



El análisis crítico de la información médica. Cómo leer artículos sobre terapéutica

EDGARDO SANDOYA

Unidad Coronaria, Hospital Español, Montevideo, Uruguay

Durante décadas la investigación y la enseñanza médica han enfatizado el conocimiento de las enfermedades a nivel básico a partir de los estudios de laboratorio. Los resultados han sido extraordinarios. Sin embargo, desde el punto de vista asistencial los descubrimientos del laboratorio deben ser considerados como hipótesis hasta que no sean validados por investigaciones clínicas cuidadosas.¹ El método empleado para la observación clínica y la experimentación con grupos de pacientes —para predecir luego el curso de pacientes individuales— se ha desarrollado extensamente en la última década. El cuerpo de conocimientos que acompaña estos métodos se ha transformado en una nueva ciencia básica en medicina: la epidemiología clínica. Esta procura integrar los avances en biología humana, ciencia del medio ambiente y medicina clínica a través del aporte de la metodología estadística con la finalidad de hacer la medicina más racional, efectiva y eficiente. La lectura crítica es parte medular de la disciplina referida y provee bases científicas para tomar decisiones acerca de la lectura. El aplicarlas puede ayudar a los médicos a obtener de ella respuestas que probablemente sean tanto válidas como aplicables en su práctica habitual.²⁻⁴

La calidad de la asistencia médica ha mejorado sustancialmente a partir de dos grandes avances científicos: el desarrollo de intervenciones terapéuticas efectivas y la demostración a través de estudios clínicos randomizados (ECR) de que las mismas reportan beneficios.

Los ECR constituyen uno de los avances médicos más importantes de este siglo.⁵ Los mismos consisten en la exposición prospectiva y organizada de pacientes a una intervención determinada (farmacológica, quirúrgica, alimentaria, o de otra naturaleza) para contestar una pregunta específica. Si bien los estudios retrospectivos, las observaciones cuidadosas de casos limitados y los análisis epidemiológicos mantienen vigencia y pueden aportar conocimien-

tos relevantes, son los ECR los únicos bajo control del investigador, no sujetos al azar ni a la disponibilidad de datos, por lo que constituyen la herramienta idónea para la evaluación de nuevos tratamientos.⁶

A pesar de que su diseño es muy bueno, ya que remedia las condiciones de experimentación en el laboratorio, los ECR no son sinónimo de infalibilidad.

Basta tener presente los resultados contradictorios de diferentes ECR sobre un mismo tópico. Estos se producen ya sea por fallas metodológicas en el diseño (que determinan diferencias entre los pacientes, en los efectos medidos o en los tratamientos administrados) o por muestras pequeñas de pacientes que no aseguran detectar diferencias entre los grupos tratados y control.

Por eso, si bien los ECR son la única herramienta válida —en el momento actual— para evaluar nuevas terapéuticas, al igual que las otras formas de investigación médica, requieren un análisis cuidadoso para determinar tanto su validez como su aplicabilidad.⁵

POR QUE LEER Y COMO HACERLO CRITICAMENTE

Diversos autores han tratado de establecer cuáles son los medios que emplean los médicos para mantenerse informados.⁷⁻⁹ En todos los casos la lectura de revistas es el método al que se dedica más tiempo y el que con más frecuencia determina cambios en la práctica clínica. El uso de revistas médicas consistentemente se prefirió a otras publicaciones periódicas, libros de texto y monografías. También se situó antes de otras formas de educación continua como congresos, cursos, rounds clínicos y concurrencia a centros.

Los médicos recibieron muchas de las revistas gratis por correo y/o fueron financiadas por empresas comerciales. Sin embargo a la mayoría de las mismas las descartaron sin haberlas leído,

prefiriendo la lectura de revistas reconocidas. Algunos, sin embargo, prefirieron la lectura de malas publicaciones. Médicos rurales, médicos generales y aquellos que no interactuaban con estudiantes era más probable que leyeran literatura no solicitada y tuvieran menos tendencia a frecuentar las bibliotecas.

Dado que la medicina es una disciplina en permanente actualización, el volumen de información que anualmente se genera es muy amplio. Para mantenerse al día con las principales publicaciones un clínico general debería leer 200 artículos y 70 editoriales por mes, número que año a año va en aumento.³ Pero el tiempo de que dispone para la lectura es el mismo o se ve reducido por otras obligaciones. La conjunción de estos dos hechos —información médica en aumento y tiempo para la lectura constante o reducido— hace difícil al médico mantenerse al día.

La lectura crítica brinda elementos para lograr —mediante la selección adecuada de los artículos— la permanente actualización de los conocimientos necesarios en la práctica.

El análisis que se propone implica un cambio de actitud ante el material impreso. Se tiene tendencia a considerar que los artículos son de buena calidad al ser publicados en revistas importantes, ya que éstas imponen revisiones previas por expertos para su aceptación.

Tan pronto el lector se interna en el análisis, ve que existen muchos artículos con defectos metodológicos y/o estadísticos, incluso en esas revistas. Por eso el médico debe conocer los criterios con los que el investigador diseña, ejecuta y presenta los resultados para poder así analizar los ECR adecuadamente.¹⁰

LAS SEIS REGLAS "FILTRO"

Al plantear estas guías se asume que los médicos están atrasados en su lectura clínica y que no van a disponer de más tiempo para leer que el que tienen actualmente. Por esta razón —y debido a que el análisis debe centrarse en la sección material y métodos— el objetivo inicial de esta estrategia es descartar rápidamente los artículos en base a la evidencia preliminar. Sólo a través del descarte de la mayoría es posible dedicarse al análisis en profundidad de los restantes para identificar los que sean tanto válidos como aplicables en la propia práctica.

Las seis reglas iniciales de la Figura 1 son el "filtro" para cumplir con este propósito.

