



Actividad Dietética

Act Diet. 2008;12(1):11-9

www.elsevier.es/dietetica



Originales

Sistema de clasificación de los estudios en función de la evidencia científica. Dietética y Nutrición Aplicada Basadas en la Evidencia (DNABE): una herramienta para el dietista-nutricionista del futuro

Eduard Baladia^a y Julio Basulto^{a,b}

^aGrupo de Revisión, Estudio y Posicionamiento de la AED-N. Barcelona. España.

^bDepartamento de Bioquímica y Biotecnología. Facultad de Medicina y Ciencias de la Salud. Universidad Rovira i Virgili. Reus. Tarragona. España.

RESUMEN

Los dietistas-nutricionistas, que trabajan (como el resto de los profesionales sanitarios) según las recomendaciones extrapoladas de estudios científicos, necesitan hacerlo conociendo el grado de confianza que pueden depositar en dichas recomendaciones. La enorme cantidad de datos publicados sobre dietética y nutrición ha creado la necesidad de resumir los hallazgos y obtener conclusiones basadas en la mejor evidencia empírica disponible.

En este sentido, es prioritario desarrollar un sistema práctico de clasificación de los niveles de evidencia para la dietética y la nutrición. Dicho sistema debe ser la herramienta que permita dar solidez a las recomendaciones y prescripciones de los dietistas-nutricionistas. La dietética y la nutrición aplicada basadas en la evidencia es, por tanto, el instrumento a utilizar, según datos racionales u objetivos, a la hora de desarrollar y revisar guías, validar alegaciones nutricionales y de salud, establecer cualquier tipo de recomendación o tomar decisiones preventivas y terapéuticas.

© 2008 AED-N. Todos los derechos reservados.

Historia del artículo:
Recibido 11-4-2008.
Aceptado 11-4-2008.

Palabras clave:
Dietética.
Nutrición.
Evidencia.
Metodología.
Herramienta.
Dietista-nutricionista.

ABSTRACT

Systematics for classifying studies according to scientific evidence. Evidence-Based Dietetics and Applied Nutrition (EBDAN): a tool for the future dietician-nutritionist

Like other health professionals, dieticians-nutritionists work on the basis of recommendations extrapolated from scientific studies and need to be aware of the degree of confidence they can place in these recommendations. The enormous amount of data published on dietetics and nutrition has created the need to summarize the findings and to draw conclusions based on the best empirical evidence available. In this sense, developing a practical system to classify the levels of evidence for dietetics and nutrition is a priority. Such a system should allow the soundness of dieticians-nutritionists' recommendations to be evaluated. Evidence-Based Dietetics and Applied Nutrition (EBDAN) is, therefore, the instrument to be used, based on rational data or objectives, when developing and revising guidelines, validating nutritional and health statements, establishing any type of recommendation, or taking preventive and therapeutic decisions.

© 2008 AED-N. All rights reserved.

Correspondencia: E. Baladia.
Consell de Cent, 314, pral. B.
08007 Barcelona. España.
Correo electrónico: info@grep-aedn.es

Introducción y antecedentes

En la práctica profesional diaria del dietista-nutricionista, es frecuente enfrentarse a un sinfín de preguntas acerca del uso de las recomendaciones dietético-nutricionales, a las cuales debe hallarse respuesta. Pese a que aún es habitual entre los profesionales sanitarios recurrir a la experiencia personal, o bien acudir a los libros de texto para solucionar dichas incógnitas^{1,2}, cabe destacar que ésta podría no ser la forma más adecuada.

Esta forma de proceder no ha sido única en la historia. Tal y como explica Doherty³, en 1937, el joven médico Sir Richard Doll observó que en Londres los nuevos tratamientos surgían de la propia experiencia de los especialistas, sin tener en cuenta los factores de confusión ni los posibles sesgos⁴.

De esta forma, durante años la medicina fue incorporando la necesidad de crear la así llamada «medicina basada en la evidencia» para poder ofrecer la mejor práctica posible, hasta que en 1993 se publicaron, en el *Journal of the American Medical Association*, dos de los primeros documentos en que se usa el concepto de medicina basada en la evidencia⁵⁻⁷. Sin embargo, no es hasta 1995 cuando se define de forma explícita este concepto: «la forma sistemática de analizar los estudios científicos publicados para ser la base de los árboles de decisión clínica»⁸. Posteriormente, en 1996, Sacket et al⁹ redefinieron el nuevo término como «el uso concienzudo y juicioso de la mejor evidencia científica actual derivada de la investigación en el manejo de pacientes individuales en atención clínica». En 1997 el National Center for Biotechnology Information (NCBI), basado en un artículo publicado en el *British Medical Journal*¹⁰, la introduce como «MeSH Term» —MeSH son las siglas de *Medical Subject Headings* (encabezados de materia médica) designados por National Library of Medicine de Estados Unidos— y la define como: «El proceso de encontrar, evaluar y usar, de manera sistemática, los hallazgos contemporáneos relativos a la investigación como la base en la toma de decisiones clínicas». Asimismo, se indica que la medicina basada en la evidencia sigue cuatro pasos: formulación de una cuestión clínica clara a partir de un problema médico, indagar en la literatura para hallar artículos clínicos relevantes, evaluar (valoración clínica) la evidencia desde el punto de vista de su validez y utilidad e implementar los hallazgos útiles en la práctica clínica¹⁰.

Como se puede apreciar, pese a que la medicina basada en la evidencia es un concepto bastante nuevo, el concepto sobre el cual se ha construido (los niveles de evidencia) es más antiguo. Ya en 1979, la Canadian Task Force on the Periodic Health Examination¹¹ elaboró una jerarquía para clasificar la evidencia. Dicha jerarquía consistía en la clasificación de los estudios científicos en 5 niveles de evidencia según el tipo de estudio realizado:

Tabla 1
Sistema de clasificación de los niveles de evidencia¹¹ y grado o solidez de las recomendaciones correspondientes¹⁴

Jerarquía para clasificar la evidencia	
Nivel de evidencia	Nivel de solidez
I	A. Hay buena evidencia para recomendar que la condición sea específicamente considerada en un examen periódico de salud
II.1	B. Hay considerable evidencia para recomendar que la condición sea específicamente considerada en un examen periódico de salud
II.2	C. No hay suficiente evidencia para recomendar a favor o en contra de la inclusión de la condición en un examen periódico de salud, pero se puede hacer recomendaciones en otros campos
II.3	D. Hay considerable evidencia para recomendar que la condición sea excluida de un examen periódico de salud
III	E. Hay buena evidencia para recomendar que la condición sea excluida de un examen periódico de salud

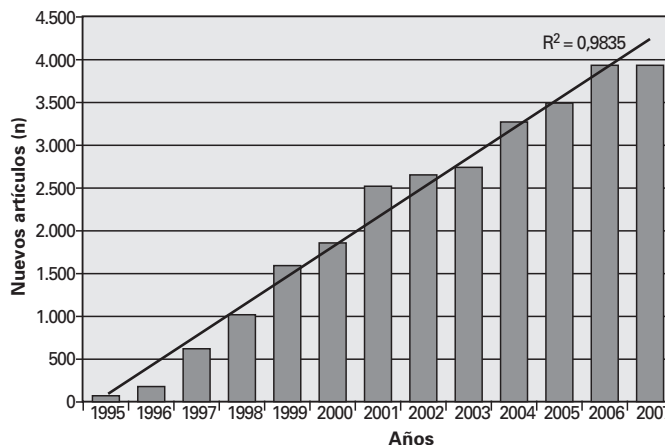


Fig. 1. Resultado de la introducción de «evidence based medicine» en PubMed por años. Nuevos artículos acerca de la medicina basada en la evidencia por año (elaboración propia).

