unitérminos o compuestos) con significado que están incluidos en la Base de Datos. Esta opción permite hacer una búsqueda seleccionando los términos directamente del listado.

Indexar: Ver indización.

Ineficiencia: Situación en la cual los recursos no son empleados de la forma en que se incrementarían los beneficios producidos.

Inferencia: Proceso de pasar de las observaciones y axiomas a la generalizaciones. En estadística, desarrollo de la generalización a partir de los datos de la muestra, generalmente con un margen de incertidumbre.

Instrumentos de medida de CVRS específicos: Los instrumentos específicos incluyen dimensiones de la CVRS de una determinada afección que se quiere estudiar, capacidad funcional, o de un determinado grupo de pacientes.

Instrumentos de medida de CVRS genéricos: Los instrumentos de calidad de vida genéricos no están relacionados con ningún tipo de enfermedad o especialidad en particular. Son aplicables tanto a la población general como a grupos específicos de pacientes.

Interpretabilidad: grado en el que uno puede asignar un significado comprensible a las puntuaciones cuantitativas del instrumento.

Intervalo de confianza: Margen o recorrido comprendidos entre dos valores extremos (el límite inferior y el superior), entre los que cabe esperar el valor medio real de la población con un margen grado prefijado de seguridad. Ver: Nivel de confianza.

Intervención: Cualquier acción (preventiva, diagnóstica, sintomática o curativa) aplicada con el fin de modificar el curso de la situación previa. Cualquier tratamiento o actuación sanitaria que queramos evaluar

ISI: El Institute for Scientific Information fue creado por Eugene Garfield en 1958 con el objeto de publicar el primer Current Contents que cubría unas 200 revistas. En 1971, cubría más de 4000 revistas, publicaba más de 650 páginas de contenidos a la semana, y era consultado por más de un millón de científicos. En 2001, se desarrolló el ISI Web of KnowledgeSM, que integra literatura científica, tanto revistas (journals) como actas (proceedings), patentes, compuestos químicos, y hasta secuencias genómicas. El ISI fue adquirido por The Thomson Corporation, que lo convirtió en el Thomson Institute for Scientific Information. Es, por tanto, una institución privada que produce diversas Bases de Datos de primer orden en calidad y utilidad en la investigación científica, pero que sobre todo ha obtenido un notable éxito por la difusión de lo que se conoce como Impact Factor o Factor de Impacto.

ISSN: International Standard Serial Number, número internación único que identifica de manera inequívoca cada revista.

Ítem: Cada uno de los elementos, preguntas o cuestiones que constituyen un test.

Key word: Ver Palabra clave.

Lectura crítica: Es el proceso de evaluar e interpretar la evidencia aportada por la literatura científica, considerando sistemáticamente los resultados que se presenten, su validez y su relevancia para el trabajo propio. Sin: Evaluación crítica.

Lenguaje documental: Lenguaje artificial creado específicamente para el análisis y recuperación de la información. Contrapuesto al Lenguaje natural, se diferencia de este especialmente por carecer de variabilidad, sinonimia y polisemia.

Lenguaje natural: Expresión de ideas o de conceptos con los términos y con las reglas propios de un idioma o de un grupo especial de hablantes (p.e. sanitarios). Se contrapone a Lenguaje artificial y, específicamente en documentación, a Lenguaje documental.

Límites: Opción que ofrecen varias bases de datos, como Medline, que permite limitar/restringir una búsqueda de acuerdo a unas características como p.ej. artículos publicados en un idioma determinado, tipo de publicación (revisión, ensayo clínico), o grupo de edad, sexo, etc.

Likehood Ratio: Véase Razón de probabilidad.

Literatura inédita: Literatura que no ha sido publicada.

Mapeo Automático de Términos: Ver Acotado Automático de Términos.

Mayor Topic: Materia principal de un artículo. Sin: Descriptor mayor.

