

ANÁLISIS BASADO EN LA EVIDENCIA

Jorge Mario Estrada



AREANDINA

Fundación Universitaria del Área Andina

MIEMBRO DE LA RED

ILUMNO

Análisis Basado en la Evidencia
Jorge Mario Estrada
Bogotá D.C.

Fundación Universitaria del Área Andina. 2018

Catalogación en la fuente Fundación Universitaria del Área Andina (Bogotá).

Análisis Basado en la Evidencia

© Fundación Universitaria del Área Andina. Bogotá, septiembre de 2018
© Jorge Mario Estrada

ISBN (impreso): **978-958-5462-57-1**

Fundación Universitaria del Área Andina
Calle 70 No. 12-55, Bogotá, Colombia
Tel: +57 (1) 7424218 Ext. 1231
Correo electrónico: publicaciones@areandina.edu.co

Director editorial: Eduardo Mora Bejarano
Coordinador editorial: Camilo Andrés Cuéllar Mejía
Corrección de estilo y diagramación: Dirección Nacional de Operaciones Virtuales
Conversión de módulos virtuales: Katherine Medina

Todos los derechos reservados. Queda prohibida la reproducción total o parcial de esta obra y su tratamiento o transmisión por cualquier medio o método sin autorización escrita de la Fundación Universitaria del Área Andina y sus autores.

BANDERA INSTITUCIONAL

Pablo Oliveros Marmolejo †
Gustavo Eastman Vélez

Miembros Fundadores

Diego Molano Vega
Presidente del Consejo Superior y Asamblea General

José Leonardo Valencia Molano
Rector Nacional
Representante Legal

Martha Patricia Castellanos Saavedra
Vicerrectora Nacional Académica

Jorge Andrés Rubio Peña
Vicerrector Nacional de Crecimiento y Desarrollo

Tatiana Guzmán Granados
Vicerrectora Nacional de Experiencia Areandina

Edgar Orlando Cote Rojas
Rector – Seccional Pereira

Gelca Patricia Gutiérrez Barranco
Rectora – Sede Valledupar

María Angélica Pacheco Chica
Secretaria General

Eduardo Mora Bejarano
Director Nacional de Investigación

Camilo Andrés Cuéllar Mejía
Subdirector Nacional de Publicaciones

ANÁLISIS BASADO EN LA EVIDENCIA

Jorge Mario Estrada



AREANDINA
Fundación Universitaria del Área Andina

MIEMBRO DE LA RED
ILUMNO

EJE 1

Introducción	7
Desarrollo Temático	8
Bibliografía	32

EJE 2

Introducción	34
Desarrollo Temático	35
Bibliografía	64

EJE 3

Introducción	66
Desarrollo Temático	67
Bibliografía	92

EJE 4

Introducción	94
Desarrollo Temático	95
Bibliografía	110

ANÁLISIS BASADO EN LA EVIDENCIA

Jorge Mario Estrada

EJE 1

Conceptualicemos



¿Qué es la práctica basada
en la evidencia?

En la práctica diaria de atención en salud desde cualquier ámbito (clínico, salud pública, auditoría y gerencia) la toma de decisiones es un factor fundamental que impacta directamente los objetivos, desde desenlaces clínicos en muchas patologías hasta la administración de recursos en salud que serán invertidos en la consecución de resultados. Esta toma de decisiones en un modelo clásico está basada en la experiencia individual (experiencia personal) o en un profesional especializado (autoridad, experto) que por su formación académica tiene altos conocimientos sobre un tema. Este modelo genera una jerarquía en el manejo de los saberes que se van transmitiendo de generación a generación de profesionales y hacen de la fundamentación de las prácticas y las decisiones un ciclo monótono, donde prima la tradición de haberlo hecho así (cotidianidad), recayendo siempre sobre lo aconsejado por el especialista en una materia.

La práctica basada en la evidencia reválúa totalmente dicho modelo y tiene sus cimientos en la verificación y la experimentación, es decir, es una práctica basada en la experimentación que, a partir de la misma, da respuesta a interrogantes prácticos y sustenta sus decisiones con el uso de los resultados de una investigación rigurosa. Una definición formal se da desde el ámbito de la práctica de la medicina, pero, acuñada al proceso de atención en salud general, no solo médico, se puede definir como un proceso para obtener y aplicar la mejor evidencia científica en el ejercicio de la práctica cotidiana en salud.

Este proceso se proyecta en los sistemas de salud actuales como una de las formas más objetivas de prestar servicios de salud. El actual Sistema General de Seguridad Social en Salud de Colombia ha

integrado métodos, prácticas y resultados de este proceso en las políticas de atención en salud. Esto es más notorio a partir del momento en que por normatividad se obliga a desarrollar o adoptar guías de práctica clínica basadas en la evidencia en los servicios de salud, la implementación de políticas frente a la depuración de un plan de beneficios bajo una escogencia con evidencia científica y la creación de entidades como el Instituto de Evaluación de Tecnologías en Salud (IETCS), el cual apoya los procesos de implementación de tecnologías en el país con el uso de metodologías de análisis basado en evidencia sobre el cuerpo de investigaciones clínicas publicadas.

La metodología del análisis basado en la evidencia se fundamenta en cinco pasos:

- Paso 1. Identificación del problema o escenario clínico.
- Paso 2. Formulación de una pregunta a partir de una situación clínica con un paciente.
- Paso 3. Búsqueda de la literatura (con una estrategia reproducible) de los estudios clínicos que dan respuesta a la pregunta planteada.
- Paso 4. Evaluación crítica de la literatura encontrada.
- Paso 5. Aplicación de las conclusiones de los estudios.

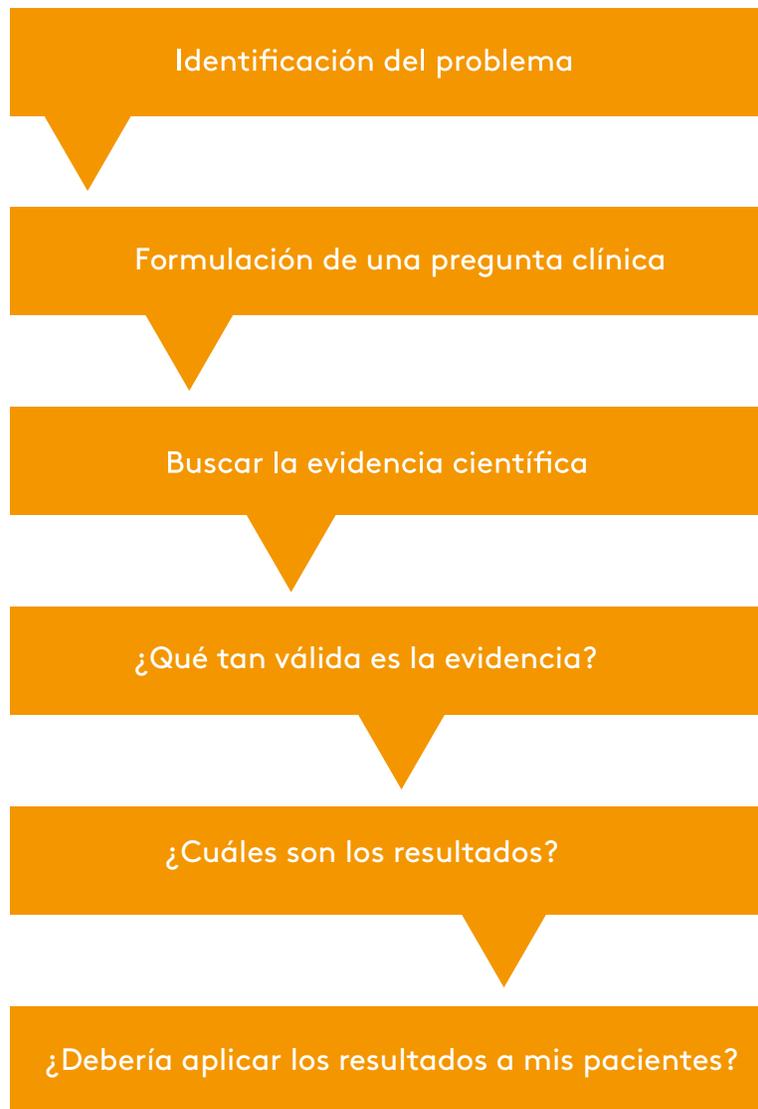


Figura 1. Uso de la literatura científica en la atención en salud
Fuente: Guyatt, Rennie, Meade y Cook (2008)



Instrucción

Lo invito a realizar la actividad de videopreguntas:
"¿Qué es la medicina basada en la evidencia?"

Diseños de estudios clínico-epidemiológicos acordes con los objetivos

La investigación clínica como proceso de investigación está fundamentada en la aplicación del método científico sobre un ámbito clínico, es decir, está orientada a dar respuesta a diversas situaciones que requieren de verificación o experimentación sobre la atención de los seres humanos. Toda la evidencia generada como parte de la solución a preguntas de investigación del ámbito clínico se da por la implementación de diferentes diseños de investigación en ciencias de la salud que permiten obtener rigurosidad en las conclusiones emitidas en dichas investigaciones.

Estos **diseños** presentan diferentes niveles en la validez de los resultados, según el objetivo de investigación. Para ser un poco más claro: imagine que se desea obtener información sobre la relación entre un factor de riesgo y un desenlace o resultado en salud negativo. Si para dar respuesta a esta pregunta que tiene como objetivo establecer la relación existente entre dos variables un investigador implementa casos, podría tener como resultado que dicho factor de riesgo prevalece entre todos los casos evaluados, y podría pensarse que con este hallazgo es suficiente para confirmar dicha relación. Sin embargo, desde el punto de vista metodológico, la ausencia de un grupo alterno o control para comparar (es decir, un grupo de individuos sin la enfermedad) debilita dicha conclusión, únicamente cuestionando con respecto a: ¿cómo es la presencia del factor de riesgo entre otros individuos sin la enfermedad?, ¿se distribuiría de la misma forma el factor de riesgo o tendría una distribución (presencia) igual o menor que en el grupo de enfermos? Según el ejemplo anterior, si bien se utiliza un diseño para valorar una pregunta de investigación, este no permite generar la conclusión con la mayor validez posible, aunque la implementación del diseño de investigación hubiese sido correcta. De esta manera, se hace necesario conocer los diferentes diseños de investigación y los casos más aconsejables en los que se pueden utilizar para dar respuesta a diferentes preguntas de investigación.



Diseños de estudios

Metodología implementada para dar respuesta a preguntas de investigación científica.

Tipos de diseños

Se pueden organizar los diseños de investigación clínico-epidemiológica en dos grandes grupos, según el papel que juega el investigador en cada uno de ellos (Manterola y Otzen, 2014). Cuando el investigador asume el papel de observador únicamente y se encuentra al margen de toda posible intervención en los participantes del estudio, estamos frente a diseños que tienen un propósito descriptivo u observacional analítico, es decir, su papel es medir exposiciones y desenlaces y, mediante métodos estadísticos, resumir la situación

observada, dando cuenta de cómo están los individuos o realizando una comparación entre grupos de ellos, algunos definidos para el estudio y otros formados tras obtener los resultados de las mediciones. El ámbito de aplicación de los diseños descriptivos u observacionales analíticos, dentro de lo epidemiológico, está dado para la estimación de prevalencias e incidencias, la identificación de factores asociados a un resultado o enfermedad, la evaluación de pruebas diagnósticas y el estudio de factores pronósticos.

Por otro lado, cuando el papel del investigador tiene que ver con la intervención de los sujetos en estudio (aplicación controlada de intervenciones farmacológicas o no farmacológicas), se está ante un diseño de investigación experimental, el cual permite la evaluación del impacto de intervenciones de manera controlada, siendo este diseño el de mayor validez ante la conclusión del estudio de la relación de dos variables: la intervención y el resultado. Como se explicó, el ámbito de aplicación de estos diseños es la evaluación de la efectividad de las intervenciones en salud.

Taxonomía	Diseño	Pregunta u objetivo que contesta
Descriptivo u observacional	Serie de casos	Descripción de un conjunto de pacientes con un hallazgo o patología común.
	Corte transversal	Estimar la prevalencia de una enfermedad o condición. Evaluar el desempeño de pruebas diagnósticas.
	Casos y controles	Establecer una relación entre un factor de riesgo o protector y una enfermedad o desenlace relevante.
	Cohorte	Establecer una relación entre un factor de riesgo o protector y una enfermedad o condición. Evaluación del pronóstico e historia natural de la enfermedad.
Experimental	Ensayo clínico	Evaluar la efectividad de intervenciones farmacológicas o no farmacológicas.

Tabla 1. Diseño de investigación clínico-epidemiológica según la taxonomía y la pregunta u objetivo
Fuente: propia



Instrucción

Para complementar esta información, revise el organizador gráfico que encontrará en los recursos de este eje.

Características principales de los diseños de investigación

Descriptivos u observacionales analíticos

- **Reporte o serie de casos**

El reporte o serie de casos es el diseño más elemental. Se limita a describir un grupo de sujetos con la condición de interés del estudio (enfermedad), por ende, carece de algún grupo de comparación. Se hace una descripción del evento desde los puntos de vista clínico y paraclínico, y de cualquier otra variable de interés para el investigador. Generalmente, son muy utilizados en el inicio de la investigación epidemiológica clínica más avanzada, ya que sus resultados se orientan más a formular hipótesis que sean comprobadas con diseños de mayor nivel de evidencia.

- **Estudio de corte transversal (pruebas diagnósticas)**

Es un diseño en el que todas las mediciones se hacen en un solo momento del tiempo, por lo que no existe seguimiento de los individuos. Generalmente, sirve para evaluar la magnitud de un problema en un momento dado. Es frecuentemente utilizado en la salud pública para la medición de indicadores de líneas base, prevalencia en enfermedades crónicas o la presencia de factores de riesgo ya conocidos dentro de una población. Dentro del ámbito clínico, su uso más frecuente está relacionado con la evaluación de pruebas diagnósticas, dado que en un mismo momento son aplicadas las pruebas diagnósticas en evaluación, seguidas de la clasificación por una prueba de oro (*gold estándar*) o de referencia.

- **Estudios de cohorte**

Son estudios observacionales analíticos. Un grupo de individuos se define sobre la base de la presencia o ausencia de una exposición, factor de riesgo o pronóstico para una enfermedad. Luego, se realiza un seguimiento por un periodo, de modo que se evalúa la ocurrencia del resultado (la enfermedad). Son estudios con los cuales se pueden estimar la incidencia o el riesgo de padecer un evento. En relación a uno de los principios que sustenta la causalidad, este diseño respeta la temporalidad en que el factor de riesgo o pronóstico está presente antes de ocurrir el desenlace estudiado.

- **Estudios de casos y controles**

Este diseño también hace parte de los observacionales analíticos con una temporalidad retrospectiva, es decir, inicia la selección de un grupo de sujetos de estudio por su condición de enfermos y lo compara con un grupo de individuos sin la condición o eventos de estudio (controles). Luego, se compara entre estos dos grupos las exposiciones que hayan ocurrido en el pasado de los individuos, garantizando que esta exposición se dio antes de haber sido caso o control.

Experimentales

- **Ensayos o experimentos clínicos**

Son estudios en los que el investigador asigna la exposición o no a una determinada variable (intervención) y observa el efecto que tiene sobre un determinado resultado. La forma más común de este tipo de estudios es tener al menos dos grupos. Uno de los grupos recibe el nuevo tratamiento o intervención en prueba, mientras que el otro recibe la intervención estándar o habitual. En ocasiones, ante la ausencia de este, se compara contra un placebo (en el caso de ensayos clínicos con medicamento corresponde a una forma igual al medicamento en prueba, pero sin acción farmacológica). La principal ventaja de este tipo de diseño frente a un estudio observacional es la fuerza de la inferencia de causalidad, siendo utilizado para evaluar la eficacia de nuevas drogas, tratamientos e intervenciones.



Lectura recomendada

Estudios observacionales: los diseños utilizados con mayor frecuencia en investigación clínica

Carlos Manterola y Tamara Otzen

Formulación de la pregunta con el método PICO

Existen diferentes tipos de preguntas que el clínico y el practicante pueden formular en los ámbitos de la atención. Lo que se pregunta incidirá directamente sobre dónde se busca la respuesta y el tipo de diseño a utilizar.

Tipos de preguntas a formular

Pueden existir múltiples preguntas y escenarios en la práctica; sin embargo, se estudiarán los cinco enfoques más comunes en la literatura sobre la práctica basada en la evidencia.

- **Terapia:** determinar el efecto de las intervenciones en los resultados importantes sobre los pacientes (síntomas, función, morbilidad, mortalidad, costos).
- **Daño:** determinar los efectos de los agentes potencialmente nocivos (incluyendo terapias del primer tipo de pregunta) sobre los resultados importantes para el paciente.
- **Diagnóstico diferencial:** en pacientes con una presentación clínica particular, establecer la frecuencia de los trastornos subyacentes.
- **Diagnóstico:** establecer el poder de una prueba para diferenciar entre aquellos con y sin una condición o enfermedad objetivo.
- **Pronóstico:** estimar el curso futuro de un paciente.

Estos tipos de preguntas tienen una relación directa con los diseños de investigación. Un punto fundamental en el análisis basado en la evidencia es la transformación de una situación clínica que amerita respuesta a una pregunta contestable a través de los estudios publicados. Esta transformación fácilmente se logra con una metodología de formulación de preguntas, la cual es conocida como formulación PICO, la cual corresponde a:



Figura 2.
Fuente: propia

Los componentes deben estar explícitos en la formulación final de la pregunta. A continuación, se justifica y desglosa cada componente de la pregunta clínica:

- **Pacientes:** se piensa en el paciente o en el entorno que se está tratando. Se busca identificar las características clínicas que influyen en el problema que son relevantes para su práctica y que afectarían la relevancia de la investigación. Esto ayudará a que sea lo más específico posible en esta etapa. Dentro de la situación clínica que amerita ser contestada, se deben identificar los pacientes protagonistas. Esta identificación se da en términos de variables como edad, sexo, severidad o estadio de la enfermedad, ámbito donde se realiza la atención, entre otras que permitan especificar con el detalle que el clínico defina para la aplicabilidad del resultado.
- **Intervención:** corresponde a la variable que dirige la situación clínica que genera la duda. Esta variable, en caso de que la pregunta sea sobre una terapia o intervención, puede ser un fármaco. En una pregunta dirigida a un diagnóstico, podría ser una prueba o test. Si es sobre el daño o la etiología, puede ser la exposición a un agente ambiental.

- **Comparación:** hace referencia a la alternativa que se tendría si no se aplicara el tratamiento. Generalmente, puede ser una terapia estándar, pero se debe pensar en esta etapa sobre las alternativas. En caso de que el escenario clínico haga referencia a un diagnóstico, este puede ser una prueba que hasta el momento está siendo usada. La comparación debe ser interpretada en el contexto de lo que sería su práctica normal.
- **Outcome (resultado):** hay que hacer una distinción importante entre el resultado que es relevante para el paciente o problema y las medidas de resultado desplegadas en estudios. Se debe dedicar un buen tiempo a pensar cuál es el resultado exacto que se espera. Muchas veces, es fácil concentrarse en la mortalidad y perder la importancia en aspectos de la morbilidad. Sin embargo, las medidas de resultado pueden ser guiadas por los propios estudios y no por su pregunta.



Ejemplo

Una mujer blanca de 55 años presenta diabetes *mellitus* tipo 2 e hipertensión. Su control glucémico es excelente con la metformina y no tiene antecedentes de complicaciones. Para controlar su hipertensión, toma una pequeña dosis diaria de un diurético tiazídico. Durante un periodo de seis meses, la presión está cerca de 155/88 mmHg (Guyatt, Rennie, Meade y Cook, 2008). Le atrae al clínico tratante de esta paciente ¿qué terapias antihipertensivas estarían relacionadas mejor con metas de control de cifras tensionales diastólicas?, ¿cercanas a 80 mmHg o a 90 mmHg? De esta forma, para dar inicio a sustentar la respuesta con evidencia, se utiliza la estrategia PICO:

- **Pacientes:** paciente hipertensa diabética tipo II sin complicaciones diabéticas.
- **Intervención/comparación:** cualquier agente antihipertensivo para obtener presiones diastólicas blanco de 90 mmHg versus una presión blanco de 80 mmHg.
- **Outcome:** enfermedad cerebrovascular, infarto de miocardio, muerte de causa cardiovascular y mortalidad general.

Lo anterior permite entablar una búsqueda de evidencia de manera tal que se recupere información científica que aporte a la respuesta y la toma de decisiones en la práctica clínica.



Lectura recomendada

Eficacia de la terapia cognitivo conductual en mujeres con cáncer de mama

Ana Lucía González García, Angelina González Hurtado y Benito Estrada Aranda

Búsqueda en bases de datos de literatura científica



Posterior a la especificación de una pregunta, nos trasladamos a la búsqueda de la literatura. Actualmente, existen diferentes bases de datos o motores de búsqueda que acumulan información de manera organizada e indexada y, sobre todo, revistas científicas examinadas por pares, lo cual nos permite confiar en gran medida en el proceso de revisión que se les realizó frente a la calidad metodológica mínima para ser publicadas, sin embargo, esto no garantiza ni exime al lector de hacer su propia lectura crítica para generalizar los resultados correctamente.

Fuentes de información

Según Guyatt, Rennie, Meade y Cook (2008), podemos clasificar las fuentes para búsquedas de evidencia en cuatro tipos:

- **Sistemas:** parecidos a un libro de texto. Son recursos que resumen e integran evidencia clínica con otros tipos de información dirigida a la práctica clínica. Algunos ejemplos son:
 - DynaMed Plus (www.acponline.org/clinical-information/clinical-resources-products/dynamed-plus)
 - UpToDate (www.uptodate.com)
 - Clinical Evidence (www.clinicalevidence.com)
 - EBM Guidelines (<http://onlinelibrary.wiley.com/book/10.1002/0470057203>)
- **Sinopsis:** resúmenes de estudios y revisiones sistemáticas que incluyen guías o consejos para la aplicación por expertos clínicos. Algunos ejemplos son: ACP Journal Club (www.acpjc.org) e InfoPOEMs (www.infopoems.com).
- **Resúmenes:** revisión sistemática de artículos y guías de práctica clínica. El ejemplo más representativo de este tipo de fuente de información es la Colaboración Cochrane en español e inglés (www.cochrane.org).
- **Estudios:** bases de datos de estudios primarios. Dentro de esta categoría se encuentra una gran cantidad de bases de datos, generalmente adquiridas por las instituciones de educación superior como fuentes de consulta para sus estudiantes y docentes. Algunas de estas funcionan con cobros y otras prestan servicios de motores de búsqueda. Entre las más comunes están: Medline, PubMed, Embase, Biblioteca Virtual en Salud (BVS), EBSCO y CINAHL.