- I. Ensayo clínico controlado, aleatorizado.
- II.1. Ensayos clínicos controlados (sin aleatorización).
- II.2. Estudios de cohortes o estudios analíticos de caso-control.
- II.3. Estudios de series de tiempo múltiples con o sin intervención.
- III. Opiniones de expertos, experiencias clínicas, estudios descriptivos u opiniones de comités de expertos.

Posteriormente se han publicado otras jerarquías^{12,13}, todas ellas basadas en una clasificación de los diferentes tipos de estudios según el nivel de evidencia, y se establecieron los grados de recomendación correspondientes. En 1996, la US Preventive Task Force, como se aprecia en la tabla 1¹⁴, añadió a los niveles de evidencia de la Canadian Task Force on the Periodic Health Examination¹¹ un sistema de clasificación de la solidez de las recomendaciones que dichas evidencias proporcionan.

En una revisión de 1.602 artículos realizada por la Agency for Healthcare Research and Quality de la US Department of Health and Human Services en 2002, se detectó el uso de hasta 120 sistemas diferentes de clasificación de la evidencia científica y de calidad de los estudios¹⁵.

De esta forma, paralelamente al desarrollo de los niveles de evidencia, ha ido creciendo la medicina basada en la evidencia. En la figura 1 se observa el crecimiento lineal ($R^2 = 0,9835$) de los nuevos estudios publicados cada año en PubMed en torno a la medicina basada en la evidencia.

De la misma forma que la medicina basada en la evidencia ha evolucionado y actualmente determina la práctica clínica, la dietética y la nutrición aplicada basadas en la evidencia (DNABE) debe hacerlo también, y dictar cuál es la mejor práctica dietético-nutricional.

Existen iniciativas cuyo objetivo es ayudar a los dietistas-nutricionistas a desarrollar e implementar guías basadas en la evidencia. Es el caso de la iniciativa del departamento de dietética de la División de Rehabilitación y Nutrición del Centro Médico Universitario Utrecht NL¹⁶, en 1998, o la iniciativa de la American Dietetic Association (ADA)¹⁷ de poner a disposición de sus socios, en 2006, una librería de análisis de la evidencia desde donde se puede acceder a las guías más importantes de la práctica dietética basada en la evidencia. En cualquier caso, no se ha hallado ningún estudio que describa, de forma detallada, un método práctico de sistematizar la clasificación de los niveles de evidencia y la solidez de sus recomendaciones en el campo de la dietética y la nutrición humana aplicada.

Es, por tanto, la finalidad del presente documento iniciar el proceso de creación de la DNABE en España, y ofrecer una metodología práctica para que los dietistas-nutricionistas, así como cualquier

otro profesional sanitario cualificado, puedan ofrecer recomendaciones actuales, útiles, válidas y sólidas en el campo de la dietética y la nutrición humana aplicada.

Sistema para establecer la solidez de un conjunto de artículos que conformen la evidencia científica

Uno de los mayores problemas actuales es que hay muchos estudios, y de muchos tipos, que nos pueden ofrecer diferentes respuestas o incluso contradictorias. Esto puede hacer difícil al profesional discernir cuál es «la verdad» o en qué medida una recomendación se «acerca a la verdad».

En este sentido, cuando hay un conjunto de artículos de los que se quiere conocer a cuál o cuáles de ellos se debe tener en mayor consideración para realizar una recomendación, hay que tener en cuenta 3 aspectos principales: el tipo de diseño de estudio, la calidad individual de los estudios¹⁸ y la cantidad relativa de muestra de población¹⁹, así como la duración del estudio. Estos aspectos se detallan a continuación.

Importancia del diseño del estudio

Para clasificar los estudios dentro de una jerarquía de niveles de evidencia y grados de recomendación, es necesario determinar el diseño del estudio. Según su diseño podríamos diferenciar cuatro grandes grupos de estudios: experimentales, observacionales, de revisión y de consenso o posicionamiento.

Los estudios experimentales se caracterizan por tener un grupo llamado «intervención», que se somete a un tratamiento, mientras que los estudios observacionales se caracterizan por observar y examinar la relación entre los factores específicos de un grupo de población y los factores o marcadores de salud o enfermedad²⁰. Por otro lado, los estudios de revisión, consenso y posicionamiento pueden ser muy heterogéneos, ya que pueden abarcar desde revisiones de la literatura de cualquier tipo (libros, artículos de revistas científicas, páginas de internet, etc.) hasta revisiones de artículos científicos de un diseño específico (por ejemplo, de ensayos clínicos controlados, aleatorizados). Entre los estudios experimentales, se puede distinguir²¹:

1. Ensayos clínicos controlados, aleatorizados (ECCA, o sus siglas inglesas RCT: *randomized controlled trial*): trabajos consistentes en un ensayo clínico que involucre, como mínimo, un grupo de intervención y un grupo control, asignación simultánea, y seguimiento de ambos grupos. Asimismo, los tratamientos a administrar se seleccionan por un proceso aleatorio.

2. Ensayos clínicos controlados (no aleatorizados): trabajos consistentes en ensayos clínicos que involucren, como mínimo, un grupo de intervención y un grupo control, y que especifiquen las medidas de los parámetros para la intervención estudiada, mediante un método libre de sesgos para asignar a los pacientes al tratamiento. Los tratamientos podrían ser fármacos, aparatos o procedimientos estudiados para valorar la efectividad del diagnóstico, la terapia o la profilaxis. Las medidas de control incluyen placebos, sustancia activa, no tratar, dosis utilizada y regímenes, comparaciones históricas, etc.

3. Ensayos clínicos (no controlados): trabajos consistentes en estudios clínicos prediseñados acerca de la seguridad, la eficacia o la planificación de la dosis óptima de uno o más fármacos, instrumentos o técnicas para realizar diagnósticos, terapias o profilaxis en humanos seleccionados en función de criterios predeterminados de elegibilidad y observados por evidencias previamente determinadas de efectos favorables o no favorables.

Entre los estudios observacionales²¹:

1. Estudios de cohortes: estudios en los que se identifican subgrupos de población. Estos grupos pueden ser o no expuestos a fac-

tores que hipotéticamente podrían influir en la probabilidad de que ocurra una enfermedad en concreto u otro acontecimiento. Las cohortes son poblaciones definidas que, como grupo, se siguen con el objetivo de determinar características distintivas en subgrupos.