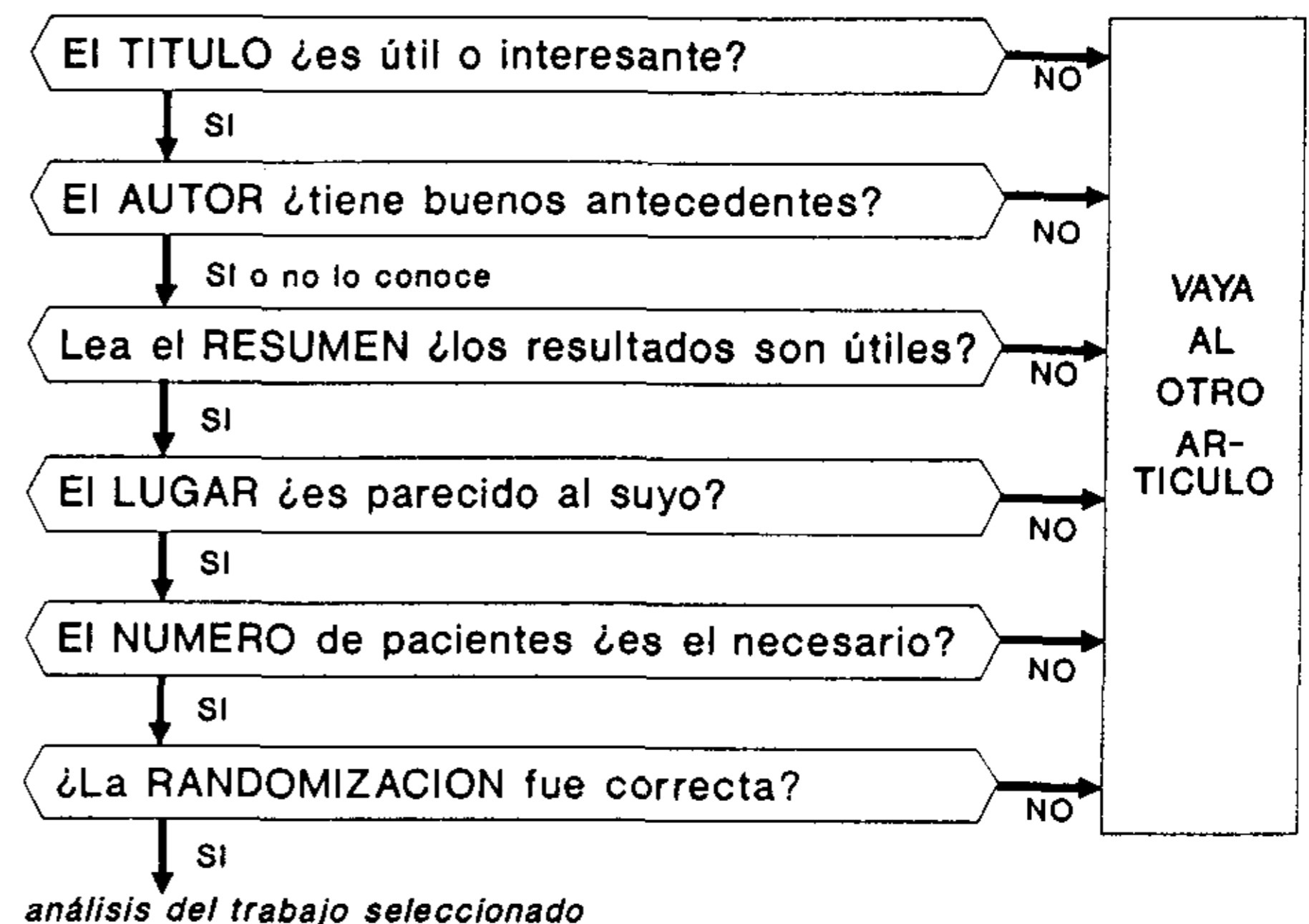


Fig. 1. Modificada de Can Med Assoc J.³

1. *Vea el título: ¿el artículo es interesante o útil en su práctica?* Si no lo es, déjelo de lado y pase al próximo.

2. *¿Los autores tienen buenos antecedentes?* Si es así o si no son conocidos para usted, siga adelante. Si sus antecedentes no son buenos, vaya al otro artículo.

3. *Lea el resumen: ¿los resultados pueden ser útiles?*

El objetivo no es analizar si los resultados son válidos —lo que habitualmente no se puede determinar leyendo el resumen— sino que en el caso de que sean correctos los mismos pueden llegar a ser útiles.

4. *¿El lugar donde se realizó es parecido al suyo?*

Es decir, ¿los resultados son fácilmente transportables a su práctica clínica? Por ejemplo, si usted actúa a nivel primario no puede extrapolar los resultados obtenidos en centros de nivel terciario. Durante años los médicos generales han sido aconsejados por los nefrólogos de centros de referencia para estudiar con urografía, ecografía y toda una batería de exámenes de laboratorio a los enfermos hipertensos para tratar de detectar los casos curables. Una de las principales razones para esta recomendación fue la cantidad relativamente alta de hipertensión arterial corregible quirúrgicamente entre los pacientes de centros de nivel terciario.

El 6% de los pacientes referidos a la Cleveland Clinic —por ejemplo— tenían hipertensión arterial de este tipo. Cuando se realizaron los mismos estudios a la población de hipertensos asistidos en la práctica general de Ontario, la

misma se encontró sólo en el 0,6% de los mismos.³ Ambas conclusiones son correctas. Su diferencia se origina a partir de los "filtros" a través de los cuales deben pasar los pacientes para llegar a la Cleveland Clinic, lo que selecciona una población diferente de la de la práctica general.

Prestando atención al lugar donde se llevó a cabo el estudio se evitará la aplicación errónea del tratamiento en poblaciones diferentes.

5. ¿El número de pacientes es el necesario?

La base para realizar un ECR correcto consiste en incluir el número adecuado de pacientes. La teoría estadística clásica utilizaba dos elementos para validar un resultado: la plausibilidad de la hipótesis y el valor de p. Actualmente se acepta que el tamaño del ensayo es un tercer criterio adicional independiente.¹¹ El número de pacientes que debe tener un ECR depende de dos variables: la prevalencia del evento analizado (por ejemplo, mortalidad) y la reducción esperable del mismo con el tratamiento.¹² Dichos datos se toman de estudios previos referidos al tema. Cuando los mismos no existen, se debe realizar un estudio piloto inicial con un número limitado de pacientes para sobre esa base diseñar el estudio. Como puede verse en la Tabla 1, cuando menor es la prevalencia del evento y menor la reducción esperada del mismo, mayor es el número de casos necesarios. Por el contrario, tanto la mayor prevalencia basal como la mayor reducción esperada del evento disminuyen los pacientes requeridos.

La mortalidad del infarto agudo de miocardio (IAM) en la etapa intrahospitalaria se halla próxima al 10%.¹³ Para detectar una reducción del 20% de la misma (es decir, que pase de 10% a 8%) es necesario incluir 8.600 pacientes en el estudio para que los resultados —positivos o negativos— sean válidos.