Para efectos de este módulo, se abordarán únicamente las herramientas de utilización de PubMed, por ser una fuente conformada por más de 9000 revistas con diferentes indexaciones, búsqueda sin costo y acceso libre. Además, puede clasificar la información según el acceso completo libre o con pago. Las herramientas que utiliza PubMed son aplicables a otras de las fuentes de información mencionadas.

El ingreso a PubMed se hace a través de la dirección, como se muestra en la figura:

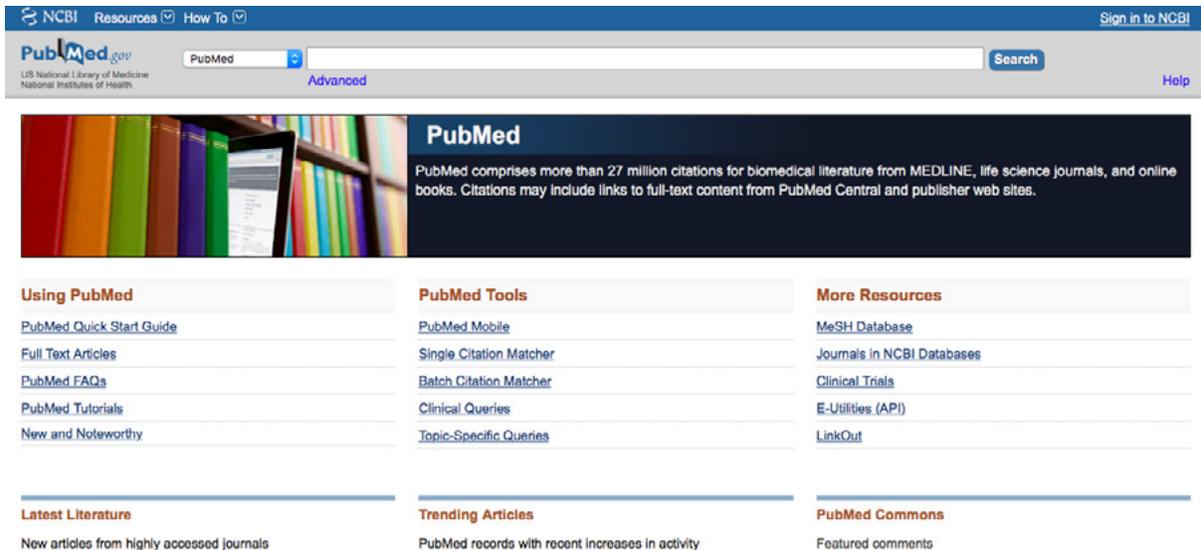


Figura 3. Fuente de información de estudios primarios PubMed
Fuente: PubMed

Búsqueda

Para iniciar la búsqueda de los estudios que den evidencias de resultados publicados sobre la pregunta formulada, debemos tener términos (palabras) que nos guíen. La elección de los términos de uso será fácil si tenemos desarrollada la pregunta en el formato PICO. PubMed utiliza el corrector ortográfico de Google y hace el trabajo de encontrar sinónimos para los términos con solo poner una frase o palabra por concepto de la formulación PICO. Generalmente, se utilizan tres o más conceptos.

Retomando el ejemplo del tratamiento de hipertensos diabéticos tenemos los términos así:

Elementos	Términos
<p>Pacientes: paciente hipertensa diabética tipo II sin complicaciones diabéticas.</p>	<ul style="list-style-type: none"> ● Hipertensión. ● Diabetes <i>mellitus</i> tipo II. ● Sin complicaciones diabéticas.
<p>Intervención: cualquier agente antihipertensivo para obtener presión diastólica blanco de 90 mmHg.</p>	<ul style="list-style-type: none"> ● Antihipertensivos. ● Presión diastólica. ● Metas.
<p>Comparación: cualquier antihipertensivo para obtener presión diastólica blanco de 80 mmHg.</p>	<ul style="list-style-type: none"> ● Antihipertensivos. ● Presión diastólica. ● Objetivo. ● Metas.
<p>Outcome (resultado): enfermedad cerebrovascular, infarto de miocardio, muerte de causa cardiovascular y mortalidad general.</p>	<ul style="list-style-type: none"> ● Mortalidad general. ● Infarto al miocardio. ● Mortalidad cardiovascular. ● Enfermedad cerebrovascular.

Tabla 2. Confirmación de términos o palabras para la búsqueda en PubMed o cualquier otra base de datos
Fuente: propia

Hay dos tipos principales de estrategias para buscar en bases de datos bibliográficas:

- La primera estrategia de búsqueda es denominada “libre”, funciona de tal manera que busca la ocurrencia de palabras o frases específicas en el registro bibliográfico del artículo. Estas palabras serán buscadas en cualquier parte del artículo (título, *abstract*, introducción, materiales y métodos, resultados, discusión y conclusiones). Esta estrategia es bastante amplia; sin embargo, tiende a traer información que puede no ser relevante para la pregunta formulada.
- La segunda estrategia, llamada “búsqueda por términos clave”, a los que PubMed denomina “MeSH”, hace referencia a términos que están conectados con las palabras clave que todos los artículos contienen y por los que se encuentran indexados bajo encabezamientos de temas. Si se busca un encabezado específico, se recuperan muchos materiales potencialmente relevantes.



Instrucción

Lo invito a realizar la actividad de juego de roles:
“Sustentando un diseño adecuado para responder una pregunta clínica específica”.

Operadores booleanos, límites, términos MeSH o claves

La mayoría de las bases de datos permiten crear una consulta escribiendo varias palabras o términos que se pueden combinar mediante operadores booleanos y otras herramientas.

Los operadores booleanos

Son aquellos que nos permiten combinar dos o más palabras, incluso con un orden jerárquico (utilizando paréntesis). Los dos más representativos son "AND" y "OR". El primero de ellos especifica que las dos palabras que se unen mediante el operador atraigan documentos en los que obligatoriamente los dos términos deben estar en alguna parte del artículo. Este conector hace que la búsqueda sea más específica y acota los resultados. El segundo operador une dos o más palabras, pero la acción que se ejecuta es la búsqueda del artículo que incluya cualquiera de los dos términos. Esto hace que los resultados sean más amplios. Es utilizado cuando se tienen varios sinónimos de un término.



Ejemplo

"Antihypertensive drugs" AND "diastolic pressure" AND Diabetes.

The screenshot shows the PubMed search interface. The search bar contains the query: "antihypertensive drugs" AND "diastolic pressure" AND Diabetes. The search results are displayed in a list format. The first three results are:

- [Profile development of noncommunicable chronic diseases in a Brazilian rural town.](#)
Di Pietro G, Cardoso DS, da Silva HM, Santos JC, Dos Santos JR, Simões RA.
J Am Coll Nutr. 2015;34(3):191-8. doi: 10.1080/07315724.2014.928162. Epub 2015 Mar 9.
PMID: 25751416
- [Level of blood pressure control among hypertensive patients on follow-up in a regional referral hospital in Central Kenya.](#)
Mutua EM, Citonga MM, Muthia B, Muiruri N, Cheptum JJ, Maingi T.
Pan Afr Med J. 2014 Aug 5;18:278. doi: 10.11604/pamj.2014.18.278.4308. eCollection 2014.
PMID: 25489372 Free PMC Article
- [Association of serum uric acid with level of blood pressure in type 2 diabetic patients.](#)
Rafieian-Kopaei M, Behradmanesh S, Kheiri S, Nasri H.
Iran J Kidney Dis. 2014 Mar;8(2):152-4.
PMID: 24685739 Free Article

The search details section shows the query: "diabetes"[All Fields] OR "diabetes insipidus"[MeSH Terms] OR ("diabetes"[All Fields] AND "insipidus"[All Fields]) OR "diabetes insipidus"[All Fields].

Figura 4. Búsqueda libre con términos y combinación con el operador booleano AND
Fuente: PubMed

Nótese en este ejemplo el uso de comillas para especificar que son frases y no palabras.



Ejemplo

("Antihypertensive treatment" OR "antihypertensive drugs") AND "diastolic pressure" AND Diabetes.

PubMed
US National Library of Medicine
National Institutes of Health

Search: ("antihypertensive treatment" OR "antihypertensive drugs") AND "diastolic pressure" AND Diabetes

Format: Summary - Sort by: Most Recent - Per page: 20 -

Send to - Filters: Manage Filters

Find related data

Database: Select

Find items

Search details

("antihypertensive treatment"[All Fields] OR "antihypertensive drugs"[All Fields]) AND "diastolic pressure"[All Fields] AND ("diabetes mellitus"[MeSH Terms] OR ...)

Search

See more...

Recent Activity

Turn Off Clear

Q ("antihypertensive treatment" OR "antihypertensive drugs") AND "d... (3) PubMed

Q "antihypertensive treatment" OR "antihypertensive drugs") AND "AN... (PubMed

Q "antihypertensive treatment" AND "diastolic pressure" AND Diabete... (23) PubMed

Search results

Items: 1 to 20 of 37

<< First < Prev Page 1 of 2 Next > Last >>

Profile development of noncommunicable chronic diseases in a Brazilian rural town.

1. Di Pietro G, Cardoso DS, da Silva HM, Santos JC, Dos Santos JR, Simões RA. J Am Coll Nutr. 2015;34(3):191-8. doi: 10.1080/07315724.2014.926162. Epub 2015 Mar 9. PMID: 25751416 [Similar articles](#)

Level of blood pressure control among hypertensive patients on follow-up in a regional referral hospital in Central Kenya.

2. Mutua EM, Gitonga MM, Mbuthia B, Muiruri N, Cheptum JJ, Maingi T. Pan Afr Med J. 2014 Aug 5;18:278. doi: 10.11604/pamj.2014.18.278.4308. eCollection 2014. PMID: 25489372 [Free PMC Article](#) [Similar articles](#)

Association of serum uric acid with level of blood pressure in type 2 diabetic patients.

3. Rafieian-Kopaei M, Behradmanesh S, Kheiri S, Nasri H. Iran J Kidney Dis. 2014 Mar;8(2):152-4. PMID: 24665739 [Free Article](#) [Similar articles](#)

[How to manage a patient with chronic arterial hypertension during pregnancy and the postpartum period].

4.

Figura 5. Búsqueda libre con términos y combinación del operador booleano OR
Fuente: PubMed

En este ejemplo es de resaltar el uso del operador OR para indicar un posible sinónimo para el término "tratamiento antihipertensivo", el cual es especificado como "drogas antihipertensivas". Además, se usa paréntesis, de tal manera que se jerarquiza para que este se ejecute primero y luego el resultado se combine con los demás términos.

En los ejemplos, la inclusión del operador OR aumenta el número de resultados, pasando de 19 a 37 artículos.

Límites

Esta opción permite limitar los resultados en PubMed mediante el uso de aspectos individuales (por ejemplo, año de publicación, sexo de los participantes, idioma inglés y tipo de artículo, como un ensayo controlado aleatorio [ECA] o metaanálisis).

Esta opción es visualizada de manera directa al lado izquierdo de los resultados emitidos por PubMed en una búsqueda (ver figura).



Ejemplo

En el ejemplo en relación con el tratamiento con antihipertensivos para lograr cifras tensionales objetivo en pacientes con diabetes tipo II sin complicaciones esta pregunta se puede situar en la categoría de terapia, es decir, se requiere decidir qué intervención o terapia sería la adecuada para este tipo de pacientes. De esta manera, se limita la búsqueda al tipo de artículo "ensayo clínico" y a sujetos como "seres humanos". La decisión de limitar a ensayos clínicos está fundamentada en que el tipo de estudio que daría la mejor evidencia frente a una decisión sería un diseño experimental (ensayo clínico).

Nótese que los resultados cambiaron de 37 a siete artículos que cumplirían los criterios de los límites ingresados.

PubMed.gov
US National Library of Medicine
National Institutes of Health

Search: ("antihypertensive treatment" OR "antihypertensive drugs") AND "diastolic pressure" AND Diabete

Format: Summary - Go by: Most Recent - Per page: 20 -

Send to - Filters: Manage Filters

Find related data
Database: Select

Find items

Search details
("antihypertensive treatment"[All Fields] OR "antihypertensive drugs"[All Fields]) AND "diastolic pressure"[All Fields] AND "diabeton mellitus"[MeSH Terms] OR...

Recent Activity
Turn Off Clear

Article types: Clinical Trial, Review, Systematic Reviews, Customize...

Text availability: Abstract, Free full text, Full text

PubMed Commons, Reader comments, Trending articles

Publication dates: 5 years, 10 years

Species: Humans, Other Animals

Language: English, Customize...

Sex: Female, Male

Search results
Items: 7
Filters activated: Clinical Trial, Humans. Clear all to show 37 items.

- Aggressive antihypertensive strategies based on hydrochlorothiazide, candesartan or lisinopril decrease left ventricular mass and improve arterial compliance in patients with type II diabetes mellitus and hypertension.
Spoelstra-de Man AM, van Ittersum FJ, Schram MT, Kamp O, van Dijk RA, Ijzerman RG, Twisk JW, Brouwer CB, Stehouwer CD.
J Hum Hypertens. 2006 Aug;20(8):599-611. Epub 2006 May 4.
PMID: 16673014
Similar articles
- [HOTEL P--Hypertension Optimal Treatment with Enalapril Lachema-Pliva. Optimal treatment of hypertension with Enalapril Lachema-Pliva].
Spinar J, Vitovec J.
Vnitr Lek. 2001 Apr;47(4):232-7. Czech.
PMID: 15635889
Similar articles
- Improved arterial compliance by a novel advanced glycation end-product crosslink breaker.
Kass DA, Shapiro EP, Kawaguchi M, Capriotti AR, Scuteri A, deGroof RC, Lakatta EG.
Circulation. 2001 Sep 25;104(13):1464-70.
PMID: 11571237 Free Article

Figura 6. Búsqueda libre con términos y combinación de los operadores booleanos OR y AND y uso de límites
Fuente: PubMed

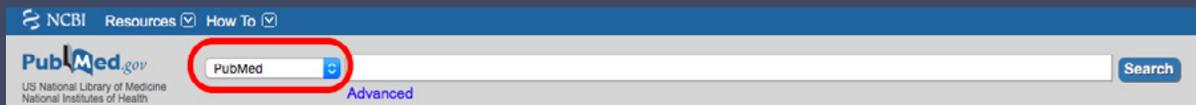
Términos MeSH o descriptores

Los términos MeSH son encabezados de términos médicos (vocabulario controlado) que se aplican a los artículos ubicados por PubMed. Este diccionario controlado está dentro de la página web de PubMed, de manera tal que el usuario puede ingresar y seleccionar el término. Se encontrará para cada término una definición que servirá al usuario para determinar si el término utilizado será aplicado en el contexto correcto, sobre todo cuando la búsqueda es en inglés, debido a que en traducción directa el significado del mismo cambia y, por ende, su contexto de uso.

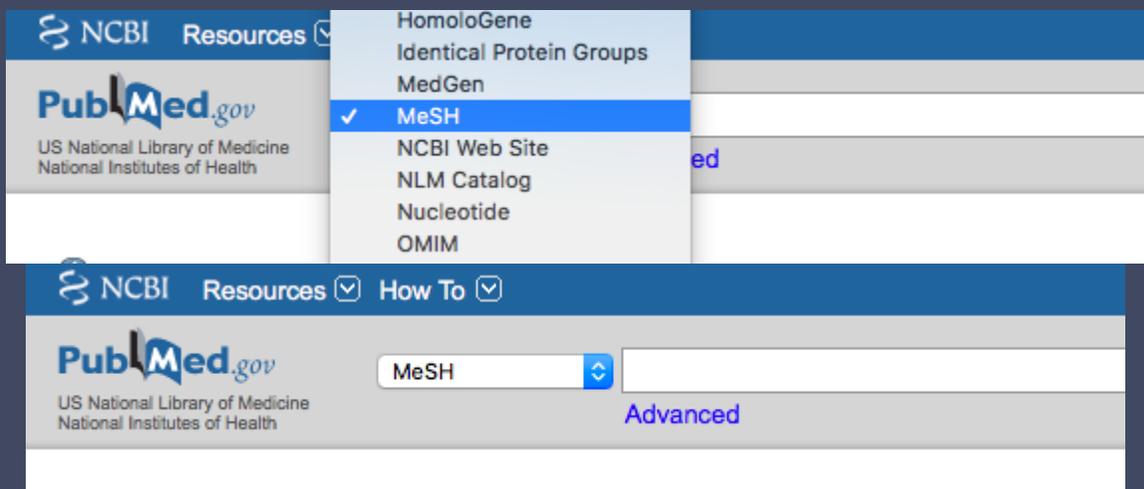
Estos términos MeSH también pueden ser combinados mediante operadores booleanos, así como con el uso de límites. La combinación de estas estrategias hace de los resultados obtenidos los mejores y más potentes, expresado esto en búsquedas con alta sensibilidad (ubicar los artículos relevantes para la pregunta) y alta especificidad (descartar o rechazar información “basura”), lo cual permite tener eficiencia en tiempo y objetivos logrados.

Para el uso de términos MeSH, se deben seguir los siguientes pasos:

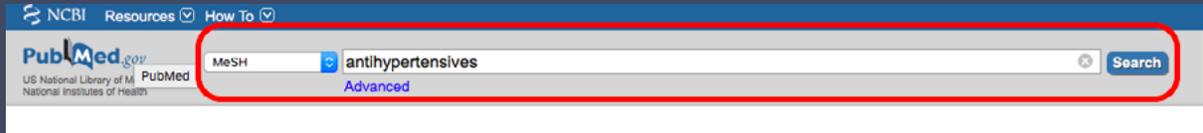
1. Ubicar en la página principal de PubMed la opción “MeSH”.



2. Seleccionar la opción “MeSH”.

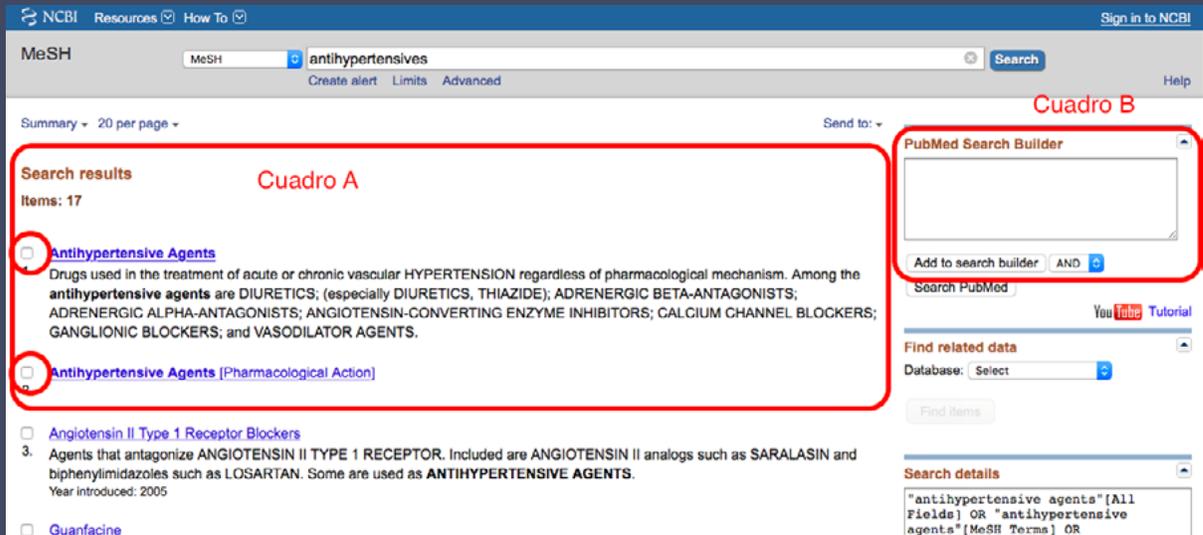


3. Colocar la palabra que se desea buscar entre los términos MeSH y hacer clic en buscar (la palabra debe estar en inglés).



4. Ubicar los términos MeSH a ser utilizados según la definición dada, la cual debe aplicar para el caso de búsqueda. Luego, hacer clic donde indican los círculos rojos (cuadro A), seleccionar el operador booleano con el cual se quieren combinar los términos (cuadro B) y llevarlos a la caja del buscador PubMed del lado izquierdo (cuadro B), haciendo clic en la opción "Add to search builder". En el caso del ejemplo, los términos que muestra el buscador MeSH que son útiles para la búsqueda están en el cuadro A.

El resultado de este paso sería:



5. Según la cantidad de términos a utilizar, se podría hacer una búsqueda de más palabras para agregar a esta estrategia que se está armando antes de lanzar la búsqueda. Como se muestra en la figura, no hay necesidad de borrar lo hecho hasta el momento. Solo se cambia la palabra nueva que quiere agregarse a la búsqueda para que sea localizada en la terminología MeSH (cuadro A). En el ejemplo se adiciona "diabetes" y, al hacer clic sobre buscar, obtenemos los resultados que se resaltan en el cuadro B de la figura. Rápidamente, ubicamos qué es relevante para nuestra búsqueda. Los pacientes con diabetes tipo II son la población objetivo de la pregunta formulada. Como en el paso anterior, se selecciona y, posteriormente, como se muestra en el cuadro C, se indica primero el operador booleano para su combinación con los términos agregados (paso 4) y se da clic en el botón "Add to search builder".