2. Estudios de casos-control: estudios que se inician con la identificación de personas con una enfermedad de interés y con un grupo control (comparación o de referencia) sin la enfermedad. Se examina la relación de un atributo con la enfermedad comparando a las personas enfermas y las no enfermas en relación con la frecuencia o los niveles del atributo en cada grupo.

3. Estudios transversales: estudios en que se valora la presencia o ausencia de una enfermedad, o de otras variables determinadas, en cada miembro de la población de estudio o en una muestra representativa de dicha población, en un momento dado.

Entre los estudios de revisión²¹:

1. Metaanálisis: trabajos que consisten en la valoración de estudios con un método cuantitativo mediante la combinación de resultados de estudios independientes (normalmente extraídos de lo publicado en la literatura científica) y sumarios sintetizados y conclusiones que podrían utilizarse para evaluar la efectividad terapéutica, planificar nuevos estudios, etc.

2. Revisión sistemática²²: una revisión de una cuestión claramente formulada, que utiliza métodos sistemáticos y explícitos para identificar, seleccionar y criticar estudios relevantes, y para recoger y analizar información incluida en la revisión. Se puede o no usar métodos estadísticos (metaanálisis) para analizar o sintetizar los resultados de los estudios que se hayan incluido.

3. Revisión: un artículo o libro publicado tras el examen de material publicado acerca de un tema. Puede ser exhaustiva, a varios niveles, y el intervalo de tiempo seleccionado para captar estudios puede ser amplio o estrecho, pero las revisiones más solicitadas son las que evalúan la literatura actual.

No todos los estudios tienen la misma finalidad. Por ejemplo, se considera que los ECCA son los estudios que estiman «la verdad» de forma más precisa²³, mientras que los estudios de casos clínicos, se considera que son más útiles para generar hipótesis²⁴. Entre estos dos tipos de diseño están los estudios de cohortes y los estudios de casos-control. Los primeros ayudan a crear una relación entre un factor de riesgo y una consecuencia potencial, mientras que los segundos a menudo son utilizados para identificar los factores que pueden pronosticar una posible consecuencia^{23,24}. Los estudios de casos-control son más susceptibles a tener algún tipo de sesgo que los estudios de cohortes²⁵.

Por un lado, la US Agency for Health Care Policy and Research (AHCPR) estableció en 1996 que, en los casos en que no es posible realizar ECCA, o bien éstos no son capaces de proporcionar la respuesta adecuada, se puede igualar el valor científico de estudios observacionales y de ECCA. La jerarquía propuesta por esta agencia está basada en 7 niveles de evidencia que se corresponden a 3 grados de recomendaciones diferentes¹⁹.

Por otro lado, investigadores del Instituto Alemán de Medicina Social, Epidemiología y Economía de la Salud y el Departamento de Epidemiología del Instituto Alemán de Nutrición Humana han puesto de manifiesto que cuando se realizan recomendaciones preventivas a la población con consejos para mejorar conductas, sin especificar cantidades (por ejemplo, recomendar el aumento del consumo de frutas y verduras sin establecer una cantidad bien definida), o eliminar conductas perniciosas (hábito de fumar o consumo de alcohol), no es necesario que se diseñe un ECCA para comprobarlo, sino que con estudios observacionales es suficiente para llegar a la conclusión²⁶.

Muchos autores han desarrollado sus propias jerarquías de niveles de evidencia. En este sentido, se ha utilizado la base de datos del

Tabla 2

Sistema de gradación de la evidencia para las recomendaciones en la práctica clínica de la American Diabetes Association²⁹

A	Evidencia clara a partir de ECCA bien diseñados, generalizables y con suficiente fuerza, como es el caso de: Evidencias a partir de un ensayo multicéntrico bien diseñado Evidencias a partir de un metaanálisis que incorpore un análisis de la calidad de los estudios evaluados Evidencias a partir de ECCA bien diseñados, y con suficiente fuerza, como es el caso de: Evidencias a partir de un ensayo bien diseñado en una o más instituciones Evidencias a partir de un metaanálisis que incorpore un análisis de la calidad de los estudios evaluados
B	Evidencia apoyada por estudios de cohortes bien diseñados: Evidencias a partir de estudios de cohortes prospectivos Evidencias a partir de metaanálisis bien diseñados de estudios de cohortes Evidencias a partir de estudios de caso-control bien diseñados
C	Evidencias a partir de estudios no controlados o mal controlados, como es el caso de: Evidencias a partir de ECCA con uno o más sesgos metodológicos importantes, o con tres o más sesgos metodológicos leves, que podrían invalidar los resultados Evidencias a partir de estudios observacionales con un alto potencial de sesgos (tales como series de casos con comparaciones con controles históricas) Evidencias a partir de series de casos Evidencias que están en conflicto con recomendaciones soportadas por un gran número de estudios que indican lo contrario que dichas evidencias
E	Consensos de expertos o experiencia clínica

ECCA: ensayo clínico controlado, aleatorizado.

National Guideline Clearinghouse (base de datos pública que ofrece acceso a guías de práctica clínica basada en la evidencia, creada por la Agency for Healthcare Research and Quality [AHRQ], US Department of Health and Human Services, y en su origen, por la American Medical Association y la American Association of Health Plans) para encontrar estudios que usen un sistema de jerarquía de la evidencia en el marco de la nutrición basada en la evidencia (*evidence based nutrition*; EBN). Debido a las limitaciones del buscador de esta base de datos, se ha utilizado Google como motor de búsqueda dentro de la misma base de datos (para más información al respecto, consultar Baladia et al²⁷, 2006). Únicamente se han hallado 2 documentos que cumplan con los requisitos establecidos: la «Guía práctica basada en la evidencia de los desórdenes en el metabolismo de los lípidos»²⁸ y los «Estándares en el cuidado médico en la diabetes»²⁹, de los cuales únicamente el segundo establece de forma clara una jerarquía de niveles de evidencia en sus recomendaciones (tabla 2).

Se han hallado, además, otras guías que, pese a no especificar que son de nutrición basada en la evidencia, tratan sobre el manejo dietético-nutricional en distintas enfermedades crónicas: obesidad³⁰, enfermedad cardiovascular³¹, osteoporosis³² y salud dental³³.

De esta forma, tomando como base únicamente el tipo de diseño de estudio, y según lo descrito en este apartado, se puede resumir y establecer un esquema general de jerarquía de los niveles de evidencia:

- Revisión sistemática con trato estadístico de ECCA.
- Metaanálisis de ECCA.
- Revisión sistemática de ECCA (sin trato estadístico).
- Revisiones de ECCA.
- ECCA independientes.
- Revisiones sistemáticas con trato estadístico de estudios de cohortes prospectivos.
- Metaanálisis de estudios de cohortes prospectivos.

– Ensayos clínicos controlados (sin aleatorización): dentro de éstos, y por este orden: revisión sistemática con trato estadístico, metaanálisis, revisión e independientes.