En la insuficiencia cardíaca grado 4 la mortalidad al año es cercana al 40%. Los inhibidores de la enzima convertidora pueden reducirla en este grupo en aproximadamente 40%, por lo que el número de pacientes que se necesita para realizar un estudio sobre este tema es de 175 en cada uno de los grupos.¹⁴

La Tabla 1 puede ser empleada para determinar el número de pacientes que deben incluirse en cada caso.

En los problemas cardiológicos comunes, donde los resultados terapéuticos que se esperan son en general modestos, habitualmente se necesita gran cantidad de pacientes.¹⁵

Tabla 1
Casos necesarios para un ECR*
(Modificado de JAMA¹²)

Reducción de riesgo %	Frecuencia de eventos %				
	10	15	20	30	40
10	36.000	22.650	16.000	9.300	6.000
20	8.600	5.300	3.900	2.300	1.500
30	3.600	2.250	1.650	1.000	650
40	2.000	1.250	850	550	350
50	1.200	750	550	300	250

* $\alpha = 0,05$; poder = 0,90.

6. ¿La randomización fue correcta?

La randomización sirve para evitar desvíos en la selección de pacientes, como seguro a largo plazo contra los desvíos accidentales, para permitir adoptar procedimientos ciegos y para dar validez a la aplicación posterior de tests estadísticos.²

Por eso no son válidos los ensayos terapéuticos sin grupo control. Cuando se realizan estudios no controlados aplicando el nuevo tratamiento a un grupo de pacientes sin otro de referencia, invariablemente los resultados van a ser mejores que cuando se realizan estudios controlados. C. Foulds, analizando 52 estudios no controlados en psiquiatría, vio que el 85% de ellos obtenían éxito terapéutico, mientras que en 20 ensayos controlados sólo el 35% mostró resultados similares.¹⁵ Esto sucede porque para dar una chance razonable al nuevo tratamiento se puede —inconscientemente o no— seleccionar los pacientes menos graves. Como consecuencia de ello —e independientemente del real valor del tratamiento— dicho grupo va a evolucionar sorpresivamente bien comparado con la práctica habitual. Además se puede tender a poner mayor énfasis en los éxitos, quizás incluso exagerando un poco, y pueden no reportarse algunos fracasos sobre las bases de que los pacientes estaban "muy enfermos" para beneficiarse con el nuevo tratamiento.

El grupo control recibirá el tratamiento estándar de real valor para esa situación. En el caso de que el mismo no exista, entonces recibirá placebo. Cuando se realizaron los primeros estudios con estreptoquinasa (SK) en el del IAM, los pacientes del grupo control recibieron placebo. Los estudios actualmente en curso GISSI-2 e ISIS-3 quieren determinar cuál es el valor del activador tisular del plasminógeno (t-PA) para la misma patología. Dado que existe

un tratamiento eficaz para el IAM, no sería ético administrar placebo a los pacientes del grupo control, por lo que en ambos estudios ese grupo recibe SK.

Cada paciente que entre en el estudio debe tener la misma posibilidad de recibir uno u otro tratamiento, es decir, de ser incluido en el grupo tratado o en el grupo control. Esto se logra solamente mediante la asignación al azar de los pacientes a uno u otro grupo. De esta forma se asegura que las poblaciones van a ser homogéneas (más lo serán en la medida que sea mayor el número). La única diferencia entre ambos va a ser el tratamiento empleado. Se puede analizar entonces si esa única variable (tratamiento) determina resultados diferentes. De no distribuirse los pacientes al azar, una serie de variables (gravedad de los casos, edad, tiempo de evolución, sexo, etc.) va a incidir sobre el resultado, impidiendo valorar el verdadero efecto del tratamiento.

Es incorrecto randomizar por hospital o equipo quirúrgico, pues esto determina importantes desvíos: no sólo del tratamiento sino también por las diferencias entre centros y del seguimiento médico. Tampoco es correcto randomizar alternativamente identificando el tratamiento del paciente anterior, o en base al número de historia o a la fecha de nacimiento. Estas formas de distribución no garantizan la ausencia de desvíos en la selección.

La randomización simple tiene desventajas en algunas circunstancias: si el tamaño del ECR es menor de 100 pacientes, en estudios multicéntricos o cuando se planean análisis interinos (cortes cada cierto número de pacientes para evaluar los resultados parciales). En todos estos casos se debe realizar randomización en bloques para permitir que los grupos se vayan balanceando permanentemente a lo largo del estudio.

Los ensayos ciegos (el médico o el paciente no conocen el tratamiento asignado) o dobles ciegos (ni médico ni paciente pueden identificar el tratamiento) son imprescindibles cuando se evalúan marcadores "blandos" como calidad de vida, en donde la evaluación dificultosa de los resultados se complica aún más al agregarle el conocimiento por parte de los actores del tratamiento recibido. Cuando se evalúa otro tipo de marcadores no es imprescindible que los ECR sean ciegos. Una ventaja adicional que tienen los estudios ciegos es que se pueden evaluar los efectos colaterales en forma más objetiva.

No pueden ser usados como control los grupos históricos. El problema radica en que no se puede contestar a la pregunta básica ¿los gru-

pos son comparables? Los pacientes de los grupos históricos es probable que no tengan criterios de inclusión claramente definidos; puede haber un cambio en los pacientes elegibles, el investigador puede haber variado sus criterios de selección, la calidad de los datos registrados puede ser inferior a la necesaria y —muy importante— pueden haber variado otras medidas complementarias al tratamiento. Por todo ello los controles históricos tienden a exagerar el valor de los nuevos tratamientos. H. Sacks, analizando la literatura, halló 50 ECR y 57 estudios con controles históricos referidos a seis problemas médicos. En el 79% de estos últimos el tratamiento nuevo era mejor mientras que lo mismo sucedió sólo en el 20% de los ECR (Fig. 2).¹⁵

Existen excepciones a la necesidad de randomización. Antes del advenimiento de la estreptomycin, la mortalidad de la meningitis tuberculosa era del 100%. Cuando en 1946 se comenzó a disponer de la droga, se vio que los pacientes tratados sobrevivían, por lo que no fue necesario un estudio randomizado para demostrar la utilidad del fármaco.¹⁶ Otro tanto ocurrió con el tratamiento de las infecciones graves con penicilina.¹⁷ Pero no son frecuentes las patologías con cifras de 100% de mortalidad y con porcentajes de curación de magnitud parecida, por lo que estas situaciones constituyen una excepción.

