



MONOGRÁFICO SOBRE REVISIONES SISTEMÁTICAS Y METAANÁLISIS

EDITORIAL

Revisiones sistemáticas, una herramienta clave para la toma de decisiones clínicas y sanitarias.
Gerard Urrutia y Xavier Bonfill. 1-3.

COLABORACIÓN ESPECIAL

Desarrollo de una hoja excel para metaanálisis de comparaciones indirectas y mixtas. **Aurelio Tobías, Ferrán Catalá-López y Marta Roqué. 5-15.**

REVISIONES SISTEMÁTICAS CON METAANÁLISIS

Neuroticismo y trastorno por estrés postraumático: un estudio meta-analítico. **Fabiola B Soler-Ferrería, Julio Sánchez-Meca, José Manuel López-Navarro y Fernando Navarro-Mateu. 17-36.**

Bloqueo dual del sistema renina-angiotensina frente a la monoterapia: revisión sistemática y metaanálisis acumulativo de ensayos clínicos y estudios observacionales. **Ferrán Catalá-López F, Diego Macías Saint-Gerons, César de la Fuente Honrubia y Gloria Martín-Serrano. 37-65.**

Evaluación de técnicas inmunológicas in vitro para el diagnóstico de alergias: metaanálisis 2000-2012.
John Fredy Cuervo-Perez, Julián Camilo Arango y Jaiberth Antonio Cardona-Arias. 67-84.

Impacto del VIH/sida sobre la calidad de vida: metaanálisis 2002-2012. **Jaiberth Antonio Cardona-Arias y Luis Felipe Higueta-Gutiérrez. 85-99.**

Metaanálisis de comparación de grupos y metaanálisis de generalización de la fiabilidad del cuestionario *State-Trait Anxiety Inventory* (STAI). **Alejandro Guillén-Riquelme y Gualberto Buela-Casal. 101-112.**

REVISIONES SISTEMÁTICAS

Modelos de colaboración entre atención primaria y salud mental en la asistencia sanitaria a las personas con depresión: resultados principales y retos metodológicos de una meta-revisión sistemática. **Carlos Calderón Gómez, Isabel Mosquera Metcalfe, Laura Balagué Gea, Ander Retolaza Balsategui, Amaia Bacigalupe de la Hera, Jon Belaunzaran Mendizabal, Álvaro Iruin Sanz y Beatriz García Moratalla. 113-133.**

Frecuentación de las consultas de medicina general y especializada por población inmigrante y autóctona: una revisión sistemática. **Rocío Carmona, Raimundo Alcázar-Alcázar, Antonio Sarria-Santamera y Enrique Regidor. 135-155.**

Efectividad de distintas terapias físicas en el tratamiento conservador de la fascitis plantar. revisión sistemática.
Ana Maria Diaz Lopez y Patricia Guzman Carrasco. 157-178.

RELACIÓN DE EVALUADORES 179.

VERSIÓN EN INGLÉS

EDITORIAL**REVISIONES SISTEMÁTICAS, UNA HERRAMIENTA CLAVE PARA LA TOMA DE DECISIONES CLÍNICAS Y SANITARIAS****Gerard Urrutia (1,2) y Xavier Bonfill (1,2,3).**

(1) Centro Cochrane Iberoamericano – Institut d'Investigació Biomèdica Sant Pau (IIB Sant Pau).

(2) CIBERESP.

(3) Universitat Autònoma de Barcelona.

Debemos felicitarnos por este monográfico que la Revista Española de Salud Pública dedica a las revisiones sistemáticas. Si bien es cierto que el número de revisiones sistemáticas publicadas en revistas españolas indexadas ha ido en aumento en los últimos años (en un intento rápido de sustentar esta afirmación, hemos estimado que serían unas 750 las revisiones sistemáticas o metanálisis publicados hasta la fecha), hasta donde nosotros sabemos se trata del primer número de una revista biomédica española dedicado específicamente a las revisiones sistemáticas. Este hecho pone bien de manifiesto la importancia creciente que las revisiones sistemáticas, como modelo de investigación (denominada secundaria) y también como tipo de publicación científica (síntesis) están adquiriendo en todo el mundo. En un editorial reciente de esta misma revista argumentábamos acerca de esta tendencia¹. Algunos datos que apoyarían esta afirmación serían la notable expansión de los recursos bibliográficos específicos (bases de datos bibliográficas de documentos de síntesis, como la *Cochrane Library* o su versión en idioma español la Biblioteca Cochrane Plus)² o de herramientas que facilitan su identificación y utilización (como la base de datos colaborativa Epistemonikos)³, así como también el elevado factor de impacto que algunas de estas fuentes han obtenido en poco tiempo desde su aparición (por ejemplo, la *Cochrane Database of Systematic Reviews* tiene actualmente un factor de

impacto de 5,703 que la ha colocado en el primer decil dentro de la categoría de revistas del área de medicina general/medicina interna). Por todo ello, esta iniciativa debe ser muy bienvenida en la medida que supone un incentivo para que otros se animen también a emprender esta labor científica tan necesaria y útil, todavía muy lejos de la actividad desarrollada en otros países aumenta la percepción social del valor de este tipo de contribuciones y difunde sus resultados para que otros puedan fundamentar mejor sus decisiones clínicas y sanitarias.

Son varios los aspectos específicos que deseamos destacar de este monográfico. En primer lugar, la presencia de un artículo de colaboración especial que trata acerca de aspectos metodológicos relacionados con las revisiones y metanálisis. Se trata este de un campo muy dinámico y prolífico donde constantemente se están investigando y proponiendo nuevos métodos y directrices para realizar este tipo de estudios con mayor calidad. A modo de ejemplo, la Colaboración Cochrane (a través de sus 16 grupos de métodos)⁴, organización referente por lo que respecta a las revisiones sistemáticas, o el grupo de trabajo internacional GRADE^{5,6} que lo es en el desarrollo de guías de práctica clínica, están impulsando numerosas iniciativas de amplio calado para refinar la metodología de las revisiones y mejorar su utilización por parte de los usuarios finales. En esta monografía se incluye un artículo

que trata sobre uno de los desarrollos más recientes y que despierta mayor interés, como son los metanálisis en red (*network meta-analysis*)^{7,8} pues permiten combinar datos de estudios en los que se han realizado comparaciones directas entre intervenciones alternativas con datos procedentes de comparaciones indirectas. Se trataría de una especie de triangulación, donde se busca aprovechar toda la información disponible, especialmente cuando, como suele ser el caso en demasiadas ocasiones, escasean las comparaciones directas entre alternativas de tratamiento. Es una buena noticia observar cómo en nuestro ámbito existen expertos que demuestran poseer experiencia en esta materia tan compleja y, además, lo divulgan de forma muy didáctica.

En segundo lugar, cabe destacar la variedad de temas, de modalidades de intervención y de métodos de análisis utilizados en las revisiones que componen este monográfico, lo que, una vez más, denota la riqueza de posibilidades que ofrece esta metodología así como también algunas controversias subyacentes. Mientras que algunas revisiones sistemáticas adoptan el enfoque clásico evaluando los efectos de una intervención a partir del análisis combinado (metanálisis) de ensayos clínicos controlados aleatorizados, otras lo hacen mediante la técnica del metanálisis acumulativo combinando datos de ensayos clínicos y estudios observacionales. También hay aquí revisiones que aplican el metanálisis para evaluar pruebas diagnósticas o para evaluar la fiabilidad de un test. Estos son aspectos de más reciente desarrollo, donde todavía existe un amplio recorrido metodológico por completar, no exento de limitaciones y controversias, pero donde la necesidad y oportunidad de hacerlo es incuestionable. A modo de ejemplo, cabe mencionar que la Colaboración Cochrane ha desarrollado solo recientemente métodos específicos para realizar revisiones con metanálisis de pruebas diagnósticas⁹ que han permitido iniciar la publicación de este tipo de revisiones en la *Cochrane Library*.

También existen en nuestro país grupos de investigación que han hecho importantes contribuciones en esta aplicación específica del metanálisis como, por ejemplo, el grupo de la Unidad de Bioestadística del Hospital Ramon y Cajal, desarrollando el software de libre distribución conocido como Meta-DiSc¹⁰.

En tercer lugar, cabe destacar que muchas de las intervenciones que aquí se evalúan son de salud pública y cabría considerarlas como complejas puesto que, a diferencia de una típica intervención médica, están compuestas por múltiples componentes que a menudo no son estandarizables y que interactúan entre sí. Sin duda, la evaluación de este tipo de intervenciones constituye todo un reto, tanto por la dificultad intrínseca para la realización de estudios primarios que evalúen su eficacia como por la dificultad que supone su evaluación de conjunto por medio de una revisión sistemática y su correcta interpretación. Este es un ámbito donde se está prestando una creciente atención por la necesidad que existe de disponer de métodos adecuados para interpretar esta información. Sólo así será posible identificar qué componentes de estas intervenciones son los que determinan unos mejores resultados, de modo que puedan tomarse decisiones bien fundamentadas en políticas de salud. A modo de ejemplo, recientemente la Colaboración Cochrane ha puesto en marcha un proyecto denominado *Methodological Investigation of Cochrane Reviews of Complex Interventions* (MICCI), cuyo objetivo es identificar modelos, herramientas, teorías y marcos conceptuales que hayan sido utilizados hasta la fecha en este tipo de revisiones, con el fin de elaborar y consensuar unas directrices específicas que puedan guiar a los futuros autores de este tipo de revisiones. Nuevamente, este ejemplo supone un pequeño botón de muestra del gran dinamismo que existe en el área del desarrollo metodológico relacionado con las revisiones. También dentro de la Colaboración Cochrane existe un grupo de revisión que, a diferencia de la

mayoría del resto de los grupos, está focalizado en el análisis de este tipo de intervenciones [*Effective Practice and Organisation of Care (EPOC) Group*]¹¹. El enfoque de este grupo está en la evaluación de intervenciones diseñadas para mejorar la prestación, la práctica y la organización de los servicios de salud, frecuentemente mediante intervenciones complejas. Desde este editorial nos atrevemos a invitar a cuantos están particularmente interesados en estos aspectos metodológicos para que participen activamente en las iniciativas mencionadas, pues sin duda serán determinantes para el desarrollo de este campo en los próximos años.

En cuarto y último lugar, cabe remarcar que para la revisión editorial de estos trabajos se han utilizado las directrices PRISMA, recientemente adoptadas por esta revista. Sin duda, esta ha sido una decisión importante en la medida que esta herramienta ayuda a estandarizar el proceso de revisión por pares, a la vez que contribuye a una mejor calidad de las publicaciones. Nunca se insistirá lo suficiente en la importancia de la calidad del reporte de los estudios de investigación (entre los cuales se hallan también las revisiones sistemáticas) que se publican en las revistas biomédicas y que se pueden consultar en la web de la red EQUATOR¹². Estas herramientas requieren, necesariamente, que los aspectos que son esenciales en cuanto al diseño del estudio y reporte de los resultados se comuniquen con el suficiente grado de detalle y además que se haga de una manera consistente, objetiva y clara para permitir su interpretación no sesgada.

En conclusión, sea muy bienvenido este número monográfico de la Revista Española de Salud Pública por el valor que tienen en sí mismo los trabajos que en él se publican y como precedente de una presencia de las revisiones sistemáticas que, en un futuro próximo, esperamos sea más habitual en las revistas biomédicas de nuestro país.

BIBLIOGRAFÍA

1. Urrútia G, Bonfill X. La Declaración PRISMA: un paso adelante en la mejora de las publicaciones de la Revista Española de Salud Pública. *Rev Esp Salud Pública* 2013;87:99-102.
2. La Biblioteca Cochrane Plus. [citado el 9-2-2014]. Disponible en: <http://www.biblioteca-cochrane.com>
3. Epistemonikos. Disponible en: <http://www.epistemonikos.org/es> [citado el 9-2-2014]
4. The Cochrane Collaboration. Disponible en: <http://www.cochrane.org/contact/methods-groups> [citado el 9-2-2014]
5. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group. Disponible en: <http://www.gradeworkinggroup.org/> [citado el 9-2-2014]
6. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924-6.
7. Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med*. 2013 Jul 16;159(2):130-7.
8. Catalá-López F, Tobías A. Síntesis de la evidencia clínica y metaanálisis en red con comparaciones indirectas. *Med Clin (Barc)* 2013;140(4):182-7.
9. Diagnostic Test Accuracy Working Group. Disponible en: <http://srdta.cochrane.org/handbook-dta-reviews> [citado el 9-2-2014]
10. Zamora J, Abraira V, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31.
11. Cochrane Effective Practice and Organisation of Care Group. Disponible en: <http://epoc.cochrane.org/> [citado el 9-2-2014]
12. Enhancing the QUALity and Transparency Of health Research. (Equator Network). Disponible en: <http://www.equator-network.org/> [citado el 9-2-2014]

COLABORACIÓN ESPECIAL**DESARROLLO DE UNA HOJA EXCEL PARA METAANÁLISIS
DE COMPARACIONES INDIRECTAS Y MIXTAS**

Aurelio Tobías (1), Ferrán Catalá-López (2,3) y Marta Roqué (4).

(1) Instituto de Diagnóstico Ambiental y Estudios del Agua (IDAEA). Consejo Superior de Investigaciones Científicas (CSIC). Barcelona.

(2) División de Farmacoepidemiología y Farmacovigilancia. Departamento de Medicamentos de Uso Humano. Agencia Española de Medicamentos y Productos Sanitarios (AEMPS). Madrid.

(3) Fundación Instituto de Investigación en Servicios de Salud. Valencia.

(4) Centro Cochrane Iberoamericano (CCIb). Hospital de la Santa Creu i Sant Pau. Barcelona.

Financiación: ninguna.

Las opiniones expresadas en este trabajo son responsabilidad de los autores, por lo que no reflejan necesariamente el punto de vista de los organismos en los que trabajan.

RESUMEN

En investigación clínica, el metaanálisis ha servido generalmente para evaluar la eficacia y la seguridad de un tratamiento a través de la comparación con un único comparador. Las comparaciones indirectas, a través del método de Bucher, permiten combinar datos procedentes de estudios primarios de una forma sencilla cuando la información para las comparaciones directas es limitada o inexistente. Las comparaciones mixtas permiten combinar las estimaciones procedentes de comparaciones directas e indirectas, aumentando el poder estadístico. En la actualidad, es necesario el desarrollo de aplicaciones informáticas para el cálculo de metaanálisis de comparaciones indirectas y mixtas. Por este motivo, se ha desarrollado una hoja de cálculo con Microsoft Office Excel para el cálculo de comparaciones indirectas y mixtas, de uso sencillo para investigadores clínicos interesados en revisiones sistemáticas con metaanálisis, pero no familiarizados con el uso de paquetes estadísticos más avanzados. La utilización de esta hoja de cálculo para las comparaciones indirectas y mixtas puede ser de gran utilidad en epidemiología clínica, con el fin de extender el conocimiento proporcionado por el metaanálisis tradicional cuando la evidencia de las comparaciones directas es limitada o inexistente.

Palabras clave: Metaanálisis como asunto. Revisión sistemática. Medicina Basada en la Evidencia. Práctica Basada en la Evidencia.

ABSTRACT**Development of an Excel Spreadsheet for
Meta-Analysis of Indirect and Mixed
Treatment Comparisons**

Meta-analyses in clinical research usually aimed to evaluate treatment efficacy and safety in direct comparison with a unique comparator. Indirect comparisons, using the Bucher's method, can summarize primary data when information from direct comparisons is limited or nonexistent. Mixed comparisons allow combining estimates from direct and indirect comparisons, increasing statistical power. There is a need for simple applications for meta-analysis of indirect and mixed comparisons. These can easily be conducted using a Microsoft Office Excel spreadsheet. We developed a spreadsheet for indirect and mixed effects comparisons of friendly use for clinical researchers interested in systematic reviews, but non-familiarized with the use of more advanced statistical packages. The use of the proposed Excel spreadsheet for indirect and mixed comparisons can be of great use in clinical epidemiology to extend the knowledge provided by traditional meta-analysis when evidence from direct comparisons is limited or nonexistent.

Keyword: Meta-Analysis as Topic. Evidence-Based Medicine. Evidence-Based Practice. Systematic review.

Correspondencia

Aurelio Tobías

Instituto de Diagnóstico Ambiental y Estudios del Agua (IDAEA)

Consejo Superior de Investigaciones Científicas (CSIC)

C/ Jordi Girona 18-26

08034 Barcelona

Tel.: 934006100

aurelio.tobias@idaea.csic.es

INTRODUCCIÓN

La revisión sistemática con metaanálisis está considerada como la principal herramienta que permite sintetizar de manera integral la evidencia científica disponible, y cuando se realiza y se reporta de una manera completa y transparente, permite estimar el efecto de un determinado tratamiento con un elevado nivel de calidad y rigor científico¹. Los metaanálisis han estado dirigidos tradicionalmente a evaluar la eficacia y la seguridad de un tratamiento a través de una comparación directa con un único comparador, habitualmente placebo o cualquier tratamiento estándar. Este hecho ha puesto en evidencia la falta de información procedente de estudios que comparan directamente alternativas terapéuticas que compiten entre sí en la práctica clínica. Así, en los últimos años se han venido desarrollando e implementado métodos cuantitativos que permiten comparar diversos tratamientos entre sí de manera indirecta cuando estos tratamientos se han comparado frente a un comparador común (por ejemplo placebo), pero la información procedente de comparaciones directas es escasa o inexistente. Asimismo, en caso de existir evidencia procedente de comparaciones directas entre tratamientos, estas se podrían combinar junto a las comparaciones indirectas a partir de comparaciones mixtas, permitiendo así incrementar la evidencia científica al respecto.

De igual forma, la consolidación del metaanálisis tradicional de comparaciones directas ha estado también relacionada con su implementación en aplicaciones informáticas y programas estadísticos. En este sentido, existen diversas utilidades desarrolladas en hojas de cálculo para Microsoft Excel^{2,3} que permiten realizar metaanálisis de una manera cada vez más sencilla y cómoda, sin necesidad de recurrir al uso de complejos programas estadísticos. En este sentido, son escasas las aplicaciones informáticas para el cálculo de comparaciones indirectas y mixtas de uso fácil para el investigador clínico.

El objetivo de este trabajo es exponer los conceptos metodológicos básicos de las comparaciones indirectas y mixtas a partir del desarrollo de una sencilla hoja de cálculo en Microsoft Excel, con objetos visuales y programación en lenguaje Visual Basic.

COMPARACIONES INDIRECTAS Y MIXTAS

Comparaciones indirectas

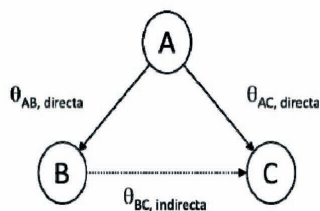
A partir de una revisión sistemática con metaanálisis de comparaciones directas podemos determinar cuanto mejor es un tratamiento B respecto a un tratamiento A ($\Theta_{AB, directa}$), y cuanto mejor es un tratamiento C respecto al mismo comparador A ($\Theta_{AC, directa}$)⁴. Brevemente, para la obtención de estimadores combinados directos, los resultados de los estudios individuales son ponderados en función de su precisión a través de diversas técnicas y modelos. La elección del modelo empleado depende fundamentalmente de la consideración de la hipótesis de homogeneidad de efectos entre los resultados de los estudios individuales incluidos en la revisión sistemática⁵. En el modelo de efectos fijos se asume que los estudios incluidos en el metaanálisis están estimando a un único, y común, efecto poblacional, por lo que la única variabilidad presente es debida al azar. El modelo de efectos aleatorios se fundamenta en que los estudios estiman una distribución de efectos en la población, asumiendo que los efectos subyacentes siguen una distribución normal con una media y varianza fijas. En otras palabras, el efecto individual de cada estudio estima a un efecto poblacional diferente.

En ausencia de estudios que presenten comparaciones directas de B respecto a C, es posible establecer una comparación indirecta y obtener una estimación de la medida del efecto ($\Theta_{BC, indirecta}$), como la diferencia entre las dos comparaciones directas (tabla 1). Este tipo de comparaciones indirectas también se suele denominar ‘comparacio-

Tabla 1

Descripción gráfica y formulación implementada en la hoja de Excel para comparaciones indirectas y mixtas⁴

Comparación indirecta entre B y C



Efecto indirecto:

$$\theta_{BC, indirecta} = \theta_{AC, directa} - \theta_{AB, directa}$$

con IC 95%:

$$\theta_{BC, indirecta} \pm 1,96 \times s[\theta_{BC, indirecta}]$$

Donde:

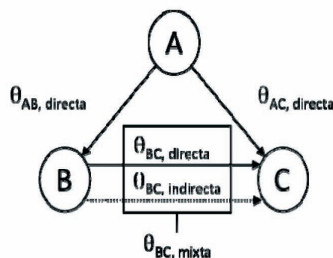
$$S^2[\theta_{AC, directa}] = [(IC_{superior} - IC_{inferior})/3,92]^2$$

$$S^2[\theta_{AB, directa}] = [(IC_{superior} - IC_{inferior})/3,92]^2$$

$$S^2[\theta_{BC, indirecta}] = S^2[\theta_{AC, directa}] + S^2[\theta_{AB, directa}]$$

$$s[\theta_{BC, indirecta}] = \sqrt{S^2[\theta_{BC, indirecta}]}$$

Comparación mixta entre B y C



Efecto mixto:

$$\theta_{BC, mixta} = \frac{\left(\theta_{BC, directa} \times \frac{1}{S^2[\theta_{BC, directa}]} \right) + \left(\theta_{BC, indirecta} \times \frac{1}{S^2[\theta_{BC, indirecta}]} \right)}{\left(\frac{1}{S^2[\theta_{BC, directa}]} + \frac{1}{S^2[\theta_{BC, indirecta}]} \right)}$$

con IC 95%:

$$\theta_{BC, mixta} \pm 1,96 \times s[\theta_{BC, mixta}]$$

Donde:

$$S^2[\theta_{BC, directa}] = [(IC_{superior} - IC_{inferior})/3,92]^2$$

$$S^2[\theta_{BC, mixta}] = \frac{1}{\left(\frac{1}{S^2[\theta_{BC, directa}]} + \frac{1}{S^2[\theta_{BC, indirecta}]} \right)}$$

$$s[\theta_{BC, mixta}] = \sqrt{S^2[\theta_{BC, mixta}]}$$

Θ : medida del efecto en escala aditiva (diferencia absoluta del riesgo (DR), logaritmo del riesgo relativo (log(RR)), logaritmo de la razón de odds (log(OR)), diferencia de medias (wmd) o diferencia de medias estandarizadas (smd))

S^2 : varianza de la medida del efecto; s : error estándar de la medida del efecto; IC: intervalo de confianza.

nes indirectas ajustadas' o método de Bucher⁴. Por ejemplo, supongamos que el tratamiento C es mejor que A y permite prevenir 5 casos adicionales de enfermedad por cada 100 pacientes tratados. Y que el tratamiento B es mejor que A y previene 2 casos por cada 100. Una estimación indirecta de la eficacia relativa del tratamiento C respecto el tratamiento B de forma aproximada, resultaría en una eficacia relativa de 3 casos prevenidos por cada 100 pacientes tratados. Así, sería posible estimar los potenciales beneficios relativos de las distintas comparaciones entre A, B y C.

Comparaciones mixtas

Asimismo, cuando se dispone de evidencia de estudios que presentan metaanálisis de comparaciones directas de B respecto a C ($\Theta_{BC, directa}$), es posible combinar la evidencia directa e indirecta ($\Theta_{BC, indirecta}$) en una comparación mixta ($\Theta_{BC, mixta}$), como un promedio ponderando por el inverso de sus respectivas varianzas⁶ (tabla 1). En el ejemplo, la comparación mixta entre C y B integra los resultados de la comparación directa de los dos tratamientos con los resultados de la comparación indirecta obtenida a partir de la información del resto de comparaciones directas (B frente a A y C frente a A). Las comparaciones mixtas pueden no solo complementar la información para aquellas comparaciones en las cuales existe poca información sino también pueden ayudar a mejorar la precisión de las estimaciones del efecto de los tratamientos.

Condiciones de aplicación: transitividad y consistencia

Es importante remarcar que la evidencia indirecta es de carácter observacional, pudiendo presentar serios problemas de validez interna⁷, por lo que se requieren unas condiciones básicas de aplicación. La primera es la transitividad^{8,9}, cuando los estudios en los que se comparan de forma directa el tratamiento B con A y el tratamiento C con A son comparables por no

diferir en relación a la distribución de factores modificadores del efecto. La transitividad puede formularse de una manera sencilla del modo que una comparación indirecta entre tratamientos estima válidamente una comparación directa no observada. Aunque la transitividad no se puede probar estadísticamente, su plausibilidad se puede evaluar conceptualmente desde una perspectiva más clínica o epidemiológica cuando hay comparaciones directas entre los diferentes tratamientos que conforman un ciclo cerrado en una red de estudios ABC (tabla 1). Cuando no hay comparaciones directas, una de las hipótesis establecidas al pretender calcular el efecto de C frente a B es que uno puede utilizar el conocimiento de C frente a B a través de un comparador común A. Esta aplicación del supuesto de transitividad permite decir que el comparador común A permite establecer una comparación válida entre los tratamientos. Para que se cumpla el supuesto de transitividad, no deben presentarse diferencias en relación con la distribución de factores modificadores del efecto entre los ensayos considerados. Ejemplos de posibles factores modificadores del efecto a nivel de estudio clínico pueden ser aquellos aspectos relacionados con el paciente (por ejemplo, la edad, el sexo, presencia de comorbilidades, etc.), la intervención (dosis, duración), el comparador y el resultado (diferencias en la definición del acontecimiento o enfermedad que se quiere prevenir o tratar). Por ejemplo, el no cumplimiento de la condición de transitividad puede producirse especialmente cuando se comparan tratamientos 'nuevos' respecto a tratamientos 'antiguos' debido a que pueden existir variables no observadas que pueden diferir entre las comparaciones, si bien es posible identificar *a priori* posibles factores modificadores del efecto y comparar sus distribuciones. También es posible que no se cumpla cuando los tratamientos han sido estudiados en indicaciones terapéuticas diferentes. Diversos autores⁸ formulan la condición de transitividad diciendo que si los estudios que comparan A con C no tienen un brazo B y los estudios que comparan

A con B no tienen un brazo C, puede considerarse que los brazos B o C han sido excluidos debidos al azar. De este modo, la transitividad no se cumplirá si la elección del grupo comparador A está asociado con el efecto relativo de las intervenciones. Esto pone de manifiesto la importancia de esta condición, ya que la elección de los tratamientos que se evalúan en los estudios son raramente elegidos al azar. La transitividad puede examinarse por clínicos e investigadores a través de una revisión cuidadosa de las características metodológicas de los estudios y de las poblaciones de pacientes estudiadas en estos estudios.

Además debe existir consistencia o coherencia entre la evidencia procedente de comparaciones directas e indirectas (tabla 1). Es decir, cuando podemos considerar que $\Theta_{BC, directa}$ es igual o similar a $\Theta_{BC, indirecta}$. La consistencia puede evaluarse estadísticamente, es decir, su diferencia puede cuantificarse a través de un factor de inconsistencia ($FI = \Theta_{BC, directa} - \Theta_{BC, indirecta}$) y utilizar una prueba Z para comprobar si $FI=0$ (hipótesis nula de consistencia) y/o calcular el intervalo de confianza para FI⁴. Es importante remarcar que la combinación de estimaciones inconsistentes procedentes de distintas fuentes puede proporcionar resultados sesgados que dejan de tener validez en los procesos de toma de decisiones. Al igual que sucede con la condición de transitividad, los estudios individuales combinados en las comparaciones directas e indirectas no deben presentar diferencias en relación a la distribución de los factores modificadores del efecto para que los resultados sean consistentes.

EJEMPLO CON LA HOJA DE CALCULO EN EXCEL

Consideraciones prácticas

Se ha implementado el cálculo de comparaciones indirectas y mixtas por el método de Bucher⁴, junto al factor de inconsistencia

en una hoja de cálculo Excel protegida, con fórmulas, objetos visuales (botones) y programación en Visual Basic for Applications (VBA) tanto de los objetos como de la propia hoja. Su funcionamiento se ha validado bajo distintos entornos de Microsoft Office, funcionando siempre correctamente. Aunque dependiendo de la versión de Microsoft Office, puede que el usuario necesite deshabilitar el uso de macros en Excel (anexo 1).

La hoja se puede descargar de forma gratuita en cualquiera de los URL que figuran en el anexo 1 o bien solicitándola por correo electrónico a los autores

Ejemplo para comparaciones indirectas

Para ilustrar el uso de la hoja de cálculo se utilizan datos de un reciente metaanálisis en red sobre intervenciones coronarias percutáneas en pacientes con enfermedad coronaria no aguda¹⁰. Los autores obtienen en primera instancia, a partir de un metaanálisis de comparaciones directas combinando los resultados de 3 ensayos clínicos (cuyos resultados se muestran en el anexo 2), un riesgo menor de revascularización coronaria en los pacientes tratados con stents convencionales (BMS) respecto los pacientes tratados con terapia médica (comparación identificada como B vs A en la hoja de cálculo, con riesgo relativo [RR]=0,78; intervalo de confianza del 95% [IC95%]=0,58-1,05). Mientras que a partir de otro metaanálisis de comparaciones directas para 7 ensayos clínicos obtienen un mayor riesgo de revascularización para pacientes con angioplastia coronaria transluminal percutánea (PTCA) respecto, de nuevo, a la terapia médica (comparación identificada como C vs A con RR=1,08; IC95%=0,74-1,56).

Asumiendo siempre el cumplimiento de la condición de transitividad, podemos obtener una comparación indirecta entre ambos tratamientos, BMS vs PTCA, desde el momento en que ambos se comparan en

sus respectivos metaanálisis con el mismo comparador común, la terapia médica.

Al abrir la hoja de cálculo, en primer lugar hay que elegir la medida del efecto que se desea utilizar (figura 1). En este ejemplo es el RR ubicado en la columna de variables de respuesta de tipo binario (revascularización si/no). A continuación, en el cuadro superior para el cálculo de comparaciones indirectas, hay que introducir los resultados procedentes de los metaanálisis de comparaciones directas anteriormente mencionados (figura 2). Los resultados muestran dos posibles comparaciones indirectas dependiendo del tratamiento que fijemos a priori como comparador, BMS vs PTCA (identificada como B vs C) o bien PTA vs BMS (identificada como C vs B). En el primer caso se observa un menor riesgo de revascularización en pacientes con BMS vs PTCA ($RR=0,72$; $IC95\%=0,45-1,16$). Lógicamente, en el segundo caso es la comparación inversa, que muestra un mayor riesgo de revascularización en pacientes con PTCA vs BMS ($R=1,38$; $IC95\%=0,86-2,23$). Aunque para ambas comparaciones

se observa que no existe una diferencia estadísticamente significativa ($p=0,18$).

Ejemplo para comparaciones mixtas

En el mismo estudio¹⁰, los autores también encontraron 3 investigaciones que comparaban de forma directa BMS vs PTCA (anexo 2). El resultado de este metaanálisis de comparaciones directas mostraba asimismo una disminución del riesgo de revascularización en los pacientes con BMS ($RR=0,79$; $IC95\%=0,59-1,06$). La comparación mixta, resultante de combinar la evidencia indirecta del apartado anterior con la nueva evidencia directa, se puede realizar introduciendo los datos de ambos tipos de comparaciones en el cuadro inferior para comparaciones mixtas (figura 3). Es importante destacar que los resultados que el usuario debe introducir han de ser coherentes en cuanto al orden del tratamiento de interés y el tratamiento que se utiliza como comparador, que para ser consistentes con los autores del ejemplo¹⁰ será BMS vs PTCA. El resultado de la comparación mixta indica que de nuevo hay un menor riesgo de revascularización en pacientes con BCM vs PTCA ($RR=0,77$;

Figura 1
Menú principal de la hoja Excel

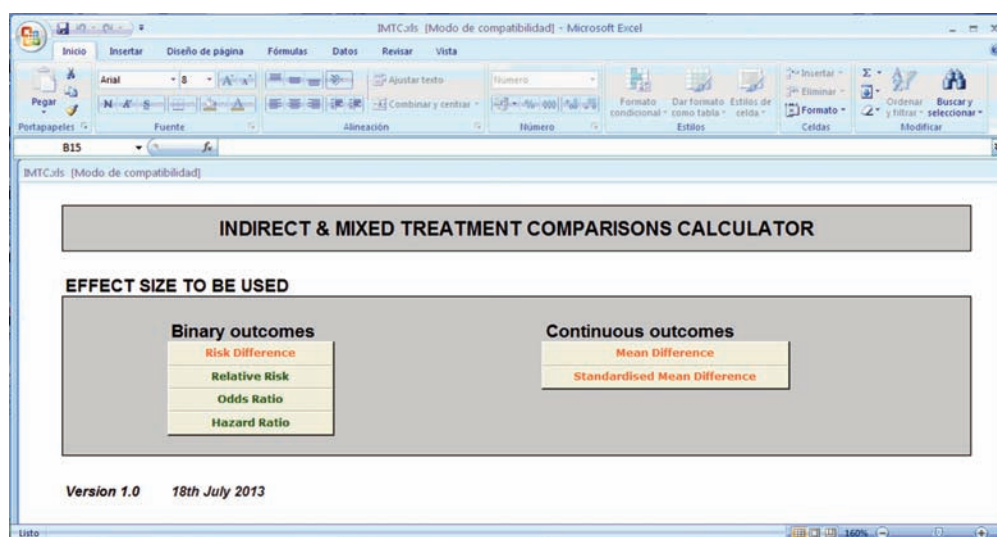


Figura 2
Cálculo de comparaciones indirectas con la hoja Excel.

CALCULATOR FOR INDIRECT COMPARISONS (for relative effects)					
		95% CI			
Direct comparisons	Effect size	Lower	Upper		
A vs B (B/A)	0.78	0.58	1.05		
A vs C (C/A)	1.08	0.74	1.56		
		95% CI			
Indirect comparison	Effect size	Lower	Upper	z	p
B vs C (C/B)	1.38	0.86	2.23	1.34	0.1808
C vs B (B/C)	0.72	0.45	1.16	1.34	0.1808

Los datos de las comparaciones directas (A vs. B, y A vs. C) se deben introducir manualmente (Riesgo Relativo, e intervalo inferior y superior del límite del confianza del 95%).

Los datos de ambas comparaciones directas serían calculados con otro software para metanálisis de comparaciones directas^{2,3} o bien obtenidos directamente de la revisión publicada¹⁰.

IC95%=0,61-0,99), aunque esta reducción del riesgo si que resulta estadísticamente significativa ($p=0,0397$). La hoja de cálculo también cuantifica el factor de consistencia en el caso del RR en términos relativos junto a su intervalo de confianza del 95% y la prueba Z para consistencia⁴. El resultado indica que hay evidencia para no rechazar la hipótesis nula de consistencia entre ambos tipos de comparaciones ($p=0,7440$), por lo que estadísticamente la comparación mixta es consistente.

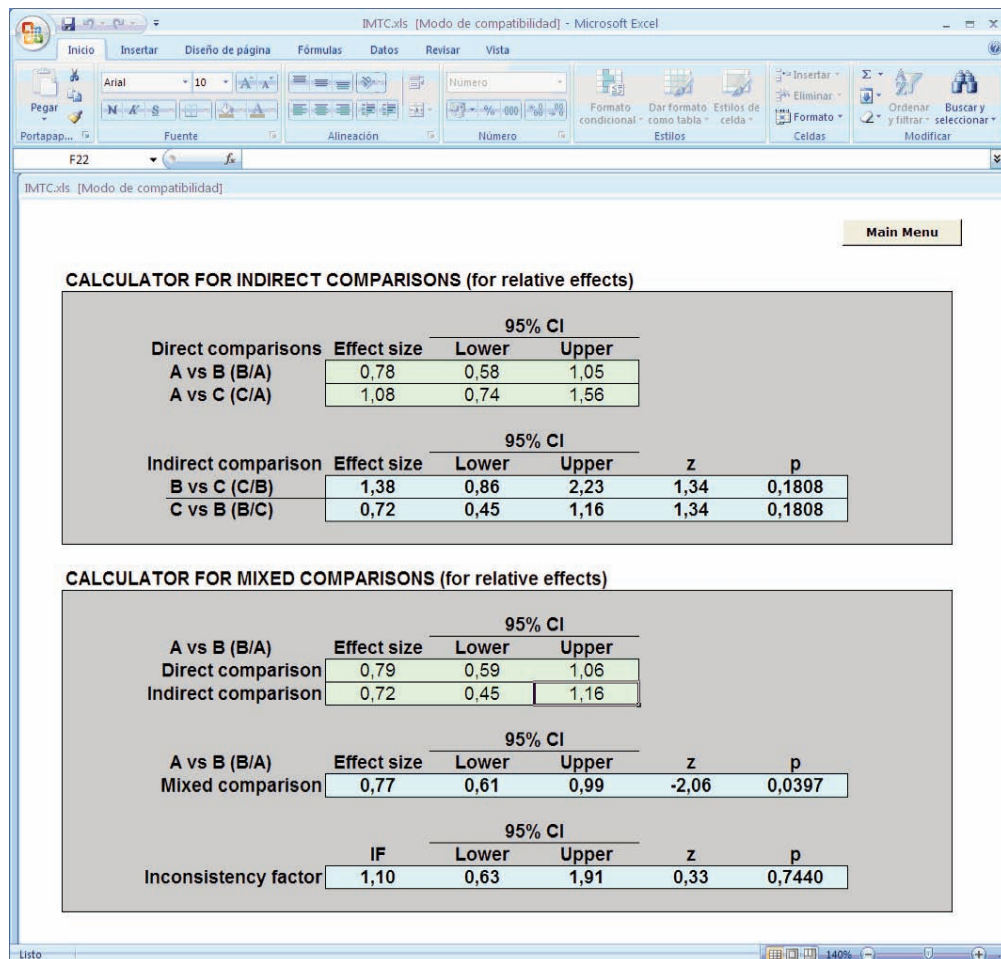
LIMITACIONES

Diversos estudios^{11,12} han puesto de manifiesto que las comparaciones indirectas y mixtas pueden sobre o infraestimar los efectos de los tratamientos en relación a la evidencia procedente de comparaciones directas, aunque otros autores sugieren lo contrario¹³. Asimismo, en la práctica existen comparaciones indirectas que podrían resultar válidas a pesar de no cumplirse las condiciones anteriormente

expuestas¹¹. Además no debemos olvidar las limitaciones comunes al metaanálisis de comparaciones directas, tales como la falta de homogeneidad, sesgo de publicación, sesgo de selección, etcétera. Por este motivo, actualmente estas técnicas siguen desarrollándose y evaluándose rigurosamente para aumentar el conocimiento disponible antes de su generalización¹⁴.

En cuanto a las limitaciones prácticas directamente relacionadas con la hoja de cálculo que presentamos en este trabajo, su aplicación al análisis de una red más extensa de los tres tratamientos que presentamos en el ejemplo, implica un aumento exponencial del número de cálculos a realizar a través de la propia hoja. En este caso, la solución habitual consiste en utilizar modelos de metaregresión, a partir de macros disponibles en programas estadísticos de carácter general, como Stata¹⁵ o R¹⁶. Asimismo, el método de Bucher implementado en la hoja de cálculo sólo es aplicable al metaanálisis

Figura 3
Cálculo de comparaciones mixtas con la hoja Excel



Los datos de las comparaciones directas e indirectas (B vs. C) se deben entrar de forma manual (Riesgo Relativo, e intervalo inferior y superior del límite del confianza del 95%).

Nota al pie: Los datos de la comparación indirecta podrían ser calculados utilizando la calculadora anterior para comparaciones indirectas (figura 2), mientras que los datos para la comparación directas serían calculados con otro software para metanálisis de comparaciones directas^{2,3} o bien obtenidos directamente de la revisión publicada¹⁰.

de ensayos de dos brazos⁴. En caso de tener ensayos con más de dos brazos es recomendable utilizar modelos de metaregresión multivariante¹⁷. Por último, en cualquiera de las dos situaciones anteriores puede resultar difícil la valoración de la condición de consistencia estadística de forma global. Para lo cual,

de nuevo, es recomendable utilizar modelos de inconsistencia como los de Lumley¹⁸ (para estudios con dos brazos de comparación) o Lu y Ades⁶ (para estudios con más de dos brazos). Estos modelos^{6,18} no permiten combinar comparaciones directas e indirectas, sino que su finalidad es contrastar la asunción

de consistencia estadística entre ambos tipos comparaciones. En caso de cumplirse la asunción de consistencia estadística, la estimación de las comparaciones directas, indirectas y mixtas puede obtenerse utilizando modelos de metaregresión (para estudios con dos brazos de comparación) o metaregresión multivariante (para estudios con más de dos brazos)¹⁷.

CONCLUSION

En este artículo se han presentado los aspectos metodológicos básicos de las comparaciones indirectas y mixtas y su implementación a través de una hoja de cálculo en Excel, con el fin de familiarizar con esta técnica al investigador de una forma sencilla, de tal forma que el uso de las comparaciones indirectas y mixtas permita ampliar el conocimiento proporcionado por la síntesis tradicional de la evidencia clínica a través del metaanálisis de comparaciones directas. Este tipo de comparaciones, y su posterior extensión al metaanálisis en red^{19,20}, son una aportación metodológica importante en la síntesis de la evidencia científica. Su potencial uso en epidemiología clínica es considerable ya que, además de proporcionar una estimación más precisa del efecto de un tratamiento respecto a alternativas terapéuticas disponibles, pueden ofrecer información relevante en situaciones en las que la evidencia procedente de comparaciones directas es escasa o inexistente. En el futuro, los resultados procedentes de este tipo de comparaciones pueden ser de gran utilidad para los procesos de evaluación y la toma de decisiones si se realizan de manera coherente, transparente y explícita, considerando las condiciones de transitividad y consistencia.

BIBLIOGRAFÍA

1. Ferreira González I, Urrútia G, Alonso-Coello P. Revisiones sistemáticas y metaanálisis: bases conceptuales e interpretación. *Rev Esp Cardiol*. 2011; 64: 688-96.
2. Kontopantelis E, Reeves D. MetaEasy: A Meta-Analysis Add-In for Microsoft Excel. *J Stat Software*. 2009; 30.
3. Neyeloff JL, Fuchs SC, Moreira LB. Meta-analyses and Forest plots using a microsoft Excel spreadsheet: step-by-step guide focusing on descriptive data analysis. *BMC Research Notes*. 2012; 5: 52.
4. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiology*. 1997; 50: 683-91.
5. Catalá-López F, Tobías A. Metaanálisis de ensayos clínicos aleatorizados, heterogeneidad e intervalos de predicción. *Med Clin (Barc)*. 2013. Disponible en: <http://dx.doi.org/10.1016/j.medcli.2013.06.013>
6. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23: 3105-24.
7. Baker SG, Kramer BS. The transitive fallacy for randomized trials: if A bests B and B bests C in separate trials, is A better than C? *BMC Med Res Methodol*. 2002; 2: 13. Erratum in: *BMC Med Res Methodol*. 2003; 3:23.
8. Catalá-López F, Hutton B, Moher D. Transitivity property across randomized controlled trials: If B is better than A, and C is better than B, will C be better than A? *Rev Esp Cardiol*. 2014 (en prensa).
9. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc*. 2006; 101: 447-59.
10. Trikalinos TA, Alsheikh-Ali AA, Tatsioni A, Nallamothu BK, Kent DM. Percutaneous coronary interventions for non-acute coronary artery disease: a quantitative 20-year synopsis and a network meta-analysis. *Lancet* 2009; 373: 911-8. Erratum in: *Lancet*. 2009; 374: 378.
11. Chou R, Fu R, Huffman LH, Korthuis PT. Initial highly-active antiretroviral therapy with a protease inhibitor versus a non-nucleoside reverse transcriptase inhibitor: discrepancies between direct and indirect meta-analyses. *Lancet*. 2006; 368: 1503-15.
12. Mills EJ, Ghement I, O'Regan C, Thorlund K. Estimating the power of indirect comparisons: a simulation study. *PLoS One*. 2011; 6: e16237.
13. Madan J, Stevenson MD, Cooper KL, Ades AE, Whyte S, Akehurst R. Consistency between direct and indirect trial evidence: is direct evidence always more reliable? *Value Health*. 2011; 14: 953-60.

14. Li T, Puhan MA, Vedula SS, Singh S, Dickersin K; Ad Hoc Network Meta-analysis Methods Meeting Working Group. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med.* 2011; 9: 79.
15. Harbord RM, Higgins JPT. Meta-regression in Stata. *Stata Journal* 2008; 8: 493-519.
16. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Soft.* 2010; 36.
17. Salanti G, Dias S, Welton NJ, Ades AE, Golfinopoulos V, Kyrgiou M, Mauri D, Ioannidis JP. Evaluating novel agent effects in multiple-treatments meta-regression. *Stat Med.* 2010; 29: 2369-83.
18. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med.* 2002; 21-2313-24.
19. Catalá-López F, Tobías A. Síntesis de la evidencia clínica y metaanálisis en red con comparaciones indirectas *Med Clin (Barc).* 2013; 140: 182-7.
20. Catalá-López F, Tobías A, Roqué M. Conceptos básicos del metaanálisis en red. *Aten Primaria.* 2014 (en prensa).

Anexo 1

URLs para el funcionamiento y descarga de la hoja IMTC.xls

Información de Microsoft para habilitar/deshabilitar el uso de macros en Excel:

<http://office.microsoft.com/es-mx/powerpoint-help/habilitar-o-deshabilitar-macros-en-documentos-de-office-HA010031071.aspx#BM12>

Direcciones para descargar la hoja de Microsoft Excel para el cálculo de comparaciones indirectas y metaanálisis en red:

- desde la web de Comparaciones Indirectas y Metaanálisis en Red:

<http://metaanalisenred.weebly.com/excel.html>

- desde la web del Centro Cochrane Iberoamericano:

http://www.cochrane.es/?q=es/elaborar_RS

Anexo 2

Comparaciones directas disponibles para los estudios incluidos en los ejemplos (a partir de Trikalinos et al 2009)¹⁰

Estudio (año)	Intervención (vs comparador)	Variable de resultado	RR (IC95%)
TOAT (2002)	BMS vs terapia médica	Revascularización	2,50 (0,70 – 8,80)
OAT (2006)	BMS vs terapia médica	Revascularización	0,80 (0,70 – 1,00)
COURAGE (2007)	BMS vs terapia médica	Revascularización	0,60 (0,60 – 0,80)
VA ACME (1992, 1998)	PTCA vs terapia médica	Revascularización	0,90 (0,60 – 1,40)
Sievers (1993)	PTCA vs terapia médica	Revascularización	0,80 (0,30 – 1,90)
MASS (1995, 1999)	PTCA vs terapia médica	Revascularización	5,20 (1,90 – 14,50)
VA ACME (2 vasos) (1997)	PTCA vs terapia médica	Revascularización	1,10 (0,90 – 1,40)
VA ACME (1 vaso) (1997)	PTCA vs terapia médica	Revascularización	19,5 (2,70 – 143)
RITA II (2003)	PTCA vs terapia médica	Revascularización	0,80 (0,60 – 0,90)
SWISS-II (2007)	PTCA vs terapia médica	Revascularización	0,60 (0,40 – 0,90)
EPISTENT (1999)	BMS vs PTCA	Revascularización	0,70 (0,50 – 0,80)
TOSCA (1999, 2005)	BMS vs PTCA	Revascularización	0,80 (0,60 – 1,00)
WIDEST (2000)	BMS vs PTCA	Revascularización	1,20 (0,80 – 1,90)

REVISIÓN SISTEMÁTICA Y METAANÁLISIS

NEUROTICISMO Y TRASTORNO POR ESTRÉS POSTRAUMÁTICO: UN ESTUDIO META-ANALÍTICO

Fabiola B Soler-Ferrería (1), Julio Sánchez-Meca (2), José Manuel López-Navarro (3) y Fernando Navarro-Mateu (4,5).

(1) Hospital Psiquiátrico Román Alberca. Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. El Palmar. Murcia.

(2) Departamento de Psicología Básica y Metodología. Facultad de Psicología. Universidad de Murcia.

(3) Centro de Salud Mental de Cieza. Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud. Murcia.

(4) Unidad de Docencia Investigación y Formación en Salud Mental (UDIF-SM). Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud

(5) CIBER de Epidemiología y Salud Pública (CIBERESP).

Los autores declaran la no existencia de conflictos de interés.

RESUMEN

Fundamentos: Existen resultados contradictorios sobre el papel del neuroticismo como factor de riesgo del trastorno por estrés posttraumático (TEPT). Los objetivos de este estudio fueron estimar la magnitud y dirección de la asociación entre neuroticismo y TEPT así como analizar la influencia de diferentes variables moderadoras sobre el tamaño del efecto.

Métodos: Se realizó una revisión sistemática y un meta-análisis. La búsqueda de estudios se efectuó en Medline, IME, PsycINFO, Trip Database y Google Scholar, siendo el límite temporal hasta octubre de 2012. También se hizo una búsqueda manual y se contactó con los autores principales. Criterios de inclusión: (a) estudios empíricos que analizaran la asociación entre neuroticismo y TEPT mediante escalas validadas o criterios diagnósticos (DSM, CIE); (b) en sujetos expuestos a un evento traumático y (c) escritos en inglés o español. Dos evaluadores independientes realizaron la selección y codificación de los datos siguiendo el protocolo previamente elaborado. Se aplicó el modelo de efectos aleatorios para obtener el tamaño del efecto medio y explorar variables moderadoras.

Resultados: De 96 artículos elegibles, 34 cumplieron los criterios de inclusión (9.941 participantes). El tamaño del efecto medio fue $r_s=0,371$ (IC95%: 0,327-0,414). Se descartó que hubiera sesgo de publicación. Se construyó un modelo predictivo con tres variables (reporte de la etnia, edad y tipo de muestra: comunitaria o clínica).

Conclusión: El neuroticismo puede considerarse un factor de riesgo para el TEPT en personas que han estado expuestas a un evento traumático. Estos hallazgos son importantes para la elaboración de actuaciones preventivas y de intervención.

Palabras clave: Trastorno por estrés posttraumático. Neuroticismo. Trastornos de ansiedad. Salud mental. Revisión sistemática. Meta-análisis. Medicina basada en evidencia.

Correspondencia

Dr. Julio Sánchez Meca

Departamento de Psicología Básica y Metodología

Facultad de Psicología

Campus de Espinardo

Universidad de Murcia

30100 Murcia.

jsmeca@um.es

<http://www.um.es/metaanalysis>

ABSTRACT

Neuroticism and Post-Traumatic Stress Disorder: A Meta-analytic Study

Background: There are conflicting results on the role of neuroticism as a risk factor for post-traumatic stress disorder (PTSD). The objectives of this study were to estimate the magnitude and direction of the association between neuroticism and PTSD, and to analyze the influence of different moderator variables on effect size.

Methods: A systematic review and a meta-analysis were carried out. The search for studies was conducted in Medline, IME, PsycINFO, Trip Database, and Google Scholar, until October 2012. A hand search was also carried out and main researchers were contacted. Inclusion criteria: (a) empirical studies of the association between neuroticism and PTSD using validated scales or diagnostic criteria (DSM, ICD), (b) in participants exposed to a traumatic event and (c) written in English or Spanish. Two independent evaluators performed the selection and extracted the data following a previously developed protocol. The random-effects model was applied to obtain the mean effect size and to explore moderators.

Results: Out of 96 potential articles, 34 fulfilled the inclusion criteria (9,941 participants). The mean effect size was $r_s = 0.371$ (95%CI: 0.327 and 0.414). Publication bias was discarded as a threat. A predictive model was elaborated with three variables (reporting of ethnicity, age, and type of sample).

Conclusion: Neuroticism can be considered as a risk factor for PTSD in people who have been exposed to a traumatic event. These findings are relevant for developing preventive interventions and treatments.

Keyword: Post-Traumatic Stress Disorders. Neuroticism. Anxiety disorders. Mental health. Meta-analysis. Systematic review. Evidence-Based Medicine.

INTRODUCCIÓN

En los últimos años, el trastorno por estrés postraumático (TEPT) ha ganado un reconocimiento progresivo como un problema de salud pública¹ por su alta prevalencia así como por tener un alto impacto personal y social y generar incapacidad y altos costes en salud, estos últimos²⁻⁴ tanto de manera directa sobre las personas afectadas como sobre sus familiares⁵. Así, según el estudio *European Study of the Epidemiology of Mental Disorders* (ESEMED), la prevalencia en población general en España a lo largo de la vida es del 1,95% (IC del 95%: 1,18-2,73)⁶ y del 1,9% en Europa (IC del 95%: 1,7-2,1)⁷.

El interés por el estudio del TEPT se ha incrementado en los últimos años a causa del impacto personal, social y mediático de las catástrofes naturales, de los accidentes y de los actos de terrorismo y guerras. Por ejemplo, en Madrid, tras los atentados del 11 de marzo de 2004, la cifra total de TEPT relacionado con ellos fue de un 2,3% en la población general de la ciudad, con un 4,3% en las zonas próximas a las explosiones⁸.

Uno de los factores de riesgo descritos para el desarrollo del TEPT tras la exposición a un evento estresante son determinados rasgos de personalidad, entre ellos, el neuroticismo⁹. Conceptualizado en la actualidad como la tendencia relativamente estable a responder con emociones negativas ante situaciones de miedo, estrés, frustración o pérdida, parece ser un predictor significativo para los síntomas de TEPT y problemas generales de salud entre las víctimas de diferentes desastres¹⁰⁻¹⁴. Recientemente se han revisado las implicaciones en salud pública del neuroticismo al ser identificado como un sólido correlato y predictor de diversos trastornos mentales (sobre todo depresión y ansiedad) y de la morbimortalidad de diversas enfermedades físicas (diabetes, trastornos cardiovasculares, problemas dermatológicos, respiratorios y gas-

trointestinales e, incluso, del cáncer) y de la comorbilidad entre ellas asociada a un mayor número de complicaciones, un aumento en la frecuencia del uso de los servicios sanitarios y peores pronóstico y calidad de vida¹⁵.

La literatura científica ha estudiado la relación entre neuroticismo y TEPT en diferentes poblaciones y tras distintos tipos de eventos traumáticos¹⁶⁻²⁸, sin embargo, existen hallazgos contradictorios. Mientras que unos estudios sugieren que el neuroticismo es un factor de riesgo para desarrollar TEPT²⁹⁻³¹ otros no han encontrado esta relación^{32,33}. Los estudios meta-analíticos permiten analizar en profundidad este tipo de situaciones en las que existen resultados discrepantes en la literatura científica³⁴.

Los objetivos de este estudio fueron: estimar la magnitud y dirección de la asociación entre neuroticismo y TEPT tras la exposición a un evento traumático, analizar la influencia de diferentes variables moderadoras sobre el tamaño del efecto y proponer un modelo predictivo con las variables moderadoras que expliquen mejor la variabilidad en los tamaños del efecto reportados por cada estudio.

MATERIAL Y MÉTODOS

Diseño. Se realizó una revisión sistemática con meta-análisis³⁴. Para su publicación se siguieron las indicaciones de la guía *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA)³⁵.

Criterios de selección de los estudios. Los criterios de selección para la inclusión de los trabajos encontrados en la revisión fueron:

- que se tratara de estudios empíricos (casos-contrroles, cohortes o transversales) que informaran de medidas de neuroticismo y de sintomatología o diagnósticos de TEPT valorados a través de escalas validadas o del

cumplimiento de los criterios diagnósticos según DSM o CIE.

- que utilizaran los mismos instrumentos para valorar el TEPT y el neuroticismo en toda la muestra expuesta al evento traumático.

- que informaran de manera cuantitativa de la relación directa entre neuroticismo y TEPT en muestras de sujetos en las que al menos el 80% habían estado expuestos a un evento traumático.

- y que estuvieran publicados en inglés o español.

Búsqueda de los estudios. Para optimizar la localización de los estudios que cumplieran con los criterios de selección se realizó una búsqueda sistemática hasta octubre de 2012 en las siguientes bases bibliográficas de datos: MEDLINE, IME, PsycINFO, Trip Database y Google Académico. Las palabras-clave utilizadas fueron: “Neuroticism AND (PTSD OR post-traumatic stress disorder)”, las cuales se buscaron en texto libre y no en *Medical Subject Headings* (MeSH). También se contactó mediante correo electrónico con los autores de referencia solicitando el envío de los estudios cuyos resultados no hubieran sido publicados y se revisó manualmente la bibliografía de los artículos incluidos en busca de artículos no identificados previamente.

La valoración de los criterios de inclusión fue realizada en dos fases por dos evaluadores independientes y a ciegas. En la primera, se preseleccionaron los artículos revisando los títulos y los resúmenes. En la segunda, se revisó el texto completo de los estudios preseleccionados. Los desacuerdos se resolvieron por consenso.

Extracción de los datos. La extracción de los datos de los estudios incluidos finalmente se realizó también por dos evaluadores independientes y a ciegas utilizando un

protocolo de codificación, con un manual previamente elaborado y consensuado³⁶. Las variables moderadoras se clasificaron en variables extrínsecas, de contexto, de los participantes y metodológicas. Como variables extrínsecas se codificaron el año de publicación del estudio y el estatus de publicación -publicado vs no publicado. La única variable de contexto que se codificó fue el país en el que se realizó el estudio. De los participantes se codificaron la edad media de la muestra (en años), la distribución por sexo (porcentaje de mujeres), la población de referencia (comunitaria vs clínica), el porcentaje de la muestra expuesto a un evento traumático (no pudiendo ser inferior al 80%), el tipo de evento traumático, el porcentaje de la muestra diagnosticado de TEPT, el porcentaje de la muestra con comorbilidad y la distribución étnica de la muestra. Las variables metodológicas fueron el diseño general de estudio (casos-contróles, de cohortes o transversal), si la asociación neuroticismo-TEPT era o no un objetivo principal del estudio, el diseño específico de la relación neuroticismo-TEPT (longitudinal vs transversal), el tamaño muestral inicial y final (considerando en ambos casos sólo a los sujetos que hubieran estado expuestos a un evento traumático), el instrumento utilizado para valorar el TEPT (transcurrido al menos un mes desde el evento) y el utilizado para valorar el neuroticismo, el tipo de medida del TEPT (diagnóstico vs sintomatología), el tiempo transcurrido entre el evento traumático y la valoración del TEPT, el tiempo transcurrido entre el evento y la valoración del neuroticismo, el momento en que se valoró el neuroticismo (antes vs después del evento traumático) y si el evento traumático fue o no el mismo para toda la muestra (homogeneidad vs heterogeneidad del evento traumático).

Evaluación de la calidad de los estudios incluidos. Para realizarla se elaboró una escala *ad hoc* a partir de una revisión de varias escalas propuestas en la literatura³⁷. Nuestra escala incluyó los siguientes crite-

rios valorados de forma dicotómica (Sí/No): (1) reporte de la etnia (si el estudio reportó o no la distribución étnica de los participantes), (2) elegibilidad (si se definieron con claridad los criterios de elegibilidad de la muestra), (3) homogeneidad del evento (si toda la muestra estuvo expuesta al mismo tipo de evento traumático), (4) comorbilidad psiquiátrica (si se reportó o no la presencia de comorbilidad psiquiátrica en los participantes), (5) muestra expuesta (si el 100% de la muestra estuvo expuesta a un evento traumático), (6) neuroticismo evaluado sin sesgo (si el neuroticismo se midió antes del evento traumático o al menos un mes después el evento), (7) incluido en los objetivos (si la relación entre el neuroticismo y el TEPT era uno de los objetivos principales del estudio), (8) representatividad (si la tasa de respuesta fue mayor del 80% y la mortalidad de los que aceptaron respecto al tamaño muestral final fue menor del 20%) y (9) diseño longitudinal (si el diseño respecto a la relación entre neuroticismo y TEPT fue prospectivo midiendo neuroticismo previamente a la valoración del TEPT). Se calculó una Puntuación Global de Calidad sumando 1 punto por cada ítem positivo (rango 0-9), si bien también se analizó la influencia de cada uno de los criterios de forma independiente, como la existencia de variables potencialmente moderadoras de la relación neuroticismo-TEPT.

Índice del tamaño del efecto. Para la obtención del tamaño del efecto individual de cada uno de los estudios seleccionados se utilizó el coeficiente r correlación de Pearson entre neuroticismo y TEPT. En los estudios en los que no se aportaba directamente este coeficiente se realizaron las transformaciones pertinentes para su obtención (a partir de una diferencia de medias o de una odds ratio). Cuando más de una medida del mismo constructo era utilizada en un mismo estudio, se priorizaron las puntuaciones obtenidas por las escalas *Penn Inventory*³⁸ y la Escala Revisada del Impacto del Evento o *Impact of Event Scale-Revised*³⁹ (IES-R) al ser las más recomendadas para valorar la sintomatología de TEPT².

Para evaluar la fiabilidad de la codificación de las variables moderadoras y de los tamaños del efecto, dos evaluadores realizaron de forma independiente la extracción de datos. Se calculó el coeficiente kappa de Cohen para las variables cualitativas y el coeficiente de correlación intraclass (CCI) para las continuas. En ambos se consideraron como satisfactorios valores iguales o superiores a 0,60.

Análisis estadístico. En primer lugar, se realizó un análisis descriptivo de las principales características de los estudios. Posteriormente, se llevó a cabo un análisis meta-analítico para la obtención del tamaño del efecto medio de la asociación entre neuroticismo y TEPT. Todos estos análisis se realizaron asumiendo modelos de efectos aleatorios, dada la alta variabilidad observada entre los estudios. Para la integración estadística, las correlaciones fueron transformadas a Z de Fisher, Z_r , con objeto de normalizar su distribución y estabilizar su varianza. Después de los análisis, los valores Z de Fisher se transformaron de nuevo a la métrica del coeficiente de correlación de Pearson para facilitar su interpretación. Valores positivos de r reflejaron que a mayor nivel de neuroticismo mayor nivel de sintomatología de TEPT y viceversa. Se calculó el tamaño del efecto medio, su intervalo de confianza al 95% (IC95%) y el contraste de su significación estadística con la prueba z . Para examinar la heterogeneidad de los tamaños del efecto se construyó un *forest plot* (con los tamaños del efecto de cada estudio, el tamaño del efecto medio y sus límites confidenciales al 95%) y se calculó el estadístico Q de Cochran y el índice I^2 ⁴⁰. Este índice se interpreta como el porcentaje de la variación total que no puede explicarse por el mero error de muestreo aleatorio, sino por la existencia de características diferenciales en los estudios meta-analizados. Valores de I^2 en torno al 25%, 50% y 75% implican una heterogeneidad baja, moderada y alta, respectivamente.

Para comprobar si el sesgo de publicación podía amenazar la validez de los resultados, se construyó un *funnel plot* de los tamaños del efecto contra sus errores típicos, con el fin de observar si el aspecto del gráfico resultante se alejaba mucho de la simetría, con el método *trim-and-fill* de Duval y Tweedie⁴¹, se aplicó el test de Egger⁴² y el índice de tolerancia a los resultados nulos⁴³.

Para el análisis de las variables moderadoras cualitativas se aplicaron ANOVAs y para las continuas modelos de meta-regresión, todos ellos asumiendo un modelo de efectos mixtos. Se examinó el influjo de cada variable moderadora con el estadístico Q_B , para los ANOVAs y con Z para los modelos de meta-regresión. Finalmente, se propuso un modelo de meta-regresión múltiple que incluyera el subconjunto de variables moderadoras más relevantes para explicar la variabilidad de los tamaños del efecto. La selección de estos predictores se hizo en función de que presentara una relación estadísticamente significativa con el tamaño del efecto, fueran conceptualmente relevantes y que apenas tuvieran datos faltantes.

Todas las pruebas estadísticas se interpretaron asumiendo un nivel de significación del 5% ($p=0,05$) y un contraste bilateral. Los análisis estadísticos se realizaron con las macros elaboradas por David B. Wilson para el paquete estadístico SPSS y con el programa *Comprehensive Meta-analysis 2.0*⁴⁴.

Este estudio no recibió financiación específica y no necesitó ser valorado por un Comité de Ética de Investigación Sanitaria al utilizar la información de estudios ya publicados.

RESULTADOS

Proceso de búsqueda, selección y codificación. La búsqueda electrónica localizó 203 referencias únicas a las que se sumaron 7 localizadas en la bibliografía de las priemras (figura 1). A pesar de nuestros intentos,

no se localizó ningún estudio no publicado. En total, se identificaron 210 artículos, de los cuales se seleccionaron los 34 artículos que cumplían los criterios de inclusión^{2,11,21,27,29,45-73} (36 estudios, ya que dos artículos^{46,50} aportaron dos muestras diferentes de participantes, que permitieron obtener dos estimaciones independientes de la relación entre neuroticismo y TEPT). En el análisis de la fiabilidad del proceso de selección se obtuvo un coeficiente kappa de Cohen de 0,589 para la valoración de los criterios de inclusión del 87% de los artículos localizados, y coeficientes kappa entre 0,635 y 0,965 en la extracción de los datos de los estudios preseleccionados, excepto para el ítem que valoraba el momento de evaluación del neuroticismo que fue de 0,431. Las discrepancias entre los codificadores se resolvieron por consenso.

Análisis descriptivo. En la tabla 1 se recogen las características de los 36 estudios incluidos en el meta-análisis. En conjunto dieron un tamaño muestral final global de 9.941 participantes. El 83,3% fueron publicados a partir del año 2000 y en distintos países, sobre todo ubicados en Europa (44,4%) y América del Norte (38,9%). El porcentaje de caucásicos presentó una mediana del 78% (rango: 0-93,2%; media: 66,4%; DT: 32,41%; $k=18$). Respecto del tipo de muestra en 32 (88,9%) estudios los participantes procedían de la población general mientras que en 4 (11,1%) se trataba de pacientes procedentes de la clínica (pacientes de salud mental o que pedían atención en dicho servicio).

En cuanto a la valoración de la calidad de los estudios, se cumplían los criterios, salvo tres ítems: 18 (50%) incluían reporte de la etnia, 15 (41,7%) la valoración del neuroticismo sin sesgo y 9 (25%) el diseño longitudinal para el estudio de la relación entre neuroticismo y TEPT. La puntuación global de calidad obtuvo una media de 5,72 (mediana: 6; DT: 1,14) y un rango de 3 a 8 (tabla 2).

Figura 1

Diagrama del proceso de búsqueda y selección de los estudios en el meta-análisis de la asociación entre neuroticismo y Trastorno de Estrés Postraumático

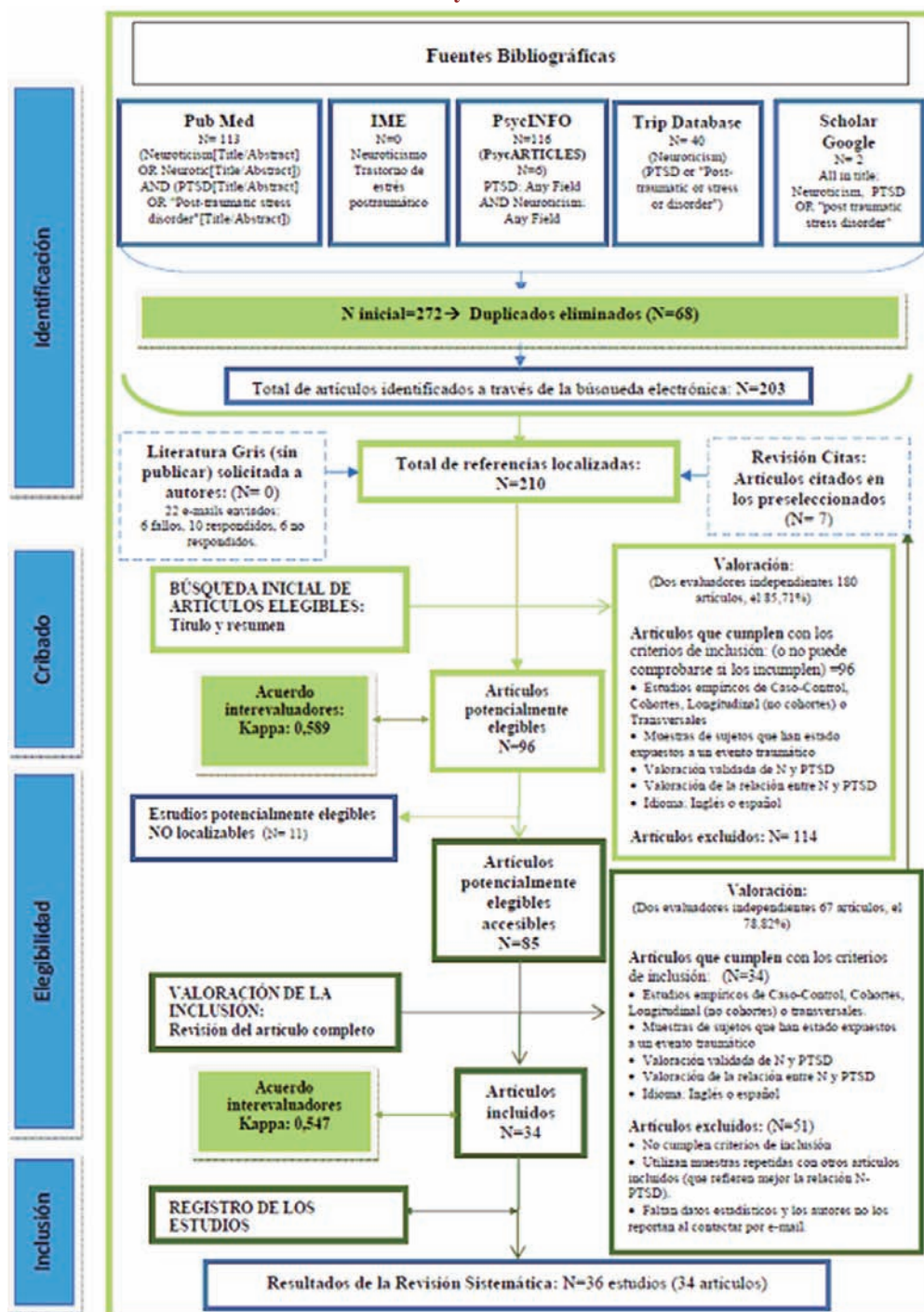


Tabla 1

Descripción de las características de los estudios incluidos en el meta-análisis de la asociación entre neuroticismo y TEPT

Estudios	Muestra				TEPT				Escala de neuroticismo [†]	Diseño Relación N-TEPT	Tamaño del efecto (r)
	n final	Mujeres (%)	Edad media	Expuesta (%)	Tipo de Medida	Instrumentos valoración TEPT i	Trauma	Diagnóstico TEPT (%)			
McFarlane, 1988	45	-	34,4	100	Diagnóstico	Entrevista clínica según DSM-III	Bomberos en una actuación ante un incendio (desastre)	24,4	EPI	T	0,488
Breslau et al, 1991	394	61,7	26	100	Diagnóstico	DIS	Heterogéneo	23,6	EPQ-R	T	0,341
Hyer et al, 1994	51	0	43,8	100	Síntomas	IES (MCRS, MMPI: PK y PS)	Veteranos de guerra (experiencias militares)	100	NEO-PI	T	0,110
Tranah y Farmer, 1994	64	0	36,5	100	Diagnóstico	Entrevista clínica según DSM-III-R	Conductores de metro que han visto a una persona caer/saltar delante.	20,3			
	EPQ	Transversal	0,558								
Carr et al, 1997	680	50,7	43,4	100	Síntomas	IES	Desastre natural (terremoto)	-	EPI	L	0,490
Sembi et al, 1998	61	-	66,4	100	Síntomas (Ambos)	PI (IES-A, CAPS)	Ataque isquémico transitorio o infarto cerebral	9,8	EPQ-R-SS	T	0,640
Fauerbach et al, 2000	70	21,1	33,4	100	Diagnóstico	SCID-NP	Supervivientes a quemaduras	25,7	NEO-PI	L	0,234
Kozaric-Kovacic et al, 2000	70	0	27,8	100	Diagnóstico	Entrevista clínica según ICD-10 y DSM-III-R	Experiencias militares	51,4	EPQ-R	CC	0,656
Stewart et al, 2000	291	100	38,7	96,3	Síntomas	PSS-SR	-	63	NEO-FFI	T	0,260
Holeva y Tarrier, 2001	265	58	35,2	100	Síntomas						
(Ambos)	PI	Accidentes de tráfico	23	EPQ-R	L	0,438					
Lauterbach y Vrana, 2001	402	44	-	80	Síntomas						
(Ambos)	PPTSD-R	Heterogéneo	3,2	EPQ-R	T	0,390					
Pedersen et al., 2003	112	30	60	100	Diagnóstico	PDS	IAM	22	EPQ-R-SS	T	0,072
Brodsky et al, 2004	100	59	74,1	-	Diagnóstico	Entrevista clínica según DSM-IV	Supervivientes del Holocausto	-	EPI	T	0,567
Cox et al, 2004a	1485	100	33,9	100	Diagnóstico	DIS	Heterogéneo		The five-item neuroticism index	T	
Cox et al. 2004a	1485	100	33,9	100	Diagnóstico	DIS	Heterogéneo		The five-item neuroticism index	T	

Tabla 1
(continuación)

Estudios	Muestra				TEPT				Escala de neuroticismo [†]	Diseño Relación N-TEPT	Tamaño del efecto (r)
	n final	Mujeres (%)	Edad media	Expuesta (%)	Tipo de Medida	Instrumentos valoración TEPT *	Trauma	Diagnóstico TEPT (%)			
Cox et al, 2004b	1.753	0	33,9	100	Diagnóstico	DIS	Heterogéneo	-	The five-item neuroticism index	T	0,327
Dimic et al, 2004	-	40,5	39,4	100	Diagnóstico	Entrevista clínica según DSM-IV	Eventos de guerra (como civiles o militares)	56,9	NEO-PI-R	T	0,281
Enrique ,2004	60	0	38,9	100	Diagnóstico	PTCI	Eventos de guerra	50	NEO-PI-R	T	0,170
Van den Hout y Engelhard, 2004	117	100	31	100	Síntomas	PSS	Aborto espontáneo.	-	EPQ	L	0,250
Engelhard y van de Hout, 2007	308	3	22,4	100	Síntomas	PSS (Y SCID).	Soldados enviados a Iraq (experiencias militares)	-	EPQ (short)	L	0,300
Michalski, 2007	409	76,1	19	96,1	Síntomas	IES-R (CMS, PI)		10,8	EPI	T	0,487
Qouta et al, 2007	65	52	17	100	Síntomas	PTSD-RI	Experiencias militares	-	EPI	L	0,280
Ribi et al, 2007	139	-	-	100	Síntomas	PDS	Padres de niños con lesiones o enfermedades graves	23	NEO-FFI	T	0,330
Dörfel et al, 2008	44	59,1	31,9	100	Síntomas	CAPS	Accidentes de tráfico	9,1	EPI	T	0,580
Boelen, 2009	254	88,6	42,2	100	Síntomas	PSS-SR	Muerte de un ser querido	-	EPQ	T	0,500
Borja et al, 2009a	51	76,5	19,8	100	Síntomas	DTS	Desastre Natural	-	NEO-PI-R	T	0,200
Borja et al, 2009b	35	97,1	21	100	Síntomas	DTS	Asalto sexual	-	NEO-PI-R	T	
Crostley, 2009	92	84,8	39	100	Síntomas	IES-R	-	100	NEO-FFI	T	0,190
Hatcher et al, 2009	102	18,6	45,7	100	Síntomas	IES	Lesión de la médula espinal	-	PANAS	T	0,620
Mason et al, 2009	367	48,3	36,9	100	Diagnóstico	IES-R	Lesiones por accidentes	-	EPI	L	0,441
Marshall-Berenz et al, 2010	81	63,1	23,4	-	Síntomas	CAPS	Heterogéneo	4,7	BFI	T	0,370
Wang et al, 2010	119	13,4	26,7	-	Síntomas	PCL-C	Su experiencia más estresante durante el servicio como policía	-	NEO-FFI	L	0,227
Caska y Renshaw, 2011	214	2,8	35,1	100	Síntomas	PCL-M	Soldados con experiencias de guerra	-	BFI	T	0,430

Tabla 1
(continuación)

Estudios	Muestra				TEPT				Escala de neuroticismo †	Diseño Relación N-TEPT	Tamaño del efecto (r)
	n final	Mujeres (%)	Edad media	Expuesta (%)	Tipo de Medida	Instrumentos valoración TEPT *	Trauma	Diagnóstico TEPT (%)			
Engelhard et al. (2011)	138	0	23,8	100	Síntomas	PSS	Soldados enviados a Afganistán (experiencias militares)	2	EPQ	L	0,390
Frazier et al. (2011)	236	86	19,5	100	Síntomas	PCL-S	-	-	Escala de 5 ítems del N	T	0,290
LaFauci Schutt y Marotta (2011)	197	35	46	100	Síntomas	PCL-C	Experiencias en trabajos de atención en emergencias.	13,2	NEO-FFI	T	0,550
Wu et al. (2011)	968	-	18,6	100	Síntomas	IES-R	Tormenta de Nieve	14,5	EPQ	T	0,312

* **Instrumentos utilizados para valorar TEPT** (entre paréntesis los que no fueron incluidos en el meta-análisis). DIS: Diagnostic Interview Schedule⁸⁵; MCRS: Mississippi Scale for Combat-Related Posttraumatic Stress Disorder⁸⁶; CMS: Civilian Mississippi Scale⁸⁷; MMPI: The Minnesota Multiphasic Personality Inventory-2 (MMPI-2;⁸⁸ (PK: Defensiveness, y PS: Superlative Self-Presentation); IES: La Escala de Impacto del Evento (Impact of Event Scale⁸⁹; IES-R: La escala Revisada del Impacto del Evento (Impact of Event Scale-Revised³⁹; IES-A: La Escala de Impacto del Evento- Evitación (Impact of Event Scale-Avoidance⁸⁹ Subescala de Ansiedad. CAPS: Escala De TEPT Administrada por un Clínico (Clinician-Administered PTSD Scale⁹⁰; PI: Penn Inventory (38); SCID: Structured Clinical Interview for the DSM-IV (91); SCID-I/P: Structured Clinical Interview for DSM-IV, patient version⁹²; SCID-NP: Structured Clinical Interview for DSM-IV, non-patient version⁹³; PSS: Escala de Síntomas Postraumáticos (Posttraumatic Symptom Scale⁹⁴; PSS-SR: Escala de Síntomas Postraumáticos, versión auto-informada (Self-reported version of the Posttraumatic Symptom Scale⁹⁴; PPTSD-R: Purdue PTSD Scale-Revised⁹⁵; PTCI: Cuestionario de Cogniciones Postraumáticas (Posttraumatic Cognition Inventory⁹⁶; PDS: La Escala de Diagnóstico Postraumático (Posttraumatic Diagnostic Scale⁹⁷; PTSD-RI: The adolescent version of the Reaction Index⁹⁸; DTS: Escala de Trauma de Davidson. (DTS: Davidson Trauma Scale⁹⁹; PCL-C: The PTSD Checklist-Civilian¹⁰⁰; PCL-M PTSD Checklist-Military (PCL-C¹⁰⁰) PCL-S: The PTSD Checklist –Specific Version^{92,100}.