Médical Subject Headings: Ver MeSH

Medicina Basada en la Evidencia (MBE): Práctica de la medicina basada en la integración de la maestría clínica de cada médico (es decir, el dominio del conocimiento y el juicio adquirido por la experiencia) con las mejores pruebas clínica externas disponibles. También ha sido definida como la utilización consciente y juiciosa de las mejores demostraciones provenientes de la investigación clínica para la toma de decisiones en el cuidado de cada paciente.

Medicamentos genéricos (Especialidad Farmacéutica Genérica, EFG): Especialidad con la misma forma farmacéutica e igual composición cualitativa y cuantitativa en sustancias medicinales que otra especialidad de referencia, cuyo perfil de eficacia y seguridad está suficientemente establecido por su continuado uso clínico. La EFG debe demostrar la equivalencia terapéutica con la especialidad de referencia mediante los correspondientes estudios de bioequivalencia.

Medicamentos huérfanos: Cuando un fármaco no es desarrollado porque el detentador de la patente considera que no recuperará los costes de este desarrollo (en el caso de enfermedades raras y otras enfermedades no raras pero que ocurren en países pobres en los que no existe un “mercado” potencial), se lo denomina medicamento huérfano.

MeSH (Médical Subject Headings, Encabezamientos de Materias Médicas): Tesauro (diccionario que permite conocer la equivalencia entre los términos del lenguaje natural y los del lenguaje documental) de Medline y de PubMed.

MeSH Subheading: Subencabezamiento propio de MeSH.

MeSH Term: Descriptor propio de MeSH.

Meta-análisis: Integración estructurada y sistemática de la información obtenida en diferentes estudios sobre un problema determinado. Consiste en identificar y revisar los estudios controlados sobre un determinado problema, con el fin de dar una estimación cuantitativa sintética de todos los estudios disponibles.

Método matricial multirasgo-multimetodo (multitrait-multimethod matriz method, MTMM): procedimiento empleado para la validación de constructo aplicando para ello los conceptos de convergencia (validez convergente) y discriminabilidad (validez discriminante).

Método test-retest: Método utilizado para la comprobación de la fiabilidad de un test. Correlación entre la puntuación de dos administraciones de una prueba a los mismos individuos. De esta manera se mide uno de los conceptos incluidos en la fiabilidad: la estabilidad del test.

Modelo de medida: Se refiere a las escalas y subescalas de un cuestionario y a los procedimientos seguidos para crear las puntuaciones.

Muestra: Fracción seleccionada de una población determinada, idealmente representativa de la población de referencia y de tamaño suficiente.

Muestreo: Selección de los sujetos de una población que tomarán parte en un estudio. El muestreo aleatorio implica que cada participante potencial de la población estudiada tiene las mismas probabilidades de ser incluido en la muestra.

NEAR: Operador booleano utilizado en las búsquedas en Librería Cochrane Plus. Buscará documentos donde las dos palabras que enlaza NEAR estén separadas entre sí por un máximo de seis palabras en cualquier orden.

NEXT: Operador booleano utilizado en las búsquedas en Librería Cochrane Plus. Recupera documentos donde las dos palabras están juntas en ese orden.

Nivel de confianza: Se expresa en términos probabilísticos y se interpreta como la probabilidad de que el verdadero parámetro poblacional esté dentro del intervalo de confianza. Ver: Intervalo de confianza.

Normograma de Fagan: Método de cálculo de la probabilidad posttest mediante la representación gráfica de una tabla. Permite conocer la probabilidad posttest mediante el trazo de una línea que una la probabilidad pretest y la razón de probabilidad, el siguiente punto de intersección con la línea situada a la derecha, señala la probabilidad posttest.

Número necesario para dañar (number needed to harm, NNH): Número de pacientes que deben tomar el tratamiento experimental para que se produzca un caso adicional de un determinado acontecimiento adverso, en comparación con los pacientes que reciben el tratamiento de referencia.

Número necesario para tratar (NNT): Es una medida de la eficacia de un tratamiento. Cuando el tratamiento experimental incrementa la probabilidad de un acontecimiento favorable (o cuando disminuye la de un acontecimiento adverso), número de pacientes que hay que tratar para dar lugar a un paciente más con mejoría (o para prevenir un acontecimiento adverso adicional). Se calcula como $1/\text{RAR}$.