Tampoco se debe desconocer que gran parte de los ECR no se publican cuando los resultados son "negativos". La pérdida de interés de los investigadores o razones de índole comercial hacen que no salgan a luz los resultados de muchos ECR con resultados diferentes a los esperados. Por tal motivo se conoce solamente una parte de la evidencia científica. Para corre-

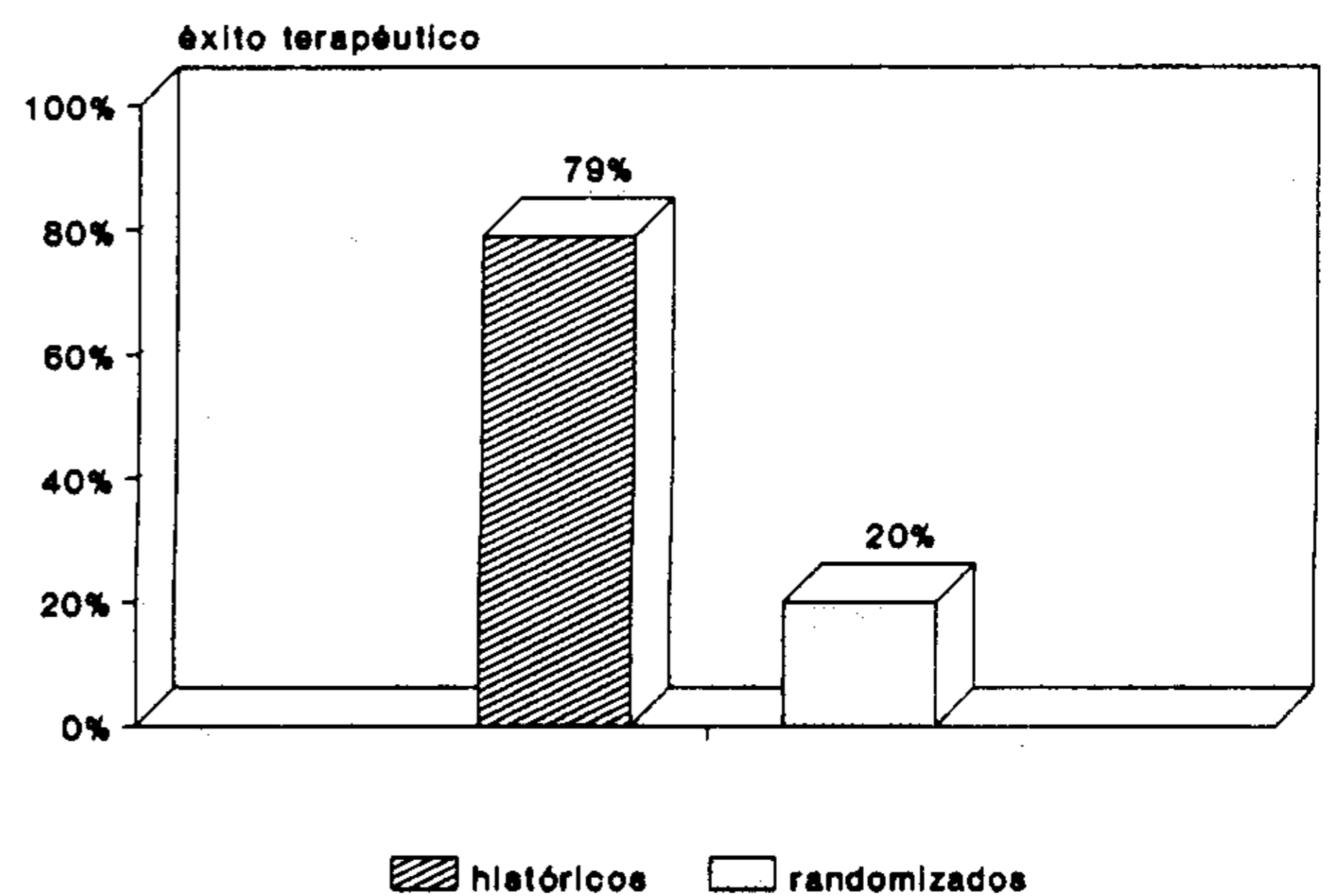


Fig. 2. Resultados y tipo de controles.

gir esto existe en algunos países la obligatoriedad de registrar las investigaciones que se comienzan y el informe periódico de sus resultados, independientemente de cuáles sean los mismos.

Cuando estamos buscando información para resolver el problema concreto de algún paciente, las cosas a veces pueden ser diferentes. En esas circunstancias muchas veces no existe ningún ECR válido y aplicable. Debe entonces realizarse una revisión de la literatura para tratar de localizar alguno no conocido. De no encontrarse ninguno, el médico debe valerse entonces de la experimentación subclínica existente. Es importante —una vez llegado a ese punto— tener en cuenta que los resultados de la experimentación animal no se pueden extrapolar a la clínica como tales, tanto sea en terapéutica como en la exploración de hipótesis fisiopatológicas o de causalidad. En muchas circunstancias sucede que los hechos originados a partir de la investigación básica en modelos animales, después no se correlacionan con lo que sucede en los pacientes.

Antes de actuar en estas situaciones el clínico debe replantearse: ¿realmente el paciente requiere alguna intervención?, y si es así, ¿cuál es la que probablemente presenta el mejor balance riesgo/beneficio? El objetivo es utilizar la mejor información disponible cuando se deben tomar decisiones frente al paciente individual.

En esas circunstancias debe tenerse presente también que los umbrales de decisión son variables. Para problemas con grandes beneficios potenciales y poco riesgo o costo podrían aceptarse menores estándares de evidencia que para problemas con beneficios potenciales menores o grandes riesgos o costos.

ANALISIS DEL TRABAJO SELECCIONADO

Una vez que el artículo ha pasado el filtro de las reglas previas, debe proceder a analizarlo en sus puntos clave para determinar si el mismo es válido y aplicable en su práctica. Para ello es necesario conocer el material y método empleados. Eso lleva a que deban preferirse los artículos originales en los que estos aspectos se describen con exactitud, información de la que habitualmente no se dispone cuando se leen revisiones o editoriales.¹⁸ Esto no implica que éstos no tengan su lugar en la lectura clínica. Los editoriales en general contienen ideas interesantes expresadas por expertos reconocidos y las revisiones son la mejor forma de acercarse a un conocimiento en el que no se es experto. Pero

1. ¿Cuál es el problema que trata de resolver?
2. ¿Cuál es la hipótesis planteada?
3. ¿Fueron considerados todos los hechos relevantes?
4. ¿Se consideraron significación clínica y estadística?
5. ¿Los pacientes fueron similares a los suyos?
6. ¿La maniobra terapéutica es factible en su práctica?
7. ¿Se incluyeron todos los pacientes en la conclusión?