- Revisiones de estudios de cohortes prospectivos.
- Estudios de cohortes prospectivos independientes.
- Estudios de casos-control (control no histórico): dentro de éstos, y por este orden: revisión sistemática con trato estadístico, metaanálisis, revisión e independientes.
- Ensayos clínicos no controlados.
- Estudios de cohortes con control histórico.
- Estudios de casos-control (con control histórico).
- Estudios transversales.
- Estudios controvertidos: estudios de diseño parecido o bien muchos estudios de diseño inferior refutados por poca cantidad de estudios de diseño de calidad superior.
- Opiniones de expertos: en caso de que éstos aporten estudios, se clasificará según su diseño.
- Otros tipos de revisiones.
- Otros documentos.

A estos niveles de evidencia, les debe corresponder unos grados de recomendación o solidez de la recomendación:

- Recomendaciones cuya variación a lo largo del tiempo es altamente improbable.
- Recomendaciones cuya variación a lo largo del tiempo es improbable.
- Recomendaciones que es probable que varíen a lo largo del tiempo.
- Hay estudios limitados o no están sustentados en suficientes estudios. No se puede establecer recomendación.
- No se han hallado estudios al respecto.

Pese a que el diseño de un estudio es muy importante para definir la evidencia científica que nos aporta la afirmación que se baraja en su hipótesis, hay que tener en cuenta los otros dos aspectos para establecer una buena jerarquía de niveles de evidencia y grados de solidez de las recomendaciones¹⁵, descritos a continuación.

Sistema para estimar la calidad individual de los artículos

Aparte de establecer una jerarquía entre los distintos tipos de diseño de estudios, también es posible establecer un sistema jerárquico dentro de un mismo tipo de estudio para permitir que los estudios de mayor calidad estén por encima de los de peor calidad; de esta forma se obtiene una mayor fiabilidad de las recomendaciones que se desprendan de dichos estudios, y se pueden evitar posibles sesgos o errores de interpretación. Según el Center of Reviews and Dissemination (CRD) de la Universidad de York (Reino Unido)²⁵, es posible que los estudios de baja calidad tiendan a sobrestimar el efecto del tratamiento³⁴, o incluso pueden llegar a conclusiones contrarias a las halladas por estudios de alta calidad³⁵.

Hay muchos puntos que se podría considerar para definir la calidad de un artículo. Tal y como muestra la tabla 3, un documento, aceptado por la Agency for Healthcare Research and Quality de la US Department of Health and Human Services en 2002, considera entre 9 y 11 dominios (por cada tipo de diseño de estudio) que pueden ser revisados cuando se intenta definir la calidad de un artículo¹⁵. Se puede apreciar que se han destacado los 19 dominios que se consideraron «críticos» para establecer dicha calidad.

Según señalan los autores de ese documento, para definir la calidad de un artículo se pueden usar los 9-11 dominios o solamente los indicados como críticos. Asimismo, advierten que debe tenerse siempre en consideración que, según el sistema elegido para establecer la calidad, puede aumentar tanto la dificultad de uso que disminuya su practicidad y su precisión¹⁵.

Tabla 3
Dominios que pueden ser revisados para establecer la calidad individual de estudios¹⁵

Revisiones sistemáticas	Ensayos clínicos controlados, aleatorizados	Estudios observacionales
Pregunta de estudio*	Pregunta de estudio	Pregunta de estudio
Estrategia de búsqueda*	Población de estudio*	Población de estudio
Criterios de inclusión-exclusión*	Aleatorización*	Comparabilidad de los sujetos*
Intervención	Forma de cegamiento*	Exposición o intervención*
Parámetros	Intervención*	Medida de los parámetros*
Extracción de los datos*	Parámetros*	Análisis estadístico*
Calidad del estudio y validez*	Análisis estadístico*	Resultados
Síntesis de los datos y análisis*	Resultados	Discusión
Resultados	Discusión	Patrocinador*
Discusión	Patrocinador*	
Patrocinador*		

*Dominios considerados críticos.

Tabla 4
Test propuesto por Detsky et al³⁶ para evaluar la calidad de los estudios clínicos controlados, aleatorizados y puntuaciones establecidas

Preguntas	Puntuación		
	Sí	Parcialmente	No
¿Los pacientes fueron asignados de forma aleatoria?	1	1	0
¿La aleatorización estuvo adecuadamente descrita?	2		0
¿El grupo intervención estuvo cegado al investigador?	1		0
¿Hay una descripción adecuada de la medida de los parámetros?	1	1	0
¿Y del objetivo de la medida de los parámetros?	2		0
¿El tratamiento fue cegado al investigador?	1		0
¿Están bien definidos los criterios de inclusión/exclusión?	2	1	0
¿Y el número de pacientes excluidos, junto con la razón?	2	1	0
¿La terapia está enteramente descrita para el grupo de intervención?	2	1	0
¿La terapia está enteramente descrita para el grupo control?	2	1	0
¿Se declara el tipo de test? ¿Se hizo un análisis estadístico y se describieron los valores de p*?	1	1	0
¿El análisis estadístico fue apropiado?	2		0
Si el estudio fue negativo, ¿se llevó a cabo cálculos de los intervalos de confianza o cálculos de potencia post hoc?	1		0
¿Se calculó el tamaño de la muestra antes del estudio?	1		0
Puntuación máxima del test: 20 (estudio positivo) o 21 (estudio negativo)			

*El valor de p asociado a un resultado observado es la probabilidad de obtener un valor como el observado o más extremo si la hipótesis nula es cierta. Las calculadoras modernas con funciones estadísticas avanzadas y los programas estadísticos proporcionan el valor de p para facilitar la interpretación de los resultados a la hora de realizar inferencia.

Estudios experimentales

En el apartado «Importancia del diseño del estudio», se ha podido comprobar que los ECCA ocupan los primeros lugares en la jerarquía del tipo de diseño. Sin embargo, no todos los ECCA son de igual calidad o igual de fiables. Como se puede observar en la tabla 3, hay 7 dominios críticos que pueden ser revisados para determinar la calidad de los ECCA. Sin embargo, para mantener la practicidad, algunos autores han descrito que para definir la calidad de los ECCA es aceptable usar, como mínimo, los siguientes indicadores: la aleatorización, el análisis estadístico y la forma de cegamiento^{19,36-38}, es decir 3 de los 7 descritos como críticos en la tabla 3. Una revisión realizada en 1995 por Moher et al³⁸ identificó 25 escalas de puntuación diferentes para definir la calidad de ECCA. Los autores identificaron, de cada una de las escalas, una serie de características para poder evaluar su calidad. Entre otras, se describió:

1. Si podía medir ECCA de cualquier materia, o bien, estaba especializada, por ejemplo, en alguna patología.
2. Si se refería a la calidad metodológica o a la calidad total del artículo.

