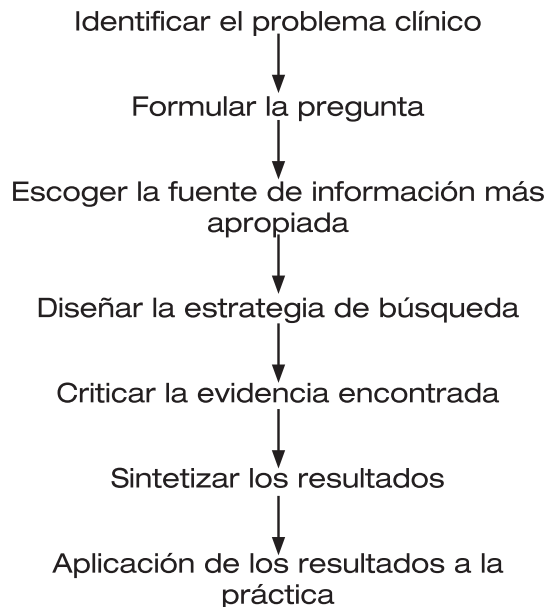


EDUCACIÓN MÉDICA CONTÍNUA

MEDICINA BASADA EN LAS EVIDENCIAS

Dr. Arturo Raúl Arévalo Barea*

LOS PASOS BASICOS PARA LA IMPLEMENTACION DE LA PRÁCTICA CLINICA BASADA EN LA EVIDENCIA



“En la situación en la que se tiene un artículo referido sobre una intervención preventiva o una intervención terapéutica se debe inicialmente valorar si se siguieron los pasos adecuados y se evitaron los sesgos que los anulen”

INTRODUCCIÓN

En este capítulo desarrollaremos guías útiles que nos permitan evaluar de forma crítica las investigaciones de una intervención terapéutica y otra de intervención preventiva. Iniciaremos revisando una situación clínica:

SITUACIÓN CLÍNICA

Paciente con embarazo de 32 semanas que acude a la consulta en emergencias de un hospital materno infantil, con inicio de contracciones uterinas y dolores en los flancos abdominales, se revisa a la paciente y se encuentra cervix uterino con 3 cm de dilatación y 1 cm de longitud. Se debe decidir si se le debe administrar

corticoide a la mujer embarazada para la maduración pulmonar fetal.

DEFINICIÓN DE LA PREGUNTA

1. Establecer claramente la pregunta que desea responder.
2. Realizar la búsqueda de la información requerida para solucionar la situación clínica.

La pregunta clínica se compone de por lo menos tres elementos:

1. Población
2. Intervención que se considera y con qué intervención se quiere comparar
3. Resultados que se desean evaluar

Para definir la pregunta clínica se puede seguir con el siguiente esquema:

ESTRUCTURA DE UNA PREGUNTA CLINICA	POBLACION	INTERVENCION	INTERVENCION A COMPARAR	RESULTADO
CONSEJO	¿Cómo podría describir al grupo de pacientes de interés? *Compensar brevedad con precisión	¿Cuál es la principal actuación que se quiere llevar a cabo? *Ser específico	¿Cuales son la otras opciones alternativas que existen? *Ser específico	¿ Que se espera que le suceda al paciente o población? *Ser específico
EJEMPLO	Mujeres con amenaza de parto pretérmino	Administración materna parenteral de corticoides	No administración de corticoides o administración de un placebo	Medición de la maduración pulmonar fetal

Adaptación al esquema de ayuda de Sackett DL et al .Medicina basada en la evidencia: Cómo ejercer y enseñar la MBE. Churchill-Livingstone, Madrid; 1997.

En esta situación la pregunta pertinente podría plantearse de la siguiente forma:

¿Qué beneficios y riesgos produce a la madre y al feto-neonato la administración materna parenteral de corticoides para la maduración pulmonar fetal?

Si desglosamos esta pregunta encontraremos los elementos necesarios para una buena elaboración de la misma, siendo:

POBLACION: Mujeres con amenaza de parto pretérmino

INTERVENCION:Administración materna parenteral de corticoides

RESULTADOS: Maduración pulmonar fetal medido por ejemplo como síndrome de distres respiratorio

REALIZAR LA BÚSQUDA SISTEMÁTICA DE LA LITERATURA

La búsqueda de la información realizada de acuerdo al capítulo previo dio un total de 125 investigaciones clínicas aleatorizadas (ECAs), sobre la administración de corticoides en mujeres con amenaza de parto pretérmino o en aquellas mujeres en las que se planificó la interrupción del embarazo previo a las 37 semanas de gestación.

TRES PREGUNTAS PRINCIPALES QUE SE DEBEN FORMULAR EN

EL PROCESO DE LA LECTURA CRÍTICA DE LOS ARTÍCULOS

1. ¿SON VÁLIDOS LOS RESULTADOS DE LA INVESTIGACIÓN?

Esta pregunta está relacionada con la validez de los resultados, y considera si el efecto del tratamiento que se informó en el artículo representa la verdadera dirección y magnitud del efecto de dicho tratamiento.

También se puede formular esta pregunta de la siguiente manera: ¿Los resultados representan una estimación **no sesgada** del efecto del tratamiento, o por el contrario, dichos resultados estuvieron desviados de manera sistemática lo que consiguientemente conduce a una conclusión falsa?

2. ¿CUÁLES SON LOS RESULTADOS DE LA INVESTIGACIÓN?

Si los resultados son válidos y proporcionan una evaluación no sesgada del efecto del tratamiento, entonces se justifica que los resultados sigan siendo examinados. Esta pregunta considera también la magnitud y la precisión del efecto del tratamiento. El mejor estimador del efecto del tratamiento serán los hallazgos del estudio en sí mismos; la precisión de dicho estimador será superior cuanto mayor sea el número de pacientes que sufren el resultado a estudiar.

3. ¿LOS RESULTADOS AYUDARÁN EN LA ATENCIÓN DE LAS PACIENTES?

La pregunta está estructurada en dos partes.

1. ¿Los resultados son aplicables a las pacientes?

1.1.- Se deberá dudar antes de iniciar el tratamiento si la paciente es muy distinta de aquellas que participaron en la investigación.

1.2.- Si el resultado sobre el que se demuestra mejoría no es importante para su paciente.

2. Si los resultados son aplicables, ¿cuál es el impacto neto del tratamiento?

El impacto depende de los beneficios y de los riesgos (efectos secundarios y efectos tóxicos) del tratamiento y las consecuencias de interrumpir el tratamiento. Por lo tanto, aún una terapia efectiva podría evitarse cuando el pronóstico de una paciente ya es bueno sin dicho tratamiento, especialmente cuando el tratamiento tiene efectos secundarios

adversos y/o efectos tóxicos.

Para poder interpretar de manera adecuada estos resultados es necesario definir la diferencia que existe entre eficacia y efectividad de una intervención.

EFICACIA: Es la magnitud en la que una intervención (tratamiento, procedimiento o servicio) mejora los resultados para los pacientes en condiciones ideales.

EFFECTIVIDAD: Es la magnitud en la que una intervención (tratamiento, procedimiento o servicio) mejora los resultados para los pacientes en una situación habitual.

¿LOS RESULTADOS DEL ESTUDIO SON VÁLIDOS?

¿La investigación responde a una pregunta claramente formulada?