Odds: Se llama odds a la razón en la que el numerador es la probabilidad de que ocurra un suceso y el denominador es la probabilidad de que tal suceso no ocurra.

Odds post-test: Equivale a los odds (proporción) de que un paciente tenga la enfermedad una vez que se ha realizado el test diagnóstico.

Odds pre-test: Definida como los odds (proporción) de que un paciente tenga la enfermedad antes de que el test diagnóstico se efectúe.

Odds ratio (OR): Resultado de dividir la odds del grupo de interés (expuesto) por la odds del grupo de referencia (control). Si el $\text{OR} = 1$ el riesgo o probabilidad de que ocurra el evento es el mismo para ambos grupos (el efecto del tratamiento experimental es el mismo que el del tratamiento control); si el $\text{OR} > 1$ se interpretará como que el tratamiento experimental favorece la aparición del evento y si el $\text{OR} < 1$ diremos que protege de la aparición del mismo.

Operadores booleanos: Partículas habitualmente expresadas en inglés (AND, OR, NOT) que indican operaciones lógicas entre conceptos. Las operaciones son las de intersección (AND = recupera los documentos que contengan ambos términos), adición (OR = recupera los documentos que contengan tanto un término como otro, o la suma de los dos) o exclusión (NOT = sirve para excluir documentos que tienen un determinado término).

P: Seguía de la abreviatura n.s. (no significativa) o del símbolo $<$ o una cifra decimal, indica la probabilidad de que la diferencia observada en una muestra haya ocurrido puramente por azar siendo los grupos comparados realmente semejantes, es decir bajo la hipótesis nula.



Palabra clave: Término propio del lenguaje natural, sin otras normas que las del habla de la cual procede, que puede utilizarse en la búsqueda de información. Con frecuencia se utiliza como sinónimo de Descriptor, pero este uso no es correcto y conviene evitarlo. En MeSH las palabras clave son el equivalente a los términos de entrada (Entry Terms) de un diccionario de idiomas. Sin: Key word.

Palabra de texto: PubMed considera “palabras de texto” a todos los números y términos con significado que aparecen, entre otros menos importantes, en los campos título (title [TI]), resumen (abstract [AB]), descriptores (MeSH terms [MH]), subencabezamientos (MeSH Subheadings [SH]), así como los nombres de sustancias químicas (Substance Name [NM]) y los nombres de personas (Personal Name [PS]) cuando unas y otras son la materia del artículo. Sin: Text Word.

Panel de expertos: Reunión de un grupo de profesionales de prestigio en una determinada materia que abordan, exponen, contrastan ... sus opiniones y experiencias sobre la actuación en un tema preocupante de interés para el que se busca la mayor uniformidad de criterios que no es posible obtener por otro tipo de estudios más eficientes.

Perdidas de seguimiento: Aquellos pacientes que son perdidos durante el estudio y en los cuales no podemos valorar los resultados finales.

Población: Grupo individuos, objetos, mediciones, etc., con determinadas características demográficas, de donde se obtiene la muestra y a la que se quiere hacer inferencia de los resultados de un estudio.

Positivismo lógico: También denominado neopositivismo. Movimiento filosófico contemporáneo que destaca en filosofía la importancia del análisis del lenguaje y de la metodología científica.

Pretest cognitivo: Conjunto de estrategias y procedimientos que se utilizan para determinar si el cuestionario funciona tal y como lo ha concebido el investigador. Debe reproducir exactamente las condiciones en las que se responderá el cuestionario definitivo y, por tanto, tiene que realizarse con el cuestionario terminado y empleando el mismo procedimiento de recogida de información, puesto que la única diferencia entre el pretest y el trabajo de campo es el número de cuestionarios contestados.

Prevalencia: Número de casos de una enfermedad en una población y en un momento dados.