The screenshot shows the MeSH search interface. At the top, the search bar contains 'diabetes' and a 'Search' button. Below the search bar, the search results are displayed, with 'Diabetes Mellitus, Type 2' highlighted in a red box labeled 'Cuadro B'. To the right, a 'PubMed Search Builder' window is open, showing a search query: '"Antihypertensive Agents"[Mesh] OR "Antihypertensive Agents"[Pharmacological Action]'. Below the query, there is an 'Add to search builder' button and an 'AND' dropdown menu. A 'Search PubMed' button is also visible. The interface includes navigation links like 'First', 'Prev', 'Page 2 of 20', 'Next', and 'Last'.

6. Como las palabras están listas, la estrategia que PubMed va a aplicar para ubicar artículos relevantes quedaría como se muestra a continuación:

The close-up shows the 'PubMed Search Builder' window. The search query is: `("Antihypertensive Agents"[Mesh] OR "Antihypertensive Agents"[Pharmacological Action]) AND "Diabetes Mellitus, Type 2"[Mesh]`. Below the query, there is an 'Add to search builder' button, an 'AND' dropdown menu, and a 'Search PubMed' button, which is highlighted with a red circle.

La estrategia que el buscador construyó es una combinación de términos MeSH, lo cual se reconoce porque después de la palabra se coloca entre corchetes la opción “[Mesh]”, lo que indica que este término ya está definido como especial. Además, se visualiza un paréntesis que combina mediante un OR dos términos MeSH sinónimos y, finalmente, lo combina con otra palabra adicionada en relación a “diabetes mellitus”, también esta última indicada como un término MeSH. Por último, se debe hacer clic en el botón “Search PubMed”, el cual ejecutará la búsqueda de información.

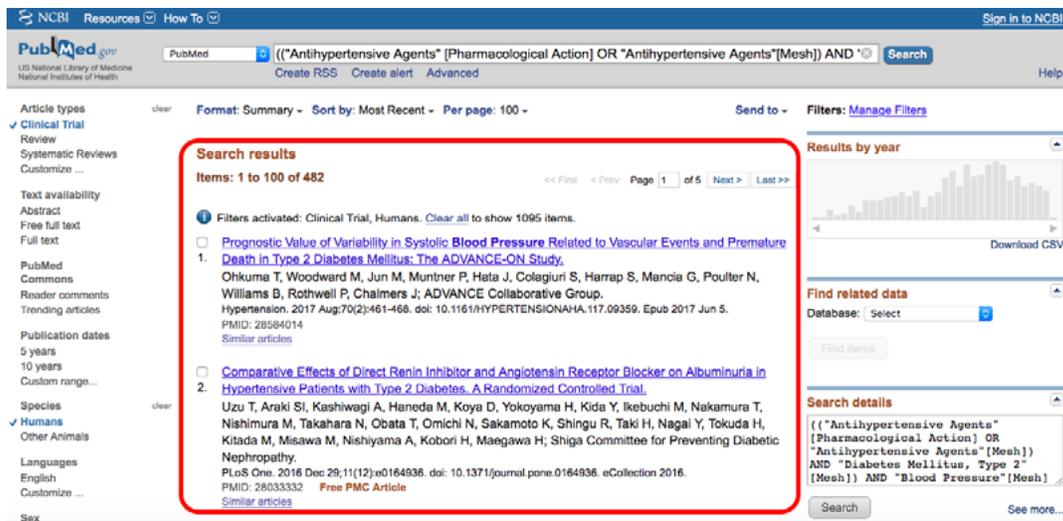


Figura 7. Resultados de la búsqueda
Fuente: PubMed

En la anterior figura, se pueden apreciar los resultados obtenidos, los cuales corresponden a 482 artículos relevantes, teniendo aún activado el uso de los límites previamente discutidos que corresponden solo a estudios en seres humanos y ensayos clínicos.

Un método a menudo acertado para enriquecer la búsqueda consiste en que, tras haber ubicado un artículo relevante, se ingresa en él haciendo clic en el título. Posteriormente, al lado derecho de su resumen (*abstract*), PubMed muestra una ventana con artículos que cree que están relacionados con el seleccionado previamente (recuadro rojo). Esta opción de artículos relacionados a menudo es fructífera para identificar más citas.

Effects of azilsartan medoxomil compared with olmesartan and valsartan on ambulatory and clinic blood pressure in patients with type 2 diabetes and prediabetes.

White WB¹, Cuadra RH, Lloyd E, Bakris GL, Kupfer S.

Author information

Abstract

BACKGROUND: Angiotensin receptor blockers (ARBs) are preferred antihypertensive therapies in patients with type 2 diabetes mellitus (T2DM). Azilsartan medoxomil (AZL-M) is a potent ARB for the treatment of stages 1-2 hypertension. We compared the efficacy, safety, and metabolic effects of AZL-M to both valsartan (VAL) and olmesartan (OLM), separately in patients with impaired fasting glucose (prediabetes mellitus) and T2DM.

METHODS: A pooled analysis of 3821 patients from three separate randomized placebo-controlled trials comparing the effects of AZL-M (40 and 80mg), OLM (40mg), VAL (320mg), and placebo on changes in ambulatory and clinic blood pressure (BP) among patients with hypertension and prediabetes mellitus or T2DM was performed. Two analysis pools were created to facilitate comparisons: Pool A included patients who received placebo, AZL-M or OLM and Pool B included those who received AZL-M or VAL. Within each pool, patients were stratified by glycemic subgroups (normoglycemic, prediabetes mellitus, or T2DM) based on hemoglobin A1c values. Changes from baseline in both 24-h and clinic SBP were the primary efficacy assessments.

RESULTS: Baseline 24-h mean SBPs were approximately 145 and 146mmHg in the prediabetes mellitus and T2DM subgroups, respectively; corresponding clinic SBPs were approximately 158 and 159mmHg. Baseline hemoglobin A1c values for each subgroup (both pools) were normoglycemic, 5.3%; prediabetes mellitus, 6.0%; and T2DM, 6.9%. Changes from baseline in 24-h or clinic SBP were significantly greater with AZL-M, 80mg compared with either OLM 40mg or VAL 320mg in all subgroups in each pool. Safety and tolerability were similar among the active treatment and placebo subgroups.

CONCLUSION: These analyses indicate that AZL-M, 80mg/day lowers SBP by a greater magnitude than OLM or VAL at maximally approved doses in patients with prediabetes mellitus and T2DM. These findings have important clinical implications for this high-risk patient group.

PMID: 26786564 PMCID: PMC4947533 DOI: 10.1097/HJH.0000000000000839

Full text links



Save items



Similar articles

The comparative effects of azilsartan medoxomil and olmesartan [J Clin Hypertens (Greenwich)...]

Effects of the angiotensin receptor blocker azilsartan medoxomil versus [Hypertension. 2011]

Comparison of the novel angiotensin II receptor blocker azilsartan [J Clin Hypertens (Greenwich)...]

Review Azilsartan medoxomil: a review of its use in hypertension. [Clin Drug Investig. 2012]

Review Antihypertensive efficacy of olmesartan medoxomil or valsartan [Curr Med Res Opin. 2009]

See reviews...

See all...

Cited by 3 PubMed Central articles

Review Blood pressure control in type 2 diabetic patients. [Cardiovasc Diabetol. 2017]

Figura 8. Artículos relacionados con la búsqueda Fuente: PubMed

Para terminar con este eje, cabe resaltar que se abordaron los tres primeros pasos en la metodología del análisis basado en la evidencia, los cuales nos permitirán la praxis del método de análisis, descrito en los pasos 4 y 5.

También se revisaron elementos fundamentales en este tema, como el conocimiento sobre los diseños de investigación utilizables para dar respuesta a diferentes preguntas clínicas que se pueden formular para llevar a cabo una práctica basada en evidencia.

Además, se exploraron los aspectos más relevantes en la búsqueda de la evidencia, a través del buscador más famoso y conocido dentro del ámbito científico y sus diferentes estrategias de búsqueda de la literatura publicada. Así, este módulo ofrece la conceptualización básica para abordar la respuesta de la pregunta articuladora.



Video

Para finalizar, lo invito a revisar el videotutorial sobre búsquedas en PubMed.

[https://vimeo.com/247524945.](https://vimeo.com/247524945)

González García, A. L., González Hurtado, A. y Estrada Aranda, B. (2015). Eficacia de la terapia cognitivo conductual en mujeres con cáncer de mama. *Psicooncología*, 12(1), 129-140.

Guyatt, G. H., Rennie, D., Meade, M. O. y Cook, D. (2008). *Users' guides to the medical literature: a manual evidence-based clinical practice*. Nueva York, Estados Unidos, McGraw-Hill.

Manterola, C. y Otzen, T. (2014). Estudios observacionales: los diseños utilizados con mayor frecuencia en investigación clínica. *International Journal of Morphology*, 32, 634-645.

ANÁLISIS BASADO EN LA EVIDENCIA

Jorge Mario Estrada

EJE 2

Analicemos la situación



Error aleatorio y sesgo en estudios clínicos



Los resultados de una investigación científica son propensos al error, es decir, se pueden obtener resultados que no reflejen el estado verdadero del fenómeno que se está investigando. La desviación del resultado puede ser atribuida a dos fuentes de error y su presencia en gran magnitud no permite tener conclusiones válidas.

Un estudio se desarrolla a partir del vacío de conocimiento en un área y con relación a un estado de un fenómeno desconocido por quien inicia el proceso de investigación. Puede afirmarse que el investigador siempre parte de la existencia de una verdad que desconoce, pero que, a través de un estudio clínico o epidemiológico, pretende aclarar (conocer) y ampliar su conocimiento del fenómeno para tomar decisiones.

En la búsqueda de aclarar esta verdad que subyace en una realidad, la forma en que se determinen los procedimientos y algunos de estos, *per se*, pueden introducir errores que lleven a subestimar o sobreestimar la verdad, esto se denomina en el ámbito de la metodología de la investigación científica sesgo o desvío del resultado.

De manera general, este desvío puede ser provocado por dos tipos de errores que se pueden cometer durante la planeación y el desarrollo de una investigación: el error aleatorio y el error sistemático o **sesgo**.



Sesgo

Desvío o falso resultado en un proceso de investigación científica.



Figura 1.

Fuente: Shutterstock/243096655

Error aleatorio

Este error es constante en toda investigación, debido a que generalmente se trabaja con una muestra y no con toda la población, es decir, con un subgrupo que representa dicha población y en el que también está representado el fenómeno sobre el cual se quiere tener conocimiento. Sin embargo, esta medición del fenómeno en una única muestra no deja de ser más que una aproximación al valor real del fenómeno. Esta aproximación es conocida en el ámbito de la estadística como una estimación puntual. En estas estimaciones, al ser generalmente calculadas en una muestra única, cabe siempre la posibilidad de que se dé un valor diferente si se selecciona una segunda muestra de la misma población. Ocurriría lo mismo si se tomara una tercera muestra, quedando claro que al obtener conclusiones sobre una estimación de una única muestra se tiene una alta relación con lo que se denomina incertidumbre. La incertidumbre se define mediante la siguiente pregunta: si tomase una muestra en la que se estimó un valor y luego tomase una segunda muestra para estimar el mismo valor, ¿los resultados serían diferentes? La pregunta que queda es: ¿cómo podría establecer que bajo condiciones de desarrollo de una investigación y solo teniendo la posibilidad de obtener una única muestra esta represente un valor acertado frente al valor real del fenómeno investigado?

El error aleatorio está directamente relacionado con trabajar con una porción de individuos (muestra) y no con el total de la población (universo) que permitiese conocer el valor real del fenómeno estudiado. Para efectos prácticos, el error aleatorio siempre está presente y puede ser causal de que las estimaciones sean alejadas de una verdad y, por ende, las conclusiones derivadas sean incorrectas. Sin embargo, la estadística antepone herramientas para controlar el problema de azar (error aleatorio) que se produce al trabajar con una única muestra de una población. Estas herramientas están encaminadas a la minimización del error aleatorio. Primero minimizándolo desde el diseño, es decir, ajustando un tamaño de muestra que sea lo suficientemente grande y permita controlar un grado de error máximo admisible por el investigador; segundo, usando técnicas de estadística de muestreo y estimación que cuantifiquen el error aleatorio presente en una medición y permitan reportarlo de manera tal que el lector pueda estar informado y lo integre a sus conclusiones.

Como un ejemplo de la presencia de este error se expone un caso:



Ejemplo

Suponga que se desea conocer el valor real del efecto (verdad subyacente que poca o ninguna vez se conoce) que tiene un nuevo antihipertensivo sobre la mortalidad por eventos cardiovasculares en una población de pacientes con hipertensión moderada a severa. Este efecto es expresado en la forma de diferencia de riesgo absoluto de muerte (existen otras medidas que serán revisadas más adelante) por eventos cardiovasculares. Este efecto puede tomar valor cero (0) indicando que no hubo diferencias en la mortalidad por estos



eventos entre quienes toman el antihipertensivo nuevo comparado contra algún medicamento convencional. Puede tomar valores de uno (1) indicando que el antihipertensivo nuevo reduce comparativamente con el antihipertensivo convencional el 100% de las muertes. Además, supongamos que este valor real del efecto es de 0,7. Posterior a esto, tenemos tres estudios diferentes con igual tamaño de muestra y probando la efectividad del nuevo antihipertensivo. Cada estudio obtuvo una muestra de 20 pacientes, 10 recibieron el nuevo fármaco y los restantes el medicamento convencional. En ambos grupos se midió la mortalidad. Los resultados reportados fueron:

Estudio 1 n = 20	Estudio 2 n = 20	Estudio 3 n = 20
DR = 0,46	DR = 0,65	DR = 0,73

Tabla 1.
Fuente: propia

Como se aprecia, los valores del efecto estimado puntualmente en cada estudio de hecho son diferentes. La población de donde fueron extraídas las muestras para estos estudios fue la misma, es decir, pacientes hipertensos susceptibles de beneficiarse de la nueva terapia.

Como se observa en el ejemplo anterior, por simple azar los resultados son diferentes en cada estudio. Si no se conoce el verdadero valor del efecto (normalmente es lo que intenta un estudio conocer), ¿cómo se puede saber que uno de estos estudios es una buena aproximación a ese valor verdadero, contando con la oportunidad de realizar el estudio una sola vez por costos? Esto deja entrever la presencia del error aleatorio en todo estudio realizado; sin embargo, de nuevo la estadística provee técnicas de estimación (serán revisadas más adelante) que se basan en la información de una única muestra, pero con la capacidad de calcular el nivel de error aleatorio o por azar que está relacionado con la medida puntual obtenida en la muestra, y permite hacer la inferencia sobre cuál sería la variabilidad observada si se hubiera realizado el mismo experimento, pero en muestras repetidas. Algunas técnicas nos permiten estimar qué valor de probabilidad está asociado a cometer un nivel de error cuando se trabaja con esta única muestra.

En conclusión, el error aleatorio está presente en todos los estudios. Puede ser identificado antes del estudio, controlado o minimizado, así como cuantificado para su asociación con el resultado puntual obtenido en la muestra. El azar y su control en los estudios clínicos y epidemiológicos tiene manejo con técnicas estadísticas.

Error sistemático o sesgo

Este error tiene un origen y un control totalmente distintos al aleatorio. Se origina cuando la implementación de procedimientos dentro de la investigación tiende a desviar sistemáticamente un resultado, alejándose del valor real del fenómeno estudiado.

Principalmente, se encuentra presente el sesgo en estudios dentro de la implementación de procedimientos de medición como tal, por ejemplo, cuando para determinar el efecto de un factor de riesgo ante una condición, esta es medida bajo protocolos o instrumentos no aptos o no validados para tal fin. Este es el caso de la determinación del uso de una prueba diagnóstica para su clasificación, la prueba *per se* puede presentar problemas con sus probabilidades de clasificación (sensibilidad y especificidad), es decir, en su capacidad para detectar la enfermedad o para descartarla. De tal manera, este error genera una mala clasificación entre los sujetos y los errores están relacionados directamente con el desvío en la estimación real del fenómeno estudiado. Las desviaciones de estos resultados son: mostrar una magnitud y fuerza más altas de las reales o, por el contrario, más bajas. En el peor de los casos, serían contrarias a las reales.

El sesgo principalmente se ve expresado cuando los procedimientos escogidos por el investigador hacen que las mediciones se desvíen de su valor real, siendo de esta manera aspectos metodológicos del diseño de la investigación los responsables del sesgo. Dichos aspectos metodológicos son controlados solo cuando se planean, no siendo posible su control posterior cuando se obtienen resultados. El problema del error sistemático es eminentemente metodológico, es decir, de proceder. No existe ningún control más que tomar o aplicar correctamente los procedimientos en una investigación.



Ejemplo

Continuando con el ejemplo anterior, la diferencia en los resultados de los tres estudios expuestos puede ser atribuida a que los grupos analizados (10 sujetos por grupo en cada estudio) hayan presentado discrepancias entre ellos por la presencia diferenciada de un factor pronóstico. Por ejemplo, uno de los grupos podría tener en promedio una edad 10 años mayor, comparado con el otro. Esta diferencia de edad puede tener como efecto que los pacientes respondan con menor efectividad a la intervención y, así, el efecto visto no sea real.

Como se puede evidenciar, el sesgo y el error aleatorio influyen sobre los resultados finales de cualquier investigación y son los responsables de la inexactitud y la imprecisión de una estimación. En términos prácticos, un paso fundamental en el análisis basado en la evidencia es tratar de identificar si los resultados observados en una investigación se vieron afectados por algunos de estos errores y, de esta manera, evaluar la validez de las estimaciones y su disponibilidad para ser aplicadas en la práctica diaria.



Como se especificó en el eje 1, cuatro escenarios prácticos permiten establecer la aplicación de métodos basados en la evidencia. Las evidencias (investigaciones clínicas) que se recuperan con búsquedas deben ser analizadas con una mirada crítica en su desarrollo metodológico y en los resultados reportados. Para cada uno de estos escenarios se desarrollan aspectos metodológicos y de interpretación de los resultados posibles que permiten establecer estándares mínimos a evaluar en la validez de los resultados en cada escenario, es decir, en cada tipo de estudio que responde a un tipo de pregunta clínica planteada.

Ensayos clínicos



Figura 2.
Fuente: Shutterstock/691186021

Cada estudio debe ser evaluado en cuanto a la validez de su resultado, lo cual puede ser evidente mediante la valoración de aspectos que pudieron haberse afectado por la presencia de error aleatorio y del sesgo.

La situación de duda o desconocimiento sobre la efectividad de los resultados clínicos de un tratamiento o intervención en el paciente está sobre la mesa. Solo puede ser abordada con evidencia de un estudio clínico experimental, es decir, un ensayo clínico en el que la intervención es comparada de manera controlada frente a una alternativa, comúnmente una terapia estándar o, en su defecto, con un placebo. En la literatura científica, habitualmente estos estudios son denominados artículos que “prueban la efectividad de una terapia” y son los que de manera mayoritaria aportan el máximo nivel de evidencia en la toma de decisiones, esto atribuido al carácter experimental del diseño en este tipo de estudio.

Como todo diseño de investigación, este está sujeto a errores aleatorios y sistemáticos que pueden detectarse bajo el no cumplimiento de aspectos metodológicos mínimos en su desarrollo, implementación y análisis. Se abordarán a continuación algunos de estos aspectos a tener en cuenta en un futuro análisis crítico de ensayos clínicos.

Validez de los resultados en un ensayo clínico

Uno de los principales aspectos a determinar en un ensayo clínico para observar una diferencia entre los grupos evaluados (intervención versus control) es que dichos grupos están conformados de manera similar. Se podría afirmar que deberían estar sin diferencias al inicio del ensayo, esto permitiría establecer fácilmente que ningún factor pronóstico con presencia mayoritaria en uno de los grupos pudiera favorecer el resultado final.

Una de las razones por las que esto sucede (diferencias en factores pronósticos entre grupos al inicio del estudio) es que actúa la preferencia del paciente o del médico en determinar si un paciente es asignado al grupo de tratamiento o control. Diferentes factores o variables pronósticos pueden estar relacionados con la preferencia a recibir una intervención u otra dentro de un estudio. La edad del paciente, la severidad de su enfermedad y la comorbilidad son algunos ejemplos de estos factores, a los que también se les puede atribuir su contribución al resultado visto en una variable de desenlace principal tomada como medida de efectividad de la terapia, es decir, pueden tener un efecto sobre la medida evaluada, por ende, si se observan diferencias entre dos grupos en un ensayo clínico, podrían no solo ser o solo ser atribuibles a estos factores pronósticos y no a la intervención o terapia en evaluación.

Aleatorización

Para dar solución al sesgo que se produce por tener grupos desbalanceados frente a factores pronósticos conocidos, el ensayo debe implementar dentro de su diseño la aleatorización, importante técnica que

consiste en la asignación aleatoria (probabilística) que da iguales posibilidades de ser asignado a un grupo de tratamiento como al grupo de control. El poder de la aleatorización es que los grupos de tratamiento y control tienen más probabilidades de ser equilibrados con respecto a los factores pronósticos conocidos y desconocidos. Aunque la aleatorización es una técnica poderosa, no siempre crea grupos con pronóstico similar. Los investigadores pueden cometer errores que comprometen la aleatorización o la misma puede fallar debido a la mala suerte. Como punto suplementario, los investigadores que están reclutando pacientes en un ensayo pueden consciente o inconscientemente distorsionar el equilibrio entre los grupos, por tal motivo, esta aleatorización debe ser también encubierta al investigador. Generalmente, se utilizan centros externos que asignan mediante diferentes estrategias a qué brazo del estudio pertenece el paciente.

Otra causa del desequilibrio que se puede dar al inicio de un ensayo clínico es la baja cantidad de pacientes, de tal manera que el bajo número no permite que los mismos se asignen de manera homogénea. Finalmente, se puede verificar cómo funcionó la aleatorización cuando se despliegan resultados para ambos grupos en términos de los factores pronósticos tomados en cuenta en el estudio. Las diferencias deben ser mínimas o nulas. Cuando esta aleatorización no funciona, el resultado no está perdido.