El lector de una investigación debe poder determinar claramente la pregunta que se busca responder con dicho estudio. Para ello debe estar especificado en forma clara: qué tipo de población se estudiará, cuáles

ATRIBUTOS	EFICACIA	EFFECTIVIDAD
Criterio de elegibilidad de pacientes	Muy estricto, limitado por alto riesgo, alta responsabilidad y alto grado de obediencia	todos
Pacientes inelegibles	Son excluidos en el análisis del estudio	Usualmente están incluidos en el análisis
Tratamiento	Realizado de manera estrictamente monitorizado para dosis y respuesta	Cuidados de rutina
Intensidad del seguimiento	Fuerte, con frecuentes visitas	Baja, no mayor frecuencia que el seguimiento de rutina
Obediencia de los pacientes	Fuertemente monitorizados obediencia	Como en la práctica rutinaria
Adherencia de los clínicos al protocolo de estudio	Fuertemente controlados, con retro comunicación en caso de incompleta ejecución	No muy controlados
Eventos de interés	Restringidos a aquellos que contesten la pregunta clínica	Todos lo eventos nocivos
Duración del seguimiento	Se detiene al momento de obtener el evento de interés	Continúa hasta la muerte
Elegibilidad de los eventos para análisis	Restringidos	Todos los resultado, desde el comienzo hasta el final del ensayo

Fuente: Dave Sackett: *Clinical Epidemiology*, 3^o edition

serán las intervenciones a evaluar, qué resultados se medirán y la magnitud del efecto diferencial esperado entre las dos intervenciones.

La pregunta primaria es aquella que los investigadores tienen intención de responder y que debe ser respondida, es la que fija el **tamaño muestral**, necesario para llevar a cabo el estudio.

¿La asignación de los tratamientos fue aleatorizada?

La asignación de los tratamientos al azar es de fundamental importancia ya que de ella depende la comparabilidad de los diversos grupos de tratamiento.

Aleatorización: es el método que utiliza el azar para generar la secuencia de asignación de los individuos al grupo tratamiento o intervención; esta forma de manejo de pacientes evita potenciales sesgos de parte de los investigadores y/o participantes **generando grupos similares** en todas sus características excepto en la intervención que se desea evaluar.

Una correcta aleatorización, acompañada de un tamaño muestral adecuado, asegura que aquellos factores que influyen en el resultado, tanto los conocidos. De esta manera las diferencias en el resultado de los distintos grupos en comparación pueden ser adjudicadas sólo al efecto del tratamiento y no a la acción de alguna variable contundente que pueda haber influido en el resultado por estar distribuida en forma desigual entre los grupos.

Ciertas “sorpresas” generadas por investigaciones aleatorizadas que contradijeron los resultados de investigaciones menos rigurosas incluyen la demostración de que un principio químico pueden incrementar (más que reducir) la mortalidad de los pacientes con diagnóstico de sepsis, otra situación podría ser, que las inyecciones de esteroides no disminuyen el dolor articular, e incluso que la terapia hormonal sustitutiva para pacientes postmenopáusicas

no previene la enfermedad cardiovascular.

Estas “sorpresas” pueden ocurrir cuando los pacientes se asignan en forma aleatoria, en lugar de que sea por una decisión consciente de los médicos y los pacientes.

Así, los resultados clínicos ocurren por muchas causas, y el tratamiento es sólo una de ellas: la severidad subyacente de la enfermedad, la presencia de condiciones concomitantes, y una multitud de otras variables pronósticas (conocidas y desconocidas) tienen a menudo mayor influencia sobre el resultado que el efecto propio de la terapia.

Debido a que estos otros aspectos también influyen en las decisiones de los clínicos cuando ofrecen el tratamiento en cuestión, los estudios no aleatorizados sobre eficacia o efectividad, están inevitablemente limitados en su capacidad para distinguir lo útil de lo inútil y/o si dicha terapia es nociva.

Como confirmación de este hecho, sucede que los estudios en los que los tratamientos se asignan por cualquier otro método que no sea aleatorizado, tienden a mostrar efectos del tratamiento de mayor magnitud (y frecuentemente estos resultados son falso-positivos) que las investigaciones aleatorizadas.

Un buen método de aleatorización debe ser:

1. Objetivo,
2. Impredecible,
3. Enmascarado,
4. Invulnerable,
5. Reproducible

La ICA resulta ser el único diseño que posee esta cualidad, ya que asegura, si el tamaño muestral es suficientemente grande, que los determinantes de resultado conocidos e incluso los desconocidos, estén distribuidos equitativamente entre los grupos tratamiento y control.

¿SE ESPECIFICA EL MÉTODO QUE SE UTILIZÓ PARA GENERAR LA SECUENCIA DE ASIGNACIÓN DE TRATAMIENTOS?

El artículo de la investigación debe aclarar la forma en la que se generó la secuencia de asignación de tratamientos ya que sólo de esta forma se podrá afirmar que la asignación de los tratamientos se realizó aleatoriamente, es decir, por azar. Esta secuencia se refiere al orden en que se asignan los tratamientos y puede generarse a través de distintos métodos.

Algunos ejemplos de buenos métodos de aleatorización son:

- a) La tabla de números aleatorios siguiendo un proceso específico de selección de números,
- b) La generación computarizada de números aleatorios utilizando programas estadísticos como EPIINFO, otros.

Otros métodos para generar la secuencia, que son menos confiables por ser más factibles de manipulación, son:

- a) moneda,
- b) orden de consulta,
- c) fecha de nacimiento,
- d) número de historia clínica.

¿DICHA SECUENCIA FUE ENMASCARADA?

Es necesario que la secuencia de asignación generada aleatoriamente esté enmascarada para todo el personal involucrado en la investigación.

Si los médicos que ingresan los pacientes al estudio conocieran dicha secuencia, podrían manipular la asignación de pacientes a uno u otro tratamiento. Generalmente el clínico tiene una opinión formada sobre cuál es el mejor tratamiento, aún cuando no haya evidencia suficiente para tal afirmación, o a qué tipo de pacientes querría administrar uno de los tratamientos.

Una posibilidad es que desee aplicar el nuevo tratamiento a los

pacientes con formas más graves de enfermedad y administrar el tratamiento estándar a aquellas formas más leves; si las personas encargadas de la administración del tratamiento conocen el orden de asignación de los tratamientos podría dar lugar a un sesgo de selección, respetando la secuencia sólo en aquellos casos en que coincida con su opinión (el nuevo tratamiento para los pacientes más severos) y no ingresando los pacientes a los que querría administrar el otro tratamiento. La falta de comparabilidad de los grupos que se generarían de esta forma, invalidaría la conclusión del estudio ya que su resultado estaría influenciado por las diferentes condiciones pronosticas de los pacientes incluidos en cada grupo.

¿LA CANTIDAD DE PACIENTES INGRESADOS EN EL ESTUDIO A LOS CUÁLES NO SE LES PUDO EVALUAR EL RESULTADO FINAL, FUE IMPORTANTE?