† **Escalas de neuroticismo:** EPI: Inventario de personalidad de Eysenck (Eysenck Personality Inventory¹⁰¹), EPQ: Cuestionario de Personalidad de Eysenck (The Eysenck Personality Questionnaire¹⁰²; EPQ-R: Cuestionario de la Personalidad de Eysenck Revisado (The Revised Eysenck Personality Questionnaire¹⁰³); EPQ-R-SS Cuestionario de Personalidad de Eysenck Revisado-Escala abreviada (The Eysenck Personality Questionnaire-R Short Scale¹⁰⁴); NEO-PI: Inventario de Personalidad NEO (The NEO Personality Inventory; (105), NEO-FFI: Inventario de Personalidad NEO-Reducido (NEO-Five Factor Inventory; (106); NEO-PI-R: Inventario de Personalidad NEO-Revisado (The Revised NEO Personality Inventory¹⁰⁷; PANAS: Escalas de Afecto Positivo y Negativo¹⁰⁸ BFI: Big Five Inventory¹⁰⁹, The five-item Neuroticism index(110); Escala de 5 ítems del neuroticismo¹¹¹.

T: Transversal. L: Longitudinal.

Tabla 2
Evaluación de la calidad de los estudios incluidos en el meta-análisis de la asociación entre neuroticismo y TEPT

Estudio	Valoración de la etnia	Criterios de elegibilidad	Evento homogéneo	Mide comorbilidad psiquiátrica	Muestra expuesta	Neuroticismo sin sesgo	Objetivo	Representatividad	Diseño N-TEPT longitudinal	Puntuación total (0-9)
McFarlane,1988	No	Sí	Sí	Sí	Sí	Sí	Sí	Sí	No	7
Breslau et al, 1991	Sí	Sí	No	Sí	Sí	No	Sí	Sí	No	6
Hyer et al, 1994	Sí	Sí	Sí	No	Sí	No	Sí	No	No	5
Tranah y Farmer, 1994	No	Sí	Sí	Sí	Sí	Sí	No	Sí	No	6
Carr et al, 1997	No	Sí	Sí	Sí	Sí	Sí	Sí	No	Sí	7
Sembi et al, 1998	No	Sí	Sí	Sí	Sí	Sí	No	No	No	5
Fauerbach et al, 2000	Sí	Sí	Sí	Sí	Sí	No	Sí	No	Sí	7
Kozaric-Kovacic et al, 2000	No	Sí	Sí	No	Sí	Sí	Sí	Sí	No	6
Stewart et al, 2000	No	Sí	No	Sí	No	No	No	Sí	No	3
Holeva y Tarrier, 2001	No	Sí	Sí	No	Sí	No	Sí	No	Sí	5
Lauterbach y Vrana, 2001	Sí	No	No	Sí	No	No	Sí	Sí	No	4
Pedersen et al, 2003	No	Sí	Sí	No	Sí	Sí	Sí	No	No	5
Brodaty et al, 2004	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	No	8
Cox et al, 2004a	Sí	Sí	No	No	Sí	No	Sí	Sí	No	5
Cox et al, 2004b	Sí	Sí	No	No	Sí	No	Sí	Sí	No	5
Dimic et al, 2004	No	Sí	Sí	No	Sí	No	Sí	Sí	No	5
Enrique, 2004	No	Sí	Sí	No	Sí	No	Sí	Sí	No	5
Van den Hout y Engelhard, 2004	No	Sí	Sí	No	Sí	Sí	Sí	Sí	Sí	7
Engelhard y van den Hout, 2007	No	Sí	No	No	Sí	Sí	Sí	Sí	Sí	6
Michalski, 2007	No	Sí	No	Sí	No	No	Sí	Sí	No	4
Qouta et al, 2007	Sí	Sí	Sí	Sí	Sí	No	Sí	No	Sí	7
Ribi et al, 2007	Sí	Sí	Sí	No	Sí	Sí	Sí	Sí	No	7
Dörfel et al, 2008	No	Sí	Sí	Sí	Sí	Sí	Sí	Sí	No	7
Boelen, 2009	No	Sí	Sí	Sí	Sí	No	No	No	No	4
Borja et al, 2009a	Sí	Sí	Sí	Sí	Sí	No	Sí	Sí	No	7
Borja et al, 2009b	Sí	Sí	Sí	Sí	Sí	No	Sí	Sí	No	7
Hatcher et al, 2009	No	Sí	Sí	No	Sí	Sí	Sí	No	No	5
Mason et al, 2009	No	Sí	Sí	Sí	Sí	No	Sí	No	Sí	6
Crostley, 2009	Sí	Sí	No	No	Sí	No	Sí	Sí	No	5
Marshall-Berenz et al, 2010	Sí	Sí	No	Sí	Sí	No	No	Sí	No	5
Wang et al, 2010	Sí	Sí	No	Sí	No	Sí	No	Sí	Sí	6
Caska y Renshaw, 2011	Sí	Sí	Sí	No	Sí	No	Sí	Sí	No	6
Engelhard et al, 2011	No	Sí	Sí	No	Sí	Sí	No	No	Sí	5
Frazier et al, 2011	Sí	Sí	No	Sí	Sí	No	Sí	Sí	No	6
LaFauci Schutt y Marotta, 2011	Sí	Sí	Sí	No	No	No	Sí	No	No	5
Wu et al, 2011	Sí	Sí	Sí	No	Sí	Sí	Sí	Sí	No	7

Tamaño medio del efecto. El tamaño medio del efecto (asociación entre neuroticismo y TEPT) fue de $r_+ = 0,371$ (IC 95%: 0,327-0,414), lo que reflejó la existencia de una relación positiva entre neuroticismo y TEPT. La prueba Z resultó estadísticamente significativa [$Z=15,23$; $p<0,001$]. Sin embargo, se observó una elevada heterogeneidad entre los estudios [$Q(35)=176,74$; $p<0,001$; $I^2=80,2\%$] que también se hizo patente en el *forest plot* (figura 2).

Análisis del sesgo de publicación. En la figura 3 se observa que no hubo asimetría en el *funnel plot* de los tamaños del efecto contra sus errores típicos. La técnica *trim-and-fill* de Duval y Tweedie⁴¹ de imputación de datos perdidos sobre el *funnel plot* no tuvo que imputar ningún dato perdido para simetrizarlo. Así mismo, el test de Egger no alcanzó la significación estadística [$b_0=0,59$; $T(34)=0,81$; $p=0,424$]. Finalmente, el índice de tolerancia a los resultados

Figura 2
Forest plot del meta-análisis (modelo de efectos aleatorios) de la asociación entre neuroticismo y TEPT

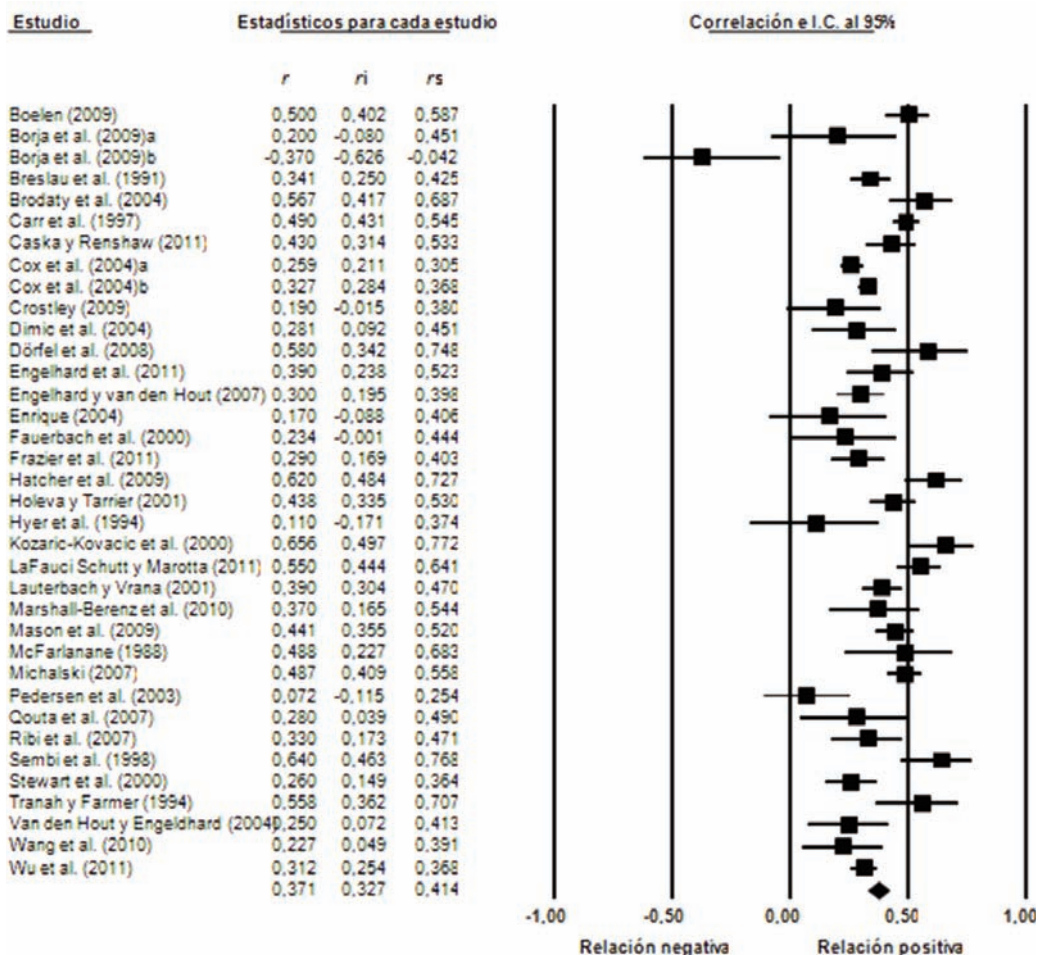
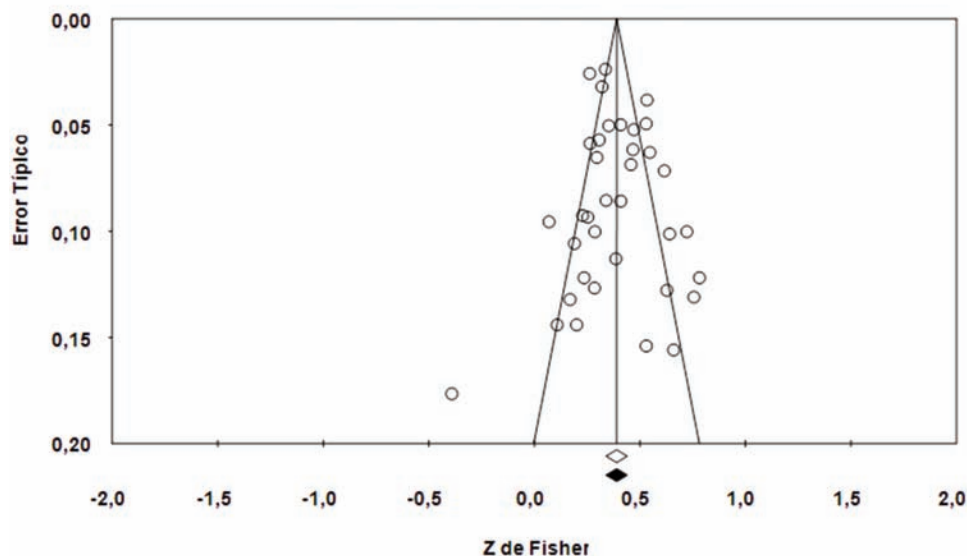


Figura 3
Funnel plot de los 36 coeficientes de correlación de Pearson
(transformados a Z de Fisher)



nulos de Orwin⁴⁵ alcanzó un valor $N=231$ estudios, superior al índice de seguridad ($N_{fs}=5k+10=5 \times 35+10=190$). Estos resultados permiten descartar el sesgo de publicación como una amenaza contra la validez de los resultados de este meta-análisis.

Análisis de variables moderadoras.

Las que presentaron una relación estadísticamente significativa con el tamaño del efecto fueron el continente en el que se había realizado el estudio, que presentó una relación estadísticamente significativa con el tamaño del efecto [$Q_B(3)=12,61$; $p<0,05$; $R^2=0,14$], de forma que Oceanía obtuvo el mayor tamaño del efecto medio ($r_+=0,514$; $k=3$) seguido de Europa ($r_+=0,423$; $k=16$), América ($r_+=0,303$, $k=15$) y Asia ($r_+=0,279$; $k=2$). Varias variables moderadoras relacionadas con las características de los participantes alcanzaron la significación estadística. En primer lugar, la población de procedencia de la muestra del estudio [$Q_B(1)=4,71$; $p<0,05$;

$R^2=0,034$], siendo el tamaño del efecto mayor cuando la muestra procedía de la población general ($r_+=0,387$, $k=32$) que cuando se trataba de una muestra de pacientes ($r_+=0,223$; $k=4$). La edad media de la muestra también alcanzó la significación estadística, de manera que a mayor edad hubo mayor asociación entre neuroticismo y TEPT ($b_f=0,005$; $Z=2,29$; $p<0,05$; $k=34$; $R^2=0,06$). Por el contrario, el porcentaje de la muestra diagnosticada de TEPT presentó una relación negativa con el tamaño del efecto, de forma que a mayor porcentaje de participantes diagnosticados fue menor el tamaño del efecto encontrado ($b_f=-0,003$; $Z=-2,30$; $p<0,05$; $k=22$; $R^2=0,11$). El resto de las variables de los participantes no resultaron estadísticamente significativas (proporción de mujeres, porcentaje de caucásicos, porcentaje de comorbilidad psiquiátrica, porcentaje de expuestos) como tampoco lo fue el año de publicación del estudio.

Varias variables metodológicas resultaron ser estadísticamente significativas. Una de ellas fue el método diagnóstico para valorar el TEPT puesto en relación con el neuroticismo [$Q_B(2)=6,88$; $p<0,05$; $R^2=0,2$], siendo mayor el tamaño medio del efecto cuando se utilizaron ambos métodos ($r_+=0,656$; $k=1$), seguido de la entrevista clínica ($r_+=0,363$; $k=8$) y finalmente por la escala estandarizada ($r_+=0,326$; $k=4$). El instrumento para valorar neuroticismo también presentó una relación estadística con el tamaño del efecto [$Q_B(2)=12,03$; $p<0,05$; $R^2=0,097$], siendo mayor la correlación cuando se utilizó alguna versión del cuestionario de personalidad de Eysenck ($r_+=0,427$; $k=19$) seguido por Otros (BFI, 5 ítems, PANAS; $r_+=0,373$; $k=6$) y menor con las del NEO ($r_+=0,256$; $k=11$). El tiempo transcurrido desde el evento traumático hasta la valoración del neuroticismo presentó una asociación positiva estadísticamente significativa con el tamaño del efecto ($b_I=0,002$; $Z=2,17$; $p<0,05$; $k=14$; $R^2=0,25$), de manera que a mayor número de meses transcurridos entre el evento traumático y la medición del neuroticismo fue mayor la asociación encontrada entre neuroticismo y TEPT.

En relación a la valoración de la calidad, tres ítems resultaron estadísticamente significativos. En primer lugar, los estudios que reportaron la distribución de la etnia [$Q_B(1)=7,08$, $p<0,05$; $R^2=0,2$] obtuvieron correlaciones menores ($r_+=0,317$, $k=18$) que los que no la reportaron ($r_+=0,426$, $k=18$). En segundo lugar, la valoración del neuroticismo sin sesgo [$Q_B(1)=4,17$; $p<0,05$] reflejó que los estudios que cumplieran este criterio obtuvieron mayores tamaños del efecto ($r_+=0,426$; $k=15$) que los que no lo cumplieran ($r_+=0,377$; $k=21$). Finalmente, cuando se cumplió la representatividad de la muestra [$Q_B(1)=4,37$; $p<0,05$] se obtuvieron mayores tamaños del efecto ($r_+=0,403$, $k=24$) que en los estudios que no la cumplieran ($r_+=0,300$; $k=12$). Sin embargo, para estas dos últimas variables la relación fue débil.

El resto de las variables de calidad no resultaron estadísticamente significativas (ni los criterios ni la puntuación total de calidad).

El tipo de diseño utilizado en el estudio fue una variable metodológica de especial relevancia para valorar la posible existencia sesgos en las estimaciones de los efectos. De los 36 estudios, 7 utilizaron un diseño de cohortes, tres de caso-control, 20 aplicaron un diseño transversal y 6 utilizaron un diseño longitudinal que no era de cohortes. Al comparar los efectos medios de estos cuatro tipos de diseño no se observaron diferencias estadísticamente significativas [$Q_B(3)=2,59$; $p=0,46$; $R^2=0$], presentando los diseños de cohortes y de caso-control el mismo efecto medio ($r_+=0,319$), los transversales un efecto medio $r_+=0,380$ y los longitudinales que no eran de cohortes un efecto medio $r_+=0,422$.

Modelo predictivo. Con objeto de proponer un modelo predictivo capaz de explicar buena parte de la variabilidad de la relación encontrada entre neuroticismo y TEPT, se seleccionaron las siguientes variables moderadoras: (a) tipo de muestra: población general (32 estudios) vs clínica (4 estudios); (b) edad media en años y (c) el reporte de la distribución étnica en la muestra (18 estudios) vs no reporte (18 estudios). El número de predictores estuvo limitado por el número de estudios disponibles ($k=36$). De los 36 estudios, 2 no reportaron la edad media^{63,68}. Para poder incluirlos se les imputó la edad media de las muestras comunitarias (34,19 años), ya que ambos utilizaron este tipo de muestra.

El reporte de la etnia vs no reporte fue la variable que tuvo la mayor relación estadísticamente significativa con el tamaño del efecto, con un 20% de varianza explicada. Sin embargo, se comprobó que esta variable estaba fuertemente correlacionada (o confundida) con el lugar geográfico en el que se realizó la investigación, de forma que la mayoría de los estudios europeos no repor-

Tabla 3**Resultados del modelo de meta-regresión múltiple de efectos mixtos aplicado sobre los tres predictores finales**

Variable moderadora	b_j	ET	Z	p
Intercepción	0,308	0,111	2,78	0,005
Población	-0,221	0,075	-2,96	0,003
Edad media	0,005	0,002	2,45	0,015
Etnia	0,113	0,045	2,51	0,012
Modelo completo:	$Q_R(3)=20,76; p<0,001; R^2=0,36$ $Q_E(32)=110,86; p<0,001$			

b_j : coeficiente de regresión parcial. ET: error típico. Z: prueba Z_p : nivel de probabilidad. Q_R : estadístico de ajuste global del modelo. Q_E : estadístico de contraste del error de especificación del modelo. R^2 : proporción de varianza explicada.

taron la etnia de sus muestras, mientras que la mayoría de los estudios americanos sí [$RV(4)=19,162; p=0,001; Phi=0,682$].

Tomando estos predictores aplicamos un modelo de meta-regresión de efectos mixtos sobre los tamaños del efecto (tabla 3). El modelo completo alcanzó una relación estadísticamente significativa con el tamaño del efecto, con un 35,9% de varianza explicada. Los tres predictores alcanzaron una relación estadísticamente significativa una vez parcializado el influjo de los otros predictores: la población de referencia ($p=0,003$), con las muestras comunitarias reflejando una relación más fuerte entre neuroticismo y TEPT que las clínicas, la edad media de la muestra ($p=0,015$), siendo más fuerte la relación entre neuroticismo y TEPT a mayor edad, y el reporte de la etnia ($p=0,012$), siendo más fuerte la relación entre neuroticismo y TEPT cuando ésta no se reportó. No obstante, el modelo no alcanzó un nivel adecuado de especificación [$Q_E(32)=110,86; p<0,001$], lo que indica la existencia de otras variables moderadoras de la fuerza de la relación entre neuroticismo y TEPT.

DISCUSIÓN

Nuestro estudio es el primer meta-análisis que investiga la asociación entre neuroticismo y TEPT. Sus resultados sugieren que el neuroticismo se asocia a la posibilidad de desarrollar un TEPT tras haber estado expuesto a un evento traumático, en consonancia con lo encontrado en publicaciones anteriores^{11,12,18}. A través de 36 estudios recoge una amplia muestra de personas adultas (9.941 sujetos) de ambos sexos y de diferentes países a lo largo de 4 continentes (América, Europa, Asia y Oceanía), expuestos a diferentes tipos de eventos traumáticos y, por tanto, sus resultados son generalizables a la población general. El tamaño medio del efecto mostró una correlación positiva de intensidad media-alta según el criterio de Cohen ($r_+=0,371$). Los análisis del sesgo de publicación descartaron su posible amenaza contra los resultados del meta-análisis. Sin embargo, la alta heterogeneidad entre los tamaños del efecto de los diferentes estudios ($I^2=80,19\%$) implica la existencia de posibles variables moderadoras. El modelo predictivo propuesto explicó un 35,9% de la varianza, siendo los predictores el tipo de muestra (comunitaria vs clínica), la edad media y si existía o no el reporte de la etnia. Así, se esperaría una mayor correlación en muestras comunitarias, de mayor edad y en estudios que no reportan la etnia (es decir, en estudios probablemente europeos). Se ha descrito un menor tamaño del efecto entre neuroticismo y TEPT en muestras procedentes de población general frente a muestras clínicas en un meta-análisis en el que se analiza la relación entre los rasgos de personalidad con los trastornos depresivos, trastornos de ansiedad y abuso de sustancias⁷⁴. Esta situación podría deberse a que se utilizaron instrumentos diagnósticos aplicados mediante entrevista clínica, mientras que en nuestro meta-análisis la mayoría de los estudios se centraron en la valoración de síntomas de TEPT mediante autoinformes.

Por otra parte, observamos que el reporte de la etnia se asoció de forma estadísticamente significativa con el continente de procedencia del estudio, de manera que los que reportaban la etnia procedían mayoritariamente de América del Norte, mientras que los que no lo hacían procedían mayoritariamente de Europa. El análisis del porcentaje de caucásicos no presentó una relación estadísticamente significativa con el tamaño del efecto, por lo que la variable reporte de la etnia estaría enmascarando diferencias culturales que sí podrían modular la expresión de determinados rasgos de personalidad, como el neuroticismo y su asociación ante el TEPT tras la exposición a un evento estresante.

Los resultados de este estudio deben tener en cuenta diversas limitaciones. En primer lugar, el hecho de que la mayoría de los estudios aplicara un diseño transversal limita la interpretación causal de la asociación entre neuroticismo y TEPT. Sin embargo, los estudios prospectivos incluidos en el meta-análisis también señalan en esta dirección, así como otros estudios posteriores⁷⁵. Además, el diseño específico de la relación entre neuroticismo y TEPT (longitudinal vs transversal) no obtuvo diferencias en los tamaños del efecto, dando más solidez a la relación encontrada. Por otra parte, la teoría de los rasgos de personalidad plantea que estos se mantienen estables en el tiempo. La correlación entre neuroticismo y TEPT podría deberse a un aumento del neuroticismo posterior al evento^{76,77}. Sin embargo, cuanto más elevado sea el neuroticismo previo a la exposición más riesgo habría de que alcance niveles asociados al desarrollo de síntomas de TEPT.

En segundo lugar, la ambigüedad en el reporte de algunos datos y el escaso número de publicaciones que incluyen algunas de las variables moderadoras impidieron su adecuado análisis. Por últi-

mo, la reciente publicación de los criterios diagnósticos DSM-V⁷⁸ incluye un cambio en la conceptualización del TEPT al separarlo de los trastornos de ansiedad e incluirlo en un nuevo capítulo referido al trauma y los trastornos relacionados según el estresor y algunos cambios en los síntomas asociados⁷⁹⁻⁸². Esta nueva conceptualización no está exenta de críticas⁸³ y tales cambios podrían afectar a los hallazgos y a la comprensión de este trastorno⁸⁴. Las repercusiones del DSM-V en este sentido tendrán que ser estudiadas en el futuro.

Los resultados del meta-análisis a favor de una asociación entre neuroticismo y TEPT suponen un nuevo apoyo empírico a la importancia de analizar este rasgo de la personalidad en diversas patologías mentales y físicas¹⁵. La detección de personas con mayor riesgo de desarrollar TEPT resulta de gran utilidad en diversos contextos. Por ejemplo, en los procesos de selección/admisión de personal expuesto a eventos altamente estresantes (militares, policías, bomberos, etcétera), así como en la orientación vocacional. El neuroticismo, y sus variables moderadoras identificadas, debería tomarse en consideración a la hora de plantear guías para la prevención del TEPT en personas que pueden exponerse a experiencias traumáticas por sus circunstancias laborales y en personas que hayan sufrido un evento traumático y que puedan necesitar apoyo para superar el trauma sin desarrollar un TEPT.

BIBLIOGRAFÍA

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. DSM-IV-TR. 4th, Rev. ed. Washington, DC: American Psychiatric Association; 2000.
2. Michalski R. Assessment of personal predictive variables and symptom expression in posttraumatic stress disorder. Dissertation Abstracts International: Section B: The Sciences and Engineering. 2007;67(10-B).

3. Westphal M, Olfson M, Gameroff MJ, Wickramaratne P, Pilowsky DJ, Neugebauer R, et al. Functional impairment in adults with past posttraumatic stress disorder: findings from primary care. *Depress Anxiety*. 2011;28(8):686-95.
4. Inoue-Sakurai C, Maruyama S, Morimoto K. Posttraumatic stress and lifestyles are associated with natural killer cell activity in victims of the Hanshin-Awaji earthquake in Japan. *Prev Med*. 2000;31(5):467-73.
5. de Burgh HT, White CJ, Fear NT, Iversen AC. The impact of deployment to Iraq or Afghanistan on partners and wives of military personnel. *Int Rev Psychiatry*. 2011;23(2):192-200.
6. Haro JM, Palacin C, Vilagut G, Martínez M, Bernal M, Luque I, et al. Prevalence of mental disorders and associated factors: results from the ESEMeD-Spain study. *Med Clin (Barc)*. 2006;126(12):445-51.
7. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl* 2004;(420):21-7.
8. Cano-Vindel A, Miguel-Tobal JJ, González-Ordi H, Iruarrizaga I. Los atentados terroristas del 11-M en Madrid: la proximidad de la residencia a las áreas afectadas. *Ansiedad y Estrés* 2004;10:181-94.
9. Lawrence JW, Fauerbach JA. Personality, coping, chronic stress, social support and PTSD symptoms among adult burn survivors: a path analysis. *J Burn Care Rehabil*. 2003;24(1):63-72.
10. Charlton PF, Thompson JA. Ways of coping with psychological distress after trauma. *Br J Clin Psychol*. 1996;35(Pt 4):517-30.
11. Hyer L, Braswell L, Albrecht B, Boyd S, Talbert S. Relationship of NEO-PI to personality styles and severity of trauma in chronic PTSD victims. *J Clin Psychol*. 1994;50(5):699-707.
12. Engelhard IM, van den Hout MA, Kindt M. The relationship between neuroticism, pre-traumatic stress, and post-traumatic stress: A prospective study. *Pers Individ Dif*. 2003;35(2):381-8.
13. Bowman ML. Individual differences in posttraumatic distress: Problems with the DSM-IV model. *Can J Psychiatry*. 1999;44(1):21-33.
14. Watson D, Pennebaker JW. Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychol Rev*. 1989;96(2):234-54.
15. Lahey BB. Public health significance of neuroticism. *Am Psychol*. 2009;64(4):241-56.
16. O'Toole BI, Marshall RP, Schureck RJ, Dobson M. Risk factors for posttraumatic stress disorder in Australian Vietnam veterans. *Aust N Z J Psychiatry*. 1998;32(1):21-31.
17. Schnurr PP, Friedman MJ, Rosenberg SD. Premilitary MMPI scores as predictors of combat-related PTSD symptoms. *Am J Psychiatr*. 1993;150(3):479-83.
18. Hyer L, Braswell L, Albrecht B, Boyd S, Boudewyns P, Tabert S. Relationship of NEO-PI to personality styles and severity of trauma in chronic PTSD victims. *J Clin Psychol*. 2003;59:1295-304.
19. Bramsen I, Dirkzwager AJ, van der Ploeg HM. Predeployment personality traits and exposure to trauma as predictors of posttraumatic stress symptoms: a prospective study of former peacekeepers. *Am J Psychiatr*. 2000;157(7):1115-9.
20. Morgan I, Matthews G, Winton M. Coping and Personality as Predictors of Post-Traumatic Intrusions, Numbing, Avoidance and General Distress: A Study of Victims of the Perth Flood. *Behav Cogn Psychother*. 1995;23(03):251-64.
21. McFarlane AC. The aetiology of post-traumatic stress disorders following a natural disaster. *Br J Psychiatry*. 1988;152:116-21.
22. McFarlane AC. The longitudinal course of posttraumatic morbidity. The range of outcomes and their predictors. *J Nerv Ment Dis*. 1988;176(1):30-9.
23. Thompson J, Solomon M. Body recovery teams at disasters: trauma or challenge? *Anxiety Res*. 1991;4:235-44.
24. Weiss DS, Marmar CR, Metzler TJ, Ronfeldt HM. Predicting symptomatic distress in emergency services personnel. *J Consult Clin Psychol* 1995;63(3):361-8.
25. Chung MC, Berger Z, Jones R, Rudd H. Posttraumatic stress disorder and general health problems following myocardial infarction (Post-MI PTSD) among older patients: the role of personality. *J Geriatr Psychiatry*. 2006;21(12):1163-74.
26. Chung MC, Berger Z, Rudd H. Comorbidity and personality traits in patients with different levels of posttraumatic stress disorder following myocardial infarction. *Psychiatry Res* 2007;152(2-3):243-52.

27. Breslau N, Davis GC, Andreski P, Peterson E. Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Arch Gen Psychiatry*. 1991;48(3):216-22.
28. Chung MC, Dennis I, Easthope Y, Werrett J, Farmer S. A multiple-indicator multiple-cause model for post-traumatic stress reactions: personality, coping, and maladjustment. *Psychosomatic Med*. 2005;67(2):251-9.
29. Wu D, Yin H, Xu S, Zhao Y. Risk factors for post-traumatic stress reactions among Chinese students following exposure to a snowstorm disaster. *BMC Public Health*. 2011;12:11-96.
30. Engelhard IM, van den Hout MA, Schouten EG. Neuroticism and low educational level predict the risk of posttraumatic stress disorder in women after miscarriage or stillbirth. *Gen Hosp Psychiatry*. 2006;28(5):414-7.
31. Kuljic B, Miljanovic B, Svcevic R. Posttraumatic stress disorder in Bosnian war veterans: analysis of stress events and risk factors. *Vojnosanit Pregl*. 2004;61(3):283-9.
32. Fitzharris M, Fildes B, Charlton J. Anxiety, acute and post-traumatic stress symptoms following involvement in traffic crashes. *Annual proceedings of the Ann Adv Automot Med*. 2006;50:297-315.
33. Engelhard IM, van den Hout MA, Weerts J, Hox JJ, van Doornen LJP. A prospective study of the relation between posttraumatic stress and physical health symptoms. *Int J Clin Health Psychol*. 2009;9(3):365-72.
34. Sánchez-Meca J, Marín-Martínez F. Meta-analysis. En: Peterson P, Baker E, McGaw B, editors. *International Encyclopedia of Education*. 3rd ed. Oxford: Elsevier; 2010. p. 274-82.
35. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Med*. 2009;6(7):e1000097.
36. Soler-Ferrera FB. Neuroticismo y Trastorno por Estrés Posttraumático: Un Estudio Meta-analítico. [Tesis doctoral]. Universidad de Murcia; 2013.
37. Jarde A, Losilla JM, Vives J. Methodological quality assessment tools of non-experimental studies: a systematic review. *An Psicol*. 28: 617-628 2012;28:617-28.
38. Hammarberg M. PennInventory for Posttraumatic Stress Disorder: Psychometric properties. *Psychol Assess*. 1992;4:67-76.
39. Weiss DS, Marmar CR. The Impact of Event Scale-Revised. En: J.P.Wilson, T.M.Keane, editors. *Assessing Psychological Trauma and PTSD: A Practitioner's Handbook*. New York: Guilford Press; 1997. p. 399-411.
40. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539-58.
41. Duval S, Tweedie R. Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*. 2000;56(2):455-63.
42. Sterne JAC, Egger M. Regression methods to detect publication and other bias in meta-analysis. En: Rothstein HR, Sutton AJ, Borenstein M, editors. *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Chichester: Wiley; 2005. p. 99-100.
43. Orwin RG. A fail-safe N for effect size in meta-analysis. *J Educ Stat*. 1983;8:157-9.
44. Borenstein M, Hedges L, Higgins J, Rothstein H. *Comprehensive Meta-analysis Version 2*. Englewood NJ: Biostat; 2005.
45. Boelen PA. The centrality of a loss and its role in emotional problems among bereaved people. *Behav Res Ther*. 2009;47(7):616-22.
46. Borja SE, Callahan JL, Rambo PL. Understanding negative outcomes following traumatic exposure: The roles of neuroticism and social support. *Psychol Trauma*. 2009;1(2):118-29.
47. Brodaty H, Joffe C, Luscombe G, Thompson C. Vulnerability to post-traumatic stress disorder and psychological morbidity in aged holocaust survivors. *Int J Geriatr Psychiatry*. 2004;19(10):968-79.
48. Carr VJ, Lewin TJ, Kenardy JA, Webster RA, Hazell PL, Carter GL, et al. Psychosocial sequelae of the 1989 Newcastle earthquake: III. Role of vulnerability factors in post-disaster morbidity. *Psychol Med*. 1997;27(1):179-90.
49. Caska CM, Renshaw KD. Personality traits as moderators of the associations between deployment experiences and PTSD symptoms in OEF/OIF service members. *Anxiety Stress Coping* 2013;26(1):36-51.
50. Cox BJ, MacPherson PSR, Enns MW, McWilliams LA. Neuroticism and self-criticism associated with posttraumatic stress disorder in a nationally representative sample. *Behav Res Ther*. 2004;42(1):105-14.

51. Crostley JT. Mediatonal roles of personality factors and vengeful rumination in the development of posttraumatic stress disorder. *Dissertation Abstracts International: Section B: Sci Eng.* 2009;70 (12-B).
52. Dimic S, Tosevski DL, Jankovic JG. The Relationship between Personality Dimensions and Posttraumatic Stress Disorder. *Psihijatrija Danas* 2004;36(1):39-50.
53. Dörfel D, Rabe S, Karl A. Coping strategies in daily life as protective and risk factors for post traumatic stress in motor vehicle accident survivors. *J Loss Trauma.* 2008;13(5):422-40.
54. Engelhard IM, van den Hout MA. Preexisting neuroticism, subjective stressor severity, and posttraumatic stress in soldiers deployed to Iraq. *Canad J Psychiatry* 2007;52(8):505-9.
55. Engelhard IM, Olatunji BO, de Jong PJ. Disgust and the development of posttraumatic stress among soldiers deployed to Afghanistan. *J Anxiety Disorders* 2011;25(1):58-63.
56. Enrique A. Neuroticismo, extraversión y estilo atribucional en veteranos de guerra: Una aproximación desde el estrés posttraumático. *Interdisc Rev Psicol Ciencias Afines* 2004;21(2):213-46.
57. Fauerbach JA, Lawrence JW, Schmidt CW, Jr., Munster AM, Costa Jr. PT. Personality predictors of injury-related posttraumatic stress disorder. *J Nerv Mental Dis.* 2000;188(8):510-7.
58. Frazier PA, Gavian M, Hirai R, Park C, Tennen H, Tomich P, et al. Prospective predictors of posttraumatic stress disorder symptoms: Direct and mediated relations. *Psychol Trauma.* 2011;3(1):27-36.
59. Hatcher MB, Whitaker C, Karl A. What predicts post-traumatic stress following spinal cord injury? *Br J Health Psychol.* 2009;14(3):541-61.
60. Holvea V, Tarrier N. Personality and peritraumatic dissociation in the prediction of PTSD in victims of road traffic accidents. *J Psychosom Res.* 2001;51(5):687-92.
61. Kozaric-Kovacic D, Ljubin T, Rutic-Puz L, Mihaljevic J, Jernej B, Cicin-Sain L. Platelet monoamine oxidase activity, ego strength, and neuroticism in soldiers with combat-related current posttraumatic stress disorder. *Croatian Med J.* 2000;41(1):76-80.
62. LaFauci Schutt JM, Marotta SA. Personal and environmental predictors of posttraumatic stress in emergency management professionals. *Psychol Trauma.* 2011;3(1):8-15.
63. Lauterbach D, Vrana S. The relationship among personality variables, exposure to traumatic events, and severity of posttraumatic stress symptoms. *J Traum Stress* 2001;14(1):29-45.
64. Marshall-Berenz EC, Vujanovic AA, Bonn-Miller MO, Bernstein A, Zvolensky MJ. Multime-thod study of distress tolerance and PTSD symptom severity in a trauma-exposed community sample. *J Traum Stress.* 2010;23(5):623-30.
65. Mason S, Farrow TFD, Fawbert D, Smith R, Bath PA, Hunter M, et al. The development of a clinically useful tool for predicting the development of psychological disorder following injury. *B J Clin Psychol.* 2009;48(1):31-45.
66. Pedersen SS, Middel B, Larsen ML. Posttraumatic stress disorder in first-time myocardial infarction patients. *Heart Lung.* 2003;32(5):300-7.
67. Qouta S, Punamaki RL, Montgomery E, El-Sarraj E. Predictors of psychological distress and positive resources among Palestinian adolescents: Trauma, child, and mothering characteristics. *Child Abuse Negl.* 2007;31(7):699-717.
68. Ribi K, Vollrath ME, Sennhauser FH, Gnehm HE, Landolt MA. Prediction of posttraumatic stress in fathers of children with chronic diseases or unintentional injuries: a six-months follow-up study. *Child Adolesc Psychiatry Ment Health.* 2007;1:1-16.
69. Sembi S, Tarrier N, O'Neill P, Burns A, Faragher B. Does post-traumatic stress disorder occur after stroke: A preliminary study. *Int J Geriatr Psychiatry.* 1998;13(5):315-22.
70. Stewart SH, Conrod PJ, Samoluk SB, Pihl RO, Dongier M. Posttraumatic stress disorder symptoms and situation-specific drinking in women substance abusers. *Alcohol Treat Q.* 2000;18(3):31-47.
71. Tranah T, Farmer RD. Psychological reactions of drivers to railway suicide. *Soc Sci Med.* 1994;38(3):459-69.
72. van den Hout M, Engelhard I. Pretrauma Neuroticism, Negative Appraisals of Intrusions, and Severity of PTSD Symptoms. *J Psychopathol Behav Assess.* 2004;26(3):181-3.
73. Wang Z, Inslicht SS, Metzler TJ, Henn-Haase C, McCaslin SE, Tong H, et al. A prospective study of predictors of depression symptoms in police. *Psychiatric Res.* 2010;175(3):211-6.

74. Kotov R, Gamez W, Schmidt F, Watson D. Linking "big" personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychol Bull.* 2010;136(5):768-821.
75. Breslau N, Schultz L. Neuroticism and post-traumatic stress disorder: a prospective investigation. *Psychol Med.* 2012;30:1-6.
76. Roberts BW, Walton KE, Viechtbauer W. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychol Bull.* 2006;132(1):1-25.
77. Jeronimus BF, Ormel J, Aleman A, Penninx BW, Riese H. Negative and positive life events are associated with small but lasting change in neuroticism. *Psychol Med.* 2013;15:1-13.
78. American Psychiatric Association. *DSM-V: Diagnostic and Statistical Manual of Mental Disorders.* Arlington, VA: American Psychiatric Publishing;2013. p.5-25;
79. Amdur RL, Liberzon I. The structure of posttraumatic stress disorder symptoms in combat veterans: a confirmatory factor analysis of the impact of event scale. *J Anxiety Disord.* 2001;15(4):345-57.
80. Asmundson GJ, Frombach I, McQuaid J, Pedrelli P, Lenox R, Stein MB. Dimensionality of posttraumatic stress symptoms: a confirmatory factor analysis of DSM-IV symptom clusters and other symptom models. *Behav Res Ther.* 2000;38(2):203-14.
81. Karam EG, Andrews G, Bromet E, Petukhova M, Ruscio AM, Salamoun M, et al. The role of criterion A2 in the DSM-IV diagnosis of posttraumatic stress disorder. *Biol Psychiatry.* 2010;68(5):465-73.
82. Friedman MJ, Resick PA, Bryant RA, Brewin CR. Considering PTSD for DSM-5. *Depress Anxiety.* 2011;28(9):750-69.
83. Zoellner LA, Rothbaum BO, Feeny NC. PTSD not an anxiety disorder? DSM committee proposal turns back the hands of time. *Depress Anxiety.* 2011; 28(10):853-6.
84. Van AM, Mancini C, Patterson B. The impact of changing diagnostic criteria in posttraumatic stress disorder in a Canadian epidemiologic sample. *J Clin Psychiatry.* 2011;72(8):1034-41.
85. Regier DA, Myers JK, Kramer M, Robins LN, Blazer DG, Hough RL, et al. The NIMH Epidemiologic Catchment Area program. Historical context, major objectives, and study population characteristics. *Arch Gen Psychiatry.* 1984;41(10):934-41.
86. Keane TM, Caddell JM, Taylor KL. Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: three studies in reliability and validity. *J Consult Clin Psychol* 1988;56(1):85-90.
87. Vreven DL, Gudanowski DM, King LA, King DW. The civilian version of the Mississippi PTSD Scale: a psychometric evaluation. *J Traum Stress* 1995;8(1):91-109.
88. Butcher JN, Dahlstrom WG, Graham J, Tellegen A&, Kaemmer B. *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring.* Minneapolis: University of Minnesota Press; 1989.
89. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. *Psychosom Med.* 1979;41(3):209-18.
90. Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, et al. The development of a clinician-administered PTSD scale. *J Traum Stress* 1995;8:75-90.
91. First MB, Caban D. *Structured Clinical Interview for DSM Diagnosis.* The Corsini Encyclopedia of Psychology. Hoboken, New Jersey: John Wiley & Sons, Inc; 2010.
92. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P).* New York: Biometrics Research, New York State Psychiatric Institute; 1997.
93. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition. (SCID-I/NP).* New York: Biometrics Research, New York State Psychiatric Institute; 2002.
94. Foa EB, Riggs DS, Dancu CV, Rothbaum BO. Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *J Traum Stress.* 1993;6(4):459-73.
95. Lauterbach D, Vrana S. Three Studies on the Reliability and Validity of a Self-Report Measure of Posttraumatic Stress Disorder. *Assess.* 1996;3(1):17-25.
96. Foa EB, Ehlers A, Clark DM, Tolin DF, Orsillo SM. *The Post-Traumatic Cognition Inventory (PTCI): Development and validation.* *Psychol. Assess.* 1999;11:303-14.
97. Foa EB, Cashman L, Jaycox L, Perry KPA. The validation of a selfreport measure for posttraumatic stress disorders: The Posttraumatic Diagnostic Scale. *Psychol Assess.* 1997;9:445-51.

98. Pynoos RS, Frederick C, Nader K, Arroyo W, Steinberg A, Eth S, et al. Life threat and posttraumatic stress in school-age children. *Arch Gen Psychiatry*. 1987;44(12):1057-63.
99. Davidson JR, Book SW, Colket JT, Tupler LA, Roth S, David D, et al. Assessment of a new self-rating scale for post-traumatic stress disorder. *Psychol Med*. 1997;27(1):153-60.
100. Weathers F, Litz B, Herman D, Huska J, Keane T. The PTSD Checklist (PCL): Reliability, Validity, and Diagnostic Utility. Annual Convention of the International Society for Traumatic Stress Studies. San Antonio: International Society for Traumatic Stress Studies; 1993.
101. Eysenck HJ. The measurement of personality: A new inventory. *J Indian Acad Applied Psychol*. 1964;1:1-11.
102. Eysenck HJ, Eysenck SBG. Manual of the Eysenck Personality Questionnaire. San Diego: Educational and Industrial Testing Service; 1975.
103. Eysenck SBG, Eysenck HJ, Barrett P. A revised version of the psychoticism scale. *Pers Individ Dif*. 1985;6(1):21-9.
104. Eysenck HG, Eysenck SBG. EPS Manual of EPScale (EPS-Adult). London: Hodder & Stoughton; 1991.
105. Costa PT, McCrae RR. The NEO Personality Inventory Manual. Odessa: Psychological Assessment Resources; 1985.
106. Costa PT, McCrae RR. The NEO-PI/NEO-FFI Manual supplement. Odessa: Psychological Assessment Resources; 1989.
107. Costa PT, McCrae RR. Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI). Professional manual. Odessa: Psychological Assessment Resources; 1992.
108. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS Scales. *J Pers Soc Psychol*. 1988;47:1063-70.
109. John OP, Srivastava S. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. En: L.A.Pervin, O.P.John, editors. *Handbook of personality: Theory and research*. 2 ed. New York: Guilford; 1999. p. 102-38.
110. Cox BJ, McWilliams LA, Enns MW, Clara IP. Broad and specific personality dimensions associated with major depression in a nationally representative sample. *Compr Psychiatry*. 2004;45(4):246-53.
111. Kashdan TB, Steger MF. Curiosity and pathways to well-being and meaning in life: Traits, states, and everyday behaviors. *Motiv Emot*. 2007;31:159-73.

REVISIÓN SISTEMÁTICA Y METAANÁLISIS

BLOQUEO DUAL DEL SISTEMA RENINA-ANGIOTENSINA FRENTE A LA MONOTERAPIA: REVISIÓN SISTEMÁTICA Y METAANÁLISIS ACUMULATIVO DE ENSAYOS CLÍNICOS Y ESTUDIOS OBSERVACIONALES

Ferrán Catalá-López, Diego Macías Saint-Gerons, César de la Fuente Honrubia y Gloria Martín-Serrano.

División de Farmacoepidemiología y Farmacovigilancia. Departamento de Medicamentos de Uso Humano. Agencia Española de Medicamentos y Productos Sanitarios (AEMPS). Madrid. España.

Financiación: ninguna.

Las opiniones expresadas en este trabajo son responsabilidad de los autores, por lo que no reflejan necesariamente el punto de vista de los organismos en los que trabajan.

RESUMEN

Fundamentos: Trabajos previos han sugerido que el bloqueo dual con inhibidores del sistema renina-angiotensina (SRA) estaría asociado a un incremento de efectos adversos comparado con la monoterapia. Se re-examina la seguridad del doble bloqueo del SRA, especialmente en pacientes de riesgo y se explora la estabilidad de la evidencia acumulada a lo largo de los años.

Método: Revisión sistemática con metaanálisis de efectos aleatorios. Se revisaron 15 metaanálisis publicados anteriormente como punto de partida. Se realizó una búsqueda en PubMed/Medline de estudios observacionales y de ensayos clínicos controlados recientes. Las variables dependientes estudiadas fueron: mortalidad (general y cardiovascular), hipotensión, insuficiencia renal, accidente cerebrovascular y la retirada del tratamiento por efectos adversos. Se calcularon los riesgos relativos (RR) y su intervalo de confianza (IC95%)

Resultados: El bloqueo dual del SRA no se asoció con una reducción del RR de la mortalidad general (RR:1.00; IC95%: 0.96-1.05; 21 estudios), ni de la mortalidad cardiovascular (RR:1.01; 0.94-1.09; 13 estudios) ni del riesgo de accidente cerebrovascular (RR:1.02; 0.94-1.11; 11 estudios) en comparación con la monoterapia. El bloqueo dual se asoció a un riesgo aumentado de hiperpotasemia (RR:1.58; 1.37-1.81; 34 estudios), hipotensión (RR:1.66; 1.41-1.95; 25 estudios), daño renal (1.52; 1.28-1.81; 29 estudios) y retirada del tratamiento por efectos adversos (RR:1.26; 1.22-1.30; 37 estudios). Los resultados fueron consistentes en las cohortes de pacientes con diabetes mellitus, enfermedad renal o insuficiencia cardíaca.

Conclusiones: El bloqueo dual del SRA vs monoterapia incrementa los riesgos de hiperpotasemia, hipotensión, insuficiencia renal e interrupción del tratamiento por efectos adversos. Además no ofrece beneficios adicionales por reducción de la mortalidad general, mortalidad cardiovascular o accidente cerebrovascular. Estos hallazgos son consistentes a lo largo del tiempo.

Palabras clave: Hipertensión arterial. Enfermedad renal crónica. Efectos adversos. Sistema renina-angiotensina. Medicina basada en la evidencia. Metaanálisis. Revisión. Terapia farmacológica, combinación.

Correspondencia

Ferrán Catalá-López

División de Farmacoepidemiología y Farmacovigilancia

Agencia Española de Medicamentos y Productos Sanitarios (AEMPS)

Parque Empresarial "Las Mercedes"

C/ Campezo, 1 - Edificio 8, 1ª planta 28022 Madrid

Correo electrónico: ferran_catala@hotmail.com

ABSTRACT

Risks of Dual Blockade of the Renin-Angiotensin System Compared with Monotherapy: A Systematic Review and Cumulative Meta-analysis of Randomized Trials and Observational Studies

Background: Previous work has suggested that dual blockade using inhibitors of the renin-angiotensin system (RAS) would be associated with an increase in side effects compared to monotherapy. We reexamined the safety of dual RAS blockade, especially in patients at risk, and explored the stability of the evidence accumulated over the years.

Method: Systematic review with random-effects meta-analyses. We reviewed 15 previously published meta-analyses as the starting point. PubMed/Medline was searched for recent evidence from both observational and randomized controlled trials. Outcomes measures were: mortality (overall and cardiovascular), hyperkalemia, hypotension, renal failure, stroke, and treatment withdrawal due to adverse effects. We calculated relative risks (RR) and confidence intervals (95% CI)

Results: Dual RAS blockade was not associated with reduced relative risk (RR) overall mortality (RR:1.00, 0.96-1.05; 21 studies), cardiovascular mortality (RR:1.01, 0.94-1.09; 13 studies) or stroke (RR:1.02; 0.94-1.11; 11 studies) compared to monotherapy. Dual blockade was associated with an increased risk of hyperkalemia (RR:1.58, 1.37-1.81; 34 studies), hypotension (RR:1.66; 1.41-1.95; 25 studies), renal failure (RR:1.52; 1.28-1.81; 29 studies) and treatment discontinuation due to adverse events (RR:1.26; 1.22-1.30; 37 studies). These results were consistent in cohorts of patients with diabetes mellitus, kidney disease or heart failure.

Conclusions: Dual RAS blockade increased (vs monotherapy) the risks of hyperkalemia, hypotension, renal failure and treatment discontinuation. Dual RAS blockade did not offer additional benefit in reducing overall mortality, cardiovascular mortality or stroke. These findings are consistent over time.

Key words: Hypertension. Kidney failure, chronic. Adverse effects. Renin-angiotensin system. Evidence-Based Medicine. Meta-analysis. Review. Drug therapy. Drug Therapy, Combination.

INTRODUCCIÓN

En la última década se han llevado a cabo importantes ensayos clínicos aleatorizados y controlados con el objetivo de determinar los efectos de las diversas estrategias antihipertensivas, más allá de controlar o reducir las cifras de presión arterial¹. Una de las opciones terapéuticas que ha suscitado especial interés es el bloqueo dual (o doble bloqueo) del sistema renina-angiotensina (SRA) en el abordaje clínico de la hipertensión y de otras enfermedades cardiovasculares. El bloqueo dual del SRA consiste en interferir simultáneamente con inhibidores de la enzima de conversión de la angiotensina (IECA), antagonistas de los receptores de la angiotensina II (ARA-II) o inhibidores directos de la renina.

Se ha propuesto como hipótesis que la acción combinada de medicamentos para el bloqueo dual del SRA pudiese mejorar los resultados obtenidos con cada uno de los medicamentos por separado (monoterapia) en la prevención cardiovascular y/o renal de los pacientes con un perfil de riesgo cardiometabólico elevado². Sin embargo, ya en el ensayo *The ONgoing Telmisartan Alone and in combination with Ramipril Global Endpoint Trial* (ONTARGET) se observó que el tratamiento mediante bloqueo dual con IECA y ARA-II no parecía ofrecer ningún beneficio adicional en la reducción de la morbimortalidad cardiovascular de los pacientes con elevado riesgo cardiovascular respecto a la monoterapia³. Del mismo modo, el ensayo *Aliskiren Trial in Type 2 Diabetes Using Cardiovascular and Renal Disease Endpoints* (ALTITUDE) se interrumpió prematuramente debido a que la terapia dual tampoco mostró un claro beneficio clínico en las personas diabéticas que recibían un inhibidor directo de la renina (aliskireno) además del tratamiento con IECA o ARA-II, mientras que se produjeron mayores complicaciones y efectos adversos⁴. En enero de 2013 se publicó una revisión sistemática y metaanálisis de ensa-

yos clínicos sobre el efecto del bloqueo dual del SRA en pacientes con diverso perfil de riesgo cardiovascular⁵. Los autores indicaban que representaba el estudio más completo de la literatura que evaluaba la seguridad del bloqueo dual del SRA. Sin embargo, no fue el primer metaanálisis en examinar el efecto de la terapia dual⁶⁻⁹. Otras revisiones también han evaluado los efectos del bloqueo dual del SRA en grupos de pacientes más específicos, como son aquellos que tienen insuficiencia cardíaca¹⁰, enfermedad renal crónica¹¹ o diabetes mellitus¹². Por otro lado, recientemente se han seguido publicando nuevos ensayos clínicos (incluyendo reanálisis de estudios previos)¹³⁻¹⁶ o estudios observacionales que, en ocasiones, presentaban resultados contradictorios^{17,18}.

El objetivo de este trabajo es examinar la seguridad del doble bloqueo del SRA en comparación con la monoterapia, especialmente en personas con diabetes mellitus, enfermedad renal o insuficiencia cardíaca, y explorar la estabilidad de la evidencia acumulada a lo largo de los años considerando la procedente de ensayos clínicos y estudios observacionales.

MATERIAL Y MÉTODOS

Diseño. Revisión sistemática y metaanálisis de ensayos clínicos y estudios observacionales

Búsqueda de la literatura. En una primera fase, como punto de partida de este trabajo, se buscaron y revisaron los metaanálisis publicados hasta noviembre de 2013. En una segunda fase, también se realizó una búsqueda en PubMed/Medline para actualizar los trabajos identificados con nuevos ensayos clínicos, pero incorporando además la evidencia de origen observacional. Para la búsqueda específica se utilizaron los siguientes términos distribuidos en 3 bloques:

a) tratamiento farmacológico (“renin-angiotensin system”, “renin-angiotensin-aldosterone system”, “angiotensin-converting enzyme inhibitor,” “angiotensin receptor blocker,” “renin inhibitor,” or “direct renin inhibitor”, “benazepril”, “captopril”, “enalapril”, “cilazapril”, “delapril”, “fosinopril”, “imidapril”, “lisinopril”, “moexipril”, “perindopril”, “quinapril”, “ramipril”, “spirapril”, “temocapril”, “trandolapril”, “zofenopril”, “candesartan”, “irbesartan”, “losartan”, “telmisartan”, “valsartan”, “olmesartan”, “eprosartan”, “azilsartan”, “aliskiren”).

b) diseño de los estudios (“systematic review and meta-analysis” OR “randomized controlled trial”, “cohort”, “case-control”).

c) combinación del tratamiento (“dual blockade”, “dual therapy”, “combination therapy”, “combined”, “addition”, “added”). La lista completa de términos utilizados se presenta en los anexos 1 y 2.

Selección de estudios. De los metaanálisis identificados se seleccionaron los ensayos clínicos aleatorizados y controlados (con un diseño que incluyera grupos paralelos) con al menos comparaciones de la terapia dual del SRA [(IECA, ARA-II, o inhibidores directos de la renina (aliskireno)] respecto a monoterapia, con datos disponibles sobre la seguridad y que presentaran un tamaño muestral de al menos 30 pacientes. A partir de una búsqueda actualizada, se incluyeron nuevos ensayos clínicos o reanálisis de estos (no disponibles en los metaanálisis previos), además de incorporar también publicaciones de estudios observacionales epidemiológicos (cohorte o caso-control anidado) que incluyeran alguna de las variables de interés. Las variables de seguridad consideradas fueron: mortalidad, mortalidad cardiovascular, hiperpotasemia, hipotensión, daño renal, accidente cerebrovascular y la retirada del tratamiento (por efectos adversos). Se utilizaron las definiciones propuestas en cada uno de los estudios identificados.

Extracción de la información. Dos investigadores identificaron los estudios potencialmente elegibles que cumplieran con los criterios de inclusión. Un tercer investigador realizó la extracción de la información de los documentos fuente en hojas de recogida de datos diseñadas *ad hoc* (disponibles mediante solicitud a los autores). Se extrajeron las características de los estudios individuales y se representaron en tablas de evidencia.

Síntesis de la evidencia y análisis estadístico. En la síntesis cuantitativa de la evidencia se utilizó un enfoque en dos etapas: en la primera se utilizaron conjuntamente los datos procedentes de ensayos clínicos y de estudios observacionales. En la segunda se incluyeron solamente datos de ensayos clínicos, lo que permitió una evaluación de la contribución adicional de la evidencia de los estudios observacionales.

Para cada una de las variables de seguridad (mortalidad, mortalidad cardiovascular, hiperpotasemia, hipotensión, daño renal, accidente cerebrovascular, retirada del tratamiento) se calcularon los riesgos relativos (RR) junto con los intervalos de confianza del 95% (IC95%). Para ello, se llevaron a cabo metaanálisis utilizando modelos de efectos aleatorios ponderando por el inverso de la varianza (DerSimonian y Laird)¹⁹. El estadístico I^2 se utilizó para analizar heterogeneidad²⁰. Para explorar fuentes de heterogeneidad se definieron subgrupos a partir del tipo de terapia dual (por ejemplo que entre los medicamentos inhibidores del SRA se incluyera o no aliskireno) y poblaciones de alto riesgo (por ejemplo personas con diabetes mellitus, enfermedad renal e insuficiencia cardíaca). Se llevaron a cabo metaanálisis acumulativos para explorar cambios en el cuerpo de evidencia a lo largo del tiempo²¹.

Todos los análisis se realizaron con el programa Stata 12 (StataCorp LP, College Station, TX, EE UU).

RESULTADOS

La búsqueda bibliográfica permitió identificar 15 metaanálisis previos (anexo 3), de los cuales se extrajeron 39 estudios individuales. Tras actualizar las búsquedas con el objetivo de identificar nuevos estudios publicados, un total de 50 trabajos se incluyeron en los análisis (figura 1, tabla 1). Las características de los estudios y de los pacientes se presentan en la tabla 2.

En las tablas 3 y 4 se presentan los principales resultados de los metaanálisis.

a) Mortalidad. Se identificaron 21 estudios (104.538 pacientes) que presentaron datos de mortalidad. El bloqueo dual del SRA no se asoció con una reducción de la mortalidad total en comparación con la monoterapia (RR=1,00; IC95%: 0,96-1,05) (tabla 3), con independencia de la estrategia de combinación utilizada (con o sin aliskireno). Los análisis de subgrupos no identifica-

Figura 1
Algoritmo de selección de artículos

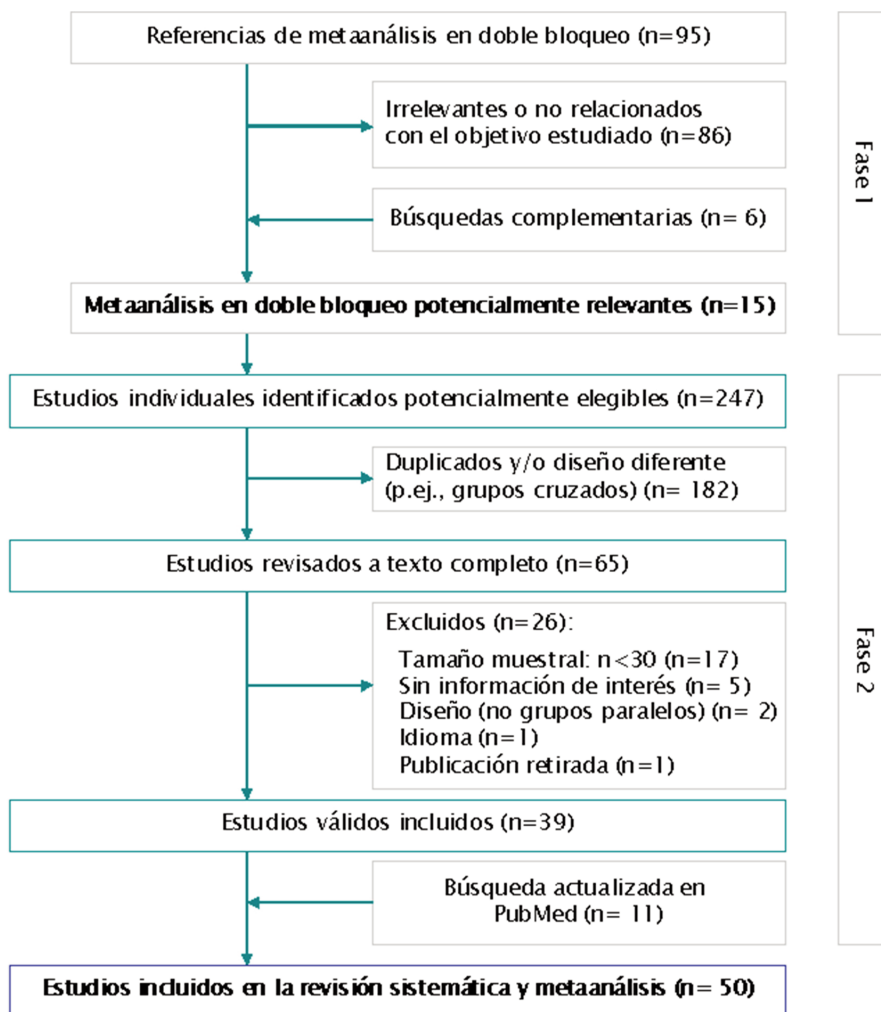


Tabla 1
Evidencia incluida en el metaanálisis,
por variable de resultado

Variable de resultado por tipo de diseño	Número de pacientes (número de estudios)
1. Mortalidad	104.538 (21)
Ensayos clínicos	75.910 (20)
Estudios observacionales	28.628 (1)
2. Mortalidad cardiovascular	97.717 (13)
Ensayos clínicos	69.042 (12)
Estudios observacionales	28.628 (1)
3. Hipertensión	140.094 (34)
Ensayos clínicos	74.342 (30)
Estudios observacionales	65.752 (4)
4. Hipotensión	100.268 (25)
Ensayos clínicos	71.640 (24)
Estudios observacionales	28.628 (1)
5. Daño renal	128.929 (29)
Ensayos clínicos	78.386 (27)
Estudios observacionales	50.543 (2)
6. Accidente cerebrovascular	57.317 (11)
Ensayos clínicos	49.034 (10)
Estudios observacionales	8.238 (1)
7. Retirada del tratamiento (por efectos adversos)	68.606 (37)
Ensayos clínicos	68.606 (37)
Estudios observacionales	-

ron diferencias significativas en el riesgo de mortalidad en pacientes con diabetes, enfermedad renal o insuficiencia cardiaca (tabla 4).

b) Mortalidad cardiovascular. Se identificaron 13 trabajos (97.717 pacientes). No se observó una reducción de la mortalidad cardiovascular con la terapia dual en comparación con la monoterapia (RR=1,01; IC95%: 0,94-1,09) (tabla 3), si bien se observó una tendencia a un incremento de riesgo del 17% cuando se incluía tratamiento con aliskireno (RR=1,17; IC95%: 0,98-1,39; P=0,07). No se identificaron cambios entre los diferentes subgrupos de pacientes de riesgo (tabla 4).

c) Hipertensión. Se identificaron 34 estudios (140.094 pacientes). El bloqueo dual se asoció a un incremento de riesgo de hipertensión (RR=1,58; IC95%: 1,37-1,81), con independencia de la estrategia de combinación utilizada (con o sin aliskireno). Los resultados fueron consistentes en los diferentes subgrupos de pacientes de riesgo (tabla 4).

d) Hipotensión. Veinticinco estudios (100.268 pacientes) permitieron observar un incremento del riesgo de hipotensión asociado a la terapia dual en comparación con la monoterapia (RR=1,66; IC95%: 1,41-1,95). Estos hallazgos fueron consistentes en los diferentes análisis de subgrupos (tablas 3 y 4).

e) Daño renal. Veintinueve trabajos (128.929 pacientes) presentaron información sobre el riesgo de daño renal. El doble bloqueo se asoció a un aumento del riesgo de daño renal en comparación con la monoterapia (RR=1,52; IC95%: 1,28-1,81) (Tabla 3). Los resultados fueron consistentes en los diferentes subgrupos de pacientes de riesgo (tabla 4).

f) Accidente cerebrovascular. Se identificaron 11 estudios (57.317 pacientes). El bloqueo dual no se asoció con una reducción del riesgo de accidente cerebrovascular (RR=1,02; IC95%: 0,94-1,11) (tabla 3). Sin embargo, la terapia dual con aliskireno se asoció a un incremento potencial del riesgo de accidente cerebrovascular en comparación con la monoterapia (RR=1,21; IC95%: 1,00-1,46). No se identificaron diferencias significativas en los subgrupos de pacientes con diabetes, enfermedad renal o insuficiencia cardiaca (tabla 4).

g) Retirada del tratamiento (por efectos adversos). En 37 estudios (68.606 pacientes) la terapia dual se asoció a un incremento en la retirada del tratamiento por efectos adversos (RR=1,26; IC95%: 1,22-1,30).