Existen técnicas estadísticas que ajustan los resultados por dichas diferencias observadas y deben ser aplicadas por los investigadores.

Enmascaramiento

Si la aleatorización funciona, los grupos de tratamiento y control comienzan con un pronóstico similar. La aleatorización, sin embargo, no ofrece garantías de que los dos grupos seguirán siendo equilibrados en los factores pronósticos medidos al inicio. El enmascaramiento es la óptima estrategia para mantener el balance del pronóstico.

En la medida que sea posible, se deben enmascarar los siguientes actores en el desarrollo del ensayo clínico: pacientes, investigadores, recolectores de información, personal que mide el o los desenlaces principales y analista de datos, cada uno con efecto distinto. La aplicación del enmascaramiento garantizará mayor validez del resultado.

Pérdida de seguimiento

La realización de un ensayo clínico suele ser demorada y difícil de completar correctamente. Si se controla adecuadamente, menos del 80 % de los pacientes, los resultados pueden verse sesgados.

Idealmente, al final de un ensayo se debe conocer en cada paciente el estado de la variable clínica de efectividad seleccionada. Cuanto mayor sea el número de pacientes al cual se les desconoce su estado final (pacientes perdidos en el seguimiento), mayor afectación en la validez tendrá el estudio. La razón es que los pacientes con peores pronósticos pueden ser retenidos, mientras que quienes tengan resultados adversos o su resultado es bueno, pueden desaparecer.

Tiempo de estudio

Los estudios deben permitir tiempo suficiente para que los resultados que se están midiendo se manifiesten. El lector debe usar su juicio clínico para decidir si esto era cierto para el estudio que está evaluando y si la duración del seguimiento era apropiada para los resultados que le interesan.

Intención a tratar

Los investigadores también pueden afectar la aleatorización si se omiten de los análisis pacientes que no reciben el tratamiento asignado o, peor aún, cuentan los eventos que ocurren en pacientes no adherentes que fueron asignados al tratamiento versus el grupo de control. Dichos análisis sesgan los resultados si las razones de la no adherencia están relacionadas con el pronóstico.

Los investigadores previenen este sesgo cuando siguen el principio de intención de tratar y analizan a todos los pacientes del grupo al que fueron asignados al azar, independientemente de su adherencia.

Resultados en un ensayo clínico

Tras evaluar en primera instancia los aspectos metodológicos más importantes que garantizan en gran medida el control de sesgos en los ensayos clínicos, pasamos a evaluar los resultados que se reportan en estos estudios, los cuales son variados de acuerdo con el tipo de variable (cuantitativa o categórica) que fue medida como desenlace. Ante todo, debe tenerse claro que las medidas utilizadas frecuentemente son de origen estadístico y están basadas en la comparación de, al menos, dos grupos (tratamiento y experimental). Para efectos de sencillez en los elementos básicos de las medidas, se utilizarán únicamente ejemplos de un ensayo con dos brazos: un grupo que recibió el tratamiento nuevo o experimental y otro que recibió un tratamiento control, estándar o, como mínimo, placebo.

Las variables que pueden ser seleccionadas para medir la efectividad de una intervención pueden clasificarse de dos formas. Las llamadas cuantitativas hacen referencia a datos de tipo numérico. Generalmente, aluden a biomarcadores o bioclínicas como la tensión arterial, la concentración de un metabolito en sangre y la carga viral, todas dan como resultado de medírselas en el sujeto un valor numérico.

La otra forma de variable es la categórica. Asigna al sujeto un atributo entre unos posibles, por ejemplo: la curación (sí, no), la mortalidad (muerto/vivo), la aparición de infarto o no, la infección o no, etc. Se pueden colocar más de dos categorías entre sus posibilidades; sin embargo, para dar sencillez didáctica, se trabajará con variables categóricas con únicamente dos posibles resultados, también llamadas dicotómicas.

Variable cuantitativa

La medida más utilizada para estimar la asociación entre un tratamiento y un desenlace es utilizar un estimador que resuma el resultado de dicha variable cuantitativa en cada grupo en evaluación (intervención vs. control) y compare la diferencia entre los resultados. Estimadores como la media (o promedio) son frecuentemente utilizados, de esta manera, se procede a calcular el promedio entre el grupo de intervención y el promedio en el grupo de control, luego, se restan ambos promedios y se obtiene la medida denominada **diferencia de medias**.

Si dicha diferencia es igual a cero (0), esto se interpreta como que no hay diferencia alguna entre los resultados de un grupo experimental comparado con el grupo control. Por el contrario, si dicha diferencia absoluta tiene un valor distinto a cero, se apreciará que existe una diferencia en el resultado clínico evaluado atribuible a la intervención o al control.



Diferencia de medias

Estimador de las diferencias de dos medias en dos poblaciones independientes, clasificado junto con otros estimadores como tamaño del efecto.



Ejemplo

Supongamos que se evalúa un nuevo régimen profiláctico de anticoagulante (única dosis/día vs. una dosis/día fraccionada) en pacientes con riesgo de trombosis venosa profunda. Para evaluar su efectividad, se tomó como variable de desenlace la TP en segundos. Al finalizar el ensayo, se estimó que, por término medio, el grupo de una dosis diaria obtuvo un TP de 25 segundos, mientras que el grupo de una dosis/día fraccionada obtuvo un TP promedio de 35 segundos. La diferencia de medias sería:

$$m_1 = 25$$

$$m_2 = 35$$

$$Diff = 35 - 25 = 10$$

Por tanto, la diferencia obtenida corresponde a que, en promedio, el régimen de una dosis fraccionada alargó 10 segundos más el TP por encima del grupo de una única dosis, mostrando un efecto mayor sobre la anticoagulación.

Variable categórica

Cuando la variable de desenlace se mide de una forma dicotómica, es decir, con la posibilidad de dos atributos de tipo infectado/no infectado, curado/no curado, vivo/muerto, los resultados del ensayo clínico se pueden representar en una tabla de contingencia de 2x2.

		Desenlace		
		SÍ	NO	
Grupo experimental	A	B	<ul style="list-style-type: none"> ■ Diferencia de riesgo (DR) = $(c/c+d) - (a/a+b)$ ■ Riesgo relativo (RR) = $(a/a+b) / (c/c+d)$ ■ Reducción de riesgo relativo (RRR) = $(c/c+d - a/a+b) / (c/c+d)$ o $1-RR$ ■ Odds ratio (OR) = $(a/b) / (c/d)$ ■ Número necesario a tratar (NNT) = $1/DR$ 	
Grupo control	C	D		

Figura 3.
Tabla de 2x2 para representar los resultados de un ensayo clínico y las medidas para resumir la efectividad de un tratamiento
Fuente: propia

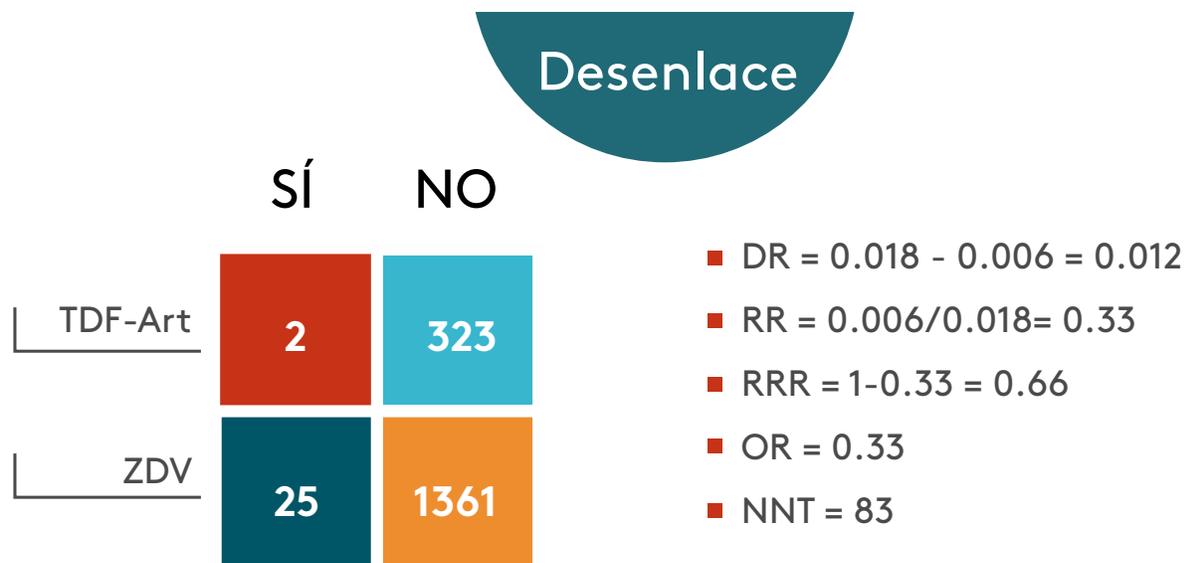


Figura 4. Resultados de un ensayo clínico comparando Tenofovir (TDF) con base en ART con Zidovudina únicamente (ZDV) para prevención de VIH perinatal
Fuente: Fowler et ál. (2016)

La primera medida a ser calculada es el riesgo. Se denomina riesgo a la posibilidad de desarrollar el evento o desenlace dentro de un grupo al que se fue asignado (experimental o control). También es llamado tasa del evento y se le asigna el nombre de riesgo de base al calculado entre el grupo de control. Es importante aclarar que cuando se denomina riesgo no obligatoriamente implica peligro, pues en algunos ensayos el desenlace principal no necesariamente es negativo (mortalidad, complicaciones), también puede ser positivo (curación, mejoramiento en severidad de la enfermedad). Entonces, no se hablaría de riesgo, sino de proporción de sujetos que hacen el evento. De manera general, se maneja el término riesgo. Se debe tener en cuenta siempre el desenlace.

Riesgo (proporción de evento) grupo experimental = $a/a+b$; valores posibles [0,1]

Riesgo (proporción de evento) grupo control = $c/c+d$; valores posibles [0,1]

Las siguientes medidas se desprenden del uso de los riesgos en ambos grupos.

Una primera medida intuitiva, como se mostró en la diferencia de medias, es obtener la diferencia entre los riesgos de ambos grupos, esta es denominada diferencia de riesgos (DR) o reducción de riesgo absoluta.

DR = $(a/a+b) - (c/c+d)$; DR con valores posibles en el intervalo cerrado [0,1]

La interpretación de la DR es la menos intuitiva, debido a que está evaluando de manera absoluta la diferencia de riesgo generada a causa de la intervención. Clínicamente, no podemos valorar de una manera fácil cuánto riesgo es suficiente o cuánto riesgo es alto o tan alto como para alarmarse. Valores de DR como en el ejemplo de la

tabla 3, el riesgo de infección perinatal de VIH para los grupos control (ZDV) y experimental (TDF-Art) fueron de 0,018 (1,8 %) y 0,006 (0,6 %), respectivamente, mostrando una diferencia de riesgo de 0,012 (1,2 %), para lo cual es claro que cuantitativamente hubo una diferencia de 1,2% en el riesgo de infección entre los dos antirretrovirales, siendo este resultado a favor de la terapia antirretroviral con TDF. Sin embargo, valorar si dicha diferencia tiene una magnitud importante se hace difícil, pues no tenemos la percepción clara cuantificada del riesgo medido. Quizá para el paciente no se ha valorado de manera importante esta diferencia como para someterse a una nueva terapia.

Otra forma de medida de resultado, pero con un carácter relativo, a diferencia de la anterior, es comparar los grupos mediante un cociente. Es medida es llamada riesgo relativo (RR) o razón de riesgos.

RR = a/a+b / c/c+d; RR con valores posibles en intervalo abierto (0,)

La interpretación del RR resulta más fácil y con mejor percepción sobre el resultado, esto tanto para el clínico como para los pacientes. Al ser un cociente entre el riesgo (proporción del evento) del grupo experimental (numerador) y el riesgo (proporción del evento) del grupo control (denominador), muestra cuántas veces más riesgo de padecer un desenlace se tiene perteneciendo a un grupo comparado con el otro. Tomando de nuevo del ensayo de Fowler et ál. (2016), se estimaron tasas de infección de VIH perinatal que dan como resultado un RR de 0,33 (0,006/0,018), lo que significa que el TDF-Art redujo en una tercera parte el riesgo de infección entre quienes lo recibieron comparado con quienes recibieron ZDV.

Las dos anteriores medidas son algunas de las formas comúnmente vistas en las publicaciones de ensayos clínicos; sin embargo, existen otras medidas más informativas, incluso para transmitir los resultados de estos estudios a los mismos pacientes en quienes serán implementados las intervenciones evaluadas. Este es el caso de la reducción del riesgo relativo (RRR).

RRR = DR/riesgo del grupo control = (c/c+d - a/a+b)/(c/c+d)

Operando con las fracciones se encuentra que:

RRR = 1- RR; RRR con valores posibles en el intervalo cerrado [0,1]

La RRR hace referencia a la proporción de riesgo de base que es removida por el tratamiento en prueba. Por ejemplo, con el ensayo de transmisión perinatal de VIH, la RRR es 0,66, es decir, se estima que el TDF reduce en un 66 % el riesgo de transmisión de perinatal de VIH comparado con ZDV o, lo que es mejor, reduce en dos terceras partes el riesgo de infección perinatal que se da con ZDV.

Otra medida similar al RR para evaluar la relación entre un tratamiento y un desenlace es la razón de posibilidades (en inglés *odds ratio*) (OR). Se estiman las *odds* (posibilidades) en cada grupo (experimental y control), se calcula el cociente de las *odds* y esto da la OR. El resultado es equivalente al RR, su expresión de cálculo es:

$OR = (a/b)/(c/d) = ad/cb$; OR con valores posibles en el intervalo abierto (0,)

También se puede expresar el impacto del tratamiento por el número de pacientes que se necesitaría tratar para prevenir un evento o desenlace. Esto es llamado el número necesario para tratar (NNT).

$NNT = 1/RD$; NNT con valores posibles en el intervalo abierto (0,)

Si se dice que un $RD = 1,2 \%$, entonces, si se tratan 100 pacientes con TDF-Art, se evita un evento de transmisión de infección perinatal. Ahora bien, ¿cuántos pacientes se requiere tratar para evitar una infección perinatal de VIH?, la respuesta sería $NNT = 1/0,012 = 83$ pacientes.

Daño o etiología



Figura 5.

Fuente: Shutterstock/146441699

En este nuevo escenario práctico, las preguntas clínicas están dirigidas a conocer el daño producido por una exposición o el papel etiológico que juega una exposición frente a una condición que, generalmente, es una enfermedad o evento negativo.

Ahora bien, los estudios que permiten dar respuesta a este tipo de pregunta bien podrían estar en el marco de un diseño de investigación tipo ensayo clínico. Aunque los investigadores lo realizan para determinar si los agentes terapéuticos son beneficiosos, también deben buscar efectos dañinos y, a veces, pueden hacerse descubrimientos sobre los efectos negativos de la intervención en sus resultados primarios. Este diseño tiene el poder de proveer la evidencia más contundente frente a una exposición, ya que incluye metodológicamente la técnica de aleatorización, la cual permitirá tener grupos balanceados, en términos de factores pronósticos conocidos y desconocidos, y controlará el efecto de estos sobre cualquier desenlace evaluado.

De otra manera, existen varias razones por las que un ensayo clínico no puede implementarse para evaluar el daño producido por una exposición. Se mencionan las tres más importantes que tienen las justificaciones más fuertes. En primer lugar, se considera que no es ético aleatorizar pacientes a exposiciones que de manera anticipada (bajo hipótesis de investigación) se espera que traigan efectos nocivos. En segundo lugar, a menudo hay preocupación por los efectos adversos raros y graves que pueden llegar a ser evidentes solo después de que decenas de miles de pacientes han consumido un medicamento durante años. Incluso ensayos clínicos muy grandes no logran detectar una asociación entre una exposición (tratamiento) y algún evento adverso. En tercer lugar, los ensayos clínicos a menudo no reportan adecuadamente la información sobre el daño.

Según lo indicado hasta el momento, el clínico o practicante interesado en contestar una pregunta clínica orientada sobre el efecto dañino de una exposición debe dirigir la búsqueda de evidencia hacia los estudios observacionales, como son los estudios de cohorte y los de casos y controles.

En un estudio de cohorte el investigador identifica a grupos expuestos y no expuestos de pacientes, cada uno de una cohorte, y luego los sigue en el tiempo, controlando la ocurrencia del resultado previsto. El diseño de la cohorte es similar a un ensayo clínico, pero sin aleatorización. La determinación de si un paciente recibió la exposición de interés resulta de la preferencia del paciente o de la casualidad.

Los estudios de casos y controles también evalúan las asociaciones entre las exposiciones y la enfermedad. Los resultados raros o los que tardan mucho tiempo en desarrollarse pueden amenazar la viabilidad de los estudios de cohorte. El estudio de casos y controles proporciona un diseño alternativo que se basa en la identificación inicial de los casos, es decir, en pacientes que ya han desarrollado la enfermedad y en la selección de controles, personas que no tienen la enfermedad de interés. Utilizando diseños de casos y controles, los investigadores evalúan la frecuencia relativa de la exposición previa a ser caso o control.

Teniendo claro lo anterior, al momento de encontrar un estudio de este tipo para tomar decisiones sobre el paciente, se debe tener, al igual que en cualquier investigación clínico-epidemiológica, la posibilidad de la presencia del error aleatorio y del sesgo. Por tanto, profundizaremos en aspectos metodológicos que tienen relación con la validez de los resultados (sesgos) y con la forma de reportar los resultados de estos estudios y la valoración de la precisión de las estimaciones (error aleatorio).



Validez

Argumentos necesarios para calificar un resultado como certero o verdadero de una realidad.

Validez de los resultados en
estudios observacionales
(cohorte y casos y controles)



Estudios de cohorte

Comparabilidad entre expuestos - no expuestos y su ajuste

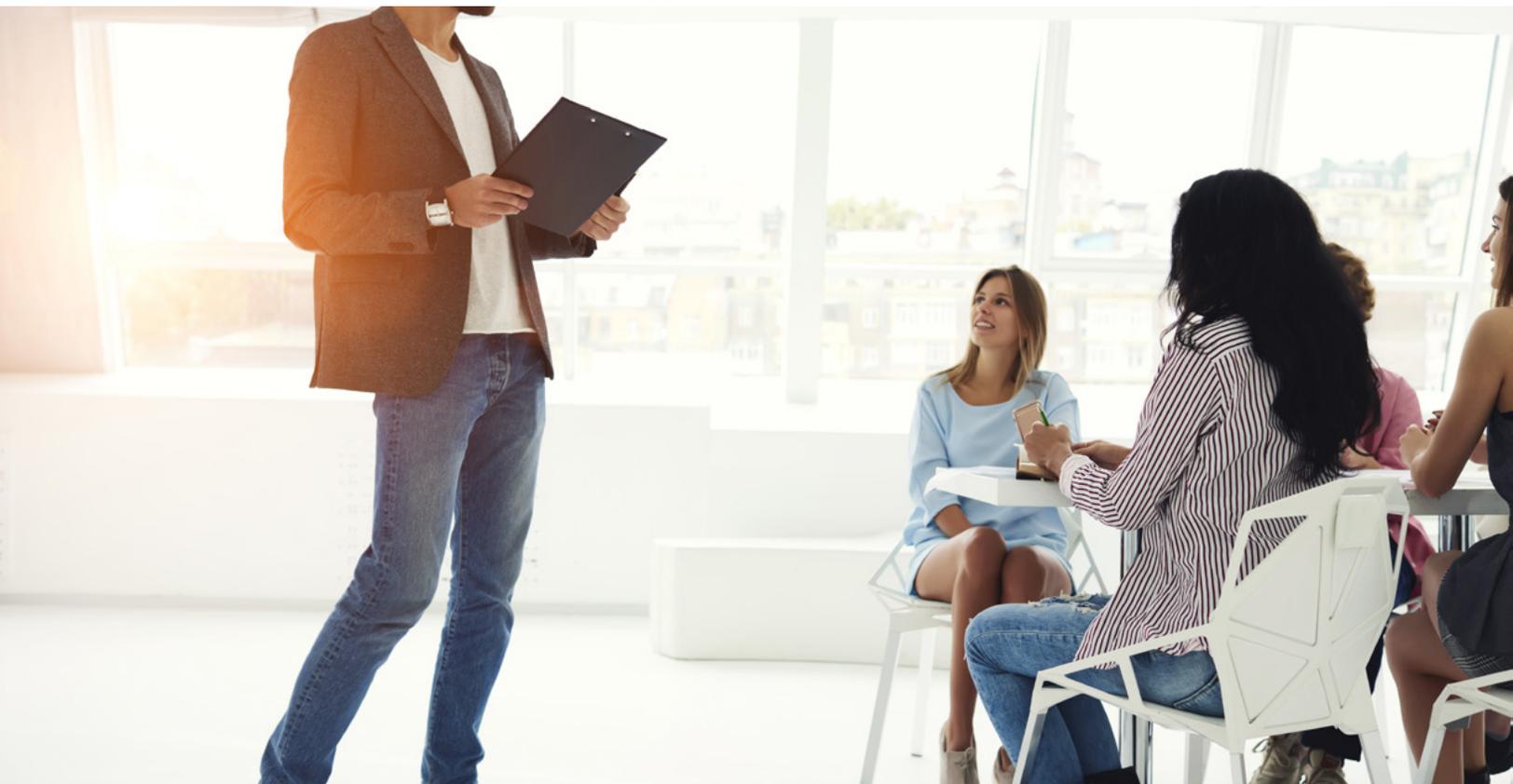


Figura 6.
Fuente: Shutterstock/566945398

El investigador puede desarrollar un estudio de cohorte retrospectiva o prospectivamente. En el primer caso, la exposición y el desenlace o daño evaluado ocurrieron antes del inicio del estudio, por tanto, el investigador debe retroceder para recolectar datos sobre el desenlace y la exposición, para lo cual debe tenerse en cuenta la calidad de las fuentes de donde serán tomados los datos. En un estudio prospectivo, el investigador inicia conformando los grupos de exposición entre individuos libres del daño. Para ambos casos, se debe evaluar qué diferencias existen en factores pronósticos, probablemente también relacionados con la aparición de la enfermedad o el daño. Este desbalance es ocasionado porque los grupos son conformados por el estado natural de cada participante, es decir, el investigador no interviene en asignar dicha exposición (como sí ocurre en los ensayos clínicos), de tal modo que los grupos son conformados de forma no aleatoria y bajo esta situación pueden ser diferentes.