Todo paciente que ingresa al estudio debe ser contabilizado y analizado en su conclusión. Si esto no se hace, o si un número substancial de pacientes se reportan como “pérdidas de seguimiento”, (aquellos pacientes en los cuales no se pudo evaluar el resultado a estudiar, la validez del estudio está sujeta a dudas. Cuanto mayor es el número de sujetos que se pierden, mayor es la probabilidad de que los resultados de la investigación se alejen de la verdad. Esto es debido a que los pacientes que se pierden tienen frecuentemente un pronóstico diferente al de aquéllos que permanecen en el estudio, y puede suceder que estos pacientes desaparezcan debido a que sufren resultados adversos (incluyendo la muerte) o porque gozan de buena salud (y por ello no regresan a la clínica para atenderse).

Los lectores pueden decidir ellos mismos, cuándo la pérdida de los pacientes en su seguimiento es excesiva. Para asumir dicha postura en las investigaciones positivas, se presume

que a todos los pacientes perdidos del grupo de tratamiento les fue mal, y que a todos los pacientes perdidos del grupo control les fue bien, y de esta manera, calcular los resultados del estudio bajo estos supuestos. Si las conclusiones de la investigación no cambian, entonces las pérdidas de seguimiento no fueron excesivas. Si las conclusiones se llegaran a modificar, mediante dicho método, la fuerza de la inferencia se debilitaría (es decir que los resultados del estudio merecen menos confianza). La medida en que la inferencia se debilita dependerá de cuán probable es que a los pacientes tratados perdidos, les haya ido mal a todos, al mismo tiempo que a los pacientes del grupo control perdidos les haya ido bien.

¿LOS PACIENTES SE ANALIZARON EN EL GRUPO DE TRATAMIENTO AL QUE FUERON ASIGNADOS?

Como sucede en la práctica diaria, los pacientes de las investigaciones clínicas aleatorizadas se olvidan algunas veces de tomar los medicamentos, rechazan su tratamiento o reciben el otro tratamiento en comparación. Los investigadores podrían estar de acuerdo, en forma intuitiva, en excluir del análisis de la eficacia del tratamiento a aquellos pacientes que nunca recibieron el tratamiento que se les asignó. Esta no es la manera correcta de analizar los resultados. Las razones para decir que esto no es correcto se basan en los siguientes hechos.

Las razones por las que las personas no toman sus medicamentos se relacionan frecuentemente con el pronóstico de la enfermedad. En una serie de investigaciones clínicas, a aquellos pacientes que no cumplieron el tratamiento les fue peor que a aquéllos que tomaron los medicamentos como se les indicó, aún teniendo en cuenta todos los factores pronósticos, y aún cuando sus medicamentos eran placebos. Excluyendo del análisis a los pacientes que no cumplieron se dejan de lado a aquéllos que pueden estar destinados a tener un resultado mejor

y se destruye la comparación sin sesgo que brinda la aleatorización.

La situación es similar en las terapias quirúrgicas. Algunos de los pacientes aleatorizados para cirugía nunca se someten a la operación porque están demasiado enfermos o porque sufren el evento a estudiar (como un ataque o un infarto de miocardio) antes de ingresar a la sala de operación. Si los investigadores incluyen estos pacientes, que están destinados a que les vaya mal, en el grupo control y no en el grupo de cirugía de una investigación, hasta una cirugía inútil puede aparentar ser efectiva. Sin embargo, la efectividad aparente no se deberá a un beneficio de los pacientes operados, sino a la exclusión sistemática de aquellos de peor pronóstico del grupo de cirugía.

Este principio de atribuir todos los pacientes al grupo al que fueron aleatorizados se denomina **“análisis según intención de tratamiento”**. Esta estrategia preserva el valor de la aleatorización: aquellos factores pronósticos que conocemos y aquellos que desconocemos, estarán distribuidos equitativamente en los dos grupos, y el efecto de las intervenciones en comparación se deberá exclusivamente al tratamiento asignado. Cabe agregar que además esta manera de analizar los pacientes nos brinda una ventaja adicional: nos dará un estimador del efecto acorde con la realidad de la aplicación de la intervención en la práctica clínica diaria.

Si la exclusión de un participante de la cohorte a la que fue asignada, se produce luego de la aleatorización se afecta la comparabilidad de los grupos y por lo tanto la validez interna del estudio.

Los desvíos del protocolo (no cumplimiento de la intervención asignada) se dan más frecuentemente cuando:

- Se ingresan pacientes que no cumplen los criterios de inclusión
- El paciente no acepta recibir el

tratamiento asignado

- El paciente no recibe el total de la dosis
- El paciente recibe otro tratamiento

Cuando en una investigación en donde por alguna situación las pacientes no continúan recibiendo el principio químico y por esta razón, no siguen siendo consideradas dentro de la investigación, esto indica que estas pacientes no se analizaron en el grupo al que pertenecían, lo que constituye un error metodológico. Debido a que no constará el grupo al que pertenecían cada una de estas pacientes y sus resultados, el análisis según intención de tratamiento no podrá realizarse.

¿EL TRATAMIENTO ASIGNADO ESTUVO ENMASCARADO PARA?

¿LOS PACIENTES?

Si el paciente tiene conocimiento del tratamiento que recibe, su opinión personal sobre dicho tratamiento, ya sea pesimista u optimista, tendrá un efecto directo en su recuperación. En diversos estudios se ha demostrado lo que se denomina “**efecto placebo**” indicando que aquellos pacientes que reciben un placebo tienen mejores resultados que aquellos que no reciben ningún tratamiento. Por ello es muy importante que los pacientes desconozcan a qué tratamiento fueron asignados para que no se produzca un sesgo que distorsione el verdadero efecto del tratamiento.

Cuando sea posible, los tratamientos experimentales deben ser administrados de la misma forma que el tratamiento estándar para que el paciente no pueda detectar cuál de ellos está recibiendo.

Ejemplo: iguales tabletas o cápsulas de igual color y sabor para que los pacientes no sospechen a cual intervención fueron asignadas.

¿LOS ENCARGADOS DEL TRATAMIENTO DE LOS PACIENTES?

Si el clínico encargado del tratamiento de los pacientes tiene conocimiento de la intervención que está administrando

puede llegar a modificar, aún inconscientemente, su conducta hacia los pacientes según al grupo que pertenezcan. Esto podría significar que a un grupo lo controle más que a otro o que administre **co-intervenciones** diferentes (intervenciones adicionales a la que está en comparación) a los distintos grupos de tratamiento. Aquí se observa cómo la opinión subjetiva de los clínicos puede sesgar los resultados cuando no se realiza un enmascaramiento apropiado.

¿LOS ENCARGADOS DE LA EVALUACIÓN DE LOS RESULTADOS?

El personal del estudio que mide los resultados en forma no enmascarada, puede proporcionar diferentes interpretaciones de hallazgos marginales o diferentes estímulos a los pacientes en la ejecución de las pruebas de medición de resultados, pudiendo de alguna de las dos maneras, distorsionar los resultados. Por ello es también muy importante que desconozcan el grupo al que pertenecen los pacientes.

Cuando usted lee informes de tratamientos (como investigaciones sobre cirugías) en los que los pacientes y los clínicos no pueden mantenerse enmascarados, usted debe fijarse si los investigadores minimizaron el sesgo, enmascarando a aquéllos que evaluaban los resultados clínicos.

En el caso en que el tratamiento no pueda enmascarse, por lo que las diferentes personas involucradas en el estudio conocen las intervenciones que están siendo administradas, la medida de resultado final debe ser lo suficientemente poderosa (Ej.: muerte) como para que la opinión subjetiva de la persona que realiza la medición no pueda influir sistemáticamente en los resultados.