Fig. 3. Guías para el análisis del trabajo seleccionado.

para actualizarse en el área de la propia práctica profesional la información debe ir a buscarse en las publicaciones originales.

Las siguientes reglas permiten determinar la validez y aplicabilidad del trabajo (Fig. 3).

1. ¿Cuál es el problema que trata de resolver?

Considere cuál es el problema que aborda el estudio. Al encarar el análisis crítico de una publicación científica, en primer lugar se debe definir el área de incertidumbre que tiende a dilucidar. De esa forma se estará situando en su correcta perspectiva al estudio en cuestión.²

El tratamiento fibrinolítico del IAM constituye uno de los avances más importantes de la cardiología en la presente década. Pero si analizamos el problema —los pacientes que mueren por IAM en la etapa aguda— vemos que la fibrinólisis no es la solución potencial para todos los casos. De los pacientes con IAM que ingresan a una unidad coronaria, sólo el 37% es pasible de recibir dicho tratamiento.¹³ Esto no invalida el tratamiento fibrinolítico, sino que define su papel con relación al problema. La investigación continúa para tratar de dar solución a la parte del problema aún no resuelta.

2. ¿Cuál es la hipótesis planteada?

Debe determinarse claramente para ver si luego es contestada. Muchas veces se planea un estudio con determinada hipótesis y luego se pretende con sus resultados responder a otra no prevista. Tal conducta es metodológicamente inadecuada y generalmente lleva a conclusiones erróneas.

El estudio TIMI (*Trombolysis in Myocardial Infarction*) fue diseñado para ver si el porcentaje de arterias (responsables del IAM) repermeabilizadas era mayor con t-PA o SK. Se vio que las arterias se reperfundían en un 31% con SK y en un 62% con t-PA. Pero la publicación no se limitó a publicar esos datos, sino que enfatizó

en los resultados de mortalidad en la etapa aguda, así como a los seis y doce meses.¹⁹⁻²⁰ Esto fue manejado posteriormente como evidencia de que la mortalidad en el IAM era menor con t-PA a partir de un estudio no diseñado para tal fin. Se cometieron en ese caso varios errores metodológicos: primero, el estudio no fue planeado para evaluar mortalidad, y por lo tanto no se pudo pretender obtener esa respuesta; en segundo lugar, para hacerlo, el número era absolutamente insuficiente, y en tercer lugar, se asumió a partir de una evidencia indirecta —arteria coronaria permeable— que este hecho era equivalente a menor mortalidad.

El estudio ASSET (*Anglo Scandinavian Study Early Thrombolysis*) estudió la mortalidad con t-PA en 5.000 pacientes. Mostró una reducción de 26% de la misma, similar a la observada con SK, lo que no confirmó la presunción de TIMI.²¹

En ISIS-2 (*International Survival Infarction Study-2*) se observó una reducción del 42% de la mortalidad al tratar con SK y aspirina a los pacientes con IAM. Cuando se realizó un análisis por subgrupos de pacientes de acuerdo con el signo zodiacal, se vio que los nacidos bajo el signo de Géminis o de Libra no se beneficiaban con el tratamiento, mientras que sí lo hacían los nacidos bajo los otros signos.²² Este caso ejemplifica otra premisa metodológica: no es válido el análisis de subgrupos dentro de un ECR.

Los dos ejemplos muestran que es incorrecto tomar los resultados de estudios planeados para otro fin y pretender validar con esos datos otra hipótesis que no fue específicamente explorada.

3. ¿Fueron registrados todos los hechos clínicamente relevantes?

Importa conocer todos los hallazgos del estudio. Aun aquellos que se mencionan superficialmente (o no se publican) pero que pueden llegar a ser relevantes. La controversia t-PA/SK en el tratamiento del IAM constituye un ejemplo a este respecto. Algunos trabajos basaron la preferencia por la primera droga en su selectividad lítica. Se asumió que t-PA produciría menos hemorragias a partir de los estudios de laboratorio que mostraban una caída más importante del fibrinógeno plasmático con SK que con t-PA, tal como puede apreciarse en la Figura 4.²³

En un estudio se encontró que los valores sanguíneos de fibrinógeno eran más bajos con SK que con t-PA pero no se registró la evolución clínica de los pacientes. Posteriormente se conocieron dichas cifras: la incidencia de AVE hemorrágico fue de 1,4% mientras que