Los estudios de cohorte de exposiciones potencialmente dañinas darán resultados sesgados si el grupo expuesto al agente nocivo y el grupo no expuesto comienzan con diferentes características basales que les dan un pronóstico diferente. Este desbalance no puede ser evitado desde el diseño del estudio (como sí ocurre en ensayos clínicos por la aleatorización), dado que es propio del diseño de observacionales, pero, desde una forma analítica (métodos estadísticos), se puede controlar el efecto de dicho desbalance en los factores pronósticos identificados.

Debe tenerse en cuenta que, para realizar dicho ajuste por diferencias entre los grupos, las variables pronósticas deben ser medidas de manera exacta, es decir, debe ponerse especial atención en los métodos de medición que el investigador utiliza para valorar su presencia. También aplica esta condición para la medición de la exposición y el mismo desenlace o daño evaluado, ya que sesgos de información relacionados a estas variables generan estimados sesgados de la relación exposición daño/enfermedad.

Medición del desenlace en los grupos

No se desea que un grupo sea estudiado más exhaustivamente que el otro porque esto podría conducir a reportar una mayor incidencia de daño/enfermedad en el grupo más estudiado. Condiciones especiales dentro de la cohorte pueden hacer que los investigadores busquen más exhaustivamente el daño/enfermedad entre ellos y hagan ver que hay un mayor riesgo de padecer el daño/enfermedad entre un grupo comparado con el otro, mostrando al final una asociación exposición-enfermedad falsa o aumentada en magnitud.



Desenlace

Medida que se toma como la variable efecto o dependiente en la comparación de dos grupos independientes o dependientes.

Seguimiento de los individuos

Al igual que sucede en los ensayos clínicos, la pérdida de seguimiento puede introducir sesgo porque los pacientes que se pierden pueden tener resultados diferentes frente a los pacientes que están disponibles para la evaluación. Se espera que el seguimiento sea lo suficientemente largo para que los eventos se expresen y lo adecuadamente completo sobre todos los individuos participantes.

Estudios de casos y controles

Al igual que los estudios de cohorte, los estudios de casos y controles son susceptibles a variables de confusión no medidas, particularmente cuando la exposición varía con el tiempo.

Medición de la exposición en los grupos

En los estudios de casos y controles la determinación de la exposición es un problema clave. Si los casos tienen un mejor recuerdo para reportar la exposición que los pacientes control, el resultado será una asociación espuria. También existen problemas en relación con los métodos con los que se valora y clasifica el grado de exposición entre casos y controles. Esto puede generar una mala clasificación y que las estimaciones de la asociación sean sesgadas.

En el caso de estudios observacionales también incluyen otros diseños como el corte transversal y la serie de casos. En el primero, a partir de una muestra aleatoria de una población, se obtienen mediciones en un momento dado del tiempo sobre la exposición y enfermedad, por tanto, el problema principal de este diseño radica en que no hay una clara dirección de la temporalidad entre exposición y daño, siendo difícil establecer el papel causal de la exposición, ya que no se puede determinar qué ocurrió primero: la exposición o la enfermedad.

En el diseño de serie de casos, el problema radica en no proporcionar un grupo de comparación, por lo que es imposible determinar si el resultado observado probablemente habría ocurrido en ausencia de la exposición.

Resultados en estudios de cohorte y de casos y controles

Los resultados de los estudios de cohorte y de casos y controles también se pueden presentar en una tabla de 2x2. Las medidas para expresar la relación exposición-enfermedad utilizadas para el análisis son RR y OR.

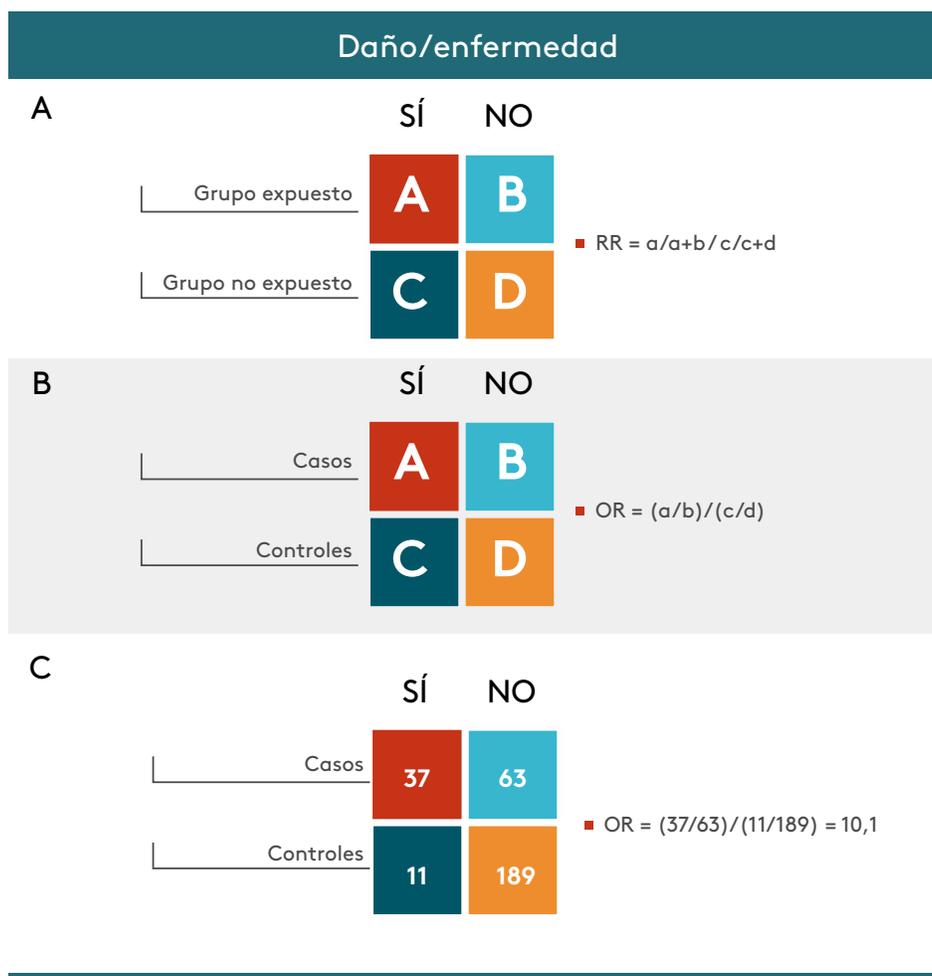


Figura 7. Tabla 2x2 para resultados de un estudio de cohorte (parte A) para estudio de casos y controles (parte B) y OR en resultados de un estudio de caso y controles entre asociación de infección oral por VPH y cáncer orofaríngeo (parte C)
Fuente: D'Souza et ál. (2007)

Diagnóstico

En este nuevo escenario o situación práctica, la pregunta PICO planteada hace referencia al ejercicio práctico del diagnóstico, es decir, la situación en que se ve involucrado el proceso de llegar a un diagnóstico en un paciente. El proceso diagnóstico proviene de un fenómeno probabilístico dirigido así: el practicante, con base en una entrevista, que incluye características del paciente, signos y síntomas, además de la epidemiología local de las enfermedades, establece un listado de posibles diagnósticos con valores de probabilidad para cada uno (probabilidad pretest). Posterior a esto, establece la realización de pruebas diagnósticas que intentan dar información adicional, lo que está relacionado con valorar de nuevo la probabilidad del diagnóstico ante esta nueva información (resultado de un test), la cual es llamada probabilidad posttest y, finalmente, tomar una decisión.

En la práctica general, muchos test diagnósticos son objeto de evaluación en su desempeño diagnóstico, es decir, la capacidad que tienen para clasificar correctamente a los sujetos. Esto hace necesario que la toma de decisiones frente a las pruebas enviadas para un diagnóstico esté fundamentada en el análisis crítico de la literatura que soporta la utilidad de la prueba, allí es cuando surge desde los estudios la evidencia con la que se cuenta. Esta evidencia debe ser evaluada frente a la ocurrencia del sesgo y del error aleatorio, lo cual permita aplicarla de manera correcta y con evidencia científica.

Clásicamente, estos estudios de evaluación de pruebas diagnósticas se desarrollan bajo un diseño de investigación de corte transversal o de casos y controles, y algunos bajo un diseño de cohorte; sin embargo, para simplificar el desarrollo teórico se tendrán en cuenta los que son implementados como estudios descriptivos de corte transversal. Las medidas que resumen los resultados de estos estudios son iguales, independientemente del diseño utilizado por el investigador.

El estudio descriptivo de corte transversal utilizado en la evaluación de pruebas diagnósticas consiste en la selección de una muestra representativa de una población con el cumplimiento de criterios mínimos para sospecha de un diagnóstico (espectro de signos y síntomas). En esta muestra de sujetos se aplica el test o prueba nueva o en evaluación y se corrobora en ellos mismos, mediante una prueba *gold* estándar o de referencia, la presencia de manera real de la enfermedad o condición de diagnóstico en cuestión. Los datos que el investigador recolecta en este tipo de estudios dependen de los resultados arrojados por la prueba en evaluación. De esta manera, si el resultado dado por el test es un valor cuantitativo, por ejemplo, diluciones de una serología, nivel de hemoglobina y concentración de un biomarcador, los datos serán de tipo numérico. Si, por el contrario, los resultados son de tipo positivo/negativo, reactivo/no reactivo, se dará un dato de tipo cualitativo dicotómico. Para facilidad del lector y según lo que frecuentemente se ve en la literatura, se tomará en cuenta para el análisis de estos tipos de estudios aquellos que entregan datos de tipo cualitativo dicotómicos.

Validez de los resultados en estudios sobre pruebas diagnósticas

Espectro de la enfermedad

Los participantes en un estudio son seleccionados o muestreados de una población objetivo de personas que presentan signos y síntomas probablemente asociados a una enfermedad que se está investigando. Idealmente, esta muestra debe reflejar la población objetivo en todas las formas importantes, de modo que la frecuencia y las posibilidades de diagnóstico sean un reflejo de lo que sucede en la población. Cuanto más representativa sea la muestra, más exacta serán las probabilidades de enfermedad resultantes. Para verificar si esta representación de la población blanco está en la muestra, existen cuatro formas. En primer lugar, encontrar la definición que los investigadores dan al problema clínico que se presenta, porque esto determina la población objetivo de la cual los pacientes del estudio van a ser extraídos.

En segundo lugar, examinar el contexto del que los pacientes son reclutados. Los pacientes con el mismo problema clínico podrían presentarse en cualquiera de los diferentes entornos clínicos, un primer nivel de atención, urgencias o clínicas de referencia. La elección de dónde buscar atención puede involucrar varios factores, incluyendo la enfermedad, la disponibilidad de los diversos entornos, los hábitos de referencia de su médico o preferencias del paciente. Los pacientes en niveles más complejos de atención involucran mayor severidad, permitiendo esto que la posibilidad de presentación de la enfermedad sea más alta que en otros niveles. En tercer lugar, los métodos de los investigadores para identificar a los pacientes en cada sitio y cómo pudieron evitar algunos de los pacientes. Idealmente, deben reclutarse una muestra de todos los pacientes que buscan atención en los sitios de estudio para el problema clínico durante un periodo especificado. Si los pacientes no se incluyen consecutivamente, la inclusión de pacientes con diferentes trastornos subyacentes reduce la representatividad de la muestra y, de esta manera, reduce la confianza en la validez de los resultados. En cuarto lugar, examinar el espectro de gravedad y características clínicas de los pacientes en la muestra del estudio. ¿Se encuentran todas las variaciones importantes de este problema clínico en la muestra?

Evaluación diagnóstica definitiva

Los artículos sobre pruebas diagnósticas proveerán evidencia válida solo si los investigadores llegan a los diagnósticos correctos para los pacientes del estudio. Esto se logra verificando diferentes aspectos dentro del estudio.

La evaluación diagnóstica debe ser capaz de detectar todas las posibles causas del problema clínico en consideración. Debe haber exhaustividad en la búsqueda del diagnóstico, si no, las probabilidades de enfermedad estimadas serán sesgadas.

Deben estar descritos suficientemente todos los procedimientos diagnósticos llevados a cabo en el paciente, así como las reglas de decisión frente a llegar al diagnóstico definitivo, esto tiene que ver con la estandarización de los procedimientos realizados, de manera tal que los haga reproducibles en un contexto fuera de la investigación. Al no haber dicha estandarización, se puede provocar inexactitud en las estimaciones de las probabilidades de diagnóstico.

Por último, el estándar de referencia utilizado para dar clasificación final a cada sujeto en el estudio debe tener características que lo hagan reproducible. Tanto la prueba que se está evaluando como la estándar de referencia deberían llevarse a cabo en todos los pacientes del estudio. Por ejemplo, si la prueba bajo investigación resulta negativa, esto puede haber llevado a que el investigador decida no realizar la prueba de oro en pacientes con este resultado, lo cual generará una situación denominada en el diseño de estos estudios, como verificación parcial de la enfermedad. Esto debe estar informado dentro del artículo, además de cómo fueron tratados los análisis de datos en estos casos. Algo importante en este punto es que tanto la evaluación por estándar de referencia como por prueba deben haberse evaluado de manera enmascarada entre sí, garantizando la independencia de los resultados por la prueba en evaluación y por la prueba de oro.

Resultados en estudio de evaluación de pruebas diagnósticas

Partiendo de que los datos obtenidos en un estudio de evaluación de pruebas diagnósticas y los valores del test en evaluación tienen la forma de positivo/negativo, además de que la condición establecida por la prueba de referencia también constituye un resultado dicotómico (enfermo/sano), estos datos pueden organizarse en una tabla de 2x2.

La tabla de ejemplo muestra los resultados de un estudio de corte transversal realizado por Miller et ál. (2008), en el que se evalúa el desempeño diagnóstico de una tomografía de multidetección angiográfica utilizada para la identificación de la enfermedad coronaria (obstrucción).



Pruebas diagnósticas

Resultados de test (cuantitativos-cualitativos) objetivos o subjetivos, que clasifican a un individuo en términos de ausencia o presencia de un resultado.

Gold estándar

	Enfermo	Sano	Total
Positivo	A	B	$r_1 = a+b$
Negativo	C	D	$r_2 = c+d$
Total	$c_1 = a+c$	$c_2 = b+d$	$n = c_1+c_2$

- Prevalencia de la enfermedad o probabilidad de enfermar: $p = c_1/n$
- Sensibilidad (Se) = a/c_1 Falsos negativos (FN) = c/c_1
- Especificidad (Sp) = d/c_2 Falsos positivos (FP) = b/c_2
- Valor predictivo positivo (VPP) = a/r_1
- Valor predictivo negativo (VPN) = d/r_2
- Razón de probabilidad positiva (RP+) = $Se/(1-Sp)$
- Razón de probabilidad negativa (RP-) = $(1- Se)/Sp$

Figura 8. Tabla de 2x2 para estudios de evaluaciones de pruebas diagnósticas y sus medidas de desempeño
Fuente: propia

Enfermedad coronaria aguda

		SÍ	NO	Total
Multidetector CT angiográfico	Positivo	139	13	152
	Negativo	24	115	139
Total		163	128	291

- Prevalencia de la enfermedad o probabilidad de enfermar $p = 0.56$
- Se = 0.85 FN = 0.15
- Sp = 0.9 FP = 0.10
- VPP = 0.91
- VPN = 0.82
- RP+ = 8.5
- RP- = 0.16

Figura 9. Resultados de un estudio evaluando el desempeño diagnóstico del multidetector CT angiográfico comparado con la angiografía convencional para el diagnóstico de enfermedad coronaria aguda
Fuente: Miller et ál. (2008)

Las medidas utilizadas están dadas en términos de las probabilidades de clasificación correcta de los sujetos, calculadas a partir de los resultados arrojados por el test en evaluación. La probabilidad de ser clasificado como positivo, dado que se está enfermo, es denominada sensibilidad y es estimada a través de la proporción de positivos entre los enfermos. Da cuenta de la capacidad del test para detectar a los sujetos en estudio como enfermos. La probabilidad de ser clasificado como negativo, dado que se está sano, es denominada especificidad y es estimada a través de la proporción de negativos entre los sanos. Lógicamente, asociadas a estas probabilidades de clasificación correcta también están las probabilidades de clasificación incorrecta: la proporción de falsos positivos y falsos negativos. Los valores posibles para estas medidas están en el intervalo cerrado $[0,1]$. Para lo cual, sensibilidad y especificidad de una prueba con valores de 100 % sería lo ideal.

Las probabilidades de clasificación correctas e incorrectas son medidas intrínsecas del desempeño de una prueba que, finalmente, solo son de utilidad e informativas para la investigación. En la praxis médica del diagnóstico hay otras medidas que suelen ser más informativas frente a la toma de decisiones para confirmar un diagnóstico sobre un individuo que consulta con una sintomatología y le es enviada una prueba diagnóstica, la cual a vuelta de consulta trae un resultado positivo o negativo. A partir de esta información, al médico o practicante le es de más utilidad conocer la probabilidad de que el paciente tenga o no la enfermedad sospechada, con base en que trae un resultado de un test. Esta es la utilidad de calcular los valores predictivos, los cuales son: predictivo positivo, que refleja la probabilidad de tener la enfermedad sospechada cuando se trae un resultado positivo, y el predictivo negativo, que se refiere a la probabilidad de estar sano cuando el resultado de la prueba fue negativo. Estos valores predictivos se estiman a través de la proporción de enfermos entre los sujetos con resultados positivos y la proporción de sanos entre sujetos que tienen un resultado negativo.

Un inconveniente ampliamente descrito por la teoría de probabilidades y que puede demostrarse empíricamente es la fuerte asociación o dependencia entre los valores predictivos y la prevalencia o probabilidad de enfermar, lo cual dificulta el uso de estos en la práctica, principalmente. Lo anterior se atribuye a que la prevalencia de la enfermedad varía dependiendo de las características de la población y el sitio donde se realiza el estudio. A mayor prevalencia, el valor predictivo positivo de la prueba aumenta. Si la prevalencia es baja, el valor predictivo negativo aumenta. Dada esta relación, y que la prevalencia varía según el sitio y las características de la población base de donde se extrae la muestra, los valores predictivos pueden cambiar fácilmente de acuerdo con el contexto de desarrollo del estudio, situación que no sucede con las probabilidades de clasificación correctas e incorrectas.

Para solventar dicho traspié, se plantean dos nuevas medidas que eliminan esta dependencia entre los valores predictivos y la prevalencia de la enfermedad. Las razones de probabilidad (*likelihood ratio*) permiten evaluar el desempeño diagnóstico de la prueba bajo la aplicación práctica de su resultado en escenarios clínicos, independientemente de la prevalencia. La razón de probabilidad positiva es el cociente entre la probabilidad de tener un resultado positivo entre enfermos y la probabilidad de tener un resultado positivo entre sanos, por tanto, su interpretación es: cuántas veces es más probable un

resultado positivo entre enfermos comparados con los sanos. En referencia a la razón de probabilidades negativa, que corresponde al cociente entre la probabilidad de tener un resultado negativo entre enfermos y la probabilidad de tener un resultado negativo entre sanos, de la misma forma su interpretación sería: cuántas veces es más probable ver un resultado negativo entre enfermos comparados con los sanos.

Precisión de los resultados en los estudios clínico-epidemiológicos

Este apartado del eje hace referencia a la cuantificación del error aleatorio presente en los resultados de un estudio. Debido a que las formas de expresar dicho error son transversales a las diferentes formas de análisis de los resultados, se explicarán de manera general, siendo su diferencia particular solo en la aplicación de fórmulas explícitas de acuerdo con la medida a la cual se le pretende calcular el nivel de error estadístico asociado a su estimación. Sin embargo, no es objeto de este documento presentar una exhaustiva formulación que se aleja del objetivo principal del análisis basado en la evidencia.

La idea de evaluar el grado de error aleatorio en los resultados de un estudio es determinar la precisión de los resultados obtenidos. Se parte de que el resultado puntual obtenido con los datos de una muestra es solo una aproximación al valor verdadero en la población de interés. Como se mencionó, esta medida puntual está asociada a un error por variabilidad muestral (si se repite el estudio con una muestra diferente de la misma población, el resultado puntual será diferente). Técnicas estadísticas de inferencia permiten calcular el grado de error asociado a esta variabilidad muestral, permitiendo obtener un posible rango de valores, en el que podría estimarse se encuentra el valor verdadero en la población, teniendo como insumos únicamente los datos de la muestra. Las dos técnicas principalmente utilizadas son los intervalos de confianza y los contrastes de hipótesis, que son aplicables a todos los tipos de análisis revisados hasta el momento (RR, OR, DR, Se, Sp, VPP, VPN, etc.).

Intervalos de confianza

Los intervalos de confianza son una técnica inferencial que permite estimar por medio de un rango posible de valores en los cuales, con cierto nivel de seguridad o confianza, se espera que esté el valor verdadero del parámetro poblacional. Para realizar esto, los intervalos de confianza integran dentro de su formulación una cantidad denominada "error estándar", la cual corresponde a la variabilidad que se da en una estimación si se calculara en muestras diferentes.



Intervalos de confianza

Técnica de inferencia estadística para generalizar resultados a una población partiendo de una muestra aleatoria de la misma.

Este error estándar es la clave para construir el intervalo, pues supone la cantidad de incertidumbre que se genera haciendo los cálculos en diferentes muestras. Asociándolo al valor puntual obtenido en la muestra del estudio, se puede hablar de que el valor aproximado se encuentra en la medida si se hubiera evaluado toda la población.