¿LOS GRUPOS ERAN SIMILARES AL COMIENZO DE LA INVESTIGACIÓN?

Para asegurarse el grado de validez del estudio, los lectores deberían estar informados si los grupos tratamiento

y control eran similares en todos los factores que determinan los resultados clínicos de interés salvo en uno: si recibían la terapia experimental. Los investigadores proporcionan esta información cuando muestran en una **tabla los factores pronósticos basales** de los pacientes de los grupos tratamiento y control, que generalmente son evaluados previos al ingreso de los pacientes en el estudio. A pesar que nunca sabremos si la similitud existe para los factores pronósticos desconocidos, tendremos confianza que ello ocurre cuando los factores pronósticos conocidos están bien balanceados.

La aleatorización no siempre produce grupos balanceados para los factores pronósticos conocidos. Cuando el número de pacientes en los grupos es pequeño, por azar puede ocurrir que los pacientes de mejor pronóstico se ubiquen en uno de los grupos. Esto se denomina desbalance de factores pronósticos. Cuando el tamaño muestral aumenta, dicha posibilidad disminuye (esto es análogo a arrojar la moneda muchas veces: uno no debería sorprenderse en ver siete caras en diez tiradas de moneda, pero debería sorprenderse muchísimo si ve 70 caras en 100 tiradas de moneda)

El tema aquí no es ver si hay diferencias estadísticamente significativas en los factores pronósticos conocidos entre los grupos en comparación (en una investigación clínica aleatorizada uno sabe de antemano que cualquier diferencia que ocurra en los factores pronósticos sucede por azar), sino en la magnitud de estas diferencias. Si estas diferencias son clínicamente relevantes como para existir la posibilidad de que los resultados se vean influenciados por este desbalance, la validez del estudio puede comprometerse. Cuanto más fuerte es la relación entre los factores pronósticos y el resultado, y más pequeña la investigación, la diferencia entre los grupos debilitará aún más cualquier inferencia sobre eficacia.

No está todo perdido si los grupos

de tratamiento no son similares en los factores pronósticos basales. Existen técnicas estadísticas que permiten ajustar los resultados del estudio para dichas diferencias. Por lo tanto, los lectores deberían buscar información sobre la similitud de las características basales relevantes y, si existen diferencias clínicamente substanciales, deberían fijarse si los investigadores realizaron un análisis que ajuste aquellas diferencias. Cuando los análisis ajustados y no ajustados arriban a las mismas conclusiones, se justifica que los lectores ganen confianza en la validez del resultado del estudio.

¿LOS GRUPOS FUERON TRATADOS DE LA MISMA FORMA SALVO EN LA INTERVENCIÓN EXPERIMENTAL?

La atención médica en los grupos experimentales y de control se puede diferenciar en una serie de aspectos distintos además de la terapia puesta a prueba, y estas diferencias en el cuidado, aparte de aquél en estudio, pueden debilitar o distorsionar los resultados. Si un grupo recibe un seguimiento más minucioso, podría resultar en una mayor probabilidad de reportar ciertos eventos, y los pacientes podrían recibir tratamientos adicionales fuera de los puestos a prueba en el estudio.

Por ejemplo, en investigaciones sobre nuevas formas de terapia para artritis reumatoide resistente, el tratamiento auxiliar con esteroides sistémicos (extremadamente efectivo para aliviar los síntomas), si se administraba más frecuentemente en el grupo control que en el grupo de tratamiento, puede obscurecer un efecto de tratamiento verdadero de la droga experimental (a menos que la exacerbación que requería la administración de esteroides fuera considerada en sí misma como un resultado)

Las otras intervenciones distintas del tratamiento en estudio, cuando se aplican en forma diferente entre los grupos control y tratamiento, se llaman frecuentemente "cointervenciones". La cointervención es un problema más

serio cuando el enmascaramiento para el paciente y el personal encargado de la atención no está presente, o cuando el uso de tratamientos muy efectivos aparte de los estudiados se permiten a discreción de los médicos. Los lectores tienen más confianza en los resultados de las investigaciones cuando las cointervenciones permitidas se describen en la sección de “Métodos” y se informan como infrecuentes en los resultados.

“Durante la evaluación de un artículo el lector siempre debe tener en cuenta si los investigadores evitaron los sesgos utilizando las estrategias apropiadas”:

1. **SESGO DE SELECCIÓN Y EL SESGO DE CONFUSIÓN:** se evitan mediante la aleatorización.
2. **SESGO DE OBSERVACIÓN Y SESGO DE INFORMACIÓN:** se evita con el enmascaramiento.
3. **SESGO DE EVALUACIÓN DE LA MEDIDA DE RESULTADOS:** se evita con el enmascarando y determinando puntos finales fuertes.
4. **Sesgo de Análisis de Datos:** se evita con el análisis según intención de tratamiento y análisis de todos los participantes

Las siete preguntas anteriores (las cinco primeras serían la más importantes ya que las dos últimas podrían resultar una consecuencia de las anteriores), aplicadas secuencialmente, ayudarán al lector a determinar si los resultados de un artículo sobre terapia son probablemente válidos. Si los resultados son válidos, entonces el lector puede proceder a considerar la magnitud del efecto y la aplicabilidad a sus pacientes.

¿CUÁLES FUERON LOS RESULTADOS?

¿CUÁN GRANDE FUE EL EFECTO DEL TRATAMIENTO?

Es sumamente frecuente que las ICAs monitoreen cuidadosamente la frecuencia con la que los pacientes experimentan algunos eventos o efectos

adversos.

Ejemplos de estos resultados dicotómicos (resultados sí o no, que ocurren o no ocurren) incluyen la recidiva de cáncer, infarto de miocardio, y muerte. Los pacientes pueden padecer o no un evento, y el artículo informa la proporción de pacientes que desarrollan estos eventos.

Los resultados de una intervención o tratamiento se miden mediante el cálculo de los riesgos.

RIESGO: Probabilidad de ocurrencia de un evento dado. Se estima por la razón entre los pacientes con un evento dentro de un grupo, con el total de pacientes de ese grupo. Si en un grupo de 100 individuos se observan 15 eventos, el riesgo estimado es 0,15 o 15%.

Una forma de expresar el efecto del tratamiento es utilizando el **RIESGO RELATIVO (RR)**. El RR se define como la razón entre el riesgo de presentar el evento en el grupo experimental y el riesgo de presentar el evento en el grupo control.

$$RR = Re/Rb \text{ o } Rc$$

Dicho en otras palabras estima la magnitud de la asociación entre la exposición y el resultado e indica la probabilidad de desarrollar el evento en el grupo expuesto con relación al grupo no expuesto.

Ejemplo: un estudio informa que el riesgo de morir en el grupo que no recibe el tratamiento (grupo control) es del 20% (0.20) y el riesgo de morir en el grupo de pacientes que recibe el tratamiento es del 15% (0.15).

¿Cómo puede expresarse el efecto del tratamiento en base a estos resultados?
¿Cuál es el RR?:

$$RR = Re/Rc = 0.15/0.20 = 0.75.$$

Cuando **RR < 1** el riesgo o la probabilidad de sufrir el evento con la exposición, es menor que el riesgo de sufrir el mismo evento sin la exposición. Indica que la intervención fue efectiva en reducir el riesgo de tal resultado.