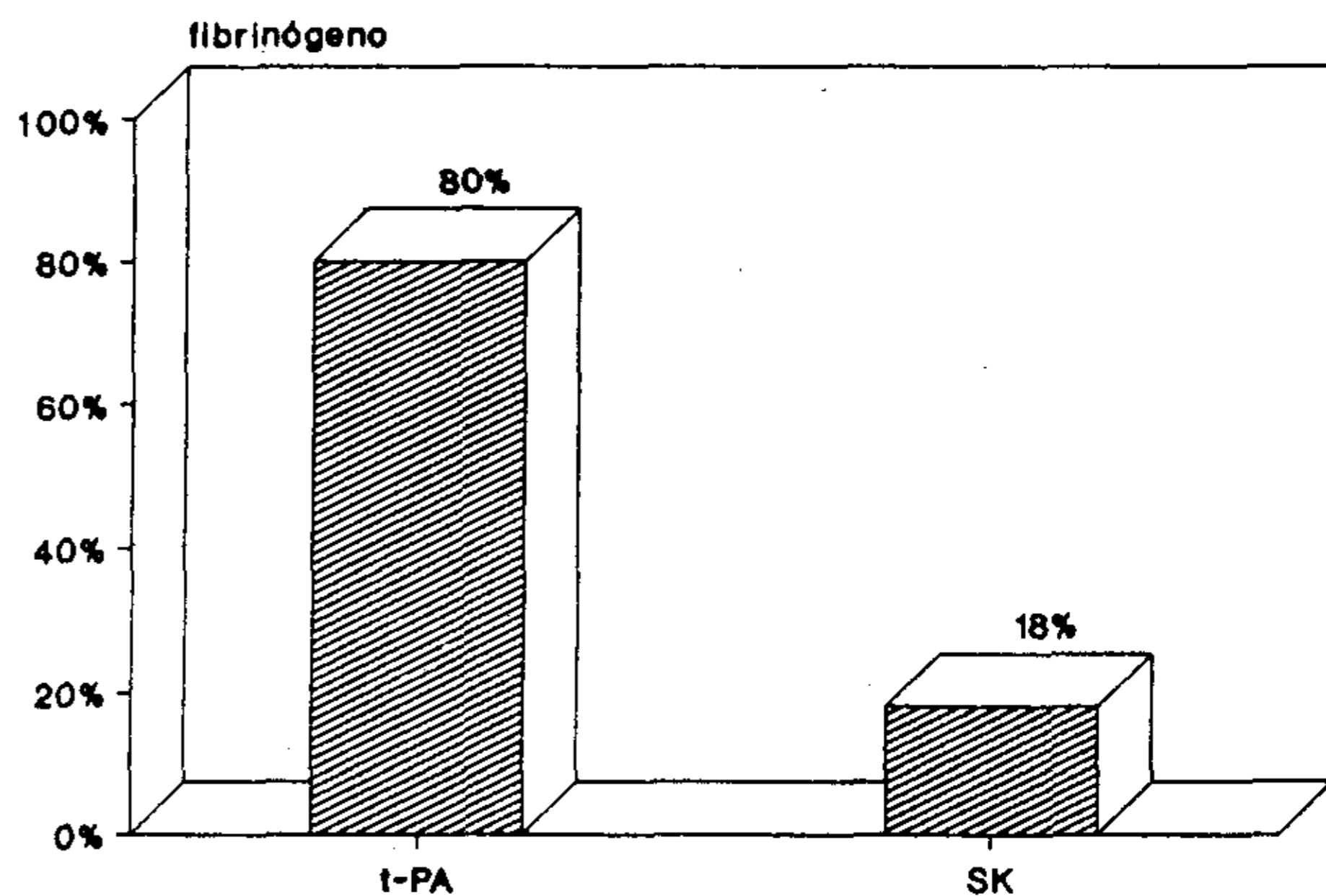


Fig. 4. Fibrinógeno a los 60 minutos de administrar el fármaco.

con SK la misma se da en el 0,2%.¹³ Es decir que con t-PA fue 7 veces superior, a pesar de lo que decían las mediciones de laboratorio (Fig. 5).

El ECR referido introduce además otro elemento a considerar: el empleo de marcadores o *end-points* indirectos.

Es importante en cada caso establecer si los marcadores que se analizan son directos o indirectos.

Veamos un ejemplo. Cuando se trata la hipertensión arterial, el objetivo es —en primer lugar— disminuir la mortalidad. No es lo mismo un estudio que evalúa cifras de presión arterial que otro que analiza mortalidad. La reducción de las cifras de PA anormalmente elevada no necesariamente implica una disminución paralela de la mortalidad, aunque esto parezca lógico. De hecho la mayoría de los fármacos disminuyen la PA y ninguno ha demostrado mejorar significativamente la sobrevida de los

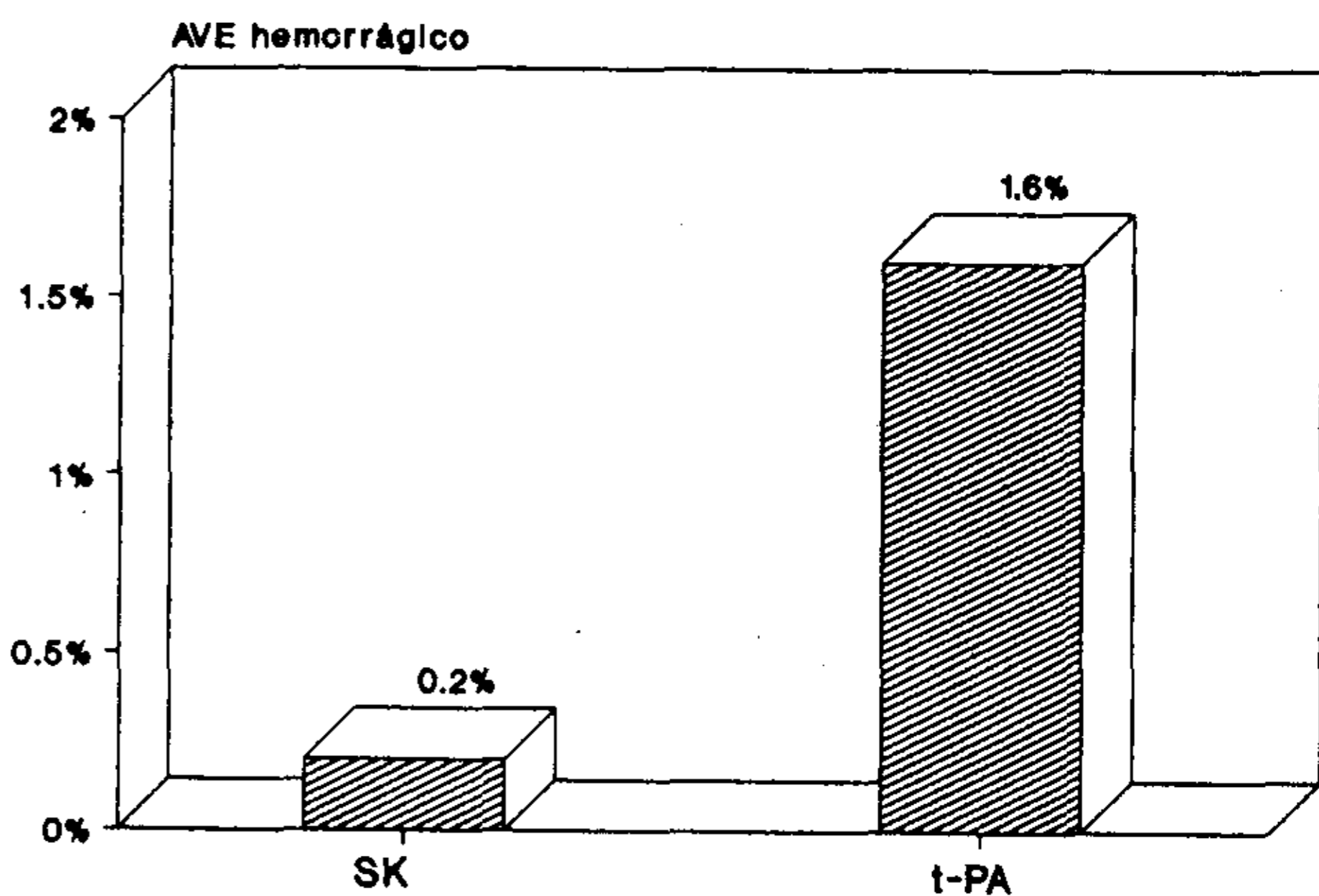


Fig. 5. AVE hemorrágico SK (GISSI) y t-PA (TIMI).

hipertensos leves o moderados.^{24, 25}

En suma, si se quiere valorar el efecto de un fármaco sobre determinado parámetro se debe realizar un estudio específico con esa finalidad y no realizar asunciones a partir de otros datos.