La forma genérica para cualquier intervalo de confianza que se le construya a una medida tiene la siguiente forma:

Estimador muestral coeficiente de confiabilidad x error estándar

Los estimadores muestrales corresponden a las medidas estudiadas anteriormente, estos son: RR, OR, RD, Se, Sp, etc., los cuales pueden ser calculados con las expresiones matemáticas mostradas en secciones anteriores. En relación al coeficiente de confiabilidad, este corresponde a un valor constante preestablecido para un nivel de confianza. Si el investigador diseñó el estudio con un nivel de confianza de 95 %, el coeficiente a utilizarse será 1,96, pero si se determinó un nivel de confianza de 90 %, dicho valor cambiará a 1,64.

En cuanto al error estándar, es diferente para cada estimador muestral, es decir, hay una expresión matemática según la medida a ser utilizada; sin embargo, no es objeto de este curso profundizar en dichos aspectos estadísticos, debido a que regularmente estos intervalos de confianza son reportados como requisitos para valorar la precisión de los resultados en la publicación de los estudios clínico-epidemiológicos.

Amplitud del intervalo de confianza

Un aspecto importante de la interpretación de los intervalos de confianza es la amplitud. El intervalo expresa el rango de posibles valores que una medida toma en la población, entre más estrecho sea, mayor precisión expresa sobre el valor verdadero, permitiendo al lector del estudio conocer el tamaño del efecto estudiado en toda población de interés a la cual se pretende generalizar el resultado.

Retomando el ejemplo del ensayo clínico comparando los antirretrovirales TDF vs. ZDV (Fowler et ál., 2016), se estimó con los datos del estudio de manera puntual que la diferencia de riesgos para TDF vs. ZDV era de 1,2 %, es decir, los pacientes que recibieron TDF disminuyen su riesgo absoluto de infección perinatal por VIH en 1,2 puntos porcentuales, comparados con los que recibieron ZDV. A esta diferencia los autores reportaron un intervalo de confianza al 95 % de 0,4 % a 2,1 %. El intervalo permite fijar que el valor del efecto en la reducción absoluta de riesgo de infección por VIH perinatal en la población general de interés oscila entre 0,4 % y 2,1 % con un 95 % de seguridad. Como se puede observar, este intervalo no incluye el valor de cero, lo que dentro del 95 % de posibilidades dice que esta reducción siempre será observada si se repitiese este estudio múltiples veces, indicando que un efecto significativo, desde el punto de vista estadístico, se da en la población.



Lectura recomendada

Intervalos de confianza

Roberto Candia y Gianella Caiozzi

Determinantes importantes de la amplitud de los intervalos de confianza son el tamaño de la muestra y el número de eventos/desenlaces que ocurren en el estudio, cuando el evento es dicotómico (muerte/vivo, infectado/no infectado, curado/no curado). Si bien los estimados con tamaños de muestra grandes proveen intervalos de confianza estrechos, también es cierto y demostrable empíricamente que tamaños de muestra pequeños donde la tasa del desenlace (número de eventos de interés por grupo) es alta, pueden proveer intervalos de confianza igual o más precisos que con los tamaños de muestra grandes.

Contrastes de hipótesis

Para cada tratamiento existe un verdadero efecto subyacente que cualquier experimento individual solo puede estimar. Los investigadores usan métodos estadísticos para avanzar en la comprensión de este verdadero efecto. Un acercamiento desde lo estadístico es comenzar con lo que se llama hipótesis nula y se trata de rechazar dicha hipótesis.

La hipótesis nula es un modelo propuesto por el investigador que se da bajo la premisa de que los datos obtenidos en una muestra siguen ese modelo. Generalmente, la afirmación subyacente en el modelo es que no hay diferencia entre los tratamientos que se comparan.

En este contexto, el procedimiento estadístico (contraste de hipótesis) implementado permite, utilizando los datos obtenidos en la muestra puntual, determinar si son consistentes con la hipótesis nula (modelo propuesto). Si los resultados muestran que los datos no son compatibles con el modelo propuesto, habría entonces evidencias en contra de la hipótesis nula y quedaría solo la opción de rechazarla. Por el contrario, si los datos muestran una compatibilidad con la hipótesis nula planteada, sería el caso en que los resultados muestran evidencia a favor de la hipótesis nula y, por ende, no podría ser rechazada, no por lo menos con los datos obtenidos en la muestra.

El procedimiento de contraste de hipótesis lleva a únicamente a la toma de dos posibles decisiones: se rechaza o no la hipótesis nula. Como consecuencia de este rechazo o no, dos errores se pueden cometer: el primer error es el rechazo de una hipótesis nula, cuando esta en realidad era verdadera. En palabras más comunes, sería concluir que un tratamiento es efectivo o que la asociación vista entre una exposición-enfermedad existe cuando en realidad no era así.

Este error es denominado error tipo I o alfa. El segundo error está asociado a la decisión de no rechazar la hipótesis nula cuando esta en realidad era falsa o lo que es lo mismo, afirmar que un tratamiento no era efectivo cuando en realidad sí lo era. Este error es denominado error tipo II o beta.

	Rechazo	No rechazo
Hipótesis nula verdadera	Error tipo I	Decisión correcta
Hipótesis nula falsa	Decisión correcta	Error tipo II

Tabla 2. Errores aleatorios tipo I y II asociados con un contraste de hipótesis
Fuente: propia

Dentro del contraste, estos errores no ocurren al mismo tiempo. En la medida en que el error tipo I disminuye, como consecuencia, aumenta la certeza (confianza) sobre el rechazo y el error tipo II se incrementa. De forma contraria, si reduzco el error tipo II, aumenta la probabilidad de un contraste con poder para detectar diferencias observadas.



Instrucción

Lo invito a que realice el control de lectura sobre intervalos de confianza.

Valores de p

Para el control del error aleatorio alfa dentro de un contraste de hipótesis se le ha establecido una frontera máxima permitida, lo cual establece que si los datos sobrepasan dicha frontera el error aleatorio asociado a una decisión sea mínimo y, por tanto, aumente la plausibilidad de una conclusión sobre la hipótesis nula. Este nivel alfa máximo permitido es arbitrario y colocado antes de llevar a cabo el estudio o experimento. El valor tomado es objeto de discusión; sin embargo, en estudios clínico-epidemiológicos convencionalmente se ha establecido en 5 % o 0,05. De tal modo que, cuando es llevado a cabo el contraste, se produce un valor de probabilidad denominado "valor de p", el cual es comparado contra el nivel alfa preestablecido y permite tomar una decisión frente a la hipótesis nula planteada. Si este valor de p del contraste sobrepasa esta frontera del nivel de error alfa, el resultado obtenido con dicha muestra es incompatible con el modelo propuesto de la hipótesis nula, siendo su observación tan extrema o más extrema que la vista en menos del 5 % de las veces si se repitiera el estudio. Esta situación da evidencia empírica para rechazar la hipótesis nula y establecer que se está ante un resultado del estudio estadísticamente significativo.

Finalmente, el contraste de hipótesis presenta serias limitaciones en el análisis de resultados desde el punto de vista práctico. La primera hace referencia al valor umbral de error alfa que se antepone en un contraste. Este valor es arbitrario, no existe ninguna fórmula o justificación práctica que permita decidir, por ejemplo, entre un 5 % o 10 % o, porque no, entre un 1 % y 5 %. Dicha discusión permanece abierta entre los expertos.

En segundo lugar, el contraste es un método que reduce la conclusión de si un efecto visto es real o no a únicamente dos radicales posibilidades: existe o no efectividad, existe o no asociación, es o no un factor pronóstico. No deja opciones en un continuo para que se den puntos intermedios también válidos, como, por ejemplo, "es casi eficaz", "sugiere una asociación" o "muy poco probable que sea efectivo".

En tercer lugar, hace más de 15 años el Comité Internacional de Editores de Revistas Biomédicas lanzó su recomendación para la publicación de resultados de estudios, dando como sugerencia el uso más frecuente de intervalos de confianza que de valores p , debido a sus claras limitaciones.

Son múltiples las perspectivas de error que pueden estar involucradas en los resultados arrojados por una investigación científica en el ámbito de las ciencias de la salud; sin embargo, uno de los elementos necesarios para garantizar una evaluación adecuada de la evidencia es el reconocimiento del error sistemático, atribuido a lo que se llama en epidemiología sesgo y que desvía el resultado de la verdad, para lo cual se entregaron elementos que deben ser evaluados en cada tipo de estudio, de acuerdo con el escenario clínico (terapia, diagnóstico, pronóstico, etc.). También está

el error aleatorio que, si bien es un tema estadístico, es necesario conocer las fuentes de este posible error y las formas de controlarlo y expresarlo en los resultados para que quien lea esté informado y tome adecuadas decisiones.



Lectura recomendada

El valor de "p" y la "significación estadística". Aspectos generales y su valor en la práctica clínica

Carlos Manterola, Viviana Pineda y Grupo Mincir



Instrucción

Para finalizar, lo invito a revisar la animación que hemos preparado sobre escenarios clínicos.

D'Souza, G. et ál. (2007). Case-control study of human papillomavirus and oropharyngeal cancer. *The New England Journal of Medicine*, 356(19), 1944-1956.

Fowler, M. et ál. (2016). Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *The New England Journal of Medicine*, 375(18), 1726-1737.

Miller, J. M. et ál. (2008). Diagnostic performance of coronary angiography by 64-Row CT. *The New England Journal of Medicine*, 359(22), 2324-2336.

ANÁLISIS BASADO EN LA EVIDENCIA

Jorge Mario Estrada

EJE 3

Pongamos en práctica

INTRODUCCIÓN

El análisis crítico es un ejercicio conceptual que permite evidenciar los aportes del artículo ubicado que proveen la evidencia para la toma de decisiones fundamentada. A partir de esto, se ha creado una guía con una relación de preguntas que permite al lector dirigir mejor su lectura y colocar la atención en los puntos trascendentales para tener certeza sobre la validez y la confiabilidad de los resultados. En este punto, se presentan los elementos necesarios para realizar un análisis crítico.

Análisis crítico



Guías de usuario

Las preguntas que se detallan a continuación hacen parte de las guías de usuario publicadas en los años 90 por el Evidence-Based Medicine Working Group de la Universidad McMaster de Canadá en la *Journal of the American Medical Association* (Richardson et al., 1999). En el año 2014, Manterola y Otzen (2014) publicaron en una versión en español.

Las guías de usuario contienen elementos que permiten valorar la calidad metodológica con la que fueron obtenidos los resultados y si estos pueden o no ser aplicados en el cuidado de los pacientes.

Se aporta una guía por tipo de estudio o pregunta que se resuelve y está compuesta por dos columnas: a la izquierda están las preguntas guía que el lector debe contestar con base en los resultados leídos del artículo, y a la derecha están “pistas” o comentarios que ayudan a responder de forma apropiada la guía.

Las respuestas deben estar argumentadas a la luz de los resultados provistos en el artículo.

Requiero decidir sobre una intervención

Este escenario está dedicado al análisis de la evidencia de artículos que publican resultados de una nueva terapia o alguna modificada. Se requiere analizar críticamente la validez y la confiabilidad del resultado, con el ánimo de implementar en los pacientes los resultados y conclusiones.

Frente al artículo que se encuentra ubicado bajo una búsqueda en las bases de datos y que responde probablemente a la pregunta PICO inicial, se debe dar respuesta a cada una de las preguntas guía, las cuales llevarán a valorar críticamente los resultados y tomar una decisión sobre su implementación.

Las preguntas están dirigidas a valorar las características metodológicas que hacen que los resultados de cada estudio sean válidos y precisos. Este tema fue socializado en el eje 2.



Instrucción

En este punto, lo invito a que revise el videorre-
sumen sobre terapia.



Aleatorización

Técnica de asignación aleatoria de indivi-
duos a dos o más grupos independientes
durante la realización de un ensayo clínico.

¿Son válidos los resultados de la investigación?

Preguntas de eliminación

Guía	Comentarios
<p>1. ¿Se orienta el artículo a una pregunta claramente definida?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Pista</p> <p>Una pregunta debe definirse en tér- minos de:</p> <ul style="list-style-type: none"> • Población en estudio. • Intervención realizada. • Resultados considerados.
<p>2. ¿Fue aleatoria la asignación de los pacientes al tratamiento?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Pista</p> <ul style="list-style-type: none"> • ¿Se mantuvo oculta la secuencia de aleatorización?
<p>3. ¿Fueron considerados hasta el final del estudio todos los pacientes?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Pistas</p> <ul style="list-style-type: none"> • ¿Fue el seguimiento completo? • ¿Los pacientes fueron analiza- dos en los grupos a los que fue- ron aleatorizados?

Tabla 1.
Fuente: propia

En este primer bloque de preguntas guía se pretende que el lector evalúe la validez de los resultados, condición *sine qua non* para continuar en la valoración del artículo. También son denominadas preguntas de eliminación, con alguna respuesta negativa se puede descartar el artículo.



Enmascarados

Técnica de ocultamiento sobre los individuos, investigadores y analistas de datos durante el desarrollo de un ensayo clínico, con el fin de controlar el efecto placebo, la subjetividad en la evaluación, entre otros.

Preguntas de detalle

Guía	Comentarios
<p>1. ¿Estuvieron los pacientes, el personal de salud y el personal involucrado en el estudio enmascarados respecto del tratamiento?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Pistas</p> <ul style="list-style-type: none"> ● Los pacientes. ● El personal que aplica las terapias. ● El personal que realiza las mediciones.
<p>2. ¿Los grupos de estudio eran similares al inicio del estudio?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Pista</p> <ul style="list-style-type: none"> ● Verificar otros factores que pudieran tener efecto sobre el resultado: edad, sexo, etc.
<p>3. Aparte de la intervención en el estudio, ¿los grupos fueron tratados de forma similar?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	

Tabla 2.
Fuente: propia

¿Son precisos los resultados del estudio?

Guía	Comentarios
1. ¿Cuál fue la magnitud del efecto del tratamiento?	Pistas <ul style="list-style-type: none">• ¿Qué resultados se midieron?• ¿Qué estimadores se utilizaron? ¿RR, OR, DR o diferencias de medias, sobrevida y NNT?
2. ¿Cuán precisa fue la estimación del efecto del tratamiento?	Pistas <ul style="list-style-type: none">• ¿Cuáles son sus intervalos de confianza?, ¿fueron lo suficientemente estrechos?• Si no fueron significativos, ¿qué sugieren?

Tabla 3.
Fuente: propia



Magnitud del efecto

Medida utilizada para expresar el efecto de un tratamiento o una exposición sobre variables de desenlace o dependientes.

¿Me ayudarán estos resultados en el cuidado de mis pacientes o en la situación problema?

Guía	Comentarios
<p>1. ¿Los resultados pueden ser aplicados en el cuidado de mis pacientes o en la situación que se me presenta?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Pista</p> <p>¿Los pacientes incluidos en el estudio son lo suficientemente parecidos a mis pacientes en términos de características generales, edad, sexo, severidad de la enfermedad, inicio de la misma y presencia de factores pronósticos?</p>
<p>2. ¿Se consideran todos los resultados de importancia clínica?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Pista</p> <p>En caso negativo, ¿en qué afecta esto la decisión a tomar?</p>
<p>3. ¿Los beneficios a obtener justifican los riesgos y los costos del tratamiento?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Pista</p> <p>Es improbable que pueda deducirse del estudio, pero ¿qué piensa usted al respecto?</p>

Tabla 4.
Fuente: propia

Requiero saber sobre un factor de riesgo, daño o etiología

Este escenario está dedicado al análisis de la evidencia de artículos que publican resultados sobre un factor de riesgo, daño o etiología.

Guía	Comentarios
1. ¿Son válidos los resultados del estudio?	<p>Guías primarias</p> <p>¿Fueron claramente identificados los grupos en comparación y similares con respecto a importantes determinantes de resultados?</p> <ul style="list-style-type: none">● Sí.● No.● No se puede decir. <p>¿Los resultados y la exposición fueron medidos en la misma forma en los grupos que están siendo comparados?</p> <ul style="list-style-type: none">● Sí.● No.● No se puede decir. <p>¿El seguimiento fue lo suficientemente largo y completo?</p> <ul style="list-style-type: none">● Sí.● No.● No se puede decir. <p>Guías secundarias</p> <p>¿La relación temporal es correcta?, ¿se identifica que la posible causa ocurrió antes del desenlace?</p> <ul style="list-style-type: none">● Sí.● No.● No se puede decir. <p>¿Hay un gradiente de dosis respuesta?</p> <ul style="list-style-type: none">● Sí.● No.● No se puede decir.

<p>2. ¿Cuáles son los resultados?</p>	<p>¿Cuán fuerte es la asociación entre la exposición y el resultado (enfermedad)?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Cuán precisa fue la estimación del riesgo?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir.
<p>3. ¿Me ayudarán estos resultados en el cuidado de mis pacientes?</p>	<p>¿Los resultados son aplicables a mi práctica?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Cuál es la magnitud del riesgo?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Debería intentar detener la exposición?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir.

Tabla 5.
Fuente: propia

Requiero conocer del desempeño de una prueba diagnóstica

Este escenario está dedicado al análisis de la evidencia de artículos que publican resultados sobre el desempeño de pruebas diagnósticas.

Guía	Comentarios
1. ¿Son válidos los resultados del estudio?	<p>Guías primarias</p> <p>¿Hubo una comparación independiente y enmascarada?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir. <p>¿La muestra de sujetos incluía un espectro apropiado de pacientes en los que la prueba diagnóstica será aplicada en la práctica usual?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir. <p>Guías secundarias</p> <p>¿Los resultados de la prueba que está siendo evaluada influyeron en la decisión de realizar el patrón de referencia?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir. <p>¿Tienen los métodos para realizar la prueba descrita suficiente detalle para permitir su reproducción?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir.

<p>2. ¿Cuáles son los resultados?</p>	<p>¿Son calculadas las razones de probabilidad o se presentan los datos necesarios para su cálculo?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir.
<p>3. ¿Me ayudarán estos resultados en el cuidado de mis pacientes?</p>	<p>¿Serán reproducibles los resultados de la prueba y su interpretación será satisfactoria en mi medio?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Son los resultados aplicables a mis pacientes, teniendo en cuenta que son similares a los incluidos en el estudio?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Los resultados cambiarán mi tratamiento?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>Como resultado de la prueba, ¿los pacientes estarán mejor?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir.

Tabla 6.
Fuente: propia



Razones de probabilidad

Medida que corresponde a la razón entre la probabilidad de tener un resultado positivo y la probabilidad de tener uno negativo entre individuos enfermos.



Lectura recomendada

Lithium use in pregnancy and the risk of cardiac malformations

Elisabetta Patorno et ál.

Pronóstico





Figura 1.
Fuente: Shutterstock/619483583

Como parte de la atención en salud, se han dejado claros los escenarios que explican la necesidad de diagnóstico, de implementar un tratamiento y de evaluar los daños asociados al tratamiento o exposición exógenos. Sin embargo, es necesario que el profesional evalúe el pronóstico de los pacientes, es decir, que examine los posibles resultados de una enfermedad y la probabilidad con la que pueden esperarse que se produzcan para complementar la toma de decisiones. El conocimiento del pronóstico de un paciente ayuda a tomar las decisiones correctas sobre su tratamiento. No se deben recomendar tratamientos caros o potencialmente tóxicos si la recuperación ocurrirá de todos modos. Por otro lado, los pacientes pueden estar destinados a tener malos resultados a pesar de cualquier tratamiento que reciban. La terapia agresiva en estos individuos solo puede prolongar el sufrimiento y malgastar recursos. Cualesquiera que sean las posibilidades de tratamiento, al comprender el pronóstico y presentar el curso futuro esperado de la enfermedad de un paciente, los profesionales también ofrecen tranquilidad y esperanza, o preparación para la muerte o discapacidad a largo plazo.

Cuando se habla de estudios que evalúan el pronóstico de una enfermedad, se hace referencia a la identificación de variables que están relacionadas con la predicción del resultado clínico en el paciente, por ejemplo, la edad, la presencia de comorbilidad o condiciones sociodemográficas que, sin ser causales, modifican el curso de la enfermedad. Estas variables son llamadas factores pronósticos y no necesariamente son factores de riesgo para la enfermedad. Se hace alusión a un factor de riesgo cuando dicha variable determina (causa) la presencia de una enfermedad. Los estudios que evalúan factores pronósticos y factores de riesgo pueden ser analizados desde la misma perspectiva de la validez del resultado.

Los mismos diseños observacionales (cohorte y de casos y controles) que los investigadores usan para valorar la asociación exposición-enfermedad pueden ser utilizados para evaluar información pronóstica en los pacientes. Implícitamente, los ensayos clínicos también abordan problemas de pronóstico. Los resultados reportados para el grupo de tratamiento y el grupo de control proporcionan información pronóstica: los datos del grupo de control informan sobre el pronóstico en pacientes que no recibieron el tratamiento experimental, mientras que los resultados de los grupos experimentales informan sobre el pronóstico en pacientes que recibieron la intervención. En este sentido, cada brazo de un ensayo clínico representa un estudio de cohorte.

Validez de los resultados en estudios de pronóstico

Muestra representativa de la población

Se debe recordar que un resultado es sesgado cuando está desviado de la verdad, por tanto, si la muestra elegida para un estudio difiere sistemáticamente de la población de interés, el resultado de un desenlace evaluado en estos puede resultar con sobreestimación o subestimación del mismo, haciendo ver que un paciente puede tener peor o mejor pronóstico que uno de la población general de interés. La muestra que tiene la característica de ser altamente seleccionada es denominada muestra no representativa.

Se puede identificar una muestra altamente seleccionada cuando ocurre un filtro antes de que los sujetos ingresen al estudio. Este es el caso del llamado sesgo de selección, como, por ejemplo, cuando los pacientes que aceptan participar en una investigación provienen de una institución que maneja patologías de mayor complejidad. Esta muestra puede sistemáticamente tener sujetos con mayor severidad de la enfermedad, que aquel que aún no acude a dicha institución, haciendo ver que el pronóstico, por ejemplo, en la sobrevida, sea mucho menor que un sujeto de la población general de interés.