Tabla 2
Características de los estudios seleccionados

Estudio (año)	Diseño	Características de los pacientes	N pacientes	Edad	Varones (%)	meses	Doble bloqueo	Monoterapia	Otros tratamientos (junto doble bloqueo)	Datos disponibles
Doble bloqueo con aliskireno										
Uresin et al 2007	ECA, DC, multicéntrico	Hipertensión y diabetes mellitus	837	60	56-61	2	Ramipril +aliskireno	Ramipril o aliskireno	No disponible	Hiperpotasemia, daño renal, retirada del tratamiento (por efectos adversos)
Oparil et al 2007	ECA, DC, multicéntrico	Hipertensión	1797	52-53	58-62	2	Valsartán +aliskireno	Valsartán	No disponible	Mortalidad (total), hiperpotasemia, daño renal, retirada del tratamiento (por efectos adversos)
Pool et al 2007	ECA, DC, multicéntrico	Hipertensión	1123	55-57	56	2	Valsartán +aliskireno	Valsartán	No disponible	Hiperpotasemia, daño renal, retirada del tratamiento (por efectos adversos)
McMurray et al 2008	ECA, DC, multicéntrico	Hipertensión e insuficiencia cardíaca	302	68	76-80	3	IECA o ARA-II +aliskireno	IECA o ARA-II	Beta-bloqueantes (94%), Antagonistas de la aldosterona (33%)	Mortalidad (cardiovascular), hiperpotasemia, hipotensión, accidente cerebrovascular, daño renal, retirada del tratamiento (por efectos adversos)
Parving et al 2008	ECA, DC, multicéntrico	Hipertensión y nefropatía diabética	599	60-62	68-74	6	Losartán +aliskireno	Losartán	Diuréticos (63%), Antagonistas del calcio (51%), Beta-bloqueantes (35%)	Mortalidad (total), hiperpotasemia, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Solomon et al 2009	ECA, DC, multicéntrico	Hipertensión y hipertrofia ventricular izquierda	465	59	58-65	9	Losartán +aliskireno	Losartán o aliskireno	Diuréticos (40%), Antagonistas del calcio (39%), Beta-bloqueantes (35%), Antagonistas de la aldosterona (0%)	Hiperpotasemia, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Scirica et al 2010	ECA, DC, multicéntrico	Síndrome coronario agudo sin insuficiencia cardíaca, péptido natriurético aumentado	1101	63	69-73	2	Valsartán +aliskireno	Valsartán o aliskireno	Diuréticos (18%), Antagonistas del calcio (27%), Beta-bloqueantes (90%), Antagonistas de la aldosterona (3%)	Mortalidad (total y cardiovascular), hiperpotasemia, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Solomon et al 2011	ECA, DC, multicéntrico	Infarto de miocardio con disfunción sistólica	820	59-61	81-85	9	IECA o ARA-II +aliskireno	IECA o ARA-II	Beta-bloqueantes (96%), Antagonistas de la aldosterona (29%)	Mortalidad (total y cardiovascular), hiperpotasemia, hipotensión, daño renal
Drummond et al 2011	ECA, DC, multicéntrico	Hipertensión y diabetes mellitus	363	57	54	3	Valsartán +aliskireno	Valsartán	No disponible	Hiperpotasemia, daño renal, retirada del tratamiento (por efectos adversos)
Flack et al 2012	ECA, DC, multicéntrico	Hipertensión	451	57	47-54	2	Valsartán +aliskireno	Valsartán	No disponible	Retirada del tratamiento (por efectos adversos)
Parving et al 2012	ECA, DC, multicéntrico	Diabetes mellitus y enfermedad renal crónica, enfermedad cardiovascular e hipertensión (controlada)	8561	65	67	32	IECA o ARA-II +aliskireno	Aliskireno	No disponible	Mortalidad (total y cardiovascular), hiperpotasemia, hipotensión, accidente cerebrovascular, daño renal, retirada del tratamiento (por efectos adversos)
Bakris et al 2013	ECA, DC, multicéntrico	Hipertensión y diabetes mellitus	1143	55	57-59	2	Valsartán +aliskireno	Valsartán	No disponible	Mortalidad (total), hiperpotasemia, daño renal, retirada del tratamiento (por efectos adversos)
Gilbert et al 2013	Observacional	Insuficiencia cardíaca, diabetes melli- tus y/o enfermedad renal crónica	30749	81	45-50	28	IECA o ARA-II +aliskireno	IECA o ARA-II	Diuréticos (45-76%), Beta-bloqueantes (38-44%)	Hiperpotasemia, accidente cerebrovascular, daño renal

Tabla 2
Características de los estudios seleccionados (continuación)

Estudio (año)	Diseño	Características de los pacientes	N pacientes	Edad	Varones (%)	meses	Doble bloqueo	Monoterapia	Otros tratamientos (junto doble bloqueo)	Datos disponibles
Doble bloqueo sin aliskireno										
McKelvie et al 1999	ECA, DC, multicéntrico	Insuficiencia cardiaca y fracción de eyección reducida	768	63	80-90	10	Enalapril + candesartán	Enalapril o candesartán	Diuréticos (84%), Antagonistas del calcio (15%), Beta-bloqueantes (13%)	Mortalidad (total), hipotensión, daño renal
Baruch et al 1999	ECA, DC, multicéntrico	Insuficiencia cardiaca sintomática	83	62-65	100	1	Valsartán +IECA	IECA	No disponible	Hipertensión, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Hamroff et al 1999	ECA, DC, multicéntrico	Insuficiencia cardiaca	33	60-62	31-65	6	Losartán + IECA	IECA	Diuréticos (100%), Antagonistas del calcio (6%), Beta-bloqueantes (6%)	Mortalidad (total), retirada del tratamiento (por efectos adversos)
Azizi et al 2000	ECA, DC, multicéntrico	Hipertensión esencial	177	ND	64	1,2	Enalapril o losartán	Enalapril o losartán	No disponible	Retirada del tratamiento (por efectos adversos)
Mogensen et al 2000	ECA, DC, multicéntrico	Hipertensión, diabetes mellitus y microalbuminuria	197	60	37-48	6	Candesartán + lisinopril	Candesartán o lisinopril	No disponible	Hipotensión
Ruilope et al 2000	ECA, A, multicéntrico	Hipertensión, enfermedad renal crónica con/sin proteinuria	108	57	66-73	1	Benazepril + valsartán	Valsartán	No disponible	Hipertensión, hipotensión, retirada del tratamiento (por efectos adversos)
Tonkon et al 2000	ECA, DC, multicéntrico	Insuficiencia cardiaca y fracción de eyección reducida	109	63	73-79	2	Irbesartán + IECA	IECA	Diuréticos (84%), Antagonistas del calcio (15%), Beta-bloqueantes (13%)	Hipotensión, retirada del tratamiento (por efectos adversos)
Cohn et al 2001 - Val-HeFT	ECA, DC, multicéntrico	Insuficiencia cardiaca sintomática estable	5010	62	80	25	Valsartán + IECA	IECA	Diuréticos (86%), Beta-bloqueantes (34%)	Mortalidad (total), hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Murdoch et al 2001	ECA, DC, unicéntrico	Insuficiencia cardiaca	36	61-66	83	2	Eprosartán + IECA	IECA	Diuréticos (44%), Antagonistas del calcio (17%), Beta-bloqueantes (50%)	Hipotensión, retirada del tratamiento (por efectos adversos)
Tylicki et al 2002	ECA, A, unicéntrico	Glomerulonefritis primaria	51	37-44	41-73	3	Enalapril + losartán	Enalapril o losartán	No disponible	Retirada del tratamiento (por efectos adversos)
McMurray et al 2003	ECA, DC, multicéntrico	Insuficiencia cardiaca y fracción de eyección reducida	2548	64	79	41	Candesartán +IECA	IECA	Diuréticos (90%), Antagonistas del calcio (11%), Beta-bloqueantes (56%), Antagonistas aldosterona (17%)	Mortalidad (total y cardiovascular), hipertensión, hipotensión, accidente cerebrovascular, daño renal, retirada del tratamiento (por efectos adversos)
Pfeffer et al 2003	ECA, DC, multicéntrico	Infarto agudo de miocardio com- plicado por insuficiencia cardiaca y/o fracción de eyección reducida	14703	65	69	24,7	Valsartán + captopril	Captopril o valsartán	Diuréticos (12%), Antagonistas del calcio (23%), Beta-bloqueantes (32%)	Mortalidad (total y cardiovascular), hipertensión, hipo- tensión, daño renal, retirada del tratamiento (por efectos adversos)
Izzo et al 2004	ECA, DC, multicéntrico	Hipertensión esencial	1096	54	58	2	Lisinopril + candesartán	Lisinopril	No disponible	Mortalidad (total), hipertensión, retirada del tratamien- to (por efectos adversos)
Yasamura et al 2004	ECA, A, multicéntrico	Insuficiencia cardiaca	106	65	77-84	6	IECA+ARA-II	IECA+ARA-II	No disponible	Mortalidad (total), hipotensión, accidente cerebrovascular

Tabla 2
Características de los estudios seleccionados (continuación)

Estudio (año)	Diseño	Características de los pacientes	N pacientes	Edad	Varones (%)	meses	Doble bloqueo	Monoterapia	Otros tratamientos (junto doble bloqueo)	Datos disponibles
Doble bloqueo sin aliskireno										
Andersen et al 2005	ECA, A, unicéntrico	Hipertensión y diabetes mellitus	75	54-56	89-92	12	Candesartán + lisinopril	Lisinopril	Diuréticos (20%), Antagonistas del calcio (9%), Beta-bloqueantes (6%)	Hiperpotasemia, accidente cerebrovascular, retirada del tratamiento (por efectos adversos)
Kanno et al 2006	ECA, A, multicéntrico	Hipertensión y enfermedad renal crónica	90	60	40	36	Candesartán + IECA	IECA	Diuréticos (15%), Antagonistas del calcio (67%), Beta-bloqueantes (7%)	Daño renal
Horita et al 2006	ECA, A, unicéntrico	Nefropatía con IgA y proteinuria	43	38-43	54-57	12	Temocapril + losartán	Temocapril o losartán	No disponible	Hipotensión, retirada del tratamiento (por efectos adversos)
Bakris et al 2007	ECA, DC, multicéntrico	Hipertensión, albuminuria y alto riesgo cardiovascular	405	66	60-64	5	Ramipril + irbesartán	Irbesartán	No disponible	Hiperpotasemia, daño renal, retirada del tratamiento (por efectos adversos)
Ogawa et al 2007	ECA, MC, -	Hipertensión y nefropatía diabética	164	61-64	48	22,1	Temocapril + candesartán	Temocapril o candesartán	No disponible	Retirada del tratamiento (por efectos adversos)
White et al 2007	ECA, DC, multicéntrico	Insuficiencia cardíaca sintomática y fracción de eyección reducida	80	63	87-93	6	Candesartán + IECA	IECA	Diuréticos (80%), Antagonistas del calcio (7%), Beta-bloqueantes (95%), Antagonistas de la aldosterona (40%)	Hiperpotasemia, hipotensión, retirada del tratamiento (por efectos adversos)
Kum et al 2008	ECA, A, unicéntrico	Insuficiencia cardíaca y fracción de eyección reducida	50	66-69	72	12	Irbesartán + IECA	IECA	Diuréticos (88%), Beta-bloqueantes (64%)	Mortalidad (total y cardiovascular)
ONTARGET 2008	ECA, DC, multicéntrico	Alto riesgo cardiovascular	25.620	66	73	54	Ramipril + elmisartán	Ramipril o telmisartán	Diuréticos (28%), Antagonistas del calcio (34%), Beta-bloqueantes (57%)	Mortalidad (total y cardiovascular), hiperpotasemia, hipotensión, accidente cerebrovascular, daño renal, retirada del tratamiento (por efectos adversos)
Menne et al 2008	ECA, DC, multicéntrico	Hipertensión y microalbuminuria	133	57-59	67-77	7,5	Lisinopril + valsartán	Lisinopril o valsartán	Diuréticos (12%), Antagonistas del calcio (23%), Beta-bloqueantes (32%)	Hiperpotasemia, hipotensión, retirada del tratamiento (por efectos adversos)
Mori-Takeyama et al (2008)	ECA, A, unicéntrico	Glomerulonefritis crónica	86	37-38	56-63	36	Benazepril + candesartán	Candesartán	No disponible	Retirada del tratamiento (por efectos adversos)
Medhi et al 2009	ECA, DC, unicéntrico	Hipertensión, diabetes mellitus y albuminuria	81	49-52	44-50	12	Lisinopril + losartán	Lisinopril	Diuréticos (92%), Beta-bloqueantes (85%)	Hiperpotasemia, daño renal, retirada del tratamiento (por efectos adversos)
Cice et al 2010	ECA, DC, multicéntrico	Insuficiencia cardíaca y fracción de eyección reducida (en hemodiálisis)	332	63	88-91	36	Telmisartán + IECA	IECA	Diuréticos (90%), Antagonistas del calcio (11%), Beta-bloqueantes (62%)	Mortalidad (total y cardiovascular), hiperpotasemia, hipotensión, accidente cerebrovascular, retirada del tratamiento (por efectos adversos)

Tabla 2
Características de los estudios seleccionados (continuación)

Estudio (año)	Diseño	Características de los pacientes	N pacientes	Edad	Varones (%)	meses	Doble bloqueo	Monoterapia	Otros tratamientos (junto doble bloqueo)	Datos disponibles
Doble bloqueo sin aliskireno										
Bianchi et al 2010	ECA, A, uni- céntrico	Glomerulonefritis crónica	128	53	64	36	Ramipril+irbe- sartán	Ramipril	Diuréticos (52%), Antagonistas de la aldosterona (100%)	Hipertensión, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Titan et al 2011	ECA, A, uni- céntrico	Nefropatía diabética con ma- croalbuminuria	56	58	54-71	4	Enalapril+losar- tán	Losartán	Diuréticos (75%), Antagonistas del calcio (61%), Beta-bloqueantes (71%)	Hipertensión, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
McAlister et al 2011	Observacional	Mayores de 65 años	32312	76	35-46	13	IECA+ARA-II	IECA+ARA-II	Diuréticos (6%), Antagonistas del calcio (23%), Beta-bloqueantes (32%)	Hipertensión, hipotensión, daño renal
Chan et al 2011	Observacional	Alto riesgo cardiovascular en hemodiálisis	28628	61	48-52	15,1	IECA+ARA-II	IECA o ARA-II	No disponible	Mortalidad (total y cardiovascular), hipertensión, hipotensión, accidente cerebrovascular
Kurmik et al 2011	Observacional	Hipertensión y alto riesgo car- diovascular	577	68	58	19	IECA+ARA-II	IECA	Diuréticos (41%), Antagonistas del calcio (63%), Beta-bloqueantes (55%)	Hipertensión
Yano et al 2012	ECA, A, uni- céntrico	Infarto agudo de miocardio	160	60-61	85-86	7	Captopril+val- sartán	Valsartán	Beta-bloqueantes (35%)	Hipertensión, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Mann et al 2013 ONTARGET diabetes	ECA, DC, multicéntrico	Diabetes mellitus (con o sin enfermedad renal crónica)	9.628	66	68	54	Ramipril + telmisartán	Ramipril o telmisartán	Diuréticos (35%), Antagonistas del calcio (37%), Beta-bloqueantes (51%)	Mortalidad (total y cardiovascular), hipertensión, hipotensión, accidente cerebrovascular, daño renal
Lesogor et al 2013	ECA, DC, multicéntrico	Insuficiencia cardíaca sintomática estable	5.010	63	80	25	Valsartán + IECA	IECA	Diuréticos (91%), Beta-bloqueantes (35%), Antagonistas de la aldosterona (5%)	Mortalidad (cardiovascular) (reanálisis - Val-HeFT)
Fernández-Juárez et al 2013	ECA, A, multicéntrico	Nefropatía diabética	133	63-69	70-78	32	Lisinopril + irbesartán	Irbesartán	No disponible	Mortalidad (total), hipertensión, hipotensión, daño renal, retirada del tratamiento (por efectos adversos)
Imai et al 2013 ORIENT Hypertension	ECA, DC, multicéntrico	Nefropatía diabética e hipertensión	414	59	67-70	36	Olmesartán + IECA	IECA	Diuréticos (35-39%), Antagonistas del calcio (66-70%), Beta-bloqueantes (15-19%)	Mortalidad (total y cardiovascular), hipertensión, hipotensión, accidente cerebrovascular, daño renal, retirada del tratamiento (por efectos adversos) (reanálisis ORIENT)
Fried et al 2013 VA-NEPHRON-D	ECA, DC, multicéntrico	Nefropatía diabética	1.448	65	99	25	Losartán + lisinopril	Losartán	Diuréticos (71%), Antagonistas del calcio (57-59%), Beta-bloqueantes (69%)	Mortalidad (total), hipertensión, hipotensión, accidente cerebrovascular, daño re

A: abierto; ARA-II: antagonistas de los receptores de la angiotensina II; DC: doble-ciego; ECA: ensayo clínico aleatorizado; ERC: enfermedad renal crónica; IECA: inhibidores enzima convertidora de angiotensina; MC: mono-ciego.

Estas diferencias no fueron estadísticamente significativas en el subgrupo que consideraba la terapia dual con aliskireno (RR=1,16; IC95%: 0,91-1,49; p=0,24) (tabla 3). El riesgo de retirada del tratamiento fue consistente en las diferentes cohortes de pacientes con diabetes, enfermedad renal o insuficiencia cardiaca (tabla 4).

Metaanálisis acumulativos. La figura 2 presenta diversos metaanálisis acumulativos por año de publicación del estudio. Los primeros efectos estadísticamente significativos se observaron en los años 2001 (hipotensión y retirada del tratamiento por efectos adversos), 2003 (daño renal) y 2007 (hipertensión). Aunque los estudios posteriores han aumentado la precisión de la estimación puntual, no se

Tabla 3
Resumen de resultados del metaanálisis comparando el bloqueo dual del SRA respecto a la monoterapia

Variable de resultado	Riesgo relativo (IC95%)			
Tratamiento	Síntesis de ensayos clínicos y estudios observacionales		Síntesis de ensayos clínicos	
1. Mortalidad	1,00 (0,96-1,05)	0,88; 24,4%	1,00 (0,94-1,06)	0,99; 28,1%
Doble bloqueo con aliskireno	1,06 (0,93-1,21)	0,40; 0,0%	1,06 (0,93-1,21)	0,84; 0,0%
Doble bloqueo sin aliskireno	1,00 (0,94-1,05)	0,88; 36,4%	0,99 (0,92-1,06)	0,32; 40,7%
2. Mortalidad cardiovascular	1,01 (0,94-1,09)	0,81; 57,6%	0,98 (0,92-1,05)	0,65; 44,3%
Doble bloqueo con aliskireno	1,17 (0,98-1,39)	0,07; 0,0%	1,17 (0,98-1,39)	0,07; 0,0%
Doble bloqueo sin aliskireno	0,99 (0,91-1,07)	0,82; 66,1%	0,96 (0,90-1,03)	0,31; 50,9%
3. Hipertensión	1,58 (1,37-1,81)	<0,001; 70,6%	1,62 (1,41-1,86)	<0,001; 53,7%
Doble bloqueo con aliskireno	1,43 (1,20-1,70)	<0,001; 18,7%	1,50 (1,22-1,85)	<0,001; 22,3%
Doble bloqueo sin aliskireno	1,69 (1,36-2,09)	<0,001; 78,7%	1,69 (1,39-2,06)	<0,001; 50,0%
4. Hipotensión	1,66 (1,41-1,95)	<0,001; 63,2%	1,71 (1,44-2,03)	<0,001; 63,5%
Doble bloqueo con aliskireno	1,46 (1,15-1,86)	0,002; 27,1%	1,46 (1,15-1,85)	0,002; 26,5%
Doble bloqueo sin aliskireno	1,75 (1,40-2,19)	<0,001; 68,8%	1,86 (1,45-2,38)	<0,001; 67,2%
5. Daño renal	1,52 (1,28-1,81)	<0,001; 74,1%	1,53 (1,26-1,85)	<0,001; 73,2%
Doble bloqueo con aliskireno	1,22 (1,00-1,49)	0,05; 18,1%	1,57 (1,04-2,38)	0,03; 23,0%
Doble bloqueo sin aliskireno	1,58 (1,26-1,98)	<0,001; 79,8%	1,52 (1,20-1,92)	<0,001; 80,0%
6. Accidente cerebrovascular	1,02 (0,94-1,11)	0,66; 0,0%	1,01 (0,92-1,10)	0,85; 0,0%
Doble bloqueo con aliskireno	1,21 (1,00-1,46)	0,06; 0,0%	1,22 (0,96-1,55)	0,10; 0,0%
Doble bloqueo sin aliskireno	0,98 (0,89-1,08)	0,64; 0,0%	0,98 (0,89-1,08)	0,64; 0,0%
7. Retirada del tratamiento (por efectos adversos)	1,26 (1,22-1,30)	<0,001; 0,0%	1,26 (1,22-1,30)	<0,001; 0,0%
Doble bloqueo con aliskireno	1,16 (0,91-1,49)	0,24; 43,9%	1,16 (0,91-1,49)	0,24; 43,9%
Doble bloqueo sin aliskireno	1,26 (1,22-1,31)	<0,001; 0,0%	1,26 (1,22-1,31)	<0,001; 0,0%

Doble bloqueo con aliskireno (IECA o ARA-II+aliskireno); Doble bloqueo sin aliskireno (IECA+ARA-II). ARA-II: antagonistas de los receptores de la angiotensina II; IECA: inhibidores de la enzima convertidora de angiotensina; IC95%: intervalo de confianza del 95%; SRA: sistema renina-angiotensina. Un riesgo relativo > 1 indica que existe asociación directa con el efecto adverso, es decir, que el bloqueo dual se asocia a una mayor frecuencia de suceder el acontecimiento.

ha producido ningún cambio sustancial en la dirección o la magnitud del efecto del tratamiento a lo largo del tiempo.

Citas a los metaanálisis previos. Se identificaron 15 metaanálisis publicados entre 2005 y

2013 sobre los efectos del doble bloqueo del SRA con IECA, ARA-II o aliskireno. En la tabla 5 se presenta un resumen de sus principales características y hallazgos.

Tabla 4
Resumen de resultados del metaanálisis comparando el bloqueo dual del SRA respecto a la monoterapia, por subgrupos o poblaciones específicas

Variable de resultado	Síntesis de ensayos clínicos y estudios observacionales		Síntesis de ensayos clínicos	
	Riesgo relativo (IC95%)	P (riesgo relativo); I ² (%)	Riesgo relativo (IC95%)	P (riesgo relativo); I ² (%)
Diabetes mellitus				
Mortalidad	1,04 (0,96-1,13)	0,30; 0,0%	1,04 (0,96-1,13)	0,30; 0,0%
Mortalidad cardiovascular	1,07 (0,88-1,30)	0,52; 60,5%	1,07 (0,88-1,30)	0,52; 60,5%
Hiperpotasemia	1,52 (1,28-1,82)	<0,001; 53,5%	1,57 (1,30-1,89)	<0,001; 56,5%
Hipotensión	1,88 (1,26-2,81)	<0,01; 70,1%	1,88 (1,26-2,81)	<0,01; 70,1%
Daño renal	1,30 (1,11-1,51)	<0,01; 31,8%	1,31 (1,09-1,59)	<0,01; 37,4%
Accidente cerebrovascular	1,01 (0,93-1,10)	0,83; 0,0%	1,01 (0,92-1,10)	0,85; 0,0%
Retirada del tratamiento (por efectos adversos)	1,27 (1,14-1,42)	<0,001; 0,0%	1,27 (1,14-1,42)	<0,001; 0,0%
Enfermedad renal				
Mortalidad	0,98 (0,88-1,09)	0,69; 50,7%	0,95 (0,80-1,12)	0,54; 57,4%
Mortalidad cardiovascular	1,05 (0,87-1,28)	0,58; 74,1%	1,00 (0,78-1,28)	0,98; 73,5%
Hiperpotasemia	1,39 (1,13-1,69)	<0,001; 77,1%	1,54 (1,18-2,01)	<0,001; 51,9%
Hipotensión	1,71 (1,26-2,31)	<0,01; 61,1%	2,01 (1,33-3,01)	<0,01; 60,5%
Daño renal	1,20 (1,04-1,40)	0,01; 27,3%	1,23 (1,03-1,49)	0,03; 33,1%
Accidente cerebrovascular	1,05 (0,92-1,20)	0,46; 0,0%	1,04 (0,91-1,19)	0,56; 0,0%
Retirada del tratamiento (por efectos adversos)	1,31 (1,17-1,46)	<0,001; 0,0%	1,31 (1,17-1,46)	<0,001; 0,0%
Insuficiencia cardíaca				
Mortalidad	0,95 (0,84-1,06)	0,36; 53,3%	0,95 (0,84-1,06)	0,36; 53,3%
Mortalidad cardiovascular	0,94 (0,84-1,05)	0,28; 49,7%	0,94 (0,84-1,05)	0,28; 49,7%
Hiperpotasemia	2,16 (1,23-3,80)	<0,01; 65,1%	2,40 (1,19-4,80)	0,01; 70,1%
Hipotensión	1,57 (1,29-1,92)	<0,001; 16,4%	1,57 (1,29-1,92)	<0,001; 16,4%
Daño renal	2,01 (1,53-2,66)	<0,001; 40,8%	2,24 (1,96-2,55)	<0,001; 1,3%
Accidente cerebrovascular	1,22 (0,65-2,27)	0,53; 2,8%	1,43 (0,68-3,01)	0,34; 3,9%
Retirada del tratamiento (por efectos adversos)	1,34 (1,23-1,45)	<0,001; 0,0%	1,34 (1,23-1,45)	<0,001; 0,0%

IC95%: intervalo de confianza del 95%. Un riesgo relativo > 1 indica que existe asociación directa del efecto adverso, es decir, que el bloqueo dual se asocia a una mayor frecuencia de suceder el acontecimiento.

Figura 2
Metaanálisis acumulativo comparando el bloqueo dual del SRA
respecto a la monoterapia

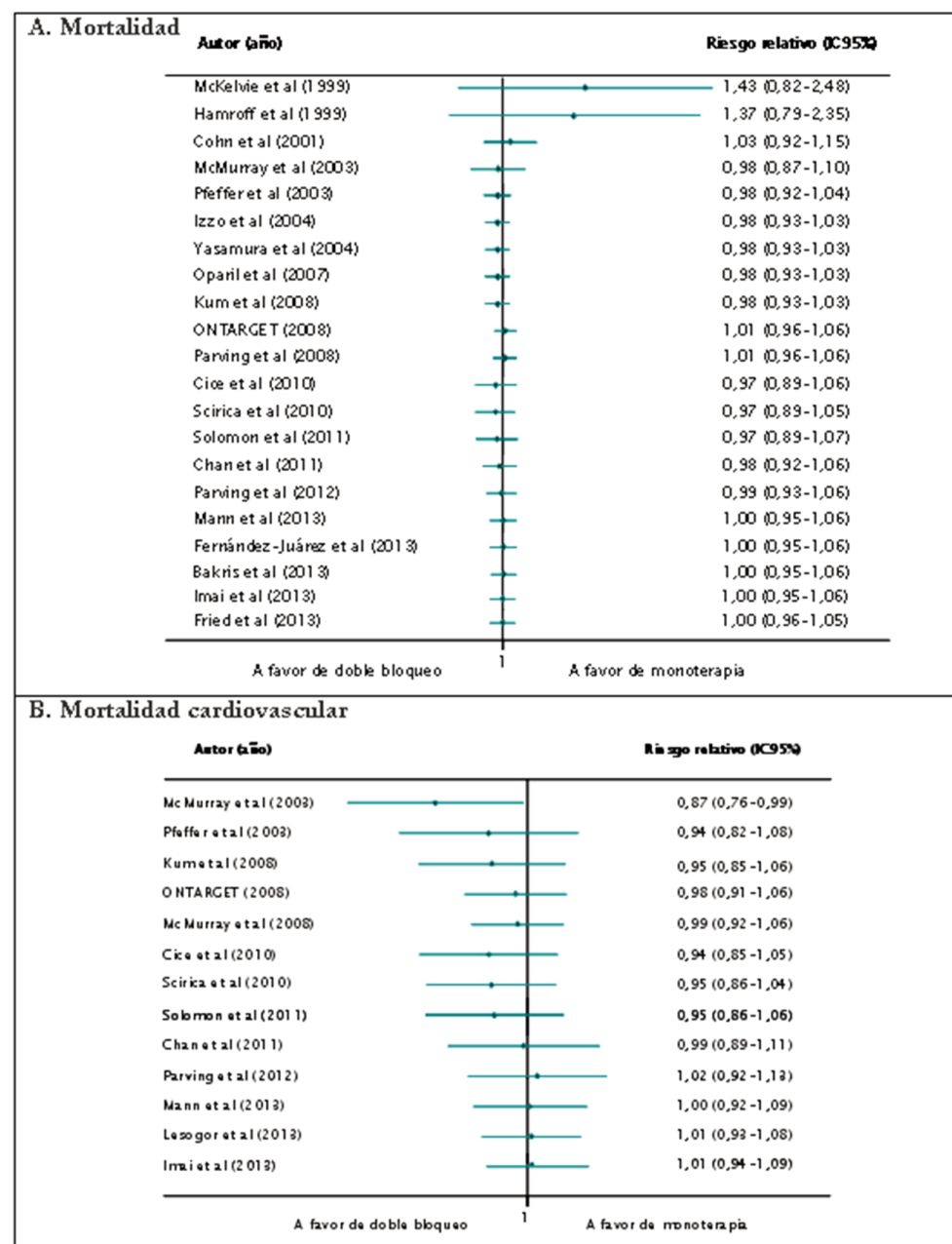


Figura 2
Metaanálisis acumulativo (continuación)

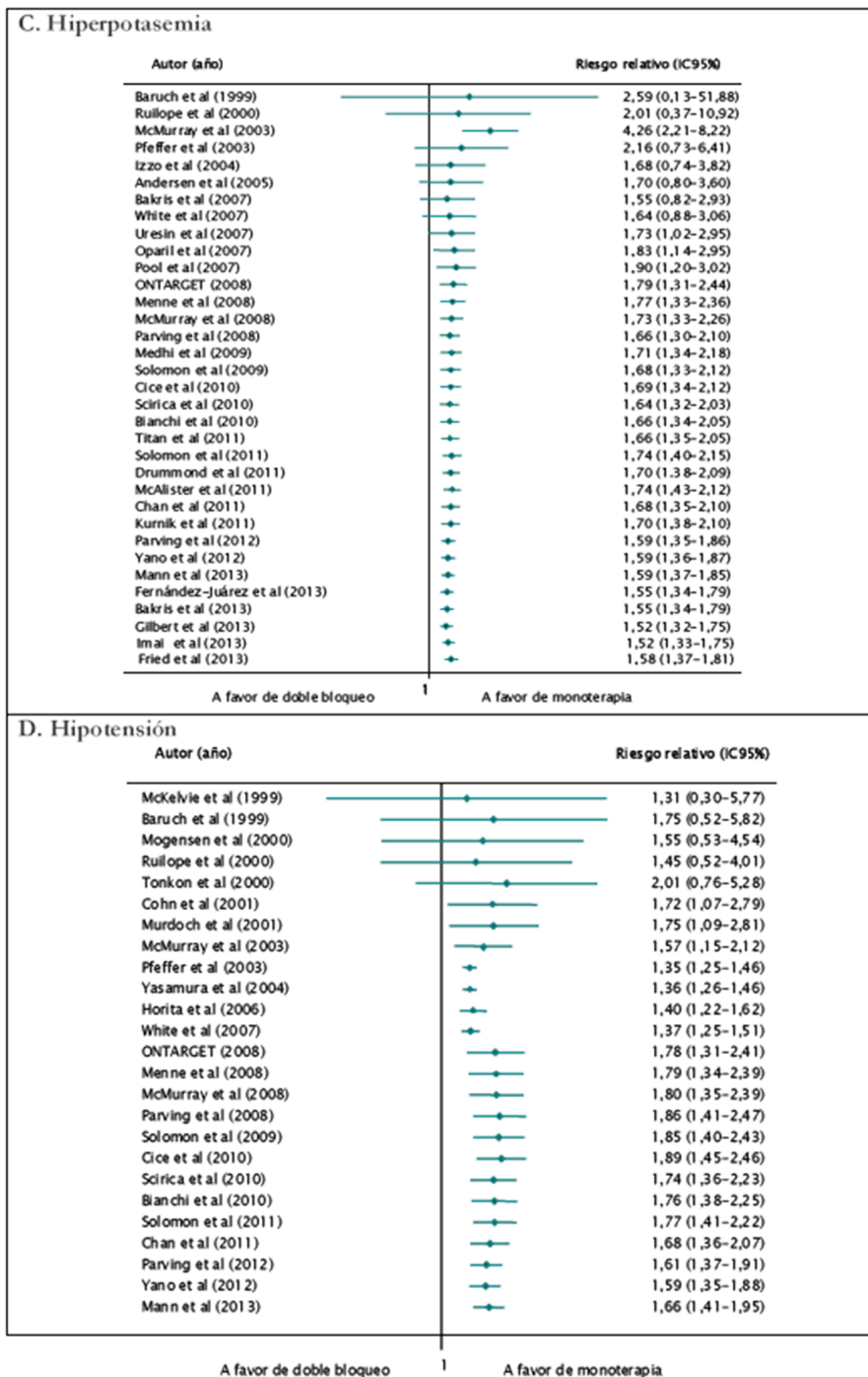


Figura 2
Metaanálisis acumulativo (continuación)

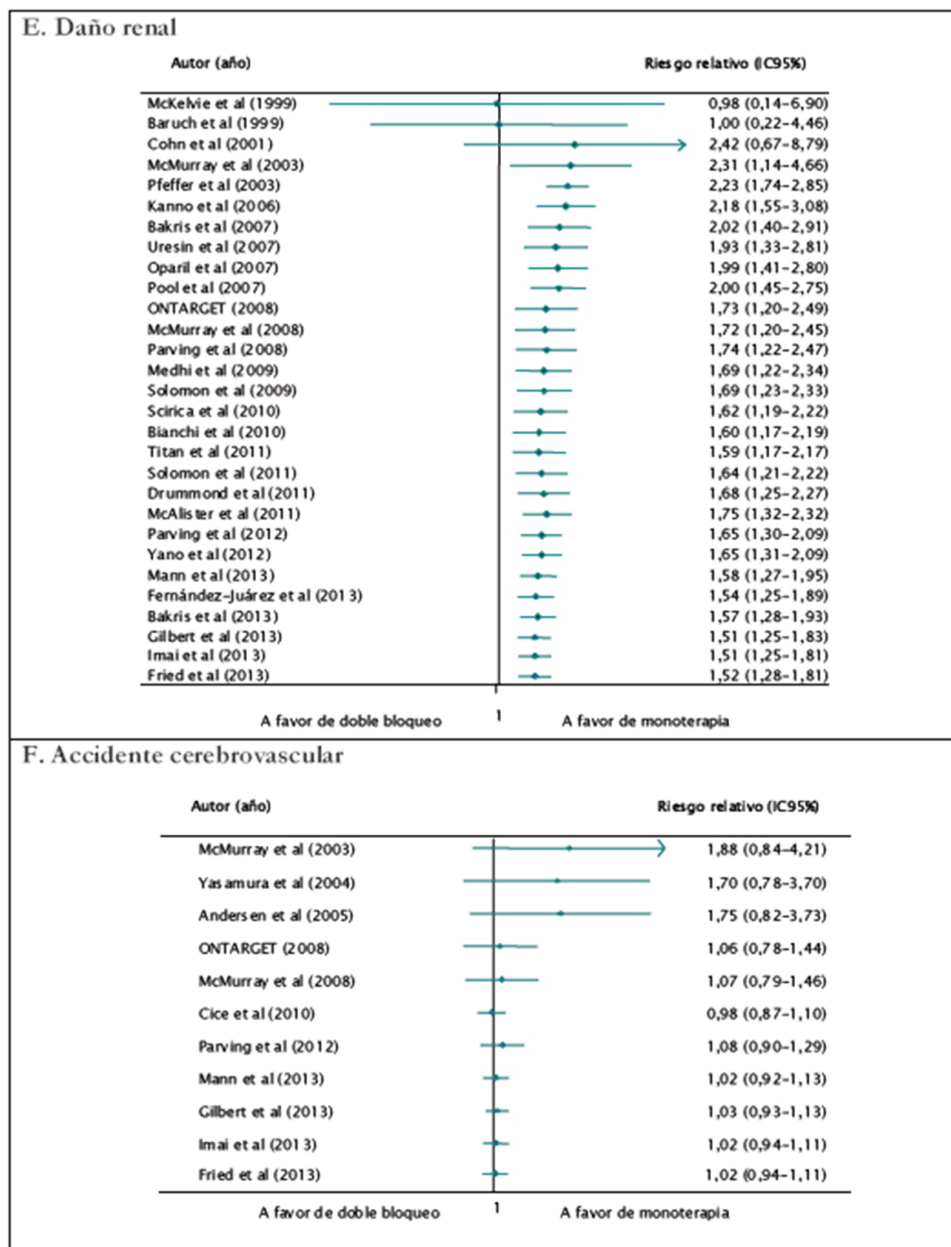


Figura 2
Metaanálisis acumulativo (continuación)

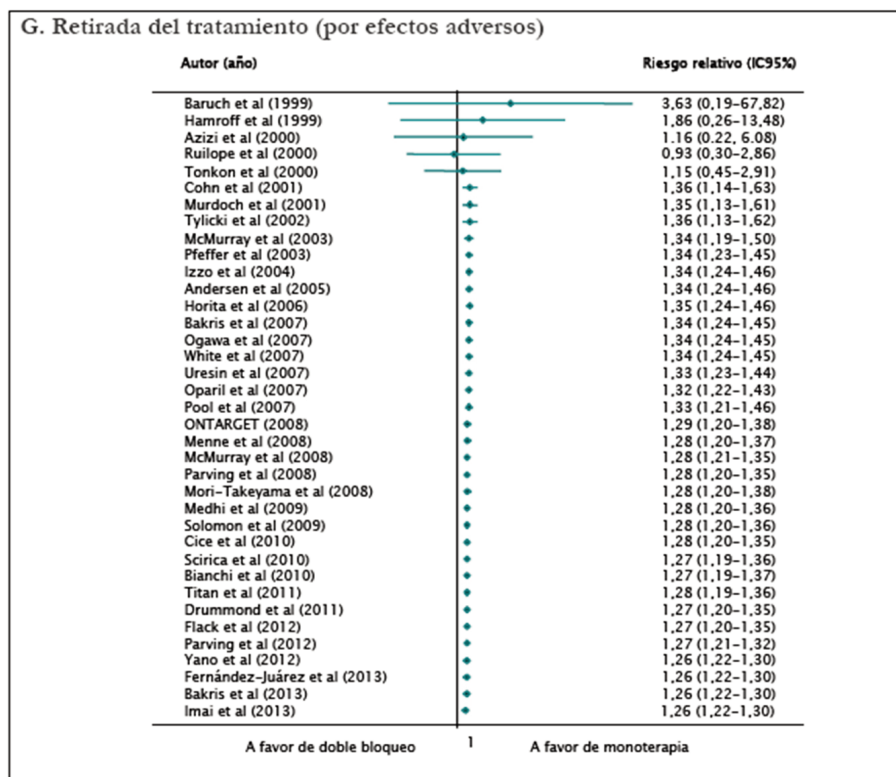


Tabla 5
Principales metaanálisis previos sobre el efecto del bloqueo dual del sistema renina-angiotensina

Autor y año	Estudios incluidos	Población estudiada	Intervenciones	Principales resultados de eficacia (para la combinación)	Principales resultados de seguridad (para la combinación)	Principales limitaciones
Doulton et al, 2005	14 ensayos clínicos (4 grupos paralelos, 10 grupos cruzados) 434 pacientes	Pacientes con hipertensión, diabetes mellitus, ERC	IECA + ARA-II frente a monoterapia	Reducción de la proteinuria y de la PAS	Hiperpotasemia	Ensayos de pequeño tamaño (solo 3 estudios con más de 30 pacientes tratados en combinación) y corta duración (solo 2 estudios tratamiento de ≥ 12 semanas)
MacKinnon et al, 2006	21 ensayos clínicos (5 grupos paralelos, 16 grupos cruzados) 654 pacientes	Pacientes con ERC y proteinuria	ECA + ARA-II frente a monoterapia	Reducción de la proteinuria	Hiperpotasemia, Reducción en TFG	Ensayos de pequeño tamaño (solo 4 estudios con más de 30 pacientes tratados en combinación) y corta duración (solo 7 estudios tratamiento de ≥ 12 semanas)
Phillips et al, 2007	4 ensayos clínicos (4 grupos paralelos) 17337 pacientes	Pacientes con IAM, insuficiencia cardíaca	ECA + ARA-II frente a monoterapia	-	Hiperpotasemia, Hipotensión, Insuficiencia renal, Retirada de tratamiento por efectos adversos	Reducido número de estudios
Jenning et al, 2007	10 ensayos clínicos (9 grupos cruzados, 1 grupo paralelos) 156 pacientes	Pacientes con nefropatía diabética	ECA + ARA-II frente a monoterapia	Reducción de la proteinuria	Hiperpotasemia Reducción en TFG Aumento CrS	Ensayos de pequeño tamaño (solo 2 estudios con más de 30 pacientes tratados en combinación) y corta duración (9 estudios de 8-12 semanas)
Lakhdar et al, 2008	9 ensayos clínicos (9 grupos paralelos) 18160 pacientes	Pacientes con insuficiencia cardíaca	ECA + ARA-II frente a monoterapia	-	Hiperpotasemia, Hipotensión, Insuficiencia renal, Aumento cualquier efecto adverso	Reducido número de acontecimientos adversos por estudio
Kunz et al, 2008	49 ensayos clínicos (23 compararon doble bloqueo frente a monoterapia) 6181 pacientes	Pacientes con o sin diabetes mellitus y microalbuminuria o proteinuria	ECA + ARA-II frente a monoterapia	Reducción de la proteinuria	Retirada de tratamiento por efectos adversos (no significativa)	La mayoría de estudios con reducido tamaño y no presentaban datos de seguridad
Catapano et al, 2008	13 ensayos clínicos (5 grupos paralelos y 8 grupos cruzados) 425 pacientes	Pacientes con glomerulonefritis	ECA + ARA-II frente a monoterapia	Reducción de la proteinuria y de la PAS	Hiperpotasemia	La mayoría de estudios con reducido tamaño

Tabla 5**Principales metaanálisis previos sobre el efecto del bloqueo dual del sistema renina-angiotensina (continuación)**

Autor y año	Estudios incluidos	Población estudiada	Intervenciones	Principales resultados de eficacia (para la combinación)	Principales resultados de seguridad (para la combinación)	Principales limitaciones
Kuenzli et al, 2010	8 ensayos clínicos (8 con grupos paralelos) 18061 pacientes	Pacientes con insuficiencia cardíaca	IECA + ARA-II frente a monoterapia	Reducción hospitalizaciones por insuficiencia cardíaca, No reducción de la mortalidad, de hospitalizaciones por cualquier causa o IAM	Hiperpotasemia (no significativa), Hipotensión, Insuficiencia renal, Retirada de tratamiento por efectos adversos	Para la definición de doble bloqueo, existía variabilidad entre estudios ya que no se exigía que todos los pacientes estuvieran tratados con terapia combinada Reducido número de acontecimientos adversos por estudio
Maione et al, 2011	85 ensayos clínicos (12 compararon doble bloqueo frente a monoterapia) 21708 pacientes	Pacientes con micro- o macroalbuminuria y otros factores de riesgo cardiovascular	IECA + ARA-II frente a monoterapia	Reducción progresión de microalbuminuria a macroalbuminuria	Hipotensión	La mayoría de estudios con reducido tamaño y no presentaban datos de seguridad para el doble bloqueo
Pham et al, 2012	16 ensayos clínicos (2 grupos paralelos y 14 grupos cruzados) 992 pacientes	Pacientes con nefropatía diabética	IECA + ARA-II o aliskireno o antagonista aldosterona frente a monoterapia	Reducción de la proteinuria Reducción de la PAS	Hiperpotasemia	La mayoría de estudios con reducido tamaño (n<30) y de corta duración
Cheng et al, 2012	6 ensayos clínicos (-) 109 pacientes	Nefropatía inmunológica (IgA)	IECA + ARA-II frente a monoterapia	Reducción de la proteinuria No reducción PAS	Hiperpotasemia (no significativa) No efecto en TFG	La mayoría de estudios con reducido tamaño y no presentaban datos de seguridad para el doble bloqueo
Harel et al, 2012	10 ensayos clínicos (10 grupos paralelos) 4814 pacientes	Pacientes con hipertensión, IAM, insuficiencia cardíaca, diabetes mellitus	Aliskireno + IECA o ARA-II frente a monoterapia	-	Hiperpotasemia Insuficiencia renal (no significativa)	Reducido número de acontecimientos adversos por estudio
Makani et al, 2013	33 ensayos clínicos (33 grupos paralelos) 68405 pacientes	Pacientes con insuficiencia cardíaca (o no), hipertensión, diabetes mellitus, ERC, IAM, etc.	Aliskireno + IECA o ARA-II, IECA + ARA-II frente a monoterapia	Reducción hospitalizaciones por insuficiencia cardíaca. No reducción de la mortalidad (cualquier causa o cardíaca)	Hiperpotasemia, Hipotensión, Insuficiencia renal, Retirada de tratamiento por efectos adversos	No incluye subgrupos para algunas subpoblaciones de interés (diabetes mellitus, ERC, etc.)

Tabla 5
Principales metaanálisis previos sobre el efecto del bloqueo dual del sistema renina-angiotensina (continuación)

Autor y año	Estudios incluidos	Población estudiada	Intervenciones	Principales resultados de eficacia (para la combinación)	Principales resultados de seguridad (para la combinación)	Principales limitaciones
Catalá-López y Macías-Saint Gerons, 2013	9 ensayos clínicos (9 grupos paralelos) 11.543 pacientes	Pacientes con diabetes mellitus	Aliskireno + IECA o ARA-II, IECA + ARA-II, frente a monoterapia	No efecto en la reducción de la mortalidad (cualquier causa)	Hiperpotasemia Hipotensión Insuficiencia renal (no significativa)	Reducido número de acontecimientos adversos por estudio No incluye reanálisis de estudios publicados recientemente y no diferencia entre opciones de bloqueo dual
Susantitaphong et al, 2013	59 ensayos clínicos (34 grupos paralelos y 25 grupos cruzados) 4.975 pacientes	Pacientes con ERC	Aliskireno + IECA o ARA-II, IECA + ARA-II frente a monoterapia	Reducción de la proteinuria y albuminuria. No efecto en la reducción de la mortalidad (cualquier causa), CrS u hospitalizaciones	Hiperpotasemia Hipotensión	La mayoría de estudios con reducido tamaño

ARA-II: antagonistas de los receptores de la angiotensina II; CrS: creatinina sérica; ERC: enfermedad renal crónica; IECA: inhibidores enzima convertidora de angiotensina; TFG: tasa filtración glomerular; IAM: infarto agudo de miocardio; PAS: presión arterial sistólica

DISCUSIÓN

En este trabajo se ha revisado la evidencia publicada sobre la seguridad de la terapia dual con inhibidores del SRA, procedente tanto de ensayos clínicos como de estudios observacionales. Los resultados obtenidos a partir de la síntesis cuantitativa con metaanálisis sugieren que el bloqueo dual del SRA no ofreció ningún beneficio adicional (respecto a la monoterapia) en la reducción de la mortalidad general, la mortalidad cardiovascular o el accidente cerebrovascular. En cambio, la terapia dual incrementó los riesgos de hiperpotasemia, hipotensión, insuficiencia renal y la interrupción del tratamiento por efectos adversos. En términos generales, se observó que estos hallazgos son consistentes en las cohortes de pacientes con diabetes mellitus, enfermedad renal o insuficiencia cardíaca y en la mayoría de casos los efectos se mantienen de forma estable durante más de una década. Cabría mencionar que los resultados del incremento de riesgos asociado a la terapia dual estarían en consonancia con los obtenidos en trabajos previos utilizando criterios de análisis más restrictivos^{5,10-13}, así como con los resultados de estudios que no fueron incluidos en la revisión por no cumplir los criterios de inclusión²². En concreto, los resultados del ensayo *The Aliskiren Trial on Acute Heart Failure Outcomes* (ASTRONAUT)²² mostraron que el inhibidor directo de la renina, aliskireno, añadido a la terapia antihipertensiva estándar (por ejemplo aproximadamente el 85% de los 1.639 pacientes recibía bloqueantes del SRA como IECA o ARA-II) no mejoró la mortalidad cardiovascular después del alta ni los reingresos en pacientes hospitalizados por insuficiencia cardíaca, mientras que se asoció con un incremento del 37% en el riesgo de insuficiencia renal, un aumento del 36% en el riesgo de hipotensión y un aumento del 19% en el riesgo de hiperpotasemia. Por otro lado, recientemente se han publicado los resultados del ensayo *The Combination Angiotensin Receptor Blocker and Angio-*

tensin-Converting Enzyme Inhibitor for Treatment of Diabetic Nephropathy (VA NEPHRON-D)¹⁶ en el que se estudiaba en 1.448 pacientes diabéticos el efecto sobre la progresión de la enfermedad renal de la combinación de ARA-II e IECA (losartán y lisinopril) en comparación con ARA-II en monoterapia. Este ensayo tuvo que finalizar prematuramente debido a un mayor número de lesiones renales e hiperpotasemia en los pacientes con nefropatía diabética tratados con terapia dual¹⁶. Existen al menos 3 estudios²³⁻²⁵ que están actualmente en curso y que se espera que generen datos adicionales: *Long-term Impact of RAS Inhibition on Cardiorenal Outcomes* (LIRICO)²³, *Preventing ESRD in Overt Nephropathy of Type 2 Diabetes* (VALID)²⁴ y *Efficacy and Safety of Aliskiren and Aliskiren/Enalapril Combination on Morbi-mortality in Patients With Chronic Heart Failure* (ATMOSPHERE)²⁵. LIRICO es un ensayo pragmático que evalúa el efecto cardiovascular y renal de la terapia dual con IECA y ARA-II respecto a la monoterapia en pacientes con albuminuria y al menos un factor de riesgo cardiovascular. El ensayo VALID tiene como objetivo evaluar el efecto del bloqueo dual con IECA (benazepril) y ARA-II (valsartán) sobre la progresión de enfermedad renal terminal en pacientes de alto riesgo con nefropatía diabética. ATMOSPHERE es un ensayo clínico que evalúa el efecto del bloqueo dual con aliskireno e IECA (enalapril) sobre la mortalidad cardiovascular y hospitalización por insuficiencia cardíaca.

Algunos estudios y metaanálisis previos en pacientes con insuficiencia cardíaca han mostrado que la combinación de dos bloqueadores del SRA puede reducir el riesgo de hospitalizaciones por insuficiencia cardíaca, pero no se han observado beneficios consistentes en la reducción de la mortalidad general, cardiovascular o en el riesgo de hospitalizaciones por otras causas^{5,9}. Concretamente, los principales datos sobre el bloqueo dual del SRA en pacientes con insuficiencia cardíaca proviene de diversos

ensayos clínicos entre los que destacan: *Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity – Added* (CHARM-Added)²⁶, *Valsartan Heart Failure Trial* (Val-HeFT)^{27,28} y *Valsartan in Acute Myocardial Infarction Trial* (VALIANT)²⁹. CHARM-Added ha sido el único ensayo en el que se ha observado una disminución de la mortalidad cardiovascular (reducción absoluta del riesgo del 3,6%) y de las hospitalizaciones por insuficiencia cardíaca (reducción absoluta del riesgo del 3,8%) con la terapia dual (candesartán junto a IECA) respecto a la monoterapia. Sin embargo, la mortalidad general no fue diferente entre los distintos grupos. Igualmente, en el ensayo Val-HeFT^{27,28} se observó una disminución de las hospitalizaciones por insuficiencia cardíaca (reducción absoluta del riesgo del 4,4%), aunque la mortalidad general y la mortalidad cardiovascular fueron similares en los grupos de bloqueo dual (93% pacientes con valsartán e IECA) y monoterapia. Mientras que en el ensayo VALIANT²⁹ la combinación de ARA-II (valsartán) e IECA (captopril) aumentó la ocurrencia de efectos adversos sin mejorar la supervivencia.

Por otro lado, un estudio reciente¹¹ ha sugerido que la terapia dual podría tener efectos beneficiosos a corto plazo, como la reducción de la proteinuria en pacientes con enfermedad renal. Sin embargo, este hallazgo sería contrario a los resultados obtenidos en grandes estudios que valoran acontecimientos renales^{3,4,21,30}. Algunas guías de práctica clínica (como es el caso de las guías europeas 2013 sobre el manejo de la hipertensión arterial y diabetes^{31,32}) desaconsejan el uso de la combinación de dos inhibidores del RAS. La ficha técnica de algunos medicamentos no recomienda el uso concomitante de IECA y ARA-II y en los casos individualmente definidos se recomienda seguir estrechamente la función renal y las concentraciones sanguíneas de potasio^{33,34}. En 2012, como consecuencia de los resultados del ensayo ALTITUDE, se contraindicó el

bloqueo dual del SRA conjuntamente con aliskireno e IECA/ARA-II en pacientes diabéticos o con insuficiencia renal moderada-grave y se desaconsejó su uso en el resto de pacientes³⁵. Más recientemente, teniendo en cuenta la información que se ha generado en los últimos años, la Agencia Europea de Medicamentos ha anunciado el inicio de una revisión de los riesgos del bloqueo dual del SRA en el tratamiento de la hipertensión y la insuficiencia cardíaca congestiva³⁶.

El SRA es importante en la regulación de la presión arterial. La inhibición del SRA produciría incrementos en la secreción de renina y, por lo tanto, de la actividad de la renina plasmática. Algunos estudios han sugerido que valores altos de actividad de renina plasmática se han asociado con mayor incidencia de infarto de miocardio, peor pronóstico de la insuficiencia cardíaca y aumento de la mortalidad cardiovascular, así como empeoramiento de la función renal^{37,38}. La aparición de efectos adversos asociados al bloqueo dual del SRA podría suceder en pacientes susceptibles, como por ejemplo aquellos con depleción de volumen y/o sodio ocasionada por un tratamiento intensivo con diuréticos o por una dieta restrictiva en sal, entre otras posibles causas^{38,39}. En pacientes cuyo tono vascular y función renal dependen más de la actividad del SRA (por ejemplo personas con insuficiencia cardíaca congestiva o enfermedad renal subyacente) el tratamiento con medicamentos que afectan al SRA también se ha asociado con hipotensión o, raramente, con fallo renal agudo³³.

Cabe señalar algunas consideraciones en relación a este trabajo. Como en toda evaluación, las inferencias que puedan hacerse de los resultados presentados dependen de la elección de las fuentes de información y de los modelos empleados para combinar los diferentes estudios. En este sentido, los autores decidieron *a priori* incluir la evidencia de los estudios observacionales además de la procedente de ensayos clínicos, asu-

miendo que un estudio observacional de gran tamaño puede facilitar información relevante a pesar de que ello pudiera introducir cierta heterogeneidad a los resultados. Sin embargo, en este trabajo, la dirección y la magnitud de los riesgos asociados al bloqueo dual del SRA fueron consistentes cuando el análisis se restringió a ECA solos. Por otro lado, se decidió emplear el modelo de efectos aleatorios con el objetivo de poder generalizar los resultados a poblaciones más amplias y heterogéneas siendo conscientes que los estudios seleccionados no necesariamente presentaban características idénticas. Como en cualquier revisión, este análisis no está exento de una posible falta de identificación de algún estudio, si bien esto se ha intentado reducir utilizando criterios de inclusión predefinidos, y una búsqueda lo más amplia posible. Por último, no es infrecuente encontrar cierto grado de variabilidad en relación a la forma en la que se describen los efectos adversos y riesgos en las publicaciones (en ocasiones de baja calidad)⁴⁰⁻⁴², lo cual, añadido al reducido número de estudios identificados en algunos subgrupos, podría haber dificultado algunas estimaciones del efecto de los tratamientos.

Concluyendo, en comparación con la monoterapia, el bloqueo dual del SRA no parece ofrecer ningún beneficio adicional en la reducción de la mortalidad general, la mortalidad cardiovascular o accidente cerebrovascular, pero incrementó los riesgos de hiperpotasemia, hipotensión, insuficiencia renal y la interrupción del tratamiento por efectos adversos. Estos hallazgos son consistentes a lo largo del tiempo. El uso de la terapia dual del SRA, por tanto, parece desaconsejable y en los casos individualmente definidos en los que el clínico decida iniciar el tratamiento, se debería monitorizar y seguir estrechamente la función renal, las concentraciones plasmáticas de potasio y las cifras de presión arterial.

BIBLIOGRAFÍA

1. Ritter JM. Angiotensin converting enzyme inhibitors and angiotensin receptor blockers in hypertension. *BMJ*. 2011;342:d1673.
2. Menard J, Campbell DJ, Azizi M, Gonzales MF. Synergistic effects of ACE inhibition and Ang II antagonism on blood pressure, cardiac weight, and renin in spontaneously hypertensive rats. *Circulation*. 1997;96:3072-78.
3. Yusuf S, Teo KK, Pogue J, Dyal L, Copland I, Schumacher H, et al. Telmisartan, ramipril, or both in patients at high risk for vascular events. *N Engl J Med*. 2008;358:1547-59.
4. Parving HH, Brenner BM, McMurray JJ, de Zeeuw D, Haffner SM, Solomon SD, et al. Cardiorenal end points in a trial of aliskiren for type 2 diabetes. *N Engl J Med*. 2012;367:2204-13.
5. Makani H, Bangalore S, Desouza KA, Shah A, Messerli FH. Efficacy and safety of dual blockade of the renin-angiotensin system: meta-analysis of randomised trials. *BMJ*. 2013;346:f360.
6. Doultou TW, He FJ, MacGregor GA. Systematic review of combined angiotensin-converting enzyme inhibition and angiotensin receptor blockade in hypertension. *Hypertension*. 2005;45:880-6.
7. Phillips CO, Kashani A, Ko DK, Francis G, Krumholz HM. Adverse effects of combination angiotensin II receptor blockers plus angiotensin-converting enzyme inhibitors for left ventricular dysfunction: a quantitative review of data from randomized clinical trials. *Arch Intern Med*. 2007;167:1930-6.
8. Kunz R, Friedrich C, Wolbers M, Mann JF. Meta-analysis: effect of monotherapy and combination therapy with inhibitors of the renin angiotensin system on proteinuria in renal disease. *Ann Intern Med*. 2008;148:30-48.
9. Kuenzli A, Bucher HC, Anand I, Arutunov G, Kum LC, McKelvie R, et al. Meta-analysis of combined therapy with angiotensin receptor antagonists versus ACE inhibitors alone in patients with heart failure. *PLoS One*. 2010;5:e9946.
10. Harel Z, Gilbert C, Wald R, Bell C, Perl J, Juurlink D, et al. The effect of combination treatment with aliskiren and blockers of the renin-angiotensin system on hyperkalaemia and acute kidney injury: systematic review and meta-analysis. *BMJ*. 2012;344:e42.

11. Susantitaphong P, Sewaralthahab K, Balk EM, Eiam-ong S, Madias NE, Jaber BL. Efficacy and safety of combined vs. single renin-angiotensin-aldosterone system blockade in chronic kidney disease: a meta-analysis. *Am J Hypertens*. 2013;26:424-41.
12. Catalá-López F, Macías Saint-Gerons D. Diabetes Mellitus and Risks of Dual Blockade of the Renin-angiotensin-aldosterone System. *Rev Esp Cardiol*. 2013;66:412-5.
13. Bakris GL, Oparil S, Purkayastha D, Yadao AM, Alessi T, Sowers JR. Randomized study of antihypertensive efficacy and safety of combination aliskiren/valsartan vs valsartan monotherapy in hypertensive participants with type 2 diabetes mellitus. *J Clin Hypertens (Greenwich)*. 2013;15:92-100.
14. Fernández Juárez G, Luño J, Barrio V, de Vinuesa SG, Praga M, Goicoechea M, et al. Effect of dual blockade of the renin-angiotensin system on the progression of type 2 diabetic nephropathy: a randomized trial. *Am J Kidney Dis*. 2013;61:211-8.
15. Mann JF, Anderson C, Gao P, Gerstein HC, Boehm M, Rydén L, et al. Dual inhibition of the renin-angiotensin system in high-risk diabetes and risk for stroke and other outcomes: results of the ONTARGET trial. *J Hypertens*. 2013;31:414-21.
16. Fried LF, Emanuele N, Zhang JH, Brophy M, Conner TA, Duckworth W, et al. Combined Angiotensin inhibition for the treatment of diabetic nephropathy. *N Engl J Med*. 2013;369:1892-903.
17. McAlister FA, Zhang J, Tonelli M, Klarenbach S, Manns BJ, Hemmelgarn BR; Alberta Kidney Disease Network. The safety of combining angiotensin-converting-enzyme inhibitors with angiotensin-receptor blockers in elderly patients: a population-based longitudinal analysis. *CMAJ*. 2011;183:655-62.
18. Gilbert CJ, Gomes T, Mamdani MM, Hellings C, Yao Z, Garg AX, et al. No increase in adverse events during aliskiren use among Ontario patients receiving angiotensin-converting enzyme inhibitors or angiotensin receptor blockers. *Can J Cardiol*. 2013;29:586-91.
19. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7:177-88.
20. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557-60.
21. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med*. 1992;327:248-54.
22. Gheorghiade M, Böhm M, Greene SJ, Fonarow GC, Lewis EF, Zannad F, et al. Effect of aliskiren on postdischarge mortality and heart failure readmissions among patients hospitalized for heart failure: the ASTRONAUT randomized trial. *JAMA*. 2013;309:1125-35.
23. Maione A, Nicolucci A, Craig JC, Tognoni G, Moschetti A, Palasciano G, et al. Protocol of the Long-term Impact of RAS Inhibition on Cardiorenal Outcomes (LIRICO) randomized trial. *J Nephrol*. 2007;20:646-55.
24. Mario Negri Institute for Pharmacological Research. Preventing ESRD in Overt Nephropathy of Type 2 Diabetes (VALID, NCT00494715). Disponible en: <http://clinicaltrials.gov/ct2/show/NCT00494715?term=VALID+trial&rank=1>
25. Krum H, Massie B, Abraham WT, Dickstein K, Kober L, McMurray JJ, et al. Direct renin inhibition in addition to or as an alternative to angiotensin converting enzyme inhibition in patients with chronic systolic heart failure: rationale and design of the Aliskiren Trial to Minimize OutcomeS in Patients with HEart failure (ATMOSPHERE) study. *Eur J Heart Fail*. 2011;13:107-14.
26. McMurray JJ, Ostergren J, Swedberg K, Granger CB, Held P, Michelson EL, et al. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function taking angiotensin-converting-enzyme inhibitors: the CHARM-Added trial. *Lancet*. 2003;362:767-71.
27. Cohn JN, Tognoni G; Valsartan Heart Failure Trial Investigators. A randomized trial of the angiotensin-receptor blocker valsartan in chronic heart failure. *N Engl J Med*. 2000;345:1667-75.
28. Lesogor A, Cohn JN, Latini R, Tognoni G, Krum H, Massie B, et al. Interaction between baseline and early worsening of renal function and efficacy of renin-angiotensin-aldosterone system blockade in patients with heart failure: insights from the Val-HeFT study. *Eur J Heart Fail*. 2013 Jun 19.
29. Pfeffer MA, McMurray JJ, Velazquez EJ, Rouleau JL, Køber L, Maggioni AP, et al. Valsartan, captopril, or both in myocardial infarction complicated by heart failure, left ventricular dysfunction, or both. *N Engl J Med*. 2003;349:1893-906.
30. Imai E, Haneda M, Yamasaki T, Kobayashi F, Hara-da A, Ito S, et al. Effects of dual blockade of the renin-angiotensin system on renal and cardiovascular outcomes in type 2 diabetes with overt nephropathy and hypertension in the ORIENT: a post-hoc analysis (ORIENT-Hypertension). *Hypertens Res*. 2013 Sep 12.

31. Mancia G, Fagard R, Narkiewicz K, Redón J, Zanchetti A, Böhm M, et al. 2013 ESH/ESC Guidelines for the management of arterial hypertension: the Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *J Hypertens*. 2013;31:1281-357.
32. Authors/Task Force Members, Rydén L, Grant PJ, Anker SD, Berne C, Cosentino F, et al. ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD: The Task Force on diabetes, pre-diabetes, and cardiovascular diseases of the European Society of Cardiology (ESC) and developed in collaboration with the European Association for the Study of Diabetes (EASD). *Eur Heart J*. 2013 Aug 30. [Epub ahead of print]
33. Agencia Europea de Medicamentos. Ficha técnica de Micardis® (telmisartán) Disponible en: http://www.ema.europa.eu/docs/es_ES/document_library/EPAR_-_Product_Information/human/000209/WC500027641.pdf
34. Agencia Española de Medicamentos y Productos Sanitarios. Ficha técnica de Renitec® (enalapril) Disponible en: <http://www.aemps.gob.es/cima/especialidad.do?metodo=verFichaWordPdf&codigo=57773&formato=pdf&formulario=FICHAS&file=ficha.pdf>
35. Agencia Española de Medicamentos y Productos Sanitarios (AEMPS). Nota informativa MUH (FV), 03/2012. Aliskireno (Rasilez®, Riprazo®, Rasilez HTC®): conclusiones de la re-evaluación del balance beneficio riesgo. Madrid, febrero de 2012. Disponible en: http://www.aemps.gob.es/informa/notasInformativas/medicamentosUsoHumano/seguridad/2012/NI-MUH_03-2012.htm
36. Agencia Europea de Medicamentos. Review started of combined use of renin-angiotensin system (RAS)-acting agents. Londres, mayo de 2013. Disponible en [http://www.emea.europa.eu/docs/en_GB/document_library/Referrals_document/Renin-angiotensin_system_\(RAS\)-acting_agents/Procedure_started/WC500143500.pdf](http://www.emea.europa.eu/docs/en_GB/document_library/Referrals_document/Renin-angiotensin_system_(RAS)-acting_agents/Procedure_started/WC500143500.pdf)
37. Morales-Olivas FJ, Estañ L. La actividad de la renina plasmática. Papel de la inhibición directa de la renina. *Rev Esp Cardiol*. 2011;Suppl 11(D):3-7.
38. Sealey JE, Alderman MH, Furberg CD, Laragh JH. Renin-angiotensin system blockers may create more risk than reward for sodium-depleted cardiovascular patients with high plasma renin levels. *Am J Hypertens*. 2013;26:727-38.
39. Laragh JH, Sealey JE. Sodium depletion in patients in clinical trials may account for the increased cardiovascular risk of dual blockade of the renin-angiotensin system. *BMJ*. 2013;346:f1685.
40. Bagul NB, Kirkham JJ. The reporting of harms in randomized controlled trials of hypertension using the CONSORT criteria for harm reporting. *Clin Exp Hypertens*. 2012;34:548-54.
41. Haidich AB, Birtsou C, Dardavessis T, Tirodimos I, Arvanitidou M. The quality of safety reporting in trials is still suboptimal: survey of major general medical journals. *J Clin Epidemiol*. 2011;64:124-35.
42. Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*. 2004;141:781-8.

Anexo 1

Estrategia de búsqueda para metaanálisis

Bloque 1. Tratamiento farmacológico: (("renin-angiotensin system"[MeSH Terms] OR ("renin-angiotensin"[All Fields] AND "system"[All Fields]) OR "renin-angiotensin system"[All Fields] OR ("renin"[All Fields] AND "angiotensin"[All Fields] AND "system"[All Fields]) OR "renin angiotensin system"[All Fields]) OR ("renin-angiotensin system"[MeSH Terms] OR ("renin-angiotensin"[All Fields] AND "system"[All Fields]) OR "renin-angiotensin system"[All Fields] OR ("renin"[All Fields] AND "angiotensin"[All Fields] AND "aldosterone"[All Fields] AND "system"[All Fields]) OR "renin angiotensin aldosterone system"[All Fields]) OR ("angiotensin-converting enzyme inhibitors"[Pharmacological Action] OR "angiotensin-converting enzyme inhibitors"[MeSH Terms] OR ("angiotensin-converting"[All Fields] AND "enzyme"[All Fields] AND "inhibitors"[All Fields]) OR "angiotensin-converting enzyme inhibitors"[All Fields] OR ("angiotensin"[All Fields] AND "converting"[All Fields] AND "enzyme"[All Fields] AND "inhibitor"[All Fields]) OR "angiotensin converting enzyme inhibitor"[All Fields]) OR (("receptors, angiotensin"[MeSH Terms] OR ("receptors"[All Fields] AND "angiotensin"[All Fields]) OR "angiotensin receptors"[All Fields] OR ("angiotensin"[All Fields] AND "receptor"[All Fields]) OR "angiotensin receptor"[All Fields]) AND blocker[All Fields]) OR ("renin"[MeSH Terms] OR "renin"[All Fields]) AND inhibitor[All Fields]) OR (direct[All Fields] AND ("renin"[MeSH Terms] OR "renin"[All Fields]) AND inhibitor[All Fields]))

Bloque 2. Revisiones sistemáticas y metaanálisis: (("review"[Publication Type] OR "review literature as topic"[MeSH Terms] OR "systematic review"[All Fields]) AND ("meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields])).

Bloque 3. Combinación del tratamiento: ((dual[All Fields] AND blockade[All Fields]) OR (dual[All Fields] AND ("therapy"[Subheading] OR "therapy"[All Fields] OR "therapeutics"[MeSH Terms] OR "therapeutics"[All Fields])) OR ("combined modality therapy"[MeSH Terms] OR ("combined"[All Fields] AND "modality"[All Fields] AND "therapy"[All Fields]) OR "combined modality therapy"[All Fields] OR ("combination"[All Fields] AND "therapy"[All Fields]) OR "combination therapy"[All Fields]) OR combined[All Fields] OR addition[All Fields] OR added[All Fields])).

Límites utilizados: hasabstract[text] AND ("2003/08/13"[PDAT] : "2013/11/13"[PDAT]) AND "humans"[MeSH Terms]) AND (English[lang] OR Spanish[lang]).

Anexo 2

Estrategia de búsqueda para estudios adicionales

Bloque 1. Tratamiento farmacológico: (((("renin-angiotensin system"[MeSH Terms] OR ("renin-angiotensin"[All Fields] AND "system"[All Fields]) OR "renin-angiotensin system"[All Fields] OR ("renin"[All Fields] AND "angiotensin"[All Fields] AND "system"[All Fields]) OR "renin angiotensin system"[All Fields]) OR ("renin-angiotensin system"[MeSH Terms] OR ("renin-angiotensin"[All Fields] AND "system"[All Fields]) OR "renin-angiotensin system"[All Fields] OR ("renin"[All Fields] AND "angiotensin"[All Fields] AND "aldosterone"[All Fields] AND "system"[All Fields]) OR "renin angiotensin aldosterone system"[All Fields]) OR ("angiotensin-converting enzyme inhibitors"[Pharmacological Action] OR "angiotensin-converting enzyme inhibitors"[MeSH Terms] OR ("angiotensin-converting"[All Fields] AND "enzyme"[All Fields] AND "inhibitors"[All Fields]) OR "angiotensin-converting enzyme inhibitors"[All Fields] OR ("angiotensin"[All Fields] AND "converting"[All Fields] AND "enzyme"[All Fields] AND "inhibitor"[All Fields]) OR "angiotensin converting enzyme inhibitor"[All Fields]) OR ("receptors, angiotensin"[MeSH Terms] OR ("receptors"[All Fields] AND "angiotensin"[All Fields]) OR "angiotensin receptors"[All Fields] OR ("angiotensin"[All Fields] AND "receptor"[All Fields]) OR "angiotensin receptor"[All Fields]) AND blocker[All Fields]) OR ("renin"[MeSH Terms] OR "renin"[All Fields] AND inhibitor[All Fields]) OR (direct[All Fields] AND ("renin"[MeSH Terms] OR "renin"[All Fields] AND inhibitor[All Fields]) OR ("benazepril"[Supplementary Concept] OR "benazepril"[All Fields]) OR ("captopril"[MeSH Terms] OR "captopril"[All Fields]) OR ("enalapril"[MeSH Terms] OR "enalapril"[All Fields]) OR ("cilazapril"[MeSH Terms] OR "cilazapril"[All Fields]) OR ("delapril"[Supplementary Concept] OR "delapril"[All Fields]) OR ("fosinopril"[MeSH Terms] OR "fosinopril"[All Fields]) OR ("imidapril"[Supplementary Concept] OR "imidapril"[All Fields]) OR ("lisinopril"[MeSH Terms] OR "lisinopril"[All Fields]) AND ("moexipril"[Supplementary Concept] OR "moexipril"[All Fields])) OR ("perindopril"[MeSH Terms] OR "perindopril"[All Fields]) AND ("quinapril"[Supplementary Concept] OR "quinapril"[All Fields])) OR ("ramipril"[MeSH Terms] OR "ramipril"[All Fields]) OR ("spirapril"[Supplementary Concept] OR "spirapril"[All Fields]) OR ("temocapril hydrochloride"[Supplementary Concept] OR "temocapril hydrochloride"[All Fields] OR "temocapril"[All Fields]) OR ("trandolapril"[Supplementary Concept] OR "trandolapril"[All Fields]) OR ("zofenopril"[Supplementary Concept] OR "zofenopril"[All Fields]) OR ("candesartan"[Supplementary Concept] OR "candesartan"[All Fields]) OR ("irbesartan"[Supplementary Concept] OR "irbesartan"[All Fields]) OR ("losartan"[MeSH Terms] OR "losartan"[All Fields]) OR ("telmisartan"[Supplementary Concept] OR "telmisartan"[All Fields]) OR ("valsartan"[Supplementary Concept] OR "valsartan"[All Fields]) OR ("olmesartan"[Supplementary Concept] OR "olmesartan"[All Fields]) OR ("eprosartan"[Supplementary Concept] OR "eprosartan"[All Fields]) OR ("azilsartan"[Supplementary Concept] OR "azilsartan"[All Fields]) OR ("aliskiren"[Supplementary Concept] OR "aliskiren"[All Fields]))

Bloque 2. Diseño de los estudios: (("randomized controlled trial"[Publication Type] OR "randomized controlled trials as topic"[MeSH Terms] OR "randomized controlled trial"[All Fields] OR "randomised controlled trial"[All Fields]) OR cohort[All Fields] OR case-control[All Fields]).

Bloque 2. Diseño de los estudios: (("randomized controlled trial"[Publication Type] OR "randomized controlled trials as topic"[MeSH Terms] OR "randomized controlled trial"[All Fields] OR "randomised controlled trial"[All Fields]) OR cohort[All Fields] OR case-control[All Fields]).

Bloque 3. Combinación del tratamiento: ((dual[All Fields] AND blockade[All Fields]) OR (dual[All Fields] AND ("therapy"[Subheading] OR "therapy"[All Fields] OR "therapeutics"[MeSH Terms] OR "therapeutics"[All Fields])) OR ("combined modality therapy"[MeSH Terms] OR ("combined"[All Fields] AND "modality"[All Fields] AND "therapy"[All Fields]) OR "combined modality therapy"[All Fields] OR ("combination"[All Fields] AND "therapy"[All Fields]) OR "combination therapy"[All Fields]) OR combined[All Fields] OR addition[All Fields] OR added[All Fields])

Límites utilizados: hasabstract[text] AND ("2003/08/13"[PDAT] : "2013/11/13"[PDAT]) AND "humans"[MeSH Terms] AND (English[lang] OR Spanish[lang])

Anexo 3

Metaanálisis previos

1. Doulton TW, He FJ, MacGregor GA. Systematic review of combined angiotensin-converting enzyme inhibition and angiotensin receptor blockade in hypertension. *Hypertension*. 2005;45(5):880-6.
2. MacKinnon M, Shurraw S, Akbari A, Knoll GA, Jaffey J, Clark HD. Combination therapy with an angiotensin receptor blocker and an ACE inhibitor in proteinuric renal disease: a systematic review of the efficacy and safety data. *Am J Kidney Dis*. 2006;48(1):8-20.
3. Phillips CO, Kashani A, Ko DK, Francis G, Krumholz HM. Adverse effects of combination angiotensin II receptor blockers plus angiotensin-converting enzyme inhibitors for left ventricular dysfunction: a quantitative review of data from randomized clinical trials. *Arch Intern Med*. 2007;167(18):1930-6.
4. Jennings DL, Kalus JS, Coleman CI, Manierski C, Yee J. Combination therapy with an ACE inhibitor and an angiotensin receptor blocker for diabetic nephropathy: a meta-analysis. *Diabet Med*. 2007;24(5):486-93.
5. Lakhdar R, Al-Mallah MH, Lanfear DE. Safety and tolerability of angiotensin-converting enzyme inhibitor versus the combination of angiotensin-converting enzyme inhibitor and angiotensin receptor blocker in patients with left ventricular dysfunction: a systematic review and meta-analysis of randomized controlled trials. *J Card Fail*. 2008;14(3):181-8.
6. Kunz R, Friedrich C, Wolbers M, Mann JF. Meta-analysis: effect of monotherapy and combination therapy with inhibitors of the renin-angiotensin system on proteinuria in renal disease. *Ann Intern Med*. 2008;148(1):30-48.
7. Catapano F, Chiodini P, De Nicola L, Minutolo R, Zamboli P, Gallo C, Conte G. Antiproteinuric response to dual blockade of the renin-angiotensin system in primary glomerulonephritis: meta analysis and metaregression. *Am J Kidney Dis*. 2008;52(3):475-85.
8. Kuenzli A, Bucher HC, Anand I, Arutiunov G, Kum LC, McKelvie R, Afzal R, White M, Nordmann AJ. Meta-analysis of combined therapy with angiotensin receptor antagonists versus ACE inhibitors alone in patients with heart failure. *PLoS One*. 2010;5(4):e9946.
9. Maione A, Navaneethan SD, Graziano G, Mitchell R, Johnson D, Mann JF, Gao P, Craig JC, Tognoni G, Perkovic V, Nicolucci A, De Cosmo S, Sasso A, Lamacchia O, Cignarelli M, Manfreda VM, Gentile G, Strippoli GF. Angiotensin-converting enzyme inhibitors, angiotensin receptor blockers and combined therapy in patients with micro- and macroalbuminuria and other cardiovascular risk factors: a systematic review of randomized controlled trials. *Nephrol Dial Transplant*. 2011;26(9):2827-47.
10. Pham JT, Schmitt BP, Leehey DJ. Effects of dual blockade of the renin-angiotensin system in diabetic kidney disease: a systematic review and meta-analysis. *J Nephrol Therapeutic* 2012;S2:003.
11. Cheng J, Zhang X, Tian J, Li Q, Chen J. Combination therapy an ACE inhibitor and an angiotensin receptor blocker for IgA nephropathy: a meta-analysis. *Int J Clin Pract*. 2012;66(10):917-23.
12. Harel Z, Gilbert C, Wald R, Bell C, Perl J, Juurlink D, Beyene J, Shah PS. The effect of combination treatment with aliskiren and blockers of the renin-angiotensin system on hyperkalaemia and acute kidney injury: systematic review and meta-analysis. *BMJ*. 2012;344:e42.
13. Makani H, Bangalore S, Desouza KA, Shah A, Messerli FH. Efficacy and safety of dual blockade of the renin-angiotensin system: meta-analysis of randomised trials. *BMJ*. 2013 ;346:f360.
14. Catalá-López F, Macías Saint-Gerons D. Diabetes Mellitus and Risks of Dual Blockade of the Renin-angiotensin-aldosterone System. *Rev Esp Cardiol*. 2013;66(5):412-415.
15. Susantitaphong P, Sewaralthahab K, Balk EM, Eiam-ong S, Madias NE, Jaber BL. Efficacy and safety of combined vs. single renin-angiotensin-aldosterone system blockade in chronic kidney disease: a meta-analysis. *Am J Hypertens*. 2013;26(3):424-41.

Anexo 4

Publicaciones incluidas en el metaanálisis

Doble bloqueo con aliskireno

1. Uresin Y, Taylor AA, Kilo C, Tschope D, Santonastaso M, Ibram G, et al. Efficacy and safety of the direct renin inhibitor aliskiren and ramipril alone or in combination in patients with diabetes and hypertension. *J Renin Angiotensin Aldosterone Syst* 2007;8:190-8.
2. Oparil S, Yarows SA, Patel S, Fang H, Zhang J, Satlin A. Efficacy and safety of combined use of aliskiren and valsartan in patients with hypertension: a randomised, double-blind trial. *Lancet* 2007;370:221-9.
3. Pool JL, Schmieder RE, Azizi M, Aldigier JC, Januszewicz A, Zidek W, et al. Aliskiren, an orally effective renin inhibitor, provides antihypertensive efficacy alone and in combination with valsartan. *Am J Hypertens* 2007;20:11-20.
4. McMurray JJ, Pitt B, Latini R, Maggioni AP, Solomon SD, Keefe DL, et al. Effects of the oral direct renin inhibitor aliskiren in patients with symptomatic heart failure. *Circ Heart Fail* 2008;1:17-24.
5. Parving HH, Persson F, Lewis JB, Lewis EJ, Hollenberg NK. Aliskiren combined with losartan in type 2 diabetes and nephropathy. *N Engl J Med* 2008;358:2433-46.
6. Solomon SD, Appelbaum E, Manning WJ, Verma A, Berglund T, Lukashevich V, et al. Effect of the direct renin inhibitor aliskiren, the angiotensin receptor blocker losartan, or both on left ventricular mass in patients with hypertension and left ventricular hypertrophy. *Circulation* 2009;119:530-7.
7. Scirica BM, Morrow DA, Bode C, Ruzyllo W, Ruda M, Oude Ophuis AJ, et al. Patients with acute coronary syndromes and elevated levels of natriuretic peptides: the results of the AVANT GARDE-TIMI 43 Trial. *Eur Heart J* 2010;31:1993-05.
8. Solomon SD, Shin SH, Shah A, Skali H, Desai A, Kober L, et al. Effect of the direct renin inhibitor aliskiren on left ventricular remodelling following myocardial infarction with systolic dysfunction. *Eur Heart J* 2011;32:1227-34.
9. Drummond W, Sirenko YM, Ramos E, Baek I, Keefe DL. Aliskiren as add-on therapy in the treatment of hypertensive diabetic patients inadequately controlled with valsartan/HCT combination: a placebo-controlled study. *Am J Cardiovasc Drugs* 2011;11:327-33.
10. Flack JM, Yadao AM, Purkayastha D, Samuel R, White WB. Comparison of the effects of aliskiren/valsartan in combination versus valsartan alone in patients with stage 2 hypertension. *J Am Soc Hypertens*. 2012;6(2):142-51.
11. Parving HH, Persson F, Lewis JB, Lewis EJ, Hollenberg NK. Aliskiren combined with losartan in type 2 diabetes and nephropathy. *N Engl J Med* 2008;358:2433-46.
12. Bakris GL, Oparil S, Purkayastha D, Yadao AM, Alessi T, Sowers JR. Randomized study of antihypertensive efficacy and safety of combination aliskiren/valsartan vs valsartan monotherapy in hypertensive participants with type 2 diabetes mellitus. *J Clin Hypertens (Greenwich)*. 2013;15:92-100.
13. Gilbert CJ, Gomes T, Mamdani MM, HELLINGS C, Yao Z, Garg AX, Wald R, Harel Z, Juurlink DN. No increase in adverse events during aliskiren use among Ontario patients receiving angiotensin-converting enzyme inhibitors or angiotensin-receptor blockers. *Can J Cardiol*. 2013;29:586-91.