3. Si había como mínimo un ítem para identificar la asignación de los pacientes.

4. Si había como mínimo un ítem acerca del enmascaramiento.

5. Si había como mínimo un ítem acerca del seguimiento del paciente.

6. Si se hacía un análisis estadístico.

7. El número de ítems a responder.

8. El tiempo necesario para completar dicha escala.

Pese a que los autores concluyeron que únicamente la escala de Jadad-Bechara³⁹ superó todo el examen propuesto, cabe destacar que la Cochrane Collaboration indica que no existe el estándar en el uso de dichas escalas para establecer la calidad de los estudios⁴⁰, y que no están soportadas por evidencia empírica^{41,42}. La misma organización recomienda que, en caso de usarlas, se haga con cautela. Los autores del presente artículo quieren recalcar, por tanto, que pueden existir otros sistemas que ofrezcan mayor fiabilidad en la evaluación de la calidad de los estudios, pero también advierten que establecer otro tipo de metodología para evaluar la calidad de los estudios podría requerir una mayor inversión de tiempo y mejores conocimientos en bioestadística. Es por ello que se aconseja usar, conociendo las limitaciones del sistema, alguna de las citadas escalas de calidad detalladas por Moher et al³⁸, antes que no usar ningún sistema para evaluar la calidad individual de los estudios. Como el actual documento pretende ser una herramienta práctica que sirva para evaluar de forma rápida la calidad de ECCA de cualquier materia, se ha valorado cualitativamente cada uno de los 8 puntos que caracterizan las diferentes escalas³⁸ para determinar cuál de ellas puede ser la más adecuada.

Pese a que hay muchas escalas, con múltiples ítems y complejos sistemas de puntuación que pueden hacer que el evaluador invierta mucho tiempo en con-

testarlos, no se ha comprobado que dichas escalas ofrezcan mayor validez que otras más sencillas⁴³. Por tanto, tras revisar las 25 escalas de puntuación del estudio de Moher et al³⁸, la escala de Detsky et al³⁶ parece ser la que mejor se adapta a las necesidades de los lectores del presente artículo. Dicha escala sirve para evaluar ECCA de cualquier materia y se refiere tanto a la calidad metodológica como a la calidad general del artículo. Hay, como mínimo, un ítem para identificar la asignación de los pacientes y otro acerca del enmascaramiento. Asimismo, hay, como mínimo, un ítem para evaluar el análisis estadístico, y es uno de los que más rápidamente se puede contestar (aproximadamente en 10 min) (tabla 4). Además, tal y como detallan Petrison et al¹⁸, el sistema desarrollado por Detsky et al³⁶ es un método validado. Otros estudios posteriores han usado dicha escala para evaluar, con éxito, ECCA de distintas materias⁴⁴⁻⁴⁶.

Esta puede ser, por tanto, una forma práctica de clasificar los ECCA según el porcentaje de cumplimiento respecto al total de puntos obtenidos en el test de Detsky et al³⁶.

Otra forma de determinar la calidad individual de un estudio es mediante las listas de control sin puntuaciones o *checklists*. El Center of Reviews and Dissemination (CRD) usa como ejemplo los 9 ítems del Consenso de Delphi (Verhagen et al⁴⁷):

1. ¿Realmente la asignación del tratamiento a los grupos fue aleatoria?

- Enfoques apropiados para la generación de secuencias: a) números aleatorios generados por computadora, y b) tablas de números aleatorios.
- Enfoques inapropiados para la generación de secuencias: uso de alternancias, números de historia clínica, fechas de nacimientos, días de la semana, etc.

2. ¿Se ocultó la asignación al tratamiento?

- Enfoques apropiados del enmascaramiento de la aleatorización: a) aleatorización centralizada o controlada; b) contenedores numerados con secuencias idénticas; c) sistema computarizado con una secuencia aleatoria no legible hasta la asignación, y d) otros enfoques con métodos sólidos para prevenir el reconocimiento de la secuencia de asignación por parte de los investigadores o los pacientes.

- Enfoques inapropiados del enmascaramiento de la aleatorización: a) uso de alternancias, números de historia clínica, fechas de nacimientos, días de la semana, etc.; b) listas legibles de números aleatorios, y c) envases numerados con secuencias (incluso los envases opacos sellados pueden estar sujetos a manipulaciones).

3. ¿Los grupos eran similares al inicio del estudio en términos de factores pronóstico?

4. ¿Se especificaron los criterios de inclusión?

5. ¿Hubo ciego simple (paciente), doble ciego (investigador) o triple ciego (quien analiza los datos)?

6. La estimación puntual y la medida de la variabilidad ¿fueron presentadas para la variable de valoración predefinida (es decir, la de mayor importancia y la que en un estudio habitualmente determina el tamaño muestral)?

7. ¿Incluyeron los análisis una evaluación por intención de tratamiento?

También, en este caso, se podría clasificar los ECCA según el porcentaje de respuestas favorables o negativas a dicho *checklist*.

Estudios observacionales

Para los estudios observacionales, la Agency for Healthcare Research and Quality de la US Department of Health and Human Services concluye que para poder inferir un resultado válido a partir de estudios observacionales, los investigadores deben asegurar, como mínimo, la comparabilidad entre los grupos de estudio (exceptuando los factores que están bajo estudio), y garantizar que se haya hecho un adecuado análisis estadístico de los resultados¹⁵.

Asimismo, el Center of Reviews and Dissemination señala que los estudios de cohorte que usan datos históricos para comparar con el grupo de estudio actual, son vistos, generalmente, como de menor calidad que aquellos que comparan dos grupos actuales, y añade, además, que los estudios prospectivos acostumbran a tener menor riesgo de sesgos que los estudios retrospectivos²⁵.

Pese a que existen distintos *checklists* para la evaluación de la calidad individual de estudios epidemiológicos, como, por ejemplo, el propuesto por Levine et al⁴⁸, el Center of Reviews and Dissemination propone una fusión de distintos *checklists*²⁵:

1. Estudios de cohortes:

- ¿Hay suficiente descripción de los grupos y de la distribución de los factores de pronóstico?
- ¿Los grupos se agruparon en función de la similitud en la progresión de su enfermedad?
- ¿Se estableció formalmente la intervención o el tratamiento?
- ¿Los grupos son comparables en relación con todos los factores de confusión?

- ¿Se ajustaron adecuadamente los efectos de dichos factores de confusión?

- ¿Hay una relación dosis-respuesta entre la intervención y los resultados?

- ¿La persona que analiza fue cegada en referencia a la intervención?

- ¿El seguimiento fue lo suficientemente largo como para que tuviesen lugar los resultados?

- ¿Qué proporción de la cohorte fue seguida?

- ¿La tasa de abandono y sus causas fueron similares entre los grupos intervención y control?

2. Estudios de casos y controles:

- ¿Es explícita la definición del grupo «caso»?

- ¿Se ha valorado y validado correctamente el estado de la enfermedad de los «casos»?

- ¿Se aleatorizó a los grupos control a partir de la fuente de población de los «casos»?

- ¿Son comparables los casos y los controles con respecto a los potenciales factores de confusión?

- ¿Se valoró de la misma manera las intervenciones y otras exposiciones tanto en los casos como en los controles?

- ¿Cómo se definió la tasa de respuesta?

- ¿Las tasas de falta de respuesta y sus razones fueron iguales en ambos grupos?

- ¿Es posible que haya habido un sobreemparejamiento en cuanto a la asignación de los casos y los controles a los factores relacionados con la exposición?

- ¿Se realizó un análisis estadístico apropiado? (emparejado o no emparejado).

3. Series de casos:

- ¿El estudio se basa en una muestra representativa seleccionada de una población relevante?