Prevalencia del test: Número total de positivos obtenidos al realizar una prueba diagnóstica en relación al total de sujetos que componen la muestra estudiada.

Probabilidad: Es el cociente entre los casos posibles de que ocurra un hecho dividido por el total de casos.

Probabilidad pretest o preprueba: Proporción de pacientes que tienen la enfermedad en la muestra estudiada, y cuyo diagnóstico es fehacientemente establecido por un estándar de referencia. Es la prevalencia de la enfermedad.

Probabilidad postest o posprueba: Probabilidad de tener una enfermedad tras la realización de una prueba diagnóstica.

PubMed: Base de datos de bibliografía sobre ciencias de la salud. Puede obtenerse de ella información acerca de publicaciones de carácter científico y de contenido médico (en su sentido más amplio, incluyendo enfermería, odontología, veterinaria...).

Randomización: Ver Aleatorización.

Razón de probabilidad: Probabilidad de un resultado en los enfermos dividido por la probabilidad de ese mismo resultado en no enfermos. Sin: Cociente de probabilidad, Likelihood Ratio.

Reacciones adversas a medicamentos (RAM): Según la OMS, “reacción nociva y no deseada que se presenta tras la administración de un fármaco, a dosis utilizadas habitualmente en la especie humana, para prevenir, diagnosticar o tratar una enfermedad, o para modificar cualquier función biológica”. Las reacciones adversas se clasifican en dos grandes grupos: RAM tipo A, de tipo cuantitativo o predecibles, y RAM tipo B, de tipo cualitativo o no predecibles. Sin: Efecto indeseado.

Record: Ver Asiento.

Reducción absoluta del riesgo (RAR): Diferencia aritmética absoluta entre la tasa de acontecimientos (riesgo) en el grupo experimental y la tasa de acontecimientos (riesgo) en el grupo control. También se puede denominar: Diferencia de riesgos.

Reducción relativa del riesgo (RRR): Cociente entre la reducción absoluta del riesgo (RAR) y el riesgo en el grupo control. A la reducción relativa del riesgo se le llama también “fracción atribuible” o “diferencia relativa del riesgo”.

Registro: Ver Asiento.

Regresión logística: Tipo de análisis de regresión (véase) en el que la variable dependiente sólo puede tomar dos valores. En el análisis de regresión la variable dependiente puede ser o bien categórica o bien continua (véase).

Repetibilidad: Referida a una prueba, es el grado en que los resultados son idénticos o estrechamente semejantes cada vez que aquella se lleva a cabo por los mismos profesionales y los mismos emplazamientos.

Reproducibilidad: Grado en que los resultados son idénticos o estrechamente semejantes cada vez que una prueba se lleva a cabo por distintos profesionales y es distintos ámbitos. Sin: Estabilidad.

Retroalimentación: Método para mantener la acción o eficacia de un sistema mediante la continua revisión de los elementos del proceso y de sus resultados con el fin de realizar las modificaciones necesarias.

Reunión educacional: Visita individualizada de un educador a un profesional de la salud en su ámbito de trabajo; en ella se suelen combinar técnicas de auditoría y retroalimentación.



Revisión: Se denomina así a todo intento que sintetice los resultados y las conclusiones de dos o más publicaciones relacionadas con un tema dado.

Revisión bibliográfica: recopilación de literatura realizada sobre un tema concreto en un intervalo de tiempo concreto del que se realiza un resumen y en el que no quedan explícitos los criterios de búsqueda y se destacan las conclusiones a criterio del autor.

Revisiónes narrativas: Se basan en el análisis que hace un experto de los estudios originales existentes sobre un determinado tema o problema para posteriormente subrayar aquellos resultados que considera más importantes.

Revisiónes sistemáticas: Revisión en las que la evidencia sobre un tema ha sido sistemáticamente identificada, criticada y resumida de acuerdo a unos criterios predeterminados.

Riesgo: Resultado predecible o incierto desfavorable de una actividad, intervención o exposición. Probabilidad de que ocurra un determinado fenómeno adverso.