Doble bloqueo sin aliskireno

14. McKelvie RS, Yusuf S, Pericak D, Avezum A, Burns RJ, Probstfield J, et al. Comparison of candesartan, enalapril, and their combination in congestive heart failure: randomized evaluation of strategies for left ventricular dysfunction (RESOLVD) pilot study. The RESOLVD Pilot Study Investigators. *Circulation*. 1999;100(10):1056-64.
15. Baruch L, Anand I, Cohen IS, Ziesche S, Judd D, Cohn JN. Augmented short- and long-term hemodynamic and hormonal effects of an angiotensin receptor blocker added to angiotensin converting enzyme inhibitor therapy in patients with heart failure. Vasodilator Heart Failure Trial (V-HeFT) Study Group. *Circulation* 1999;99(20):2658-64.
16. Hamroff G, Katz SD, Mancini D, Blaufarb I, Bijou R, Patel R, et al. Addition of angiotensin II receptor blockade to maximal angiotensin-converting enzyme inhibition improves exercise capacity in patients with severe congestive heart failure. *Circulation* 1999;99(8):990-2.
17. Azizi M, Linhart A, Alexander J, Goldberg A, Menten J, Sweet C, et al. Pilot study of combined blockade of the renin-angiotensin system in essential hypertensive patients. *J Hypertens* 2000;18:1139-47.
18. Mogensen CE, Neldam S, Tikkanen I, Oren S, Viskoper R, Watts RW, Cooper ME. Randomised controlled trial of dual blockade of renin-angiotensin system in patients with hypertension, microalbuminuria, and non-insulin dependent diabetes: the candesartan and lisinopril microalbuminuria (CALM) study. *BMJ* 2000;321:1440-4.

Anexo 4

Publicaciones incluidas en el metaanálisis (continuación)

19. Ruilope LM, Aldigier JC, Ponticelli C, Oddou-Stock P, Botteri F, Mann JF. Safety of the combination of valsartan and benazepril in patients with chronic renal disease. European Group for the Investigation of Valsartan in Chronic Renal Disease. *J Hypertens*. 2000;18:89-95.
20. Tonkon M, Awan N, Niazi I, Hanley P, Baruch L, Wolf RA, et al. A study of the efficacy and safety of irbesartan in combination with conventional therapy, including ACE inhibitors, in heart failure. Irbesartan Heart Failure Group. *Int J Clin Pract*. 2000;54(1):11-4, 16-8.
21. Cohn JN, Tognoni G; Valsartan Heart Failure Trial Investigators. A randomized trial of the angiotensin-receptor blocker valsartan in chronic heart failure. *N Engl J Med* 2001;345:1667-75.
22. Murdoch DR, McDonagh TA, Farmer R, Morton JJ, McMurray JJ, Dargie HJ. ADEPT: Addition of the AT1 receptor antagonist eprosartan to ACE inhibitor therapy in chronic heart failure trial: hemodynamic and neurohormonal effects. *Am Heart J* 2001;141:800-7.
23. Tylicki L, Rutkowski P, Renke M, Rutkowski B. Renoprotective effect of small doses of losartan and enalapril in patients with primary glomerulonephritis. Short-term observation. *Am J Nephrol*. 2002;22:356-62.
24. McMurray JJ, Ostergren J, Swedberg K, Granger CB, Held P, Michelson EL, et al. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function taking angiotensin-converting-enzyme inhibitors: the CHARM-Added trial. *Lancet*. 2003;362:767-71.
25. Pfeiffer MA, McMurray JJ, Velazquez EJ, Rouleau JL, Køber L, Maggioni AP, et al. Valsartan, captopril, or both in myocardial infarction complicated by heart failure, left ventricular dysfunction, or both. *N Engl J Med* 2003;349(20):1893-906. Erratum in: *N Engl J Med*. 2004;350(2):203.
26. Izzo JL Jr, Weinberg MS, Hainer JW, Kerkering J, Tou CK; AMAZE. Antihypertensive efficacy of candesartan-lisinopril in combination vs. up-titration of lisinopril: the AMAZE trials. *J Clin Hypertens (Greenwich)*. 2004;6:485-93.
27. Yasumura Y, Miyatake K, Okamoto H, Miyauchi T, Kawana M, Tsutamoto T, et al. Rationale for the use of combination angiotensin-converting enzyme inhibitor and angiotensin II receptor blocker therapy in heart failure. *Circ J*. 2004;68:361-6.
28. Andersen NH, Poulsen PL, Knudsen ST, Poulsen SH, Eiskjaer H, Hansen KW, Helleberg K, Mogensen CE. Long-term dual blockade with candesartan and lisinopril in hypertensive patients with diabetes: the CALM II study. *Diabetes Care* 2005;28:273-7.
29. Kanno Y, Takenaka T, Nakamura T, Suzuki H. Add-on angiotensin receptor blocker in patients who have proteinuric chronic kidney diseases and are treated with angiotensin-converting enzyme inhibitors. *Clin J Am Soc Nephrol* 2006;1:730-7.
30. Horita Y, Taura K, Taguchi T, Furusu A, Kohno S. Aldosterone breakthrough during therapy with angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers in proteinuric patients with immunoglobulin A nephropathy. *Nephrology (Carlton)*. 2006;11:462-6.
31. Bakris GL, Ruilope L, Locatelli F, Ptaszynska A, Pieske B, de Champlain J, et al. Treatment of microalbuminuria in hypertensive subjects with elevated cardiovascular risk: results of the IMPROVE trial. *Kidney Int* 2007;72:879-85.
32. Ogawa S, Takeuchi K, Mori T, Nako K, Tsubono Y, Ito S. Effects of monotherapy of temocapril or candesartan with dose increments or combination therapy with both drugs on the suppression of diabetic nephropathy. *Hypertens Res* 2007;30:325-34.
33. White M, Lepage S, Lavoie J, De Denus S, Leblanc MH, Gossard D, et al. Effects of combined candesartan and ACE inhibitors on BNP, markers of inflammation and oxidative stress, and glucose regulation in patients with symptomatic heart failure. *J Card Fail* 2007;13:86-94.
34. Kum LC, Yip GW, Lee PW, Lam YY, Wu EB, Chan AK, et al. Comparison of angiotensin-converting enzyme inhibitor alone and in combination with irbesartan for the treatment of heart failure. *Int J Cardiol* 2008;125:16-21.
35. ONTARGET Investigators, Yusuf S, Teo KK, Pogue J, Dyal L, Copland I, et al. Telmisartan, ramipril, or both in patients at high risk for vascular events. *N Engl J Med* 2008;358:1547-59.
36. Menne J, Farsang C, Deák L, Klebs S, Meier M, Handrock R, et al. Valsartan in combination with lisinopril versus the respective high dose monotherapies in hypertensive patients with microalbuminuria: the VALERIA trial. *J Hypertens* 2008;26:1860-7.

Anexo 4

Publicaciones incluidas en el metaanálisis (continuación)

37. Mori-Takeyama U, Minatoguchi S, Murata I, Fujiwara H, Ozaki Y, Ohno M, et al. Dual blockade of the renin-angiotensin system versus maximal recommended dose of angiotensin II receptor blockade in chronic glomerulonephritis. *Clin Exp Nephrol* 2008;12:33-40
38. Mehdi UF, Adams-Huet B, Raskin P, Vega GL, Toto RD. Addition of angiotensin receptor blockade or mineralocorticoid antagonism to maximal angiotensin-converting enzyme inhibition in diabetic nephropathy. *J Am Soc Nephrol* 2009;20:2641-50.
39. Cice G, Di Benedetto A, D'Isa S, D'Andrea A, Marcelli D, Gatti E, et al. Effects of telmisartan added to Angiotensin-converting enzyme inhibitors on mortality and morbidity in hemodialysis patients with chronic heart failure a double-blind, placebo-controlled trial. *J Am Coll Cardiol* 2010;56:1701-8.
40. Bianchi S, Bigazzi R, Campese VM. Intensive versus conventional therapy to slow the progression of idiopathic glomerular diseases. *Am J Kidney Dis* 2010;55:671-81.
41. Titan SM, M Vieira J Jr, Dominguez WV, Barros RT, Zatz R. ACEI and ARB combination therapy in patients with macroalbuminuric diabetic nephropathy and low socioeconomic level: a double-blind randomized clinical trial. *Clin Nephrol* 2011;76:273-83.
42. McAlister FA, Zhang J, Tonelli M, Klarenbach S, Manns BJ, Hemmelgarn BR; Alberta Kidney Disease Network. The safety of combining angiotensin-converting-enzyme inhibitors with angiotensin-receptor blockers in elderly patients: a population-based longitudinal analysis. *CMAJ* 2011;183:655-62.
43. Chan KE, Ikizler TA, Gamboa JL, Yu C, Hakim RM, Brown NJ. Combined angiotensin-converting enzyme inhibition and receptor blockade associate with increased risk of cardiovascular death in hemodialysis patients. *Kidney Int* 2011;80:978-85.
44. Kurnik D, Vesterman-Landes J, Bialik M, Katzir I, Lomnický Y, Halkin H, et al. Hyperkalemia and renal function during monotherapy and dual renin-angiotensin blockade in the community setting. *Clin Ther* 2011;33:456-64.
45. Yano H, Hibi K, Nozawa N, Ozaki H, Kusama I, Ebina T, et al. Effects of valsartan, an angiotensin II receptor blocker, on coronary atherosclerosis in patients with acute myocardial infarction who receive an angiotensin-converting enzyme inhibitor. *Circ J* 2012;76:1442-51.
46. Mann JF, Anderson C, Gao P, Gerstein HC, Boehm M, Rydén L et al. Dual inhibition of the renin-angiotensin system in high-risk diabetes and risk for stroke and other outcomes: results of the ONTARGET trial. *J Hypertens*. 2013;31:414-21.
47. Lesogor A, Cohn JN, Latini R, Tognoni G, Krum H, Massie B et al. Interaction between baseline and early worsening of renal function and efficacy of renin-angiotensin-aldosterone system blockade in patients with heart failure: insights from the Val-HeFT study. *Eur J Heart Fail* 2013 Jun 19. [Epub ahead of print]
48. Fernández Juárez G, Luño J, Barrio V, de Vinuesa SG, Praga M, Goicoechea M, et al; PRONEDI Study Group. Effect of dual blockade of the renin-angiotensin system on the progression of type 2 diabetic nephropathy: a randomized trial. *Am J Kidney Dis* 2013;61:211-8.
49. Imai E, Haneda M, Yamasaki T, Kobayashi F, Harada A, Ito S, et al. Effects of dual blockade of the renin-angiotensin system on renal and cardiovascular outcomes in type 2 diabetes with overt nephropathy and hypertension in the ORIENT: a post-hoc analysis (ORIENT-Hypertension). *Hypertens Res*. 2013 Sep 12.
50. Fried LF, Emanuele N, Zhang JH, Brophy M, Conner TA, Duckworth W, Leehey DJ, McCullough PA, O'Connor T, Palevsky PM, Reilly RF, Seliger SL, Warren SR, Watnick S, Peduzzi P, Guarino P; VA NEPHRON-D Investigators. Combined Angiotensin inhibition for the treatment of diabetic nephropathy. *N Engl J Med*. 2013;369:1892-903.

REVISIÓN SISTEMÁTICA Y METAANÁLISIS

EVALUACIÓN DE TÉCNICAS INMUNOLÓGICAS IN VITRO PARA EL DIAGNÓSTICO DE ALERGIAS: METAANÁLISIS 2000-2012 (*)

John Fredy Cuervo-Perez (1), Julián Camilo Arango (1) y Jaiberth Antonio Cardona-Arias (2,3).

(1) Escuela de Microbiología Universidad de Antioquia. Grupo de investigación de alergología clínica y experimental. Universidad de Antioquia. Medellín. Colombia.

(2) Facultad Medicina. Universidad Cooperativa de Colombia.

(3) Grupo de Investigación Salud y Sostenibilidad, Escuela de Microbiología. Universidad de Antioquia. Medellín. Colombia.

(*) El estudio fue financiado con recursos de la Universidad de Antioquia.

Ninguno de los autores declara conflicto de intereses para la publicación del manuscrito.

RESUMEN

Fundamentos: Las alergias presentan elevada prevalencia, afectan a todos los grupos etarios, generan impactos negativos sobre los sistemas de salud, educativo y económico. Se desconoce la utilidad diagnóstica de las pruebas de tamización. El objetivo del estudio fue evaluar la validez, el desempeño, la seguridad y la efectividad diagnóstica de las técnicas inmunológicas in vitro para alergias.

Métodos: Revisión sistemática con metaanálisis. Se aplicó una estrategia de búsqueda de estudios en Pubmed, Sciencedirect y Wiley con los términos de búsqueda *activation basophil test, lymphocyte transformation test, specific IgE immunoassay*. Período estudiado: 2000-2012. Se determinó la reproducibilidad de la selección, extracción y evaluación de la calidad de los artículos. Se calculó sensibilidad, especificidad, cocientes de probabilidad, valores predictivos, proporción de resultados falsos, exactitud, razón de odds, índice J de Youden y curva ROC con los software Meta-DiSc(es) y Epidat 3.0.

Resultados: Se incluyeron 18 estudios con 3.520 individuos, 58% enfermos y 42% sanos. La activación de basófilos presentó sensibilidad del 78% (IC95%:74-81), especificidad 95% (IC95%:83-100), cociente de probabilidad positivo 9,9 (IC95%:6,8-14,4) y negativo de 0,20 (IC95%:0,13-0,30), OR diagnóstica de 70,8 (IC95%:40,2-124,8) y un área bajo la curva de 0,97. En la inmunoglobulina E específica la sensibilidad fue 72% (IC95%:69-75), especificidad 90% (IC95%:88-92), cociente de probabilidad positivo 12,9 (IC95%:4,0-41,6) negativo 0,32 (IC95%:0,23-0,43), OR diagnóstica 41,6 (IC95%:11,6-148,9) y área bajo la curva 0,87.

Conclusión: Se evidenció que la activación de basófilos y la IgE específica son pruebas útiles en el diagnóstico de alergias.

Palabras claves: Diagnóstico. Alergia. Test inmunológicos.

ABSTRACT

Evaluation of Techniques in Vitro Immune to the Diagnosis of Allergy: Meta-analysis 2000-2012

Background: Allergies have high prevalence, affecting all age groups, generate negative impacts on health, educational and economic systems, and they are unknown the diagnostic utility of screening tests. The objective of the study was to evaluate the validity, performance, safety and diagnostic efficiency of in vitro immunological techniques for allergies, 2000-2012.

Methods: Systematic review with meta-analysis. We applied a search strategy studies in PubMed, Sciencedirect and Wiley, with search terms activation basophil test, lymphocyte transformation test, specific IgE immunoassay. We determined the reproducibility of the selection, extraction and quality assessment of articles. We calculated sensitivity, specificity, likelihood ratios, predictive values, proportion of false, accuracy, odds ratio, Youden index J and ROC curve in Meta-DiSc(es) and Epidat 3.0 software.

Results: We included 18 studies with 3520 individuals, 58% patients and 42% healthy. Activation of basophils showed sensitivity of 78% (95% CI :74-81), specificity 95% (95% CI: 83-100), positive likelihood ratio 9.9 (95% CI: 6.8 to 14.4) and negative of 0.20 (95% CI = 0.13 to 0.30) a diagnostic OR 70.8 (IC95: 40.2 to 124.8) and area under the curve of 0.97. In specific immunoglobulin E sensitivity was 72% (95% CI: 69-75), specificity 90% (95% CI : 88-92), positive likelihood ratio 12.9 (95% CI = 4.0 to 41.6) negative 0.32 (95% CI:0.23-0.43), diagnostic OR 41.6 (95% CI :11.6 to 148.9) and area under the curve 0.87.

Conclusion: We showed that activation of basophils and specific IgE are useful tests for diagnosing allergies.

Key words: Diagnosis. Immunologic tests. Allergy and immunology.

Correspondencia

Jaiberth Antonio Cardona Arias

Calle 67 Número 53 – 108, Bloque 5, oficina 103
Medellín, Colombia

Correo electrónico: jaiberthcardona@gmail.com

INTRODUCCIÓN

El sistema inmune tiene la capacidad de diferenciar moléculas extrañas de un patógeno de otras que siendo igualmente extrañas no generan una amenaza para la salud, estas últimas generalmente se derivan de plantas, animales y del ambiente y constituyen potenciales alérgenos. En la mayoría de las personas el contacto con sustancias ajenas no genera una respuesta inmune. Sin embargo, en algunas circunstancias, moléculas inocuas pueden desencadenar una respuesta inmunitaria en determinados individuos, originando lo que se conoce como enfermedad alérgica¹⁻².

Las alergias se producen por la entrada de antígenos por la piel o las mucosas del trato gastrointestinal y respiratorio, los cuales son captados por las células presentadoras de antígeno (Ag) que estimulan a los linfocitos Th₂ a secretar citoquinas para activar a los linfocitos B y producir Inmunoglobulina E (IgE) específica contra el Ag. La IgE se fija a receptores ubicados en la membrana celular de los mastocitos, basófilos y eosinófilos, generando la sensibilización al alérgeno. Cuando el individuo se expone al Ag, éste se une a la IgE de las células sensibilizadas lo que conduce a una desgranulación celular y a la liberación de mediadores vasoactivos e inflamatorios como histaminas, factores quimiotácticos, leucotrienos y factores activadores de plaquetas que causan los signos de las alergias, como vasodilatación, aumento de la permeabilidad capilar, hipersecreción glandular, espasmo de músculo liso e infiltración tisular de eosinófilos y otras células inflamatorias responsables de la sintomatología^{1,2}. Estos mecanismos explican las manifestaciones clínicas de enfermedades como rinitis, conjuntivitis, asma, urticaria, dermatitis atópica, alergias físicas y alergias a alimentos, fármacos y picaduras de insectos³.

Aunado a lo anterior, las alergias presentan un origen genético, algunos autores afir-

man que si ambos progenitores son alérgicos existe una probabilidad del 50% de tener un hijo alérgico y cuando solo lo es uno de ellos la probabilidad disminuye al 30%⁴⁻⁵.

La importancia de las enfermedades alérgicas radica en que son procesos muy frecuentes, que afectan en especial a niños y personas jóvenes en las fases de sus vidas en las que la producción laboral o académica es más intensa e interfieren de forma significativa en las actividades cotidianas. Además, en algunos casos más del 60% de los pacientes admite tener algún tipo de limitación en actividades físicas y laborales. Con elevada frecuencia presentan alteraciones en el sueño cuando el paciente padece enfermedades como la rinitis o asma, lo que conduce a que el 40% de estos individuos sufran somnolencia diurna, afectando a su rendimiento académico y laboral. La Organización Mundial de la Salud (OMS) estima que el 1% de los recursos sanitarios en el ámbito mundial se invierten en el asma y que en algunos países el tratamiento de dicha enfermedad asciende a 150 millones de euros al año y sus gastos indirectos, como incapacidades médico-laborales, superan los 300 millones de euros anuales. De forma adicional, estudios de la OMS dan cuenta de que muchos individuos que padecen enfermedades alérgicas tienen sus proyectos de vida condicionados al padecimiento de su enfermedad, como encontrar limitaciones al elegir su carrera profesional⁶.

Sumado a las consecuencias citadas, se debe explicitar que las enfermedades alérgicas mediadas por IgE constituyen una de las enfermedades más prevalentes en todo el mundo. Se estima que el 20% de la población mundial sufre alguna enfermedad mediada por la IgE. La OMS ha indicado que las enfermedades alérgicas se encuentran entre las seis enfermedades más frecuentes en el ámbito mundial. Particularmente en el asma, se ha documentado que más del 50% de los casos en adultos y más del 80% en niños son de origen alérgico,

estando afectada entre el 5% y el 15% de la población infantil mundial⁷.

Las técnicas *in vitro* para el diagnóstico de alergias han permitido un gran avance en su correcto diagnóstico al proporcionar mediciones más exactas de la actividad de la IgE y las células efectoras en los procesos alérgicos, lo cual ha mejorado la calidad analítica y ha servido para monitorizar al paciente y permitir un mejor tratamiento. Estas pruebas incluyen la determinación de IgE específica, la activación de basófilos y la linfoproliferación⁸.

A pesar de la disponibilidad y ventajas de las pruebas *in vitro*, en la actualidad no son muy utilizadas debido a diversas causas como los costos, falta de estandarización y, en especial, por la gran divergencia en validez y desempeño diagnóstico que se ha reportado en estudios previos, con valores de sensibilidad y especificidad entre 50% y 90%. Adicional a esto, la mayoría de los estudios que han evaluado estas técnicas sólo aluden a los valores de validez diagnóstica (sensibilidad y especificidad) obviando otros parámetros de gran relevancia en la evaluación de pruebas, como los valores predictivos, los cuales se relacionan con la prevalencia de la enfermedad, la razón de verosimilitud que determina la interacción entre la sensibilidad y la especificidad, la eficiencia de la prueba que define la proporción de pacientes correctamente diagnosticados y la proporción de resultados falsos, tanto positivos como negativos⁹⁻¹⁰.

Los aspectos anteriores justifican la realización de una revisión sistemática de artículos que abordan este tópico. Además, se debe precisar que la elección de esta modalidad de investigación radica en que presenta múltiples ventajas, como el poseer una metodología científica claramente explícita para la identificación, selección y elaboración de estudios sintetizando sus resultados, eliminar los sesgos de las revisiones narrativas, brindar información amplia, exacta y resu-

mida que puede ser asimilada con rapidez y facilidad por los profesionales de la salud, permitir plantear nuevas hipótesis, detectar áreas en las que la evidencia científica es escasa, comparar los resultados de diferentes estudios para establecer la generalización de sus hallazgos y su consistencia así como reducir el tiempo entre los descubrimientos y la implementación de las estrategias terapéuticas y diagnósticas para las distintas enfermedades¹¹⁻¹³.

El objetivo de esta investigación fue evaluar la validez, el desempeño, la seguridad y la efectividad diagnóstica de las técnicas inmunológicas *in vitro* para el estudio de alergias entre los años 2000 y 2012.

MATERIAL Y MÉTODOS

Tipo de estudio: Revisión sistemática de la literatura con metaanálisis.

Identificación, tamización, elección e inclusión de los estudios: Con el fin de garantizar la exhaustividad de la revisión, se realizó una búsqueda por sensibilidad de los artículos originales de investigaciones publicadas en las bases de datos ScienceDirect, Wiley y Medline (Pubmed). La búsqueda por sensibilidad, es decir, sin circunscribirla a Descriptores de Ciencias de la Salud (DeCS) o a las cabeceras del Medical Subject Headings (MeSH), permitió la obtención de un mayor número de estudios frente a la búsqueda por especificidad. Se utilizaron los siguientes términos: *activation basophil test*, *lymphocyte transformation test*, *specific IgE immunoassay* y sus equivalentes en español y portugués. Previamente a la delimitación de estas tres pruebas, se realizó una búsqueda con términos más amplios, como “diagnóstico *in vitro* de alergias”, “pruebas *in vitro* para detección de alergias” e “identificación *in vitro* de alergias”, con el fin de identificar las pruebas empleadas en el ámbito internacional, concluyendo que los tres test incluidos son los más frecuentes en investigación clínica.

Se aplicaron los siguientes criterios de inclusión: i) artículos con los términos de búsqueda en título y resumen, ii) publicados entre el 2000 y el 2012, iii) observacionales o longitudinales, iv) donde se reportaran sensibilidad o especificidad de las pruebas de diagnóstico in vitro para alergias y v) que especificaran el protocolo de ejecución de las técnicas diagnósticas. Se excluyeron los artículos que cumplieren los siguientes criterios: i) donde la población o muestra evaluada fuesen personas sin diagnóstico previo de algún tipo de alergia, ii) estudios que no contribuyeran al logro del objetivo general de esta revisión, como artículos en los que se abordaran otras anomalías de la IgE y iii) publicaciones con problemas de validez interna por los estadísticos empleados para el análisis y mal control de sesgos de selección e información. Para determinar la calidad de los estudios elegibles e incluidos se emplearon los criterios de la guía *Quality Assessment of Diagnostic Accuracy Studies* (QUADAS)^{14,15}.

Recolección de la información. Se desarrolló por tres investigadores de forma independiente para garantizar la reproducibilidad de la búsqueda y selección de los artículos. Las discrepancias se resolvieron por consenso.

En la extracción de la información de los estudios incluidos se aplicó un manual que contenía las variables a extraer de cada uno, lo que fue diligenciado de forma independiente por cada investigador en dos ocasiones diferentes con una diferencia de 2 semanas para garantizar la reproducibilidad intra e interobservador de la extracción de la información. Para las variables lugar de estudio, técnica empleada, número de sujetos estudiados, número de positivos y de negativos o controles y tipo de alergia se obtuvo un índice kappa mayor a 0,95 y para la variable edad los coeficientes de correlación intraclass entre investigadores presentaron un valor de 0,98.

Análisis de la información. Para la evaluación de la validez de las pruebas se determinó la sensibilidad, la especificidad y las razones

de verosimilitud o coeficientes de probabilidad (CP). El desempeño se estimó a partir de los valores predictivos y la proporción de resultados falsos, la exactitud con el porcentaje de pacientes correctamente diagnosticados y la razón de odds diagnóstica (ORD). La seguridad de la prueba se estudió con el Índice J de Youden. Para todos los resultados se calcularon sus respectivos intervalos de confianza del 95%¹⁶.

Se realizaron curvas ROC (Receiver Operating Characteristic) con base en los valores de sensibilidad y especificidad y se calculó el área bajo la curva de las pruebas, estableciendo en su interpretación las siguientes categorías: i) mala para valores menores a 0,6; ii) regular entre 0,60 y 0,74; iii) bueno entre 0,75 y 0,89; iv) muy bueno entre 0,90 y 0,95 y v) excelente a partir de 0,96.

En la interpretación del Índice J de Youden se tomó como base el hecho de que cuanto más se acerque a 1,0 mayor sería la calidad del resultado obtenido al realizar la prueba en un paciente, mientras que para la ORD se precisó que valores 1 o cercanos a 1 indicarían que la prueba no tenía capacidad para discriminar sujetos sanos y enfermos. Los valores más altos se tomaron como indicadores de mayor utilidad diagnóstica. Los cocientes de probabilidad se interpretaron con base en las siguientes categorías: i) excelente ayuda en el diagnóstico: CP- <0,1 y CP+ >10, ii) buena ayuda diagnóstica o de importancia clínica: CP- entre 0,1-0,2 y CP+ 5-10, iii) poca ayuda al clínico: CP- entre 0,21 y 0,5 CP+ entre 2 y 5, y iv) la prueba no presenta capacidad discriminante para CP- entre 0,51 y 1 o CP+ entre 1 y 2¹⁶.

Los datos relacionados con los falsos positivos y negativos fueron almacenados y analizados en una base de datos en el software *Meta-analysis of studies of evaluations of Diagnostic and Screening tests Meta-DiSc* (es). En éste se empleó el método de aglomeración con el estadístico Prueba Q (χ^2) de DerSimonian-Laird (REM), se agregó 0,5 a las celdas con cero, se calcularon los intervalos de confianza

del 95% con corrección por sobredispersión, en las curvas ROC el área bajo la curva se calculó con la constante del modelo de Moses y el método de estimación del modelo fue con mínimos cuadrados no ponderados. Se determinó la heterogeneidad u homogeneidad de los metaanálisis de cada parámetro de la evaluación diagnóstica a través del estadístico Q de Cochran, que sigue una distribución χ^2 y para el análisis de sensibilidad se estableció el porcentaje de peso de cada estudio individual sobre el resultado global del CP+, CP- y ORD. Además, en el software se realizaron los *forest plot* para la sensibilidad, especificidad, CP+, CP- y ORD de las pruebas identificadas, como resultado total del metaanálisis para evidenciar efectos individuales y efecto global con sus respectivos intervalos de confianza¹⁷. Se debe precisar que el Meta-DiSc(es) no presenta los resultados relacionados con los valores predictivos, la proporción de resultados falsos, el porcentaje de pacientes correctamente diagnosticados ni el índice J de Youden, por lo que estos parámetros, con sus respectivos intervalos de confianza, fueron calculados con la versión 3.0 del programa para análisis epidemiológico de datos tabulados de la Organización Panamericana de la Salud (EPIDAT), tanto para los estudios individuales como para el resultado global de las pruebas analizadas.

En todos los análisis se tomó un nivel de significación estadística de 0,05.

RESULTADOS

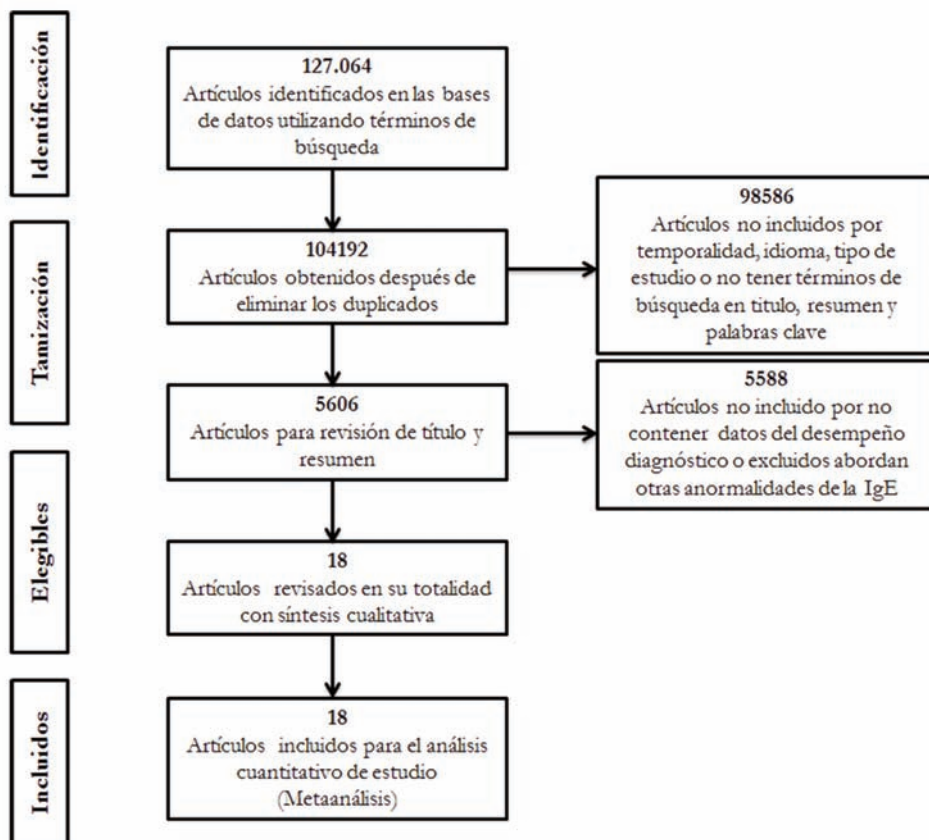
Se identificaron 127.064 artículos con todas las fuentes y estrategias de búsqueda empleadas, a partir de la eliminación de duplicados se tamizaron 104.192 manuscritos. Tras la aplicación de los criterios de inclusión y exclusión y el análisis de las investigaciones elegibles, se incluyeron 18 estudios (figura 1). Después de excluir los que no presentaban datos de validez o desempeño diagnóstico o que abordaban otras anomalías de la IgE, se aplicó la guía

QUADAS a 18 estudios. Todos ellos obtuvieron una buena o aceptable calidad metodológica (los artículos evaluados cumplían entre 10 y 12 ítem de los 14 puntos de la guía) y fueron incluidos en el análisis. No se realizó meta-análisis por subgrupos de estudios en función de su calidad metodológica.

Las 18 investigaciones incluyeron un total de 3.520 individuos con un 57,6% de personas alérgicas y 42,4% de personas sanas. 14 las investigaciones estaban realizadas en Europa, con diferentes tipos de alergia, en personas con un rango de edad entre 1 y 85 años, a las que se les evaluó los marcadores de membrana CD63 de los basófilos activados. Se debe precisar que el estudio de Eberlin²² incluyó dos grupos, uno con marcadores CD203c y otro CD63 mientras que el estudio de Ocmant³⁰ aplicó al mismo grupo cuatro pruebas diferentes que comprendían los marcadores CD63 y CD203 para personas alérgicas a maní y huevo. Un estudio evaluó la linfoproliferación, lo que representó el 0,5% de la población incluida, cuatro la IgE específica con el 68,0% de individuos y los demás artículos empleaban la prueba de activación de basófilos, que dio resultado positivo en el 31,5% de la población de estudio (tabla 1). Los estudios con la prueba de activación de basófilos incluyeron a 1.108 individuos (700 personas alérgicas y 408 sanas), en la linfoproliferación fueron 7 alérgicos y 11 sanos, mientras que en la IgE específica fueron 1.322 alérgicos y 1.072 sanos.

El tipo de estudio en todas las investigaciones incluidas correspondieron a evaluaciones diagnósticas. El criterio de clasificación en sanos y enfermos en los estudios que emplearon pruebas cutáneas fue la presencia de pápulas cutáneas con diámetro mayor a 3mm después de 15 minutos como marcador de un proceso alérgico. En las que emplearon IgE específica fue la presencia de concentraciones mayores a 0,35UI/L y en los que emplearon como estándar de referencia el diagnóstico clínico se evaluó la

Figura 1
Algoritmo de selección de los artículos



presencia de exantema maculopapular, urticaria y angiodema.

En la tabla 2 se presenta la descripción de los valores para la evaluación diagnóstica, sumado al porcentaje del peso que presenta cada estudio individual en el valor global del CP+, el CP- y la ORD. En esta se observó que el estudio de Sturm et al²⁰ con alérgeno de avispas presentó el mayor peso en los resultados del CP+,

mientras que para el CP- fue el de Sanz M et al¹⁹ (tabla 2).

El estudio de Tsuge I et al.³¹ fue el único que empleó el estudio de linfoproliferación, y halló una sensibilidad de 100% (IC95:93-100), especificidad de 91% (IC95:69-100), CP+ 11 (IC95%:2-71) e Índice J de Younden 0,91 (IC95%:0,74-1,08).

Tabla 1
Descripción de los estudios

Autor	Año	País	Edad	Grupo	Prueba de referencia	Alergeno
Activación de basófilos						
Apostolou ¹⁸	2006	N.Zelanda	57-85	Pacientes con anafilaxia al Gelofusin	Cutáneas	Medicamentos
Sanz ¹⁹	2006	España	5-17	Niños sin intervenciones quirúrgicas	Cutáneas	Látex
Sturm ²⁰	2004	Austria	-	Pacientes con historia de alergia al veneno de avispas	Cutáneas e IgE específica	Abejas
Sturm ²⁰	2004	Austria	-	Pacientes con historia de alergia al veneno de abejas	Cutáneas e IgE específica	Avispas
Rubio ²¹	2010	Francia	1-12	Niños con diagnóstico de alergia a la leche de vaca	Cutáneas e IgE específica	Leche de vaca
Eberlein ²²	2006	Alemania	-	Pacientes con historia de alergia al veneno de insectos	Cutáneas	Veneno insecto CD203c
Eberlein ²²	2006	Alemania	-	Pacientes con historia clínica al veneno de insectos	Cutáneas	Veneno insecto(CD63)
Sanz ²³	2002	España	51.5	Pacientes con tratamientos previos con betalactámicos	Cutáneas e IgE específica	Betalactámicos
Frezzolini ²⁴	2009	Italia	24-54	Pacientes con urticaria crónica	Cutáneas e IgE específica	Parásitos de peces
Pâris ²⁵	2000	Francia	21-64	Pacientes con historia de alergias al polen de cipre	Cutáneas	Polen
Monneret ²⁶	2002	Francia	30-61	Pacientes con reacción alérgica en anestesia	Cutáneas e IgE específica	Medicamentos
Eberlein ²⁷	2012	Alemania	-	Pacientes con historia de alergia a veneno de abejas y avispas	Cutáneas e IgE específica	Veneno avispas, abejas
Sanz ²⁸	2003	España	51.5	Pacientes con tratamientos previos con betalactámicos	Cutáneas e IgE específica	Medicamentos
Wolanczyk ²⁹	2009	Polonia	21-40	Pacientes tratados previamente con cetirizina	Cutáneas	Alimentos
Ocmant ³⁰	2009	Bélgica	1-12	Pacientes con antecedentes de alergia al huevo	Cutáneas e IgE específica	Maní CD63
Ocmant ³⁰	2009	Bélgica	1-12	Pacientes con antecedentes de alergia al huevo	Cutáneas e IgE específica	Maní CD203
Ocmant ³⁰	2009	Bélgica	1-11	Pacientes con antecedentes de alergia al maní	Cutáneas e IgE específica	Huevo CD63
Ocmant ³⁰	2009	Bélgica	1-11	Pacientes con antecedentes de alergia al maní	Cutáneas e IgE específica	Huevo CD203
Linfoproliferación						
Tsuge ³¹	2007	Japón	2-13	Pacientes alérgicos a fenitoína	Clínica	Medicamentos
Inmunoglobulina E específica						
Dang ³²	2012	Australia	--	Paciente son historial clínico a alergia al maní	Clínica	Maní
Corey ³³	2000	USA	17-72	Pacientes alérgicos a diferentes tipos de alérgenos	Cutáneas	Caspa de gato y ambrosia
Petersen ³⁴	2004	Dinamarca	15-41	Pacientes con historial alérgico a ácaros	Cutáneas	Ácaros del polvo
Biagini ³⁵	2001	USA	-	Pacientes alérgicos, reclutados por la FDA	Historia clínica	Medicamentos y alimentos

Tabla 2
Descripción del aporte de los estudios individuales al resultado global

Activación de basófilos					
Estudio	VP/(VP+FN)	VN/(VN+FP)	% de peso		
			CP+	CP-	OR
Apostolou ¹⁸	6/6	7/8	6,47	1,78	2,66
Sanz ¹⁹	22/23	10/10	1,92	3,64	2,80
Sturm ²⁰	21/23	27/30	12,06	4,24	7,60
Sturm ²⁰	29/34	25/30	21,36	5,85	12,73
Rubio ²¹	102/112	14/15	3,93	6,57	6,13
Eberlein ²²	42/43	22/25	12,48	2,81	5,27
Eberlein ²²	38/43	25/25	1,87	5,99	3,44
Sanz ²³	29/58	28/30	7,58	7,43	10,62
Frezzolini ²⁴	38/57	20/20	1,87	7,24	3,62
Pâris ²⁵	31/34	33/33	1,86	5,28	3,30
Monneret ²⁶	20/39	21/21	1,85	7,34	3,59
Eberlein ²⁷	12/22	9/10	3,91	6,89	5,66
Sanz ²⁸	28/58	29/30	3,72	7,46	6,50
Wolanczyk ²⁹	20/22	19/19	1,88	4,71	3,12
Ocmant ³⁰	28/32	27/29	7,81	5,53	8,32
Ocmant ³⁰	29/32	28/29	3,79	5,00	5,27
Ocmant ³⁰	28/31	22/22	1,87	5,30	3,28
Ocmant ³⁰	20/31	21/22	3,77	6,94	6,10
Inmunoglobulina E específica					
Dang ³²	60/100	98/100	25,63	31,29	26,18
Corey ³³	321/458	693/792	39,31	35,94	38,96
Biagini ³⁵	243/311	172/180	35,06	32,77	34,86

VP: Verdaderos Positivos. FP: Falsos Positivos. FN: Falsos Negativos. VN: Verdaderos Negativos. CP+: Coeficiente de Probabilidad Positiva. CP-: Coeficiente de Probabilidad Negativa. OR: Razón de Odds Diagnóstica.

En la evaluación de la validez diagnóstica de la prueba de activación de basófilos se halló una sensibilidad global de 78% (IC95%:74%-81%) con un rango de 48%-100%. Una especificidad de 95% (IC95%:83-100), CP+ de 9,9 (IC95%:6,8-14,4), CP- 0,20 (IC95%:0,13-0,30) y una ORD de 70,8 (IC95:40,2-124,8). En comparación con el resultado global, en la sensibilidad se hallaron tres estudios con resultados estadísticamente menores y dos mayo-

res, lo que puso de manifiesto la heterogeneidad de los resultados de los estudios individuales ($p=0,000$), al igual que en el CP- donde tres estudios presentaron resultados mayores al global y el análisis de heterogeneidad fue estadísticamente significativo (Q de Cochran=139,26; $p=0,000$). En los demás parámetros los hallazgos de los estudios individuales y el global fueron estadísticamente iguales con valores de $p>0,05$ en el análisis de heterogeneidad (figura 2).

Figura 2
Evaluación diagnóstica de la prueba de activación de basófilos

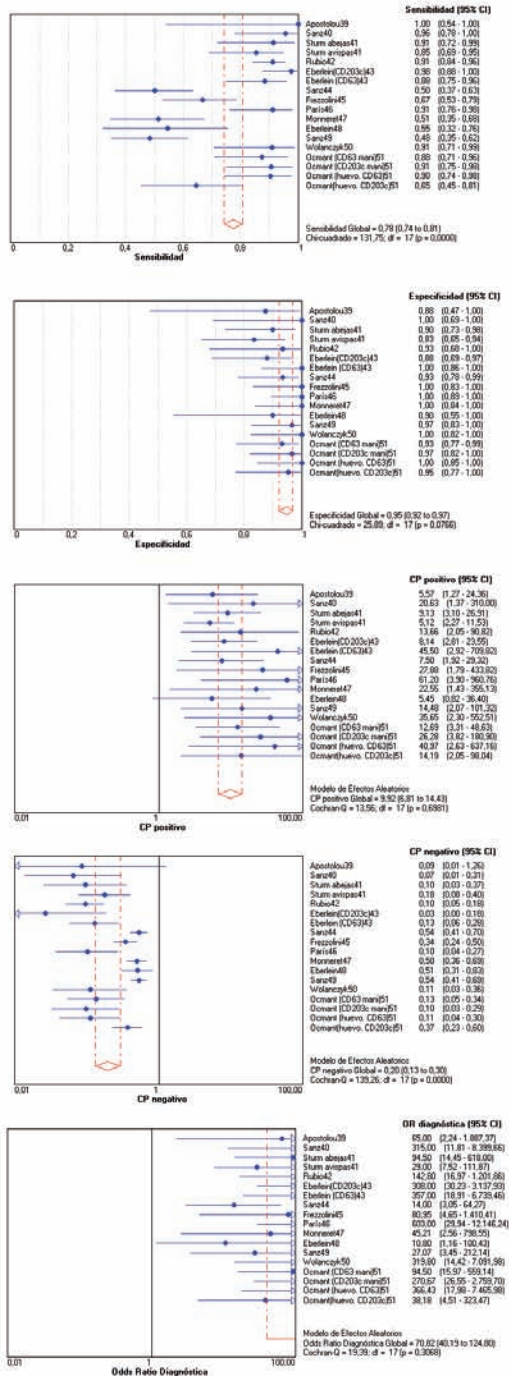
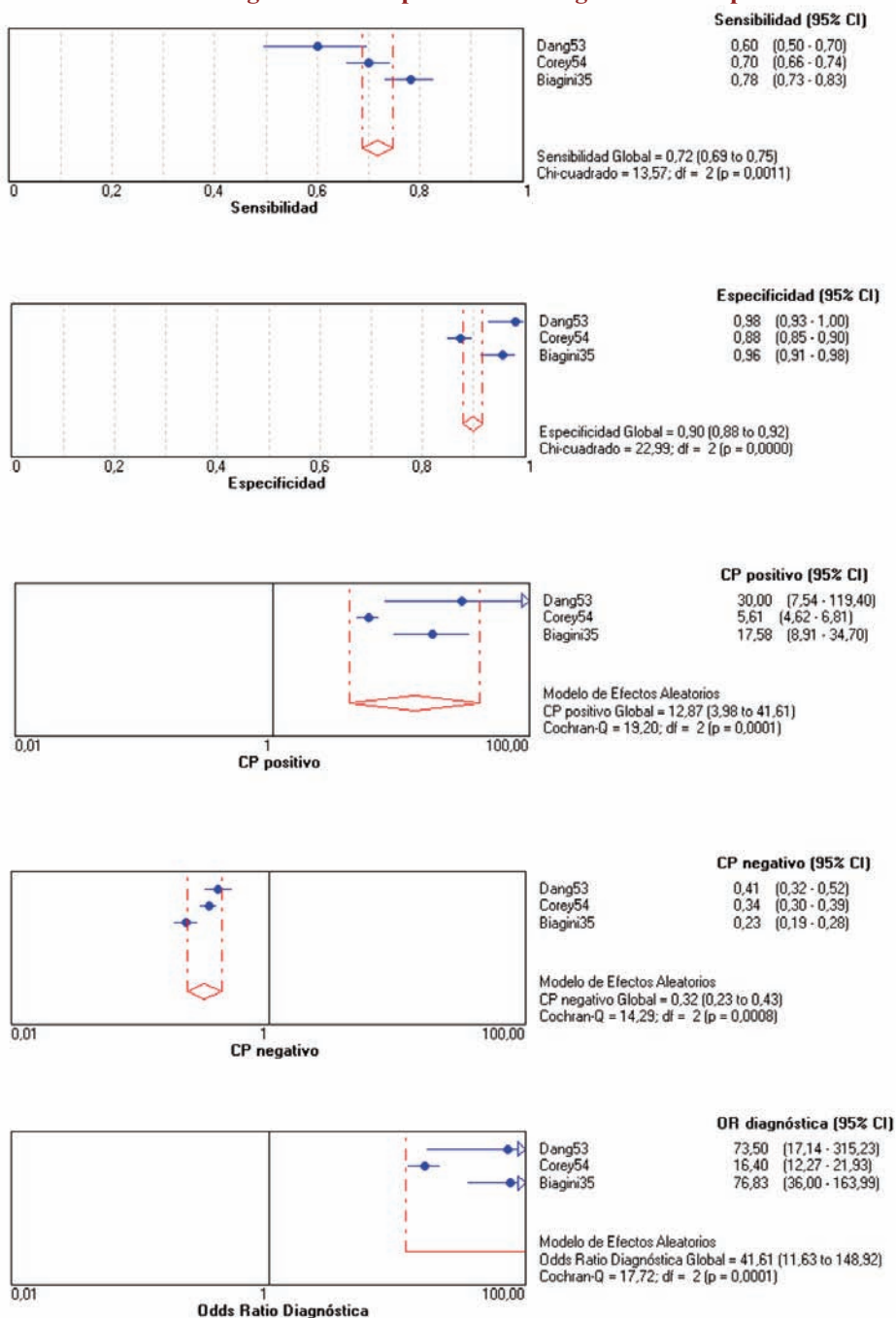


Figura 3
Evaluación diagnóstica de la prueba Inmunoglobulina E específica



En la determinación de inmunoglobulina E específica, la sensibilidad global fue 72% (IC95%:69-75), la especificidad 90% (IC95%:88-92), el CP+ 12,9 (IC95%:4,0-41,6), el CP- 0,32 (IC95%:0,23-0,43), la ORD 41,6 (IC95%:11,6-148,9) y se observó heterogeneidad en todos los parámetros evaluados (figura 3).

En la tabla 3 se presentan los valores predictivos, la proporción de resultados

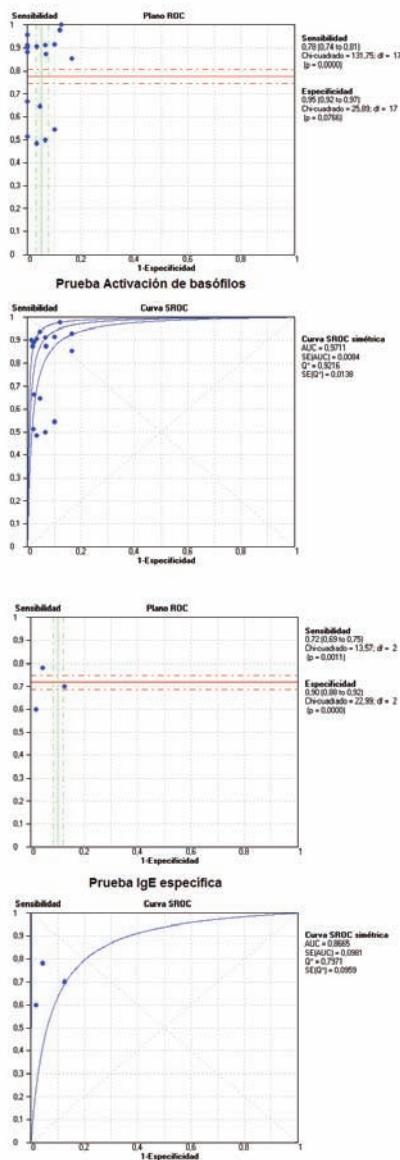
falsos, el índice J de Youden y la eficiencia de la prueba de activación de basófilos, la linfoproliferación y la IgE específica. Se observó que la activación de basófilos fue estadísticamente mayor en el valor predictivo positivo, la IgE específica fue mayor en la proporción de falsos positivos, mientras que los demás parámetros fueron estadísticamente iguales en las pruebas analizadas.

Tabla 3
Análisis de sensibilidad

Autor	VPP	VPN	PFP	PFN	PPCD	I JdeYouden
Activación de basófilos						
Apostolou ¹⁸	86 (53;100)	100 (93;100)	12 (2;47)	0 (0;39)	93 (68;99)	0,9 (0,7;1,1)
Sanz ¹⁹	100 (97;100)	91(69;100)	0 (0;28)	4,3 (1;21)	97 (85;99)	0,96 (0,87;1,04)
Sturm (abejas) ²⁰	87 (72;100)	93 (82;100)	10 (3;26)	9 (2;27)	91(80;96)	0,81 (0,66;0,97)
Sturm (avispas) ²⁰	85 (72;99)	83 (68;98)	17 (7;34)	15 (6;30)	84 (74;91)	0,69 (0,51;0,87)
Rubio ²¹	99 (97;100)	58 (36;80)	7 (1;30)	9 (5;16)	91(85;95)	0,84 (0,71;0,98)
Eberlein (CD203) ²²	93 (84;100)	96 (85;100)	12 (4;30)	2 (0;12)	94 (86;98)	0,86 (0,72;0,99)
Eberlein (CD63) ²²	100 (99;100)	83 (68;98)	0 (0;13)	12 (5;24)	93 (84;97)	0,88 (0,79;0,98)
Sanz ²³	95 (83;100)	49 (35;63)	7 (2;21)	50 (37;62)	65 (54;74)	0,43 (0,29;0,59)
Frezzolini ²⁴	100 (99;100)	51(34;68)	0 (0;16)	33 (22;46)	75 (65;84)	0,67 (0,54;0,79)
Páris ²⁵	100 (98;100)	92 (81;100)	0 (0;10)	9 (3;23)	95 (88;98)	0,91 (0,82;1,01)
Monneret ²⁶	100 (97;100)	52 (35;69)	0 (0;15)	48 (34;64)	68 (56;79)	0,51 (0,36;0,67)
Eberlein ²⁷	92 (74;100)	47 (22;72)	10 (2;40)	45 (27;63)	66 (48;80)	0,45 (0,17;0,72)
Sanz ²⁸	96 (88;100)	49 (35;63)	3 (1;17)	52 (39;54)	65 (54;74)	0,45 (0,31;0,59)
Wolanczyk ²⁹	100 (97;100)	90 (75;100)	0 (0;17)	9 (2;28)	95 (84;99)	0,91 (0,79;1,03)
Ocmant (CD63) ³⁰	93 (83;100)	87 (74;100)	7 (2;22)	12 (5;28)	90 (80;95)	0,81 (0,66;0,95)
Ocmant (CD203c) ³⁰	97 (88;100)	90 (78;100)	3 (1;17)	9 (3;24)	93 (84;97)	0,87 (0,75;0,99)
Ocmant (huevo, CD63) ³⁰	100 (98,2-100)	88 (73;100)	0 (0;15)	10 (3;25)	94 (85;98)	0,90 (0,80;1,01)
Ocmant (huevo,CD203) ³⁰	95 (84;100)	66 (47;83)	4 (1;22)	35 (21;53)	77 (64;86)	0,60 (0,41;0,79)
Subtotal	96 (94;98)	71 (67;75)	5 (3;8)	22 (19;26)	84 (82;86)	0,72 (0,71;0,73)
Linfoproliferación						
TsugeI ⁵²	87(58;100)	100(95;100)	9(2;38)	0(0;35)	94(74;99)	0,91 (0,74;1,08)
Inmunoglobulina E específica						
Dang ⁵³	100(99;100)	71(63;79)	2(1;7)	40(31;50)	79(73;84)	0,58 (0,48;0,68)
Corey ⁵⁴	76(72;80)	83(81;86)	12(10;15)	30(26;34)	81(79;83)	0,58 (0,53;0,62)
Petersen ⁵⁵				10(7;13)	89(86;92)	
Biagini ⁵⁶	96(94;99)	71(66;78)	4(2;8)	22(18;27)	85(81;87)	0,74 (0,68;0,79)
Subtotal	90(89;92)	77(74;79)	10(8;12)	22(20;24)	83(82;85)	0,69 (0,68;0,70)

VPP: Valor Predictivo Positivo. VPN: Valor Predictivo Negativo.

Figura 4
Comparación de la validez diagnóstica de la prueba de de activación de basófilos y la Inmunoglobulina E específica



Al comparar las prueba de activación de basófilos frente a la IgE específica se halló una mayor área bajo la curva para la primera (0,97) frente a la segunda (0,87), evidenciando una mejor validez diagnóstica (figura 4).

DISCUSIÓN

En el diagnóstico de alergias, las pruebas inmunológicas *in vitro* más empleadas son la linfoproliferación, la activación de basófilos y la IgE específica. Con respecto a dichas técnicas, un número creciente de estudios ha demostrado que la cuantificación por citometría de flujo de la activación *in vitro* de basófilos puede ser una herramienta válida y fiable para medir las respuestas alérgicas en pacientes. Debido a que la técnica de activación de basófilos acopla procesos semejantes a los que suceden *in vivo*, como lo es la fijación de antígenos a la IgE de los receptores FcεRI y FcγRII de la superficie de los basófilos. Cuando existe contacto de un basófilo sensibilizado con un alérgeno específico que se une a la IgE, los basófilos secretan una amplia variedad de mediadores bioactivos y generan la expresión de marcadores de activación en su membrana como CD63 y CD203c que tras la activación celular se expresan en la membrana citoplasmática, lo cual permite su medición mediante citometría de flujo utilizando anticuerpos monoclonales específicos Anti CD63 y CD203c marcados con fluorocromos. Los mejores resultados de esta técnica se han visto cuando se hace diagnóstico de alergias a medicamentos³⁶⁻³⁸.

La técnica de la IgE específica y los métodos *in vivo* tienen un papel muy relevante en el diagnóstico de los procesos alérgicos y son los más empleados. Los anticuerpos IgE tienen un papel fundamental en las reacciones alérgicas a través de la unión con sus receptores de alta afinidad (FcεRI y FcγRII) presentes en mastocitos, basófilos y eosinófilos. Tras la

unión de la IgE con el alérgeno para la que es específica, dichos mediadores producen vasodilatación y broncoconstricción en los órganos diana¹⁻².

Desde el punto de vista inmunológico, las pruebas de activación de basófilos e IgE específica deberían generar resultados similares en la tamización de alergias, debido al rol de la IgE en ambas. En la primera la IgE se fija a la superficie del basófilo permitiendo la reacción con antígenos específicos y en la segunda la IgE se fija al antígeno implantado en una matriz donde se realizara la prueba de identificación³⁹⁻⁴².

Con base en el número de estudios, la técnica más empleada en la tamización y diagnóstico de alergias es la activación de basófilos. Sin embargo, al considerar el número de individuos, es la IgE específica y la técnica con menor uso (en número de estudios e individuos incluidos) es la linfoproliferación. Esto podría atribuirse al hecho de que esta técnica presenta una estandarización compleja y existe poco conocimiento sobre la concentración de antígenos necesaria en cada prueba, lo que puede generar que no se dé una reacción celular adecuada o que concentraciones elevadas de los antígenos generen muerte de las células de la prueba. No están claros los tiempos de incubación ni la modelación *in vitro* de los procesos de regulación celular, y la prueba presenta un alto costo. Estos factores han restringido su uso a laboratorios de alto nivel, especialmente de investigación⁴³⁻⁴⁴.

Junto con lo anterior, es oportuno aclarar que las pruebas *in vitro* tienen como fin identificar el alérgeno responsable de los síntomas y determinar el mecanismo inmunológico de la reacción alérgica. Dependiendo de la fracción sanguínea que se analice, se pueden clasificar en serológicos y celulares. Los primeros tienen como finalidad identificar y cuantifi-

car anticuerpos contra alérgenos específicos y medir la presencia de mediadores de reacción alérgica como la histamina y la triptasa. Las técnicas celulares utilizan células efectoras de la reacción alérgica que son aisladas en sangre periférica del paciente, siendo más empleados los linfocitos y basófilos. La prueba de transformación de linfocitos mide específicamente la proliferación de los linfocitos T de memoria en presencia de un antígeno específico y la expresión de marcadores en su superficie, mientras que la prueba de activación de basófilos mide la expresión de marcadores de activación ante un alérgeno para el cual están sensibilizados⁴⁴⁻⁴⁵.

Por otra parte, es oportuno precisar que en el diagnóstico de alergias las pruebas epicutáneas y la provocación con alimentos o medicamentos son ampliamente usadas⁴⁶, particularmente la provocación puede desarrollarse con la administración de la sustancia por ingestión, inhalación o inyecciones subcutáneas, intramusculares o intravenosas. Estos métodos son buenos en cuanto al diagnóstico pero presentan grandes inconvenientes con las reacciones adversas que pueden presentar los pacientes, como diarrea, vómitos y mareos, entre otros⁴⁷.

En la evaluación de la validez, desempeño, seguridad y eficiencia diagnóstica de la prueba de activación de basófilos y la IgE específica se halló una sensibilidad mayor de 70%, especificidad superior a 90%, CP+ de 10 o más, CP- inferior a 0,3; ORD mayor a 40 y pacientes correctamente diagnosticados mayor al 83%. Estos resultados evidencian que ambas pruebas son de gran utilidad clínica y diagnóstica en la medida que discriminan de forma correcta a las personas de las sanas y enfermas. Aunque se debe precisar que en la activación de basófilos fue estadísticamente mayor el valor predictivo positivo, en la IgE específica la proporción de falsos positivos y, en términos

globales, la activación de basófilos presentó una mayor área bajo la curva frente a la IgE específica, lo que sugiere su mayor validez diagnóstica¹⁶.

Es pertinente precisar que la sensibilidad y la especificidad permiten dimensionar la validez de una prueba y son descriptores diagnósticos *per se* (inherentes a la prueba, independientemente de la prevalencia de la enfermedad en la población a la cual se aplica). No obstante, su utilidad clínica es limitada dado que frente a un paciente se requiere resolver dudas en sentido contrario al de la probabilidad *a priori*, es decir, se busca conocer la probabilidad de que el paciente esté enfermo dado que el resultado de la prueba fue positivo o que esté sano dado que el resultado diagnóstico fue negativo. En este sentido, resultan de mayor utilidad clínica los valores predictivos positivo y negativo, los cuales presentan la limitación de ser dependientes de la prevalencia de la enfermedad en la población de personas de similares características al individuo consultante⁴⁸.

Frente a las limitaciones de la sensibilidad, especificidad y valores predictivos, el cálculo y aplicación de cocientes de probabilidad (razones de verosimilitud o *likelihood ratios*) constituye una herramienta de gran utilidad clínica en la medida que ayudan a cuantificar la probabilidad pre-prueba de ocurrencia de un diagnóstico. Cada resultado posible de una prueba posee un CP a partir del cual es posible estimar cuánto aumenta o disminuye la probabilidad pre-prueba, permitiendo al clínico descartar un diagnóstico, confirmarlo o continuar su investigación mediante la solicitud de nuevos exámenes⁴⁹.

A pesar de las consideraciones anteriores, la mayoría de estudios sobre pruebas diagnósticas en inmunología son incompletas, ya que sólo aluden a los aspectos

de validez (sensibilidad y especificidad) y obstan los parámetros de desempeño, seguridad y eficiencia diagnóstica, lo que sustenta la necesidad de ampliar el estudio de dichas pruebas entre los clínicos, académicos e investigadores para poder orientar el tratamiento de las alergias de forma idónea.

En la investigación y praxis clínico-epidemiológica no existen pruebas diagnósticas ideales que contribuyan a una correcta clasificación de la totalidad de pacientes. En el caso específico de las alergias, los principales factores que sustentan este hecho son el intervalo de tiempo transcurrido entre la sensibilización del paciente y la realización de las pruebas diagnósticas que puede generar falsos negativos y las reacciones cruzadas que derivan en falsos positivos cuando hay sensibilización de los individuos a antígenos cuaternarios con estructura similar a algunos antígenos⁵⁰.

Aunado a la utilidad diagnóstica hallada para la activación de basófilos y la IgE específica, se debe explicitar que estas presentan otras características que evidencian su relevancia en la tamización, como presentar buena disponibilidad, seguridad, bajo costo y aceptabilidad cultural, identificar individuos para los cuales el fácil brindar tratamiento y que no pueden pasar inadvertidos por su impacto en la salud individual y colectiva⁵¹.

Se debe aclarar que una prueba parte del planteamiento de una probabilidad pre-prueba cuya precisión depende de la clínica, por ello cobra importancia el conocimiento teórico, la experiencia clínica y la evaluación del paciente para recolectar información que optimice el rendimiento de la prueba diagnóstica bajo estudio. Además, en el marco de la medicina basada en la evidencia, en la actualidad se dispone de múltiples herramientas para analizar críticamente diferentes grados de evidencia que orientan la toma de decisiones en diversas áreas, inclusive la evaluación de pruebas diagnósticas⁵²⁻⁵⁴.

Cabe aclarar que este estudio presenta varias limitaciones relacionadas con la heterogeneidad de las investigaciones individuales, tanto en los valores de sensibilidad y especificidad como en el tipo de población y alergia estudiada. Además, en los estudios en que la identificación de los enfermos fue bajo criterios clínicos, las categorías diagnósticas no se describen de forma detallada, y en el caso de la determinación de la IgE específica no se explicitan aspectos relacionados con la técnica de medición, lo que en conjunto puede afectar el grado de inferencia de los resultados del metaanálisis. En esta revisión sistemática se observó que la mayor parte de los artículos provienen de Europa, esto puede atribuirse a tres factores: i) en este continente se encuentran los centros de estudios más importantes sobre alergias, entre ellos la Academia Europea de Alergología e Inmunología Clínica (EAA-CI), considerada la asociación médica mas grande en el estudio de las alergias en el mundo, la cual cuenta con medios de difusión específicos del área como las revista *Allergy*, *Pediatric allergy and immunology*, *Clinical and translational allergy*, la Federación Europea de Alergias (EFA) entre otras organizaciones ii) en Europa se documentan las mayores cifras de ocurrencia de la enfermedad, en la actualidad cerca de 113 millones de ciudadanos en Europa sufren rinitis alérgica y unos 68 millones de asma alérgico, iii) se estima que en Europa, a pesar de contar con buenos programas para el manejo de las alergias, existe un 45% de pacientes alérgicos no diagnosticados, lo que motiva la realización de estudios en este tópico, buscando disminuir el sub-registro, y iv) existen amplias divergencias en los criterios y pruebas diagnósticas lo que sustenta la necesidad de este tipo de investigaciones⁵⁵.

En conclusión, el estudio evidenció que la activación de basófilos y la IgE específica son pruebas útiles en el diagnóstico de alergias. Además, generó información de gran relevancia para los clínicos, académicos e investigadores interesados en la temática.

Resulta determinante reunir, evaluar y sistematizar los estudios disponibles sobre la validez, el desempeño, la seguridad y la efectividad de las pruebas diagnósticas disponibles para alergias, con el fin de esclarecer su utilidad, mejorar y aumentar su uso y por conexidad, orientar su selección o utilización para generar evidencia que sustente las decisiones clínicas y epidemiológicas.

La revisión incluyó personas en un amplio rango etario y con diversos tipos de alergias, corroborando el mejoramiento en la validez externa que se obtiene con esta modalidad de investigación, en la medida que los hallazgos de validez, desempeño, seguridad, eficiencia y utilidad diagnóstica no están circunscritos a un único grupo etario o tipo de alergia.

El metaanálisis presentó múltiples ventajas frente a los estudios individuales entre los cuales se destacan i) la evaluación de la utilidad diagnóstica completa con lo que superaron los hallazgos de los estudios individuales que sólo refieren los valores de sensibilidad y especificidad, ii) la generación de un resultado global de utilidad diagnóstica que permite superar las divergencias en los estudios individuales, iii) el aumento de la validez interna y externa de las conclusiones al utilizar métodos estadísticos más robustos y una población con mayor grado de generalización frente a los estudios individuales, iv) generación de información sin sesgos potenciales de selección y de extracción de datos, válida y reproducible, v) mayor grado de evidencia y menor costos que otros estudios lo suficientemente grandes para llegar al mismo nivel de significación estadística⁵⁶.

BIBLIOGRAFÍA

1. Palomo I, Ferreira A, Roseblatt M, Sepúlveda C, Vergara U. Células y órganos del sistema inmune. En: Palomo I, Pereira Jaime, Koenig C. editores. Fundamentos de inmunología básica y clínica. Santiago de Chile: Editorial Universidad de Talca; 2002. p. 53-86.
2. Holgate S, Church M, Broide D, Martinez F. Allergy. Londres: Elsevier; 2012; 4. p. 1-32.
3. Romero G, Pereira Q, Zini R, Canteros G. Reacciones de hipersensibilidad. Rev Posgrado de la VIa Cátedra de Medicina. 2007; 167:11-6.
4. Cardona R, Serrano C. Aspectos básicos de las reacciones de hipersensibilidad y alergia: Alergia abordaje clínico diagnóstico y tratamiento. México DF: Editorial Médica Panamericana. 2010. p. 3-27.
5. Kindt T, Goldsby R, Osborne B. Reacciones de hipersensibilidad: Inmunología de Kuby. 6a ed. California: Mc Graw Hill; 2007; 6. p. 371-400.
6. Zubeldia J, Baeza L, Jáuregui R, Senet C. Importancia actual de las enfermedades alérgicas: Libro de las enfermedades alérgicas fundación BBVA. España: Fundación BBVA; 2012. p.41-47.
7. Méndez I, Huerta J, Bellanti J, Ovilla R, Escobar A. Epidemiología de las enfermedades alérgicas: Alergia enfermedad multisistémica fundamentos básicos y clínicos. 1a .ed. México: Editorial Médica Panamericana; 2008. p.7-18.
8. Holgate S, Church M, Broide D, Martinez F. Allergy. Londres: Elsevier; 2012; 4. p. 129-146.
9. Cox L, Williams B, Sicherer S, Oppenheimer J, Sher J, Hamilton R, et al. Pearls and pitfalls of allergy diagnostic testing: report from the American College of Allergy, Asthma and Immunology/American Academy of Allergy, Asthma and Immunology Specific IgE Test Task Force. Ann allergy asthma immunoh 2008; 101(6):580-92.
10. Malbran A, Rey G, Galassi N, Yeyati E. El laboratorio en el diagnóstico de alergias. Disponible en http://www.alergia.org.ar/profesionales/emc/prodaai2008/laboratorio_diagnostico_alergia.pdf
11. Ortiz Z. Qué son las revisiones sistemáticas. Centro de investigaciones epidemiológicas. 2005. Disponible en http://www.epidemiologia.anm.edu.ar/pdf/publicaciones_cie/2005/Que_son_revisiones_sistematicas_2005.pdf.
12. Manterola C. Revisión sistemática de la literatura síntesis de la evidencia. Rev méd Clín Las Condes. 2009; 20(6):897-903.
13. Letelier M. Revisiones sistemáticas y metaanálisis: ¿son la mejor evidencia? Rev Med Chile. 2005; 133(2):246-9.

14. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003; 3: 25.
15. Zamora J, Abaira J. Análisis de la calidad de los estudios de evaluación de pruebas diagnósticas. *Nefrología.* 2008; 28(2):42-5.
16. Universidad de Granada. Valoración de pruebas diagnósticas. Disponible en: www.ugr.es/~bioest/medicina/Pruebas_Diagnosticas.xls.
17. Zamora J, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol.* 2006; 6:31.
18. Apostolou E, Deckert K, Puy R, Sandrini A, De León M, Douglass J, et al. Anaphylaxis to Gelofusine confirmed by in vitro basophil activation test: a case series. *Anaesthesia.* 2006; 61(3):264-8.
19. Sanz M, García M, Tabar A, Anda M, García B, Barber D, et al. Basophil Activation Test and specific IgE measurements using a panel of recombinant natural rubber latex allergens to determine the latex allergen sensitization profile in children. *Pediatr Allergy Immunol.* 2006; 17: 148-56.
20. Sturm G, Böhm E, Trummer M, Weiglhofer I, Heinemann A, Aberer W. The CD63 basophil activation test in Hymenoptera venom allergy: a prospective study. *Allergy* 2004; 59: 1110-7.
21. Rubio A, Bourrier, Vivinus T, Saggio B, Albertini M, Bernard A. Benefit of the basophil activation test in deciding when to reintroduce cow's milk in allergic children. *Allergy.* 2011; 66:92-100.
22. Eberlein B, Varga R, Mempel M, Darsow U, Behrendt H, Ring J. Comparison of basophil activation tests using CD63 or CD203c expression in patients with insect venom allergy. *Allergy* 2006; 61: 1084-5.
23. Sanz M, Gamboa P, Antépara I, Uasuf C, Vila L, García C, et al. Flow cytometric basophil activation test by detection of CD63 expression in patients immediate-type reactions to betalactam antibiotics. *Clin Exp allergy.* 2002; 32(2):277-86.
24. Frezzolini A, Cadoni S, Pita O. Usefulness of the CD63 basophil activation test in detecting Anisakis hypersensitivity in patients with chronic urticaria: diagnosis and follow-up. *Clin Exp Dermatol.* 2010;35(7):765-70.
25. Pâris A, Demoly P, Persi L, Lebel B, DPharma, Bousquet J, et al. In vitro diagnosis of cypress pollen allergy by using cytofluorimetric analysis of basophils (Basotest). *J Allergy Clin Immunol.* 2000; 105(2):339-45.
26. Monnereta G, Benoit Y, Debarde A, Gutowska M, Topenot I, Bienvenu J. Monitoring of Basophil Activation Using CD63 and CCR3 in Allergy to Muscle Relaxant Drugs. *Clin Immunol.* 2002; 102(2):192-9.
27. Eberlein B, Krischan L, Darsow U, Ollert M, Ring J. Double positivity to bee and wasp venom: Improved diagnostic procedure by recombinant allergen-based IgE testing and basophil activation test including data about cross-reactive carbohydrate determinants. *J Allergy Clin Immunol.* 2012; 130(1): 155-61.
28. Sanz M, García M, Caballero M, Dieguez I, Gamboa P. Test de activación de basófilos en el diagnóstico de alergia a medicamentos. *An. Sis.Sanit. Navar.* 2003; 26 (2): 39-47.
29. Wolanczyk A, Gogolewski G, Liebhart J, Gomulka K, Litwa M, Panaszek B. A New Variant of the Basophil Activation Test for Allergen-Induced Basophil CD63 Upregulation. The Effect of Cetirizine. *J Investig Allergol Clin Immunol.* 2009; 19(6): 465-73.
30. Ocmant A, Mulier S, Hanssens L, Goldman M, Casimir G, Mascart F, et al. Basophil activation tests for the diagnosis of food allergy in children. *Clin Exp Allergy.* 2009; 39(8): 1234-45.
31. Tsuge I, Okumura A, Kondo Y, Itomi S, Kakami M, Kawamura M, et al. Allergen-specific T-cell Response in Patients with Phenotypic Hypersensitivity; Simultaneous Analysis of Proliferation and Cytokine Production by Carboxyfluorescein Succinimidyl Ester (CFSE) Dilution Assay. *Allergol Int.* 2007;56(2):149-55.
32. Dang, T, Mimi T, Choo S, Licciardi P, Koplin J, Martin P, et al. Increasing the accuracy of peanut allergy diagnosis by using Ara h 2. *J Allergy Clin Immunol* 2012; 129(4):1056-63.
33. Corey J, Mamikoglu B, Akbar I, Houser S, Gungor A. ImmunoCAP and HY•TEC enzyme immunoassays in the detection of allergen-specific IgE compared with serial skin end-point titration by receiver operating characteristic analysis. *Otolaryngol Head Neck Surg.* 2000; 122(1): 64-70.
34. Petersen A, Gudmann P, Milvang P, Mørkeberg R, Bøgestrand S, Linneberg A, et al. Performance evaluation of a specific IgE assay developed for the ADVIA centaur immunoassay system. *Clin Biochem.* 2004; 37(10):882-92.