- ¿Los criterios de inclusión son explícitos?

- ¿Todos los individuos empezaron la investigación en un punto similar de la progresión de su enfermedad?

- ¿El seguimiento fue lo suficientemente largo como para que ocurriesen eventos importantes?

- ¿Los parámetros se valoraron siguiendo criterios objetivos o se utilizó un método de cegamiento?

- Si se han hecho comparaciones de subgrupos, ¿hubo suficiente descripción de las series y de la distribución de los factores pronóstico?

Otros criterios aplicables a todos los tipos de estudios

Además de los dominios descritos en los apartados correspondientes a cada tipo de diseño de estudio, también se puede verificar, como propone la Agency for Healthcare Research and Quality de la US Department of Health and Human Services¹⁴: los patrocinadores y los conflictos de intereses.

Por tanto, entre estudios de igual diseño, la calidad individual puede hacer que los de más calidad estén, científicamente hablando, por encima de los estudios de menor calidad. De la misma forma, estudios de mala calidad que teóricamente ofrecerían, según su diseño, mayor nivel de evidencia, puede que estén por debajo de otros que serían de calidad científica inferior. A este respecto no se ha hallado ningún estudio que establezca de forma clara a partir de qué puntuación se puede decir que un estudio es de menor calidad individual que otro. Vista la falta de información, los autores del presente documento proponen clasificar la calidad de los estudios, de forma transitoria hasta tener más datos, de la siguiente manera:

- Calidad alta: estudios cuya puntuación o porcentaje de cumplimiento de las escalas establecidas en este punto supere el 75%.

Tabla 5
Clasificación del GREP-AEDN de los niveles de evidencia y grados de recomendación

Nivel de evidencia	Tipo de estudio	Grado de recomendación
Ia	Revisión sistemática con trato estadístico de ECCA Metaanálisis de ECCA	A1
Ib	Revisión sistemática de ECCA (sin trato estadístico) Revisiones de ECCA ECCA de más de 100 pacientes	A2
Ila	Revisiones sistemáticas con trato estadístico de estudios de cohortes prospectivos Metaanálisis de estudios de cohortes prospectivos	B1
Ilb	ECCA de menos de 100 pacientes Ensayos clínicos controlados (sin aleatorización): dentro de éstos, y por este orden: revisión sistemática con trato estadístico, metaanálisis, estudios de revisión y ensayos aislados Revisiones de estudios de cohortes prospectivos Estudios de cohortes prospectivos aislados Estudios de casos-control (control no histórico): dentro de éstos, y por este orden: revisión sistemática con trato estadístico, metaanálisis, estudios de revisión y estudios aislados Estudios controvertidos. Si hay varios estudios de igual diseño y calidad, pertenecientes desde el nivel Ia hasta el nivel Ila, y de los cuales se desprenden resultados o conclusiones contradictorias, dichos estudios serán catalogados en este nivel de evidencia (Ilb) ^{53*}	B2
III	Ensayos clínicos no controlados Estudios de cohortes con control histórico Estudios de casos-control (con control histórico) Estudios transversales Estudios de cualquier tipo (experimentales, observacionales, de revisión o consenso/posicionamiento) de «calidad media»	C
IV	Opiniones de expertos. Si dichas opiniones no se acompañan de referencias bibliográficas, se incluirán en este nivel de evidencia (IV). Sin embargo, si tales opiniones han sido clasificadas, a su vez, en niveles de evidencia y se quiere aprovechar esta clasificación, se debe consultar la bibliografía citada por dicho experto o comité de expertos y comprobar: 1. Que la clasificación ha sido correctamente realizada 2. Que la clasificación se equipara a lo descrito en esta tabla 3. Si hay estudios posteriores a la fecha de publicación de dicha opinión, se deberá analizarlos según lo descrito en el presente documento Otros tipos de revisiones Otros documentos (libros, capítulos de libros, etc.)	D
Grado de recomendación	Descripción	
A1	Recomendaciones cuya variación a lo largo del tiempo es altamente improbable	
A2	Recomendaciones cuya variación a lo largo del tiempo es improbable	
B1	En caso de que no se pueda aprobar un estudio de diseño tipo Ia o Ib, se podrá considerar a este grupo de un grado de recomendación A2, en caso contrario será considerado como B2	
B2	Recomendaciones que es probable que varíen a lo largo del tiempo	
C	Hay estudios limitados o no están sustentados en suficientes estudios. No se puede establecer recomendación	
D	No se han hallado estudios al respecto	

ECCA: ensayo clínico controlado, aleatorizado.

*Los autores concluyen, tras consultar la opinión de algunos expertos, que se deberá evaluar cuidadosamente las evidencias cuando muchos estudios catalogados como nivel de evidencia bajo sean refutados por un solo estudio de nivel de evidencia superior.

- Calidad media: cuya puntuación o porcentaje de cumplimiento de las escalas establecidas en este punto estén entre el 50 y el 75%.
- Calidad baja: cuya puntuación o porcentaje de cumplimiento de las escalas establecidas en este punto estén por debajo del 50%.

Así, ¿cómo puede afectar la calidad de un estudio individual a la jerarquía de niveles de evidencia? Se propone que las recomendaciones procedentes de estudios de calidad media tengan un formato igual o similar al

siguiente: «hay pocos estudios o las conclusiones no están sustentadas en suficientes estudios. No se puede establecer una recomendación», mientras que los estudios de calidad baja automáticamente serán desechados.

Cantidad relativa de muestra de población y duración de la intervención

La cantidad de muestra de población de los estudios también es uno de los factores que debe tenerse en cuenta para establecer el peso que se le tiene que dar a dicho estudio. Hay sistemas para establecer el tamaño de muestra que debe usarse según las características del estudio. Pese a que la explicación de dichos sistemas excede los objetivos fijados para esta revisión, cabe resaltar que la US Agency for Health Care Policy and Research clasifica los ECCA de menor evidencia científica¹⁹. Es decir, que entre estudios de igual diseño, y cuya calidad individual sea parecida, serán de mayor calidad científica aquellos que tengan una muestra mayor de población.

En este sentido, cuando se comparan estudios de diseño distinto, cabe destacar que los estudios observacionales y los metaanálisis, por sus características, pueden tener una mayor cantidad relativa de muestra de población respecto a los ECCA. Esto no implica que dichos estudios sean directamente de mayor calidad que los ECCA. Únicamente los metaanálisis de ECCA, con su correspondiente análisis estadístico bien realizado, serán de mayor evidencia científica que los ECCA de más de 100 pacientes de forma individual. Sin embargo, las revisiones sistemáticas con trato estadístico o metaanálisis de estudios de cohortes prospectivos podrán ser de igual evidencia científica que los ECCA con menos de 100 pacientes^{18,19}.

De igual forma, al comparar estudios de igual diseño, calidad individual parecida y tamaño de muestra parecido, serán de mayor calidad científica aquellos cuya duración de la intervención sea mayor.

Conclusiones. Creando una metodología

Actualmente existe la necesidad de que el dietista-nutricionista, o cualquier otro profesional que se dedique a la dietética y nutrición aplicada (tanto clínica como comunitaria), utilice la DNABE en sus métodos para obtener guías de práctica clínica que incorporen recomendaciones basadas en la evidencia científica⁴⁹⁻⁵¹.