Cuando **RR = 1** el riesgo de sufrir el

evento con o sin exposición es el mismo, no varía, indica la no diferencia entre los grupos que se están comparando.

Cuando **RR>1** el riesgo de sufrir el evento con la nueva exposición es mayor, por lo tanto estaríamos frente a un efecto perjudicial de la intervención.

Otra manera de expresar el efecto de la intervención es utilizando la **REDUCCIÓN DEL RIESGO RELATIVO**

(**RRR**), que es el complemento del **RR**.

La **RRR** se expresa como un porcentaje:

$$[1 - RR] \times 100\% = [1 - 0.75] \times 100\% = 25\%$$

Una **RRR** de 25% significa que el nuevo tratamiento reduce el riesgo de muerte en un 25% relativo a aquél que ocurre entre los pacientes del grupo control; cuanto mayor es la **RRR**, más efectiva es la terapia.

Introducción de algunas medidas de los efectos de una terapia

Riesgo sin la terapia (riesgo basal o riesgo control) Rb	Riesgo con la terapia: Rt	Reducción del riesgo absoluto (RRA) RRA= Rb-Rt	Riesgo relativo (RR) RR= Rt/Rb	Reducción del riesgo relativo (RRR): RRR= 1-RR x100% ó RRR=RRA/Rb x 100%	Intervalo de confianza del 95% para la RRR:
20/100=0.20 ó 20%	15/100=0.15 ó 15%	0.20-0.15 = 0.05	0.15/0.20= 0.75	[1-0.75] x100%=25% [0.05/0.20]x 100%=25%	-38% a +59%

¿CUÁN PRECISA FUE LA ESTIMACIÓN DEL EFECTO DEL TRATAMIENTO?

La verdadera reducción del riesgo nunca puede conocerse; todo lo que tenemos es la estimación proporcionada por investigaciones clínicas aleatorizadas, y la mejor estimación del verdadero efecto del tratamiento es aquella observada en la investigación. Esta estimación se llama “estimación puntual” con el objeto de recordarnos que a pesar de que el verdadero valor se encuentra en algún lugar alrededor de ella, es improbable que sea el “verdadero valor” en forma precisa. Mediante la estrategia estadística del cálculo de los **INTERVALOS DE CONFIANZA** (ICs) los investigadores nos informan los límites de confianza entre los que el verdadero valor probablemente se encuentra, por lo tanto corresponde al rango en el cual probablemente se encuentre con cierto grado de confianza la verdadera magnitud del efecto.

Usualmente (aunque en forma arbitraria) usamos el **IC** del 95%, que puede interpretarse simplemente definiéndolo como el rango que incluye el verdadero valor del **RR** el 95% de las veces. El verdadero valor del RR se encontrará fuera de los extremos del **IC** solamente el 5% de las veces.

Si en una investigación se **aleatorizan** 100 pacientes en los grupos de control y de tratamiento, y ocurren 20 muertes en el grupo control y 15 en el grupo tratamiento, los investigadores podrían calcular una estimación puntual para el **RR** de 0,75 (siendo Rb=20/100 ó 0.20, Re=15/100 ó 0.15) y para el **RRR** de 25%. Sin embargo, usted podría suponer que el verdadero **RR** podría ser más pequeño o más grande que este 0,75, basado en una diferencia de solamente 5 muertes. En realidad, usted puede suponer que el tratamiento podría no proporcionar **ningún beneficio (una RR de 1)** o aún que resulte **perjudicial (una RR mayor que 1)**.

De hecho se puede estar en lo cierto, estos resultados son consistentes con una **RRR** de -0.38 (es decir, que los pacientes a los que se le administró el nuevo tratamiento podrían tener un 38% más de probabilidad de morir que aquellos pacientes de control), y una **RRR** de casi 59% (es decir, que los pacientes que han recibido el nuevo tratamiento podrían tener un riesgo de morir casi un 60% menor que el riesgo de aquellos que no han sido tratados).

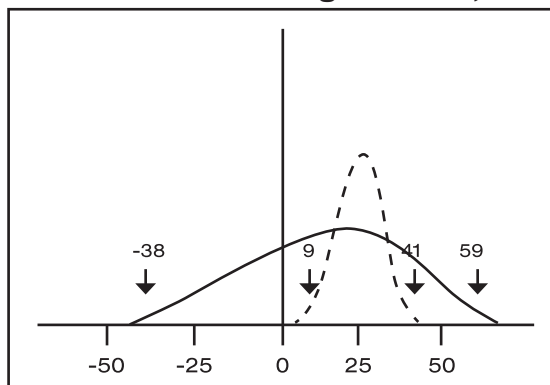
Se ve claramente en este ejemplo que el resultado es tan ambiguo que nos dice que el nuevo tratamiento puede

estar perjudicando o beneficiando al paciente.

En otras palabras, el **IC** del 95% de esta **RRR** es de -38% a +59%, y la investigación realmente no nos ha ayudado a decidir si brindamos o no el nuevo tratamiento.

¿QUÉ CLASE DE ESTUDIO SERÁ MÁS ÚTIL?

Reducción del riesgo relativo, %



Qué pasa si la investigación enroló, en lugar de 100 pacientes por grupo, 1000 pacientes por grupo, y se observó la misma razón de eventos que antes, por lo tanto hubo 200 muertes en el grupo control ($R_b = 200/1000 = 0.20$) y 150 muertes en el grupo experimental ($R_t = 150/1000 = 0.15$). Nuevamente la estimación puntual de la **RRR** es de 25%.

En esta investigación más grande, usted podría pensar que la verdadera reducción en el riesgo está más cercana a 25%, y nuevamente estaría en lo cierto; el **IC** del 95% de la **RRR** para estos resultados está todo en el lado positivo del 0 y va desde 9% a 41%.

La línea entera representa el **IC** alrededor del primer ejemplo en el que había 100 pacientes por grupo y en el que el número de eventos en el grupo de tratamiento y de control fueron dos y cuatro respectivamente.

La línea discontinua representa el intervalo de confianza alrededor del segundo ejemplo en el que había 1000 pacientes por grupo y el número de pacientes en el grupo tratamiento y control fueron 20 y 40 respectivamente.

Lo que estos ejemplos muestran es que para tamaños muestrales mayores, en general, pero no siempre, mayor es el número de eventos de resultado, y por consiguiente mayor nuestra confianza de que el verdadero valor de la **RRR** (o cualquier otra medida de la eficacia) esté cerca de aquél que hemos observado.

En el segundo ejemplo, el menor valor posible de la **RRR** fue 9% y el valor más alto fue 41%. La estimación puntual (en este caso 25%) es el valor que representa con mayor probabilidad la verdadera **RRR**.

La Figura representa los ICs alrededor de la estimación puntual de una **RRR** de 25% en estos dos ejemplos, donde una reducción del riesgo igual a 0 representa que no hay efecto tratamiento.

En ambas situaciones la estimación puntual de la **RRR** es de 25%, pero el IC es mucho más estrecho en la segunda situación.

El **IC** nos da pautas sobre:

- Qué efectos tiene el tamaño muestral en los resultados obtenidos. con tamaños muestrales pequeños, mayor o más amplio es el IC y mayor es la variabilidad en los resultados encontrados.
- Cuando el tamaño muestral es adecuado se estrecha el IC y más precisa es la estimación del efecto del tratamiento.