35. Biagini R, Krieg E, Pinkerton L, Hamilton R. Receiver Operating Characteristics Analyses of Food and Drug Administration-Cleared Serological Assays for Natural Rubber Latex-Specific Immunoglobulin E Antibody. *Clin Diagn Lab Immunol*. 2001;8(6):1145-9.
36. Valovirta E. Libro de EFA sobre Alergias Respiratorias. Aumentar la conciencia y aliviar la carga. Italia: officine Grafiche Francesco Giannini & Figli s.p.a. 2011.
37. Ebo D, Bridts C, Hagendorens M, Aerts N, Clerck L, Stevens W. Basophil activation test by flow cytometry: present and future applications in allergology. *Cytometry B Clin Cytom*. 2008;74(4):201-10.
38. Chirumbolo S. Use of basophil activation test in the investigation of adverse effects to vaccines. *Hum Vaccin*. 2011;7(8): 878-80.
39. Galassi N, Riera N, Rey G, Bracco M, Malbran A. Urticaria crónica evolución clínica, prueba del suero antólogo, recuento y activación de basófilos. *Medicina (Buenos Aires)* 2003; 63(1):15-20.
40. Mayorga C, Sanz M, Gamboa P, García B, García J, Labrador M et al. In Vitro Diagnosis of Immediate Allergic Reactions to Drug An Update. *J Investig Allergol Clin Immunol*. 2010; 20(2): 103-9.
41. Jing L, Bardina L, Shreffler W, Andreae D, Ge Y, Wang J, et al. Development of a novel peptide microarray for large-scale epitope mapping of food allergens. *J Allergy Clin Immunol*. 2009; 124(2):315-22.
42. Shyur S, Jan R, Wester J, Chang P, Jung lu Y, Yao Wang J. Determination of multiple allergen-specific IgE by microfluidic immunoassay cartridge in clinical settings. *Pediatric Allergy Immunol*. 2010; 21(4):623-33.
43. Kenny Y, Jacqueline S. Kwong, Jacqueline S. Sabet E, Mendoza G, Platts-Mills; T Ryan. The Benefits of Specific Immunoglobulin E Testing in the Primary Care Setting. *Am J Manag Care*. 2011; 17(S17):447-59.
44. Kastylevsky G, Vial T, Descotes J. Drug allergy diagnosis in humans possibilities and pitfalls. *Toxicology*. 2001 158; 1-10.
45. Zubeldia J, Baeza L, Jáuregui R, Senet C. Análisis de sangre para el estudio de alergias: Libro de las enfermedades alérgicas fundación BBVA. 1a. ed. Madrid: Fundación BBVA; 2012.p.371-8.
46. Chivato T. Estudio alérgico in vitro. Guía rápida para residentes de alergología. Madrid. Luzan 5 ediciones. 2009. p. 47-55.
47. Méndez I, Huerta J, Bellanti J, Ovella R, Escobar A. Apoyos diagnósticos de las alergias: Alergia enfermedad multisistémica fundamentos básicos y clínicos. 1a .ed. México DF: Panamericana; 2008.p. 375-82.
48. Zubeldia J, Baeza L, Jáuregui R, Senet C .Pruebas de provocación con alimentos y medicamentos: Libro de las enfermedades alérgicas fundación BBVA. España: Fundación BBVA; 2012. p. 379-85.
49. Cerda J, Cifuentes L. Uso de tests diagnósticos en la práctica clínica (Parte 1): Análisis de las propiedades de un test diagnóstico. *Rev. chil. infectol*. [online]. 2010;27(3):205-8.
50. Cifuentes L, Cerda J. Uso de tests diagnósticos en la práctica clínica. Aplicación clínica y utilidad de un test diagnóstico. *Rev. Chil. Infectol* [online]. 2010; 27(4):316-9.
51. Ebo G, Leysen J, Mayorga C, Rozieres A; Knol E; Terreehorst I. The in vitro diagnosis of drug allergy status and perspectives. *Allergy*. 2011; 66: 1275-86.
52. Alonso E, Zapatero L Martínez M. Educación del paciente alérgico prevención de riesgos y tratamiento. Disponible en: <http://www.aeped.es/sites/default/files/documentos/8-educacion-alergico.pdf>
53. Guyatt G, Drummond R Meade M, Cook D. Users's Guides to the Medical Literature. Essentials of Evidence-Based Clinical Practice. 2 a .ed. Nueva York: McGraw Hill Professional. 2002.
54. Vera C, Letelier L M, Carvajal J. Guía para el análisis crítico de estudios que evalúan exámenes diagnósticos. *Rev Chil Obstet Ginecol*. 2005; 70: 196-202.
55. Valenzuela L, Cifuentes L. Validez de estudios de tests diagnósticos. *Rev Méd Chile* 2008; 136: 401-4.
56. Marin F; Meca J, López J. El metaanálisis en el ámbito de las Ciencias de la Salud una metodología imprescindible para la eficiente acumulación del conocimiento. *Fisioterapia*. 2009; 31(3):107-14.

REVISIÓN SISTEMÁTICA Y METAANÁLISIS

IMPACTO DEL VIH/SIDA SOBRE LA CALIDAD DE VIDA: METAANÁLISIS 2002-2012 (*)

Jaiberth Antonio Cardona-Arias y Luis Felipe Higueta-Gutiérrez

Escuela de Microbiología Universidad de Antioquia. Facultad de Medicina Universidad Cooperativa de Colombia. Grupo de investigación Salud y Sostenibilidad. Universidad de Antioquia. Medellín. Colombia.

(*) El estudio fue financiado con recursos de la Universidad de Antioquia.

Ninguno de los autores declara conflicto de intereses para la publicación del manuscrito.

RESUMEN

Fundamentos: El VIH/Sida es una infección crónica que trasciende el dominio biológico, afectando a las relaciones sociales, la salud mental, los aspectos económicos y la calidad de vida. El objetivo del estudio fue evaluar la percepción de la calidad de vida relacionada con la salud (CVRS) en personas con VIH/sida.

Métodos: Revisión sistemática con meta-análisis. Se realizó una búsqueda bibliográfica en las bases Pubmed, Science direct, Lilacs y Scielo con los siguientes términos de búsqueda: WHOQOL-BREF, WHOQOL-HIV-BREF, WHOQOL-BREF-HIV. Se evaluó la calidad de los artículos por dos investigadores para garantizar la exhaustividad y reproducibilidad. La CVRS se midió en una escala de 0 (peor resultado) a 100 puntos. El análisis se basó en frecuencias, medidas de resumen, intervalos de confianza y prueba t student.

Resultados: Se incluyeron 19 investigaciones con 3.348 personas de 11 países. Se obtuvo un promedio de CVRS entre 59 y 64 puntos, el menor puntaje se encontró en la dimensión ambiental (59) y el mayor en la salud psicológica (64). En el análisis de sensibilidad se hallaron cambios entre 0 y 5 puntos entre el puntaje global y el observado al excluir los datos de cada estudio individual. Se halló buena validez convergente en las dimensiones comunes a los dos instrumentos evaluados, con diferencias menores a 3 puntos.

Conclusión: Los puntajes de CVRS en personas con VIH/sida fueron mayores a 50, lo que implica que su adecuado manejo puede disminuir los efectos negativos sobre la salud física, psicológica, social, ambiental, el nivel de independencia y la espiritualidad.

Palabras clave: VIH. Sida. Calidad de vida. Salud. Organización Mundial de la Salud.

Correspondencia

Jaiberth Antonio Cardona Arias
Calle 67 Número 53 – 108, Bloque 5, oficina 103
Medellín, Colombia
Teléfono 2198486. Fax 2195486.
Correo electrónico jaiberthcardona@gmail.com

ABSTRACT

Impact of HIV / AIDS on Quality of Life: Meta-Analysis 2002-2012

Background: HIV/AIDS is a chronic infection that transcends the biological domain and affects social relations, mental health, financial aspects and quality of life. The objective of the study was to assess the perception of Health Related Quality Of Life (HRQOL) in people with HIV / AIDS.

Methods: Systematic review with meta-analysis. We applied a research protocol in four multidisciplinary databases: Pubmed, Science direct, Lilacs y Scielo; with search terms WHOQOL-BREF, WHOQOL-HIV-BREF, WHOQOL-BREF-HIV. Quality assessment of articles was done by two researchers to ensure the completeness and reproducibility. HRQL was measured on a scale of 0 (worst outcome) to 100 points. The analysis was based on frequencies, summary measures, confidence intervals, and Student t test.

Results: The study included 19 studies with 3,348 people from 11 countries. The average of HRQL was between 59 and 64 points, the lowest score was the environmental dimension with 59 points and the greater psychological health with 64 points. In the sensitivity analysis we found changes of 0 and 5 points between overall score and the observed data by excluding each individual study. We found good convergent validity on the dimensions common to both instruments evaluated, with differences of less than 3 points.

Conclusion: HRQL scores in people with HIV/aids were good (greater than 50) which implies that proper management can reduce the negative effects on the physical health, psychological health, social relations, the environment, independence level and spirituality.

Keyword: HIV. AIDS. Quality of life. Health. World Health Organization.

INTRODUCCIÓN

La infección por el VIH/sida afecta a cerca de 34 millones de personas en el mundo de las que 23,5 millones se encuentran en África subsahariana y los 10,5 millones restantes se distribuyen en Oceanía (53.000), Oriente Medio, África septentrional (300.000), Europa (2,3 millones), las Américas (2,8 millones) y Asia (4,8 millones)¹. Debido al elevado número de personas y familias afectadas, a su carácter epidémico y al elevado costo de su atención, el VIH/sida es considerado como una prioridad en salud pública. Además, todas las actividades de diagnóstico, tratamiento, seguimiento y control están contempladas en los sistemas de seguridad social, lo que ha derivado en la ampliación del acceso a la terapia antirretroviral y con ello a un descenso cercano al 50% de los fallecimientos relacionados con el sida^{1,2}.

De acuerdo con lo anterior, la expectativa de vida de los sujetos infectados ha aumentado considerablemente y la infección ha pasado a ser una enfermedad crónica tratable, cuya atención trasciende el dominio biológico, tomando mayor importancia para el manejo integral del paciente y el mejoramiento de su calidad de vida y la de sus familiares, aspectos como la estigmatización, la discriminación, la desigualdad entre sexos, las relaciones sociales, la salud mental, la independencia económica, la seguridad del entorno, entre otros, que constituyen grandes necesidades y en algunos contextos obstáculos^{1,3,4}.

Existen divergencias respecto a la definición del concepto calidad de vida y también respecto a su medición. Algunos investigadores la conciben como la calidad de las condiciones de vida de una persona, otros hacen alusión a la satisfacción experimentada por la persona con dichas condiciones vitales y un ter-

cer grupo combina los componentes objetivos y subjetivos, es decir, las condiciones materiales de vida de una persona y la satisfacción que ésta experimenta⁵. Estos aspectos aplicados a personas enfermas dan lugar al concepto de Calidad de Vida Relacionada con la Salud (CVRS). En este orden de ideas, se han establecido unas características transversales a la definición de calidad de vida y CVRS que permiten abordar de manera multidimensional la percepción del sujeto frente a su salud física, mental, sus relaciones sociales y el entorno en el que vive⁵. Estos lineamientos han sido utilizados como referencia para el diseño de instrumentos que permiten medir la CVRS y que son utilizados en la práctica clínica para dar cuenta del impacto de una intervención médica, detectar cambios en el estado de salud, facilitar la comunicación entre personas afectadas y el personal sanitario y generar información relevante sobre la adherencia terapéutica⁶.

Entre los múltiples instrumentos que se han diseñado para medir la CVRS destacan el *World Health Organization Quality Of Life* versión breve (WHOQOL-BREF) y el *World Health Organization Quality Of Life* versión breve para personas con VIH/sida (WHOQOL-HIV-BREF), los cuales presentan una excelente estructura conceptual y operativa, buen desarrollo psicométrico y confiabilidad, han sido adaptados a múltiples países y diferentes idiomas, presentan validación transcultural e incluyen las dimensiones de la CVRS que la Organización Mundial de la Salud (OMS) considera más relevantes⁷⁻⁸.

El WHOQOL-BREF es un instrumento genérico que produce un perfil de CVRS en 4 dimensiones: salud física, psicológica, relaciones sociales y medio ambiente. Se puede aplicar en personas enfermas y sanas, lo que permite realizar

comparaciones, establecer valores de referencia, identificar desviaciones en los valores esperados en una población y se utiliza como marco de referencia para establecer objetivos terapéuticos⁷. El WHOQOL-HIV-BREF es un instrumento específico para personas con VIH/SIDA, incluye las 26 preguntas del WHOQOL-BREF más 5 ítems específicos para VIH/sida: 1) problemas físicos relacionados con la infección, 2) culpabilización, 3) temor al futuro, 4) preocupación por la muerte y 5) aceptación. Con estos se genera un perfil de CVRS con 2 dimensiones adicionales al WHOQOL-BREF: el nivel de independencia y la espiritualidad, lo que lo hace más sensible para detectar cambios y particularidades asociadas a la infección⁸.

Diferentes investigaciones han reportado que el abordaje de la CVRS en personas con VIH/sida, a través de este tipo de instrumentos resulta relevante para dar cuenta de aspectos importantes que circundan la infección y para identificar las dimensiones que pueden condicionar la adherencia al tratamiento, la forma de percibir la enfermedad, el desarrollo de enfermedades concomitantes y la progresión o curso de la infección⁹⁻¹².

A pesar de que existe un elevado número de investigaciones realizadas en CVRS de personas que viven con VIH/sida, presentan resultados divergentes en cuanto al impacto de la infección en las dimensiones del WHOQOL-BREF y el WHOQOL-HIV-BREF y discrepancias en lo referido a las dimensiones más afectadas. Asimismo, los estudios individuales tienen algunas limitaciones relacionadas con la validez externa de las conclusiones por pequeño tamaño de muestra y potencia estadística, no dan cuenta de diferencias relacionadas con el lugar de estudio o no presentan comparaciones entre los resultados del instrumento genérico y el específico. Razones que

sustentan la necesidad de combinar los resultados de estudios individuales y en general resumir las publicaciones sobre este tema.

El objetivo de este trabajo fue analizar el impacto del VIH/sida sobre la calidad de vida relacionada con la salud durante el período 2002-2012.

MATERIAL Y MÉTODOS

Tipo de estudio: Revisión sistemática de la literatura con meta-análisis.

Instrumentos de medición de calidad de vida: El WHOQOL-HIV-BREF es un instrumento genérico derivado del WHOQOL-100, que consta de 26 puntos, 24 de ellos generan un perfil de calidad de vida en 4 dimensiones: salud física, psicológica, relaciones sociales y medio ambiente, en una escala de 0-20 ó 0-100, siendo 0 el peor resultado y 20 ó 100 el mejor. Es aplicable a sujetos sanos y enfermos y su puntaje ha sido utilizado para observar los cambios en la calidad de vida debido a la enfermedad, para establecer valores pronósticos de la enfermedad y para valorar algunas intervenciones⁷. El WHOQOL-HIV-BREF es un instrumento específico que incluye 5 preguntas adicionales al WHOQOL-HIV-BREF con las cuales aumenta la sensibilidad para detectar cambios asociados a la infección por VIH/sida. En total contiene 31 puntos que se agrupan en 6 dominios: salud física, salud psicológica, relaciones sociales, ambiente, nivel de independencia y espiritualidad, proporcionando una puntuación similar a la descrita con el WHOQOL-HIV-BREF; sin embargo, se debe precisar que la forma de calcular el puntaje de cada dimensión en ambos instrumentos es diferente, por ejemplo la salud física en el WHOQOL-HIV-BREF se evalúa con siete ítems: energía, dolor, sueño, movilidad, capacidad de trabajo, dependencia a medicamentos y activida-

des de la vida diaria. De estos, sólo los tres primeros conforman el dominio de salud física en el WHOQOL-HIV-BREF y los cuatro últimos contribuyen al dominio de independencia⁷⁻⁸. Ambos instrumentos evalúan las percepciones del individuo en las dos o cuatro semanas previas a su aplicación, han sido traducidos a varios idiomas y validados transculturalmente⁷⁻⁸. La validación de la escala de la OMS en personas con VIH/sida presentó un α de Cronbach para la consistencia interna entre 0,74 y 0,85 en las dimensiones evaluadas, fiabilidad test-retest entre 0,64 y 0,79 y con respecto a la validez convergente/divergente se hallaron correlaciones mayores a 0,5 frente a escalas que avalúan estado de salud y felicidad⁸.

Protocolo de investigación: Se realizó una búsqueda de literatura científica en las bases de datos PubMed, Science direct, Lilacs y Scielo utilizando los términos de búsqueda WHOQOL-BREF, WHOQOL-HIV-BREF, WHOQOL-BREF-HIV. De manera adicional se revisaron las referencias bibliográficas de los artículos seleccionados para identificar otros que no se encontraron en la búsqueda inicial y se enviaron correos electrónicos a varios autores para recuperar artículos no disponibles en las bases de datos, todo ello con el fin de garantizar la exhaustividad de la investigación.

Los criterios de inclusión fueron los siguientes: 1) investigaciones con términos de búsqueda en el título y/o resumen; 2) publicadas entre enero de 2002 y octubre de 2012; 3) estudios publicados en español, inglés o portugués; 4) que fueran artículos originales; 5) podían ser estudios observacionales transversales y longitudinales, 6) que evaluaran todas las dimensiones del WHOQOL-BREF y/o el WHOQOL-HIV-BREF y 7) investigaciones realizadas en personas con VIH/sida. Como criterios de exclusión se definieron: 1) estudios en los que no se docu-

mente la puntuación de cada dominio del instrumento, 2) estudios de validación o en los que los instrumentos fueran empleados para validar otras escalas, 3) artículos con problemas de validez interna por un manejo estadístico inadecuado, resultados incongruentes con el objetivo del estudio, sin información sobre control de sesgos o que no presentaran todos los componentes de un estudio epidemiológico según lo planteado en las guías *Strengthening the Reporting of Observational studies in Epidemiology* (STROBE)¹³ y 4) investigaciones en personas con dos o más infecciones.

Recogida de la información: Para garantizar la exhaustividad del protocolo de investigación se realizó una búsqueda por sensibilidad, sin circunscribirla a términos de los tesauros Descriptores de Ciencias de la Salud (DeCS) o *Medical Subject Headings* (MeSH), lo que permitió la obtención de un mayor número de estudios frente a la búsqueda por especificidad. Los artículos obtenidos fueron exportados al programa Endnote-Web para la eliminación de duplicados. La aplicación del protocolo de investigación se llevó a cabo por dos investigadores de forma independiente para garantizar la reproducibilidad de la revisión, las discrepancias se resolvieron por consenso y con referencia de un tercero. La extracción de la información se realizó con base en unas instrucciones que contenían las variables a evaluar, con su naturaleza, codificación y nivel de medición y se almacenó en una base de datos Excel, lo que realizó cada investigador en dos ocasiones diferentes (en el intervalo de un mes) de forma independiente, con el fin de garantizar la reproducibilidad intra e inter-observador de la información recolectada y analizada.

Análisis de la información. Para evaluar el acuerdo inter e intra-codificadores se calculó el coeficiente kappa para las variables año de publicación (kap-

pa=1,00) y lugar de estudio (kappa=1,00), y coeficiente de correlación intraclase (CCI) para los puntajes de salud física (CCI=0,98), salud psicológica (CCI=0,99), relaciones sociales (CCI=0,96), ambiente (CCI=0,97), independencia (CCI=0,98) y espiritualidad (CCI=0,98).

Para describir los artículos se calcularon frecuencias absolutas, relativas e intervalos de confianza del 95% para proporciones, según las variables de tiempo (año de estudio), lugar (país) y persona (frecuencia de sujetos infectados incluidos) y promedio ponderado para la edad de los individuos incluidos. Para calcular el promedio ponderado se empleó el siguiente estadístico

$$\bar{X} = \frac{\sum_{i=1}^n W_i X_i}{\sum_{i=1}^n W_i}$$

Dado que algunos estudios reportaban el puntaje promedio de cada dimensión del WHOQOL-BREF y el WHOQOL-HIV-BREF en una escala de 0 a 20 se unificó en un rango de 0 a 100 según recomendaciones del Grupo de Salud Mental de la OMS⁷. Para los artículos que presentaban la desviación estándar de los puntajes de CVRS en una escala de 0-20 se transformó a una de 0-100 a partir del siguiente estadístico:

$$\text{Desviación estándar} = \text{Coeficiente de variación} * \text{Media}$$

con la media unificada en escala 0-100 según la guía de la OMS previamente citada.

Para describir la CVRS de la población de estudio se calcularon promedios ponderados para los puntajes de salud física, psicológica, social, ambiental, independencia y espiritualidad, cada uno de estos se acompañó de la desviación estándar ponderada. En cuatro estudios que no reportaban la medida de dispersión esta fue imputada a través del método

de “imputación simple con estimación por máxima verosimilitud”¹⁴.

Para comparar el puntaje ponderado de cada dimensión de CVRS con el resultado de los estudios individuales y determinar su homogeneidad o heterogeneidad, se calcularon intervalos de confianza para la media con el siguiente estadístico:

$$IC95\% = \text{Media} \pm [Z_{\alpha/2} * (\text{Desviación estándar} / \sqrt{n})]$$

Con estos resultados se realizó el Forest Plot (gráfico de cotizaciones en Excel). Previamente se hizo un análisis de sensibilidad para explorar la influencia de cada estudio sobre el puntaje global, para lo que se calculó la media ponderada y el cambio frente a la puntuación total (con su intervalo de confianza) excluyendo cada investigación.

Se realizaron intervalos de confianza para la diferencia de medias bajo el estadístico:

$$(\text{Media 1} - \text{Media 2}) \pm \left(Z_{\alpha/2} * \sqrt{\frac{\text{Desviación estándar 1}^2}{n1} + \frac{\text{Desviación estándar 2}^2}{n2}} \right)$$

con el fin de comparar los puntajes de las dimensiones comunes a ambos instrumentos: salud física, salud psicológica, relaciones sociales y ambiente, para evaluar la validez convergente/divergente de ambos instrumentos.

Los análisis se realizaron con el programa para análisis epidemiológico de datos tabulados de la Organización Panamericana de la Salud (Epidat) versión 3.0 y con Statistical Package for the Social Sciences for Windows, software SPSS versión 21,0[®] con una significación del 0,05.

RESULTADOS

En la búsqueda inicial se identificaron 164 artículos de los cuales se elimi-

naron 131 debido a la identificación de duplicidades y a la aplicación de los criterios de temporalidad, idioma y palabras claves en título y/o resumen. Se realizó la revisión completa de 33 estudios. De estos se eliminaron 14 porque no incluyeron las puntuaciones del instrumento, no se circunscribieron a sujetos con VIH/sida o abordaron personas con coinfecciones (figura 1).

El estudio incluyó 19 investigaciones que describieron la CVRS de 3.348 perso-

nas con VIH/sida, de 11 países diferentes, con un promedio de edad de 38 años y un rango entre 18 y 81 años (tabla 1).

Del total de artículos incluidos, 12 evaluaron CVRS con el WHOQOL-HIV-BREF, 6 utilizaron el WHOQOLBREF y 1 aplicó ambos instrumentos. En las investigaciones que emplearon el WHOQOL-HIV-BREF se obtuvo un promedio ponderado para el dominio de salud física de $60 \pm 8,3$ con un rango entre 50 y 72 puntos, en salud psicológica $61 \pm 7,7$

Figura 1
Algoritmo de selección de artículos

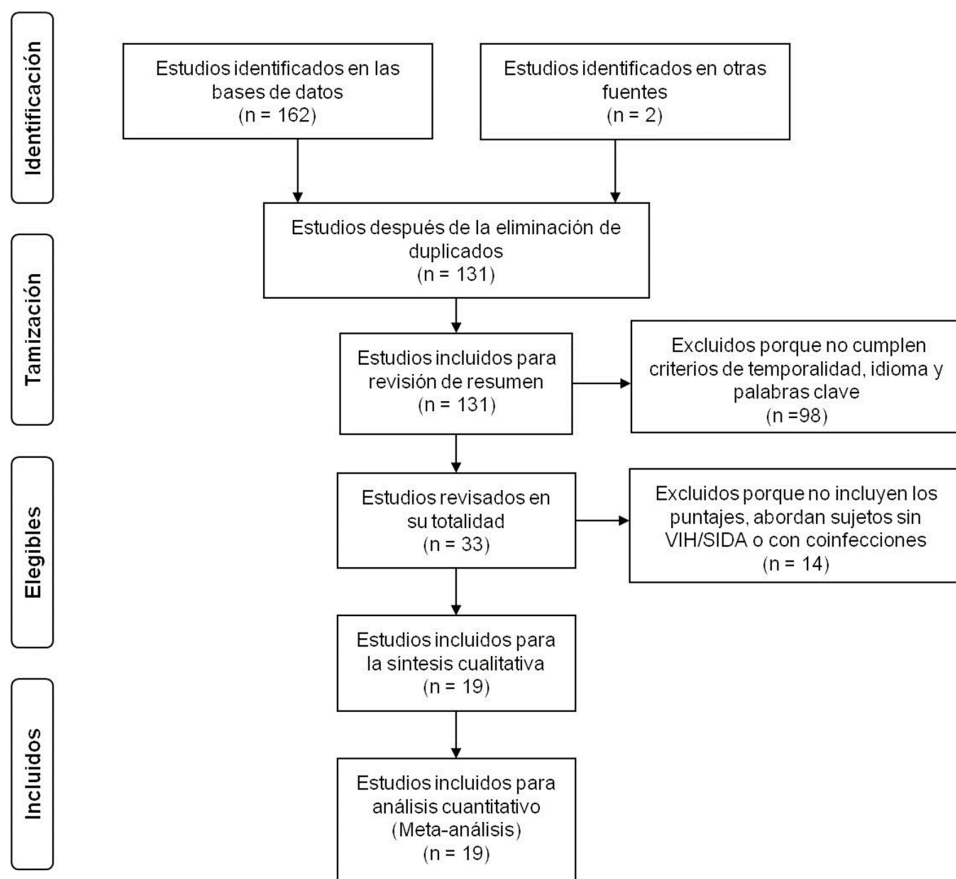


Tabla 1
Descripción de los estudios

Autor	Año	Lugar	Edad media (rango)	Sujetos	
				n	% (IC5)
Pereira M ¹⁵	2011	Portugal	40,7 (18-81)	1.191	35,6 (33,9-37,2)
Morandi E ¹⁶	2007	Brasil	39,3 (sin dato)	365	10,9 (9,8 - 11,9)
Mahalaskshmy T ¹⁷	2011	India	35 (>18)	200	5,9 (5,1 - 6,8)
Ucker C ¹⁸	2008	Brasil	38,5 (19-67)	200	5,9 (5,1 - 6,8)
Cardona J ¹⁹	2011	Colombia	>18	187	5,6 (4,8 - 6,4)
Bach X ²⁰	2012	Vietnam	31,4 (sin dato)	155	4,6 (3,9 - 5,4)
Mkangara O ²¹	2009	China	31,5 (18-55)	136	4,1 (3,4 - 4,7)
Sakthong P ²²	2009	Tailandia	36,2 (≥18)	120	3,6 (2,9 - 4,2)
Kovacevic S ²³	2006	Croacia	39,6 (sin dato)	111	3,3 (2,7 - 3,9)
Reis R ²⁴	2012	Brasil	≥20	106	3,1 (2,6 - 3,8)
Sheng Y ²⁵	2010	China	41,8 (sin dato)	93	2,8 (2,2 - 3,3)
Da Silva J ²⁶	2012	Brasil	43 (20-81)	85	2,5 (2,0 - 3,1)
Chandra P ²⁷	2006	India	30,7 (20-44)	82	2,4 (1,9 - 3,0)
Imam M ²⁸	2011	Bangladesh	34,8 (18-40)	82	2,4 (1,9 - 3,0)
Naveet W ²⁹	2006	India	32,4 (19-59)	68	2,0 (1,5 - 2,5)
Souza N ³⁰	2012	Brasil	40 (18-62)	58	1,7 (1,3 - 2,2)
Yen C ³¹	2004	China	33,4 (22-66)	41	1,2 (0,8 - 1,6)
Allen J ³²	2009	Australia	Sin dato (≥18)	41	1,2 (0,8 - 1,6)
Zaidi W ³³	2012	Malasia	43,9 (sin dato)	27	0,8 (0,5 - 1,1)
TOTAL			38,2 (18-81)	3.348	

(rango=50-81), en relaciones sociales $61 \pm 8,2$ (rango=56-69), en ambiente $59 \pm 6,7$ (rango=50-75), en independencia $64 \pm 7,5$ (rango=50-75) y en espiritualidad $63 \pm 8,2$ (rango=56-75). En referencia al WHOQOL-BREF, el promedio ponderado en el aspecto físico fue $62 \pm 13,6$ (rango=46-69), en el psicológico $64 \pm 14,4$ (rango=49-69), en las relaciones sociales $60 \pm 16,5$ (rango=44-64) y en el dominio ambiental $59 \pm 14,6$ (rango=50-68) (tabla 2).

En la tabla 3 se presenta el análisis de sensibilidad del estudio en el que se observó que el promedio global de las dimensiones CVRS cambian al excluir algunos estudios, aunque dichos cambios no fueron superiores a 3 puntos en el WHOQOL-HIV-BREF ni

a 5 puntos en el WHOQOL-BREF, lo que resultó clínicamente bajo, dado que la escala de ambos instrumentos oscila entre cero y cien y variaciones de 5 o menos puntos no resultan significativas.

En la figura 2 se presentan los promedios y sus intervalos de confianza de los dominios de CVRS del WHOQOL-HIV-BREF, tanto los globales como el de los estudios individuales, los cuales presentaron valores globales entre 59 (ambiente) y 64 (independencia), mientras que los estudios individuales presentaron puntajes en un rango entre 50 y 75 puntos, en las seis dimensiones se observaron diferencias estadísticamente significativas (dadas las diferencias en los intervalos de confianza) entre la puntuación

Tabla 2
Perfil de CVRS según el WHOQOL-BREF y el WHOQOL-HIV-BREF

Autor	n	WHOQOL-HIV-BREF (Media± Desviación estándar)					
		Física	Psíquica	Social	Ambiente	Indep ^a .	Espir ^b .
Pereira M ¹⁵	1.191	63±2	59±2	60±2	56±1	64±2	60±2
Mahalaskshmy T ¹⁷	200	56±19	56±18	63±17	63±13	69±19	56±11
Ucker C ¹⁸	200	72±15	58±11	63±12	63±9	61±12	75±14
Cardona J ¹⁹	187	60±16	65±17	66±24	69±19	72±18	71±23
Bach X ²⁰	155	56±8	63±12	56±9	50±11	60±7	59±8
Kovacevic S ²³	111	63±8	69±7	63±8	63±6	69±7	63±8
Reis R ²⁴	106	65±21	61±19	60±20	55±15	59±18	66±24
Sheng Y ²⁵	93	63±14	63±10	69±12	50±13	50±9	63±13
Da Silva J ²⁶	85	69±8	81±7	63±8	63±6	75±7	75±8
Chandra P ²⁷	82	63±12	63±10	63±8	69±7	69±12	56±14
Imam M ²⁸	82	50±12	50±11	56±10	50±7	50±9	56±13
Souza N ³⁰	58	69±15	69±13	63±14	63±13	63±12	69±16
Zaidi W ³³	27	69±8	63±7	63±8	75±6	69±7	63±8
Subtotal	2.577	60±8	61±8	61±8	59±7	64±7	63±8
WHOQOL-BREF							
Morandi E ¹⁶	365	69±13	69±14	63±16	63±10	NA	NA
Cardona J ¹⁹	187	65±16	67±17	64±27	68±19	NA	NA
Mkangara O ²¹	136	50±11	56±13	56±15	50±10	NA	NA
Sakthong P ²²	120	63±14	63±14	56±16	56±15	NA	NA
Naveet W ²⁹	68	50±11	56±13	56±15	50±10	NA	NA
Yen C ³¹	41	56±8	56±11	56±10	63±9	NA	NA
Allen J ³²	41	46±19	49±20	44±25	55±21	NA	NA
Subtotal	958	62±13	64±14	60±16	59±15		

^a Independencia. ^bEspiritualidad. NA: No aplica.

global y las individuales, demostrando una gran diversidad de perfiles de CVRS y heterogeneidad entre las mediciones. La mejor puntuación se obtuvo en las dimensiones de independencia con 64 puntos y espiritualidad con 63 puntos. De igual manera, en la figura 3 se ilustran dichos estadísticos para las investigaciones que aplicaron el WHOQOL-BREF. En este caso también se hallaron diferencias entre el puntaje global y algunos estudios individuales (lo que ponía de manifiesto la heterogeneidad de las mediciones). Así, en la salud fisi-

ca se obtuvo una puntuación global de 62 puntos, con un rango entre 46 y 69 en los estudios individuales e intervalos de confianza que evidencian la diferencia estadística entre éstos, hallazgo similar al obtenido en las dimensiones de salud psicológica (rango=49-69), relaciones sociales (44-64) y ambiente (50-68).

Finalmente, la comparación entre la puntuación global de los dominios comunes al WHOQOL-HIV-BREF y al WHOQOL-BREF presentó diferencias estadísticamente significativas en la salud física

Tabla 3
Análisis de sensibilidad

Autor	WHOQOL-HIV-BREF											
	Físico		Psíquica		Social		Ambiente		Independencia		Espiritualidad	
	X	Cambio	X	Cambio	X	Cambio	X	Cambio	X	Cambio	X	Cambio
Pereira M ¹⁵	62	2,0* (1,2;2,8)	62	3,0* (2,2;3,7)	62	1,0* (0,2;1,8)	61	2,0* (1,3;2,7)	64	0 (-0,6;0,6)	65	2,0* (1,2;2,8)
Mahalaskshmy T ¹⁷	63	3,0* (2,5;3,4)	61	0 (-0,4;0,4)	61	0 (-0,4;0,4)	58	-1,0* (-1,3;-0,6)	64	-0,3 (-0,7;0,1)	63	0 (-0,4;0,4)
Ucker C ¹⁸	62	2,0* (1,5;2,4)	61	0 (-0,4;0,4)	61	0 (-0,4;0,4)	58	-1 (-1,3;0,6)	64	0 (-0,4;0,4)	61	-1,5* (-1,9;-1,1)
Cardona J ¹⁹	63	3,0* (2,5;3,4)	60	-1,0* (-1,4;-0,5)	61	0 (-0,4;0,4)	58	-1 (-1,3;0,6)	63	-1,0* (-1,3;-0,6)	62	-1,0* (-1,4;-0,5)
Bach X ²⁰	63	3,0* (2,5;3,4)	61	-0,3 (-0,7;0,1)	62	1,0* (0,5;1,4)	59	0 (-0,3; 0,3)	64	0 (-0,4;0,4)	63	0 (-0,4;0,4)
Kovacevic S ²³	62	2,0* (1,5;2,4)	60	-1,0* (-1,4;-0,5)	61	0 (-0,4;0,4)	58	-1,0* (-0,3;-0,6)	64	0 (-0,4;0,4)	62	-1,0* (-1,4;-0,5)
Reis R ²⁴	62	2,0* (1,5;2,4)	61	0 (-0,4;0,4)	61	0 (-0,4;0,4)	59	0 (-0,3;0,3)	64	0 (-0,4;0,4)	62	-1,0* (-1,4;-0,5)
Sheng Y ²⁵	62	2,0* (1,5;2,4)	61	0 (-0,4;0,4)	61	0 (-0,4;0,4)	59	0 (-0,3;0,3)	65	1,0* (0,5;1,4)	62	-1 (-1,4;-0,5)
Da Silva J ²⁶	62	2,0* (1,5;2,4)	60	-1,0* (-1,4;-0,6)	61	0 (-0,4;0,4)	58	-1,0* (-1,3;-0,6)	64	0 (-0,4;0,4)	62	-1,0* (-1,4;-0,5)
Chandra P ²⁷	58	-2,0* (-2,4;1,5)	61	0 (-4,2;4,2)	61	0 (-4,2;4,2)	58	-1,0* (-1,4;-0,6)	64	0 (-0,4;0,4)	63	0 (-0,4;0,4)
Imam M ²⁸	63	3,0* (2,5;3,4)	61	0 (-0,4;0,4)	59	-2,0* (-2,4;-1,5)	59	0 (-0,3;0,3)	64	0 (-0,4;0,4)	63	0 (-0,4;0,4)
Souza N ³⁰	62	3,0* (2,5;3,4)	61	0 (-0,4;0,4)	61	0 (-0,4; 0,4)	58	-1,0* (-1,4;-0,6)	64	0 (-0,4;0,4)	62	-0,6* (-1,0;-0,1)
Zaidi W ³³	62	2,0* (1,5;2,4)	61	0 (-0,4;0,4)	61	0 (-0,4; 0,4)	58	-1,0* (-1,3;-0,6)	64	0 (-0,4;0,4)	63	0 (-0,4;0,4)
WHOQOL-BREF												
Morandi E ¹⁶	57	-5,0* (-6,4;-3,5)	60	-4,0* (-5,5;-2,5)	58	-2,0* (-3,9;-0,1)	58	-1,0 (-2,5;0,5)	NA	NA	NA	NA
Cardona J ¹⁹	61	-1 (-2,2;0,2)	63	-1 (-2,3;0,3)	59	-1 (-2,5;0,5)	55	-4,0* (-5,2;-2,7)	NA	NA	NA	NA
Mkangara O ²¹	64	2,0* (0,7;3,2)	65	1 (-0,3;2,3)	60	0 (-1,6;1,6)	59	0 (-1,3;1,3)	NA	NA	NA	NA
Sakthong P ²²	62	0 (-1,2;1,2)	64	0 (-1,3;1,3)	60	0 (-1,6;1,6)	58	-1,0* (-2,2;-0,2)	NA	NA	NA	NA
Naveet W ²⁹	63		64	0 (-1,3;1,3)	60	-1,0* (-2,6;-0,6)	58	-1,0* (-2,2;-0,2)	NA	NA	NA	NA
Yen C ³¹	62	0 (-1,2;1,2)	64	0 (-1,3;1,3)	60	0 (-1,6;1,6)	57	-2 (-3,2;-0,7)	NA	NA	NA	NA

*Se hallaron cambios estadísticamente significativos en el 0,01. NA: No aplica.

Figura 2
Análisis de la CVRS según las dimensiones del WHOQOL-HIV-BREF

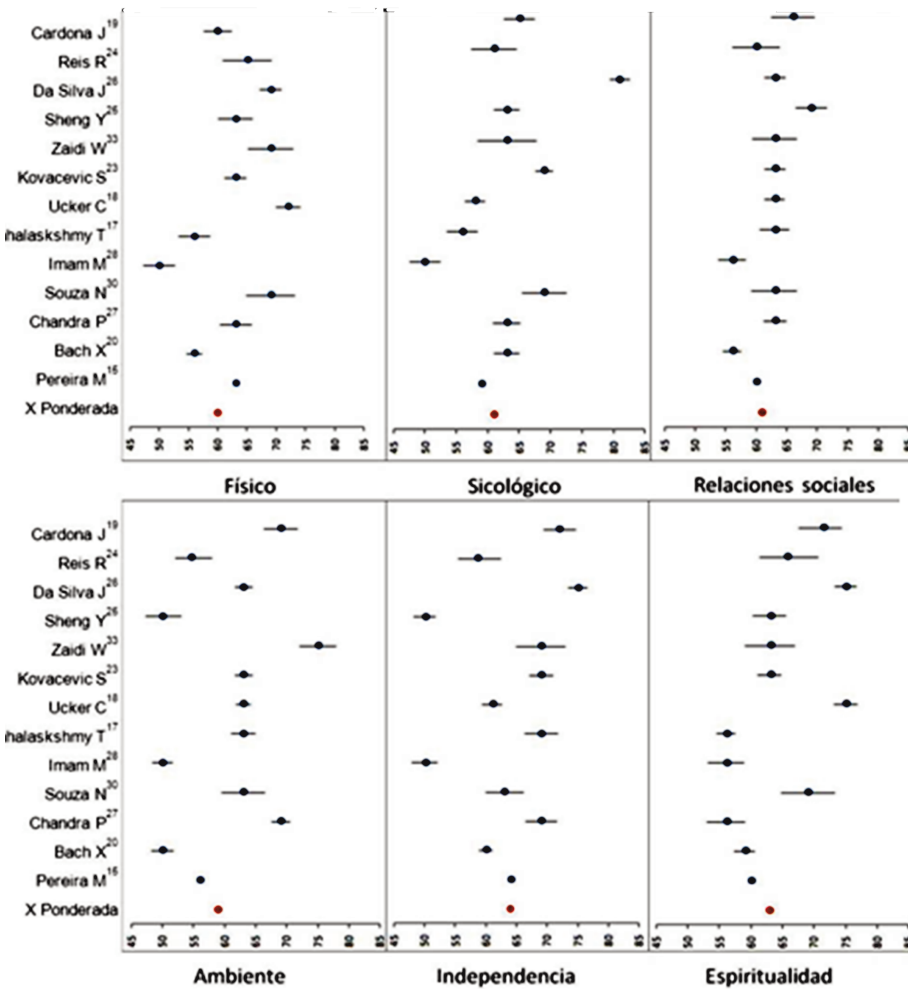
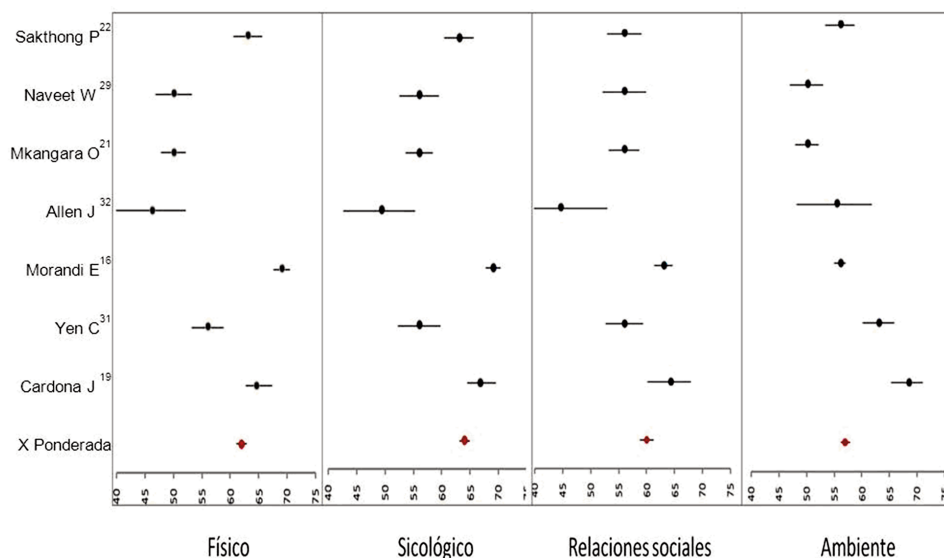


Figura 3
Análisis de la CVRS según las dimensiones del WHOQOL-BREF



con 2 puntos (IC95%=1,1;2,9. $p=0,000$) y la psicológica con 3 puntos (IC95%=2,1;3,9. $p=0,000$), siendo mayor el obtenido con el WHOQOL-BREF; mientras que en relaciones sociales (IC95%=-0,1;2,1. $p=0,064$) y ambiente (IC95%=-1,0;1,0. $p=1,000$) no se hallaron diferencias estadísticamente significativas.

DISCUSIÓN

Con los resultados de este estudio se evidencia que la CVRS en personas con VIH/sida es un tema de interés para los investigadores en el ámbito mundial, ya que se ha estudiado en un número importante de sujetos de 11 países diferentes y con instrumentos tanto genéricos como específicos. Estos hallazgos ponen de manifiesto que en la actualidad la infección por el VIH se considera que trasciende el plano físico-biológico e impacta aspectos económicos, socioculturales, psicológicos y ambientales de quien padece la infección.

En ese contexto, se evidencia la necesidad de ampliar las medidas clásicas del proceso salud-enfermedad que se centran en la morbilidad, la mortalidad y la discapacidad, para abordar instrumentos que den cuenta de la complejidad de la infección y que permitan identificar las dimensiones que condicionan el cumplimiento del tratamiento, la forma de percibir la enfermedad, el desarrollo de enfermedades concomitantes y la progresión de la infección⁹⁻¹².

El estudio evidenció que existe predilección por la evaluación de la CVRS con instrumentos específicos (el WHOQOL-HIV-BREF presentó mayor frecuencia de resultados frente al WHOQOL), debido a que son valorados por su sensibilidad para detectar cambios y particularidades asociadas a la infección³⁴. Sin embargo, en la comparación de los puntajes obtenidos en las cuatro dimensiones comunes a los dos instrumentos evaluados se observó excelente validez convergente/divergente para los dominios

de relaciones sociales y ambiente y buena para la salud física y psicológica. En estas últimas, a pesar de hallar diferencias estadísticamente significativas, no fueron superiores a 3 puntos, lo que resulta bajo, dado el amplio rango de las escalas empleadas (0 a 100).

Lo anterior implica que ambas escalas de CVRS se pueden emplear en VIH/sida y que la elección de una u otra depende de los objetivos del estudio, es decir, si la finalidad es la comparación con personas sanas o comparar el efecto de varias enfermedades sobre la CVRS se debe optar por el WHOQOL-BREF. Mientras que en estudios que busquen la evaluación de los efectos atribuibles a la infección se debe emplear el WHOQOL-HIV-BREF, dado que agrega dos dimensiones de gran significado para los sujetos infectados, el nivel de independencia y la espiritualidad.

En este estudio el puntaje de la espiritualidad fue bueno (65), lo que resulta relevante, ya que otros autores han referido que la espiritualidad impacta la calidad de vida porque mitiga situaciones de estrés, provee ímpetu y esperanza para sobreponerse a las dificultades, induce el autocuidado por la concepción del cuerpo como templo y la participación en cultos grupales genera calma, paz interior, facilita y fortalece las relaciones sociales³⁵.

El dominio de independencia subsume la movilidad, la capacidad de trabajo, la dependencia de los medicamentos y las actividades de la vida diaria. Algunos autores sostienen que ésta es una dimensión fundamental para las personas que viven con VIH/sida porque en estadios avanzados de la enfermedad presentan altas frecuencias de morbilidad, lenta respuesta al tratamiento y requieren cuidadores, lo que coincide con los hallazgos de este estudio, ya que los sujetos que presentaron puntajes más bajos en este dominio ($50 \pm 8,8$) eran personas sintomáticas

según la clasificación o que habían desarrollado sida^{25,28}.

El dominio de la salud física del WHOQOL-BREF evalúa 7 puntos que incluyen los 4 del WHOQOL-HIV-BREF que conforman el dominio de independencia, más 3 adicionales, energía, dolor y sueño. Los puntajes en ambos instrumentos fueron buenos, lo que reafirma la concepción que el acceso oportuno a los servicios de salud y a la terapia anti-retroviral de gran actividad (TARGA) hacen del VIH/sida una infección crónica manejable³⁶.

En el dominio psicológico se obtuvo un puntaje elevado con ambos instrumentos. Algunas investigaciones sugieren que los puntajes de este dominio están relacionados con el tiempo de diagnóstico de la infección debido a que se ha descrito que en las etapas que circundan el diagnóstico se presentan sentimientos negativos con mucha frecuencia, como shock, miedo, ideas suicidas y negación. Sin embargo, en las fases posteriores al diagnóstico asumen su identidad como sujetos infectados, participan en comunidades de personas que viven con VIH, realizan actividades de integración social y superan creencias erróneas sobre la infección, lo que mejora la percepción de la calidad de vida³⁷.

En las relaciones sociales el promedio ponderado no superó los 61 puntos. En este dominio se encontraron investigaciones con el menor puntaje como el estudio de Allen J *et al.* (44 ± 20), lo que podría explicarse por el estigma y la discriminación, ya que el VIH/sida generalmente se desarrolla en contextos donde prima la connotación moral y/o religiosa y al comunicar el diagnóstico se deteriora la interacción con los miembros de la familia, se presentan rupturas abruptas de la relación con los amigos, se afecta la relación con los compañeros de trabajo e incluso se presenta la pérdida del empleo,

lo que deriva en abandono y alteración de la funcionalidad familiar y social de los sujetos infectados³⁸.

El dominio ambiental presentó el puntaje global más bajo entre las dimensiones de los dos instrumentos, este resultado supone que los aspectos financieros se reflejan en este dominio debido a que en algunos casos la infección deriva en abandono de la actividad económica, inhabilidad para encontrar empleo y aumento de los gastos de prestación de servicios de salud^{38,39}.

Entre las principales limitaciones de este estudio destaca que no fue posible realizar análisis bivariado entre la CVRS y aspectos importantes de la infección, como el recuento de linfocitos CD4, el tiempo de diagnóstico, el acceso y adherencia al tratamiento antirretroviral, el soporte social y la estabilidad económica, dado que la mayoría de estudios no reportaban dichas variables. No se evaluó la heterogeneidad u homogeneidad entre los estudios individuales con pruebas como Dersimonian y Laird's, el gráfico de Galbraith o el gráfico de L'Abbé, dado que la naturaleza de los datos de las investigaciones originales no lo posibilitaron. Algo similar ocurrió con el sesgo de publicación, el cual no se pudo establecer a través del *funnel plot*.

Finalmente, se concluye que los puntajes de CVRS en personas con VIH/sida fueron superiores a 50, lo que implica que su adecuado manejo puede disminuir los efectos negativos sobre la salud física, psicológica, social, ambiental, el nivel de independencia y la espiritualidad. Además, se reafirmó la concepción de la infección como una enfermedad crónica tratable que se debe abordar desde un enfoque multidimensional y que la CVRS es de interés mundial para el seguimiento de las personas que viven con VIH/sida.

BIBLIOGRAFÍA

1. Programa conjunto de las Naciones Unidas sobre el VIH/SIDA (ONUSIDA). Informe mundial: informe de ONUSIDA sobre la epidemia mundial de SIDA 2012. Ginebra: ONUSIDA; 2012. [Citado 21 de Marzo de 2013]. Disponible en: <http://www.unaids.org/es/>.
2. Ministerio de salud y protección social República de Colombia. Informe mundial de avances en la lucha contra el SIDA. Bogotá: Ministerio de salud y protección social; 2012. [Citado 21 de Marzo de 2013]. Disponible en: <http://www.unaids.org>.
3. Programa conjunto de las Naciones Unidas sobre el VIH/SIDA (ONUSIDA). Reducir el estigma y la discriminación por el VIH: una parte fundamental de los programas nacionales del sida. Ginebra: ONUSIDA; 2008. [Citado 21 de Marzo de 2013]. Disponible en: http://data.unaids.org/pub/Report/2009/jc1521_stigmatisation_es.pdf.
4. Yen CF, Tsai JJ, Lu PL, Chen YH, Chen TC, Chen PP, et al. Quality of life and its correlates in HIV/AIDS male outpatients receiving highly active antiretroviral therapy in Taiwan. *Psychiatry Clin Neurosci*. 2004; 58(5):501-506.
5. Cardona AD, Agudelo HB. Construcción cultural del concepto calidad de vida. *Revista Facultad Nacional de Salud Pública*. 2005; 23(1): 79-90.
6. Alonso J. La medida de la calidad de vida relacionada con la salud en la investigación y la práctica clínica. *Gac sanit*. 2000; 14(2): 163-167.
7. World Health Organization. WHOQOL-BREF Introduction, administration, scoring and generic version of the assessment. Geneva: World Health Organization; 1996.
8. Fang C, Hsiung P, Yu C, Chen M, Wang J. Validation of the World Health Organization quality of life instrument in patients with HIV infection. *Quality of Life Research*. 2002;11(8): 753-762.
9. Phaladze NA, Dlamini SB, Hulela EB, Hadebe IM, Sukati NA, Makoe LN, et al. Quality of life and the concept of living well with HIV/AIDS in Sub-Saharan Africa. *J NursScholarsh*. 2005; 37(2):120-126.
10. Espinosa O. Apuntes sobre calidad de vida, desarrollo sostenible y sociedad de consumo: una mirada desde América Latina. *Contribuciones (Buenos Aires)*. 1999; 16:119-148.

11. Heckman TG, Somlai EM, Sikkema KJ, Stephen JK. Psychosocial predictors of life satisfaction among persons living with HIV infection and AIDS. *J Assoc Nurses AIDS Care*. 1997;8(5):21-30.
12. Webb A, Norton M. Clinical assessment of symptom focused health-related quality of life in HIV/AIDS. *J Assoc Nurses AIDS Care*. 2004; 15(2):67-81.
13. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0. The Cochrane Collaboration, 2011. [Citado 10 de enero de 2013]. Disponible en: www.cochrane-handbook.org.
14. Medina F, Galván M. Estudios estadísticos y prospectivos. Imputación de datos: Teoría y práctica. División Estadística y proyecciones Económicas Naciones Unidas. CEPAL. 2007 [Citado 4 de Mayo de 2013] Disponible en: www.eclac.org/publicaciones/xml/9/29949/LCL2772e.pdf
15. Pereira M, Canavarro M. Gender and age differences in quality of life and the impact of psychopathological symptoms among HIV-infected patients. *AIDS Behav*. 2011; 15:1857-1869
16. Morandi E, Junior I, Lopes F. Quality of life of people living with HIV/AIDS in Saopaulo, Brazil. *Rev Saúde Pública*. 2007; 41(Supl.2): 64-71.
17. Mahalakshmy T, Premarajan KC, Hamide A. Quality of life and its determinants in people living with human immunodeficiency virus infection in Puducherry, India. *Indian J Community Med*. 2011;36(3):203-207.
18. Ucker CP, Campo MM, Tiellet NM. Qualidade de vida e bem-estar espiritual em pessoas vivendo com HIV/AIDS. *Psicologia em Estudo*, Maringá. 2008; 13(3): 523-530.
19. Cardona AJ. Calidad de vida relacionada con la salud en personas con VIH/SIDA: Comparación del MOSSF-36, WHOQOL-BREF y WHOQOL-HIV-BREF, Medellín, Colombia, 2009. *Colombia Médica*. 2011; 42(4): 438-447.
20. Bach X, Ohinmaa A, Nguyen LT, Oosterhoff P, Vu PX, Vu TV, et al. Gender differences in quality of life outcomes of HIV/AIDS treatment in the latent feminization of HIV epidemics in Vietnam. *AIDS Care*. 2012;24(10):1187-1196.
21. Mkangara OB, Wang C, Xiang H, Xu Y, Nie S, Liu L, et al. The univariate and bivariate impact of HIV/AIDS on the quality of life: a cross sectional study in the Hubei Province-Central China. *J Huazhong Univ Sci Technolog Med Sci*. 2009; 29(2):260-264.
22. Sakthong P, Schommer JC, Gross CR, Prasithsirikul W, Sakulbumrungsil R. Health utilities in patients with HIV/AIDS in Thailand. *Value Health*. 2009; 12(2):377-384.
23. Kovacevic SB, Vurusic T, Duvancic K, Macek M. Quality of life of HIV-infected persons in Croatia. *CollAntropol*. 2006; 30(suppl 2): 79-84.
24. Reis RK, Santos CB, Gir E. Quality of life among Brazilian women living with HIV/AIDS. *AIDS Care*. 2012; 24(5): 626-634.
25. Sheng Y, Qui Z-Q, He Y, Juniper N, Zhang Y. Living conditions and palliative care needs among end-of-life former commercial plasma donors affected with HIV/AIDS in rural Henan of China. *Biomedical and environmental sciences*. 2010; 23(4): 279-286.
26. Da Silva, Bunn K, Bertoni R, Neves O, Traebert J. Quality of life of people living with HIV. *AIDS Care*. 2013;25(1):71-76.
27. Chandra PS, Gandhi C, Satishchandra P, Kamat A, Desai A, et al. Quality of life in HIV subtype C infection among asymptomatic subjects and its association with CD4 counts and viral loads – a study from South India. *Quality of Life Research*. 2006; 15(10): 1597-1605.
28. Imam MH, Karim MR, Ferdous C, Akhter S. Health related quality of life among the people living with HIV. *Bangladesh Med Res Counc Bull*. 2011; 37(1):1-6
29. Naveet W, Raja L, Hemraj P, Vivek A, Chander M, Sunil K. The impact of HIV/AIDS on the quality of life: A cross sectional study in north India. *Indian J MedSci*. 2006; 60(1): 3-12.
30. Souza NLA, Marin CSR, Karina RR, dos Santos CB, Gir E. Aids e tuberculose: a coinfeccao vista pela perspectiva da qualidade de vida dos individuos. *RevEscEnferm USP*. 2012; 46(3): 704-710.
31. Yen CF, Tsai JJ, Lu PL, Chen YH, Chen TC, Chen PP, et al. Quality of life and its correlates in HIV/AIDS male outpatients receiving highly active antiretroviral therapy in Taiwan. *Psychiatry ClinNeurosci*. 2004; 58(5):501-506.
32. Allen J, Hamilton A, Nunn R, Crock L, Frecker J, Burn N. Evaluation of a community-based mental health drug and alcohol nurse in the care of people living with HIV/AIDS. *J PsychiatrMent Health Nurs*. 2009; 16(2): 129-136.

33. Zaidi WAAW, BaharudinAq, Jamalut Y, Nor BM, Zulkapli FH, Hanapiah FA. Quality of Life amongst HIV Positive Patients in Malaysian HIV Shelter Homes.Procedia - Social and BehavioralSciences. 2012; 35:213-220.
34. Velarde E, Avila C. Consideraciones metodológicas para evaluar la calidad de vida. SaludPública Mex. 2002; 44(5): 448-463.
35. Holt C, McClure S. Perceptions of the religion-health connection among African American church members. Qualitative health research 2006; 16(2):268-281.
36. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. N Engl J Med. 1998; 338(13): 853-60.
37. Baumgartner LM. The incorporation of the HIV/AIDS identity into the self over time.Qualitative health research 2007; 17(7):919-931.
38. Varas-Díaz N, Serrano-García I, Toro-Alfonso J. AIDS-related stigma and social interaction: Puerto Ricans living with HIV/AIDS. QualHealth Res. 2005; 15(2): 169-87.
39. ONUSIDA. La Epidemia de SIDA, Situación en diciembre de 2007. Ginebra: Programa conjunto de las Naciones Unidas sobre el VIH/SIDA; 2007.

REVISIÓN SISTEMÁTICA Y METAANÁLISIS

METAANÁLISIS DE COMPARACIÓN DE GRUPOS Y METAANÁLISIS DE GENERALIZACIÓN DE LA FIABILIDAD DEL CUESTIONARIO *STATE-TRAIT ANXIETY INVENTORY* (STAI)

Alejandro Guillén-Riquelme y Gualberto Bucla-Casal.

Departamento Personalidad, evaluación y tratamiento psicológico. Universidad de Granada.

RESUMEN

Fundamentos: Desde su creación, el STAI se ha citado en más de 14.000 documentos, contando con más de 60 adaptaciones en diversos países. En algunas de ellas este cuestionario no cuenta con puntuaciones clínicas. El objetivo de este trabajo es determinar si el cuestionario *State-Trait Anxiety Inventory* (STAI) tiene puntuaciones superiores en personas diagnosticadas de ansiedad respecto a la población general. Además, se pretende analizar si la consistencia interna es adecuada en personas con ansiedad.

Método: Se realizó una búsqueda bibliográfica en Tripdatabase, Cochrane, Web of Knowledge, Scopus, PsycINFO y Scholar Google de documentos publicados entre 2008 y 2012. Se seleccionaron 131 artículos para la comparación entre pacientes diagnosticados de ansiedad respecto a la población general y 25 para la generalización de la fiabilidad. En los análisis se utilizó la *d* de Cohen para la comparación de medias (método de efectos aleatorios) y para la generalización de la fiabilidad el alfa de Cronbach (método de efectos fijos).

Resultados: En la comparación entre grupos, en la ansiedad estado ($d = 1,39$; IC95%: 1,22-1,56) y en la ansiedad rasgo ($d = 1,74$; IC95%: 1,56-1,91) las diferencias fueron estadísticamente significativas. Además de ello, la fiabilidad para pacientes con algún trastorno de ansiedad abarcó de 0,87 a 0,93.

Conclusiones: El STAI es un cuestionario sensible para medir el nivel de ansiedad y fiable en personas diagnosticadas de crisis de angustia, fobia específica, fobia social, fobia social generalizada, trastorno de ansiedad generalizada, trastorno de estrés post-traumático, trastorno obsesivo compulsivo o trastorno por estrés agudo.

Palabras clave: Ansiedad. Metaanálisis. Fiabilidad.

Correspondencia

Alejandro Guillén-Riquelme

Centro de Investigación Mente, Cerebro y Comportamiento-CIMCYC

Universidad de Granada

Campus de Cartuja s/n

18011 Granada.

agr@ugr.es

ABSTRACT

Meta-Analysis of Group Comparison and Meta-Analysis of Reliability Generalization of the *State-Trait Anxiety Inventory* Questionnaire (STAI)

Background: Since its creation the STAI has been cited in more than 14,000 documents, with more than 60 adaptations in different countries. In some adaptations this instrument has no clinical scores. The aim of this work is to determine if the State-Trait Anxiety Inventory (STAI) has higher scores in patients diagnosed with anxiety than in general population. In addition, we want to examine if the internal consistency is adequate in anxious patient samples.

Methods: We performed a literature search in Tripdatabase, Cochrane, Web of Knowledge, Scopus, PsycINFO and Scholar Google, for documents published between 2008 y 2012. We selected 131 scientific articles to compare between patients diagnosed with anxiety and general population, and 25 for the generalization of reliability. For the analysis we used Cohen's *d* for means comparisons (random-effects method) and Cronbach's alpha for the reliability generalization (fixed-effects method).

Results: In the groups comparison the differences in state anxiety ($d=1.39$; CI95%: 1.22-1.56) and in the trait anxiety ($d=1.74$; CI95%:1.56-1.91) were significant. The reliability for patients of some anxiety disorder was between 0.87 and 0.93.

Conclusions: So it seems that the STAI is sensitive to the level of anxiety of the individual and reliable for patients with diagnosis of panic attack, specific phobia, social phobia, generalized social phobia, generalized anxiety disorder, post-traumatic stress disorder, obsessive compulsive disorder or acute Stress disorder.

Keywords: Anxiety. Meta-analysis. Reliability.

INTRODUCCIÓN

Los trastornos de ansiedad son uno de los problemas psicológicos con mayor prevalencia. En una muestra de más de 9.000 estadounidenses se observaron tasas de prevalencia que alcanzaban el 15,6% en el caso de la fobia específica¹. Estos resultados coinciden con medidas epidemiológicas previas². Además, pese a que existen diferencias entre los continentes, la tasa de diagnósticos de ansiedad es muy elevada³⁻⁴.

Entre los distintos instrumentos disponibles para evaluar la ansiedad general se encuentra el cuestionario *State-Trait Anxiety Inventory* (STAI)⁵ que evalúa dos facetas de la ansiedad. En primer lugar, la ansiedad rasgo, entendida como los factores personales que influyen en la percepción del nivel de ansiedad de los estímulos que se presentan. Es decir, una persona con alta ansiedad rasgo percibirá un estímulo como más amenazante y ansiógeno que una persona con baja ansiedad rasgo. Por el contrario, la ansiedad estado se corresponde con los estímulos desencadenantes de ansiedad que hay en el entorno cercano del sujeto. La ansiedad rasgo y estado son teóricamente independientes⁵. Tras su creación se modificaron algunas preguntas y la escala de respuesta del STAI original, llamando a la nueva versión forma Y y nombrando desde entonces a la versión original como forma X⁶. En lo que respecta a su consistencia interna, el alfa de Cronbach medio para 52 estudios incluidos en un metaanálisis fue de 0,91 en ansiedad estado y para un metaanálisis con 51 artículos fue de 0,89 para la ansiedad rasgo⁷.

El STAI es uno de los cuestionarios más empleados, contando con más de 60 adaptaciones culturales y lingüísticas y con más de 14.000 citas en artículos de sus versiones para adultos⁷. Además, el STAI ha sido utilizado para medir el nivel de ansiedad en pacientes con diversos trastornos físicos y psicológicos⁸. Por ejemplo, el STAI se utilizó en un metaanálisis para evaluar el mejor tra-

tamiento para el trastorno de ansiedad generalizada, empleando la subescala de rasgo como principal medida de eficacia del tratamiento⁹. En dicho metaanálisis se incluyeron seis estudios en los que se empleó el STAI con esta muestra. Otro de los ejemplos, en los que se hace uso del STAI para evaluar los niveles de ansiedad en personas que sufren alguno de los trastornos de ansiedad es el artículo de Kvaal y cols¹⁰. En este estudio se aplicó el STAI a personas ancianas con diversos trastornos de ansiedad, hallando que las que tenían trastornos de ansiedad puntuaron significativamente más que los que no tenían ningún trastorno psicológico.

Por último, en un metaanálisis en el que se incluyeron casi 200 estudios se concluyó que el STAI estado es una medida sensible del nivel de ansiedad en grupos clínicos, así como en personas con diversos trastornos físicos¹¹. Además de ello, este instrumento es lo suficientemente preciso como para detectar diferencias en las puntuaciones de sujetos al ser sometidos a estímulos estresantes o procedimientos para aumentar sus niveles de ansiedad¹¹.

Los objetivos de este trabajo son en primer lugar analizar si el STAI detecta las diferencias entre personas con diferentes trastornos de ansiedad y sujetos sanos. En segundo lugar, determinar si la consistencia interna del STAI en muestras de personas con ansiedad es adecuada y si está influida por la forma del STAI (X o Y), por el país o por el tipo de trastorno de ansiedad diagnosticado.

MATERIAL Y MÉTODOS

Tipo de estudio. Con el fin de llevar a cabo los objetivos propuestos en este estudio, clasificado como un estudio teórico de revisión con metaanálisis¹³, se siguieron las recomendaciones PRISMA, ya que este documento presenta una guía para este tipo de artículos¹⁴ y para su redacción se adoptaron las propuestas presentadas por Hartely¹⁵.

Búsqueda de los estudios. En primer lugar se localizaron revisiones sistemáticas o metaanálisis mediante las opciones de búsqueda avanzada de las bases de datos en línea Tripdatabase y Cochrane. Tras no identificar ninguna revisión previa que emplease únicamente el STAI, se pasó a la segunda fase de la búsqueda. Para ello se buscó en diversas fuentes generales y específicas, con el fin de hallar todos los documentos relevantes¹⁶⁻¹⁷. Así pues, se emplearon la Web of Knowledge, Scopus y PsycINFO (a través de la plataforma OVID). Se acotó la búsqueda a los artículos publicados entre los años 2008 y 2012, ya que al ser un cuestionario muy utilizado, la inclusión de cinco años permitía encontrar un número representativo de artículos. Los términos de búsqueda se incluyeron únicamente en inglés.

La ecuación de búsqueda se formó, por una parte, con el nombre completo del STAI así como con el acrónimo y, por otra, con cada uno de los trastornos de ansiedad, indicados en el DSM-IV-TR¹². Se utilizó el DSM-IV-TR en lugar de la nueva versión DSM-5 debido a que en el momento de las búsquedas todavía no estaba publicada y además los trabajos incluidos empleaban la cuarta edición o anteriores. El término STAI se buscó en todo el artículo, mientras que el trastorno de ansiedad únicamente en título, resumen y palabras clave. Se realizó una búsqueda independiente para cada uno de los trastornos. Se incluyeron truncadores siempre que era posible (por ejemplo *phobi** o *obsessi**) e igualmente se utilizó el nombre completo del trastorno y el acrónimo del mismo en el caso de que éste fuese de uso común (por ejemplo PTSD o GAD). La última fase consistió en la búsqueda en Scholar Google con el fin de localizar la literatura gris y evitar un sesgo de publicación. Para realizar esta búsqueda se identificó el artículo original del STAI y, a través de las citas del mismo, se filtraron los años de interés y se verificaron los artículos resultantes. Este proceso se realizó desde el 21 de julio al 30 de septiembre de 2012. Tras descartar los

artículos que se podían excluir por título, en total se encontraron 674 referencias.

Tras la búsqueda, se trató de localizar el texto completo de los artículos. Para ello, se empleó su versión electrónica (incluyendo manuscritos publicados) mediante préstamo interbibliotecario. Para los artículos a los que no se pudo tener acceso se escribió un correo electrónico al autor, donde se informaba del objetivo del estudio y se solicitaba una copia del artículo. Los trabajos no obtenidos fueron descartados (3,12% respecto al total de artículos).

Criterios de selección de los estudios.

Una vez seleccionados los artículos, para determinar su inclusión definitiva se emplearon diversos criterios explícitos¹⁸: a) debían estar escritos en inglés, español o portugués, incluyendo muestras de cualquier país y publicados en cualquier región geográfica; b) únicamente se incluyeron artículos cuasi-experimentales y experimentales, excluyendo revisiones teóricas y estudios de caso único. Las revisiones sistemáticas y los metaanálisis también se incluyeron inicialmente, no encontrando ningún estudio de este tipo. En las citas de estos trabajos se buscaron artículos que reuniesen los criterios para ser incluidos en este estudio; c) en lo que respecta a la muestra sólo se seleccionaron aquellos estudios realizados con personas mayores de edad cuyo diagnóstico principal fuese uno de los trastornos de ansiedad establecidos en el DSM-IV-TR¹² (crisis de pánico, agorafobia, trastorno de angustia sin agorafobia, trastorno de angustia con agorafobia, agorafobia sin historia de trastorno de angustia, fobia específica, fobia social, trastorno obsesivo-compulsivo, trastorno por estrés postraumático, trastorno por estrés agudo y trastorno de ansiedad generalizada; excluyendo los trastornos de ansiedad debido a sustancia y a causas médica e incluyendo la fobia social generalizada). Además, dicho diagnóstico no podía estar justificado mediante las puntuaciones de STAI como único criterio y d)

se descartaron los artículos donde los problemas ansiosos de las personas de la muestra fuesen secundarios a causas médicas.

Respecto a la versión aplicada del STAI, en los artículos seleccionados se podía emplear la versión completa o una de las dos subescalas (estado o rasgo) indistintamente. De la misma forma, la versión del STAI aplicada podía ser tanto la original como una adaptación, siempre que hubiese sido publicada y respetase el número de ítems y la escala de respuesta se mantuviese en cuatro alternativas. En diversas adaptaciones del STAI se recodifica la escala original (de 1 a 4) a una nueva escala de 0 a 3. Tal y como se indica en la literatura⁶, para realizar comparaciones entre diferentes adaptaciones del STAI o con la versión original se debe sumar 20 a la puntuación obtenida, en cada una de las subescalas. Así pues, en los casos en los que se empleaban adaptaciones con este tipo de escala de respuesta se realizó esta transformación.

Por otra parte, también se incluyeron algunos criterios específicos para cada metaanálisis. Para la generalización de la fiabilidad se seleccionaron artículos en los que se incluye un grupo clínico sin importar o no la existencia de un grupo control. Mientras en el caso de la comparación entre grupos únicamente se seleccionaron artículos con la existencia de un grupo control. Si los grupos se establecían mediante puntuaciones de un cuestionario de algún trastorno de ansiedad estos artículos eran descartados si el grupo control estaba formado únicamente por personas con puntuaciones bajas en dicha medida. En los artículos que informaban de un rango del valor de alfa (para varios trastornos de ansiedad sin especificar un valor para cada uno de ellos) se tomó el menor de los valores. En el caso de contar con medias pre y post-tratamiento siempre se seleccionó la media pre-tratamiento de ambos grupos. Además de todo ello, se excluyó un artículo en el que se preguntaba a familiares y amigos de suicidas la puntuación que el fallecido habría puesto en el STAI, y otra investiga-

ción cuyas puntuaciones medias eran superiores a 80 (puntuación teórica máxima de la escala). En la figura 1 se muestra un resumen del proceso de selección de artículos.

En los estudios en los que se emplearon varias muestras clínicas la inclusión fue diferente en función de la finalidad del análisis (generalización de la fiabilidad o comparación entre grupos). En la generalización de la fiabilidad se incluyeron los trastornos de forma independiente, ya que al tratarse de muestras independientes no implicaba un incumplimiento de los supuestos del metaanálisis. En el caso de la comparación de medias no se podía optar por esta solución, ya que al mantenerse el mismo grupo control habría que duplicarlo y el cálculo del efecto de este grupo no cumpliría la independencia de las medidas. En este caso se optó por eliminar el grupo de pacientes de ansiedad con menor número de sujetos.

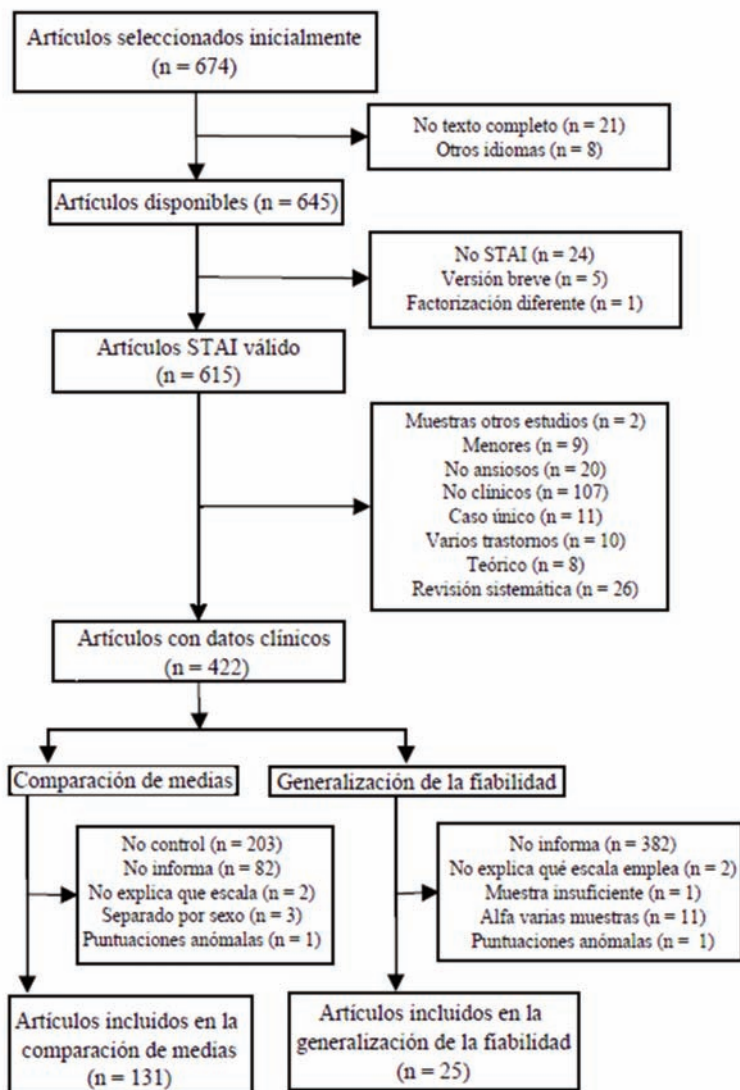
Índices del tamaño del efecto. Para el análisis de la comparación de medias se utilizó la *d* de Cohen y para la generalización de la fiabilidad el alfa de Cronbach.

Extracción de datos. De cada uno de los artículos seleccionados se extrajeron las siguientes medidas para efectuar la media en el metaanálisis:

- Medidas *d* de Cohen para el tamaño del efecto en las comparaciones entre grupos. En el caso de no informar del mismo, se extrajeron de los valores *t*-student o *F* y los grados de libertad, para poder estimarlo. En la mayoría de los casos se emplearon las medias, las desviaciones típicas o varianzas y los tamaños de cada grupo, con el fin de estimar el tamaño del efecto para cada uno de los estudios.

- En el caso de la generalización de la fiabilidad se extrajeron los valores del alfa de Cronbach en todos los casos, ya que no se hallaron estudios que informasen de otras medidas similares para el cálculo de la consistencia interna.

Figura 1
Algoritmo de selección de los artículos



Además de las variables para evaluar el efecto, también se incluyeron diversos moderadores para ver su influencia sobre una posible homogeneidad de la generalización de la fiabilidad o del tamaño del efecto. En primer lugar, se empleó el país donde se seleccionó a los participantes. En segundo

lugar se codificó la forma del STAI utilizada, pudiendo ser la forma X o Y. No se incluyó el uso de adaptaciones a los diferentes países (usando sólo forma X o Y) ya que estadísticamente se analizó en la variable del país de la muestra. Por último, se empleó el tipo de trastorno de ansiedad estudiado,

empleando la codificación del DSM-IV-TR¹². Como moderadores el tamaño muestral de los estudios se extrajo la edad media de los participantes y su distribución por sexo (codificada como porcentaje de mujeres en la muestra). Las diferentes medidas fueron extraídas por dos codificadores independientes con el fin de valorar la fiabilidad del proceso de codificación. Para ello se empleó el índice Kappa y el coeficiente de correlación intraclase, siendo el valor medio de todos los índices 0,844.