Considerando los tres dominios principales (el tipo de diseño de estudio, la calidad individual de los estudios¹⁸ y la cantidad relativa de muestra de población¹⁹ y la duración del estudio), se recomienda la siguiente metodología de trabajo para crear dichas recomendaciones:

1. Definir la pregunta o cuestión a tratar: cuanto más abierta, más difícil será recuperar (mediante una estrategia de búsqueda) estu-

dios relevantes. Pero si la pregunta es demasiado concisa, se podría estar dejando de lado consideraciones potencialmente importantes en la decisión final⁴⁰.

2. Ejecutar una correcta estrategia de búsqueda a través de una base de datos especializada. Para más información se propone consultar el artículo: «Introducción a la búsqueda de información en internet de interés para dietistas-nutricionistas (2): bases de datos electrónicas especializadas, el primer paso hacia la evidencia científica»⁵².

3. Agrupar los estudios según su diseño:

- Experimental: ECCA, estudio clínico controlado (no aleatorizado) o estudio clínico (no controlado).
- Observacional: estudio de cohortes (identificar si es prospectivo o retrospectivo), estudio caso-control o estudio transversal.
- De revisión: revisión sistemática (con o sin trato estadístico), metaanálisis o revisión.
- Otros: documentos de consenso, posicionamiento u otros tipos. Tanto en los artículos de revisión como en otros hay que establecer si dichos estudios han seleccionado ECCA, o bien, en su caso, identificar el tipo de estudio observacional.

4. Establecer la calidad individual de los estudios según lo descrito en el apartado «Sistema para estimar la calidad individual de los artículos» en: alta, media o baja.

- ECCA (usar ambos sistemas): sistema de Detsky et al³⁶ y sistema de consenso Delphi⁴⁷.
- Observacionales: sistema del CRD²⁵ (véase el apartado «Sistema para estimar la calidad individual de los artículos», concretamente el subapartado «Estudios observacionales»).
- Todos¹⁵: patrocinadores y conflictos de intereses. Desechar los estudios de calidad baja.

5. Clasificar la evidencia según la jerarquía descrita en la tabla 5.

6. Dentro de cada uno de los tipos de diseño de un estudio, ordenar por cantidad de muestra y tiempo que dura la intervención u observación.

7. Crear la recomendación, añadiendo una frase que identifique el espíritu del grado de recomendación.

Reflexiones finales

Tal y como indica Doherty³, pese a que la medicina basada en la evidencia se ha ido consolidando, con el objetivo de disminuir fuentes de error en la elección del tratamiento o para prevenir la prescripción de tratamientos peligrosos, hay algunos estudios contemporáneos que indican que más del 25% de los pacientes aún reciben tratamientos que les podrían perjudicar⁵⁴, y que más de un 40% no recibe tratamientos de comprobada efectividad⁵⁵.

Según Doherty³, hay algunas razones para que las buenas evidencias no sean realmente adoptadas en la práctica clínica⁵⁶. Éstas se pueden resumir en:

- Características de la propia evidencia (por ejemplo, la dificultad de realizarla).
- Barreras o resistencias en la capacidad de adopción de la evidencia, que se pueden dar en el paciente, el profesional, el grupo de trabajo, la institución y otros.
- Transferencia inadecuada de la evidencia al paciente o también llamada «traducción errónea del conocimiento».

Con el presente documento se ha propuesto un sistema para poner en marcha la dietética y la nutrición aplicada basadas en la evidencia en España. Ha sido la intención de los autores crear una herramienta práctica para que los dietistas-nutricionistas puedan in-

volucrarse en dicha tarea y así mejorar el nivel de sus recomendaciones, sin caer en la primera de las barreras que pueden existir a la hora de aplicar las buenas evidencias: un sistema excesivamente complicado.

Investigaciones futuras deberán establecer si se ha podido implementar el sistema propuesto de niveles de evidencia y, en todo caso, establecer cuáles han sido los puntos fuertes o frágiles del sistema planteado.

Bibliografía

1. James EL, Fraser C, Anderson K, Judd F. Use of research by the Australian health promotion workforce. *Health Educ Res.* 2007;22:576-87.
2. Freeman AC, Sweeney K. Why general practitioners do not implement evidence: qualitative study. *BMJ.* 2001;323:1100-2.
3. Doherty S. History of evidence-based medicine. Oranges, chloride of lime and leeches: barriers to teaching old dogs new tricks. *Emerg Med Australas.* 2005;17:314-21.
4. Doll R. Controlled trials: the 1948 watershed. *BMJ.* 1998;317:1217-20.
5. Claridge JA, Fabian TC. History and development of evidence-based medicine. *World J Surg.* 2005;29:547-53.
6. Oxman AD, Sackett DL, Guyatt GH. Users guides to the medical literature. I. How to get started; the Evidence-Based Medicine Working Group. *JAMA.* 1993;270:2093-5.
7. Guyatt GH, Sackett DL, Cook DJ. Users guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA.* 1993;270:2598-601.
8. Sackett DL, Rosenberg WM. On the need for evidence-based medicine. *Health Econ.* 1995;4:249-54.
9. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *BMJ.* 1996;312:71-2.
10. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *BMJ.* 1995;310:1122-6.
11. Canadian Task Force on the Periodic Health Examination. The periodic health examination. *Can Med Assoc J.* 1979;121:1193-254.
12. Eccles M, Freemantle N, Mason J. North of England evidence based guideline development project: guideline on the use of aspirin as secondary prophylaxis for vascular disease in primary care. North of England Aspirin Guideline Development Group. *BMJ.* 1998;316:1303-9.
13. Scottish Intercollegiate Guidelines Network (SIGN). SIGN 50: A Guideline Developer's Handbook. Section 6 [citado 23 Nov 2001]. Disponible en: <http://www.sign.ac.uk/guidelines/fulltext/50/sections5.html>
14. US Preventive Services Task Force. Guide to Clinical Preventive Services. 2.^a ed. Washington: US Department of Health and Human Services; 1996.
15. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al. Systems to rate the strength of scientific evidence. *Evid Rep Technol Assess (Summ).* 2002;47:1-11.
16. Wammes BM, Blom CA, Koelen M, De Groot SD, Remijnse-Meester WT, Van Staveren WA. Implementation research for 'evidence-based' guideline development by dietitians: a pilot study to test an instrument. *J Hum Nutr Diet.* 2002;15:243-54.
17. Blumberg-Kason S, Lipscomb R. Evidence-based nutrition practice guidelines: A valuable resource in the evidence analysis library. *J Am Diet Assoc.* 2006;106:1935-6.
18. Petrisor BA, Keating J, Schemitsch E. Grading the evidence: levels of evidence and grades of recommendation. *Injury.* 2006;37:321-7.
19. Hadorn DC, Baker D, Hodges JS, Hicks N. Rating the quality of evidence for clinical practice guidelines. *J Clin Epidemiol.* 1996;49:749-54.
20. Internacional Food Information Council (IFIC). Cómo entender e interpretar los estudios científicos que se relacionan con la salud y la alimentación. Informe IFIC; 2001 [citado 27 Feb 2008]. Disponible en: <http://www.ific.org/sp/publications/reviews/scientificirsp.cfm>
21. National Center for Biotechnology Information (NCBI). MeSH (Medical Subject Heading) US National Library of Medicine's controlled vocabulary used for indexing articles [citado Feb 2008] Disponible en: <http://www.ncbi.nlm.nih.gov/sites/entrez>
22. Laupacis A, Straus S. Systematic reviews: time to address clinical and policy relevance as well as methodological rigor. *Ann Intern Med.* 2007;147:273-4.
23. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine.* Boston: Little, Brown, and Company; 1991.