Riesgo relativo (RR): Cociente entre el riesgo de padecer un acontecimiento entre los individuos expuestos a un factor de riesgo determinado y el riesgo de padecerlo entre los no expuestos. También se puede denominar: Cociente de riesgos.

Seguimiento: Periodo de duración de un estudio durante el cual se mantienen contactos periódicos con los participantes en el mismo con el fin de evaluar las variables de estudio.

Sensibilidad: Referida a una prueba diagnóstica, es la proporción de personas verdaderamente enfermas que ha sido catalogadas como tales mediante dicha prueba.

Sensibilidad al cambio (de un instrumento): se refiere a la medida en que detecta y mide de manera fiable las variaciones más pequeñas de un atributo, rasgo, dimensión o constructo.

Serie de casos: Tipo de estudio que describe una serie de características poco conocidas de un proceso o tratamiento y facilitan el conocimiento del mismo y su evolución en el tiempo.

Sesgo: Error sistemático producido en la estimación de los resultados, a causa de defectos en el diseño del estudio, recogida de datos o seguimiento.

Sesgo de desgaste: Diferencias sistemáticas que se producen entre los grupos de un ensayo clínico por la pérdida de participantes durante el estudio.

Sesgo de información: Sesgo relacionado con la forma en la que se obtuvieron los datos. Entre ellos se incluyen los errores cometidos en la valoración del estado de los pacientes al inicio, durante o al final del estudio; también los que ocurren cuando los datos se obtienen de fuentes sin que se conozca la calidad de las mismas. Algunos de los sesgos de información serían: el sesgo del recuerdo o de memoria, sesgo de sospecha diagnóstica, sesgo del entrevistador...

Sesgo de interpretación de las pruebas: Error sistemático que surge debido a la realización de las pruebas de forma no independiente y a ciegas con la consecuencia de que el conocer los resultados de una prueba influye en los de la otra.

Sesgo de publicación: Sesgo que surge debido a la tendencia a publicar sólo estudios con resultados positivos.

Sesgo de selección: Error sistemático consecuencia de que las características de los sujetos incluidos en un estudio son diferentes de las características de los no incluidos, de modo que la muestra no es representativa de la población de referencia. Se incluyen en este grupo el sesgo de los voluntarios, el de prevalencia-incidencia, el de los emigrantes, el de selección no aleatorizada, el de Berkson, el de pertenencia a grupo, el del trabajador sano...

Sesgo de sospecha diagnóstica: Sesgo que ocurre cuando la interpretación de una prueba se haga conociendo el resultado de otra.

Sesgo de reproducibilidad: Sesgo debido a la ausencia de información detallada de las condiciones en que se ha aplicado la prueba, así como las definiciones de los resultados positivos y negativos.

Sesgo de resultados no interpretables: Error sistemático consecuencia de no incluir pacientes en diferentes estadios de la enfermedad.

Sesgo de verificación o de confirmación diagnóstica: Sesgo que ocurre cuando el estudio se limita a individuos sobre los que se realizó con anterioridad la prueba "gold Standard", que suelen ser los de mayor probabilidad de padecer enfermedad.

Sesgo en el espectro de la enfermedad: ¿

Sesgo por ausencia de "gold Standard": Error sistemático que se produce cuando entre los criterios de referencia que se emplean en la comparación de la utilidad de una prueba diagnóstica (en ausencia de "gold Standard") se incluye la prueba de estudio.

Sesión bibliográfica: Herramienta docente consistente en reuniones entre profesionales que realizan una lectura crítica de un artículo científico.

Significación clínica: Probabilidad de que una diferencia observada tenga una repercusión sobre el curso del problema o enfermedad tratados que sea relevante para un paciente dado o para un conjunto de pacientes.

Significación estadística: Probabilidad de que una diferencia observada sea resultado de la casualidad y no de los determinantes causales en un estudio. El hallazgo de una significación estadística no implica necesariamente significación clínica.

Síntesis: Fuentes de información que recogen resúmenes estructurados de revisiones sistemáticas de alta calidad o de artículos originales.