Análisis estadístico. Respecto a los análisis, en el caso de la comparación de grupos, para seleccionar un método de estimación de efectos fijos o aleatorios no se empleó ningún estadístico¹⁹. Así pues, para el presente estudio, se puede asumir variabilidad entre cada uno de los tamaños del efecto debido a que proceden de poblaciones diferentes¹⁹. Esta asunción se basa en que los participantes proceden de diferentes países, por tanto han cumplimentado diferentes adaptaciones del STAI e, igualmente, cada estudio se conforma por un grupo clínico, incluyendo diferentes trastornos de ansiedad en la base final. Por ello se optó por utilizar un método de efectos aleatorios. La medida del tamaño del efecto, seleccionada para realizar el metaanálisis, fue la *d* de Cohen²⁰. El método seleccionado es el DerSimonian-Laird²¹ que se considera adecuado para modelos de efectos aleatorios con una medida del tamaño del efecto continua como es la *d* de Cohen²². Se empleó el estadístico *Q* para comprobar la homogeneidad de los datos y se complementó con el índice I^2 ¹⁸. En el caso de hallar heterogeneidad en los resultados y determinar la influencia de diversos moderadores sobre la misma se empleó ANOVAs para las variables cualitativas y se aplicó un modelo de regresión para las variables moderadoras continuas.

Para la generalización del alfa de Cronbach, se transformó este estadístico a valores *T* mediante la fórmula: $T=(1-\alpha)^{1/3}$, donde α es el coeficiente obtenido en cada estudio.

Estos valores permiten transformar el alfa de Cronbach en una medida con distribución normal, requisito necesario en los metaanálisis²³. Además, estas puntuaciones se ponderaron por el inverso de la varianzas de los estudios para calcular el tamaño medio de las puntuaciones *T*. Tras obtenerlo, los valores finales se volvieron a convertir en puntuaciones α con el fin de lograr un índice más fácilmente interpretable. El método fue de efectos fijos, ya que al contar con un número reducido de artículos no cabe asumir el cumplimiento de los supuestos necesarios para utilizar un modelo de efectos aleatorios. Para la realización de ambos metaanálisis se empleó el programa R, concretamente el paquete Metafor²⁴.

RESULTADOS

El número de participantes en total fue de 2.476 (mínimo 6 y máximo 122) en el grupo clínico y 3.780 (mínimo 7 y máximo 1.329) en el grupo control en el caso de la ansiedad estado. De 3.104 (mínimo 6 y máximo 122) en el grupo clínico y 4.513 (mínimo 7 y máximo 1.324) en el grupo control, para la ansiedad rasgo. De 1.182 (mínimo 20 y máximo 406) en el grupo clínico de generalización de la fiabilidad estado y 1.631 (mínimo 20 y máximo 406) en la fiabilidad de la ansiedad rasgo. La media de edad osciló entre 30,78 y 40,94 en el conjunto de los cuatro metaanálisis. La media de ansiedad estado fue de 44,94 y en ansiedad rasgo de 50,26. En el caso de la fiabilidad, el α medio sin transformar fue 0,9 para la ansiedad estado y 0,88 para ansiedad rasgo.

Con el test de Egger hubo indicios de un posible sesgo en el caso de la ansiedad rasgo ($Z=7,03$, $p<0,001$), pero no en la ansiedad estado ($Z=1,76$, $p=0,08$). Así pues, en el caso de la ansiedad estado se observó que el tamaño medio del efecto fue de $d=1,39$ (IC95%: 1,22-1,56) resultando estadísticamente significativo ($p<0,001$). Tras ello, se analizó la homogeneidad de los tamaños del efecto con el fin de determinar si había variabilidad entre los estudios. El test de

heterogeneidad resultó estadísticamente significativo ($Q(93)=634,51$ $p<0,001$). El indicador I^2 fue igual a 85,34%, lo que implicó una heterogeneidad alta.

En el caso de la ansiedad rasgo el tamaño del efecto medio fue de $d=1,74$ (IC95%: 1,56-1,91) siendo esta medida estadísticamente significativa ($p<0,001$). En las pruebas de heterogeneidad se observó que había variabilidad en las aportaciones de los diferentes estudios a la media ($Q(109)=910,52$; $p<0,001$). Este resultado se confirmó mediante el resto de estadísticos empleados ($I^2=88,03\%$).

En cuanto a los resultados de los AOVAs, en lo que respecta a ansiedad estado, el país ($Q_m(17)=49,02$; $p<0,001$; $Q_e(76)=419,08$; $p<0,001$) y la forma del STAI ($Q_m(1)=7,61$; $p=0,006$; $Q_e(91)=596,45$; $p<0,001$) resultaron estadísticamente significativos, mientras que el tipo de trastorno no ($Q_m(7)=6,48$; $p=0,49$; $Q_e(86)=602,57$; $p<0,001$). En el caso de la ansiedad rasgo el país ($Q_m(18)=22,69$; $p=0,2$; $Q_e(91)=742,25$; $p<0,001$) y la forma del STAI ($Q_m(1)=0,85$; $p=0,356$; $Q_e(107)=902,61$; $p<0,001$) no fueron estadísticamente significativos, al contrario que el tipo de trastorno ($Q_m(7)=6,72$; $p<0,001$; $Q_e(102)=176,39$; $p<0,001$). Seguidamente, se analizaron los niveles en los que se producía la variabilidad. En la tabla 1 se recoge un resumen de los coeficientes obtenidos para cada una de las categorías de dichas variables. Además se emplearon moderadores continuos: el tamaño muestra, la media en la puntuación del STAI y su desviación típica, la edad media y el porcentaje de mujeres en la muestra. De las variables continuas empleadas únicamente se observó capacidad predictiva de la variabilidad en la media para la ansiedad estado ($b=0,08$ [CI95%: 0,06-0,11]; $Q=44,01$, $p<0,001$; $Q_e=507,65$, $p<0,001$; $R^2=0,22$), y de la media ($b=0,09$ [CI95%: 0,08-0,11]; $Q_m=106,54$, $p<0,001$; $Q_e=510,46$,

$p<0,001$; $R^2=0,5$) y la desviación típica en el caso de la ansiedad rasgo ($b=-0,10$ [CI95%: -0,16-0,14]; $Q_m=9,89$, $p=0,002$; $Q_e=827,42$; $p<0,001$; $R^2=0,09$).

El segundo tipo de metaanálisis incluido fue el de generalización de la fiabilidad. En este caso se obtuvieron 13 artículos para la escala de ansiedad estado y 19 artículos para la ansiedad rasgo, en los que se informa del α de Cronbach. De nuevo se comenzó por el análisis de la influencia de un posible sesgo de selección, observando problemas tanto en estado ($Z=4,93$, $p<0,001$) como en rasgo ($Z=6,09$, $p<0,001$). En los artículos sobre ansiedad estado el valor medio de α calculado sin ningún tipo de ponderación fue de 0,87 (DT=0,08); de 0,89 (DT=0,05) para la ansiedad rasgo.

Seguidamente, se calculó la media de las puntuaciones T ponderadas, observando un valor de 0,43 en ansiedad estado (EE=0,006) y de 0,45 (EE=0,005) para la ansiedad rasgo. Al transformar esta puntuación de nuevo a valores α de Cronbach, se obtuvo un $\alpha=0,92$ en ansiedad estado (IC95%: 0,91-0,93) y de 0,91 en ansiedad rasgo (IC95%: 0,90-0,92). En ambos casos este valor resultó estadísticamente significativo (ansiedad estado: $Z=71,24$, $p<0,001$; ansiedad rasgo: $Z=83,76$, $p<0,001$). Tras ello, se analizó la homogeneidad de los valores α en los diferentes artículos. En ambos casos se observa heterogeneidad: ansiedad estado: $Q(12)=106,19$ ($p<0,001$) y ansiedad rasgo: $Q(18)=70,86$ ($p<0,001$).

El siguiente paso fue analizar las fuentes de variabilidad. En la ansiedad estado todas las variables analizadas resultaron significativas: país ($Q_m(4)=65,09$; $p<0,001$; $Q_e(8)=41,1$; $p<0,001$), forma del STAI ($Q_m(1)=6,88$; $p=0,009$; $Q_e(11)=99,31$; $p<0,001$) y tipo de trastorno ($Q_m(3)=19,26$; $p=0,001$; $Q_e(9)=86,93$; $p<0,001$). Mientras, en la ansiedad rasgo, el país ($Q_m(6)=40,11$;

Tabla 1
Coeficientes de los moderadores en el meta-ANOVA de comparación de grupos
para la ansiedad estado y ansiedad rasgo

Variables (k estado k rasgo) significativo estado /significativo rasgo	Ansiedad estado				Ansiedad rasgo			
	<i>d</i> de Cohen	Error estándar	Z	Intervalo de confianza 95%	<i>d</i> de Cohen	Error estándar	Z	Intervalo de confianza 95%
País del estudio								
Alemania (17;27)	1,07	0,19	5,76 [†]	0,71-1,44	1,51	0,18	8,55 [†]	1,17-1,86
Australia (6;6)	1,30	0,33	3,98 [†]	0,66-1,95	1,89	0,39	4,85 [†]	1,13-2,66
Canadá (1;1)	3,19	0,85	3,76 [†]	1,53-4,85	0,65	0,94	0,7	—1,18-2,49
China (2;2)	1,38	0,53	2,6 [‡]	0,34-2,42	1,39	0,64	2,18 [§]	0,14-2,63
EEUU (21;29)	1,49	0,17	8,76 [†]	1,15-1,81	2,09	0,17	12,19 [†]	1,75-2,42
España (7;3)	1,02	0,28	3,62 [†]	0,47-1,57	1,14	0,52	2,2 [§]	0,12-2,15
Francia (4;4)	2,53	0,42	6,01 [†]	1,71-3,37	2,71	0,5	5,44 [†]	1,73-3,69
Holanda (2;1)	2,16	0,56	3,88 [†]	1,07-3,25	2,34	0,92	2,55 [§]	0,54-4,13
Israel (1;2)	1,14	0,72	1,6	—0,26-2,55	1,63	0,62	2,62 [‡]	0,41-2,85
Italia (4;3)	0,95	0,41	2,32 [§]	0,15-1,74	0,87	0,54	1,61	—0,19-1,92
Japón (6;7)	1,76	0,32	5,47 [†]	1,13-2,40	1,93	0,35	5,45 [†]	1,24-2,63
Korea (3;3)	2,34	0,47	5,01 [†]	1,43-3,26	1,51	0,53	2,85 [‡]	0,47-2,56
Nueva Zelanda (0;1)	---	---	---	---	3,22	0,94	3,43 [†]	1,38-5,06
Países Bajos (2;2)	1,62	0,55	2,96 [‡]	0,55-2,69	1,16	0,64	1,8	—0,1-2,42
Polonia (1;1)	1,57	0,77	2,05 [§]	0,07-3,07	1,52	0,91	1,66	—0,27-3,3
Suecia (1;2)	1,16	0,76	1,54	—0,32-2,64	1,91	0,66	2,89 [‡]	0,62-3,2
Suiza (1;1)	5,01	0,82	6,13 [†]	3,41-6,61	4,57	0,94	4,84 [†]	2,72-6,41
Turquía (3;4)	0,54	0,44	1,23	—0,32-1,41	2,08	0,47	4,41 [†]	1,15-3,0
UK (11;13)	1,6	0,23	6,87 [†]	1,14-2,05	2,01	0,26	7,72 [†]	1,5-2,52
Forma del STAI								
Forma Y (39;50)	1,78	0,14	12,56 [†]	1,51-2,06	2,01	0,14	14,46 [†]	1,74-2,28
Forma X (53;61)	1,24	0,12	10,28 [†]	1,01-1,48	1,71	0,13	13,66 [†]	1,47-1,96
Trastorno psicológico								
Crisis de angustia (16;16)	1,54	0,23	6,67 [†]	1,09-1,99	1,62	0,21	7,66 [†]	1,2-2,03
Fobia específica (11;17)	1,18	0,28	4,25 [†]	0,64-1,73	1,05	0,21	5,06 [†]	0,64-1,46
Fobia social (16;19)	1,89	0,23	8,07 [†]	1,43-2,35	2,33	0,2	11,56 [†]	1,93-2,72
Fobia social generalizada (4;4)	1,32	0,46	2,9 [‡]	0,43-2,21	1,97	0,44	4,49 [†]	1,11-2,83
Ansiedad generalizada (2;12)	0,79	0,63	1,25	—0,44-2,02	2,76	0,26	10,71 [†]	2,26-3,27
Estrés post-traumático (16;19)	1,59	0,23	6,99 [†]	1,15-2,04	1,79	0,2	9,08 [†]	1,4-2,17
Obsesivo-compulsivo (26;23)	1,33	0,18	7,48 [†]	0,98-1,68	1,84	0,18	10,22 [†]	1,49-2,19
Trastorno por estrés agudo (2;2)	1,25	0,63	2 [§]	0,03-2,48	0,61	0,58	1,05	—0,53-1,75

Z=puntuación Z del coeficiente; [†] p <0,001; [‡] p <0,01; [§] p <0,05. ^{||} la d de Cohen media para dicha categoría es estadísticamente diferente de la d de Cohen media total.

Tabla 2
Coefficientes de los moderadores en el meta-ANOVA de generalización de la fiabilidad para la ansiedad estado y rasgo

Variables (k estado k rasgo) significativo estado/significativo rasgo	Ansiedad estado				Ansiedad rasgo			
	α^*	Error estándar	Z	Intervalo de confianza 95%	α^*	Error estándar	Z	Intervalo de confianza 95%
País del estudio								
Alemania (-;3)	---	---	---	---	0,91	0,02	21,65 [†]	0,89-0,94
Canadá (2;1)	0,89	0,01	44,72 [†]	0,87-0,90	0,94	0,01	41,59 [†]	0,93-0,95
EEUU (7;10)	0,94	0,01	47,5 [†]	0,93-0,95	0,91	0,01	55,87 [†]	0,9-0,92
Francia (-;2)	---	---	---	---	0,87	0,02	24,08 [†]	0,83-0,90
Holanda (1;1) /	0,93	0,02	22,89 [†]	0,91-0,95	0,81	0,03	22,89 [†]	0,76-0,85
Italia (1;1)	0,83	0,04	13,17 [†]	0,74-0,90	0,92	0,03	13,17 [†]	0,88-0,95
Turquía (-;1)	---	---	---	---	0,92	0,03	14,7 [†]	0,88-0,95
Forma del STAI								
Forma Y (7;13)	0,9	0,03	15,33 [†]	0,87-0,94	0,9	0,02	24,89 [†]	0,88-0,93
Forma X (4;6)	0,93	0,04	11,24 [†]	0,88-0,96	0,89	0,03	17,32 [†]	0,85-0,92
Trastorno psicológico								
Fobia específica (0;3)	---	---	---	---	0,93	0,04	11,35 [†]	0,88-0,96
Fobia social (0;2)	---	---	---	---	0,9	0,05	10,49 [†]	0,82-0,95
Fobia social generalizada (0;1)	---	---	---	---	0,84	0,06	8,65 [†]	0,70-0,93
Obsesivo-compulsivo (2;5)	0,95	0,05	7,12 [†]	0,90-0,98	0,9	0,03	16,1 [†]	0,86-0,93
Estrés post-traumático (6;5)	0,9	0,03	15,05 [†]	0,84-0,93	0,91	0,03	15,71 [†]	0,88-0,95
Ansiedad generalizada (2;1)	0,91	0,05	8,3 [†]	0,83-0,96	0,84	0,06	8,62 [†]	0,70-0,93
Varios trastornos (1;2) /**	0,93	0,07	5,72 [†]	0,83-0,98	0,88	0,04	11,91 [†]	0,8-0,93

*Los coeficientes han sido retransformados mediante $\alpha=1-T_3$. Z: puntuación Z del coeficiente; [†] $p<0,001$; ** Incluye artículos en los que se emplean varios trastornos juntos siempre que éstos sean exclusivamente trastornos de ansiedad. ^{||} el α medio para dicha categoría es estadísticamente diferente del α medio total.

$p<0,001$; $Q_e(12)=30,76$; $p=0,002$) y el tipo de trastorno ($Q_m(5)=33,68$; $p<0,001$; $Q_e(13)=37,18$; $p<0,001$) resultan significativas pero no la forma del STAI ($Q_m(1)=0,1$; $p=0,751$; $Q_e(17)=70,76$; $p<0,001$). Un resumen de los estadísticos por cada nivel de las variables se recoge en la tabla 2. Para los moderadores continuos, debido al bajo número de estudios, los resultados fueron estadísticamente significativos pese a lo cual las betas eran inferiores a 0,01, salvo en el caso de la desviación típica como moderador de la variabilidad en la fiabilidad de la ansiedad rasgo ($b=-0,03$ [CI95%: -0,4--0,02]; $Q_m=31,62$, $p<0,001$; $Q_e=38,55$, $p<0,001$).

DISCUSIÓN

Tras el análisis se observa que, en ambos casos, ansiedad estado y rasgo, los tamaños del efecto medio son elevados. Así pues, el STAI es sensible a los niveles de ansiedad clínicos, detectando puntuaciones significativamente más elevadas que en un grupo control con sujetos sanos. En cuanto a la fiabilidad media de la escala para pacientes diagnosticados de ansiedad los valores también son elevados, observando una fiabilidad media superior a 0,9 en ambas subescalas. Por ello, se puede concluir que el STAI es un instrumento fiable a la hora de evaluar los niveles de ansiedad en pacientes con

ansiedad y que es sensible a la hora de detectar los niveles superiores de ansiedad (estado y rasgo) de los mismos, a través de sus diferentes formas y en diversos países. Además de ello, el STAI es sensible a la ansiedad provocada por diversos trastornos ansiosos.

Al realizar el metaanálisis sobre el efecto medio se ha observado variabilidad de las medidas a lo largo de los estudios. A la hora de explicar la variación de los tamaños del efecto hay diferencias entre países. No obstante, los tamaños del efecto son adecuados (abarcando desde 0,44 a 5,01 en ansiedad estado y de 0,54 a 4,57 en ansiedad rasgo). Además, salvo en los artículos de Alemania, Italia y Turquía, todos los tamaños de efecto medios son superiores a 1. Así pues el STAI detecta diferencias entre los sujetos con ansiedad y los sujetos sanos tal y como cabría esperar^{3,5}. En lo que respecta a la forma del STAI, en ambos casos los tamaños del efecto son similares, pese a que la forma X tiene un tamaño del efecto medio menor que la forma Y, tanto en ansiedad estado como en ansiedad rasgo. Lo que indica que ambas formas son adecuadas para detectar los niveles de ansiedad, pudiéndose utilizar el STAI para la evaluación de la ansiedad⁸⁻⁹. El último moderador de la variabilidad empleado fue el tipo de trastorno, mostrando que el STAI permite detectar ansiedad general de forma adecuada con puntuaciones mayores entre diversos trastornos, respecto a población general, siendo las diferencias estadísticamente significativas.

La generalización de la fiabilidad permite inferir si un test es fiable a lo largo de muestras procedentes de diversas poblaciones²³. En este caso se observa que la fiabilidad media es elevada. Además, al analizar por países, la fiabilidad media es adecuada. Al estimar la fiabilidad media por cada uno de los trastornos los valores superan el 0,86 en todos los casos. Estos resultados son coherentes ya que se encuentra un notable número de artículos en los que se afirma que el

STAI tiene una elevada fiabilidad¹¹ y un metaanálisis en el que se establecen valores medios de alfa de Cronbach de 0,91 en ansiedad estado y de 0,89 en ansiedad rasgo⁷. Así pues en población general la fiabilidad es similar a la obtenida, pese a que en el trastorno de ansiedad generalizada en el caso de la subescala de ansiedad rasgo el alfa sea ligeramente inferior. Por último, al analizar la forma del STAI se observa que la forma Y tiene 0,91 para ambas escalas, mientras que la forma X tiene una fiabilidad de 0,93 para ansiedad estado y de 0,91 para ansiedad rasgo. Generalmente, se afirma que la versión Y del STAI tiene una mayor fiabilidad⁵⁻⁶. El hecho de que en este metaanálisis se observe lo contrario puede deberse a que el tamaño de las muestras de sujetos con diagnóstico de ansiedad suele ser menor que el de las muestras de población general, lo que puede afectar a la fiabilidad.

En lo que respecta a España se observa que hay diversos artículos en los que se emplea una muestra clínica, obteniendo un tamaño del efecto medio elevado. Así pues, la adaptación española del STAI detecta puntuaciones superiores en sujetos con ansiedad. Pese a ello, dicha adaptación no cuenta con una muestra de personas con diagnóstico de ansiedad ni de baremos específicos para ellas²⁵. De la misma forma no hay fiabilidad para sujetos con ansiedad en la adaptación española. De hecho, en la revisión realizada no se ha encontrado ningún artículo con muestra española que informe de la fiabilidad. Pese a ello, en población general la fiabilidad del STAI es adecuada²⁶.

En resumen, el STAI parece ser un instrumento fiable en muestras formadas por sujetos ansiosos detectando mayores niveles de ansiedad en estas respecto a muestras control. Por ello, sería conveniente que las diversas adaptaciones del STAI tuviesen baremos con personas con ansiedad así como versiones breves del mismo donde la fiabilidad varía²⁷.

Como limitación del estudio cabe destacar la posible influencia de otros moderadores que expliquen la variabilidad hallada en los tamaños del efecto y los coeficientes de fiabilidad medios. Otra limitación es el hecho de que pese a haber incluido búsquedas que permitiesen localizar la literatura gris, en los resultados se observa un sesgo de selección, por lo que sería interesante replicar este metaanálisis incluyendo nuevas bases y búsquedas. Pese a ello, este estudio permite valorar la fiabilidad media del STAI, uno de los más utilizados en la práctica psicológica.

BIBLIOGRAFÍA

1. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res.* 2012; 21: 169-184. <http://dx.doi.org/10.1002/mpr.1359>
2. Kessler RC, Ruscio AM, Shear K, Wittchen HU. Epidemiology of Anxiety Disorders. En: *Behavioral Neurobiology of Anxiety and Its Treatment*. Libro electrónico, 2009, http://dx.doi.org/10.1007/7854_2009_9
3. Somers JM, Goldner EM, Waraich P, Hsu L. Prevalence and Incidence Studies of Anxiety Disorders: A Systematic Review of the Literature. *Can J Psychiatry.* 2006;51:100-113.
4. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet.* 2007;370:841-850.
5. Spielberger CD, Gorsuch R, Lushene R. Manual for the State-Trait Anxiety Inventory. Palo Alto: Consulting Psychologist Press; 1970.
6. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. State-Trait Anxiety Inventory (Form Y). Palo Alto: Consulting Psychologists Press; 1983.
7. Barnes LLB, Harp D, Jung WS. Reliability generalization of scores on the Spielberger State-Trait anxiety Inventory. *Educ Psychol Meas.* 2002;62:603-618.
8. Spielberger CD, Reheier EC. Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Appl Psychol Health Well Being.* 2009;1:271-302. <http://dx.doi.org/10.1111/j.1758-0854.2009.01017.x>.
9. Fisher PL, Durham RC. Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychol Med.* 1999;29:1425-1434.
10. Kvaal K, Ulstein I, Nordhus IH, Engedal K. The Spielberger State-Trait Anxiety Inventory (STAI): the state scale in detecting mental disorders in geriatric patients. *Int J Geriatr Psychiatry.* 2005;20:629-634.
11. Rossi V, Pourtois G. Transient state-dependent fluctuations in anxiety measured using STAI, POMS, PANAS or VAS: a comparative review. *Anxiety Stress Coping.* 2012; 25: 603-645. <http://dx.doi.org/10.1080/10615806.2011.582948>
12. American Psychological Association. Manual diagnóstico y estadístico de los trastornos mentales. DSM-IV-TR. Barcelona: Masson; 2002.
13. Montero I, León OG. A guide for naming research studies in Psychology. *Int J Clin Health Psychol.* 2007;7:847-862.
14. Urrútia G, Bonfill X. La Declaración PRISMA: un paso adelante en la mejora de las publicaciones de la Revista Española de Salud Pública. *Rev Esp Salud Pública.* 2013;87: 99-102.
15. Hartley J. New ways of making academic articles easier to read. *Int J Clin Health Psychol.* 2012;12:143-160.
16. Perestelo-Pérez L. Standards on how to develop and report systematic reviews in Psychology and Health. *Int J Clin Health Psychol.* 2013;13:49-57.
17. Fernández-Ríos L, Buela-Casal G. Standards for the preparation and writing of Psychology review articles. *Int J Clin Health Psychol.* 2009;9:329-344.
18. Sánchez-Meca J. Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta.* 2010;38:53-64.
19. Sánchez-Meca J, Marín Martínez F, Huedo Medina TB. Modelos de efectos fijos y modelo de efectos aleatorios. En: *Revisiones sistemáticas en las ciencias de la vida*. Castilla la Mancha: Fundación para la Investigación Sanitaria en Castilla-La Mancha; 2006. p. 189-204.
20. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1969.
21. DerSimonian R. y Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7:177-188.

22. Espallargues i Carreras M, Tebé Cordoní C. Qué resultados combinar y cómo combinarlos. En: Revisiones sistemáticas en las ciencias de la vida. Castilla la Mancha: Fundación para la Investigación Sanitaria en Castilla-La Mancha; 2006. p. 173-187
23. Sánchez-Meca J, López-Pina JA. El enfoque meta-analítico de generalización de la fiabilidad. *Acción Psicológica*. 2008;31:7-17.
24. Viechtbauer, W. metafor: Meta-Analysis Package for R." R package version 1.4-0. 2010. <http://CRAN.R-project.org/package=metafor>.
25. Buéla-Casal G, Guillén-Riquelme A, Seisdedos Cubero N. Cuestionario de Ansiedad Estado-Rasgo: Adaptación española. 8ª ed. Madrid: TEA Ediciones; 2011.
26. Guillén-Riquelme A, Buéla-Casal G. Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory. *Psicothema*. 2011;23:510-515.
27. Guillén-Riquelme A, Buéla-Casal G. Versión Breve del STAI en Adolescentes y Universitarios Españoles. *Ter Psicol*. 2013;31:293-299.

REVISIÓN SISTEMÁTICA

MODELOS DE COLABORACIÓN ENTRE ATENCIÓN PRIMARIA Y SALUD MENTAL EN LA ASISTENCIA SANITARIA A LAS PERSONAS CON DEPRESIÓN: RESULTADOS PRINCIPALES Y RETOS METODOLÓGICOS DE UNA META-REVISIÓN SISTEMÁTICA (*)

Carlos Calderón Gómez (1,2), Isabel Mosquera Metcalfe (2), Laura Balagué Gea (3,2), Ander Retolaza Balsategui (4), Amaia Bacigalupe de la Hera (5), Jon Belaunzaran Mendizabal (6), Álvaro Iruin Sanz (7), y Beatriz García Moratalla (8).

(1) Centro de Salud de Alza. Comarca Gipuzkoa. Osakidetza.

(2) Unidad de Investigación de Atención Primaria-OSIs Gipuzkoa. Osakidetza.

(3) Centro de Salud de Iztietia. Comarca Gipuzkoa. Osakidetza.

(4) Centro de Salud Mental de Basauri. Red de Salud Mental de Bizkaia. Osakidetza.

(5) Departamento de Sociología 2. Universidad del País Vasco/Euskal Herriko Unibertsitatea Bizkaia.

(6) Centro de Salud Mental de Zarautz. Red de Salud Mental de Gipuzkoa. Osakidetza.

(7) Red de Salud Mental de Gipuzkoa. Osakidetza.

(8) Centro de Salud Mental de Sestao. Red de Salud Mental de Bizkaia. Osakidetza.

(*) Beca de Investigación Comisionada 2010. Departamento de Salud. Gobierno Vasco. N.º Expediente 2010COM112007. Beca BIOEF 2009. Departamento de Salud. Gobierno Vasco. N.º Expediente BIO09/EM/014.

No existe conflicto de interés por parte de ninguno de los autores.

RESUMEN

Fundamentos: Las carencias en la colaboración entre Atención Primaria (AP) y Salud Mental (SM) constituyen un problema relevante en la atención a los pacientes con depresión. Resulta necesario analizar y evaluar los modelos de colaboración existentes para valorar su aplicabilidad en el sistema de salud español. El objetivo del presente estudio es conocer las principales características de los distintos modelos de colaboración AP-SM en la atención a los pacientes diagnosticados de depresión y la calidad de la evidencia científica acerca de su efectividad.

Métodos: Meta-revisión sistemática de los estudios secundarios publicados entre 2001 y 2010 en MEDLINE, PsycINFO, Embase, LILACS, IBECs, IME y la Biblioteca Cochrane. Las revisiones se evaluaron mediante la herramienta AMSTAR. Se realizó una síntesis aproximativa de la calidad de las evidencias encontradas.

Resultados: Se evaluaron 69 estudios. La variabilidad según contextos y las carencias metodológicas condicionan que la calidad de las evidencias sea en general baja o dudosa. Las estrategias más efectivas integran intervenciones de responsabilización en el seguimiento de los pacientes, rediseños en la gestión, e información y comunicación compartidas. Las meta-revisiones de estudios secundarios sobre modelos colaborativos favorecen la accesibilidad a las evidencias publicadas, pero conllevan importantes retos metodológicos.

Conclusiones: La calidad de la evidencia sobre la efectividad de los modelos de colaboración AP-SM durante la atención sanitaria a las personas con depresión es predominantemente baja o dudosa y su significado y aplicabilidad son menores cuanto más se simplifica el análisis de sus componentes, procesos y circunstancias de implementación.

Palabras clave: Revisión sistemática. Meta-revisión. Depresión. Atención primaria de salud. Salud mental. Continuidad de la atención al paciente. Organización y administración. Planificación en salud.

Correspondencia

Carlos Calderón Gómez.
Centro de Salud de Alza
Avda. de Larratxo s/n
20017-Donostia-San Sebastián
ccalderong@telefonica.net

ABSTRACT

Models for Primary Care and Mental Health Collaboration in the Care of People with Depression: Main Results and Methodological Challenges of a Systematic Overview

Background: Weaknesses in the collaboration between Primary Care (PC) and Mental Health (MH) are a relevant problem in the care of depressed patients. It is necessary to analyse and appraise the existing models of collaboration to assess their applicability to the Spanish Health System. The aim of this study is to know the main characteristics of the different models of collaboration between PC and MH in the care of patients with depression and the quality of their effectiveness evidence.

Methods: Systematic overview of secondary studies published from 2001 to 2010 in MEDLINE, PsycINFO, Embase, LILACS, IBECs, IME and The Cochrane Library. Assessment of reviews applying the AMSTAR tool. Approximative synthesis of the quality of evidences.

Results: A total of 69 studies were assessed. Quality of evidences is generally low or inconclusive due to the great variability among contexts and the methodological weaknesses. The most effective strategies integrate interventions for assigning responsibility for patient follow-up, redesigning management and communication/information sharing. Overviews of secondary studies on collaborative models facilitate access to published evidence, but entail important methodological challenges.

Conclusion: The quality of evidences on effectiveness of PC-MH collaboration models in depression care is mainly low or inconclusive, and the more simplified are the analysis of components, processes and implementation conditions, the less meaningful and applicable they are.

Keyword: Systematic review. Overview. Depression. Primary Health Care. Mental Health. Continuity of Patient Care. Organization and Administration. Health planning.

INTRODUCCIÓN

Entre los problemas detectados en la asistencia a las personas con depresión, las carencias en la relación entre los profesionales de Atención Primaria (AP) y Salud Mental (SM) configuran un campo de especial importancia, sobre todo si se considera el carácter crónico de la enfermedad o, al menos, muy prolongado en el tiempo¹. Tanto a nivel español como internacional se ha subrayado la importancia de un enfoque coordinado y complementario para este grupo de trastornos^{2,4}, constatándose en los últimos años un auge en la publicación de experiencias colaborativas y propuestas a favor de su difusión^{5,6}. De las intervenciones orientadas principalmente a la formación de los profesionales de AP y a los canales de derivación a SM se ha pasado a la promoción de diferentes modelos de colaboración (consulta-enlace, gestión de casos y de enfermedades, programas de mejora de la calidad, *Chronic Care Model*, etcétera) que incorporan nuevas intervenciones y figuras como la difusión de Guías de Práctica Clínica (GPC), cuestionarios de cribado, telepsiquiatría, o “gestores de casos”⁷⁻⁹. Sin embargo, estos modelos e intervenciones no han sido suficientemente evaluados^{10,11} ni son bien conocidos en nuestro medio. Los problemas de la colaboración entre AP y SM participan asimismo de las necesidades de mejora que afectan en general a las funciones y vínculos de la AP con los otros niveles asistenciales de atención especializada y hospitalaria. De hecho, algunos de los modelos e intervenciones comentadas han comenzado ya a implantarse en nuestros servicios de salud como respuesta al incremento de la multimorbilidad y las enfermedades crónicas¹². Por nuestra parte, hemos venido trabajando en los últimos años con profesionales de AP y SM y con pacientes diagnosticados de depresión con la intención de diseñar y ensayar experiencias piloto para mejorar la colaboración entre los distintos agentes y niveles asistenciales^{4,13,14}. Para ello se consideró necesario llevar a cabo un análisis

y evaluación de las propuestas colaborativas existentes en el ámbito de los trastornos depresivos mediante una meta-revisión sistemática de los estudios secundarios publicados al respecto, cuya versión completa y pormenorizada puede consultarse en el informe publicado en Osteba-Servicio de Evaluación de Tecnologías Sanitarias¹⁵. El presente artículo pretende centrarse en el análisis de los modelos y en los retos metodológicos del proceso de revisión y evaluación de propuestas de mejora complejas, como es este caso, a partir del objetivo de conocer las principales características de los distintos modelos de colaboración AP-SM en la atención a los pacientes diagnosticados de depresión y la calidad de la evidencia publicada acerca de su efectividad.

MATERIAL Y MÉTODOS

El diseño de la revisión respondió a su finalidad pragmática, orientada al contraste y aplicabilidad de los hallazgos en los servicios de salud de nuestro país. En una primera fase de revisión general se incluyeron todos los estudios publicados entre 2001 y 2010 en inglés, español o portugués sobre modelos de coordinación, interrelación, integración y colaboración entre AP y SM en la atención a pacientes con depresión. Se excluyeron los estudios referidos a aspectos específicos del diagnóstico o tratamiento. En una segunda fase se llevó a cabo una meta-revisión más exhaustiva de los estudios secundarios encontrados, es decir, de las revisiones y trabajos de síntesis que reunían y analizaban estudios primarios previamente publicados. Las meta-revisiones de revisiones previas (también denominadas *overviews*) resultan en principio de gran utilidad ante la dificultad de procesar la ingente cantidad de información publicada en determinados ámbitos de la asistencia sanitaria en los que ya se cuenta con revisiones de estudios primarios. Estas meta-revisiones presentan sus propias dificultades y requisitos metodológicos y estadísticos y su todavía limitado desarrollo se ha centrado

principalmente en las meta-revisiones de revisiones sistemáticas de ensayos clínicos previos^{16,17}. Como se verá a continuación, las dificultades se incrementan en el caso de la evaluación de propuestas e intervenciones complejas cuyo desarrollo desborda los requisitos habituales de los ensayos clínicos y metanálisis. En consecuencia, el desarrollo de las dos fases de la revisión no fue lineal. La complejidad del objeto de revisión determinó una dinámica iterativa de búsqueda y consulta de trabajos primarios así como la inclusión en la meta-revisión de estudios secundarios cuyo análisis y evaluación se consideraron asimismo necesarios para la consecución del objetivo del estudio, aun no siendo propiamente revisiones sistemáticas.

Primera fase: Estrategia de búsqueda y revisión general

Se realizó una búsqueda de la literatura en MEDLINE, PsycINFO, Embase, LILACS, IBECs, IME y la Biblioteca Cochrane publicada desde enero de 2001 a diciembre de 2010 y se complementó con búsquedas manuales. Los términos incluyeron “depresión”, “atención primaria”, “psiquiatría”, “colaboración/coordinación”, “co-localización” e “integración”, adecuados a los requisitos e idiomas de las bases consultadas. En la **tabla 1** se exponen, a modo de ejemplo, la estrategias completas de búsqueda electrónica seguidas en el caso de Embase, pudiendo consultarse los términos detallados correspondientes a las 7 bases bibliográficas en el anexo I del Informe Osteba¹⁵. Tras la lectura de los títulos y resúmenes se eliminaron los artículos duplicados y los que no se correspondían con el objeto de la revisión. En un segundo cribado se incorporaron estudios encontrados mediante búsqueda manual y finalmente se seleccionaron los artículos de estudios secundarios para su análisis pormenorizado (**figura 1**).

Segunda fase: meta-revisión de los estudios secundarios

De las 69 publicaciones disponibles a texto completo, se extrajeron y analizaron los siguientes datos: objeto de estudio, contexto de realización, población de estudio y tipo de depresión, características de las revisiones y resultados en relación a los objetivos de la revisión. Ante el limitado desarrollo de los instrumentos de evaluación en las meta-revisiones de intervenciones complejas¹⁶, conscientes de su orientación inicialmente dirigida a las revisiones sistemáticas, se optó por la aplicación de la herramienta AMSTAR¹⁸ para evaluar la calidad metodológica de los 57 artículos correspondientes a revisiones. Esta herramienta recorre las diferentes dimensiones a tener en cuenta en el proceso de evaluación a través de 11 ítems teniendo como resultado máximo 11 puntos. La clasificación de las revisiones se llevó a cabo según las categorías de calidad utilizadas por la *Canadian Agency for Drugs and Technologies in Health* (CADTH) (baja: 0 a 3 puntos; media: 4 a 7 y alta: 8 a 11)¹⁹. Como complemento de lo anterior, se identificaron y agruparon las principales limitaciones metodológicas señaladas explícitamente por los propios autores de los estudios secundarios, para ser tenidas en cuenta en la evaluación de la calidad de las evidencias de los distintos modelos incluidos en la meta-revisión. Por último, tras el análisis pormenorizado de los distintos modelos e intervenciones de colaboración, se llevó a cabo una síntesis aproximativa de las evidencias encontradas. Para ello se elaboraron tablas resumen conforme a cuatro grados de calidad de las evidencias de mejora: elevada, baja, dudosa y no mejora. Los factores que se tuvieron en cuenta para su clasificación fueron: 1) los propios resultados de los estudios secundarios y la duración en su efectividad atendiendo a las características, evolución y contextos de los modelos e intervenciones analizados; 2) la calidad metodológica de las revisiones según la herramienta AMSTAR; y 3) las limitaciones metodológicas de los estudios secundarios identificadas por los autores en la evaluación

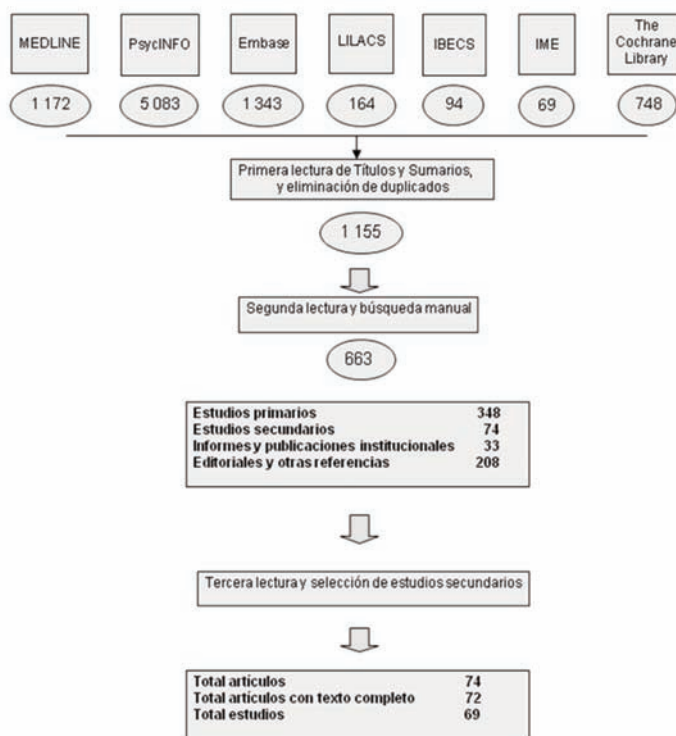
Tabla 1
Estrategias de búsqueda empleadas en Embase vía Ovid (1996 a semana 4 2011)

Depresión + atención primaria + psiquiatría
1. exp depression/dm, pc, th [Disease Management, Prevention, Therapy] 2. exp primary health care/ 3. exp PSYCHIATRY/ 4. 1 and 2 and 3 5. limit 4 to yr="2001 - 2010" 6. limit 5 to (english or portuguese or spanish)
Depresión + colaboración/coordinación
1. exp depression/dm, pc, th [Disease Management, Prevention, Therapy] 2. collaborat*.mp. or collaborative care/ 3. (shared adj3 care).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 4. (stepped adj3 care).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 5. coordina*.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 6. 2 or 3 or 4 or 5 7. 1 and 6 8. limit 7 to yr="2001 - 2010" 9. limit 8 to (english or portuguese or spanish)
Depresión + co-localización
1. exp depression/dm, pc, th [Disease Management, Prevention, Therapy] 2. colocat*.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 3. co?locat*.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 4. 2 or 3 5. 1 and 4 6. limit 5 to yr="2001-2010" 7. limit 6 to (english or portuguese or spanish)
Depresión + integración
1. exp depression/dm, pc, th [Disease Management, Prevention, Therapy] 2. integrated health care system/ or integrat*.mp. 3. 1 and 2 4. limit 3 to yr="2001 - 2010" 5. limit 4 to (english or portuguese or spanish)

de los estudios primarios. La clasificación en las categorías “elevada”, “baja” o “no mejora” requería el apoyo de referencias de mayor solidez metodológica, reservándose la categoría “dudosa” para las de menor calidad o de resultados no concluyentes. Este proceso se llevó a cabo de forma independiente por tres de los investigadores, siendo necesario su pleno acuerdo para la clasificación de los estu-

dios en las diferentes categorías. Posteriormente, todos los investigadores revisaron y discutieron la propuesta, acordándose por todos ellos los resultados alcanzados. El presente artículo se centra en los resultados y retos metodológicos relativos a la revisión de los 6 modelos principales evaluados, pudiendo consultarse el trabajo completo en el informe ya mencionado¹⁵.

Figura 1
Selección de artículos



RESULTADOS

a) Primera fase: revisión general

Las estrategias iniciales de búsqueda identificaron un total de 8.673 referencias, las cuales fueron objeto de sucesivos cribados conforme a los objetivos de la meta-revisión, resultando 74 artículos correspondientes a estudios secundarios (figura 1). En 2 artículos no fue posible acceder a los textos completos^{20,21}. Uno de los estudios secundarios se describía en 3 artículos²²⁻²⁴ por lo que se incluyó el que lo describía de forma más global²² y otra de las revisiones²⁵ incluía y actualizaba una revisión previa de los mismos autores²⁶, por lo que se excluyó la más

antigua. Finalmente, el número total de estudios seleccionados fue de 69.

b) Segunda fase: meta-revisión de los estudios secundarios

b.1.) Características de los estudios

De los 69 estudios, 57 se presentaban como revisiones. De ellas, 23 eran revisiones sistemáticas, 29 no sistemáticas y 5 meta-revisiones. Los otros 12 artículos incluidos en la meta-revisión se correspondían con 3 descripciones sobre la evolución histórica de la interrelación AP-SM en EEUU^{27,28} y Holanda²⁹, 3 análisis descriptivos de modelos de integración AP-SM vigentes a nivel internacional y en

EEUU^{9,30,31}, 1 valoración de iniciativas y proyectos de atención a la depresión en Europa³, 3 descripciones de programas de distintos proyectos en EEUU³²⁻³⁴, 1 evaluación del impacto de 2 reformas de SM primaria en Australia³⁵ y 1 propuesta de organización de la evidencia de la integración de servicios de SM en AP³⁶.

b.2.) Contextos geográficos, poblaciones y trastornos depresivos

En cuanto a los países de origen de los estudios primarios incluidos en los 69 artículos seleccionados, el 88,41 % incluía estudios de EEUU, el 49,28 % de Reino Unido, y el 18,84 % de Australia y Holanda. En menor proporción había estudios de Canadá (15,94 %), Alemania (10,14 %), Chile (7,25 %), Italia (7,25 %), Suecia (5,80 %) y Finlandia (5,80 %). En 26 artículos los estudios primarios eran exclusivamente de EEUU y en otros 7 de EEUU y Reino Unido.

Respecto a la procedencia de los 69 estudios secundarios, en 33 (47,83 %) los autores eran de EEUU, en 11 de Reino Unido (15,94 %), en 4 de Australia (5,80 %), en 4 de Holanda (5,80 %), en 3 de Canadá (4,35 %), en 2 de Italia (2,90 %), en 1 (1,45 %), de Irlanda, de China, de Nueva Zelanda y en 8 (11,59 %) había autores de varios países. En 43 (62,32%) de las revisiones no se explicitaban suficientemente las poblaciones a estudio, 18 (26,09 %) se referían a población adulta en general^{32,37-40}, incluyéndose en 13 de ellas a niños y personas mayores de 60 años^{8,9,41-51} y en 8 (11,59 %) hacían referencia total o principalmente a ancianos^{25,52-58}. En 8 (11,59 %) de los estudios, de EEUU, se trataba de poblaciones seleccionadas, como usuarios de *Health Maintenance Organizations*, *Managed Care Organizations* y centros de veteranos^{37,40,42,52,59,60}, militares³⁴ y fundaciones³³. En 36 (52,17 %) no se especificaba el tipo de trastorno depresivo. En 30 (43,48 %) se aludía a varios tipos de de-

presión y en 5 (7,25 %) se referían explícitamente a depresión mayor^{32,37,38,41,61}.

b.3.) Calidad metodológica de los estudios

La aplicación de la herramienta AMSTAR permitió una primera valoración de la calidad de las revisiones (tabla 2). Su agrupación en función del tipo de revisión muestra la baja puntuación obtenida por 23 de las 29 revisiones no sistemáticas incluidas en el estudio. Hay que considerar además los 12 estudios secundarios no considerados como revisiones^{3,9,27-36} a los que, por tanto, no se pudo aplicar esta herramienta de evaluación.

Asimismo, se identificaron y agruparon las principales limitaciones metodológicas señaladas por los propios autores de los estudios secundarios:

- Dificil comparabilidad de los estudios primarios por sus diferentes contextos y poblaciones^{10,11,27,34,49,58,59,62-64}. Las intervenciones presentes en los modelos de colaboración son plurales y a menudo insuficientemente descritas^{11,36,37,46,54,65}, lo cual dificulta la identificación de los componentes más efectivos^{42,66}. Los problemas en el diseño de las comparaciones⁶⁷ son mayores en los modelos más complejos, como la atención colaborativa^{10,68} o las intervenciones multifaceta en las que participan diferentes grupos profesionales⁶⁷.

- Escasez de estudios aleatorizados controlados debido a la complejidad de las intervenciones^{10,29} y variabilidad en la calidad metodológica de los estudios primarios^{8,41,47,53,66,68}. La atención habitual con la que se compara la efectividad de los modelos e intervenciones con frecuencia es heterogénea⁵¹ o poco detallada^{8,66,69}.

- Dificultades para la utilización de la misma definición de depresión, con utilización de gran variedad de instrumentos diagnósticos^{3,37,38,41,58}.

Tabla 2
Evaluación de la calidad de las revisiones

Calidad metodológica (AMSTAR-CADTH)	Revisiones sistemáticas	Revisiones no sistemáticas	Meta-revisiones
Alta (8-11)	Gilbody 2002 ²² Badamgarav 2003 ⁸⁵ Neumeyer-Gromen 2004 ⁴² Tsai 2005 ⁷² Bower 2006 ¹¹ Gensichen 2006 ³⁷ Gilbody 2006 ⁶⁹ Gunn 2006 ⁴⁷ Smith 2007 ⁶⁶ Williams 2007 ⁴⁹ Butler 2008 ⁵⁰ Christensen 2008a ⁸ Harkness 2009 ⁶⁴ Huang 2009 ⁵⁸ Cape 2010 ⁴⁰ van der Feltz-Cornelis 2010 ⁶⁰	Skultety 2006 ⁵⁴	Rubenstein 2009 ⁵¹
Media (4-7)	Weingarten 2002 ⁶⁵ Vergouwen 2003 ⁶² Bijl 2004 ⁴¹ Craven 2006 ⁴⁶ Smit 2007 ³⁸ Hall 2008 ⁷³ Lu 2008 ⁵⁹	Craven 2002 ¹⁰ Beaudin 2005 ⁴³ Doughty 2006 ⁶⁷ Christensen 2008b ⁷¹ Parrish 2009 ³⁹	Bartels 2002 ⁵³ Bower 2005 ⁶⁸
Baja (0-3)		Anfinson 2001 ⁷⁵ Callahan 2001 ⁵² Bower 2002a ⁹⁸ Bower 2002b ⁸¹ Ford 2002 ⁹⁴ Hilty 2002 ³² Aitken 2004 ³³ Oxman 2005 ²⁵ Adli 2006 ⁴⁴ Byrne 2006 ⁹⁵ Hilty 2006 ⁴⁸ Katon 2006 ¹³⁴ Berardi 2007 ⁶³ Kates 2007 ⁷ Magruder 2007 ⁷⁰ McEvoy 2007 ⁵⁵ Thielke 2007 ⁹⁶ Vannoy 2007 ⁵⁶ Katon 2008 ⁶¹ Skultety 2008 ⁵⁷ Meyer 2009 ⁷⁶ Simon 2009 ⁹⁷ Bauer 2010 ⁸²	Barbui 2006 ⁴⁵ Griffiths 2008 ¹⁰⁸

- Ausencia de consenso en los indicadores utilizados para la evaluación de la efectividad de las intervenciones^{47,54,70}. Necesidad de mejores mediciones, tanto de los procesos como de los resultados³³ y de las revisiones de “literatura gris”⁷¹. Tendencias a reflejar los impactos según las audiencias a las que van destinados³⁶. En ocasiones, falta de coherencia entre los objetivos o metas de los programas y lo que realmente se evalúa¹⁰.

- Dificil generalización de los resultados^{10,70-72}, sobre todo en los modelos que conllevan reorganizaciones asistenciales, de personal y de financiación^{27,30,51} y en los que precisan personal experto^{11,37,42,46,47,49,67,73}.

- Duración inadecuada de los tiempos de seguimiento⁶⁶ con predominio de las evaluaciones a corto plazo^{11,54,69}. Se subrayaba la necesidad de investigar la efectividad a largo plazo^{42,69} o al menos durante 2 o más años^{6,59}.

A pesar de dichas deficiencias y de la baja puntuación de 25 de las revisiones evaluadas con la herramienta AMSTAR (todas ellas no sistemáticas) se optó por su inclusión en el análisis descriptivo de los diferentes modelos e intervenciones, sin dejar de tener en cuenta la calidad de las evidencias sobre su efectividad.

b.4.) Clasificaciones y delimitación de modelos.

La revisión confirmó la diversidad de modelos así como de propuestas para su clasificación, con frecuentes solapamientos entre los distintos modelos y las intervenciones incluidas. Aunque en algunos casos las clasificaciones obedecían a una finalidad descriptiva en países concretos^{27-29,35,63}, en general pretendían ofrecer esquemas comparativos, identificándose tres principales:

- El denominado esquema de los cuatro cuadrantes, en el que el papel de AP y SM se distribuye según la gravedad en las dimensiones física y mental de la salud, a partir del cual se desarrollan diversas tipologías de modelos^{9,30,31,74}.

- Un segundo esquema, más frecuentemente utilizado, sitúa a AP y SM en los extremos de un eje que representa su grado de implicación en la asistencia a los trastornos depresivos: formación de profesionales de AP, consulta-enlace, modelos colaborativos y sustitución-derivación^{10,40,68,75,76}.

- Un tercer esquema distribuye los modelos según las fases en las investigaciones sobre la colaboración AP-SM en depresión, al menos en Norteamérica^{25,28,46}.

Se delimitaron y seleccionaron para su revisión 6 modelos y fueron ordenados de los más simples a los más complejos: atención escalonada, consulta-enlace, gestión de enfermedades y gestión de casos, programas de mejora de calidad (QIC) y *Breakthrough Series* (BTS), *Chronic Care Model* (CCM) y un último grupo denominados genéricamente como modelos de “atención colaborativa”. Si bien el presente artículo se limita a la revisión de los modelos de colaboración, es importante subrayar la inexistencia de una correspondencia cerrada entre estos y las distintas intervenciones. Es decir, modelos diferentes pueden presentar intervenciones comunes, y determinadas intervenciones pueden estar o no presentes en las aplicaciones de un mismo modelo¹⁵.

b.5.) Análisis y evaluación de los modelos de colaboración.

Atención escalonada. Consiste en la disposición de los recursos en niveles de atención en función de la gravedad y evolución de los trastornos depresivos. En general es evaluado como parte de otras propuestas colaborativas más complejas^{7,22,61}. El modelo se corresponde con las recomendaciones

de la GPC del *National Institute for Health and Care Excellence* (NICE)⁷⁷, pero su organización concreta y el tipo de asistencia varían según los sistemas sanitarios^{3,78}. La evidencia sobre su efectividad es limitada. Se sugiere que su integración en modelos más complejos de atención colaborativa podría asociarse con mejores resultados⁷⁷, mayor concordancia con la medicación y mejora en las tasas de recuperación²².

Consulta-enlace. Se enfoca a la vinculación-derivación entre los profesionales de AP y SM con especial atención al componente formativo^{22,79,80}. La responsabilidad en el seguimiento de los pacientes se mantiene principalmente en AP^{68,81}, aunque su práctica varía según las condiciones de acceso, cobertura y organización de los servicios de salud^{45,63,81}. Muestra una evidencia inconsistente en la mejora de la atención a la depresión^{60,67,68,81}. Su inclusión en modelos más complejos (CCM) podría mejorar la efectividad y permitir un uso más eficiente del tiempo del especialista⁸¹. El conocimiento recíproco entre los profesionales de ambos niveles y su grado de motivación parecen ser fundamentales^{81,82}, si bien se señala su falta de impacto como modelo aislado en los resultados finales⁴⁰. Tampoco parecen encontrarse diferencias claras entre este modelo y otros de mayor grado de integración⁵⁰.

Gestión de enfermedades y gestión de casos. Se trata de modelos muy variables según su configuración como programas más o menos integrados. Destaca como componente característico la existencia de un *care manager* o *case manager* responsable de la coordinación y el seguimiento de los pacientes^{83,84}, con gran variabilidad de profesiones, roles y poblaciones implicados^{41,65,85}, lo cual dificulta su evaluación^{8,33,37,42,59}. Los programas de gestión de enfermedades parecen ser efectivos en la atención a la depresión⁴². La inclusión del cribado en el modelo aporta resultados positivos en la depresión mayor en estudios de

EEUU, pero no en los de Reino Unido⁴¹. Los efectos de estos modelos conllevan una pequeña mejoría en los síntomas, satisfacción de los pacientes, aumento de visitas a AP y adecuación y adherencia al tratamiento. La mejoría no resulta significativa en la funcionalidad física, estado de salud, hospitalizaciones, costes, resultados derivados de la adherencia de profesionales y pacientes ni derivación a psiquiatría⁸⁵. Los modelos de gestión de casos se asocian también con mejoras en los resultados si el gestor de casos se vincula directamente a los médicos de AP y participa en actividades de terapia psicológica). El incremento en la adherencia al tratamiento podría intervenir como mecanismo de acción clave en la mejora de los resultados⁸, incluso en plazos de 6-12 meses³⁷.

Programas de mejora de calidad y Breakthrough series. Los denominados *Quality Improvement Collaboratives* (QIC) representan un nivel más complejo de modelos, ya que actúan simultáneamente sobre diferentes dimensiones y agentes en la organización de la asistencia. De ahí que su efectividad se reconozca condicionada por los contextos^{32,43,52} y plazos temporales^{22,38} en que se han llevado a cabo, con gran heterogeneidad de experiencias y ámbitos de aplicación e importantes inversiones de tiempo y recursos^{86,87}. Algunos estudios muestran mejoras a los 6 meses y en algún caso hasta los 5 años^{38,43}, mientras que en otros los resultados favorables se evidencian a los 12 meses para desaparecer a los 2 años²². Las mejoras se vinculan a la combinación de actividades de formación a profesionales y pacientes, gestión de casos por enfermeras y monitorización de la medicación²², es decir, actividades integradas y complejas⁵² cuya evaluación estadística resulta dificultosa³². El modelo *Breakthrough Series* (BTS) del *Institute for Healthcare Improvement* (IHI) representa una propuesta particular de programa de QIC para disminuir la distancia entre lo que se sabe y lo que se hace. Su enfoque de intervención es

Tabla 3
Resumen de la calidad de las evidencias sobre efectividad de los modelos de colaboración

Modelos	Elevada	Baja	Dudosa	No mejora
Atención escalonada		Gilbody 2002 ²²	Kates 2007 ⁷ Katon 2008 ⁶¹	
Consulta-enlace		Butler 2008 ⁵⁰	Bower 2002b ⁸¹ Bower 2005 ⁶⁸ Doughty 2006 ⁶⁷ Bauer 2010 ⁸² van der Feltz–Cornelis 2010 ⁶⁰	Barbui 2006 ⁴⁵ Cape 2010 ⁴⁰
Gestión de enfermedades		Bijl 2004 ⁴¹ (depresión mayor) Neumeyer-Gromen 2004 ⁴²	Weingarten 2002 ⁶⁵ Badamgarav 2003 ⁸⁵ Lu 2008 ⁵⁹	
Gestión de casos		Gensichen 2006 ³⁷ Christensen 2008a ⁸		
Programas de mejora de la calidad BTS		Gilbody 2002 ²² (corto plazo) Beaudin 2005 ⁴³ Smit 2007 ³⁸	Callahan 2001 ⁵² (ancianos) Rost 2001 ³² Kates 2007 ⁷	
CCM		Tsai 2005 ⁷² Williams 2007 ⁴⁹ Butler 2008 ⁵⁰ Rubenstein 2009 ⁵¹	Bower 2002b ⁸¹ Ford 2002 ⁹⁴ Kates 2007 ⁷ Magruder 2007 ⁷⁰ McEvoy 2007 ⁵⁵ Katon 2008 ⁶¹	
Modelos colaborativos	Gilbody 2006 ⁶⁹ (mayoría EEUU)	Anfinson 2001 ⁷⁵ (depresión mayor) Bartels 2002 ⁵³ (ancianos, enfermedades mentales) Gilbody 2002 ²² Vergouwen 2003 ⁶² Bower 2005 ⁶⁸ Craven 2006 ⁴⁶ (depresión mayor) Skultety 2006 ⁵⁴ (ancianos) Smit 2007 ³⁸ (corto plazo) Smith 2007 ⁶⁶ Huang ⁵⁸ (ancianos)	Craven 2002 ¹⁰ Oxman 2005 ²⁵ Adli 2006 ⁴⁴ (ancianos) Barbui 2006 ⁴⁵ Byrne 2006 ⁹⁵ Doughty 2006 ⁶⁷ (ancianos, depresión mayor) Gunn 2006 ⁴⁷ Thielke 2007 ⁹⁶ Vannoy 2007 ⁵⁶ Katon 2008 ⁶¹ Lu 2008 ⁵⁹ (farmacéuticos) Skultety 2008 ⁵⁷ (ancianos, depresión mayor) Meyer 2009 ⁷⁶ (EEUU) Simon 2009 ⁹⁷ Collins 2010 ⁹	

a corto plazo, lo cual conlleva limitaciones ante cambios organizativos de cierta complejidad⁸⁸ y quizás explique lo incierto de las evidencias de su implementación^{7,88-90}.

Chronic care model (CCM). El modelo de atención a la cronicidad (CCM) del IHI ha sido ampliamente difundido como referente de políticas sanitarias nacionales y regionales⁹¹. La simultaneidad de sus 6 elementos esenciales y el rediseño en la gestión de los servicios explican su complejidad y variabilidad según los países, servicios y programas. De ahí también las dificultades en la evaluación del modelo y en su aplicación en contextos diferentes al de EEUU con sistemas de AP más consolidados^{46,55,92,93}. En el ámbito de la depresión los acentos son diferentes en uno u otro componente (gestión de casos, atención escalonada, desarrollo de GPC, tecnología de la información, apoyo y educación de los pacientes)⁷ si bien, en general, la figura del gestor de cuidados (*care manager*), principalmente orientada a la enfermería, tiende a considerarse como uno de sus elementos clave^{49,51}. Se ha sugerido que el rediseño en la gestión y el apoyo al autocontrol por los pacientes tendrían una mayor vinculación con los resultados, mientras que el apoyo en la toma de decisiones afectaría más a los procesos⁷². En cualquier caso, los efectos parecen ser discretos^{72,94}, advirtiéndose mejoras en la adherencia a la medicación⁴⁹, la evolución de los síntomas^{49,50} y la satisfacción de pacientes y médicos de AP⁶¹. No se describen beneficios en pacientes con depresión menor⁴⁹. En algunas revisiones las experiencias de CCM se presentan como muy concretas y locales^{33,94} y de difícil sostenibilidad en condiciones rutinarias^{70,81}, con incrementos en los costes asistenciales⁷⁰.

Modelos de “atención colaborativa”. Finalmente, como modelos de “atención colaborativa” se ha agrupado un amplio número de estudios secundarios que incor-

poran modelos complejos de colaboración^{11,56,76,95} que incluyen intervenciones simultáneas en gran medida comunes a las descritas en los QIC y sobre todo en el CCM^{45,61,68}, pero sin identificarse explícitamente como tales. Resulta llamativa la diversidad en sus definiciones e intervenciones, a menudo insuficientemente descritas y evaluadas³⁶. Ello conlleva interrogantes acerca del papel de los distintos componentes⁹⁶ y de la gran variedad de agencias de servicios y sistemas de financiación involucrados, particularmente en el caso de EEUU⁹. La mayoría de las revisiones muestran su mayor efectividad con mejoras a 6^{38,58,69} y 12 meses⁵⁸, en particular en depresión mayor^{44,62} a través de una mayor toma de medicación⁵⁸ y mejora en la adherencia⁶². Si bien se señalan evidencias de beneficio incluso hasta los 5 años^{61,69}, los resultados no son consistentes en todos los estudios^{47,66} y en general se reconocen modestos. También se constatan efectos positivos en el estado de salud⁶⁸, el grado de satisfacción de pacientes^{10,61,68} y médicos de AP^{10,61}, la adherencia a la medicación^{61,62,68} y la reducción de ideas suicidas en ancianos⁵⁸. En ocasiones los resultados favorables se sustentan en referencias a otros modelos como “gestión de enfermedades” integrables en la atención colaborativa^{45,68}. Según los estudios, los beneficios se relacionan con la experiencia profesional en SM del gestor de casos y su labor de supervisión^{25,69}, la existencia de relaciones clínicas previas, GPC compartidas y cambios organizativos⁴⁶, potenciación de la enfermería²² e incorporación de recursos adicionales^{22,62}, desarrollo de equipos multidisciplinarios^{53,54}, en algún caso con inclusión de farmacéuticos⁵⁹, y más intervenciones psicosociales, principalmente en ancianos⁵⁴. Los efectos de los modelos colaborativos resultan inferiores y más ambiguos en la depresión menor^{57,67}. No parece existir un claro incremento en los costes^{58,97} y se advierte una mayor eficiencia si se combina con la atención escalonada⁷⁵. Los estudios colaborativos realizados fuera de EEUU son

más heterogéneos y no parecen aportar resultados favorables significativos⁶⁹. Las explicaciones se han vinculado a las diferencias en la fortaleza de los sistemas de AP⁴⁷, en particular en el caso de Reino Unido⁹⁸ y al hecho de que los estudios en EEUU se han dirigido principalmente a grupos de riesgo, como pacientes con comorbilidad física o psiquiátrica⁶⁹.

b.6.) Síntesis aproximativa de evidencias.

A partir de los resultados previamente analizados, se elaboró una síntesis aproximativa de la calidad de las evidencias encontradas en relación a la efectividad de los diferentes modelos (tabla 3). Tal como se avanzó en el apartado de metodología, en su ordenación se establecieron cuatro niveles (“elevada”, “baja”, “dudosa” y “no mejora”) según la fortaleza de la evidencia en los efectos a medio y largo plazo y los condicionantes metodológicos en cada caso. El 92,3% (24/26) de las referencias clasificadas como “elevada”, “baja” o “no mejora” correspondieron a revisiones de calidad alta y media conforme a la aplicación de la herramienta AMSTAR y únicamente 2 (7,6%) revisiones fueron no sistemáticas. En la categoría “dudosa” solo 7 (20,6%) de las 34 referencias fueron revisiones sistemáticas o meta-revisiones de calidad alta y media, mientras que de las 27 restantes 24 (88,9%) eran no sistemáticas y 21 (87,5%) de 24 de calidad baja. La gran heterogeneidad en la configuración e implementación de los modelos e intervenciones, su fuerte vinculación a los contextos de origen y las dificultades metodológicas derivadas de la complejidad del objeto de las investigaciones explican el carácter necesariamente aproximativo de la propuesta de síntesis expuesta en la tabla 3 pero, a su vez, su ordenación resumida permite ponderar la calidad y la dirección del efecto de las evidencias publicadas e identificar determinados elementos importantes de cara a la discusión de propuestas de mejora en nuestro medio.

DISCUSIÓN

Los resultados del análisis han permitido conocer las principales características de los 6 modelos principales de colaboración AP-SM en la atención de la depresión identificados en la literatura y aproximarnos a la evaluación de la calidad de las evidencias publicadas sobre su efectividad, generando al menos tres ejes principales para su discusión.

En primer lugar, los resultados de la meta-revisión corroboran, por un lado, la gran variabilidad existente en la práctica de los modelos colaborativos y su condicionamiento por las características de los servicios de salud de sus respectivos países y, por otro, nos muestran que la calidad de las evidencias, aunque modesta, es mayor en los modelos caracterizados como más complejos, principalmente en los “modelos de atención colaborativa” y en el CCM. Los estudios secundarios publicados más recientemente reproducen su procedencia mayoritaria de EEUU⁹⁹⁻¹⁰⁶ y remarcan la importancia de las particularidades locales así como el carácter complejo de las iniciativas integradoras más efectivas^{99,100}. El enfoque multiprofesional en la atención al paciente, la necesidad de planes de gestión, el seguimiento de los pacientes y la potenciación de la comunicación interprofesional parecen confirmarse como componentes clave de dichos modelos colaborativos^{47,103}. Sin embargo, no debemos olvidar que dichos planes de gestión incorporan con frecuencia elementos de otros modelos^{78,104}, como la atención escalonada⁷, la gestión de casos/cuidados^{25,49,51,69} o el apoyo a la educación y autogestión de los pacientes en el caso del CCM^{7,72}. Por otro lado, la comunicación interprofesional y las actividades de formación y de difusión de las GPC son las intervenciones más frecuentemente implementadas, pero también las más cuestionadas cuando se llevan a cabo de manera aislada^{8,22,45,53,68,75,96,107,108}. Nos encontramos, por tanto, con la paradoja de que los mode-

los más efectivos son los más complejos, lo cual a su vez les hace más difícilmente evaluables, más dependientes de las singularidades de sus contextos y, por lo tanto, menos transferibles. Las primeras consecuencias se sitúan en un doble plano. Por un lado, debería cuestionarse la difusión e importación mecánica de modelos de colaboración AP-SM en la atención a las personas con depresión sin analizar previamente las circunstancias y contextos donde se han desarrollado y se ha evaluado su efectividad. Y por otro, habría que tener en cuenta la deseable complementariedad de intervenciones diferentes, lo cual exige atender en su diseño y evaluación no solo al desarrollo efectivo de cada componente sino también a la propia gestión integradora^{100,102}.

En segundo lugar, las carencias en la calidad de las evidencias sobre su efectividad refuerzan la necesidad de precisar el significado de los términos a utilizar en la evaluación de modelos de colaboración^{109,110}. El ejemplo más llamativo es el de la denominación genérica de modelos de “atención colaborativa” o *collaborative care*. Aun teniendo en cuenta su frecuente vinculación al CCM, la variabilidad en la aplicación de sus componentes dificulta sobremanera su evaluación^{11,87,99,102}. La confusión conceptual no queda resuelta con la reducción numérica de estos componentes en algunas de las nuevas revisiones publicadas¹⁰³⁻¹⁰⁵, sino que requiere una explicación más precisa de las actividades que se llevan a cabo. Así, en lo que respecta a la implantación de gestores de casos/cuidados como elemento básico de los modelos colaborativos^{104,111} no debería olvidarse la diversidad de profesiones, nivel de formación y condiciones laborales ni los consiguientes efectos en la evaluación de su labor^{8,11,33,57,69,70,101,102,104,112}. Por lo tanto, resulta difícil concluir si las evidencias de mejora derivadas de su intervención son debidas a su figura como gestores o a las funciones de seguimiento, coordinación y responsabilidad continuada que llevan a cabo. La definición y evaluación de funcio-

nes resulta sin duda más laboriosa que identificar la presencia/ausencia de un nuevo profesional, pero en la asistencia a las personas con depresión en AP esa tarea resulta imprescindible, sobre todo si tenemos en cuenta la responsabilidad de los actuales profesionales de AP en la continuidad de la atención integral del paciente y la gran importancia de evitar el riesgo de su fragmentación¹¹³. La falta de precisión afecta también a la propia definición de la depresión y a lo que ello conlleva, por ejemplo, en los resultados vinculados a la adherencia a los tratamientos farmacológicos. Las particularidades de la depresión como problema de salud a nivel diagnóstico, clasificatorio y terapéutico^{3,114,115} han de ser por tanto tenidas en cuenta en el diseño y evaluación de estrategias de colaboración orientadas a su mejora. Otra de las limitaciones de las evaluaciones revisadas es la falta de mantenimiento a largo plazo de las mejoras, resultando excepcionales a partir del segundo año. Esta deficiencia, señalada también en publicaciones más recientes^{100,102,103}, estaría condicionada en parte por su vinculación a investigaciones con recursos extra que tienden a desaparecer una vez finalizado el estudio y, sobre todo, porque la sostenibilidad a largo plazo de las estrategias de colaboración requiere la implicación mantenida tanto de la gestión como de los profesionales^{43,51,99,101}. Los retos no son fáciles de resolver. Las políticas sanitarias se ven frecuentemente condicionadas por las presiones a corto plazo de los grupos interesados, mientras que la colaboración a largo plazo entre niveles asistenciales exige compromisos estables y consensuados^{116,117}. Además, resulta imprescindible “bajar” al terreno concreto de las interrelaciones entre servicios, estamentos y especialidades, teniendo en cuenta sus “culturas” y trayectorias, para poder identificar las mejores prácticas colaborativas desde la realidad de los profesionales^{13,14,99,101,118}. En nuestro sistema sanitario se ha constatado repetidamente que las estrategias de colaboración requieren fundamentarse en la AP y en la capacitación de

los profesionales que trabajan en ella para asistir correctamente a las personas que acuden con síntomas depresivos, compartiendo en caso necesario dicha asistencia con los servicios de SM^{4,119-121}. Pero para su práctica mantenida es necesario un reforzamiento y readecuación de la AP y de la SM en los servicios públicos de salud de nuestro país que contemple la valoración integral del paciente, el compromiso continuado en su seguimiento y la colaboración interprofesional y de aprendizaje recíproco (no solo unidireccional, ni limitado a lo farmacológico) como referentes prioritarios en la asignación, evaluación e incentivación de los profesionales implicados.

Y, por último, es preciso llamar la atención sobre la necesidad de continuar desarrollando nuevos instrumentos y enfoques metodológicos de evaluación y revisión que se correspondan mejor con la complejidad de los modelos y experiencias de colaboración entre niveles asistenciales. Las meta-revisiones de estudios secundarios (también denominadas “overviews”) favorecen la accesibilidad y el análisis de la evidencia respecto de un determinado objeto de estudio dado su mayor nivel de agregación, lo cual resulta especialmente útil cuando se parte de una finalidad pragmática, como en este caso¹²². Pero a su vez conllevan el riesgo de acumular los sesgos y carencias de las revisiones previas¹⁶, como la ya comentada variabilidad de los contextos sanitarios (mayoritariamente EEUU), o las insuficiencias de las herramientas evaluadoras utilizadas. Las repercusiones metodológicas de la variabilidad contextual, ya señaladas en el ámbito de la Salud Pública¹²³, cobran especial importancia en la evaluación de “intervenciones complejas” como es el caso de la colaboración AP-SM^{112,124} donde han de tenerse en cuenta tanto los procesos “exógenos” como “endógenos” en los que intervienen la interacción, la confianza y la complicidad entre gestores y profesionales, entre profesionales entre sí y entre estos y los pacientes¹²⁵. En consecuencia, en la evaluación de las experiencias co-

laborativas AP-SM en la atención a las personas con depresión adquiere especial trascendencia conocer además del “qué”, las circunstancias acerca de “quién”, “dónde”, “cómo” y “por qué”¹²¹. Sin embargo, los instrumentos habitualmente utilizados no contemplan suficientemente los componentes dinámicos y plurifactoriales que contribuyen a la complejidad de las intervenciones colaborativas^{126,127}. Las propuestas publicadas respecto a las revisiones sistemáticas de intervenciones complejas se han orientado a la necesidad de ampliar el campo de búsqueda, definir mejor las intervenciones e incorporar evidencias complementarias (cualitativas) o bien a cuestionar el modelo tradicional de revisión sistemática proponiendo enfoques más dinámicos y teóricos^{124,128-130}. El desarrollo práctico de la presente revisión confirma la justificación de avanzar en ambos sentidos. La efectividad de los distintos modelos difícilmente puede ser evaluada partiendo únicamente de los resultados estadísticos de las revisiones sistemáticas previas. En nuestro caso fue imprescindible recurrir a otras fuentes que aportaran información sobre las características, contextos y circunstancias en los que dichos modelos se habían desarrollado y evaluado, optándose por un criterio extenso en la inclusión de los estudios secundarios identificados. Entendemos, por tanto, que los componentes narrativo e iterativo resultan necesarios en el proceso de análisis y búsqueda de evidencias junto a las consiguientes exigencias de rigor y de transparencia¹³⁰. Por otra parte, la búsqueda, identificación y síntesis de las carencias metodológicas señaladas por los propios autores de los estudios secundarios constituyó, en nuestro caso, un objeto de estudio específico que resultó de gran ayuda en el intento de llevar a cabo una gradación de la calidad de las evidencias encontradas, aunque también aquí la insuficiente adecuación de los instrumentos y estándares habituales¹³¹ afecta a las meta-revisiones de estudios secundarios^{16,17}. Como ya se ha comentado previamente, el diseño original de la herramienta AMSTAR responde a la evaluación de revisiones sistemáticas y metaná-

lisis, y no recoge las particularidades contextuales y de proceso de este tipo de intervenciones, lo cual conlleva que los resultados de su aplicación deban ser interpretados en este caso con prudencia. Recomendación asimismo extensible a la ordenación, necesariamente aproximativa, de los distintos modelos conforme a grados de calidad de las evidencias encontradas. Teniendo en cuenta lo anterior, la presente meta-revisión aporta también en este campo una experiencia a considerar en las evaluaciones y revisiones de modelos de colaboración, no solo entre AP y SM. Las dificultades y los retos señalados a propósito de la búsqueda, análisis y síntesis de las evidencias precisan de marcos teóricos y herramientas metodológicas que se correspondan mejor con la naturaleza compleja de las estrategias colaborativas y nos prevengan de los errores que pueden derivarse de la aplicación precipitada de modelos insuficientemente evaluados en nuestros servicios de salud.

AGRADECIMIENTOS

Los autores desean agradecer la financiación del proyecto al Departamento de Salud del Gobierno Vasco (BIOEF BIO09/EM/014; OSTEBA 2010COM112007).