Finalmente otro elemento a tener presente es que los marcadores "duros" como mortalidad son más fácilmente cuantificables (y por ende más confiables) que los "blandos" como calidad de vida, marcadores éstos más difíciles de evaluar aunque no por ello menos importantes.²

4. ¿Se consideraron significación clínica y estadística?

No existe consenso sobre los métodos estadísticos a emplear para analizar los resultados de un ECR.⁶ Sí lo existe en cuanto a lo que se considera estadísticamente significativo. Se acepta que la diferencia entre dos tipos de tratamiento es estadísticamente significativa cuando la posibilidad de que la misma obedezca al azar es igual o menor de 5%.²⁶

Actualmente se prefiere expresar los resultados de un ECR como un determinado valor de reducción de riesgo (RR) y su intervalo de confianza (IC) más que por el valor de p.²⁷ Dicha preferencia se basa en que la p divide los resultados en dos categorías: significativos o no significativos, es decir, da una respuesta de tipo sí o no a la pregunta: ¿los tratamientos son diferentes?

El IC permite ver entre qué límites se extiende el beneficio —si es que éste existe—, dando respuesta cuantitativa a la pregunta ¿cuán grande es el beneficio?

Habitualmente el IC que se emplea es el IC 95%, siendo el que se corresponde con p menor o igual a 0,05.

El estudio noruego con timolol luego del IAM tuvo los siguientes resultados: RR -40 e IC 95% (-55 -21).²⁸ Esto expresa que los pacientes que recibieron el betabloqueante tuvieron un 40% menos de posibilidad de morir que los que recibieron placebo y que los límites del beneficio se extendieron entre -55% y -21%.

Siempre que la RR es de signo negativo habla de beneficio, mientras que cuando la misma es de signo positivo habla de perjuicio.

La trombólisis tardía en el IAM (pacientes ingresados entre las 6 y 24 horas de evolución del mismo) muestra: RR -13 e IC 95% (-25 -2). Si bien el tratamiento es beneficioso y estadísticamente significativo, el límite de confianza superior no excluye que el beneficio pueda ser casi nulo.²⁹

Para poder determinar si existe o no real be-

neficio es necesario disponer de más evidencia, lo que implica realizar nuevos estudios en tal sentido.*

La significación estadística dice si la diferencia observada entre los tratamientos es real y no obedece al azar (siempre y cuando se haya cumplido con los prerequisites de número adecuado y correcta randomización). Pero no habla de la relevancia clínica.

La significación clínica (o importancia clínica) se refiere a la relevancia médica que tienen los resultados obtenidos.

5. ¿Los pacientes del estudio fueron similares a los suyos?

Para saberlo es necesario conocer los criterios de inclusión y exclusión. Deben asimismo analizarse las características basales de quienes ingresaron al estudio y ver si son similares a las de sus propios pacientes.

También es importante conocer el total de la población que se ha analizado y qué parte de la misma ha entrado finalmente en el ECR. El siguiente ejemplo ilustra la importancia de esta consideración.

El CASS (*Coronary Surgery Study*) fue un estudio multicéntrico para comparar tratamiento médico *versus* quirúrgico en pacientes con lesiones coronarias.³⁰ Se incluyeron en él 780 pacientes sobre un total de 24.959 candidatos analizados en los 15 centros participantes. Constituye un ejemplo de ECR en el cual, debido a rígidos criterios de inclusión, se conforma un grupo tan seleccionado que es difícil encontrar pacientes en la práctica habitual similares a los del estudio.

A diferencia de éste, los estudios clínicos de poblaciones —que responden a otra filosofía de investigación— ingresan gran cantidad de pacientes partiendo de criterios de inclusión no restrictivos. Eso hace que reflejen la realidad clínica habitual pues incluyen pacientes similares a los que se ven en la práctica corriente. Sus resultados entonces son aplicables a grupos extensos de población, dado que la mayoría de la misma se ha visto contemplada en ellos.³¹

* Para contestar ese interrogante se están realizando en este momento dos estudios a nivel mundial: LATE (*Late Assessment Thrombolytic Essay*) en los Estados Unidos y EMERAS (Estudio Multicéntrico Estreptoquinasa Repúblicas de América del Sur). En este último estudio participan 220 unidades coronarias de cinco países (Argentina —que es el coordinador—, Brasil, Chile, Uruguay y Venezuela) bajo la supervisión de la Universidad de Oxford (Gran Bretaña). Se piensa incluir en él aproximadamente 5.000 pacientes hasta fines de 1990.³³

Una ventaja adicional de esta forma de investigación consiste en que permite reducir el tiempo que habitualmente media entre la demostración del beneficio de una nueva terapéutica y la incorporación de la misma a la práctica corriente.

Otro aspecto metodológico a considerar del CASS —si bien no vinculado al punto actual— es la fecha de ingreso de pacientes. La misma se extendió de 1974 a 1979, por lo que, teniendo en cuenta el avance que ha experimentado la cirugía de by-pass desde entonces, sus resultados no son directamente extrapolables a la realidad actual.

El estudio CAST (*Cardiac Arrhythmia Suppression Trial*) valoró mortalidad en pacientes post-IAM randomizando entre placebo, encainida, flecainida y morizicina.³²

Para participar en el mismo los pacientes debían tener arritmias ventriculares medianamente severas y ser asintomáticos o moderadamente sintomáticos. Asimismo su fracción de eyección debía ser inferior a 55% o a 40%, dependiendo del tiempo que había transcurrido desde el IAM.

El mismo se suspendió, luego de haber incluido 2.309 pacientes, cuando se vio que la mortalidad era mayor en los grupos tratados con flecainida o encainida que en el grupo control. A partir de estos resultados no es válido afirmar que los fármacos referidos tienen más riesgo que beneficio luego de un IAM. Sí se lo puede afirmar en aquellos pacientes de características similares a los incluidos en el estudio.