Síntesis: Fuentes de información que incluyen revisiones sistemáticas de alta calidad metodológica.



Sistema: Fuente de información que se caracteriza por abarcar una gran variedad de temas, proporcionar un resumen de los resultados de revisiones de alta calidad y ser actualizado con frecuencia.

Subencabezamiento: Término que delimita el significado del descriptor al que acompaña. P.e. “Drug therapy” (tratamiento farmacológico) como subencabezamiento unido al descriptor Hypertension nos informaría que un artículo trata, entre todos los aspectos posibles de la hipertensión, su tratamiento farmacológico. Sin: Subheading.

Subheading: Ver subencabezamiento.

Tag: Ver Etiqueta.

Tamaño de la muestra: Número mínimo necesario de sujetos que participan en un estudio, pero a su vez suficiente para proporcionar una respuesta fiable al objetivo final de dicho estudio. La determinación de dicho tamaño mínimo es de naturaleza estadística y depende de varios factores.

Técnica de grupos conocidos: Procedimiento utilizado para la validación de constructo según la cual el instrumento se aplica a los grupos que, según se presume, diferirán respecto del atributo crítico a causa de una característica conocida.

Tecnología sanitaria: Conjunto de medicamentos, aparatos, procedimientos médicos y quirúrgicos usados en la atención médicos y los sistemas organizativos y de soporte con los que se presta la atención sanitaria.

Teoría estadística de Bayes o Teorema de Bayes: Establecimiento inductiva de la probabilidad de una hipótesis a partir de algunas observaciones. Según el teorema de Bayes, la probabilidad posterior (o a posteriori) de una hipótesis es igual a una constante, multiplicada por la verosimilitud o posibilidad de la hipótesis (definida como la probabilidad de observar unos acontecimientos determinados si la hipótesis considerada fuera de hecho cierta) y por la probabilidad anterior de la hipótesis.

Término MeSH: Ver MeSH Term

Término de entrada: Relacionado con Palabra clave (véase).

Tesaurus: Es un repertorio de términos con expresión de las relaciones semánticas que los unen. Habitualmente consta de dos partes, una sección categorizada, que expresa las relaciones jerárquicas, genérico-específicas, entre los términos normalizados (descriptores), lo que nos permite conocer el campo semántico (significado) que abarcan los distintos términos. La sección alfabética es una ordenación de todos los términos, normalizados (descriptores) o no normalizados (palabras clave), que de forma similar a un diccionario de idiomas, permite conocer las relaciones de equivalencia entre unos y otros términos. Sin: Thesaurus.

Test Chi cuadrado (2): Prueba estadística utilizada para determinar el grado de asociación entre variables categóricas (véase).

Test de McNemar: Prueba de significación estadística para probar la hipótesis nula de inexistencia de cambios en la proporción de sujetos que experimentan un acontecimiento, cuando cada individuo es evaluado dos veces (en condiciones diferentes) y los datos están emparejados.

Test T de Student: Prueba de significación estadística paramétrica (véase) para contrastar la hipótesis nula respecto a la diferencia entre dos medias.

Test U de Mann-Whitney: Prueba de significación estadística no paramétrica (véase) para probar la hipótesis nula de que el parámetro de localización (generalmente la mediana) es el mismo cuando se comparan dos grupos independientes, cualquiera que sea el tipo de distribución de la variable (distribución normal o de otro tipo).

Test de Wilcoxon: El contraste de Wilcoxon es la técnica no paramétrica paralela al test T de Student para muestras apareadas.

Text Word: Ver Palabra de texto.

Thesaurus: Ver Tesoro.

Transferabilidad: Término acuñado en investigación cualitativa sugiere que en la medida en que el investigador ha respetado el criterio de maximizar la diversidad en la obtención de la muestra y realice una detallada descripción del contexto y participantes, los hallazgos pueden ser aplicables en contextos similares.