Es evidente que cuanto mayor el número de eventos de resultado observado, más estrecho es el **IC**. ¿Cuándo el tamaño muestral es suficientemente grande? En un estudio "positivo" -un estudio en el que los investigadores concluyen que el tratamiento es efectivo- uno puede mirar el límite inferior del **IC**.

En el segundo ejemplo, el límite inferior fue 9%. Si esta reducción del riesgo (la menor que es consistente con el resultado del estudio) es aún importante, o "clínicamente significativa", (es decir, que la magnitud del efecto es lo suficientemente grande como para ofrecerlo a su paciente), entonces los investigadores han enrolado suficientes

pacientes. Si, por otro lado, usted no considera que una RRR del 9% sea clínicamente significativa, entonces el estudio no puede considerarse como definitivo, aún si sus resultados son estadísticamente significativos (es decir, que ellos excluyen una reducción del riesgo del 0%). Recuerde, que la probabilidad de que el verdadero valor sea menor que el límite inferior del IC es de 2.5% y que un criterio diferente para el IC (por ejemplo un IC del 90%) podría ser más apropiado.

El IC también nos ayuda a interpretar estudios “negativos” en los que los autores concluyeron que el tratamiento experimental no es mejor que la terapia control. Todo lo que necesitamos es mirar el límite superior del IC. Si la RRR en este límite superior podría, si fuera cierta, ser clínicamente importante, el estudio no ha podido excluir un efecto de tratamiento importante. En el primer ejemplo que presentamos en esta sección, el límite superior del IC fue una RRR del 59%. Claramente, si esto representara la verdad, el beneficio del tratamiento sería substancial, y concluiríamos que aunque los investigadores no pudieron probar que el tratamiento experimental era mejor que el placebo, ellos tampoco pudieron probar que no lo era; ellos no pudieron excluir un efecto de tratamiento positivo grande. Una vez más el clínico debe tener en cuenta la arbitrariedad de la elección de los límites del 95% para el IC. Una alternativa razonable, un IC del 90%, podría ser algo más estrecho.

¿QUÉ PUEDE HACER EL CLÍNICO SI EL IC DE LA RRR NO SE EXHIBE EN EL ARTÍCULO?

HAY TRES PROPUESTAS, Y LAS PRESENTAMOS EN ORDEN A SU COMPLEJIDAD:

LA PROPUESTA MÁS SENCILLA ES EXAMINAR EL VALOR DE P.

El valor P o significancia estadística; se define como la probabilidad de que un efecto tan extremo como el que se observó en el ejemplo anterior, podría haber ocurrido sólo por azar y por lo

tanto no habría una relación entre la exposición y el resultado.

Si el valor de P es exactamente 0.05, entonces el límite inferior del IC del 95% para la RRR tiene que caer exactamente en el 0 (un RR de 1) y usted no puede excluir la posibilidad de que el tratamiento no tiene efecto.

A medida que el valor de P decrece por debajo del 0.05, el límite inferior del IC del 95% para la RRR **se eleva por encima del 0**, estaríamos frente a **una asociación estadísticamente significativa**, diríamos entonces que existe al menos un 5% o 1 en 20 la posibilidad de observar una asociación de la magnitud del efecto del estudio únicamente por azar

SEGUNDA PROPUESTA, envuelve algunos cálculos aritméticos mentales rápidos o un lápiz y un papel, puede usarse cuando el artículo incluye el valor del error standard (ES) de la RRR (o del RR). Esto se debe, a que los límites superior e inferior del intervalo de confianza del 95% para una RRR es la estimación puntual más y menos dos veces el ES.

TERCERA PROPUESTA involucra el cálculo del IC en sí mismo, o solicitar la ayuda de alguien más (por ejemplo un estadístico) para hacerlo. Una vez que usted obtiene los ICs, usted conoce cuán alta o baja puede ser la RRR (es decir, usted conoce la precisión de la estimación del efecto del tratamiento) y puede interpretar los resultados como se describe arriba.

No todas las investigaciones aleatorizadas tienen resultados dicotómicos, ni deberían. Por ejemplo, un nuevo tratamiento para pacientes con enfermedad pulmonar crónica puede enfocarse en el incremento de su capacidad de ejercicio. Por lo tanto, en un estudio de entrenamiento del músculo respiratorio para pacientes con limitaciones crónicas del flujo respiratorio, una medida primaria del resultado podría ser la distancia recorrida en 6 minutos en un corredor cerrado. Estos 6 minutos de caminata

mejoran en un promedio de 406 a 416 metros (10 metros arriba) en el grupo experimental que recibe entrenamiento del músculo respiratorio, y de 409 a 429 (20 metros arriba) en el grupo control. Por lo tanto, la estimación puntual de la mejoría en los 6 minutos de caminata debida al entrenamiento del músculo respiratorio, fue negativa, de -10 metros (ó una diferencia de 10 metros a favor del grupo control).

Aquí también usted debería mirar el IC del 95% alrededor de esta diferencia en los cambios de la capacidad de ejercicio y considerar sus consecuencias. Los investigadores nos dicen que el límite inferior fue -26 metros (es decir, que los resultados son consistentes con la diferencia de 26 metros a favor del grupo del tratamiento control) y el límite superior fue de +5 metros. Aún en las mejores circunstancias, agregar 5 metros a los 400 registrados al comienzo de la investigación, no serían importantes para el paciente, y este resultado excluye efectivamente un beneficio clínicamente significativo del entrenamiento del músculo respiratorio como se lo aplicó en este estudio.

Cuando se ha determinado la magnitud y la precisión del efecto del tratamiento, los lectores pueden pasar a la pregunta final de cómo aplicar los resultados del artículo a sus pacientes y a su práctica clínica.

¿LOS RESULTADOS ME AYUDARÁN EN LA ATENCIÓN DE MIS PACIENTES?

¿LOS RESULTADOS SON APLICABLES A MIS PACIENTES?

El primer aspecto a tratar es, cuánta confianza tiene usted en que pueda aplicar los resultados a un paciente en particular o a los pacientes de su práctica. Si los pacientes hubieran sido enrolados en el estudio y hubieran estado allí - es decir, si cumplen con todos los criterios de inclusión, y no violan ningún criterio de exclusión - hay pocas dudas de que los resultados sean aplicables. Si este no es el caso, y ellos no hubiesen sido seleccionables para

el estudio, se requiere una evaluación. Probablemente, el resultado del estudio es aplicable aún si, por ejemplo, ellos tuvieron 2 años de más para el estudio, tuvieron una enfermedad más severa, fueron tratados previamente con una terapia competente, o tuvieron una condición comórbida.

Una propuesta mejor, que aplicar rígidamente los criterios de inclusión y de exclusión del estudio, es preguntarse si existe una razón importante por la cual los resultados no deberían aplicarse a sus pacientes. Usualmente no se encontrará una razón restrictiva, y frecuentemente usted podrá generalizar con relativa confianza los resultados a sus pacientes.

Un aspecto final surge cuando nuestros pacientes son compatibles con las características de un subgrupo de pacientes del informe de la investigación.