BIBLIOGRAFÍA

1. Tylee A, Haddad M. Managing complex problems: treatment for common mental disorders in the UK. *Epidemiol Psichiatr Soc.* 2007 Oct;16(4):302-8.
2. Funk M, Ivbijaro G, Saraceno B, Freeman M, Epping-Jordan J, Faydi E, et al. Integrating mental health into primary care. A global perspective. Singapore: WHO Press; 2008.
3. Maxwell M, Pratt R. Prevention and management of depression in primary care in Europe: a holistic model of care and interventions - position paper of the European Forum for Primary Care. *Qual Prim Care.* 2008;16(3):187-96.
4. Calderón C, Retolaza A, Bacigalupe A, Payo J, Grandes G. Médicos de familia y psiquiatras ante el paciente con depresión: la necesidad de readecuar enfoques asistenciales y dinámicas organizativas. *Aten Primaria.* 2009;41(1):33-40.
5. Katon W, Unützer J. Collaborative Care Models for Depression: time to move from evidence to practice. *Arch Intern Med.* 2006;166:2301-3.
6. Simon G. Collaborative care for depression. *BMJ.* 2006 Feb 4;332:249-50.
7. Kates N, Mach M. Chronic disease management for depression in primary care: a summary of the current literature and implications for practice. *Can J Psychiatry.* 2007 Feb;52(2):77-85.
8. Christensen H, Griffiths KM, Gulliver A, Clack D, Kljakovic M, Wells L. Models in the delivery of depression care: a systematic review of randomised and controlled intervention trials. *BMC Fam Pract.* 2008;9:25.
9. Collins C, Hewson D, Munger R, Wade T. Evolving Models of Behavioral Health Integration in Primary Care. New York: Milbank Memorial Fund; 2010.
10. Craven MA, Bland R. Shared mental health care: a bibliography and overview. *Can J Psychiatry.* 2002 Apr;47(2:Suppl 1):Suppl-viiiS.
11. Bower P, Gilbody S, Richards D, Fletcher J, Sutton A. Collaborative care for depression in primary care. Making sense of a complex intervention: systematic review and meta-regression. *Br J Psychiatry.* 2006 Dec;189:484-93.
12. Eusko Jaurlaritz - Gobierno Vasco. Dpto. de Sanidad y Consumo, editor. Estrategia para afrontar el reto de la cronicidad en Euskadi. 2010.
13. Calderón C, Retolaza A, Payo J, Bacigalupe A, Zallo E, Mosquera I. Perspectivas de los pacientes diagnosticados de depresión y atendidos por médicos de familia y psiquiatras. *Aten Primaria.* 2012;44(10):595-602.
14. Balagué L, Calderón C, Retolaza A, Bacigalupe A, Payo J, Mosquera I. Los "otros profesionales" en la atención a personas con depresión: psicólogos, trabajadores sociales y enfermeras. *Index Enferm (edición digital).* 2013;22(1-2):35-9. Disponible en: <http://www.index-f.com/index-enfermeria/v22n1-2/8008.php>. Consultado el 19 de Julio de 2013.
15. Calderón C, Mosquera I, Balagué L, Retolaza A, Bacigalupe A, Belaunzaran J, et al. Modelos e intervenciones de colaboración entre Atención Primaria y Salud Mental en la atención a los pacientes diagnosticados de depresión: resultados de una meta-revisión. Investigación Comisionada. Informe Osteba D-13-13. Vitoria-Gasteiz: Departamento de Salud. Gobierno Vasco; 2013.

16. Whitlock EP, Lin JS, Chou R, Shekelle P, Robinson KA. Using existing systematic reviews in complex systematic reviews. *Ann Intern Med.* 2008 May 20;148(10):776-82.
17. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med Res Methodol.* 2011;11(1):15.
18. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol.* 2009 Oct;62(10):1013-20.
19. Canadian Agency for Drugs and Technologies in Health (CADTH). Methods for Development. CADTH 2012. Disponible en: <http://www.cadth.ca/en/resources/rx-for-change/methods-for-development>.
20. Barrientos X, Fritsch R, Solís J, Rojas G. Tratamiento de la depresión en la atención primaria. *Rev psiquiatr clín (Santiago de Chile).* 2005 Mar;42(1):49-59.
21. Agius M, Murphy CL, Zaman R. Does shared care help in the treatment of depression? *Psychiatr Danub.* 2010 Nov;22(SUPPL. 1):S18-S22.
22. Gilbody S, Whitty P, Grimshaw J, Thomas R. Improving the recognition and management of depression in primary care. *Effective Health Care* 2002;7(5).
23. Gilbody SM, Whitty PM, Grimshaw JM, Thomas RE. Improving the detection and management of depression in primary care. *Qual Saf Health Care.* 2003 Apr;12(2):149-55.
24. Gilbody S, Whitty P, Grimshaw J, Thomas R. Educational and organizational interventions to improve the management of depression in primary care: a systematic review. *JAMA.* 2003 Jun 18;289(23):3145-51.
25. Oxman TE, Dietrich AJ, Schulberg HC. Evidence-based models of integrated management of depression in primary care. *Psychiatr Clin North Am.* 2005 Dec;28(4):1061-77.
26. Oxman TE, Dietrich AJ, Schulberg HC. The depression care manager and mental health specialist as collaborators within primary care. *Am J Geriatr Psychiatry.* 2003 Sep;11(5):507-16.
27. Druss BG. The mental health/primary care interface in the United States: history, structure, and context. *Gen Hosp Psychiatry.* 2002 Jul;24(4):197-202.
28. Katon W, Unutzer J, Wells K, Jones L. Collaborative depression care: history, evolution and ways to enhance dissemination and sustainability. *Gen Hosp Psychiatry.* 2010 Sep;32(5):456-64.
29. Peters AH, de Leeuw RJ, Schrijvers GJ. Integrating care for people with depression: developments in the Netherlands. *Int J Integr Care.* 2010;10:e62.
30. Wulsin LR, Sollner W, Pincus HA. Models of Integrated Care. *Med Clin North Am.* 2006 Jul;90(4):647-77.
31. Mauer BJ, Druss BG. Mind and body reunited: improving care at the behavioral and primary healthcare interface. *J Behav Health Serv Res.* 2010 Oct;37(4):529-42.
32. Rost KM, Duan N, Rubenstein LV, Ford DE, Sherbourne CD, Meredith LS, et al. The Quality Improvement for Depression collaboration: general analytic strategies for a coordinated study of quality improvement in depression care. *Gen Hosp Psychiatry.* 2001 Sep;23(5):239-53.
33. Belnap BH, Kuebler J, Upshur C, Kerber K, Mokrín DR, Kilbourne AM, et al. Challenges of implementing depression care management in the primary care setting. *Adm Policy Ment Health.* 2006 Jan;33(1):65-75.
34. Higgins MJ, Lacy TJ. The role of psychiatry in U.S. Air Force primary care: research-based models for clinical consultation and education. *Mil Med.* 2006 Nov;171(11):1117-22.
35. Bassilios B, Pirkis J, Fletcher J, Burgess P, Gurrin L, King K, et al. The complementarity of two major Australian primary mental health care initiatives. *Aust N Z J Psychiatry.* 2010 Nov;44(11):997-1004.
36. Blount A. Integrated Primary Care: Organizing the Evidence. *Fam Syst & Health.* 2003 Jun;21(2):121-33.
37. Gensichen J, Beyer M, Muth C, Gerlach FM, von Korff M, Ormel J. Case management to improve major depression in primary health care: a systematic review. *Psychol Med.* 2006 Jan;36(1):7-14.
38. Smit A, Tiemens BG, Ormel J. Improving long-term outcome of depression in primary care: A review of RCTs with psychological and supportive interventions. *Eur J Psychiatry.* 2007 Jan;21(1):37-48.
39. Parrish E, Peden A. Clinical outcomes of depressed clients: a review of current literature. *Issues Ment Health Nurs.* 2009 Jan;30(1):51-60.

40. Cape J, Whittington C, Bower P. What is the role of consultation-liaison psychiatry in the management of depression in primary care? A systematic review and meta-analysis. *Gen Hosp Psychiatry*. 2010 May;32(3):246-54.
41. Bijl D, van Marwijk HW, de Haan M, van Tilburg W, Beekman AJ. Effectiveness of disease management programmes for recognition, diagnosis and treatment of depression in primary care. *Eur J Gen Pract*. 2004 Mar;10(1):6-12.
42. Neumeyer-Gromen A, Lampert T, Stark K, Kallischnigg G. Disease management programs for depression: a systematic review and meta-analysis of randomized controlled trials. *Med Care*. 2004 Dec;42(12):1211-21.
43. Beaudin CL, Kramer TL. Evaluation and treatment of depression in managed care (part II) quality improvement programs. *Dis Manage Health Outcomes*. 2005;13(5):307-16.
44. Adli M, Bauer M, Rush AJ. Algorithms and collaborative-care systems for depression: are they effective and why? A systematic review. *Biol Psychiatry*. 2006 Jun 1;59(11):1029-38.
45. Barbui C, Tansella M. Identification and management of depression in primary care settings. A meta-review of evidence. *Epidemiol Psichiatri Soc*. 2006 Oct;15(4):276-83.
46. Craven MA, Bland R. Better practices in collaborative mental health care: an analysis of the evidence base. *Can J Psychiatry*. 2006 May;51(6 Suppl 1):7S-72S.
47. Gunn J, Diggins J, Hegarty K, Blashki G. A systematic review of complex system interventions designed to increase recovery from depression in primary care. *BMC Health Serv Res*. 2006;6:88.
48. Hilty DM, Yellowlees PM, Cobb HC, Bourgeois JA, Neufeld JD, Nesbitt TS. Models of telepsychiatric consultation-liaison service to rural primary care. *Psychosomatics*. 2006 Mar;47(2):152-7.
49. Williams JW, Jr., Gerrity M, Holsinger T, Dobscha S, Gaynes B, Dietrich A. Systematic review of multifaceted interventions to improve depression care. *Gen Hosp Psychiatry*. 2007 Mar;29(2):91-116.
50. Butler M, Kane R, McAlpine D, Kathol R, Fu S, Hagedorn H, et al. Integration of Mental Health/Substance Abuse and Primary Care No. 173 (Prepared by the Minnesota Evidence-based Practice Center under Contract No. 290-02-0009). AHRQ Publication No. 09-E003. Rockville, MD: Agency for Healthcare Research and Quality; 2008 Oct.
51. Rubenstein LV, Williams JW, Jr., Danz M, Shekelle P. Determining key features of effective depression interventions. Los Angeles: 2009.
52. Callahan CM. Quality improvement research on late life depression in primary care. *Med Care*. 2001 Aug;39(8):772-84.
53. Bartels SJ, Dums AR, Oxman TE, Schneider LS, Arean PA, Alexopoulos GS, et al. Evidence-based practices in geriatric mental health care. *Psychiatr Serv*. 2002 Nov;53(11):1419-31.
54. Skultety KM, Zeiss A. The treatment of depression in older adults in the primary care setting: An evidence-based review. *Health Psychol*. 2006 Nov;25(6):665-74.
55. McEvoy P, Barnes P. Using the chronic care model to tackle depression among older adults who have long-term physical conditions. *J Psychiatr Ment Health Nurs*. 2007 May;14(3):233-8.
56. Vannoy S, Powers D, Unützer J. Models of care for treating late-life depression in primary care. *Aging Health*. 2007 Feb;3(1):67-75.
57. Skultety KM, Rodriguez RL. Treating geriatric depression in primary care. *Curr Psychiatry Rep*. 2008 Feb;10(1):44-50.
58. Huang C, Dong B, Lu Z, Zhang Y, Pu Y, Liu Q. Collaborative care interventions for depression in the elderly: a systematic review of randomized controlled trials. *J Investig Med*. 2009 Feb;57(2):446-55.
59. Lu CY, Ross-Degnan D, Soumerai SB, Pearson S-A. Interventions designed to improve the quality and efficiency of medication use in managed care: A critical review of the literature - 2001-2007. *BMC Health Serv Res*. 2008;8:Article Number: 75.
60. van der Feltz-Cornelis CM, Van Os TWDP, Van Marwijk HWJ, Leentjens AFG. Effect of psychiatric consultation models in primary care. A systematic review and meta-analysis of randomized clinical trials. *J Psychosom Res*. 2010 Jun;68(6):521-33.
61. Katon WJ, Seelig M. Population-based care of depression: Team care approaches to improving outcomes. *J Occup Environ Med*. 2008 Apr;50(4):459-67.
62. Vergouwen ACM, Bakker A, Katon WJ, Verheij TJ, Koerselman F. Improving adherence to antidepressants: A systematic review of interventions. *J Clin Psychiatry*. 2003 Dec;64(12):1415-20.

63. Berardi D, Bortolotti B, Menchetti M, Bombi A, Tarricone I. Models of collaboration between general practice and mental health services in Italy. *Eur J Psychiatry*. 2007 Jan;21(1):79-84.
64. Harkness EF, Bower PJ. On-site mental health workers delivering psychological therapy and psychosocial interventions to patients in primary care: effects on the professional practice of primary care providers. *Cochrane Database Syst Rev*. 2009;(1):CD000532.
65. Weingarten SR, Henning JM, Badamgarav E, Knight K, Hasselblad V, Gano J, et al. Interventions used in disease management programmes for patients with chronic illness - Which ones work? Meta-analysis of published reports. *BMJ*. 2002 Oct 26;325(7370):925-8.
66. Smith SM, Allwright S, O'Dowd T. Effectiveness of shared care across the interface between primary and specialty care in chronic disease management. *Cochrane Database Syst Rev*. 2007;(3):CD004910.
67. Doughty C. Effective models of mental health service provision and workforce configuration in the primary care setting. *NZHTA Technical Brief* 2006 Feb;5(1).
68. Bower P, Gilbody S. Managing common mental health disorders in primary care: conceptual models and evidence base. *BMJ*. 2005 Apr 9;330(7495):839-42.
69. Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. *Arch Intern Med*. 2006 Nov 27;166(21):2314-21.
70. Magruder KM, Yeager DE. Mental health problems in primary care: Progress in North America. *Eur J Psychiatry*. 2007 Jan;21(1):55-61.
71. Christensen H, Griffiths KM, Gulliver A. Plenty of activity but little outcome data: a review of the "grey literature" on primary care anxiety and depression programs in Australia. *MJA*. 2008 Jun;188(12 Suppl):S103-S106.
72. Tsai AC, Morton SC, Mangione CM, Keeler EB. A meta-analysis of interventions to improve care for chronic illnesses. *Am J Manag Care*. 2005 Aug;11(8):478-88.
73. Hall J, Callaghan P. Developments in managing mental health care: a review of the literature. *Issues Ment Health Nurs*. 2008 Dec;29(12):1245-72.
74. Mauer B. Behavioral Health/Primary Care Integration. The Four Quadrant Model and evidence-based practices. National Council for Community Behavioral Healthcare; 2006 Feb.
75. Anfinson TJ, Bona JR. A health services perspective on delivery of psychiatric services in primary care including internal medicine. *Med Clin North Am*. 2001 May;85(3):597-616.
76. Meyer F, Peteet J, Joseph R. Models of care for co-occurring mental and medical disorders. *Harv Rev Psychiatry*. 2009;17(6):353-60.
77. National Collaborating Centre for Mental Health. Depression. The NICE guideline on the treatment and management of depression in adults (updated edition). Leicester & London: The British Psychological Society and The Royal College of Psychiatrists; 2010.
78. Gask L, Khanna T. Ways of working at the interface between primary and specialist mental health-care. *Br J Psychiatry*. 2011 Jan;198:3-5, sup.
79. Gask L, Sibbald B, Creed F. Evaluating models of working at the interface between mental health services and primary care. *Br J Psychiatry*. 1997 Jan;170:6-11.
80. Fuller JD, Perkins D, Parker S, Holdsworth L, Kelly B, Fragar L, et al. Systematic review on service linkages in primary mental health care: informing Australian policy and practice. Canberra: Australian Primary Health Care Research Institute (APHCRI); 2009.
81. Bower P, Gask L. The changing nature of consultation-liaison in primary care: bridging the gap between research and practice. *Gen Hosp Psychiatry*. 2002;24(2):63-70.
82. Bauer AM, Fielke K, Brayley J, Araya M, Alem A, Frankel BL, et al. Tackling the global mental health challenge: A psychosomatic medicine/consultation-liaison psychiatry perspective. *Psychosomatics*. 2010 May;51(3):185-93.
83. Care Continuum Alliance (CCA). Definition of disease management. CCA 2013. Disponible en: http://www.carecontinuumalliance.org/dm_definition.asp.
84. Case Management Society of America (CMSA). What is a case manager? CMSA 2012. Disponible en: <http://www.cmsa.org/Home/CMSA/WhatisaCaseManager/tabid/224/Default.aspx>.

85. Badamgarav E, Weingarten SR, Henning JM, Knight K, Hasselblad V, Gano A, Jr., et al. Effectiveness of disease management programs in depression: a systematic review. *Am J Psychiatry*. 2003 Dec;160(12):2080-90.
86. Mittman BS. Creating the evidence base for quality improvement collaboratives. *Ann Intern Med*. 2004 Jun 1;140(11):897-901.
87. Schouten LM, Hulscher ME, van Everdingen JJ, Huijsman R, Grol RP. Evidence for the impact of quality improvement collaboratives: systematic review. *BMJ*. 2008 Jun 28;336(7659):1491-4.
88. Solberg LI. If you've seen one quality improvement collaborative... *Ann Fam Med*. 2005 May;3(3):198-9.
89. Nease DE, Jr., Nutting PA, Dickinson WP, Bonham AJ, Graham DG, Gallagher KM, et al. Inducing sustainable improvement in depression care in primary care practices. *Jt Comm J Qual Patient Saf*. 2008 May;34(5):247-55.
90. Franx G, Meeuwissen JA, Sinnema H, Spijker J, Huyser J, Wensing M, et al. Quality improvement in depression care in the Netherlands: the Depression Breakthrough Collaborative. A quality improvement report. *Int J Integr Care*. 2009;9:e84.
91. Coleman K, Austin BT, Brach C, Wagner EH. Evidence on the Chronic Care Model in the new millennium. *Health Aff*. 2009 Jan;28(1):75-85.
92. Burton C, Maxwell M. Depression as chronic disease. *Br J Gen Pract*. 2007 Jul;57(540):582.
93. Gervas J. El Modelo de Atención a Crónicos (Chronic Care Model). ¿Qué puede aportar y qué inconvenientes tiene? *Salud 2000*. 2010 Sep;128:12.
94. Ford DE, Pincus HA, Unützer J, Bauer MS, Gonzalez JJ, Wells KB, et al. Practice-based interventions. *Ment Health Serv Res*. 2002 Dec;4(4):199-204.
95. Byrne N, Regan C, Livingston G. Adherence to treatment in mood disorders. *Curr Opin Psychiatry*. 2006;19(1):44-9.
96. Thielke S, Vannoy S, Unützer J. Integrating mental health and primary care. *Prim Care*. 2007 Sep;34(3):571-92, vii.
97. Simon G. Collaborative care for mood disorders. *Curr Opin Psychiatry*. 2009 Jan;22(1):37-41.
98. Bower P. Primary care mental health workers: models of working and evidence of effectiveness. *Br J Gen Pract*. 2002 Nov;52(484):926-33.
99. Butler M, Kane RL, McAlpine D, Kathol R, Fu SS, Hagedorn H, et al. Does integrated care improve treatment for depression? A systematic review. *J Ambulatory Care Manage*. 2011 Apr;34(2):113-25.
100. Fuller JD, Perkins D, Parker S, Holdsworth L, Kelly B, Roberts R, et al. Effectiveness of service linkages in primary mental health care: a narrative review part 1. *BMC Health Serv Res*. 2011;11:72.
101. Fuller JD, Perkins D, Parker S, Holdsworth L, Kelly B, Roberts R, et al. Building effective service linkages in primary mental health care: a narrative review part 2. *BMC Health Serv Res*. 2011 Mar;11:66.
102. Kelly BJ, Perkins DA, Fuller JD, Parker SM. Shared care in mental illness: A rapid review to inform implementation. *Int J Ment Health Syst*. 2011 Nov 21;5(1):31.
103. Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, et al. Collaborative care for depression and anxiety problems. *Cochrane Database Syst Rev*. 2012;(10):CD006525.
104. Thota AB, Sipe TA, Byard GJ, Zometa CS, Hahn RA, McKnight-Eily LR, et al. Collaborative care to improve the management of depressive disorders: a community guide systematic review and meta-analysis. *Am J Prev Med*. 2012 May;42(5):525-38.
105. Woltmann E, Grogan-Kaylor A, Perron B, Georges H, Kilbourne AM, Bauer MS. Comparative effectiveness of collaborative chronic care models for mental health conditions across primary, specialty, and behavioral health care settings: systematic review and meta-analysis. *Am J Psychiatry*. 2012 Aug 1;169(8):790-804.
106. Beekman AT, van der Feltz-Cornelis C, van Marwijk HW. Enhanced care for depression. *Curr Opin Psychiatry*. 2013 Jan;26(1):7-12.
107. Sikorski C, Luppá M, König HH, van den Bussche H, Riedel-Heller SG. Does GP training in depression care affect patient outcome? - A systematic review and meta-analysis. *BMC Health Serv Res*. 2012 Jan 10;12(1):10.
108. Griffiths KM, Christensen H. Depression in primary health care: from evidence to policy. *MJA*. 2008 Apr;188(8 Suppl):S81-S83.
109. MacFarlane D. Current State of Collaborative Mental Health Care. Mississauga, ON: Canadian Collaborative Mental Health Initiative; 2005.

110. Miller B, Kessler R, Peek C, Kallenberg G. A National Agenda for Research in Collaborative Care: Papers From the Collaborative Care Research Network Research Development Conference. AHRQ Publication No. 11-0067. Rockville, MD: Agency for Healthcare Research and Quality; 2011 Jul.
111. Community Preventive Services Task Force. Recommendation from the community preventive services task force for use of collaborative care for the management of depressive disorders. *Am J Prev Med*. 2012 May;42(5):521-4.
112. Byng R, Norman I, Redfern S, Jones R. Exposing the key functions of a complex intervention for shared care in mental health: case study of a process evaluation. *BMC Health Serv Res*. 2008;8(274).
113. Starfield B. Primary care: an increasingly important contributor to effectiveness, equity, and efficiency of health services. *SESPAS report 2012*. *Gac Sanit*. 2012 Mar;26 Suppl 1:20-6.
114. Thomas-MacLean R, Stoppard J, Miedema BB, Tatemichi S. Diagnosing depression: there is no blood test. *Can Fam Physician*. 2005 Aug;51:1102-3.
115. Dowrick C. Beyond depression. A new approach to understanding and management. 2nd ed. New York: Oxford University Press; 2009.
116. Rosen R, Asaria P, Dixon A. Improving chronic disease management. An Anglo-American exchange. London, UK; 2007 Nov.
117. Suter E, Oelke ND, Adair CE, Armitage GD. Ten key principles for successful health systems integration. *Healthcare Q*. 2009 Oct;13:Spec-23.
118. Pautler K. Annotated Bibliography of Collaborative Mental Health Care. Mississauga, ON: Canadian Collaborative Mental Health Initiative; 2005.
119. Téllez JM, Cerecedo MJ, Pascual P, Buitrago F. La salud mental en el umbral del siglo XXI. Protagonismo de la atención primaria. ¿Un reto a nuestro alcance? *Aten Primaria* 2005; 35(2):61-3.
120. Ciurana R. La coordinación con los servicios de salud mental. Un tema pendiente, manifiestamente mejorable. *Aten Primaria* 2009; 41(3):157-8.
121. Ortiz A, Ibáñez V. Iatrogenia y prevención cuaternaria en salud mental. *Rev Esp Salud Pública* 2011; 85(6):513-23.
122. Moat KA, Lavis JN, Wilson MG, Rottingen JA, Barnighausen T. Twelve myths about systematic reviews for health system policymaking rebutted. *J Health Serv Res Policy*. 2013 Jan;18(1):44-50.
123. Rychetnik L, Frommer M, Hawe P, Shiell A. Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health*. 2002 Feb;56(2):119-27.
124. Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, et al. Can we systematically review studies that evaluate complex interventions? *PLoS Medicine*. 2009 Aug;6(8):e1000086.
125. May C. A rational model for assessing and evaluating complex interventions in health care. *BMC Health Serv Res*. 2006;6:86.
126. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. *BMJ*. 2007 Mar 3;334(7591):455-9.
127. Simons H. Utilizing Evaluation Evidence to Enhance Professional Practice. *Evaluation*. 2004;10(4):410-29.
128. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Soc Sci Med*. 2005 Jul;61(2):417-30.
129. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review - a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy*. 2005 Jul;10(Suppl 1):21-34.
130. Popay J, (Ed). Moving beyond effectiveness in evidence synthesis. Methodological issues in the synthesis of diverse sources of evidence. London: National Institute for Health and Clinical Excellence; 2006.
131. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ*. 2010 Aug 21;341(7769):384-7.
132. Hilty DM, Luo JS, Morache C, Marcelo DA, Nesbitt TS. Telepsychiatry: An overview for psychiatrists. *CNS Drugs*. 2002;16(8):527-48.
133. Aitken JB, Curtis R. Integrated Health Care: Improving Client Care while Providing Opportunities for Mental Health Counselors. *J Ment Health Couns*. 2004 Oct;26(4):321-31.

134. Katon WJ, Zatzick D, Bond G, Williams JJ. Dissemination of Evidence-Based Mental Health Interventions: Importance to the Trauma Field. *J Trauma Stress*. 2006 Oct;19(5):611-23.

REVISIÓN SISTEMÁTICA

FRECUENTACIÓN DE LAS CONSULTAS DE MEDICINA GENERAL Y ESPECIALIZADA POR POBLACIÓN INMIGRANTE Y AUTÓCTONA: UNA REVISIÓN SISTEMÁTICA

Rocío Carmona (1), Raimundo Alcázar-Alcázar (1), Antonio Sarriá-Santamera (1, 2, 3) y Enrique Regidor (4, 5).

(1) Agencia de Evaluación de Tecnologías Sanitarias. Instituto de Salud Carlos III. Ministerio de Economía y Competitividad.

(2) Departamento de Medicina Preventiva y Salud Pública. Universidad de Alcalá. Alcalá de Henares. Madrid.

(3) Red de Investigación en Servicios de Salud y Enfermedades Crónicas (REDISSEC). Madrid.

(4) Departamento de Medicina Preventiva, Salud Pública e Historia de la Ciencia. Universidad Complutense de Madrid.

(5) CIBER Epidemiología y Salud Pública (CIBERESP). Madrid.

No existen conflictos de interés

RESUMEN

Fundamentos: España pasó a ser en 2006 el país europeo con mayor saldo migratorio. Un aspecto poco estudiado es el acceso y utilización de los servicios sanitarios de esta población. El objetivo de este estudio fue describir la evidencia científica sobre si existen diferencias en la utilización de los servicios de medicina general y especializada entre la población inmigrante con respecto a la autóctona.

Métodos: Revisión sistemática. Se llevó a cabo una búsqueda de la literatura científica nacional e internacional de estudios que comparaban la utilización de los servicios de medicina general y especializada entre inmigrantes y autóctonos desde 1994 hasta 2013. Se utilizó para ello la base de datos MEDLINE además de una búsqueda manual, sin límite de lengua ni tipo de estudio. Se evaluó la calidad metodológica de los 29 trabajos finalmente incluidos. Se recopilaron características de los sujetos, del contexto, metodológicas y extrínsecas para comparar los estudios incluidos.

Resultados: Se seleccionaron 29 estudios relativos a consulta de medicina general (9 realizados en España) y 15 sobre consulta a especialistas (7 de España) los cuales utilizan principalmente las encuestas de salud como fuente de información. Analizaron tanto la frecuentación como el contacto con el médico general y/o especialista según la nacionalidad o el país de nacimiento (entre otros), ajustando en su mayoría por variables de necesidad y/o socioeconómica.

Conclusiones: En líneas generales, en España la población inmigrante hace el mismo uso de los servicios de medicina general que la autóctona y un menor o igual uso de los servicios de medicina especializada. Los resultados de los estudios realizados en otros países van en la misma línea.

Palabras clave: Revisión sistemática. Emigrantes e inmigrantes. Accesibilidad a los Servicios de Salud.

ABSTRACT -

Use of Health Services for Immigrants and Native Populations: a Systematic Review

Background: Spain was among the top immigration destinations globally between 1990 and 2005, becoming in 2006 in the European country with the highest net migration. As a result of the migration process and the living conditions in the host countries, immigrants' health may be affected. Limited research has investigated access and use of health services for this population. The aim of this study was to describe the scientific evidence on the use of general and specialist medical services for the immigrant population compared to the native.

Methods: Systematic review. It has carried out a search of the national and international scientific literature of comparative studies on the use of general and specialist medical services among immigrant and native since 1994-2013. It was used the MEDLINE database as well as a manual search, no language limit or type of study. The methodological quality of the 29 studies included was evaluated. Subject, context, methodological and extrinsic characteristics were collected for comparison of the included studies.

Results: We selected 29 studies on the general practitioners' (9 from Spain) and 15 of specialist physician (7 from Spain), they mainly used health surveys as a source of information. Analyze both the attendance and contact with the general practitioner / specialist by nationality or country of birth (among others), mostly by adjusting variables of need and / or socioeconomic.

Conclusion: Overall, the immigrant population in Spain have a similar use of general medical services than the native population, and less or similar use of the specialist physician services. These results are in line with studies in other countries.

Keyword: Systematic review. Emigrants and Immigrants. Health Services Accessibility.

Correspondencia

Rocío Carmona

Agencia de Evaluación de Tecnologías Sanitarias. Instituto de Salud Carlos III
Avenida Monforte de Lemos n.º 5. Pabellón 4
Madrid 28029. España
rocio.carmona@isciii.es

INTRODUCCIÓN

La inmigración es un fenómeno que ha ido en aumento en España desde la década de los noventa. Según los datos proporcionados por el Instituto Nacional de Estadística (INE), el porcentaje de extranjeros se incrementó de un 1,6% en 1998 a un 12,2% a 1 de enero de 2011¹, convirtiéndose España en el año 2006 en el país de la Unión Europea con mayor saldo migratorio².

Este hecho pone de manifiesto la necesidad de tener un mejor conocimiento de la situación de esta población a nivel demográfico, económico, socio-sanitario y epidemiológico³. La literatura de otros países sugiere que las personas que emigran suelen ser sanas^{4,5} aunque su salud se puede ver afectada por distintos factores de riesgo al llegar al país de destino, lo cual ocurre principalmente en los procedentes de países más desfavorecidos^{3,6}. Uno de los factores es su desconocimiento del funcionamiento de los servicios sanitarios⁷. Así, existen estudios que indican que el uso de servicios sanitarios por parte de la población inmigrante es relativamente menor que en la población autóctona y que parte de las posibles diferencias en salud entre ambas poblaciones podrían tener relación con un acceso diferente al sistema sanitario y un uso menor de las actividades preventivas por parte de la población inmigrante⁸.

Igualmente, en España, algunos autores^{5,6,9} han descrito que la población inmigrante utiliza con menor frecuencia que la autóctona la mayoría de los servicios sanitarios y acude cuando las enfermedades se encuentran más avanzadas¹⁰. No obstante, es preciso introducir matices en los resultados en función del tipo de servicio estudiado (atención primaria, especializada, hospitalizaciones y urgencias), del colectivo de inmigrantes estudiado y del ámbito geográfico considerado. Por ejemplo, Aerny¹¹ halló una relación entre el tiempo de residencia en España y la frecuentación en

medicina general, incrementándose esta al aumentar el tiempo de residencia, igualando e incluso superando a la de los autóctonos. Un estudio en Canadá llegó a las mismas conclusiones¹².

El objetivo de este estudio fue describir la evidencia científica sobre si existen diferencias en la utilización de los servicios de medicina general y especializada entre la población inmigrante con respecto a la autóctona.

MATERIAL Y MÉTODOS

Se realizó una revisión sistemática de la literatura científica en España y en otros países publicada que comparara la utilización de los servicios de medicina general y especializada entre inmigrantes y autóctonos desde el año 1994 hasta 2013.

Los trabajos seleccionados para la revisión debían cumplir los siguientes requisitos: (1) estudios sin límite de idioma ni tipo de estudio; (2) Con fechas de publicación entre enero de 1994 y septiembre de 2013; (3) los sujetos estudiados podían ser de cualquier etnia y sexo y no limitarse a población infantil; (4) ser estudios comparativos entre población inmigrante y nativos de la región estudiada, quedando claramente definidos ambos grupos poblacionales; (5) que analizaran como variable dependiente la utilización de los servicios sanitarios de medicina general y especializada. Además, se seleccionaron los trabajos incluidos en una revisión previa más amplia⁶ siempre que cumplieran los criterios de inclusión, aunque sus límites temporales fueron entre 1994 y 2007. Se excluyeron los estudios restringidos a pacientes con patologías concretas.

La estrategia de selección de los estudios consistió en una búsqueda bibliográfica exhaustiva en la base de datos electrónica Medline a través de Pubmed (realizada por uno de los autores en calidad de documentalista, siguiendo los criterios establecidos por

los autores) y una búsqueda manual a partir de las referencias de los artículos seleccionados para localizar estudios no hallados en la búsqueda anterior (llevada a cabo por otro de los autores). Se realizó una búsqueda en junio 2013 para el período 01/01/1994 a 01/06/2013 y una búsqueda de actualización final para el período 01/06/2013 a 01/09/2013.

Se llevaron a cabo dos estrategias de búsqueda, una para la localización de estudios comparativos sobre consultas realizadas en medicina general y otra para las realizadas en especializada (anexos 1 y 2).

Para obtener en las búsquedas los estudios que relativos a España se aplicó un filtro geográfico específico desarrollado a partir del creado por Valderas¹³ (anexo 3).

Tras las búsquedas, se procedió a la eliminación de artículos que no cumplían los criterios de inclusión, tanto a través de la lectura del título y del resumen como de la lectura crítica completa de los trabajos potencialmente relevantes. Dos autores (RC y ER) llevaron a cabo este proceso de forma independiente. Se seleccionó un total de 44 estudios, 29 sobre consultas en medicina general (9 de España) y 15 estudios sobre consultas a medicina especializada (7 de España) que cumplían los criterios de inclusión (figura 1). Los estudios incluidos fueron valorados en cuanto a su calidad metodológica a través de las recomendaciones de la iniciativa STROBE con una evaluación de tipo cualitativa (sí/ no) que llevaron a cabo dos autores de forma independiente (RC y AS).

Se determinaron las variables a recopilar de cada uno de los estudios para su posterior comparabilidad. Así, se registraron características de los sujetos (edad de la población estudiada), del contexto (país y año de realización del estudio), metodológicas (tamaño muestral del estudio, fuente de información, estudio longitudinal o transversal, frecuen-

cia en la utilización de los servicios sanitarios investigados, clasificación del sujeto como autóctono o inmigrante, ajuste por variables de necesidad y/o socioeconómicas, principales resultados) y extrínsecas (autor y año de publicación).

RESULTADOS

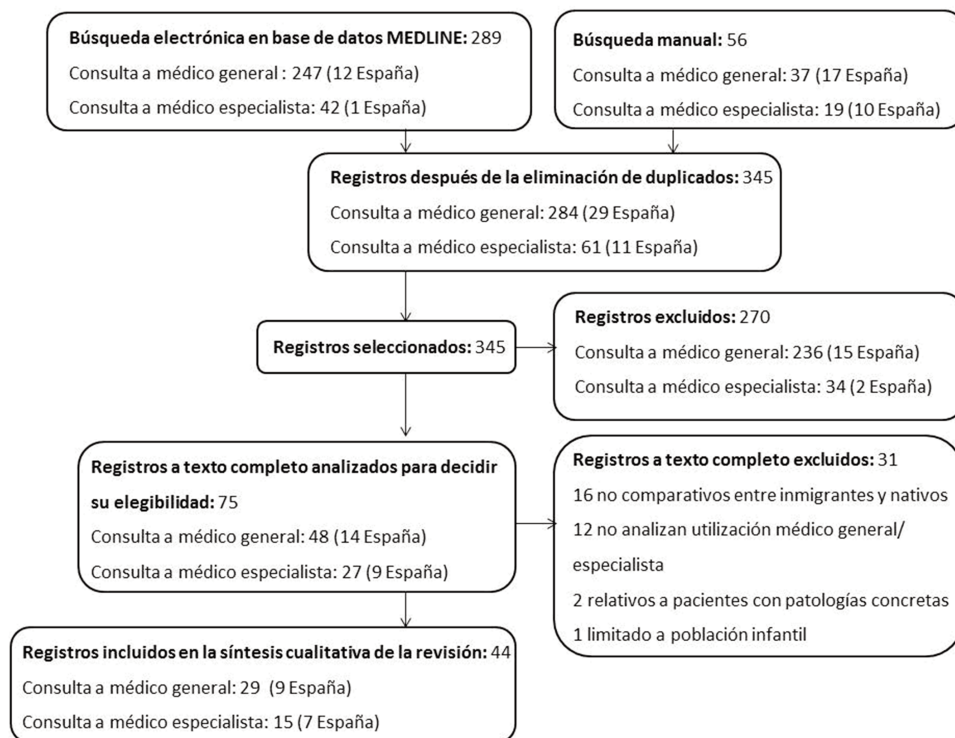
Las tablas 1 y 2 recogen los trabajos de investigación que en otros países analizaron las consultas realizadas en medicina general y especializada, respectivamente, por inmigrantes y autóctonos.

Fueron estudios anglosajones y del norte de Europa e Italia, que utilizaron como fuente de información la proporcionada por el propio individuo a través de encuestas de salud principalmente (análisis transversal) llevadas a cabo entre 1990 y 2007 (excepto un estudio¹⁴ que utilizó información del año 1985) y cuya variable dependiente (número de consultas/visitas a medicina general/especializada) varió en cuanto a la referencia temporal entre el mes anterior y los últimos doce meses. La pregunta se formuló tanto de forma continua (especificar número de consultas) como binaria (si tuvo o no algún contacto previo con medicina general).

Los autores de los estudios estimaron resultados a nivel global (5 estudios de 21 encontrados), esto es, comparando la población inmigrante definida previamente frente a la nativa del país correspondiente y/o desglosando por alguno de los criterios siguientes: inmigrante de primera o segunda generación (2), tiempo en el país de residencia (3), raza o etnia (11) y país de nacimiento (3).

Las publicaciones incluidas en la revisión compararon la utilización de los servicios de medicina general y especializada por ambos colectivos a través de análisis descriptivos (contraste de medias y porcentajes) y de modelos de regresión logística

Figura 1
Búsqueda bibliográfica de los estudios incluidos en la revisión sistemática



binaria ajustando los resultados por variables de necesidad (salud percibida, enfermedades crónicas), posición socioeconómica (educación, ocupación, ingresos) o por ambas. Otras variables de ajuste de los modelos fueron el estado civil, la limitación de la actividad diaria, ser fumador, ser bebedor, realizar ejercicio físico, tiempo de residencia en el país y cobertura sanitaria.

De esta búsqueda con datos de otros países se obtuvo, en líneas generales, que la población inmigrante hace igual uso de los servicios de medicina general y un menor o igual uso de las especialidades.

Los inmigrantes de origen marroquí en Italia y Ámsterdam, los chipriotas e irlandeses en Reino Unido, los nativos no blancos y los inmigrantes con mayor tiempo de resi-

dencia en Canadá y los inmigrantes de primera generación en Alemania, presentaron mayor utilización de los servicios de medicina general con respecto a los nativos en Italia, Ámsterdam y Reino Unido, nativos blancos y de menor tiempo de residencia en Canadá, y nativos e inmigrantes de segunda generación en Alemania, respectivamente. Mientras que asiáticos y nacionalizados de primera generación en Italia, no blancos mayores de 65 años y nativos de EEUU no blancos en EEUU, negros nacidos fuera de EEUU, afroamericanos e hispanos en EEUU, presentaron una menor utilización de los servicios de medicina general con respecto a los nativos de Italia, nativos blancos de EEUU, blancos o negros de EEUU y blancos americanos de EEUU, respectivamente.

Tabla 1

Estudios que analizan las diferencias en el uso de servicios de medicina general entre inmigrantes y autóctonos en el contexto internacional

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
De Luca et al ²¹ , 2013	Italia 2004-2005	102857 ≤ 64 años	Transversal (Encuesta de Salud)	Consulta al médico general (sí/no) y n.º de visitas en las últimas 4 semanas.	Nacionalidad (nativos, inmigrante de 1 y 2 generación, nacionalizados), país de nacimiento de los inmigrantes de 1 generación.	Salud percibida, enfermedades crónicas, discapacidad, fumador, control del peso, actividad deportiva	Riqueza del hogar (subjética), educación, ocupación	Inmigrantes menor probabilidad (no significativa) de consultar al médico general vs nativos. Marroquíes y asiáticos de 1 generación de inmigrantes presentaron una mayor/menor probabilidad vs nativos. Igual frecuencia de consultas entre inmigrantes y nativos. Los nacionalizados e inmigrantes asiáticos 1 generación menor frecuencia que los nativos.
Muggah et al ²² , 2012	Canadá 2005-2006	5269 18 años	Transversal (Encuesta)	N.º de visitas al médico general en el último año	País de nacimiento (no nacido en Canadá)	Salud percibida, número de años como paciente	No	Los inmigrantes que llevan <5 años en Canadá presentaron mayor media de consultas vs nativos y que los de mayor tiempo. Inmigrantes mayor o igual utilización de servicios médico general tras ajuste.
Glaesmer et al ²³ , 2011	Alemania 2007	2510 14-93 años	Transversal (Encuesta)	Consulta al médico general último año y número de visitas	País de nacimiento	No	No	La primera generación de inmigrantes utilizó con mayor frecuencia los servicios de médico general que la población nativa y que la segunda generación de inmigrantes.
Setia et al ¹⁹ , 2011	Canadá 1994-2006	7272 18 años	Longitudinal (Encuestas de Salud)	Consulta al médico general	País de nacimiento (no nacido en Canadá) y raza	Enfermedades crónicas	Educación, ingresos	No hubo diferencias estadísticamente significativas en las consultas al médico general entre los inmigrantes (blancos y no blancos) vs nativos por sexo.
Lum et al ²⁴ , 2010	EEUU 1992	7345 70 años	Transversal (Encuesta)	Consulta al médico general los últimos 12 meses	País de nacimiento (no nacido en EEUU), raza y etnia	No	Educación, ingresos	No hubo diferencias estadísticamente significativas en la utilización del médico general entre autóctonos e inmigrantes.
Prus et al ²⁵ , 2010	Canadá y EEUU 2002	8688 45 años	Transversal (Encuesta de Salud)	Acceso a un médico general	País de nacimiento (no nacido en Canadá) y raza	No	Educación, ingresos, cobertura sanitaria	Los inmigrantes no blancos 65 años en EEUU presentaron 4 veces más probabilidad de no tener acceso al médico general vs blancos nativos. No diferencias significativas para otras edades, inmigrantes blancos ni en Canadá.

Tabla 1 (continuación)

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
Newbold ¹⁸ , 2009	Canadá 1994-2001	8409 20 años	Longitudinal Encuesta Salud (análisis transversal para 1994-95 y 2000-01)	Consulta al médico general el año previo	País de nacimiento (no nacido en Canadá) y raza	Fumador, bebedor	Educación, ingresos	No hay diferencias significativas en la utilización del médico general entre población inmigrante (no nacida en Canadá) y nativa.
Peters et al ²⁶ , 2009	Inglaterra -	516 16 años	Transversal (Encuesta)	Consulta al médico general año previo	Grupo étnico	No	No	No diferencias significativas entre africanos/caribeños ni musulmanes/paquistaníes vs autóctonos.
Hargreaves S ²⁷ , 2006	Londres	1611 Todas edades	Transversal	Contacto permanente con el médico general	País de nacimiento y Nacionalidad	No	No	No diferencias entre británicos e inmigrantes (de países que generan refugiados políticos).
Lasser et al ²⁸ , 2006	Canadá y EEUU 2002	8688 18 años	Transversal (Encuesta de Salud)	Consulta al médico en último año	País de nacimiento y raza	No	Ingresos	Menor uso en inmigrantes vs población nativa, significativo en EEUU pero no en Canadá. No significativo entre razas.
Quan et al ²⁹ , 2006	Canadá 2001	121312 12 años	Transversal (Encuesta de Salud de Canadá)	Consulta al médico general último año	Raza, país de nacimiento	Salud percibida, enfermedades crónicas	Educación, ingresos	Mayor uso del médico general por parte de minorías étnicas con respecto a población blanca.
Newbold ¹² , 2005	Canadá 1994-2001	1305 Todas edades	Longitudinal Encuesta Salud 1994-1995 2000-1 (transversal 1994-95 y 2000-01)	Consulta al médico general último año	País de nacimiento	No	No	Mayor uso del médico general en los inmigrantes que en la población autóctona. Este uso fue mayor cuanto mayor fue el tiempo de residencia del inmigrante.
Lucas et al ³⁰ , 2003	EEUU 1997-2000	97345 Hombres 18 años	Transversal (Encuesta de Salud)	Tiempo desde la última consulta al médico general	Grupo étnico (autoreportado) y país de nacimiento	No	No	Menor uso del médico en sujetos negros nacidos fuera de EEUU que en blancos o negros nacidos en EEUU.

Tabla 1(continuación)

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
Livingston et al ³¹ , 2002	Londres	1085 65 años	Transversal (Encuesta)	Consulta médico general últimos 3 meses	País de nacimiento	Problemas de salud, Limitación actividades de la vida diaria	No	Mayor uso del médico general en sujetos de origen chipriota e irlandés respecto a los nacidos en el Reino Unido.
Stronks et al ³² , 2001	Amsterdam 1992-1993	2012 16-64 años	Transversal (enlace registros)	Consulta médico general últimos 2 meses	País de nacimiento	Salud percibida, enfermedad crónica	Educación, ocupación, ingresos	Mayor uso del médico en inmigrantes, sólo significativo en marroquíes.
Hjern et al ³³ , 2001	Suecia 1996	4342 27-60 años	Transversal (Encuesta)	Consulta al médico 3 meses anteriores	País de nacimiento	Salud percibida	Educación, ocupación	No hubo diferencias entre suecos e inmigrantes al ajustar por nivel socioeconómico y necesidad.
Hargraves JL et al ³⁴ , 2001	EEUU 1996-1997	41927 ≤ 65 años	Transversal (Encuesta hogares)	Consulta al médico último año	Raza	No	No	Afro-Americanos e Hispanos menor uso del médico general que los blancos Americanos.
Laroche M ¹⁴ , 2000	Canadá 1985/1991	13124 15 años	Transversal (Encuesta)	N de visitas al médico general en el último año	País de nacimiento	No	Ingresos, educación ocupación	Menor uso del médico general en inmigrantes.
Reijneveld ³⁵ , 1998	Amsterdam -	3296 16-64 años	Transversal (Encuesta)	Consulta al médico general en los últimos 2 meses	País de nacimiento (grupo étnico)	Salud percibida	Educación, ocupación, ingresos	No diferencias en uso del médico al ajustar por necesidad y nivel socioeconómico.
Leclerc et al ³⁶ , 1994	EEUU 1990	18 años	Transversal (Encuesta de Salud)	Número de contactos con el médico general último año.	Raza y País de nacimiento	Enfermedad crónica	Ingresos, educación	Menor uso del médico en la población inmigrante que la autóctona. Éstas se redujeron al aumentar el tiempo de residencia. Los sujetos de raza blanca utilizaron más el médico general.

Tabla 2**Estudios que analizan las diferencias en el uso de servicios de medicina especializada entre inmigrantes y autóctonos en el contexto internacional**

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
De Luca et al ²¹ , 2013	Italia 2004-2005	102857 ≤ 64 años	Transversal (Encuesta de Salud)	Consulta al médico especialista (sí/no) y n de visitas en las últimas 4 semanas	Nacionalidad (nativos, inmigrante de 1 y 2 generación, nacionalizados), país de nacimiento de los inmigrantes de 1 generación	Salud percibida, enfermedades crónicas, discapacidad, fumador, control del peso, actividad deportiva	Riqueza del hogar (subjetiva), educación, ocupación	Menor probabilidad de tener una consulta con el médico especialista entre los inmigrantes vs nativos. Menor probabilidad entre los inmigrantes de 1 y 2 generación vs nativos. Menor probabilidad de consultar entre los inmigrantes de 1 generación procedentes de Albania, Rumanía, no UE, otros africanos, Asia. Inmigrantes igual frecuencia de visitas vs nativos. Inmigrantes de 1 y 2 generación e inmigrantes de 1 generación de Marruecos y Rumanía, menor frecuencia vs nativos.
Glaesmer et al ²³ , 2011	Alemania 2007	2510 14-93 años	Transversal (Encuesta)	Consulta al médico especialista último año	País de nacimiento	No	No	La primera generación de inmigrantes menor utilización del médico especialista que la segunda generación de inmigrantes y que los nativos, controlando por edad y sexo.
Quan et al ²⁹ , 2006	Canadá 2001	121312 12 años	Transversal (Encuesta de Salud de Canadá)	Consulta al médico especialista último año	Raza y País de nacimiento	Salud percibida, Enfermedades crónicas	Educación, ingresos	No hubo diferencias significativas entre blancos y minorías étnicas en cuanto a la utilización de los servicios del médico especialista.
Burt et al ³⁷ , 2004	EEUU 1999-2000	295221 -	Transversal (Encuesta Salud)	Consulta al médico especialista	Raza	Número de enfermedades	Educación, ingresos, cobertura sanitaria	Los individuos de raza negra o Afro-Americanos menor uso de los servicios del médico especialista que los de raza blanca.
Livingston et al ³¹ , 2002	Londres (Islington)	1085 Todas las edades	Transversal (Encuesta)	Consulta al médico especialista	País de nacimiento	Salud percibida	No	Los individuos de origen Chipriota tuvieron un mayor uso de los servicios del médico especialista que los nativos nacidos en el Reino Unido.

Tabla 2
(continuación)

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
Hargraves JL et al ³⁴ , 2001	EEUU 1996-1997	41927 ≤ 65 años	Transversal (Encuesta hogares)	Consulta al médico especialista último año	Raza	No	No	Hispanos menor uso del médico especialista que Afro-Americanos y blancos Americanos.
Stronks et al ³² , 2001	Amsterdam 1992-1993	2012 16-64 años	Transversal (enlace registros)	Médico especialista últimos dos meses	País de nacimiento	Salud percibida, enfermedad crónica	Educación, ocupación, ingresos	Menor uso estadísticamente significativo de este tipo de servicios en los inmigrantes respecto a los nativos holandeses.
Laroche M ¹⁴ , 2000	Canadá 1985/1991	13124 15 años	Transversal (Encuesta)	Número visitas especialista en el año anterior	País de nacimiento	No	Ingresos, educación, ocupación	No hubo diferencias entre inmigrantes y población nativa.

De los 20 estudios que analizaron el uso de los servicios del médico general, 14 estudios ajustaron los resultados por alguna variable de necesidad y/o posición socioeconómica (2 estudios por variables de necesidad, 4 por variables de nivel socioeconómico y en 8 por ambas). En cuanto al uso de las especialidades hubo 8 estudios, de los cuales 6 proporcionaron resultados ajustados (1 por variables de necesidad, 1 por las variables de nivel socioeconómico y 4 por ambas). Las conclusiones de los estudios que no proporcionaron ningún tipo de ajuste no variaron con respecto a los que sí lo hicieron.

Las tablas 3 y 4 recogen los trabajos de investigación realizados en España que analizaron las diferencias en el uso de las consultas en medicina general y especializada entre inmigrantes y autóctonos respectivamente.

La fuente de información en general fue común en todos los artículos: Encuesta Nacional de Salud o Encuestas regionales (Comunidad de Madrid, Cataluña, Canarias, Comunidad Valenciana), excepto en el trabajo de López et al.⁹ que fusionó dos bases de datos de registros administrativos procedentes del Servicio Murciano de Salud, en el trabajo de Soler et al.¹⁵ que utilizó la información de los pacientes atendidos por 15 médicos de atención primaria en 5 áreas básicas de salud de la ciudad de Lleida de marzo a agosto de 2005, y en el trabajo de Muñoz et al.¹⁶ que obtuvo la información de los registros electrónicos de la población atendida entre 2005 y 2008 en centros sanitarios de atención primaria de una comarca de Barcelona.

La variable dependiente se analizó desde distintas perspectivas: si consultó o no en medicina general/especializada, frecuencia en medicina general/ especializada y tiempo desde la última consulta en medicina general, variando en cuanto a la referencia temporal entre las dos últimas semanas y los

cinco últimos años en los estudios relativos a medicina general y en el último año en los relativos a especializada.

Se encontraron 10 estudios realizados en España de los cuales 6 comparaban a la población inmigrante de manera global *versus* la autóctona, 3 que proporcionaban resultados desglosados según el país de nacimiento del sujeto, 2 según país de origen (uno de ellos consideraba solo los de renta media/baja), 2 según nacionalidad y 1 según tiempo de residencia en España de los inmigrantes que habían nacido en países de renta media/baja.

La metodología utilizada consistió en análisis descriptivos^{5,11, 15-17, 38-41} (contraste de medias y porcentajes), descomposición de Oaxaca^{9, 38}, modelos de regresión binomial⁵, binomial negativa³⁹, Poisson multinivel¹⁶, logística binaria^{17,38,40} y multinomial¹⁵. Los modelos se ajustaron por las mismas variables que las de los estudios realizados en otros países, incluyendo además la clase social, metros cuadrados de la vivienda, región de residencia, médicos por mil habitantes de la provincia de residencia, tamaño del municipio de residencia, y nunca se tuvo en cuenta el estado civil.

Así, de los 9 estudios que analizaron el uso de los servicios del médico general, 6 de ellos ajustaron sus resultados (1 estudio por variables de necesidad¹⁶ y 5 por variables de necesidad y posición socioeconómica^{5,17, 38-40}). En cuanto a las especialidades, los 5 estudios de 7 que proporcionaron resultados ajustados lo hicieron por variables de necesidad y nivel socioeconómico^{5,17, 38-40}. Los estudios que no ajustaron por ninguna de estas variables mostraron conclusiones en el mismo sentido que los que sí lo hicieron.

En líneas generales, los estudios indicaron ausencia de diferencias estadísticamente significativas entre ambas poblaciones en cuanto al uso de los servicios de medicina general y un menor uso de los servicios de

Tabla 3

Estudios que analizan las diferencias en el uso de servicios de medicina general entre inmigrantes y autóctonos en el contexto nacional

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición socioeconómica	Principales resultados
Muñoz et al ¹⁶ , 2012	España 2008	63257 15 años	Transversal (Registro electrónico de población atendida en AP de una comarca de Barcelona)	Número de consultas solicitadas con el médico general	País de origen	Número de problemas de salud	No	De forma general, mayor utilización de los servicios del médico general por la población inmigrante (mayor utilización: Magreb, resto de África y América Latina; menor utilización: Asia y Europa del Este).
Sanz et al ¹⁷ , 2011	España 2006	26728 16-74 años	Transversal (Encuesta Nacional de Salud 2006)	Consulta al médico general últimas 4 semanas	País de nacimiento	Salud percibida, número de enfermedades crónicas	Educación, clase social, ingresos, tipo de cobertura sanitaria	No hubo diferencias estadísticamente significativas entre los hombres inmigrantes y los autóctonos, excepto para Europa del Este con una menor utilización que los españoles. Las mujeres de África Sub-sahariana tuvieron mayor uso que las españolas.
Jiménez et al ¹⁸ , 2011	España 2003 y 2006	51086 16 años	Transversal (Encuesta Nacional de Salud 2003 y 2006)	Número de visitas al médico general en últimas 2 (ENS-2003) y 4 semanas (ENS-2006)	Nacionalidad	Salud autopercebida, limitación, enfermedades crónicas, accidente	Ingresos, educación, ocupación, cobertura sanitaria	Inmigrantes menor probabilidad de utilizar médico general que españoles.
Aerny et al ¹¹ , 2010	España 2007	12190 16 años	Transversal (Encuesta Regional de Salud de la CM)	Ha consultado al médico general en las últimas: 2 semanas, un mes, año, 5 años, más de 5 años/ nunca.	País de nacimiento	No	No	En ambos sexos, los que llevan menos tiempo de residencia en España acuden con menos frecuencia que los autóctonos, mientras que los de mayor tiempo igualan e incluso superan la frecuentación de los nacidos en España.
Muñoz de Bustillo et al ¹⁹ , 2010	España 2006	24206 16 años con única cobertura sanitaria la Seguridad Social	Transversal (Encuesta Nacional de Salud 2006)	Número de visitas al médico general en el último mes	País de nacimiento	Enfermedades crónicas, accidente el último año	Educación, ocupación	Ausencia de diferencias significativas entre inmigrantes latinoamericanos y la población local en relación con las visitas al médico de familia.

Tabla 3
(continuación)

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
Hernández et al ⁴⁰ , 2009	España 2003 y 2006	51086 16 años	Transversal (Encuesta Nacional de Salud 2003 y 2006)	Número de visitas al médico general en últimas 2 y 4 semanas	Nacionalidad	Salud autopercebida, limitación, enfermedades crónicas, accidente	Ingresos, educación, ocupación, cobertura sanitaria	Africanos mayor probabilidad de utilizar médico general que españoles. Europeos y Norte Americanos menor probabilidad de utilizar médico general que españoles.
Regidor et al ⁵ , 2009	España 2004 - 2006	29379 16 años	Transversal (Encuestas de Salud de Cataluña 2006, Ciudad de Madrid 2005, Canarias 2004, Comunidad Valenciana 2005)	Consulta al médico general (Encuesta Salud Cataluña y Madrid: últimos 15 días; Canarias y C. Valenciana: últimos 12 meses)	País de nacimiento y nacionalidad (si no se recogía el país de nacimiento)	Percepción subjetiva de la salud, limitación de la actividad diaria por problema de salud en el último año	Clase social, educación	El uso de los servicios de médico general en conjunto de población inmigrante es muy similar o ligeramente inferior al de la población española. La población africana presenta el patrón más heterogéneo: en Madrid utiliza más que la española el médico general, Cataluña menos, y Canarias y C. Valenciana es similar.
Soler et al ¹⁵ , 2008	España 2005	5755 Todas edades	Transversal (Población atendida en AP)	Frecuentación al médico general últimos 6 meses	País de origen de renta baja/media	No	No	Los inmigrantes mayor probabilidad de realizar de 3 a 11 visitas a AP vs 1 ó 2 de nativos. Latinoamericanos y magrebíes mayor probabilidad de realizar más visitas vs autóctonos.
Carrasco-Garrido et al ⁴¹ , 2007	España 2003	1506 16 años	Transversal (Encuesta Nacional de Salud 2003)	Tiempo desde la última consulta al médico general	Nacionalidad	No	No	No se encontraron diferencias entre la frecuencia de visitas al médico general de la población inmigrante y la española si bien los resultados no se ajustaron por ninguna variable sociodemográfica ni de necesidad.

Tabla 4

Estudios que analizan las diferencias en el uso de servicios de medicina especializada entre inmigrantes y autóctonos en el contexto nacional.

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
Sanz et al ¹⁷ , 2011	España 2006	26728 16-74 años	Transversal (Encuesta Nacional de Salud 2006)	Consulta al médico especialista últimas 4 semanas	País de nacimiento	Salud percibida, número de enfermedades crónicas	Educación, clase social, ingresos, tipo de cobertura sanitaria	Los hombres inmigrantes utilizan menos los servicios del médico especialista que los nativos, excepto los de frica Sub-sahariana (mayor utilización). Las mujeres inmigrantes utilizan menos (Latino-américa, frica del Norte) o igual el médico especialista, excepto las de frica Sub-sahariana (mayor utilización).
Jiménez et al ³⁸ , 2011	España 2003 y 2006	51086 16 años	Transversal (Encuesta Nacional de Salud 2003 y 2006)	Número de visitas al médico especialista en últimas 2 y 4 semanas	Nacionalidad	Salud autopercebida, limitación, enfermedades crónicas, accidente	Ingresos, educación, ocupación, cobertura sanitaria	Inmigrantes menor probabilidad de utilizar médico especialista que los españoles.
Muñoz-de Bustillo et al ³⁹ , 2010	España 2006	24206 16 años con única cobertura sanitaria la Seguridad Social	Transversal (Encuesta Nacional de Salud 2006)	Número de visitas al médico especialista en el último mes	País de nacimiento	Enfermedades crónicas, accidente el último año	Educación, ocupación	Ausencia de diferencias significativas entre inmigrantes latino-americanos y la población local en relación con las visitas al médico especialista.
Hernández et al ⁴⁰ , 2009	España 2003 y 2006	51086 16 años	Transversal (Encuesta Nacional de Salud 2003 y 2006)	Número de visitas al médico especialista en últimas 2 y 4 semanas	Nacionalidad	Salud autopercebida, limitación, enfermedades crónicas, accidente	Ingresos, educación, ocupación, cobertura sanitaria	Los inmigrantes presentan menor probabilidad de visitar al médico especialista que los españoles. Los procedentes de América Latina, Asia y UE tienen menor probabilidad de visitar al médico especialista que los españoles.
López et al ¹⁹ , 2009	España 2005	1333810 Todas las edades	Transversal (fusión de dos ficheros de registros administrativos del Servicio Murciano Salud)	Número de consultas externas último año	Nacionalidad	No	No	Las tasas de utilización de las consultas externas de los autóctonos superan las correspondientes tasas de los inmigrantes. Igual ocurre al estratificar por sexo, edad y nacionalidad.

Tabla 4
(continuación)

Referencia	Lugar y año	Población estudiada	Tipo de estudio	Variable dependiente	Variable independiente	Control necesidad	Control posición Socioeconómica	Principales resultados
Regidor et al ⁵ , 2009	España 2004-2006	29379 16 años	Transversal (Encuesta de salud de Cataluña 2006, Ciudad de Madrid 2005, Canarias 2004, C. Valenciana 2005)	Consulta al médico especialista (sí/no) (Encuesta de Salud de Cataluña y Madrid: últimos 15 días; Canarias y C. Valenciana: últimos 12 meses)	País de nacimiento y nacionalidad (si no se recogía el país de nacimiento)	Percepción subjetiva de la salud, limitación de la actividad diaria por problema de salud en el último año	Clase social, educación	El uso de los servicios de médico especialista en el conjunto de la población inmigrante es inferior al de la población española en todos los ámbitos geográficos, excepto los procedentes de América Central y del Sur en la C. Valenciana.
Soler et al ¹⁵ , 2008	España 2005	5755 Todas edades	Transversal (población atendida en AP)	Frecuentación al médico especialista últimos 6 meses	País de origen de renta baja/media	No	No	Los inmigrantes mayor probabilidad que los autóctonos de realizar de 1 a 3 visitas al médico especialista.

especializada por la población inmigrante. Aunque al tener en cuenta el tiempo de estancia en el país, cuando era menor se asoció a menor frecuentación en medicina general con respecto a los nativos, mientras que a medida que se incrementaba el tiempo se igualaba e incluso superaba la frecuentación de los autóctonos. También se hallaron diferencias entre los sujetos procedentes de América Latina y África con respecto a los autóctonos, con una mayor utilización por parte de los primeros. Destacó un estudio en el que las mujeres de África Subsahariana presentaban una probabilidad muy superior a las nativas de utilizar los servicios de medicina general¹⁷ y otro en el que esta población presentó un patrón muy heterogéneo, variando en función de la comunidad autónoma de residencia⁵. Concretamente, en Cataluña se observó una menor frecuentación, en Madrid mayor frecuentación y en Canarias y Comunidad Valenciana fue similar a la de la población nativa. Se observó una menor frecuentación en medicina general por los inmigrantes procedentes de Asia, Europa del Este y América del Norte.

Sanz et al.¹⁷ y Regidor et al.⁵ hallaron una mayor utilización de los servicios de especialidades por parte de ciertos grupos de inmigrantes. Los de África Subsahariana, tanto hombres como mujeres, considerando la población española como referencia¹⁷ y los de América Central y del Sur en la Comunidad Valenciana⁵. Además, Soler et al.¹⁵ hallaron una mayor probabilidad de realizar más visitas a los especialistas en los inmigrantes que en los autóctonos, aunque estos resultados no se ajustaron por ninguna variable.

Globalmente, existe una discrepancia moderada respecto a los principales ítems evaluados para determinar la calidad de los estudios incluidos en la revisión. Así, tanto las variables de respuesta, la definición de inmigrante y las variables de ajuste consideradas quedaron claramente descritas en los estudios, aunque existió heterogeneidad

entre ellos por analizar diferentes tipos de variables. Por otra parte, los estudios basados en encuestas en general proporcionaron tasas globales de respuesta del cuestionario, pero sin desglosarlas en función de las variables de interés principal, por lo que no se dispone de información sobre ellas. Tampoco se indicó si se llevó a cabo algún método estadístico para el tratamiento de los datos ausentes. El tamaño de la muestra para ciertos colectivos de inmigrantes es poco representativo, ofreciéndose estimaciones con intervalos de confianza amplios afectando a la precisión y comparabilidad de los resultados.

DISCUSIÓN

En líneas generales, tantos los estudios realizados en España como los realizados en otros países, muestran que, con respecto a la población autóctona, la población inmigrante realiza una utilización similar de los servicios de medicina general, pero un menor uso de los servicios de las especialidades.

Es importante señalar que en la mayoría de las investigaciones sobre utilización de los servicios sanitarios por parte de la población inmigrante predomina la falta de explicaciones a los resultados obtenidos. Esto lleva a los autores incluidos en esta revisión a plantear como hipótesis la existencia de otros factores que no han podido controlar, debido a la ausencia de información y que podrían estar afectando el patrón de utilización de los servicios de medicina general y especializada. Entre ellos hay que destacar a los siguientes:

a) Factores geográficos y culturales: en España, la proximidad cultural de los inmigrantes procedentes de América Latina puede que condicione la mayor frecuencia de consultas a medicina general, mientras que en los de Asia y Europa del Este las diferencias culturales pueden suponer una menor frecuentación. En cuanto a la pobla-

ción africana, la mayor utilización puede ser debida a mayores problemas de salud los cuales no se han recogido adecuadamente con las variables analizadas. En otros países se ha señalado que pueden existir igualmente diferencias culturales en cuanto a la percepción del estado de salud y que colectivos como los asiáticos pueden estar recurriendo a otras formas de atención de la salud (medicina tradicional, apoyo familiar)¹⁸. Otros factores, como la imposibilidad de ser atendido por un médico que entienda el idioma puede conducir a estas desigualdades, principalmente en atención especializada¹⁸.

b) Diferencias socioeconómicas: las condiciones laborales de los inmigrantes a causa de su peor cualificación puede conducir a determinados colectivos, como los asiáticos, a la pérdida de las citas previas, explicando así la menor frecuentación tanto de los servicios de medicina general como especializada.

c) Desconocimiento de la organización del sistema sanitario del país de acogida.

d) Tiempo de residencia: un mayor tiempo en el país puede asociarse a mayor integración y conocimiento del sistema sanitario¹⁹. Además, con el tiempo la población inmigrante puede presentar peor salud autopercebida y mental, lo que puede representar mayor necesidad y utilización de los servicios sanitarios¹⁸. Este puede ser un factor que explique las diferencias halladas en los resultados entre los distintos autores.

En cuanto a las limitaciones de esta investigación hay que tener en cuenta que se excluyeron estudios que no eran comparativos por lo que se pudo perder información relevante del colectivo inmigrante no recogida en los estudios de la revisión y que posiblemente ayuden a explicar algunos de los resultados obtenidos en los mismos.

La búsqueda bibliográfica se llevó a cabo solo en una base de datos, Medline, aunque se completó con una búsqueda manual y se incluyeron trabajos que procedían de una revisión anterior. La estrategia de búsqueda llevada a cabo cubrió prácticamente la totalidad de los artículos que habían sido seleccionados en la revisión anterior.

No se tuvieron en cuenta los estudios de especialidades concretas en la revisión, aunque las estrategias de búsqueda quedaron abiertas a este tipo de estudios, aunque se descartaron posteriormente debido tanto a sus características como al elevado número de estudios hallados de diferentes especialidades, cuya inclusión podría implicar tal grado de heterogeneidad entre los mismos que afectara a la comparabilidad y consistencia de los resultados.

Tampoco hay que olvidar el posible sesgo de representatividad en los estudios que utilizan como fuente de información encuestas de salud. En ellos se ha observado una baja representación de la población inmigrante, variando entre el 5 y el 15%, presentando además ciertos colectivos de inmigrantes un pequeño tamaño de muestra, lo que podría sesgar los resultados al compararlos con los nativos. También hay que tener en cuenta que en dichas encuestas de salud la variable operativa para clasificar a los individuos no es homogénea. Así, en España se ha considerado tanto el país de nacimiento como la nacionalidad (según el año de realización) mientras que en otros países identificaron a los sujetos inmigrantes a través del país de nacimiento e incorporaron otras variables de clasificación tales como la raza y la etnia. Además, las encuestas de salud que utilizan el padrón municipal de habitantes como marco muestral para seleccionar al informante, dejan fuera a los inmigrantes que se encuentran en situación irregular¹¹. También se ha observado una infrarrepresentación de la población inmigrante con menos recursos económicos en algunas encuestas de salud³⁸.

Los estudios incluidos en la revisión analizaron poblaciones de países con sistemas sanitarios diferentes, por lo que es difícil llevar a cabo una comparación sobre la utilización de los servicios sanitarios. Además, dichos estudios están condicionados por el diseño y el cuestionario utilizados en la mayoría de ellos, por lo que el uso de los servicios sanitarios queda fundamentalmente limitado a lo que considera el entrevistado cuando se le pregunta si va al médico de familia/médico general *versus* especialista.

Los estudios revisados se caracterizan, por tanto, por la heterogeneidad existente entre los mismos, principalmente debido a las diferentes metodologías empleadas. Sería recomendable que estos estudios mejoraran la calidad general mediante la inclusión de información más completa, tanto descriptiva como metodológica, sobre los datos ausentes de las variables de interés y que consideraran muestras más representativas tanto de la población inmigrante como de ciertos colectivos.

En cuanto a las limitaciones de los resultados de los estudios incluidos en la revisión hay que tener en cuenta principalmente la subjetividad que conlleva trabajar con encuestas, donde las respuestas de los entrevistados están basadas en lo que recuerdan en el momento de su realización.

Solo se ofrecieron resultados desagregados por sexo en 3 de los 10 estudios realizados en España, de manera que la influencia de esta variable se ha podido ver enmascarada al dar resultados globales. Por otro lado, en la mayoría de los estudios se utilizó a toda la población sin tener en cuenta la estructura de edad de la población inmigrante. En muchos estudios era población mayoritariamente en edad laboral, por lo que la comparación con el conjunto de la población autóctona puede distorsionar los resultados.

Los trabajos incluidos en la revisión analizaron el uso de los servicios sanitarios desde diferentes perspectivas, principalmente de forma binaria (si se consultó o no) y como frecuentación al médico general/ especialista, lo cual podría influir en los resultados. Sin embargo, se ha comprobado que los estudios de otros países consideraron generalmente la variable dependiente de forma binaria (16 de 20 relativos al médico general y 7 de 8 relativos a las consultas de especialistas) mientras que los de España a través de la frecuentación (5 estudios de 9 relativos a medicina general y 5 estudios de 7 relativo a medicina especializada), sin que los resultados variaran en líneas generales con respecto a la minoría de estudios que la analizaron desde la otra óptica.

En los estudios incluidos en la revisión predomina la falta de explicaciones a los resultados obtenidos, probablemente por limitaciones de las fuentes de información utilizadas. Estudios con otras fuentes de información, como los registros clínicos-administrativos, permitirían obtener mayor representatividad de ciertos colectivos de población inmigrante.

En conclusión, los resultados de la revisión indican que, tanto en España como en otros países, la población inmigrante hace igual uso de los servicios de medicina general, mientras que este es menor o igual en cuanto a las especialidades, aunque existan diferencias respecto a los sistemas sanitarios y los colectivos de inmigrante considerados. Con la entrada en vigor en España del nuevo Real Decreto-ley 16/2012²⁰ se deja de ofrecer atención sanitaria con carácter universal a los inmigrantes no regularizados, interponiendo barreras de acceso a los servicios sanitarios a los colectivos más vulnerables y marginados, de manera que los resultados aquí mostrados pueden variar en el futuro.

BIBLIOGRAFÍA

1. INE. Padrón municipal. Disponible en: <http://www.ine.es/prensa/np648.pdf>.
2. EUROSTAT. Disponible en: <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>.
3. Jansà JM. Immigration of foreigners into Spain. Public health considerations. *Rev Esp Salud Publica*. 1998;72(3):165-8.
4. Berra S, Elorza-Ricart JM, Bartomeu N, Hausmann N, Serra-Sutton V, Rajmil L. Necessitats en salut i utilització dels serveis sanitaris en la població immigrant a Catalunya. Revisió exhaustiva de la literatura científica. Barcelona: Agència d'Avaluació de Tecnologia i Recerca Mèdiques CatSalut Departament de Sanitat i Seguretat Social Generalitat de Catalunya; 2004.
5. Regidor E, Sanz B, Pascual C, Lostao L, Sanchez E, Diaz Olalla JM. Health services utilization by the immigrant population in Spain. *Gac Sanit*. 2009;23 Suppl 1:4-11.
6. Regidor E, Diaz Olalla JM, Lostao L. Diferencias en la utilización de servicios sanitarios entre la población inmigrante y la población española. Madrid: Fundación de Ciencias de la Salud; 2008.
7. Fuertes C, Martín Laso MA. The immigrant in primary care consultations. *An Sist Sanit Navar*. 2006;29 Suppl 1:9-25.
8. Berra S, Elorza-Ricart JM. Salud y uso de los servicios sanitarios en población inmigrante y autóctona de España. Colección: Informes, estudios e investigación. Ministerio de Ciencia e Innovación / Informes de Evaluación de Tecnologías Sanitarias; AATRM 2007/08 Barcelona: Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya; 2009.
9. López NA, Ramos Parreño JM. Utilización de servicios sanitarios por parte de las poblaciones inmigrante y nativa en la Región de Murcia. *Gac Sanit*. 2009;23 Suppl 1:12-8.
10. Díaz Olalla JM. Podrá la población inmigrante superar el impacto de nuestro nivel de salud? *Aten Primaria* 2005;36(Supl 1):38-9.
11. Aerny Perreten N, Ramasco Gutierrez M, Cruz Maccin JL, Rodríguez Rieiro C, Arabato Gonzalez S, Rodríguez Laso A. La salud y sus determinantes en la población inmigrante de la Comunidad de Madrid. *Gac Sanit* 2010;24(2):136-44.
12. Newbold B. Health status and health care of immigrants in Canada: a longitudinal analysis. *J Health Serv Res Policy*. 2005 Apr;10(2):77-83.
13. Valderas JM, Mendivil J, Parada A, Losada-Yanez M, Alonso J. Development of a geographic filter for PubMed to identify studies performed in Spain. *Rev Esp Cardiol* 2006;59(12):1244-51.
14. Larocche M. Health status and health services utilization of Canada's immigrant and non-immigrant populations. *Can Public Policy*. 2000;26(1):51-73.
15. Soler-Gonzalez J, Serna AC, Rue MM, Bosch GA, Ruiz Magaz MC, Gervilla CJ. Use of primary care resources by immigrants and the autochthonous persons who contact the care services in the city of Lleida, Spain. *Aten Primaria*. 2008;40(5):225-31.
16. Munoz MA, Pastor E, Pujol J, Del Val JL, Cordomi S, Hermosilla E. Primary health care utilization by immigrants as compared to the native population: a multilevel analysis of a large clinical database in Catalonia. *Eur J Gen Pract*. 2012;18(2):100-6.
17. Sanz B, Regidor E, Galindo S, Pascual C, Lostao L, Diaz JM, et al. Pattern of health services use by immigrants from different regions of the world residing in Spain. *Int J Public Health*. 2011;56(5):567-76.
18. Newbold KB. Health care use and the Canadian immigrant population. *Int J Health Serv* 2009;39(3):545-65.
19. Setia MS, Quesnel-Vallee A, Abrahamowicz M, Toussaint P, Lynch J. Access to health-care in Canadian immigrants: a longitudinal study of the National Population Health Survey. *Health Soc Care Community*. 2011;19(1):70-9.
20. Real Decreto-ley 16/2012, de 20 de abril. Disponible en: <http://www.boe.es/boe/dias/2012/04/24/pdfs/BOE-A-2012-5403.pdf>.
21. De Luca G, Ponzo M, Andrés AR. Health care utilization by immigrants in Italy. *Int J Health Care Finance Econ*. 2013;13(1):1-31.
22. Muggah E, Dahrouge S, Hogg W. Access to primary health care for immigrants: results of a patient survey conducted in 137 primary care practices in Ontario, Canada. *BMC Fam Pract*. 2012;28:13:128.
23. Glaesmer H, Wittig U, Brähler E, Martin A, Mewes R, Rief W. Health care utilization among first and second generation immigrants and native-born Germans: a population-based study in Germany. *Int J Public Health*. 2011;56(5):541-8.