24. Brighton B, Bhandari M, Tornetta III P, Felson DT. Hierarchy of evidence: from case reports to randomized controlled trials. *Clin Orthop Relat Res.* 2003;413:19-24.
25. Khan KS, Riet G, Glanville J, Sowden AJ, Kleijnen J. Undertaking systematic reviews of research on effectiveness: CRD's Guidance for those Carrying Out or Commissioning Reviews. Center of Reviews and Dissemination (CRD) Report Number 4. 2.^a ed. York: March; 2001.
26. Kroke A, Boeing H, Rossnagel K, Willich SN. History of the concept of 'levels of evidence' and their current status in relation to primary prevention through lifestyle interventions. *Public Health Nutr.* 2004;7:279-84.
27. Baladia E, Basulto J. Introducción a la búsqueda de información en internet de interés para dietistas-nutricionistas. *Act Diet.* 2006;30:15-6.
28. Disorders of Lipid Metabolism Evidence-Based Guideline Workgroup. Disorders of lipid metabolism evidence-based nutrition practice guideline. National Guideline Clearinghouse. 2006 [citado 11 Mar 2008]. Disponible en: http://www.guideline.gov/summary/summary.aspx?ss=15&doc_id=9533&nbr=5083#s23
29. Standards of medical care in diabetes - 2008. *Diabetes Care.* 2008;31 Suppl 1:S12-54.
30. Robeson P, Albanese ML, Ciliska D, et al. Primary prevention of childhood obesity. National Guideline Clearinghouse; 2005 [citado 10 Mar 2008]. Disponible en: http://www.guideline.gov/summary/summary.aspx?doc_id=7092&nbr=4265#s24
31. Evidence-Based Clinical Practice Guideline Development & Revision Team. Evidence-based clinical practice guideline. Cardiovascular health for women: primary prevention. 2.^a ed. National Guideline Clearinghouse. 2001 [citado 10 Mar 2008]. Disponible en: http://www.guideline.gov/summary/summary.aspx?ss=15&doc_id=4036&nbr=3139#s24
32. Donald T, Ennis J, Mackenzie L, et al. Management of osteoporosis. A national clinical guideline. National Guideline Clearinghouse. 2004 [citado 10 Mar 2008]. Disponible en: http://guideline.gov/summary/summary.aspx?doc_id=3876&mode=full&ss=15#s24
33. Evans D, Andrews F, Campbell I, et al. Prevention and management of dental decay in the pre-school child. A national clinical guideline. National Guideline Clearinghouse. 2006 [citado 10 Mar 2008]. Disponible en: http://www.guideline.gov/summary/summary.aspx?doc_id=8395#s24
34. Khan KS, Daya S, Jadad A. The importance of quality of primary studies in producing unbiased systematic reviews. *Arch Intern Med.* 1996;156:661-6.
35. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet.* 1998;352:609-13.
36. Detsky AS, Naylor CD, O'Rourke K, et al. Incorporating variations in the quality of individual randomized trials into metaanalysis. *J Clin Epidemiol.* 1992;45:255-65.
37. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17:1-12.
38. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials.* 1995;16:62-73.
39. Jadad-Bechara AR. Meta-Analysis of randomised clinical trials in pain relief. DPhil thesis, University of Oxford. 1994.
40. Higgins JPT, Green S, editores. *Cochrane Handbook for Systematic Reviews of Interventions 4.2.6* [actualizado Sep 2006; citado 6 Oct 2006]. Disponible en: <http://www.cochrane.org/resources/handbook/hbook.htm>
41. Schulz KF, Chalmers I, Hayes RJ, Altman D. Empirical evidence of bias. *JAMA.* 1995;273:408-12.
42. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care.* 1996;12:195-208.
43. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA.* 1999;282:1054-60.
44. Sánchez-Thorin JC, Rocha G, Yelin JB. Meta-analysis on the recurrence rates after bare sclera resection with and without mitomycin C use and conjunctival autograft placement in surgery for primary pterygium. *Br J Ophthalmol.* 1998;82:661-5.
45. Bhandari M, Richards RR, Sprague S, Schemitsch EH. The quality of reporting of randomized trials in the *Journal of Bone and Joint Surgery* from 1988 through 2000. *J Bone Joint Surg Am.* 2002;84-A:388-96.
46. Dulai SK, Slobogean BL, Beauchamp RD, Mulpuri K. A quality assessment of randomized clinical trials in pediatric orthopaedics. *J Pediatr Orthop.* 2007;27:573-81.
47. Verhagen AP, De Vet HC, De Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol.* 1998;51:1235-41.
48. Levine M, Walters S, Lee H, Haines T, Holbrook A, Moyer V, et al. Users' guides to the medical literature. IV: How to use an article about harm. *JAMA.* 1994; 271:1615-9.
49. Serra-Majem LI. Nutrición Basada en la Evidencia (NUBE). *Rev Esp Nutr Comunitaria.* 2000;6:5-6.
50. Doreste JL, Serra-Majem LI. Nutrición Basada en la Evidencia RESPYN. 2005; 6:sin identificar [citado 2 Abr 2008]. En: http://www.respyn.vanl.mx/vi/2/ensayos/NUBE_Indexado.htm
51. Doreste Alonso JL, Serra L. Nutrición basada en la evidencia. En: Serra Majem L, Aranceta Bartrina J, editores. *Nutrición y salud pública.* 2.^a ed. Madrid: Masson; 2006.
52. Baladia E, Basulto J. Introducción a la búsqueda de información en Internet de interés para dietistas-nutricionistas (2): bases de datos electrónicas especializadas, el primer paso hacia la evidencia científica. *Act Diet.* 2006;32:22-7.
53. Grupo de trabajo sobre GPC. Elaboración de Guías de Práctica Clínica en el Sistema Nacional de Salud. Manual Metodológico. Guías de Práctica Clínica en el SNS: I+CS. N.º 2006/01. Madrid: Plan Nacional para el SNS del MSC, Instituto Aragonés de Ciencias de la Salud-I+CS; 2007.
54. Grol R. Successes and failures in the implementation of evidencebased guidelines for clinical practice. *Med Care.* 2001;39 Suppl 2:II46-54.
55. McGlynn EA, Asch SM, Adams J, et al. The quality of care delivered to adults in the United States. *N Engl J Med.* 2003;348:2635-45.
56. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet.* 2003;362:1225-30.