En los artículos que informan los resultados de una investigación (especialmente cuando el tratamiento no parece ser eficaz para el paciente medio), los autores pueden haber examinado un largo número de subgrupos de pacientes en diferentes estados de su enfermedad, con diferentes condiciones comórbidas, con diferentes edades en el ingreso, etc. Muy a menudo estos **análisis de subgrupos** no fueron planeados con anterioridad, y los datos son simplemente “escarbados” para ver qué podría aparecer. Los investigadores a veces sobreinterpretan estos análisis “dato-dependientes” como demostración de que el tratamiento realmente tiene un efecto diferente en un subgrupo de pacientes, aquellos que son más viejos o más enfermos, por ejemplo, pueden señalarse como beneficiados substancialmente más o menos que otros subgrupos de pacientes de la investigación.

Usted puede encontrar guías para decidir si creer o no en estos análisis de subgrupos, resumidos de esta forma: es realmente probable que el

tratamiento beneficie al subgrupo más o menos que a los otros pacientes si la diferencia de los efectos del tratamiento en los subgrupos

- (1) es grande;
- (2) es muy improbable que ocurra por azar;
- (3) resulta de un análisis especificado como una hipótesis antes de que el estudio haya comenzado;
- (4) fue uno de los pocos análisis de subgrupos que fueron realizados; y
- (5) se repite en otros estudios.

En la medida que el análisis de subgrupos no cumple con estos criterios, los clínicos deben ser cada vez más escépticos para aplicarlos a sus pacientes

En la medida que el análisis de subgrupos no cumple con estos criterios, los clínicos deben ser cada vez más escépticos para aplicarlos a sus pacientes.

¿SE CONSIDERARON TODOS LOS RESULTADOS CLÍNICAMENTE IMPORTANTES?

Los tratamientos se ponen en práctica cuando proporcionan beneficios importantes. Demostrando que un broncodilatador produce pequeños incrementos en el volumen expirado forzado en pacientes con limitaciones crónicas del flujo respiratorio, que un vasodilatador mejora el gasto cardíaco en pacientes con fallas cardíacas, o que un agente para bajar los lípidos mejora los perfiles lípidos, no son razones suficientes para administrar estas drogas. Lo que se requiere es evidencia de que los tratamientos mejoran resultados que son importantes para los pacientes, como reducir la disnea durante las actividades que se requieren para la vida diaria, evitar la hospitalización por fallas cardíacas, o disminuir el riesgo de infarto de miocardio.

Los autores han sustituido los resultados importantes (disnea, hospitalización, o infarto de miocardio) por medidas fisiológicas no tan estrictas, generalmente para confirmar los beneficios (en “puntos finales

importantes”). Probablemente si se analizaran los puntos finales importantes deberían haberse enrolado muchos más pacientes y seguirlos por períodos de tiempo mucho más prolongados.

Un dramático ejemplo reciente del peligro de sustituir puntos finales, consistió en la evaluación de la utilidad de las drogas antiarrítmicas después del infarto de miocardio. Debido a que estas drogas habían mostrado que reducían la despolarización ventricular anormal (el punto final sustituto) a corto plazo, tenía sentido pensar que ellas debían reducir la ocurrencia de arritmias con amenaza de muerte a largo plazo.

Un grupo de investigadores realizó investigaciones aleatorizadas sobre tres agentes (encainida, flecainida, y morcizina) que previamente habían demostrado que eran efectivos para suprimir el punto final sustituto de despolarizaciones ventriculares anormales, para determinar si reducían la mortalidad de pacientes con arritmias asintomáticas o poco sintomáticas después de un infarto de miocardio. Los investigadores tuvieron que suspender las investigaciones cuando ellos descubrieron que la mortalidad era substancialmente mayor en aquellos que recibían tratamiento antiarrítmico que en aquellos que recibían placebo. Los clínicos creyendo en el punto final sustituto de la supresión de la arritmia hubieran continuado administrando las tres drogas, con el considerable perjuicio para sus pacientes.

Aún cuando los investigadores informan sobre efectos favorables de tratamiento en un resultado clínicamente importante, los clínicos deben tener cuidado de que no haya efectos deletéreos en otros resultados. Por ejemplo, cuando esta serie estaba en preparación, continuaba la controversia sobre si reducir los lípidos incrementa inesperadamente las causas de muerte no cardiovasculares. La quimioterapia para el cáncer puede alargar el período de vida, pero puede también disminuir su calidad. Finalmente, la

investigaciones quirúrgicas documentan frecuentemente la prolongación de la vida para aquellos que sobreviven a la operación (mostrando 3 años más de supervivencia en aquellos a los que se les efectuó la cirugía), pero con un riesgo inmediato de morir durante o

poco después de la operación. Por lo tanto, los usuarios de los informes de las investigaciones quirúrgicas deberían buscar información sobre la mortalidad temprana e inmediata (típicamente más alta en el grupo de cirugía), además de los resultados de la prolongación del término de vida.

Dos hombres con pronósticos contrastantes después de un infarto de miocardio

Si el riesgo de muerte en un año sin terapia (riesgo basal) es: X	Y el riesgo relativo de muerte con terapia (un beta-bloq.) es: Y/X	Y la reducción del riesgo relativo es: [1-(Y/X)]x100% ó [(X-Y)/X]x100%	Entonces el riesgo de muerte con el tratamiento es: Y	Y la reducción del riesgo absoluto es: X-Y	Y el número necesario a tratar para prevenir un evento es: 1/(X-Y)
1% ó 0.01 10% ó 0.1	75% ó 0.75 75% ó 0.75	25% 25%	0.01 x 0.75= 0.0075 0.10 x 0.75= 0.075	0.01-0.0075= 0.0025 0.10-0.075= 0.025	1/0.0025= 400 1/0.025= 40

Incorporación de los efectos secundarios al número necesario a tratar

Si el riesgo de muerte en un año sin terapia (riesgo basal) es: X	Y el riesgo de muerte con propanolol es: Y	Entonces la reducción del riesgo absoluto es: X-Y	Y el número necesario a tratar para prevenir un evento es: 1/(X-Y)	Y la incidencia de fatiga clínicamente importante con propanolol es:	Entonces el número de pacientes fatigados por cada vida salvada es:
1% ó 0.01 10% ó 0.10	0.01x 0.75= 0.0075 0.10 x 0.75= 0.075	0.01-0.0075= 0.0025 0.10-0.075= 0.025	1/0.0025= 400 1/0.025= 40	10% ó 0.10	400x 0.1= 40 40 x 0.10= 4

¿LOS BENEFICIOS PROBABLES DEL TRATAMIENTO JUSTIFICAN LOS DAÑOS POTENCIALES Y LOS COSTOS?

Una forma de presentar los efectos de una terapia, podría ser la diferencia absoluta, conocida como la **Reducción del riesgo absoluto** o **diferencia de riesgo**, que se expresa como la diferencia de riesgo entre la proporción de los que murieron en el grupo control (Rc) y la proporción de los que murieron en el grupo experimental (Re), ó $Rc-Re = 20\%-15\%= 5\%$.