Triangulación: Técnica empleada en investigación cualitativa para controlar sesgos. Consiste en obtener distintas perspectivas del fenómeno estudiado utilizando diferentes métodos de investigación (cualitativos y cuantitativos). La triangulación puede ser simultánea o secuencial.

Truncado: El término truncar significa en español “cortar el extremo de una cosa”. Aplicado a un término de búsqueda consiste en separar del término su terminación y sustituirla por un símbolo comodín (en PubMed es un asterisco *-). El resultado es la selección de todos los términos que contienen la raíz superviviente.

Umbral de decisión: Nivel de probabilidad a partir del cual el beneficio previsible de toma una decisión de intervenir iguala o mejora al beneficio de no tomarla

Validez: Capacidad de un instrumento para medir aquello para lo que ha sido diseñado, denota el grado en que el instrumento mide lo que se supone que debe medir.

Validez aparente o facial: Forma especial de la validez de contenido que se refiere al hecho de que el instrumento parezca, de la impresión a las personas evaluadas, de que es adecuado para medir lo que se pretende medir.

Validez concurrente: Denota la capacidad de un instrumento de distinguir a los individuos que difieren en su condición actual respecto de un criterio.

Validez convergente: Aspecto de la validez de constructo que se refiere al grado en que diferentes formas de medir el mismo constructo (o dimensiones de un constructo) se correlacionan la una con la otra.



Validez de constructo: Grado en que una prueba o instrumento mide un determinado rasgo, característica o constructo.

Validez de contenido: Determina si el instrumento contiene ítems representativos de todas las dimensiones que forman la definición del concepto o constructo en estudio.

Validez de criterio: Concordancia entre los resultados de la medida del instrumento y los resultados de la medida realizada mediante un criterio o patrón de referencia (gold standard) de lo que se quiere medir.

Validez discriminante o divergente: Aspecto de la validez de constructo que demuestra que una medida no se correlaciona con otras medidas cuya finalidad es medir rasgos, dimensiones o constructos diferentes a los que ella mide.

Validez externa: Grado en que las conclusiones obtenidas con la muestra de población que participa en un estudio pueden ser generalizadas a su población de referencia o a otras poblaciones, lugares, momentos e investigaciones.

Validez interna: Grado en que los resultados de un estudio miden la realidad para la muestra de población que participa en un estudio. Grado en que un estudio está libre de sesgos.

Validez predictiva: Se refiere a la idoneidad de un instrumento para diferenciar entre el desempeño o las conductas de los individuos respecto de un criterio futuro.

Valor global de un test: Proporción de sujetos correctamente clasificados por el test en relación al total de sujetos que componen la muestra estudiada.

Valor predictivo negativo: Referido a una prueba diagnóstica, la probabilidad de que una persona con un resultado negativo sea un verdadero negativo.

Valor predictivo positivo: Referido a una prueba diagnóstica, la probabilidad de que una persona con un resultado positivo sea un verdadero positivo.

Variabilidad: Variación entre individuos debido a determinantes genéticos, ambientales, patológicos... (variabilidad interindividual) u oscilación de variables biológicas en un mismo individuo según las circunstancias (variabilidad intraindividual).

Variable: Cualquier atributo, fenómeno o hecho que pueda tener diferentes valores.

Variable categórica: Se dice de la variable que presenta interrupciones. Una variable es categórica si, entre varios valores potencialmente observables, hay un valor que no puede ser observado. Ejemplo: variables binarias.

Variable continua: Se dice de la variable que no presenta interrupciones. Una variable es continua si, entre dos valores observables, siempre existe la posibilidad de que haya otro valor observable. Ejemplos: edad, peso, altura.

Verbatim: Su significado es: Literalmente, o palabra por palabra. Anglicismo utilizado en investigación cualitativa para denominar a la transcripción exacta, en la fase de exposición de resultados, de las expresiones surgidas durante el grupo de discusión de los participantes, es decir conservado la expresión literal grabada.

Verdaderos negativos: Grupo de pacientes sanos en los que el test es negativo.

Verdaderos positivos: Grupo de pacientes con enfermedad y que además son detectados por el test.