24. Lum TY, Vanderaa JP. Health disparities among immigrant and non-immigrant elders: the association of acculturation and education. *J Immigr Minor Health*. 2010;12(5):743-53.
25. Prus SG, Tfaily R, Lin Z. Comparing racial and immigrant health status and health care access in later life in Canada and the United States. *Can J Aging*. 2010;29(3):383-95.
26. Peters J, Parry GD, Van CP, Moore J, Cooper CL, Walters SJ. Health and use of health services: a comparison between Gypsies and Travellers and other ethnic groups. *Ethn Health*. 2009;14(4):359-77.
27. Hargreaves S, Friedland JS, Gothard P, Saxena S, Millington H, Eliahoo J, et al. Impact on and use of health services by international migrants: questionnaire survey of inner city London A&E attenders. *BMC Health Serv Res*. 2006;6:153.
28. Lasser KE, Himmelstein DU, Woolhandler S. Access to care, health status, and health disparities in the United States and Canada: results of a cross-national population-based survey. *Am J Public Health* 2006;96(7):1300-7.
29. Quan H, Fong A, De Coster C, Wang J, Musto R, Noseworthy TW, et al. Variation in health services utilization among ethnic populations. *CMAJ*. 2006;14;174(6):787-91.
30. Lucas JW, Barr-Anderson DJ, Kington RS. Health status, health insurance, and health care utilization patterns of immigrant Black men. *Am J Public Health*. 2003;93(10):1740-7.
31. Livingston G, Leavey G, Kitchen G, Manela M, Sembhi S, Katona C. Accessibility of health and social services to immigrant elders: the Islington Study. *Br J Psychiatry*. 2002;180:369-73.
32. Stronks K, Ravelli AC, Reijneveld SA. Immigrants in the Netherlands: equal access for equal needs? *J Epidemiol Community Health*. 2001;55(10):701-7.
33. Hjern A, Haglund B, Persson G, Rosen M. Is there equity in access to health services for ethnic minorities in Sweden? *Eur J Public Health*. 2001;11(2):147-52.
34. Hargraves JL, Cunningham PJ, Hughes RG. Racial and ethnic differences in access to medical care in managed care plans. *Health Serv Res*. 2001;36(5):853-68.
35. Reijneveld SA. Reported health, lifestyles, and use of health care of first generation immigrants in The Netherlands: do socioeconomic factors explain their adverse position? *J Epidemiol Community Health*. 1998;52(5):298-304.
36. Leclere FB, Jensen L, Biddlecom AE. Health care utilization, family context, and adaptation among immigrants to the United States. *J Health Soc Behav*. 1994;35(4):370-84.
37. Burt CW, Schappert SM. Ambulatory care visits to physician offices, hospital outpatient departments, and emergency departments: United States, 1999-2000. *Vital Health Stat* 13. 2006;(159):1-66.
38. Jimenez-Rubio D, Hernandez-Quevedo C. Inequalities in the use of health services between immigrants and the native population in Spain: what is driving the differences? *Eur J Health Econ*. 2011;12(1):17-28.
39. Munoz-de BR, Anton Perez JI. Use of public health services by Latin American immigrants in Spain. *Salud Publica Mex*. 2010;52(4):357-63.
40. Hernandez-Quevedo C, Jimenez-Rubio D. A comparison of the health status and health care utilization patterns between foreigners and the national population in Spain: new evidence from the Spanish National Health Survey. *Soc Sci Med*. 2009;69(3):370-8.
41. Carrasco-Garrido P, De Miguel AG, Barrera VH, Jimenez-Garcia R. Health profiles, lifestyles and use of health resources by the immigrant population resident in Spain. *Eur J Public Health*. 2007;17(5):503-7.

Anexo 1

Estrategia de búsqueda para la localización de estudios comparativos sobre consultas realizadas al médico general

1. Health Services Needs and Demand/
2. Health Status/
3. Health Services Accessibility/
4. 1-3 / OR
5. health care[Title]
6. health disparities[Title]
7. access to care[Title]
8. health resources[Title]
9. health profiles[Title]
10. health status[Title]
11. health services[Title]
12. 5 11/ OR
13. 4 and 12
14. Emigration and Immigration/
15. Emigrants and Immigrants/
16. 14 or 15
17. immigrant*[Title]
18. migrant*[Title]
19. Ethnic Groups[Title]
20. 17-19 / OR
21. 16 and 20
22. Health AND utilization AND immigrant*[Title]
23. 13 AND 21
24. 22 or 23

Anexo 2

Estrategia de búsqueda para la localización de estudios comparativos sobre consultas realizadas al médico especialista

1. Health Services/utilization/
2. Health Services Accessibility/
3. Health Status/
4. 1-3 / OR
5. Emigration and Immigration/
6. Emigrants and Immigrants/
7. Ethnic Groups
8. 5-7 / OR
9. Specialization/
10. speciali*[TI]
11. 9 OR 10
12. 4 AND 8 AND 11

Anexo 3

Filtro geográfico específico para la obtención de los estudios relativos a España

(spain OR espagne OR espana OR spagna) OR (spain[tiab] OR espagne[tiab] OR espana[tiab] OR spanien[tiab] OR spagna[tiab]) OR (catalunya[tiab] OR catalonia[tiab] OR catalogne[tiab] OR cataluna[tiab] OR catala[tiab] OR barcelon*[tiab] OR tarragona[tiab] OR lleida[tiab] OR lerida[tiab] OR girona[tiab] OR gerona[tiab] OR sabadell[tiab] OR hospitalet[tiab] OR l hospitalet[tiab]) OR (valencia*[tiab] OR castello*[tiab] OR alacant[tiab] OR alicant*[tiab]) OR (murcia*[tiab] OR cartagen*[tiab] NOT indias[tiab])) OR (andalu*[tiab] OR sevilla*[tiab] OR granad*[tiab] OR huelva[tiab] OR almeria[tiab] OR cadiz[tiab] OR jaen[tiab] OR malaga[tiab] OR (cordoba[tiab] NOT argentin*[tiab])) OR (extremadura[tiab] OR caceres[tiab] OR badajoz[tiab] OR madrid[tiab]) OR (castilla[tiab] OR salamanca[tiab] OR zamora[tiab] OR valladolid[tiab] OR segovia[tiab] OR soria[tiab] OR palencia[tiab] OR avila[tiab] OR burgos[tiab]) OR (leon[tiab] NOT (france[tiab] OR clermont[tiab] OR rennes[tiab] OR lyon[tiab] OR USA[tiab] OR mexic*[tiab])) OR (galicia[tiab] OR gallego[tiab] OR compostela[tiab] OR vigo[tiab] OR corun*[tiab] OR ferrol[tiab] OR orense[tiab] OR ourense[tiab] OR pontevedra[tiab] OR lugo[tiab]) OR (oviedo[tiab] OR gijon[tiab] OR asturia*[tiab]) OR (cantabr*[tiab] OR santander[tiab]) OR (vasco[tiab] OR euskadi[tiab] OR basque[tiab] OR bilbao[tiab] OR bilbo[tiab] OR donosti*[tiab] OR san sebastian[tiab] OR vizcaya[tiab] OR bizkaia[tiab] OR guipuzcoa[tiab] OR gipuzkoa[tiab] OR alava[tiab] OR araba[tiab] OR vitoria[tiab] OR gasteiz[tiab]) OR (navarr*[tiab] OR nafarroa[tiab] OR pamplona[tiab] OR iruna[tiab] OR irunea[tiab]) OR (logron*[tiab] OR rioj*[tiab]) OR (aragon*[tiab] OR zaragoza[tiab] OR teruel[tiab] OR huesca[tiab]) OR (mancha[tiab] OR ciudad real[tiab] OR albacete[tiab] OR cuenca[tiab]) OR (toledo[tiab] NOT (ohio[tiab] OR us[tiab] OR usa[tiab] OR OH[tiab])) OR (guadalajara[tiab] NOT mexic*[tiab]) OR (balear*[tiab] OR mallorca[tiab] OR menorca[tiab] OR ibiza[tiab] OR eivissa[tiab]) OR (palmas[tiab] OR lanzarote[tiab] OR canari*[tiab] OR tenerif*[tiab]) OR (ceuta[tiab] OR melilla[tiab])) OR (osasunbide*[tiab] OR osakidetza[tiab] OR insalud[tiab] OR sergas[tiab] OR cat-salut[tiab] OR sespa[tiab] OR osasunbidea[tiab] OR imsalud[tiab] OR sescam[tiab] OR ib-salut[tiab])

REVISIÓN SISTEMÁTICA

EFECTIVIDAD DE DISTINTAS TERAPIAS FÍSICAS EN EL TRATAMIENTO CONSERVADOR DE LA FASCITIS PLANTAR. REVISION SISTEMATICA

Ana María Díaz López y Patricia Guzmán Carrasco.

Servicio de Rehabilitación. Hospital General Ntra. Sra. del Prado. Talavera de la Reina. Toledo.

No existen conflictos de interés

RESUMEN

Fundamentos: La fascitis plantar es la afección más frecuente de dolor no traumático en el tobillo-pie. Se presenta más en mujeres entre los 40- 70 años con comienzo progresivo y difuso en planta del pie o tobillo que poco a poco se agudiza impidiendo la marcha. El objetivo de este trabajo es determinar si las distintas terapias físicas utilizadas en el tratamiento conservador de la fascitis plantar de al menos un mes de evolución son efectivas individualmente y/o combinadas entre sí.

Métodos: Revisión sistemática en las bases de datos The Cochrane Library Plus, Medline, Lilacs, IBECs, IME, PEDro y Enfispo sin restricción de fecha, en español e inglés. Se incluyeron ensayos clínicos controlados aleatoriamente con personas adultas diagnosticadas de fascitis plantar, estudios de intervención, prospectivos y revisiones sistemáticas. La evaluación de la elegibilidad de los estudios fue desarrollada por dos revisores de manera independiente, estandarizada y no cegada. Para clasificarlos, se utilizó la escala crítica PEDro, formulario de calidad metodológica y una revisión crítica de cada resumen y si esto no fue concluyente evaluación del texto completo.

Resultados: Se revisó un total de 32 artículos a texto completo. Las técnicas más empleadas fueron los estiramientos y las ondas de choque, aunque los mejores resultados se obtuvieron combinando varias técnicas. Las ondas de choque fueron efectivas cuando otras técnicas fracasaron.

Conclusiones: Las terapias físicas empleadas en los distintos estudios han demostrado ser eficaces aunque en distinta medida ya sea para disminuir el dolor o aliviar los síntomas de la fascitis plantar.

Palabras clave: Fascitis plantar. Dolor. Terapéutica. Fisioterapia. Revisión sistemática.

Correspondencia
Ana María Díaz López.
Servicio de Rehabilitación.
Departamento de Fisioterapia.
Hospital General Ntra. Sra. Del Prado.
45600 Talavera de la Reina. TOLEDO.
adiazl@sescam.jccm.es

ABSTRACT

Effectiveness of Different Physical Therapy in Conservative Treatment of Plantar Fasciitis. Systematic Review

Background: Plantar fasciitis is the most common disease of non-traumatic pain in the ankle-foot. It is more common in women aged 40 - 70 years and diffuse progressive start the foot or ankle that gradually worsens preventing progress. The aim of this work is to determine whether different physical therapies used in the conservative treatment of plantar fasciitis of at least one month duration in adults are effective individually and / or in combination.

Methods: A systematic review databases in The Cochrane Library, Medline, Lilacs, IBECs, IME, PEDro and ENFISPO no date restriction, in Spanish and English languages. Randomized controlled trials were included of adult patients diagnosed with plantar fasciitis, intervention studies, prospective and systematic reviews. Assessment of study eligibility was developed by two reviewers independently and unblinded standardized. To classify, we used the PEDro scale critical, form of methodological quality plus a critical review of each summary and if this was not conclusive assessment of the full text.

Results: 32 full-text articles were reviewed. Most used techniques are the stretches and shock waves, although the best results are obtained by combining several techniques. Shock waves are effective when other techniques have failed.

Conclusion: Physical therapies used in the various studies have proven effective to varying degrees either to reduce pain or relieve the symptoms of plantar fasciitis.

Keyword: Systematic review. Plantar fasciitis. Pain. Therapeutics. Physical Therapy Modalities.

INTRODUCCIÓN

La fascitis plantar (FP) o talalgia plantar, se caracteriza por un dolor localizado en la zona antero-interna del calcáneo que puede irradiarse hacia el borde interno del pie¹. Es la causa más habitual de dolor en esa zona del cuerpo en las personas adultas^{2,3}. La fascia plantar es una aponeurosis fibrosa que proporciona un apoyo fundamental al arco longitudinal del pie⁴. La palabra "fascitis" sugiere la inflamación de la fascia plantar pero investigaciones recientes indican que no siempre la hay. La presencia de microrroturas en la fascia derivadas de traumatismos de repetición, producen la degeneración progresiva del colágeno ocasionando tendinosis, degeneración de la fascia y necrosis⁴.

En cuanto a su prevalencia se estima que aproximadamente el 10% de las personas sufre esta patología a lo largo de su vida^{5,6}. Es un trastorno auto-limitado, pues en el 80-90% de los casos los síntomas desaparecen dentro de los 10 meses posteriores. Sin embargo, este intervalo de tiempo resulta frustrante tanto para el paciente como para el especialista⁷.

Los síntomas son el dolor, más intenso al comenzar a deambular por la mañana o después de un período de inactividad física, que aumenta con la bipedestación prolongada o con actividades que requieren cargar pesos. No son frecuentes las parestesias ni el dolor nocturno⁸. La sensación dolorosa a veces aparece mediante la dorsi-flexión pasiva forzada del antepie o mediante la posición de puntillas.

Entre los factores predisponentes a sufrir esta patología se pueden distinguir tanto causas intrínsecas (pie cavo o plano, pronación excesiva del pie, deficiencias musculares o pérdida de elasticidad de la musculatura flexora plantar) y causas extrínsecas (calzado inadecuado, sobrecargas posturales y funcionales o marchas prolongadas por terrenos irregulares)⁴.

El diagnóstico principal se realiza a través de la clínica con la anamnesis y la exploración física adecuadas. Entre las pruebas complementarias, el estudio radiológico del pie en carga (que puede ser normal o mostrar la presencia de espolón calcáneo que a veces aparece en personas sin antecedentes de dolor en el talón), el estudio ecográfico (para mostrar las características histo-patológicas de la fascia) y el estudio biomecánico (para valorar la respuesta funcional a la deambulación)⁴.

El tratamiento de primera elección del dolor plantar es el conservador^{9,10} destacando los tratamientos ortopedológicos^{9,11-15}, fisioterápicos como vendajes¹⁶⁻¹⁹, estiramientos^{5,9,11,15,17,20-23}, ultrasonidos^{5,12,13,21,22,24,25}, ondas de choque^{21,23,26-32}, iontoforesis^{12,17,22}, láser³³ y magnetoterapia²⁵ entre otras^{9-13,20}, así como la acupuntura^{15,20,34}. A menudo se combina con fármacos antiinflamatorios e infiltraciones. No se suele realizar tratamiento quirúrgico, ya que no siempre da buenos resultados y tiene una recurrencia del 30% de los casos^{35,36}.

En la actualidad pocos estudios han evaluado la efectividad de las terapias físicas en el tratamiento conservador de la fascitis plantar. Existen muy pocas guías de práctica clínica^{42,43} sobre el tema y las investigaciones existentes resultan escasas⁴⁴.

El objetivo de esta investigación es determinar si las distintas terapias físicas utilizadas en el tratamiento conservador de la fascitis plantar de al menos un mes de evolución en adultos son efectivas individualmente y/o combinadas entre sí.

MATERIAL Y MÉTODOS

Diseño: Se realizó una revisión sistemática de estudios científicos que abordaran el tratamiento de la fascitis plantar mediante agentes físicos.

Estrategia de búsqueda: Se realizó una búsqueda en las bases de datos The Cochrane Library Plus, Medline, donde se utiliza-

ron los descriptores Mesh: *plantar fasciitis*, *physical therapy specialty* y *treatment* y las bases Lilacs IBECS, IME, PEDro y Enfispo, empleando como estrategia de búsqueda los descriptores DeCS: fascitis plantar, terapia, dolor y fisioterapia. Sin restricción en la fecha de los artículos, se realizó la última búsqueda en junio de 2013. Para la búsqueda de estudios originales en la base de datos Medline, por ejemplo, se usaron las siguientes ecuaciones de búsqueda: “plantar fasciitis” [Mesh] AND “physical therapy specialty” [Mesh], treatment AND “plantar fasciitis” [Mesh].

No se limitó por año de publicación. Se analizaron además las referencias bibliográficas de los artículos seleccionados con el fin de incluir otros estudios potencialmente válidos para la revisión.

Criterios de inclusión y exclusión. Se decidió que los estudios incluidos fueran ensayos clínicos controlados aleatoriamente, con un tamaño muestral superior a 20 sujetos. También se incluyeron revisiones sistemáticas que investigaran la efectividad de las distintas técnicas conservadoras para el tratamiento de la fascitis plantar, ya fuera de forma individual o combinadas entre sí. La evaluación de la elegibilidad de los estudios fue realizada por dos revisores de manera independiente, estandarizada y no cegada. Para clasificar los ensayos clínicos se utilizó la escala de lectura crítica PEDro además de una revisión crítica de cada resumen y evaluación del texto completo en el caso de los trabajos cuyo resumen no fuera concluyente. En caso de existir desacuerdo entre los revisores se resolvió mediante consenso.

La escala PEDro consta de 11 ítems (el primero no se ha tenido en cuenta) que evalúan la calidad metodológica de los ensayos clínicos aleatorios y controlados dando prioridad a dos puntos: la validez interna y la inclusión de datos estadísticos que faciliten la interpretación de resultados. Moseley et

al³⁷. Indican que los estudios con una puntuación igual o mayor a 5 son calificados como de alta calidad metodológica y bajo riesgo de sesgo. Por esta razón se decidió que todos los ensayos clínicos tuvieran una puntuación de al menos 5 puntos para ser incluidos en esta revisión.

Para la elección de estudios de intervención y prospectivos se postuló por parte de los evaluadores un total de 7 criterios de calidad metodológica (**anexo 1**), decidiéndose como puntuación de elegibilidad un mínimo de 4.

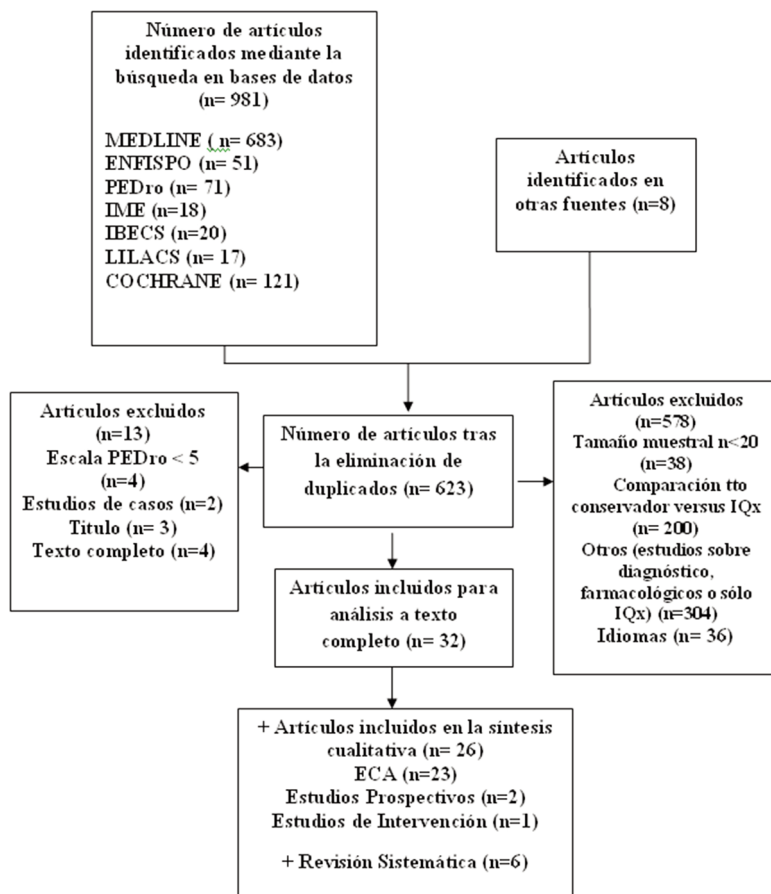
Los criterios de inclusión de los pacientes fueron: ser mayores de 18 años y con un diagnóstico de fascitis plantar de al menos un mes. Como criterios de exclusión se tuvieron en cuenta que los sujetos no sufrieran enfermedades reumáticas o inflamatorias, osteomielitis, infecciones agudas o crónicas a nivel del pie, insuficiencia neurológica o vascular, síndrome del tarso, problemas de coagulación. En el caso de las mujeres también se excluyó a las embarazadas.

Los tratamientos incluidos en la revisión fueron las terapias físicas manuales, estiramientos, distintas modalidades de electroterapia, ultrasonidos, iontoforesis, uso de material ortopodológico, vendajes, laser, magnetoterapia y acupuntura, utilizados de forma aislada o combinados entre sí. Se excluyeron los tratamientos farmacológicos exclusivos y los quirúrgicos. También se excluyeron los estudios que comparaban tratamientos conservadores con tratamientos quirúrgico y los que no estuvieran publicados en inglés o en español.

Además se tuvieron en cuenta los resultados de las revisiones anteriormente publicadas.

Extracción de datos: Dos investigadores identificaron de forma individual los estudios potencialmente elegibles por cumplir los criterios de inclusión. Para la recogida

Figura 1
Diagrama de flujo para la selección de artículos



de datos relevantes para el estudio se diseñó *ad hoc* por los propios investigadores una plantilla que facilitara la evaluación. En caso de desacuerdo esta se resolvió mediante consenso.

RESULTADOS

Se analizaron 26 artículos, 1 estudio de intervención (n=84), 2 estudios prospectivos (n=91) y 23 ensayos clínicos controlados (n=1.683). Además se consultaron 6 revisiones sistematicas^{7,8,37,39-41} (figura 1)

En los 23 ensayos clínicos revisados, la media obtenida en la escala PEDro fue de 7,17 puntos. 3 ensayos^{29,31,32} obtuvieron la puntuación más alta (9), mientras que 4 ensayos^{11,16,21,25} alcanzaron la puntuación mínima para ser incluidos (5). (tabla 1). Los dos estudios prospectivos revisados^{12,19} obtuvieron una puntuación de 5 mientras que el estudio de intervención²⁶ alcanzó una puntuación de 6 (anexo 1) según los criterios metodológicos de calidad utilizados para evaluar estudios de intervención y prospectivos.

Tabla 1
Resultados según la escala PEDro

ENSAYOS CLINICOS	Criterio de selección	Asignación aleatoria	Asignación oculta	Grupos homogéneos	Pacientes ciegos	Terapeuta ciego	Evaluadores ciegos	Seguimiento al menos 85%	Datos analizados	Estadística de al menos 1 dato comparado	Medidas puntuables al menos de 1 resultado	Total
Albornoz J ¹⁴	1	1	0	1	0	0	0	1	1	1	1	6/10
Greve A ²¹	0	1	0	1	0	0	0	1	0	1	1	5/10
Zanon R ⁵	1	1	0	1	1	0	1	1	1	1	1	8/10
Babak V ²⁷	1	1	0	1	1	0	0	1	1	1	1	7/10
Tsai CT ¹⁶ .	0	1	1	0	0	0	1	0	0	1	1	5/10
Cleland J ²²	1	1	0	1	0	0	1	1	1	1	1	7/10
Lohrer H ²⁸	1	1	0	1	1	0	1	1	1	1	1	8/10
Crawford F ²⁴	1	1	0	1	1	1	0	1	1	1	1	8/10
Kummerddee W ¹⁵	1	1	1	1	0	0	0	1	0	1	1	6/10
Karagounis P ²⁰	1	1	1	1	1	0	1	1	0	1	1	8/10
Gerald T ¹³	0	1	1	0	0	0	0	1	1	1	1	6/10
Gerdesmeyer L ²⁹	1	1	1	1	1	0	1	1	1	1	1	9/10
Haake M ³⁰	1	1	0	1	1	0	1	1	1	1	1	8/10
Ibrahim M ³¹	0	1	1	1	1	0	1	1	1	1	1	9/10
Kiritisi O ³³	1	1	0	1	1	1	1	0	0	1	1	7/10
Kudo P ³²	1	1	1	1	1	0	1	1	1	1	1	9/10
Osborne H ¹⁷	1	1	0	1	1	0	1	1	1	1	1	8/10
Radford J ¹⁸	1	1	1	1	1	0	1	1	1	1	1	9/10
Rompe J ²³	0	1	1	0	0	0	0	1	1	1	1	6/10
Stratton M ¹¹	1	1	0	1	0	0	0	1	0	1	1	5/10
Di Giovanni B ^{F9}	1	1	1	1	0	0	1	1	1	1	1	8/10
Zhang SP ³⁴	1	1	0	1	1	0	1	1	1	1	1	8/10
Telleria T ²⁵	1	1	0	0	0	0	0	1	1	1	1	5/10

Características de la población: El número total de sujetos estudiados fue de 1.858. Respecto a las variables de los sujetos, en algunos estudios se recoge el tiempo de evolución de la FP, incluyéndose sujetos con diagnóstico de fascitis plantar de cuatro semanas^{14,18}, tres meses^{21,27,34} y seis meses de evolución^{5,26,29,31}. Otra variable de la población que se estudió fue el método diagnóstico de la FP, que en su gran mayoría se basó en la anamnesis y exploración física y radiografía^{12,14,17,18,21-23,29,31,32,34}, sólo unos pocos utilizaron la ecografía^{17,20,21,26}, la gammagrafía o la resonancia magnética^{20,23}.

Análisis de los datos: La principal medida de resultado evaluada fue el dolor, sobre todo el matutino, para lo que se emplearon medidas validadas como la Escala Visual Analógica (EVA)^{11,13,15,17,18,21,24-26,29,31-34} y escalas de puntuación numérica²⁷. Otros resultados secundarios midieron la funcionalidad del paciente y su capacidad para realizar las actividades de la vida diaria. Algunas de las medidas de resultado secundarias utilizadas fueron: escala *Activities of daily living/Foot and Ankle Ability Measure* (ADL/FAAM)^{11,12}, cuestionario Índice funcional del pie (FFI)^{9,15,16,23,28}, escala *American Orthopaedic Foot and Ankle Society* (AOFAS)^{15,32}, escala de Roles and Maudsley²⁹⁻³¹, cuestionario de dolor de McGill Melnick¹⁶ e incluso uso de algómetro y medición cuantitativa del grosor de la fascia mediante ecografía.

En la **tabla 2** se muestran las características de cada uno de los 26 estudios incluidos en esta revisión clasificados con el nombre del primer autor, año de publicación, tamaño muestral, tipo de estudio, técnica empleada, tratamiento realizado y resumen de los resultados obtenidos, con las escalas de valoración empleadas y los datos estadísticos.

En cuanto al idioma, 22 (84,62%) estaban publicados en lengua inglesa y 4 (15,38%) en castellano.

Agrupando los estudios según las técnicas empleadas, tanto los estiramientos y las terapias manuales como el empleo de ondas de choque OC radiales (OCR) o focales (OCF) fueron los más frecuentes (39,13%). Respecto al nivel de mejoría, las OCR tuvieron una eficacia significativa en 4^{21,26,29,31} de los 5 ensayos clínicos en los que se utilizaron, mientras que la efectividad de las OCF no fue concluyente combinadas con otros tratamientos^{30,32}, sin embargo cuando se comparó su efectividad al ser empleadas exclusivamente ofrecieron los mejores resultados²⁸.

Los estiramientos y terapias manuales resultaron eficaces en 4 de los 9 ensayos realizados^{21-23,49}, siendo más significativo en cuanto mejoría el estiramiento de la fascia plantar respecto al estiramiento del tendón de Aquiles^{9,38}.

La siguiente técnica más utilizada fueron los ultrasonidos (30,43%) con los que se obtuvieron buenos resultados en dos ensayos clínicos^{12,13}. En tres estudios no fueron efectivos^{5,22,24} y en los otros dos estudios la mejoría fue similar cuando se compararon con el uso de OC²¹ y con magnetoterapia²⁵.

Tanto el uso de materiales ortopodológicos (26,08%) como la aplicación de vendajes, en particular la técnica de lowdye (17,39%) fueron más eficaces cuando se combinaron con otras técnicas.

En cuanto al uso de Iontoforesis (13,04%), el empleo de ácido acético¹⁷ parece revelarse como el más efectivo en la disminución del dolor matutino, encontrándose en otros estudios buenos resultados¹² pero en desventaja frente a otras técnicas²².

Otra técnica utilizada con buenos resultados fue la acupuntura (13,04%)^{15,20,34}. Con resultados moderados encontramos el uso de láser³³ y sin mejoría evidente el uso de magnetoterapia, que se estudió en un trabajo comparada con los ultrasonidos²⁵.

Tabla 2
Análisis de los estudios incluidos en la revisión

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Rumbaut R. et al ¹⁹ 2009	28 (21 mujeres)	prospectivo	Vendaje adhesivo (funcional) + tto farmacológico analgésico	Colocación de vendaje para eliminar la tensión en los fascículos aferentes de la fascia plantar.	El 100% de los pacientes experimento mejoría a las 24 horas de dolor agudo. 88,8% mejoró sintomatología tras 3 semanas e inmovilización.
Albornoz J ¹⁴ 2004	21 talones grupo experimental. 22 talones grupo control.	prospectivo experimental controlado. No doble ciego. Intervalo de confianza 95%.	Férula de uso nocturno.	Grupo experimental Aines, estiramientos y uso nocturno de férulas. Grupo control Aines y estiramientos.	Se evaluaron los resultados empleando el Score de Roles y Maudsley para la evaluación subjetiva del dolor y el numero de AINES consumidos semanalmente. La evaluación de los resultados se realizó empleando las pruebas estadísticas de T de Student y Chi ² . En ambos test se empleó un nivel de significancia de 95% (p=0.05) Grupo experimental: 16 excelentes (ausencia de dolor), 6 buenos (disminución significativa del dolor). Grupo control: 8 excelentes, 8 buenos, 5 regular (disminución leve del dolor).
Greve A ²¹ 2009	2 grupos de 16 sujetos. Total 32	Estudio clínico aleatorio prospectivo y comparativo	- Ondas de choque. - Ultrasonidos estiramientos y movilizaciones	Grupo 1 10 sesiones de ultrasonidos, estiramientos y movilizaciones. Grupo 2 3 aplicaciones de ondas de choque (1 a la semana) de 2000 impulsos y 6Hz y ejercicios de estiramientos domiciliarios.	Se realizó una evaluación antes, al terminar el tto y a los tres meses posteriores mediante una anamnesis, valoración del dolor utilizando la escala visual analógica (EVA) y algómetro de Fischer para su cuantificación, tanto en la inserción calcáneo de la fascia como en el tercio medio del gastronemio. En la evaluación inicial los dos grupos presentaron similares resultados en la presentación, duración y periodicidad del dolor así como en las escala y el algómetro. Grupo 1: Los resultados al terminar el tto eran significativamente mejores en cuanto a la disminución de la intensidad del dolor. Los resultados a los tres meses de finalizar el tto en ambos grupos no presentaban diferencias estadísticas, lográndose un descenso de la intensidad del dolor, de los valores en la escala EVA y en el algómetro de Fischer así como una disminución de la ingesta de analgésicos.

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Zanon R ⁵ 2005	Total 22 (27 pies) Grupo 1: 13. Grupo 2: 14.	Estudio clínico aleatorio y prospectivo y doble ciego.	Ultrasonidos continuos de 2w/cm2 y estiramientos.	Grupo 1 15 sesiones. Estiramientos más placebo de Us. Grupo 2 15 sesiones. Estiramientos más Us 2w/cm2.	La evaluación inicial consistió en anamnesis y una evaluación funcional que se realizó también al finalizar el tto, usando el cuestionario AOFAS. La escala visual analógica (EVA) se empleó al principio, a las ocho semanas y al final del tto. Para la intensidad del dolor se utilizó el test de Friedman y el test de Dunn El análisis estadístico fue realizado a través de la T de Student y el test de Wilcoxon por comparación antes y después del tto. Grupo 1: Los estiramientos resultan ser eficaces como tto. A las ocho sesiones se obtuvo una mejoría del 54,64% Grupo 2: El empleo de Us continuos, no implica una mayor disminución del dolor. A las ocho sesiones la mejoría era del 46,52%. Estadísticamente tanto en la escala AOFAS como en la escala EVA $p \leq 0,001$ en ambos grupos a las ocho sesiones y al final del tto.
Garcia E et al ²⁶ . 2005.	74 pacientes, 10 bilaterales (84 pies)	Estudio de intervención.	Ondas de choque	3 sesiones con ondas de choque de 2000 impulsos con intensidad progresiva.	Como medida de evaluación se utilizó la radiografía lateral de pie, la ecografía diagnóstica y la escala visual analógica (EVA) a la semana, a las 6 semanas y a las doce semanas de concluir el tratamiento. 77, 38% buenos resultados (EVA 0-2, incorporación a las actividades habituales y sin otro tratamiento conservador) 7,17% resultados regulares (EVA 3-4, incorporación a las actividades habituales pero se requirió de otro tto conservador). 15,48% malos resultados (EVA > 4 y no hubo incorporación a las actividades habituales, requiriéndose tratamiento quirúrgico también.)

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS																		
Rioja J 2001	49 sujetos (33 mujeres) 14 bilaterales. 63 pies en total	prospectivo	Iontoforesis + Us + transferencia eléctrica capacitativa + plantilla	15 sesiones: Iontoforesis con fosfato sódico de dexametasona al 0,4%; ultrasonidos continuos; diatermia con 0,8 MHz de frecuencia Transfereencia eléctrica capacitativa; plantillas silicona.	Los pacientes rellenaron una escala de valoración basada en los criterios de evaluación de Young con tres graduaciones: leve moderada e intensa obteniéndose los siguientes resultados <table><tr><td></td><td>15 sesiones</td><td>30 sesiones</td></tr><tr><td>pies asintomáticos</td><td>19</td><td>65</td></tr><tr><td>afectación leve</td><td>30</td><td>17</td></tr><tr><td>afectación moderada</td><td>35</td><td>3</td></tr><tr><td>afectación intensa</td><td>6</td><td>2</td></tr><tr><td>abandonos</td><td>10</td><td>13</td></tr></table> La escala visual analógica no se tuvo en cuenta por su discordancia con la funcional		15 sesiones	30 sesiones	pies asintomáticos	19	65	afectación leve	30	17	afectación moderada	35	3	afectación intensa	6	2	abandonos	10	13
	15 sesiones	30 sesiones																					
pies asintomáticos	19	65																					
afectación leve	30	17																					
afectación moderada	35	3																					
afectación intensa	6	2																					
abandonos	10	13																					
Babak V ²⁷ 2012	20 pacientes en cada grupo	Ensayo clínico aleatorio controlado	Ondas de choque: - Focales. - Radiales.	Grupo 1 3 sesiones/semanal. 2000 impulsos focales +2000 radiales en total: 0.2 mJ/mm2. Grupo 2 3 sesiones/semanal. Placebo 0,04 mJ/mm2	Ecografía antes y después del tto para valora el grosor de la fascia. Para la valoración subjetiva del dolor se cumplimentó la escala numérica NRS (0 no dolor, 10 dolor más severo en las actividades diarias). La valoración se repitió a los tres meses de finalizar el tto. El análisis estadístico se realizó con el programa SPSS versión 16.0, empleándose la T de Student para los cambios de grosor y la valoración subjetiva del dolor entre los dos grupos. El coeficiente de correlación de Pearson se empleó para investigar la relación entre las variables continuas considerándose significativas si el valor $p < 0,05$. Al inicio no se apreciaron diferencias entre los dos grupos en el grosor $P = 0,95$. Los resultados fueron: Grupo 1: Mejoría significativa con disminución del grosor de la fascia $p < 0,001$ Grupo 2: No se aprecia disminución del grosor incluso se apreció un aumento de grosor $P = 0,03$.																		

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Tsai CT ¹⁶ 2010	Total 52 (19 hombres, 33 mujeres) 5 bilaterales.	Ensayo clínico aleatorio controlado simple ciego	Kinesiotaping	Grupo 1 o Control US + termoterapia + Tens. Grupo 2 Us + termoterapia +Tens + kinesiotaping (1 semana)	Ecografía al principio y una semana después de acabar el tto. La valoración subjetiva del dolor se realizó con el cuestionario de McGill Meldnack y el índice funcional del pie (FFI siete ítem con una puntuación entre 0-100). El análisis estadístico se realizó con el programa SPSS versión 11.0. Para comparar las diferencias entre los grupos, se normalizo los resultados en porcentajes. Para las variables continuas se empleó el test de Mann-Whitney y para otras variables el test exacto de Fisher ambos con una diferencia significativa P= 0,05. Grupo experimental después del tto se observó una disminución del grosor de la fascia a nivel de la inserción, y una disminución en la intensidad del dolor ambos con una p> 0,05. En ambos grupos se observó una mejoría del índice funcional del pie. En ningún grupo hubo mejoría del grosor a 5 cm de la inserción, ni de densidad.
Cleland J ²² 2009	Total: 60 30 en cada grupo.	Ensayo clínico aleatorio ciego	Electroterapia iontoforesis + Us pulsátil + estiramientos. Terapia manual.	Grupo 1 Iontoforesis (dexametasona) + Us pulsátil 1,5w/cm2+ estiramientos. Grupo 2 Terapia manual + ejercicios.	Se utilizaron los cuestionarios de escalas funcionales LEFS y FAAM y el NPRS para la intensidad del dolor así como el BAI para el índice de ansiedad. Se rellenaron al principio, a las 4 semanas y a los 6 meses. El programa estadístico empleado fue el SPSS utilizando medida de tendencia central y de dispersión para las variables continuas así como la T de Student y Chi cuadrado, considerando P= 0,05. Los efectos sobre el dolor y discapacidad se examinaron con el modelo mixto de varianza ANOVA y un índice de confianza de 95%. Grupo 2: Los resultados obtenidos fueron mejores tanto a las cuatro semanas como a los seis meses en LEFS P= 0,001 y 0,027, en FAMM P= 0,004 y P= 0,012 y en NPRS P= 0,008 y P= 0,39.
Lohrer H ²⁸ 2010	Total: 39. Grupo radial: 19. Grupo focal: 20.	prospectivo aleatorio doble ciego	Ondas de choque: -Focales. -Radiales	Grupo focal: 0,20 mj/mm2.. 3sesiones/semana. 2000 impulsos 10Hz. Grupo radial 0,17 mj/mm2. 3sesiones/semana. 2000 impulsos 10Hz.	Se emplearon la escala funcional del pie (FFI con 8 ítem) y pruebas neuromusculares de salto sobre una pierna, salto de longitud, estabilidad postural y test isocinético articular analizándose a principio, a las 2 y 12 semanas después del tratamiento. El análisis estadístico se realizó usando Testimate versión 6.4, el metaanálisis a través de Metasub versión 1.3.6. Se utilizó además el test Wilcoxon-Mann Whitney (invariable y multivariable). La significación estadística se determinó con un índice de confianza del 95%. Grupo Focal: Combinando todas las variables analizadas se obtienen mejores resultados en este grupo en 6 de los ocho ítems obteniéndose P= 0,0027 en FFI frente a P= 0,0013 del grupo radial.

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Crawford F ²⁴ 1996	Total: 19 7 bilaterales	Ensayo clínico aleatorio doble ciego	Ultrasonidos	- Ultrasonido pulsátil (1:4) 0,5w/cm2 8 minutos. - Placebo: solo tiempo.	Empleo de la escala visual analógica (EVA) al principio y final del tratamiento. Para el análisis estadístico se utilizó el test de Wilcoxon y el test exacto de Fischer, para una significación estadística se consideró que debería hallarse al menos en 26 resultados una mejoría del 50%. Grupo ultrasonidos: Reducción del 30% del dolor. Grupo placebo: Reducción del 25% del dolor. Por lo tanto los datos no son estadísticamente significativos, en ningún grupo hubo resolución completa del dolor, empeoro en 4 (2 en cada grupo) y no hubo cambios en otros 2 (1 en cada grupo)
Kumnerddee W ¹⁵ 2012	Total: 30	Ensayo clínico aleatorio controlado	Tratamiento convencional: Estiramientos Analgésicos. Plantillas. Electroacupuntura en puntos locales	Grupo control Estiramientos + plantillas y analgésicos. Grupo experimental Estiramientos + analgésicos + plantillas + electroacupuntura. (2 sesiones/semana)	Se realizó un estudio radiográfico previo para descartar patología ósea. Se emplearon la escala visual analógica (EVA) y el índice funcional del pie (FFI con 23 ítems). Para el análisis estadístico se utilizó el programa estadístico SPSS versión 13.0, para las variables continuas se utilizó La T de Student con valor $p < 0,005$ para ser significativo. También se empleó el test exacto de Fischer y Chi cuadrado. Se calculó un mínimo del 50% de disminución del dolor para tener significación estadística en la escala EVA. Grupo experimental: Disminución del 69% en el 80% de los casos en la escala EVA. Disminución significativa en FFI. Grupo control: Disminución del 14% en la escala EVA que no es considerada significativa. En FFI no se observó ninguna disminución significativa.
Karagounis P ²⁰ 2011	Total: 38 hombres. 19 en cada grupo. (atletas)	Ensayo clínico aleatorio controlado. Doble ciego. Intervalo de confianza 81%.	Crioterapia. Aines. Estiramientos Fortalecimiento. Acupuntura en puntos distales.	Grupo control Crioterapia, Aines, estiramientos y fortalecimiento. Grupo experimental Ídem al control + acupuntura (16 sesiones 2/semana)	La escala utilizada es la escala para el dolor e incapacidad de Willis et al (PFPS) evaluándose al principio, a las 4 semanas y a las ocho semanas desde el tratamiento. El análisis estadístico se realizó con el programa SPSS considerándose $P = 0,05$ estadísticamente significativa. El análisis se realizó a través del modelo mixto de varianza ANOVA para determinar los diferentes resultados entre los dos grupos. Grupo experimental: Disminución $p < 0,05$ en PFPS tanto a las 4 semanas como a las 8 semanas. Grupo control: Disminución $p < 0,05$ en PFPS tanto a las 4 semanas como a las 8 semanas. No hay diferencias estadísticas entre ambos grupos a las 4 semanas, mientras que a las 8

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Gerald T ¹³ 2011	Total: 100 4 grupos de 25.	Ensayo clínico prospectivo aleatorio.	-Arcos plantares. - Ultrasonidos. -Infiltración. -Órtesis.	Grupo 1 Us. 6 sesiones de 5 m con 1,5 w 2 semanas. Grupo 2 Arcos plantares. Grupo 3 Infiltración de triamcinolona + Marcaine (5%). Grupo 4 Órtesis	Se utilizó la escala visual analógica (EVA) al principio y final del tto. Se calculó la mejoría subjetiva del dolor, el número de pacientes que habían obtenido completa resolución y así como el porcentaje de pacientes con mejoría. Grupo ultrasonidos: reducción subjetiva de 3,97, mejoría del 81% de los pacientes y resolución en el 20%. Grupo arcos plantares: Sin dolor solo un paciente, mejoría del 35%. Grupo infiltración: Mejoría del 72% de los pacientes, resolución en el 7%. Grupo órtesis: reducción subjetiva de 2,92, 64% de pacientes con mejoría y resolución en el 32%.
Gerdesmeyer L ²⁹ 2008	Total: 245. Grupo experimental: 129. Grupo placebo: 122.	Ensayo clínico aleatorio controlado doble ciego. Nivel de evidencia 1. Intervalo de confianza 90%	-Ondas de choque radiales - Placebo	Grupo experimental 3 sesiones de 2000 impulsos y 0,16 mj/mm2. Grupo control Placebo.	Se ha utilizado la escala visual EVA en criterios como dolor matutino, dolor en actividades diarias, aplicación del dolormetro con una puntuación mínima al comenzar de 5 ó más. También se empleó la escala Roles and Maudsley y las escalas SF-36 de satisfacción física y psicológica. La evaluación se realizó a las 12 semanas y a los 12 meses después del tratamiento. Grupo experimental: a las 12 semanas, la mejoría en la escala EVA en la media de todos los criterios fue de 72% frente al 44,7% del grupo placebo P=0,0103; a los 12 meses se obtuvo un 84,8% frente 43,2% del placebo P=0,0086. En criterios secundarios en Roles and Maudsley se obtuvo P=0,031 a la doce semanas y en la SF-36 física una p=0,0013.
Haake M ³⁰ 2003	Total: 272. Grupo experimental: 135. Grupo placebo: 137.	Ensayo clínico aleatorio doble ciego. Intervalo de confianza 95%	Ondas de choque focales	Grupo experimental 3 sesiones de 4000 impulsos y 0,08 mj/mm2. Se aplican con anestesia. Cada dos semanas. Grupo control Placebo + anestesia.	La escala Roles and Maudsley fue utilizada como criterio principal, aplicándose a las 6, 12 y al año después del tratamiento. Para el análisis estadístico se utilizó la OOD ratio y el test exacto de Fischer para comparar resultados entre grupos. Se consideró un 20% de diferencia entre estos para que hubiera relevancia clínica. Grupo experimental: se obtuvo una disminución del 15,1% frente al 8,0% del placebo con una p=0,5927 y una ODD=1,18 por lo que no se alcanzó el 20% exigido para obtener relevancia clínica. Al año en la escala Roles and Maudsley se obtuvo un 81% frente al 76% en el grupo placebo.

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Ibrahim M ³¹ 2010	Total: 50 todos unilaterales.	Ensayo clínico aleatorio controlado doble ciego.	Ondas de choque radiales	Grupo experimental 2 sesiones de 2000 impulsos y 0,16 mj/mm2. Semanal. Grupo placebo Se puso un aislante para evitar las ondas.	Las escalas utilizadas fueron la escala visual analógica (EVA) y la escala modificada de Roles and Maudsley (RM), evaluándose a las 4, 12 y 24 semanas tras el tto. Para el análisis estadístico se empleó el programa SPSS versión 16.00, empleándose el modelo mixto de varianza ANOVA así como el índice de corrección de Bonforromi para comparar resultados a través de secuencia temporales. Se consideró el tto efectivo si había una disminución de al menos el 60% en la escala EVA. Se empleo Chi cuadrado y una $p < 0,05$ para tener significancia estadística. Grupo experimental: La disminución en la escala EVA fue de 92,5%, 87,3% y 93,9% respectivamente en las evaluaciones. La RM disminuyó el 68,1%, 61,7% y 64,9%. Grupo control: La disminución en la escala EVA fue mucho menos efectiva 15,2%, 13,5% y 17,0%. Respecto a la RM 6,3% 15,8% y 16,8%.
Kiritisi O ³³ 2010	Total: 30 todos unilaterales	Ensayo clínico aleatorio controlado doble ciego.	Laser de baja frecuencia in- frarrojos	Grupo experimental Tres sesiones por semana durante seis semanas con una dosis de 8,4 J y 157,5 sg.	Se efectuó ecografía antes y después de tto para comprobar cambios en el grosor de la fascia. La escala utilizada para la valoración del dolor fue la EVA. Para el análisis estadístico se empleo el programa SPSS versión 16.00, se utilizó la T de Student y Chi cuadrado considerándose $p < 0,05$ para tener relevancia estadística, el coeficiente de Pearson para ver la relación entre el grosor de la fascia y el nivel del dolor en EVA. Grupo experimental: A las 6 semanas la disminución del dolor fue del 59% frente al 26% del grupo placebo $p = 0,001$ El cambio de grosor fue significativo en los dos grupos comparando antes y después del tto, sin embargo comparando entre los dos grupos se obtuvo una $P = 0,12$ no siendo significativa la diferencia.

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Kudo P ³² 2006	Total: 114. Grupo experimental: 58. Grupo placebo: 56.	Ensayo clínico aleatorio controlado doble ciego	- Ondas de choque focales - Placebo	Grupo experimental: Una sesión de siete minutos incrementándose la intensidad cada minuto. Ondas de choque a 3800 impulsos y 2,330mj/mm ² de media + anestésico	Se utilizaron las escalas con valores mínimos al principio del tto EVA (>5) y la RM (3-4), la escala AOFAS, la SF-12, la del dolor a la palpación (PTM). Las evaluaciones se realizaron al principio, a las 6 semanas y a los 3 meses. El programa estadístico utilizado fue el SAS System considerando P= 0,05 y que al menos se consiga una disminución del 60% del dolor para ser significativa a los 3 meses. Se uso el Chi cuadrado, T de Student y el test exacto de Fischer (evaluar efectos secundarios) Grupo experimental: A los 3 meses la disminución en la escala EVA en el 49,1% de los sujetos fue de más del 60%. p<0,0001. Grupo placebo: el 23% de los sujetos alcanzo el 60% de disminución en la escala EVA p= 0,0099. Así pues, las diferencia entre grupos fue significativa ya que se obtuvo P= 0,0124, en la escala AOFA y SF12 las diferencias entre grupos no fueron significativas, en la RM se obtuvo p= 0,0121 y en PTM P= 0,0027.
Osborne H ¹⁷ 2006	Total 31. (42 pies) 11 bilaterales. Grupo placebo: 12. Grupo dexametasona: 15. Grupo ácido acético: 15.	Ensayo clínico aleatorio controlado doble ciego. Intervalo de confianza 95%	-Iontoforesis. -Vendaje funcional. (técnica lowdye) -Estiramientos	Grupo placebo 6 sesiones iontoforesis con CINA 0,9% + lowdye + estiramientos. En días alternos. 2 semanas. Grupo dexametasona 6 sesiones de iontoforesis al 0,4% + lowdye + estiramientos. En días alternos. 2 semanas. Grupo ácido acético 6 sesiones de iontoforesis al 5% + lowdye + estiramientos. En días alternos. 2 semanas.	Se evaluó la rigidez matutina y el dolor al principio, final y a las semanas después del tratamiento, mediante la escala visual analógica para el dolor y la rigidez. El análisis estadístico se llevo a cabo mediante el programa SAS y Excel, la diferencia entre grupos se determino con el análisis de la covarianza respecto a las escalas iniciales, además se utilizo el test de Fischer (PLSD). Se consideró p<0,05 y el 95% de intervalo de confianza como significancia estadística. Grupo dexametasona: Al finalizar el tto la disminución de la rigidez fue similar al grupo acético pero menor la disminución del dolor. A las 4 semanas la rigidez mostraba menos reducción frente al grupo acético. Grupo acético: Mostró una significativa mejoría tanto en el dolor como en la rigidez tras el tto. Los resultados sobre la rigidez se mantuvieron a las 4 semanas de finalizar el tto con p=0,031. Grupo placebo: El uso solo de taping fue menos eficaz que cuando se combino con la iontoforesis.

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Radford J ¹⁸ 2006	Total: 92. 46 en cada grupo.	Ensayo clínico aleatorio controlado. Simple ciego. Intervalo de confianza del 95%	- Vendaje funcional. (Técnica lowdye) -Placebo de Ultrasonidos.	Grupo experimental Vendaje funcional + placebo de Us. durante 3 minutos 1 sesión diaria/ 1 semana. Grupo control Placebo de Us. 3 minutos 1 sesión diaria/1 semana.	Se evaluó al principio y a la semana el dolor matutino con la escala visual analógica (EVA) y el cuestionario Foot Health Status previamente validado sobre el dolor, funcionalidad y la salud del pie. No se especifica el programa estadístico utilizado pero se toma el 80% de disminución del dolor como significancia estadística. Los resultados se analizaron por intención de tratar y se realizó una análisis de la covarianza (ANCOVA). Grupo experimental: Se obtuvo una significativa mejoría en dolor matutino comparado con el grupo placebo $P=0,017$. La diferencia en la valoración con el cuestionario Foot Health Status no fue lo suficientemente significativa. Un 28% de los sujetos experimento efectos secundarios como alergia, nuevo dolor y vendaje demasiado apretado.
Rompe J ²³ 2010	Total: 102. Grupo estiramiento: 54. Grupo ondas de choque: 48.	Ensayo clínico aleatorio controlado. Nivel de evidencia 1. Intervalo de confianza 95%	- Ondas de choque radial. - Programa de estiramientos	Grupo ondas de choque 3 sesiones de ondas de choque radial de 2000 impulsos y 0,16 mj/mm2/semanal. Grupo estiramientos Programa de ejercicios a realizar tres veces día/8semanas.	Se utilizaron la escala funcional del pie (PS-FFI con 8 ítems) y la SROM de satisfacción. La evaluación se realizó al principio a los 2,4 y 15 meses. Los principales resultados medidos fueron el ítem 2 (dolor matutino) de PS-FFI y el 6 de SROM a los dos meses. No se especifica el programa estadístico utilizado. Para todo el análisis se considero $p<0,025$ y para la relevancia clínica al menos un 20% de reducción en el ítems 2 de PS-FFI con la corrección de Bonferromi. Para el análisis SROM se considero $p<0001$ con el test exacto de Fischer. Grupo estiramientos: Tanto en el total de la escala PS-FFI, como en el ítem 2 hubo diferencias significativas respecto al grupo de ondas con $p<0,001$ a los dos y 4 meses, sin embargo a los 15 no se hallaron diferencias. El 65% del grupo de estiramientos frente al 29% del grupo

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Stratton M ¹¹ 2009	Total: 26 Dos grupos de 13	Ensayo clínico aleatorio controlado	- Órtesis - Estiramientos - Estimulación de baja frecuencia (10Hz/sg. Tens?)	Grupo control Uso de órtesis + programa de estiramientos 4 semanas. Grupo experimental Uso de órtesis + programa de estiramientos + electroestimulación 4 semanas.	Se utilizó la escala visual analógica EVA para el dolor y la escala ADL/FAAM para la funcionalidad, haciéndose la evaluación al principio a las 4 semanas y a los tres meses. No se especifica ningún programa estadístico. Se empleo la T de Student y el coeficiente de correlación de Pearson. Para comparar los dos grupos se realizó un análisis mixto de varianza de las dos escalas considerándose P= 0,05 como significancia estadística en todas las evaluaciones. Grupo experimental: No se hallaron diferencias significativas respecto al grupo control en la escala EVA aunque respecto a la evaluación inicial a las 4 semanas p< 0,005 y a los tres meses p< 0,001. Así mismo, no se hallaron diferencias significativas en los valores de la escala ADL/FAAM entre los dos grupos. Respecto a la evaluación inicial a las 4 semanas p< 0,001 y desde las 4 semanas a los tres meses no se observaron diferencias significativas.
Di Giovanni B ^{F9} 2003	Total: 101	Ensayo clínico prospectivo aleatorio. Nivel de evidencia 1. Intervalo de confianza del 95%	- Estiramientos del tendón de Aquiles. -Estiramientos de la fascia. -Plantillas. - Aines (celebrex)	Grupo A o Fascia Programa de estiramiento de la fascia (3 veces al día) + uso de las plantillas + Aines. 8 semanas Grupo B Programa de estiramientos del tendón de Aquiles (tres veces al día) + uso de las plantillas + Aines 8 semanas.	Las escalas utilizadas fueron la FFI (7 ítems) y la SROM de funcionalidad y satisfacción con el tto. Se realizó la evaluación al principio y a las 8 semanas. Se realizó procedimiento estadístico estándar (no especifica programa), empleándose la T de Student para datos continuos y el test exacto de Fischer. Al realizar el análisis de la varianza se consideró p< 0,05 como significancia clínica y al menos un 18% de disminución respecto a la evaluación inicial. Grupo fascia: Se obtuvieron mejores resultados en la escala FFI en el ítem 1 (peor dolor) p=0,002 y en el 2 (dolor matutino) P=0,006 respecto al grupo B. Al considerar todos los ítems, no se detectan diferencias entre grupos P= 0,17 a las ocho semanas. En el cuestionario SROM se halló diferencias significativa en todos los ítems variando entre p=0,006 y P=0,025 respecto al grupo B.

Tabla 2 (continuación)

AUTOR	SUJETOS	DISEÑO DE ESTUDIO	TECNICA EMPLEADA	INTERVENCION	RESULTADOS
Zhang SP ³⁴ 2011	Total 53. Grupo experimental: 28. Grupo control: 25.	Ensayo clínico aleatorio controlado. Doble ciego.	- Aplicación de agujas de acupuntura de forma distal (en muñeca).	Grupo experimental Aplicación de la aguja en el punto PC7. 10 sesiones /2 semanas. Grupo control Aplicación de la aguja en el punto L14. 10 sesiones / 2 semanas.	Se utilizaron la escala visual analógica EVA para el dolor matutino y la escala Bar-korec and Nau para satisfacción. Se aplicó el algómetro para el dolor a la presión y en general. Se realizó la evaluación a al principio, 3 meses y seis meses. El programa estadístico utilizado fue el SPSS versión 13.0. Se uso la T de Student, Chi cuadrado análisis de la varianza y covarianza con la corrección de Bonforromi y análisis regresivo. Se considero significancia estadística una disminución de al menos 33% entre los dos grupos en la escala EVA y P=0,05 en todos los puntos. Grupo experimental: Respecto al grupo control se obtuvieron resultados significativos en el dolor matutino p=0,44, dolor general P= 0,49 y dolor a la presión P=0,007 al mes de tratamiento. Respecto a la evaluación inicial, no se observo cambios en el punto de dolor a la presión pero en todos los otros si, siendo el dolor matutino el más significativo con p< 0,001 tanto al mes como a los seis meses. Grupo control: no se observaron cambios significativos respecto a la evaluación inicial en dolor matutino aunque si en el dolor en actividad y dolor en general p<0,05
Telleria T ²⁵ 2009	Total:60. Cada grupo 30.	Ensayo clínico aleatorio prospectivo.	- Magnetoterapia 		

Otras técnicas mencionadas asociadas a las anteriores fueron la transferencia eléctrica capacitativa¹², la crioterapia²⁰, la electroestimulación neuromuscular transcutánea (TENS)¹¹, los antiinflamatorios no esteroideos⁹ y las infiltraciones¹³ aunque su efectividad aislada no fue objeto de estudio.

DISCUSIÓN

Las principales terapias físicas utilizadas para el tratamiento conservador de la fascitis plantar en los estudios prospectivos y ensayos clínicos evaluados fueron los ultrasonidos, la iontoforesis, material ortopodológico, vendajes, estiramientos, ondas de choque, láser, magnetoterapia o acupuntura entre otras. A causa de la heterogeneidad de técnicas no se ha podido realizar un análisis estadístico cuantitativo, encontrándonos como limitación la dificultad de homogeneizar la información sobre las distintas técnicas al aplicarse diferentes parámetros, número distinto de sesiones o diferente duración de tratamientos.

El uso de ultrasonidos obtiene buenos resultados cuando se le asocia a la iontoforesis, transferencia eléctrica capacitativa y plantillas de silicona¹² con un 85% de mejoría, y cuando se compara su uso con infiltraciones, órtesis y arcos plantares¹³ con un 81% de mejoría, sin embargo Crawford²⁴ no encuentra diferencia cuando los compara con placebo, al igual que Cleland²², que obtiene mejores resultados con los estiramientos e iontoforesis, y que Zanon⁵ al compararlos con estiramientos.

La iontoforesis es más efectiva cuando se utiliza ácido acético¹⁷ para la disminución del dolor matutino, combinada con estiramientos y vendaje.

En el uso de material ortopodológico es difícil saber el nivel de eficacia ya que no se ha empleado como técnica única sino

como complemento de otras, salvo en el estudio de Gerald¹³ que compara el uso de órtesis (mejoría del 64%) el uso de arcos plantares (35% de mejoría) y aplicación de ultrasonidos (81% de mejoría) aunque el tamaño muestral es pequeño para ser representativo.

Dentro de los diferentes tipos de vendajes utilizados, tanto Osborne¹⁷ como Radford¹⁸ en sus ensayos clínicos encuentran mejoría en el dolor matutino al utilizar la técnica de lowdye, tanto en combinación¹⁷ como sola¹⁸.

Respecto al uso de estiramientos, una de las técnicas más utilizada (9 estudios), los resultados más significativos son la comparación entre el estiramiento del tendón de Aquiles frente al estiramiento de la fascia que realiza Di Giovanni⁹, con un intervalo de confianza del 95% y uno de los pocos estudios con nivel de evidencia 1 que concluye que el estiramiento de la fascia es más eficaz para tratar el dolor y la recuperación funcional. En ningún estudio se constataron malos resultados, si bien siempre en combinación con otras técnicas. Así parecen ser más eficaces cuando se comparan con ondas de choque radiales según Rompe²³. También aparecen más efectivos que los ultrasonidos continuos según Zanon⁵. En otros estudios donde se combinan con técnicas de acupuntura^{15,20} resultan ser igualmente eficaces, pero el tamaño muestral en ambos casos es reducido y en el ensayo de Karagounis²⁰ los sujetos eran atletas por lo que es difícil extrapolar estos resultados.

Al igual que la técnica anterior las ondas de choque son empleadas en nueve estudios, utilizándose las ondas de choque focales en dos estudios^{30,32} de forma única, la combinación de ondas de choque focales y radiales en otros dos^{27,28} mientras que la modalidad radial es la más utilizada en un total de cinco estudios. Gerdesmeyer²⁹ e Ibrahim³¹ utilizan los mismos

parámetros de impulsos pero el número de sesiones empleadas por Gerdesmeyer es tres obteniendo una $p=0,002$ frente a las dos de Ibrahim³¹ con una $p<0,001$ aunque en el caso primero el nivel de confianza es del 90% mientras que el tamaño muestral del segundo es significativamente menor (25 pacientes frente a 129) con lo cual se incrementa el índice de probabilidad de error tipo II donde se eleva la posibilidad de producir un falso negativo.

Respecto al uso de ondas de choque focales Haake³⁰ las emplea en un estudio frente a placebo en 272 sujetos no obteniendo evidencia de mejoría con el uso de estas, por el contrario Kudo³² encuentra mejores resultados con el uso de ondas de choque focales (49% frente al 33% del placebo) en una muestra de 114 sujetos pero el tiempo e intensidad empleados son distintos en ambos estudios.

Por último Babak²⁷ combina las ondas de choque radiales y focales frente a un grupo placebo tomando como resultado primario la medición ecográfica del grosor de la fascia y obteniendo una mejoría significativa en el grupo experimental, pero como limitación al estudio aparece de nuevo el reducido tamaño muestral (20 sujetos en cada grupo).

La aplicación de láser en el estudio de Kiritsi³³ revela una disminución del dolor del 59% frente al 26% en el grupo placebo aunque el tamaño muestral es pequeño (30 sujetos).

Así mismo el uso de la magnetoterapia por Telleria²⁵ comparándolo con el empleo de ultrasonidos no muestra diferencias significativas entre ambos grupos aunque existe una mejoría significativa, necesitando al menos 15 sesiones para lograr un 90% de disminución del dolor.

Las limitaciones que se han encontrado para realizar esta revisión tienen que ver

con la dificultad para homogeneizar la información sobre las distintas técnicas al aplicarse o bien con diferentes parámetros o en número distinto de sesiones o duración del tratamiento.

No existe una estandarización de los instrumentos usados para medir los resultados, usando una gran variedad de escalas lo que dificulta realizar una comparación entre tratamientos.

La mayoría de los autores (salvo Crawford F²⁴ y Kiritsi O³³), no informa del cegamiento a los terapeutas con respecto a la asignación del tratamiento a realizar, con lo cual se puede dar una falta de objetividad y puede llevar a afectar la validez de sus conclusiones.

Los estudios con un tamaño muestral pequeño, pueden influir en los resultados al provocar estimaciones sesgadas de los efectos del tratamiento, con riesgo en la validación externa.

En conclusión, para el tratamiento de la fascitis plantar se han descrito muchas intervenciones con una gran variedad de técnicas fisioterápicas siendo más frecuente la combinación de más de una de estas técnicas. Su efectividad ha resultado demostrada en la mayoría de los casos. La velocidad de recuperación de la fascitis plantar se cree que aumenta con el inicio del tratamiento en las primeras 6 semanas de aparición de los síntomas, sin embargo, esto aún no ha podido ser demostrado⁷.

La combinación de varias técnicas parece ser más efectiva que cualquier técnica utilizada de forma aislada por sí, salvo el uso de ondas de choque cuando otras técnicas han fracasado.

BIBLIOGRAFÍA

1. Sánchez R, Martínez A, Gómez B, Fraile PA, Fascitis Plantar. Tratamiento ortopodológico. Fisioterapia. 2007; 29(2): 106-12.

2. Alcántara S, Ortega E, García F, Sánchez I, Ferrero A, Aguilar JJ, et al. En: Dolor de tobillo y pie. Manual Sermef de Rehabilitación y Medicina Física. Madrid: Ed Médica Panamericana; 2006. p. 445-51.
- 3 - Goolsby MJ. Diagnosis and treatment of heel pain. J Am Acad Nurse Pract. 2003; 15: 485-6.
4. Manejo terapéutico de la fascitis plantar. Tribuna Médica [revista en internet] 2013 enero [acceso 8 mayo 2013];54(1). Disponible en: http://www.ibermutuamur.es/IMG/pdf/13_01_Tribunas_Medicas_bip54.pdf
5. Zanon RG, Kundrat A, Imamura M. Ultra-som continuo no tratamento da fasciite plantar crônica. Acta Ortop Bras. 2006;14:137-40.
6. Roxas M. Plantar Fasciitis: diagnosis and therapeutic considerations. Altern Med Rev. 2005;10:83-93.
7. Santa Coloma E, Khoury MA. Medicina basada en la evidencia: evidencia en el manejo no quirúrgico de la fascitis plantar. Rev. Asoc. Argent. Traumatol. Deporte. 2011; 18(2):81-89.
8. Lafuente A, O'Mullony I, Escribá M, Cura P. Fascitis plantar: revisión del tratamiento basado en la evidencia. Reumatol Clin. 2007;3(4):159-65.
9. DiGiovanni BF, Nawoczenski DA, Lintal ME, Moore EA, Murray JC, Wilding GE, et al. Tissue-specific plantar fascia-stretching exercise enhances outcomes in patients with chronic heel pain: a prospective, randomized study. J Bone Joint Surg Am. 2003;85:1270-77.
10. Gill LH, Kiezbak GM. Outcome of nonsurgical treatment for plantar fasciitis. Foot Ankle Int. 1996;17:527-32.
11. Stratton M, McPoil TG, Cornwall MW, Patrick K. Use of low frequency electrical stimulation for the treatment of plantar fasciitis. J Am Podiatric Med Assoc. 2009; 99 (6):481-88.
12. Rioja J, González A, Romo M, Cantalapiedra E. Tratamiento combinado de la fascitis plantar crónica en el adulto de edad superior a los 50 años. Rehabilitación (Madr. Ed. impr.). 2001; 35(2):90-94.
13. Gerald T. A prospective randomized trial using four treatment modalities for the treatment of plantar fasciitis. O J Foot Ankle. 2011;4(8): 1210-15.
14. Albormoz JC, Rodríguez A. Tratamiento de la fascitis plantar con férulas de uso nocturno. Rev. Venez. Cir. Ortop. Traumatol. 2004;36(2):98-101.
15. Kumnerddee W, Pattapong N. Efficacy of electroacupuncture in chronic plantar fasciitis: a randomized controlled trial. Am. J. Chin. Med. 2012;40(6):1167-76.
16. Tsai CT, Chang WD, Lee JP. Effects of short-term treatment with kinesiotaping for plantar fasciitis. J Musculoskelet Pain. 2010; Vol. 18 (1): 71-80.
17. Osborne HR, Allison GT. Treatment of plantar fasciitis by low-dye taping and iontophoresis: short term results of a double blinded, randomised, placebo controlled clinical trial of dexamethasone and acetic acid. Br J Sports Med. 2006;40(6):545.
18. Radford JA, Landorf KB, Buchbinder R, Cook C: Effectiveness of low-dye taping for the short-term treatment of plantar heel pain: a randomised trial. BMC Musculoskelet Disord. 2006;7: 64.
19. Rumbaut M, Cañizares D, Cuní R, Uranga JC, García E. El strapping como coadyuvante en el tratamiento ortopédico de la fascitis plantar. Rev. Habanera Cienc. Méd. 2009;8(2).
20. Karagounis P, Tsironi M, Prionas G, Tsiganos G, Baltopoulos P. Treatment of plantar fasciitis in recreational athletes: two different therapeutic protocols. Foot Ankle Spec. 2011;4(4): 226-34.
21. Greve JM, Grecco MV, Santos PR. Comparison of radial shockwaves and conventional physiotherapy for treating plantar fasciitis. Clinics (Sao Paulo). 2009;64(2):97-103.
22. Cleland JA, Abbot JH, Kidd MO, Stockwell S, Cheney S, Gerrard DF, et al. Manual physical therapy and exercise versus electrophysical agents and exercise in the management of plantar heel pain: a multicenter randomized clinical trial. J Orthop Sports Phys Ther. 2009;39(8):573-85.
23. Rompe JD, Cacchio A, Weil L Jr, Furia JP, Haist J, Reiners V, et al. Plantar fascia-specific stretching versus radial shockwave therapy as initial treatment of plantar fasciopathy. J Bone Joint Surg Am. 2010; 92(15):2514-22.
24. Crawford F, Snaith M. How effective is therapeutic ultrasound in the treatment of heel pain? Ann Rheum Dis. 1996; 55: 265-267.
25. Telleria T, Gonzalez JL, Rubio E. Magnetoterapia en la fascitis plantar. Rev Cuba Med fis reh [revista en internet] 2009. [acceso 10 de mayo de 2013]; 1(2): Disponible en: http://bvs.sld.cu/revistas/mfr/vol1_2_09/mfr03209.htm.

26. García EM, Álvarez CR, Rodríguez MI, Valdés A, González N. Fascitis plantar tratada con ondas de choque extracorpóreas. *Rev. Cuba. Ortop. Traumatol.* 2005;19(1).
27. Babak V, Sepideh S, Vahid B, Mehdi K, Hamidreza S. Extracorporeal shock wave therapy in patients with plantar fasciitis. A randomized, placebo-controlled trial with ultrasonographic and subjective outcome assessments. *Res J Med Sci.* 2012 ; 17(9): 834-838.
28. Lohrer H, Nauck T, V. Dorn-Lange N, Schöll J, Vester JC. Comparison of radial versus focused extracorporeal shock waves in plantar fasciitis using functional measures. *Foot Ankle Int.* 2010; 31(1):121
29. Gerdesmeyer L, Frey C, Vester J, Maier M, Weil L Jr, Weil L Sr et al. Radial extracorporeal shock wave therapy is safe and effective in the treatment of chronic recalcitrant plantar fasciitis: Results of a confirmatory randomized placebo-controlled multicenter study. *Am J Sports Med.* 2008; 36 (11):2100-09.
30. Haake M, Buch M, Schoellner C, Goebel F, Vogel M, Mueller I, et al. Extracorporeal shock wave therapy for plantar fasciitis: randomised controlled multicentre trial. *BMJ.* 2003;327:75-7.
31. Ibrahim MI, Donatelli RA, Schmitz C, Hellman MA, Buxbaum F. Chronic plantar fasciitis treated with two sessions of radial extracorporeal shock wave therapy. *Foot Ankle Int.* 2010; 31 (5):391-7.
32. Kudo P, Dainty K, Clarfield M, Coughlin L, Lavoie P, Lebrun C. Randomized placebo-controlled, double-blind clinical trial evaluating the treatment of plantar fasciitis with an extracorporeal shockwave therapy (ESWT) Device: A North American confirmatory study. *J Orthop Res.* 2006;115-23.
33. Kiritsi O, Tsitas K, Malliaropoulos N, Mikroulis G. Ultrasonographic evaluation of plantar fasciitis after low-level laser therapy: results of a double-blind, randomized, placebo-controlled trial. *Lasers Med Sci.* 2010; 25:275-81.
34. Zhang SP, Yip TP, Li QS. Acupuncture treatment for plantar fasciitis: a randomized controlled trial with six months follow-up. *Evid Based Complement Alternat Med.* 2009;10:1-10.
35. Weil LS Jr, Roukis TS, Weil LS, Borrelli AH. Extracorporeal shockwave treatment of chronic plantar fasciitis: indication, protocol intermediate results and comparison of results to fasciotomy. *J Foot Ankle Surg.* 2002;41:166-72.
36. Wang CJ, Chen HS, Huang TW. Shockwave therapy: a one-year follow-up study. *Foot Ankle Int.* 2002;23:204-7.
37. Moseley AM, Herbert RD, Sherrington C, Maher CG. Evidence for physiotherapy practice: a survey of the Physiotherapy Evidence Database (PEDro). *Aust J Physiother.* 2002;48: 43-9.
38. García J, Pascual R, Ortega E, Martos D, Martínez F, Hernández S. Estiramientos del tendón de Aquiles para la fascitis plantar. ¿Son efectivos? *Rehabilitación (Madr. Ed. impr.).* 2011;45(1):57-60.
39. Cotchett M, Landorf K, Munteanu S. Effectiveness of dry needling and injections of myofascial trigger points associated with plantar heel pain: a systematic review. *J Foot Ankle Res.* 2010; 3:18.
40. Sweeting D, Parish B, Hooper L, Chester R. The effectiveness of manual stretching in the treatment of plantar heel pain: a systematic review. *J Foot Ankle Res.* 2011; 4:19
41. Crawford F, Thomson C. Interventions for treating plantar heel pain (Cochrane Review). The Cochrane Database, Issue 3,2003.Oxford: Update Software Ltd; 1998.[citado 21 mayo de 2013]. Disponible en: <http://www.update-software.com/cochrane/>
42. Castiñeira Perez MC, Louro González A, Casariego Vales E. Fascitis plantar. Atención Primaria en la Red. [citado el 14 de julio de 2013]. Disponible en: <http://www.fisterra.com/guias-clinicas/fascitis-plantar/>
43. Buchbinder R. Clinical practice. Plantar fasciitis. *N Engl J Med.* 2004; 350(21): 2159-66.
44. Stadler TA, Johnson ED, Stephen MB. Clinical inquiries.What is the best treatment for plantar fasciitis? *J Fam Pract.* 2003; 52(9):714-7.

Anexo 1

Criterios metodológicos de calidad utilizados para evaluar estudios de intervención y prospectivos

A. Representatividad de la muestra final del estudio.

1. Aparición de al menos de uno de los siguientes elementos en el estudio: población diana completa, muestra seleccionada al azar o muestra indicada para representar a la población.
2. Al menos uno de los siguientes aspectos: descripción de la población elegida, método diagnóstico utilizado, tiempo de evolución de la patología.
3. Tasa de abandono o tasa de fracasos.

B. Calidad de los datos.

1. Si especifica las escalas de evaluación utilizadas.
2. Se realizaron evaluaciones comparativas antes y después de la intervención o entre grupos.
3. Se especifica resultados estadísticos.
4. Informe consentido firmado por el paciente o aprobación por un comité ético para desarrollar el estudio.