Si los resultados del artículo son generalizables a su paciente y sus resultados son importantes, la próxima pregunta se refiere a si es probable que los beneficios del tratamiento justifiquen el esfuerzo que usted y su paciente deben poner en la empresa. Una reducción del 25% en el riesgo de

muerte puede sonar muy impresionante, pero el impacto en su paciente y en la práctica nunca debe minimizarse. Esta noción se ilustra utilizando un concepto llamado "**número necesario a tratar**" (NNT).

$$NNT = 1/(Rb-Rt) \text{ o } NNT = 1/ RRA$$

El impacto de un tratamiento no está relacionado solamente con la RRR, sino también con el riesgo del resultado adverso que está diseñado para prevenir. Los beta-bloqueantes reducen el riesgo de muerte después de un infarto de miocardio en un 25% aproximadamente, y esta RRR es consistente a través de los grupos, incluyendo a aquellos con riesgos "basales" de recurrencia y muerte más altos y más bajos cuando no son tratados.

En dos pacientes con infartos de miocardio recientes se considera:

Primero consideremos un hombre de 40 años con un infarto pequeño, capacidad de ejercicio normal, y sin signos de arritmia ventricular, que está dispuesto a dejar de fumar, comenzar ejercicios, perder peso, y tomar aspirina diariamente. Su riesgo de muerte individual en el primer año después del infarto puede ser tan bajo como el 1%. Los beta-bloqueantes reducirían este riesgo en un cuarto, a 0.75%, para una reducción del riesgo absoluto de un 0,25% o de 0.0025. La inversa de esta reducción del riesgo absoluto (es decir, 1 dividido por la reducción del riesgo absoluto) es igual al número pacientes de estos que deberíamos tratar con el objeto de prevenir un evento (en este caso, para prevenir una muerte después de un ataque al corazón suave en un paciente de bajo riesgo).

En este caso, deberíamos tratar 400 pacientes de estos por un año para salvar una sola vida ($1/0.0025=400$).

Un hombre de mayor edad con capacidad de ejercicio limitada y extrasístoles ventriculares frecuentes que continúa fumando después del infarto, puede tener un riesgo de morir en el año posterior al infarto tan alto como un 10%. Un 25% de reducción en el riesgo de muerte en un paciente de alto riesgo como este genera una reducción absoluta del riesgo de 2.5% o de 0.025, y tendríamos que tratar sólo 40 individuos de estos por un año para salvar una vida ($1/0.025=40$).

Estos elementos subrayan un elemento clave en la decisión de iniciar la terapia: antes de decidir el tratamiento, debemos considerar el riesgo del evento adverso en nuestro paciente si no se lo trata. Para cualquier RRR, cuanto mayor es la probabilidad de que un paciente experimente un resultado adverso si no lo tratamos, mayor será la probabilidad de que el paciente se beneficie con el tratamiento, y menor será la cantidad necesaria a tratar de esos pacientes para prevenir un evento. Por lo tanto, ambos pacientes y nuestra propia eficiencia clínica se benefician cuando el

NNT para prevenir un evento es bajo.

Nosotros no podríamos vacilar en tratar, aún tantos pacientes como 400 para salvar una vida, si el tratamiento fuera barato, fácil de aplicar y cumplir, e inocuo. Sin embargo, en realidad, los tratamientos son habitualmente caros y acarrean riesgos. Cuando estos riesgos de resultados adversos están documentados en informes de investigaciones, los usuarios pueden aplicar el NNT para juzgar la relación costo-beneficio de la terapia. Si, por ejemplo, los beta-bloqueantes causan una fatiga importante en el 10% de los pacientes que los usan, el NNT para causar fatiga es $1/0.10$ ó 10. Esto se muestra en la Tabla 4, donde se ve que una política de tratamiento para pacientes de bajo riesgo después de un infarto de miocardio (NNT=400 para prevenir una muerte) implicará que 40 estén fatigados por cada vida que se salve. Por otro lado, una política de tratamiento sólo para los pacientes de alto riesgo implicará que 4 estén fatigados por cada vida salvada.

Sin embargo, los clínicos, no tratan los grupos de pacientes uniformemente. Por el contrario, nosotros consideramos las respuestas individuales y ajustamos nuestras terapias de acuerdo a ellas. Una respuesta al problema de los efectos secundarios comunes y relativamente menores (como la fatiga) es suspender la terapia en los pacientes que sufren ese problema. Si nosotros pensamos en los pacientes de bajo riesgo fatigados como un grupo en total, nosotros produciríamos 40 pacientes con fatiga para salvar una vida, lo cual medido en relación riesgo beneficio no se justificaría. Suspendiendo el tratamiento en esas personas, podríamos tratar el resto sin hacer que nadie se fatigue.

Sin embargo, nosotros no podemos aplicar esta propuesta en eventos episódicos severos. Los ejemplos incluyen el riesgo de hemorragia en pacientes a los que se les da anticoagulantes, agentes trombolíticos, o aspirina, o el riesgo de reacciones a la droga raras pero devastadoras.

En cada uno de estos ejemplos el número de efectos adversos por cada vida salvada (o, si los eventos son suficientemente raros, el número de vidas salvadas por cada evento adverso) puede proporcionar un cuadro competente de los costos asociados con la intervención.

CONCLUSIONES

1. Primero, defina el problema claramente y utilice una de las estrategias para obtener la mejor evidencia disponible.
2. Una vez que ha encontrado un artículo relevante del tema terapéutico, determine la calidad de la evidencia. Para la circunstancia en que la calidad de la evidencia es pobre, cualquier inferencia subsiguiente (y la decisión clínica que genera) será insegura.
3. Si la calidad de la evidencia es adecuada, determine el rango en el cual el verdadero efecto de tratamiento cae probablemente.
4. Entonces considere la medida en que los resultados son generalizables al paciente en mano, y si los resultados que se midieron son importantes.
5. Si la generalización está en duda, o la importancia de los resultados es cuestionable, el soporte para la recomendación del tratamiento se debilita.

6. Finalmente, teniendo en cuenta el riesgo del paciente de resultados adversos, determine los resultados probables de la intervención.
7. Esto significa la realización de un balance mirando la hoja de las probabilidades de beneficio y los costos asociados (incluyendo los costos monetarios) y los riesgos.
8. La última línea de la hoja del balance lo guiará a su decisión de tratamiento.

Si bien esto puede parecer una ruta desafiante para decidir un tratamiento, esto es lo que los clínicos hacen implícitamente cada vez que administran una terapia. Realizar el proceso explícitamente y estar en condiciones de aplicar guías de ayuda para determinar la fuerza de la evidencia resultará en una mejor atención del paciente.

Nota.- En el siguiente capítulo se han extractado partes del documento utilizado en el Diplomado de Metodología de la MEDICINA BASADA EN LA EVIDENCIA del CLAP/OPS/OMS. Uruguay – Montevideo/2002

- Diplomado en Medicina Basada en las Evidencias.
- Diplomado en Metodología de la Investigación.
- Diplomado en Meta-análisis y Elaboración de Guías Clínicas.

REFERENCIAS

1. CLAP/OPS/OMS. METODOLOGIA DE LA MEDICINA BASADA EN LA EVIDENCIA. Uruguay – Montevideo. 2002.
2. Liggins GC, and Howie RN. A controlled trial of antepartum glucocorticoid treatment for prevention of the respiratory distress syndrome in premature infants. *Pediatrics* 1972, 50, no.4:515-525.
3. Wingo A., Higgins E, Rubin L, Zahniser C; *Epidemiología aplicada a la salud reproductiva*. OMS 1996