Homogeneidad frente al riesgo en el pronóstico

Al igual que ocurre en los estudios de cohorte, las diferencias entre los dos grupos al inicio del seguimiento pueden tener una relación con el desenlace en el seguimiento. Lo mismo ocurre en los estudios de pronóstico, por tanto, es necesario que los pacientes comiencen a ser observados en un mismo punto de la enfermedad.

Esto no indica que sea necesariamente en sus inicios, sino más bien que otras variables que afectan el pronóstico del desenlace tengan una distribución homogénea entre los grupos que definen los factores pronósticos. Factores como la edad, el sexo y las condiciones sociales de base pueden hacer que el desenlace (pronóstico en una variable) sea diferente y no necesariamente atribuible al factor pronóstico en estudio. En ocasiones, se debe tener en cuenta la interacción que pueda haber entre ellos y que modifique en diferentes niveles el pronóstico observado. Para controlar este efecto, se deben incluir análisis separados que permitan describir dichas modificaciones de efecto. Este análisis generalmente se llama análisis de ajuste y se realiza mediante modelos estadísticos multivariados.



Figura 2.
Fuente: Shutterstock/261087332

Seguimiento de los sujetos

Como todo diseño de estudio que implique seguimiento, el resultado final se puede ver comprometido cuando los sujetos en estudio se pierden. La magnitud del sesgo que se produce en el estimado del pronóstico depende de si los sujetos que se pierden tenían mayor o menor riesgo de hacer el desenlace. Esto producirá que se sobrestime o subestime el resultado final. Ejemplo: si de un seguimiento en la supervivencia de pacientes con coinfección TB-VIH ocurre que en cerca del 15 % de los pacientes no se pudo establecer el estado final frente a la mortalidad tras cinco años de seguimiento, la tasa obtenida de supervivencia que al final es estimada de 65 % podría estar en realidad entre 50% y 70%, implicando esto un rango de valores que no afectan de manera importante la conclusión sobre la supervivencia de los pacientes con TB-VIH. El impacto que pueda tener la proporción de pacientes que se pierden en el seguimiento sobre la estimación del estudio debe ser analizado con el alcance del resultado para la toma de decisiones frente al paciente.

Medición objetiva del desenlace (pronóstico)

La medición de la variable principal sobre la que se desea determinar el pronóstico debería estar muy bien definida. Esta puede ser tan fácilmente identificable (mortalidad por cualquier causa), requerir de una valoración especializada (mortalidad por causa cardiovascular) o, en casos más especiales, se da una medida bastante refinada que requiere de instrumentos especiales para su valoración (calidad de vida). En todos los casos, dichas variables y la medición deben estar bien definidas por los investigadores.

Resultados en estudios de pronóstico o riesgo

Las medidas utilizadas en estos estudios son propias del diseño de cohorte o experimental que se utilice, por ende, medidas como RR y OR son frecuentemente utilizadas; sin embargo, un tipo de resultado especial es examinado en estos estudios y la probabilidad de ocurrencia de un evento dado un periodo. Una forma informativa para describir estos resultados es una curva de supervivencia, que no es más que un gráfico que relaciona la proporción de pacientes que hacen un evento (o permanecen libres del mismo) y el tiempo en el que ocurre dicho evento. Los métodos estadísticos utilizados son denominados de análisis tiempo-evento. Los datos que se registran son generalmente desenlaces dicotómicos (sí/no, infectado/no infectado) y se registra el tiempo en que ocurre dicho estado.



Instrucción

Revise el organizador gráfico que se encuentra en los recursos de este eje para complementar esta información.

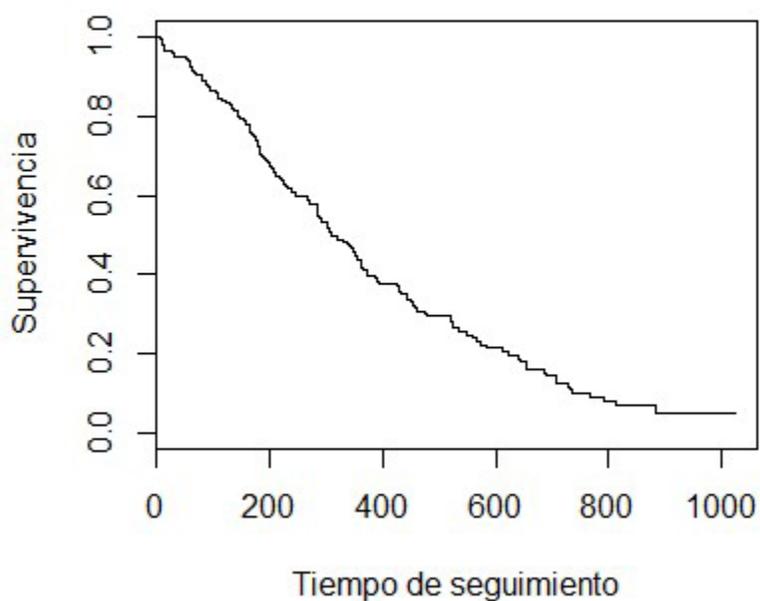


Figura 3. Curva de supervivencia de pacientes con cáncer de pulmón avanzado del North Central Cancer Treatment Group
Fuente: Ravishanker (2004)



Supervivencia

Probabilidad de sobrevivir de un individuo más allá de un tiempo determinado.

Requiero conocer sobre el pronóstico de una enfermedad

Este escenario está dedicado al análisis de la evidencia de artículos que publican resultados sobre el pronóstico de una enfermedad.

Guía	Comentarios
1. ¿Son válidos los resultados del estudio?	<p>Guías primarias</p> <p>¿Hubo una muestra de pacientes representativa y bien definida en un punto similar del curso de la enfermedad?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir. <p>¿El seguimiento fue suficientemente largo y completo?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir. <p>Guías secundarias</p> <p>¿Fueron usados resultados objetivos y no sesgados?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir. <p>¿Hubo ajustes por factores pronósticos relevantes?</p> <ul style="list-style-type: none">• Sí.• No.• No se puede decir.

<p>2. ¿Cuáles son los resultados?</p>	<p>¿Cuán grande es la probabilidad del evento en un periodo específico?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Cuán precisa fue la estimación de dicha probabilidad de ocurrencia del evento resultado?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir.
<p>3. ¿Me ayudarán estos resultados en el cuidado de mis pacientes?</p>	<p>¿Fueron los pacientes similares a mis pacientes?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Los resultados conducirán directamente a seleccionar o evitar la terapia?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. <p>¿Los resultados son útiles para aconsejar a los pacientes?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir.

Tabla 7.
Fuente: propia

Aplicación práctica de las guías

Se realizará la aplicación de la guía de usuario de análisis crítico a un artículo publicado en revistas prestigiosas del medio científico. Corresponde a un artículo de Ovalle et ál. (2000) sobre un ensayo clínico sobre la efectividad de una intervención farmacológica en **pielonefritis** aguda en el embarazo.



Pielonefritis

La pielonefritis o infección urinaria alta es una infección del riñón que involucra el parénquima renal, la pelvis renal y los cálices renales.

Guía para la lectura crítica de un artículo sobre terapia o prevención

¿Son válidos los resultados de la investigación?

Preguntas de eliminación

Guía	Comentarios
<p>1. ¿Se orienta el artículo a una pregunta claramente definida?</p> <ul style="list-style-type: none">● Sí.● No.● No se puede decir.	<p>Se identificó una pregunta clara que hacía referencia al efecto de dos regímenes antibióticos para el tratamiento de pielonefritis en mujeres en estado de embarazo.</p> <p>Los autores indican que las pacientes fueron tratadas por 14 días y su egreso hospitalario se dio según condición clínica. Se realizaron cultivos de orina durante el control y fin del seguimiento a los 28 días. Se realizó un seguimiento de 101 pacientes incluidos en el estudio pérdida tres pacientes, los cuales fueron excluidos del estudio.</p>
<p>2. ¿Fue aleatoria la asignación de los pacientes al tratamiento?</p> <ul style="list-style-type: none">● Sí.● No.● No se puede decir.	<p>Se describe que se realiza la asignación aleatoria simple. En el estudio fueron aleatorizadas alternadamente a recibir cefradina o cefuroxima. Se describieron claramente las dos intervenciones a comparar.</p>

<p>3. ¿Fueron considerados hasta el final del estudio todos los pacientes?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Se excluyeron del estudio por no cumplir el criterio de inclusión, suspensión de tratamiento o hipersensibilidad al mismo grupo de cefradina: 49/55 (seis perdidas), y en el grupo de cefuroxima: 52/56 (cuatro pérdidas). Se incluyeron 111 pacientes y completaron el tratamiento en el grupo de cefradina: 49/55 (seis perdidas) y en el grupo de cefuroxima: 52/56 (cuatro pérdidas).</p>
<p>4. ¿Estuvieron los pacientes, el personal de la salud y el personal involucrado en el estudio enmascarados respecto del tratamiento?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>No se especifica ningún tipo de control de cegamiento.</p>
<p>5. ¿Los grupos de estudio eran similares al inicio del estudio?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Los autores proveen una tabla con las características generales de la población en estudio, donde se observa que son similares en relación a su condición de ingreso, edad gestacional y paridad.</p> <p>Además, no hubo diferencias en la distribución de las distintas variables obstétricas de los grupos. Cuatro pacientes del grupo de cefradina (8,7 %) y tres (5,8 %) del grupo de cefuroxima tuvieron parto prematuro. Dos pacientes de este último grupo tuvieron abortos espontáneos.</p>
<p>6. Aparte de la intervención en estudio, ¿los grupos fueron tratados de forma similar?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>No queda claro el tratamiento recibido por 15 pacientes infectadas por <i>E. coli</i> ni a qué grupo pertenecían, por lo que no se puede asegurar que todas las embarazadas fueron tratadas de forma similar a diferencia de la intervención experimental y la posible presencia de cointervenciones, a lo que los autores no hacen referencia.</p>

Tabla 8.
Fuente: propia

¿Son precisos los resultados del estudio?

Guía	Comentarios
<p>7. ¿Cuál fue la magnitud del efecto del tratamiento?</p>	<p>Se indica que las pacientes que recibieron cefuroxima tuvieron menos días de fiebre y lograron la curación clínica en menor número de días que las tratadas con cefradina. Fiebre (media 1,7 variación 1-4 vs. media 2,2 días variación 1-7); $p < 0,05$. Curación clínica (media 2,7 variación 1-5 vs. media 3,1 días variación 1-8); $p < 0,05$. No hubo diferencias significativas en los días de estadía y en los días de administración del antibiótico por vía intravenosa que tuvieron las mujeres de ambos grupos.</p> <p>En cuanto a la evaluación de la eficacia clínica y microbiológica, se indica que la curación bacteriológica fue significativamente mayor en las pacientes que usaron cefuroxima que en las pacientes que recibieron cefradina (78,8 % [41/52] vs. 59,2 % [29/49], $p < 0,05$). El fracaso bacteriológico y la reinfección fueron significativamente menores en el grupo de cefuroxima que en el grupo de cefradina (21,2 % [11/52] vs. 40,8 % [20/49], $p < 0,05$). No hubo diferencias significativas en la curación clínica ni en el fracaso y mejoría clínica, ni en la recurrencia que presentaron las mujeres de ambos grupos.</p>
<p>8. ¿Cuán precisa fue la estimación del efecto del tratamiento?</p>	<p>Se indican con significancia estadística ($p < 0,05$) las siguientes mediciones:</p> <ul style="list-style-type: none"> ● Fiebre. ● Curación clínica. ● Curación bacteriológica. ● Fracaso bacteriológico-reinfección. ● Costos de antibióticos y de hospitalización. <p>No se indicaron intervalos de confianza.</p>

Tabla 9.
Fuente: propia

¿Me ayudarán estos resultados en el cuidado de mis pacientes o en la situación problema?

Guía	Comentarios
<p data-bbox="293 863 639 1052">9. ¿Los resultados pueden ser aplicados al cuidado de mis pacientes o situación que se me presenta?</p> <ul data-bbox="334 1167 630 1325" style="list-style-type: none"><li data-bbox="334 1167 399 1199">● Sí.<li data-bbox="334 1230 415 1262">● No.<li data-bbox="334 1293 630 1325">● No se puede decir.	<p data-bbox="672 537 1425 1644">Los resultados tienden a sugerir que en la pielonefritis aguda del embarazo la administración de cefuroxima constituye una terapia más eficiente y de costo total similar al obtenido con el uso de cefradina. Mostrándose que la cefuroxima usada por 14 días, 3,4 días de promedio por vía intravenosa, logró ser una terapia más eficiente al erradicar el patógeno causal de la vía urinaria en el 78,8 % de los casos. En comparación, la cefradina administrada por los mismos días, 3,8 días de promedio por vía intravenosa, logró la curación bacteriológica en el 59,2 %. Por consiguiente, el fracaso bacteriológico y la reinfección en el grupo con cefuroxima fueron menores que en el grupo con cefradina. El hecho de que la aleatorización fuera simple (por oportunidad y llegada al centro hospitalario) y los autores del estudio no aplicaran un <i>matching</i> individual hace dudar de los resultados, ya que los resultados clínicos tienen su origen en muchas causas y el tratamiento no es más que una de ellas. Creemos que puede existir un sesgo en la descripción y el análisis de los resultados, ya que se indica en el estudio que se perdió el seguimiento de 10 pacientes (tres por resistencia), pero también que hubo dentro de las infectadas por <i>E. coli</i> 14 resistentes a cefradina y una a cefuroxima. No se indica el grupo de aleatorización al que pertenecían, cuál fue su tratamiento final ni en qué grupo se analizaron. Por lo anterior, sugerimos que este estudio no posee una validez interna adecuada.</p>

<p>10. ¿Se consideran todos los resultados de importancia clínica?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Aunque los autores indican que los datos recolectados en la muestra incluían infección cérvicovaginal y la vaginosis bacteriana se incluye como factor de riesgo en pielonefritis o como causante de recidiva, se podría presuponer que dichos factores actúan como variables confusoras o modificadoras y debían tomarse en cuenta en el proceso de análisis pudiéndose realizar un análisis estratificado de la información; sin embargo, no se realizó así y solo se indican: la coexistencia de infección cérvicovaginal (76,9 % vs. 52,9 %; $p < 0,05$) y el aumento de la respuesta leucocitaria en la muestra vaginal (76,9 % vs. 47,1 %; $p < 0,01$). Esto sugiere que, en casos de fallas, habría un deterioro mayor de la respuesta inmune y que las pacientes requerirían de más días de terapia.</p>
<p>11. ¿Los beneficios a obtener justifican los riesgos y los costos del tratamiento?</p> <ul style="list-style-type: none"> ● Sí. ● No. ● No se puede decir. 	<p>Aunque los autores aseguran los supuestos beneficios, se considera que se deben realizar más estudios con una metodología más estricta, mejor aleatorización y seguimiento de las pacientes, y análisis de los resultados. Algunos resultados no fueron significativos, lo que sugiere que tal vez se debió aumentar la potencia del estudio (mayor muestra). No se indican los valores de NNT, al realizar los cálculos para curación bacteriológica: cefuroxima 78,8 % (41/52) y cefradina 59,2 % (29/49). El NNT fue de -5.09 con IC95% de -2.68 a -50.46.</p>

Tabla 10.
Fuente: propia



Lectura recomendada

Oxygen therapy in suspected acute myocardial infarction

Robin Hofmann et ál.



Lectura recomendada

Para finalizar, lo invito a realizar el control de lectura sobre:

Eficacia de un programa educativo estructurado en población diabética chilena

Roberto Bächler et ál.

En este eje se hizo entrega del material para analizar los resultados de estudios que evalúan una intervención, una prueba diagnóstica, el daño de una exposición y el pronóstico de una enfermedad. Estas guías constituyen el pilar fundamental en el análisis basado en la evidencia que permite al profesional hacer una aplicación práctica de los resultados de investigación para integrarlos a la práctica diaria.

Manterola, C. y Otzen, T. (2014). Análisis crítico de la literatura biomédica, *International Journal of Morphology*, 32(2), 599-607. Doi: <http://doi.org/10.4067/S0717-95022014000200037>

Ovalle, A. et ál. (2000). Estudio prospectivo, randomizado, comparativo de la eficacia, seguridad y costos de cefuroxima vs. cefradina en la pielonefritis aguda del embarazo. *Revista Médica de Chile*, 128(7), 749-757.

Ravishanker, N. (2004). Survival analysis. *Journal Technometrics*, 46(1), 1-16. Doi: <http://doi.org/10.1198/tech.2004.s740>

Richardson, W. et ál. (1999). Users' guides to the medical literature. *Jama*, 281(13), 1214-1219. Doi: <http://doi.org/10.1001/jama.281.13.1214>

ANÁLISIS BASADO EN LA EVIDENCIA

Jorge Mario Estrada

EJE 4

Propongamos



Aplicación de los resultados de las investigaciones clínicas



Evaluando la exactitud de una prueba diagnóstica



Estudio de caso

Un hombre de 60 años acude al servicio de urgencias con dificultad para respirar, ortopnea, disnea paroxística nocturna, presión venosa yugular elevada y edemas de miembros inferiores. Tiene antecedentes de enfermedad pulmonar obstructiva crónica. De acuerdo con el nivel de atención, se sospecha falla cardíaca congestiva en el paciente; sin embargo, la prueba ideal es el ecocardiograma, el cual implica un costo alto y no es fácil su realización en servicios de urgencias, de tal manera que se piensa en pruebas rápidas que permitan fácilmente diferenciar si el paciente tiene falla cardíaca congestiva o cualquier otra entidad patológica.

En este sentido, se requiere conocer la existencia de pruebas rápidas en el servicio de urgencias que puedan ser implementadas, según su desempeño diagnóstico.

Ante la situación anterior, se requiere revisar la evidencia disponible sobre algunas de estas pruebas y definir su implementación de manera local.

La pregunta PICO que el médico tratante se plantea es:

- P: paciente de 60 años con sintomatología sospechosa de insuficiencia cardiaca congestiva.
- I: prueba rápida para la detección de falla cardiaca congestiva.
- C: *gold estándar*: ecocardiografía.
- O: desempeño diagnóstico en términos de sensibilidad y especificidad de la prueba.

En los pacientes mayores con sospecha de insuficiencia cardiaca congestiva (ICC), ¿cuál es la precisión de una prueba rápida de cribado para diagnosticar ICC comparada con la ecocardiografía?

Buscando la evidencia

El médico tratante selecciona las palabras “diagnóstico”, “cribado” e “insuficiencia cardiaca congestiva” en la página de PubMed, utilizando los términos de búsqueda. La búsqueda da 48 citas. Al limitar los estudios al idioma inglés y a los seres humanos, en los últimos cinco años, se acorta la lista a 21.

Revisa los resúmenes, buscando artículos que se centren en pacientes con sospecha de ICC. Ubica un artículo que reporta los resultados de una prueba denominada péptido natriurético tipo B1 (Maisel et ál., 2002). Recupera electrónicamente el texto completo y comienza a leerlo, esperando que sus métodos y resultados justifiquen el uso de la prueba en el servicio de urgencias.

Se procede a examinar bajo la guía de usuario de análisis para estudios de evaluación de pruebas diagnósticas y se obtiene:

Cómo interpretar un artículo sobre pruebas diagnósticas

Guía	Comentarios
<p>1. ¿Son válidos los resultados del estudio?</p> <p>¿Hubo una comparación independiente y enmascarada?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí la hubo, pues el estándar de referencia fue el diagnóstico clínico que incluía otras pruebas. Este fue hecho por dos cardiólogos independientes, lo cual se hizo así porque el gold estándar de esta serie de pacientes puede ser la ecocardiografía. Hacen una aclaración con la no disponibilidad de la misma en todos los servicios, así como también la necesidad de realizar un Dx rápido de falla cardiaca. En el servicio de urgencias la falta el gold único para falla cardiaca congestiva. Además, estos cardiólogos tuvieron cegamiento ante los resultados de la medición del BNP.</p>
<p>¿La muestra de sujetos incluía un espectro apropiado de pacientes en los que la prueba diagnóstica será aplicada en la práctica usual?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Se incluye una muestra de 1586 pacientes y siete centros multinacionales, lo cual incluye una muestra amplia y variada de pacientes con disnea, aunque llama la atención la alta prevalencia de falla cardiaca (47 %).</p>
Guía secundarias	
<p>¿Los resultados de la prueba que está siendo evaluada influyeron en la decisión de realizar en el patrón de referencia?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>No, el test diagnóstico no modificó la realización del estándar de referencia, dado que se presenta un cegamiento para todos estos pacientes.</p> <p>Sin embargo, no se especifica el grado de concordancia en cuanto a la confirmación del diagnóstico por los cardiólogos, esto daría mayor rigor al grado exactitud de la clasificación correcta por parte de patrón de referencia, siendo una fuente de posible sesgo.</p>
<p>¿Fueron los métodos para realizar la prueba descrita con suficiente detalle que permita su replicación?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí, se explica el tiempo de realización, el kit utilizado con suficiente detalle y la clase de muestra a extraer, así como la forma de toma de la muestra.</p>

<p>2. ¿Cuáles son los resultados?</p> <p>¿Son calculadas las razones de probabilidad o se presentan los datos necesarios para su cálculo?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>No se presentan los cocientes de probabilidad, pero se presentan la sensibilidad y la especificidad para diferentes puntos de corte, así como el análisis de la curva ROC y, con estos datos, y los odds definidos en el estudio es posible calcular los cocientes de probabilidad.</p> <p>Para el punto de corte escogido por el autor, según curva ROC, el cociente de probabilidad para un resultado positivo es de 3,75, lo que quiere decir que es 3,75 veces más común un valor de 100 ng/ml de péptido natriurético tipo B en un paciente con falla cardiaca congestiva que uno sin falla cardiaca.</p> <p>A mayor valor de BNP, mayor es la probabilidad de encontrar estos niveles en pacientes.</p>
<p>3. ¿Me ayudarán estos resultados en el cuidado de mis pacientes?</p> <p>¿Serán reproducibles los resultados de la prueba y su interpretación será satisfactoria en mi medio?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Con lo observado en el estudio, se piensa que los resultados pueden ser reproducidos adecuadamente, pues se estableció en el estudio cómo tomarlo, con qué analizarlo y el valor límite para considerarlo positivo para falla cardiaca. Además de lo anteriormente escrito, fue posible calcular los cocientes de probabilidad con los datos aportados.</p>
<p>¿Son los resultados aplicables a mis pacientes, teniendo en cuenta que son similares a los incluidos en el estudio?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí, en un servicio de urgencias de atención de adultos para descartar la causa de disnea de origen desconocido puede ser útil, teniendo en cuenta los diagnósticos excluidos.</p>
<p>¿Los resultados cambiarán mi tratamiento?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí, cambiaría el manejo la realización de dichos test, dado que se iniciaría más rápidamente el manejo del paciente, aunando a dicho test a la historia clínica del paciente y a la realización de pruebas básicas, como radiografía de tórax.</p>
<p>Como resultado de la prueba, ¿los pacientes estarán mejor?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí, es beneficioso dado que es una patología que diagnosticada de forma rápida y adecuadamente nos permite hacer un manejo temprano que cambia el pronóstico del paciente y puede disminuir la comorbilidad y los costos en la atención.</p>

Tabla 1.
Fuente: Maisel et ál. (2002)



Lectura recomendada

Cómo interpretar un artículo sobre pruebas diagnósticas

María Burgos y Carlos Manterola

Como se puede evidenciar bajo el análisis de esta publicación, se justificaría la utilización de esta nueva prueba rápida en los servicios de urgencias para el diagnóstico de ICC, así como también la validez de los resultados obtenidos, dejando entrever la calidad metodológica del mismo, de tal manera que impactaría su aplicación en la práctica actual.

Evaluando la evidencia de una exposición



Estudio de caso

Un hombre de 75 años se encuentra hospitalizado con diagnóstico de colitis ulcerativa, enfermedad grave que causa úlceras e irritación en el revestimiento interno del colon y del recto.

Dentro de las explicaciones al paciente sobre las causas de dicha enfermedad, surge la duda si la exposición a cigarrillo en el pasado puede haber influido de manera importante en la aparición de la enfermedad.

Dada la situación anterior, se requiere revisar la evidencia disponible sobre el papel que pudo jugar la exposición al cigarrillo para el desarrollo de la enfermedad.

La pregunta PICO que el médico tratante se plantea es:

- P: paciente de 75 años con diagnóstico de colitis ulcerativa.
- I: exposición a cigarrillo.
- C: no exposición.
- O: colitis ulcerativa.

¿Cuál es el efecto que puede tener el consumo de cigarrillo para el desarrollo de colitis ulcerativa en pacientes adultos mayores?

Buscando la evidencia

El médico selecciona las palabras “tabaquismo” y “colitis ulcerativa” en la página de PubMed, utilizando los términos de búsqueda. La búsqueda da 23 citas. Al limitar los estudios al idioma inglés y a los seres humanos, se acorta la lista a 12.

Revisa los resúmenes, buscando artículos que se centren en pacientes con colitis. Recupera un artículo que reporta los resultados de la evaluación de la exposición a cigarrillo en la infancia en relación a la colitis ulcerativa (Sandler, Sandler, McDonnell y Wurzelmann, 1992) y comienza a leerlo, esperando que sus métodos y resultados justifiquen la respuesta a su paciente.

Se procede a analizar el estudio bajo la guía de usuario de artículos sobre evaluación de daño.

Guía	Comentarios
<p>1. ¿Se identificó claramente que los grupos a comparar fueran similares con respecto a determinantes importantes de resultados, además de aquel de principal interés?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Los grupos son homogéneos frente a diferentes variables, aunque los casos seleccionados no son representativos de la población, además de la forma de establecer si realmente pueden haberse generado sesgos de mala clasificación sobre la condición caso, y la confirmación que los autores hacen del caso sobre una muestra del 15 % de posibles casos no se justifica de una forma estadística.</p>
<p>2. ¿Se midieron de la misma forma las exposiciones y los resultados en los grupos a comparar?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>La exposición fue medida bajo la misma definición tanto para casos, como para controles; sin embargo, la definición utilizada no es clara porque no se justifica por qué el criterio de más de 100 cigarrillos para fumador activo. Además, no hay claridad en el criterio de fumador pasivo (temporalidad, dosis, efectos de hacinamiento), esto podría llevar a sesgos de mala clasificación sobre la exposición y sesgos de memoria por el tiempo tan amplio de la valoración de la exposición y el desenlace.</p> <p>Por el recuerdo diferente entre los casos y los controles en la exposición y la eventual búsqueda y registro diferente de la exposición por el entrevistador que no se sabe si era ciego y, además, no se sabe si se verificó la exposición.</p>
<p>3. ¿El seguimiento fue lo suficientemente largo y completo?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>No hay seguimiento, es un estudio de casos y controles.</p>
<p>4. ¿La relación temporal es correcta?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Aunque se estipula que la exposición estudiada ocurrió en la infancia (<15 años), las exposiciones intermedias entre ese tiempo y el inicio de síntomas pueden ser distintos, por ende, no permite atribuirle únicamente a esa primera exposición el desarrollo de la colitis.</p> <p>Mezclan exposición en la niñez con riesgos del fumador adulto e incluso lo comparan con el no fumador (0,87), independientemente de su exposición en la niñez y lo comparan también con el no fumador no expuesto en la niñez (0,53).</p>

<p>5. ¿Hay gradiente dosis-respuesta?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Hay que tener en cuenta que no todos los estimativos puntuales son significativamente diferentes de uno y que algunos intervalos de confianza alcanzan o contienen el valor de nulo de no asociación.</p> <p>Veamos los OR puntuales comparando la exposición en la niñez y el estatus de fumador o no con el nivel de referencia en los casos y los controles.</p> <p>OR</p> <p>No expuesto - no fumador 1.00</p> <p>Expuesto - no fumador 0.50</p> <p>p = 0,041</p> <p>No expuesto - fumador 0.34</p> <p>p = 0,016</p> <p>Expuesto - fumador 0.72</p> <p>p = 0,379</p> <p>En la evaluación de la hipótesis de que la exposición en la niñez disminuye el riesgo de colitis ulcerativa en el adulto no se evidencia gradiente en la fuerza de asociación.</p> <p>La exposición en la niñez disminuye el riesgo en un 50 % en no fumadores y un 28 % en fumadores.</p> <p>Paradójicamente ante la hipótesis examinada, la ausencia de exposición en la niñez en fumadores adultos está asociada con la mayor disminución del riesgo colitis ulcerativa: lo disminuye hasta en un 66 %.</p>
<p>6. ¿Cuán fuerte es la asociación entre la exposición y el resultado?</p>	<p>La magnitud del efecto encontrado fue alta, con una reducción de riesgo de desarrollar colitis ulcerativa en un 50 % (OR = 0,5); sin embargo, no es estadísticamente significativo, ya que el intervalo contiene valor nulo 1.</p>
<p>7. ¿Cuán precisa es la estimación del riesgo?</p>	<p>Es imprecisa por el intervalo tan amplio (IC 95% 0,25 - 1), lo cual hace perder validez a la asociación. Además, el intervalo contiene el 1 de no asociación, es decir, el verdadero valor poblacional del OR estaría desde OR 0,25 hasta la no asociación OR = 1.</p>
<p>8. ¿Los resultados son aplicables a mi práctica?</p>	<p>No, aunque el estudio muestra un efecto protector de fumar frente al desarrollo de colitis ulcerativa, puede estar sesgado por diferentes aspectos en el diseño, como mala clasificación y sesgos de recuerdo. Además, la mayoría de los casos y controles fueron en personas de altos ingresos y que pueden tener hábitos muy distintos en otros aspectos que están relacionados con la colitis y no fueron evaluados en el estudio.</p> <p>Este estudio carece de validez interna que permita generalizar.</p> <p>Se hizo un diseño pareado y no se realizó un análisis pareado.</p>

<p>9. ¿Cuál es la magnitud del riesgo?</p>	<p>Alta, con una reducción del riesgo del 50 %. Con cada dos exposiciones en la niñez se evitaría un caso de colitis ulcerativa en el adulto. Esto es según los resultados asumidos como válidos y significativos, sin embargo, dada la evaluación del estudio con todos sus errores metodológicos, esto no es cierto.</p>
<p>10. ¿Debería intentar detener la exposición?</p>	<p>No, de este estudio no se puede establecer una relación causal como se colige de la aplicación de las reglas de evidencia para causalidad.</p> <ul style="list-style-type: none"> - En general, los estudios de casos y controles son débiles para establecer relación causal. - Contextualmente, la asociación no es consistente de estudio a estudio. - La secuencia temporal de la exposición y el resultado están en la dirección correcta. - No existe gradiente dosis-respuesta. - La asociación no es admisible. - No se excluyeron otras explicaciones plausibles. - Los resultados del artículo no están en armonía con nuestra comprensión actual de las respuestas biológicas. - Los resultados del artículo no concuerdan con nuestra comprensión actual de las distribuciones de causas y desenlaces en el ser humano. - No se ha establecido que haya una causa única.

Tabla 2.
 Fuente: Sandler, Sandler, McDonnell y Wurzelmann (1992)

Si bien el artículo tiene apartados importantes que soportan la validez del resultado, hay aspectos, como la valoración de la exposición, que pueden sesgar el resultado y no dejan de manera tácita poder dar al paciente certeza sobre si la exposición al humo de cigarrillo tiene algo que ver con su actual enfermedad.



Instrucción

Lo invito a que revise la actividad práctica en los recursos de este eje para afianzar lo aprendido hasta ahora.



Estudio de caso

Un hombre adulto, quien es diagnosticado de nuevo con hipertensión sistémica, con prescripción de antihipertensivos manejados según presiones objetivo de control de la hipertensión, se encuentra en control para su enfermedad. Durante la consulta, se le recomienda realizar ejercicio cardiovascular entre moderado y alto, ya que esto impacta en su enfermedad; sin embargo, se pregunta el médico cuál será el pronóstico de la enfermedad bajo los efectos de este ejercicio cardiovascular recomendado sobre la mortalidad de origen cardiovascular y la mortalidad general.

Ante la situación anterior, se requiere revisar la evidencia disponible sobre el pronóstico de la hipertensión arterial sistémica con un factor específico como la realización de ejercicio cardiovascular.

La pregunta PICO que el médico tratante se plantea es:

- P: paciente adulto joven con diagnóstico de hipertensión arterial sistémica.
- I: ejercicio cardiovascular.
- C: no realización o niveles bajos de la misma.
- O: mortalidad general y de origen cardiovascular.

En los pacientes adultos jóvenes con diagnóstico de hipertensión arterial, ¿qué pronóstico tiene la hipertensión arterial en la mortalidad por cuenta de la realización de ejercicio cardiovascular?.

Buscando la evidencia

El médico selecciona las palabras "ejercicio cardiovascular", "fitness cardiovascular" e "hipertensión sistémica" en la página de PubMed, utilizando los términos de búsqueda. La búsqueda arroja 158 citas. Al limitar los estudios al idioma inglés y a los seres humanos, se acorta la lista a 80.

Revisa los resúmenes, buscando artículos que se centren en pacientes con hipertensión. Un artículo reporta los resultados sobre el efecto del *fitness* cardiovascular en hombres con hipertensión arterial sistémica (Church, Kampert, Gibbons, Barlow y Blair, 2001).

Guía para la lectura crítica de un artículo sobre pronóstico

Guía	Comentarios
<p>1. ¿Son válidos los resultados del estudio?</p> <p>Guías primarias</p> <p>¿Hubo una muestra de pacientes representativa y bien definida en un punto similar del curso de la enfermedad?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí, se toman tres grupos definidos en forma clara como hipertensos, no hipertensos y personas con presión alta sin hipertensión. Los hombres fueron predominantemente blancos no hispanicos, con buen nivel de educación y empleados o habían sido empleados en posiciones profesionales. Todos eran residentes de USA con un rango de edad entre 20 y 85 años. Se excluyeron hombres con historia de ECV, infarto o cáncer.</p> <p>La muestra la constituyeron 22167 hombres, aunque no se especificó tiempo de diagnóstico de hipertensión entre el grupo de hipertensos ni el grado o severidad de la hipertensión, por tanto, no se sabe si este grupo estaba en el mismo punto del curso clínico de la enfermedad.</p>
<p>¿El seguimiento fue suficientemente largo y completo?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí, fue lo suficientemente largo en el desenlace de mortalidad de cualquier causa o de origen cardiovascular, 224.173 años hombre, lo cual muestra seguimiento extenso (intervalo de seguimiento de 23 años). Se identificaron 628 eventos y, dado el sistema de vigilancia empleado y las características del desenlace, se considera que no hubo pérdidas.</p> <p>Sin embargo, los seguimientos fueron individuales, por ende, las personas que entraron al final del estudio tuvieron un tiempo de seguimiento mucho menor para la detección del desenlace.</p> <p>La mortalidad de toda causa no tiene problema para su medición y clasificación.</p> <p>La mortalidad cardiovascular no está claramente especificada ni se describieron los criterios objetivos, pero debemos asumir que, dado el sistema de vigilancia de mortalidad, es plausible que haya sido correctamente clasificada.</p>

Guía secundarias

¿Fueron usados resultados objetivos y no sesgados?

- Sí.
- No.
- No se puede decir.

Se encontró sesgo en la selección, lo cual hace que se pierda validez externa. Se trata de un grupo de pacientes circunscrito en un área de población estrecha para nuestra población y esto hace que otras variables a tener en cuenta sean importantes en el análisis de los datos de poblacionales en forma demográfica y social.

No se hace la descripción de cómo se recolecta ni por quién el desenlace en estas poblaciones (sesgo de sospecha diagnóstica o de expectativa), aunque se hace de manera objetiva (certificados de defunción).

¿Hubo ajustes por factores pronósticos relevantes?

- Sí.
- No.
- No se puede decir.

Sí. Se realiza ajuste para factores pronósticos (edad, hábitos, historia de tabaquismo, entre otros), controlando las variables alcohol, historia familiar de enfermedad cardiovascular, colesterol, glicemia en ayuno e índice de masa corporal y, muy importante, por cigarrillo, con la diferencia significativa en el hábito entre fallecidos y sobrevivientes en los tres grupos.

TABLE 3. Adjusted RRs for All-Cause and Cardiovascular Disease Mortality by Fitness Group in Normotensive Men, Men With Elevated Clinic Blood Pressure, and Hypertensive Men

Blood Pressure Groups*	Fitness Group	Men, Years (%)	All-Cause Mortality			Cardiovascular Disease Mortality		
			No. of Deaths	RR _{age} (95% CI)	RR _{all} (95% CI)	No. of Deaths	RR _{age} (95% CI)	RR _{all} (95% CI)
Normotensive	Low	26,912 (17)	88	1.0 (referent)	1.0 (referent)	29	1.0 (referent)	1.0 (referent)
	Mod.	66,296 (42)	144	0.68 (0.52, 0.89)	0.73 (0.56, 0.97)	32	0.45 (0.28, 0.77)	0.54 (0.31, 0.91)
	High	64,605 (41)	93	0.50 (0.37, 0.68)	0.57 (0.41, 0.80)	20	0.34 (0.19, 0.61)	0.44 (0.23, 0.84)
		157,812						
Elevated clinic blood pressure	Low	8,572 (25)	60	1.0 (referent)	1.0 (referent)	19	1.0 (referent)	1.0 (referent)
	Mod.	14,958 (45)	58	0.49 (0.34, 0.70)	0.59 (0.40, 0.88)	24	0.65 (0.35, 1.19)	0.91 (0.46, 1.80)
	High	10,139 (30)	34	0.44 (0.29, 0.68)	0.59 (0.36, 0.96)	8	0.34 (0.15, 0.79)	0.59 (0.23, 1.52)
		33,669						
Hypertensives	Low	9,399 (29)	70	1.0 (referent)	1.0 (referent)	24	1.0 (referent)	1.0 (referent)
	Mod.	14,017 (43)	51	0.45 (0.31, 0.65)	0.56 (0.38, 0.82)	21	0.53 (0.29, 0.96)	0.81 (0.42, 1.55)
	High	9,275 (28)	30	0.42 (0.27, 0.66)	0.57 (0.35, 0.95)	11	0.45 (0.22, 0.94)	0.91 (0.39, 2.10)
		32,691						

*Men with no history of hypertension and a normal clinic blood pressure (systolic BP <140 mm Hg and diastolic BP <90 mm Hg) were classified as normotensive. Men with an elevated clinic blood pressure (systolic BP ≥140 mm Hg or diastolic BP ≥90 mm Hg) but no previous diagnosis of high blood pressure were designated as elevated clinic blood pressure group. Men who reported a history of physician-diagnosed hypertension were classified as hypertensive.

RR_{age} = RR and 95% confidence intervals adjusted for age and year of examination; RR_{all} = RR and 95% confidence intervals adjusted for age, year of examination, alcohol use, smoking, and family history of CVD, total cholesterol, body mass index, and fasting glucose.

Figura 1.

Fuente: Church, Kampert, Gibbons, Barlow y Blair (2001)

2. ¿Cuáles son los resultados?

¿Cuán grande es la probabilidad del evento resultado en un periodo específico?

- Sí.
- No.
- No se puede decir.

Se observa una reducción en la probabilidad de muerte (riesgo) en los pacientes de todos los grupos con actividad física de moderada a alta para este factor protector. En la reducción de mortalidad cardiovascular algunos intervalos muestran que la reducción no es significativa, incluso para actividad física moderada y alta entre el grupo hipertensos RR_{all} = 0,91 IC 95 % (0,39-2.10), esto por los pocos eventos cardiovasculares en ese subgrupo y las múltiples comparaciones.

<p>¿Cuán precisa fue la estimación de dicha probabilidad de ocurrencia del evento resultado?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Son precisas por los intervalos de confianza estrechos que están brindando, a pesar de ser un grupo muy específico de una población. Sin embargo, la precisión evaluada con la amplitud del intervalo de confianza depende del número de eventos y, dado que se realizan comparaciones múltiples, el número de eventos por estrato es bajo, particularmente en cuanto a su relación con mortalidad cardiovascular, por lo cual vemos intervalos de confianza amplios que denotan poca precisión</p> <p>Según la figura, la reducción es significativa para todas las causas de muerte, ajustada por varios factores, sin embargo, la estimación de los intervalos es bastante amplia; por ejemplo, en el grupo normotenso la verdadera reducción poblacional en la categoría de alta intensidad estaría con un 95 % de confianza entre 20 %-59 %, y para hipertensos estaría entre 5 % y 65 %. Como se puede notar, estos efectos son bastante imprecisos, no dejando justificar si someter a un individuo hipertenso a rutinas de actividad física alta para obtener una reducción del riesgo de muerte por cualquier causa en tan solo un 5 %.</p>
<p>3. ¿Me ayudarán estos resultados en el cuidado de mis pacientes?</p> <p>¿Fueron los pacientes similares a mis pacientes?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>No son similares, dado que, como se menciona en el mismo estudio, las características sociodemográficas de estos pacientes son altamente seleccionadas (bien educados, blancos, hombres, empleados, profesionales). Esto hace que haya un posible sesgo de selección que entorpezca la generalización del estudio. Sin embargo, es tan apreciable el beneficio del ejercicio en este grupo de pacientes que se puede considerar la exploración de esta actividad de ejercicio como coadyuvante en la prevención de mortalidad y eventos cardiovasculares.</p>
<p>¿Los resultados conducirán directamente a seleccionar o evitar la terapia?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>La actividad física basada en el resultado del estudio favorece la reducción de muerte de cualquier causa e incluso la cardiovascular. Se podría recomendar en personas no hipertensas e hipertensas.</p>
<p>¿Los resultados son útiles para aconsejar a los pacientes?</p> <ul style="list-style-type: none"> • Sí. • No. • No se puede decir. 	<p>Sí son útiles en una población culta y profesional con un nivel de ingresos alto; sin embargo, para el paciente en malas condiciones, pobre y sin acceso a servicios médicos esta información aislada no es pertinente para definir la evolución de la patología.</p> <p>Es necesario tomarlo en contexto con el conjunto de la evidencia clínica y si consideramos válido el modelo de predicción, para lo cual primero debe validarse en otra muestra independiente de pacientes y considerando su plausibilidad biológica.</p>

Tabla 3.

Fuente: Church, Kampert, Gibbons, Barlow y Blair (2001)

En el análisis del artículo sobre pronóstico se puede afirmar con evidencia el efecto que tiene el ejercicio cardiovascular sobre la mortalidad general y la mortalidad específica por causa cardiovascular en pacientes hombres hipertensos, esto conllevaría que el profesional pueda recomendar con mayor seguridad la práctica de ejercicio en este hombre joven diagnosticado como hipertenso.



Lectura recomendada

Cómo valorar e interpretar un artículo sobre pronóstico

Carlos Manterola, Carmen Santander y Tamara Otzen



Instrucción

Para complementar esta información, revise el organizador gráfico dentro de los recursos de aprendizaje.

Como se puede notar, los tres escenarios clínicos presentados responden a situaciones de la práctica diaria. Con elementos básicos del análisis basado en la evidencia, pueden utilizarse los resultados científicos en la práctica habitual para una toma de decisiones más fundamentada en dichos resultados.

Church, T., Kampert, J., Gibbons, L., Barlow, C. y Blair, S. (2001). Usefulness of cardiorespiratory fitness as a predictor of all-cause and cardiovascular disease mortality in men with systemic hypertension. *American Journal of Cardiology*, 88(6), 651-656.

Maisel, A. et ál. (2002). Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *The New England Journal of Medicine*, 347(3), 161-167.

Sandler, R. S., Sandler, D. P., McDonnell, C. W. y Wurzelmann, J. I. (1992). Childhood exposure to environmental tobacco smoke and the risk of ulcerative colitis. *American Journal of Epidemiology*, 135(6), 603-608.

Esta obra se terminó de editar en el mes de Septiembre 2018
Tipografía BrownStd Light, 12 puntos
Bogotá D.C,-Colombia.



AREANDINA

Fundación Universitaria del Área Andina

MIEMBRO DE LA RED

ILUMNO