Recuerde entonces analizar cuidadosamente los criterios de inclusión y exclusión para poder aplicar la terapéutica a aquellos de sus pacientes que cumplan con los mismos requisitos. De otra forma estaría extrapolando resultados a pacientes en los que no ha sido determinado el efecto del tratamiento.

6. ¿La maniobra terapéutica es realizable en su práctica?

Para poder contestar este interrogante son necesarios varios elementos. En primer lugar, que se haya descrito con detalle para poder repetirlo en los mismos términos. Deben conocerse dosis (de tratarse de fármacos) así como esquema de administración, vía utilizada y duración del tratamiento. Si se trata de terapéuticas no farmacológicas, debe conocerse la técnica quirúrgica, los catéteres o los equipos empleados, así como el momento de la evolución en que se realizó el procedimiento. Después debe plantearse si la maniobra es realizable en sus condi-

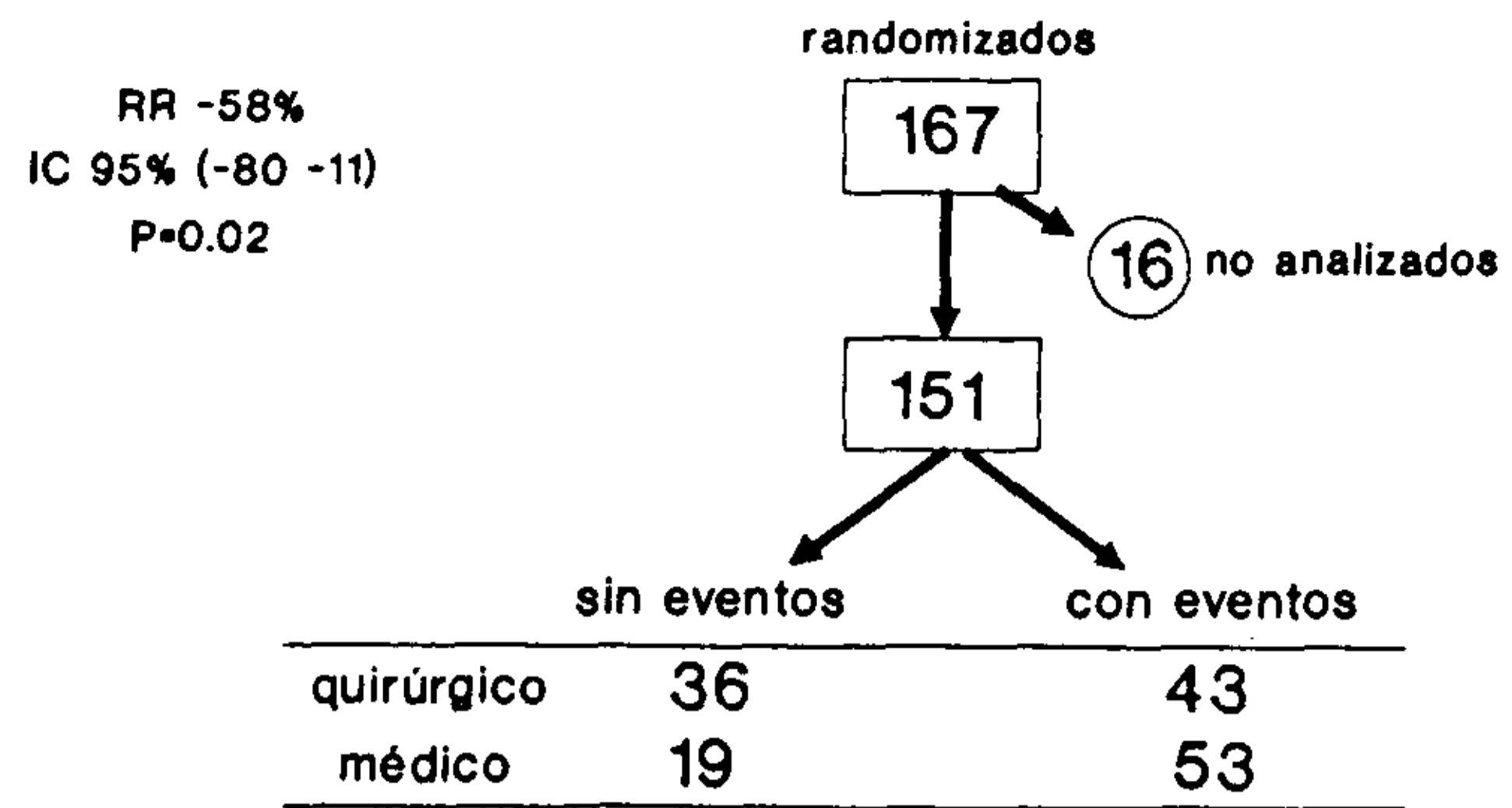


Fig. 6. Pacientes disponibles para el seguimiento.

ciones de trabajo. ¿Dispone usted del fármaco, de la pericia técnica, del equipo necesario? Si esto es así, ¿el paciente está dispuesto a aceptar el nuevo tratamiento?

7. ¿Se incluyeron todos los pacientes en la conclusión final?

El lector debe revisar con cuidado cuántos pacientes se incluyeron en el estudio y cuántos se consideran en las conclusiones finales. Un estudio realizado para valorar el tratamiento médico *versus* quirúrgico en la estenosis carotídea bilateral mostró los resultados que se ven en la Figura 6.

Entre los pacientes disponibles para el seguimiento se demostró una reducción del 58% ($p = 0,02$) del riesgo de AIT, AVE o muerte en los del grupo quirúrgico. Sin embargo, analizando cuidadosamente el trabajo se ve que entraron en el estudio 167 pacientes y no 151, como se dio a conocer en las conclusiones. Los 16 que no se incluían en los resultados tuvieron un AVE o fallecieron durante su hospitalización inicial, por lo que fueron excluidos del análisis posterior. De los 16, 15 habían sido randomizados a cirugía. Los resultados de su reinclusión en el análisis final muestran que la reducción de eventos en el grupo quirúrgico es 44%, no siendo los resultados estadísticamente significativos: IC 95% (-73 +16%) y $p = 0,13$, tal como se ve en la Figura 7.³⁴

Este estudio ejemplifica la importancia de saber la evolución del total de los pacientes, lo que debe figurar en todos los casos.

Actualmente se acepta que los resultados deben ser analizados en base al principio de intención de tratar, y la mayoría de los trabajos así lo hacen.¹¹ Este establece que todos los pacientes randomizados son incluidos en el análisis independientemente del cumplimiento con el

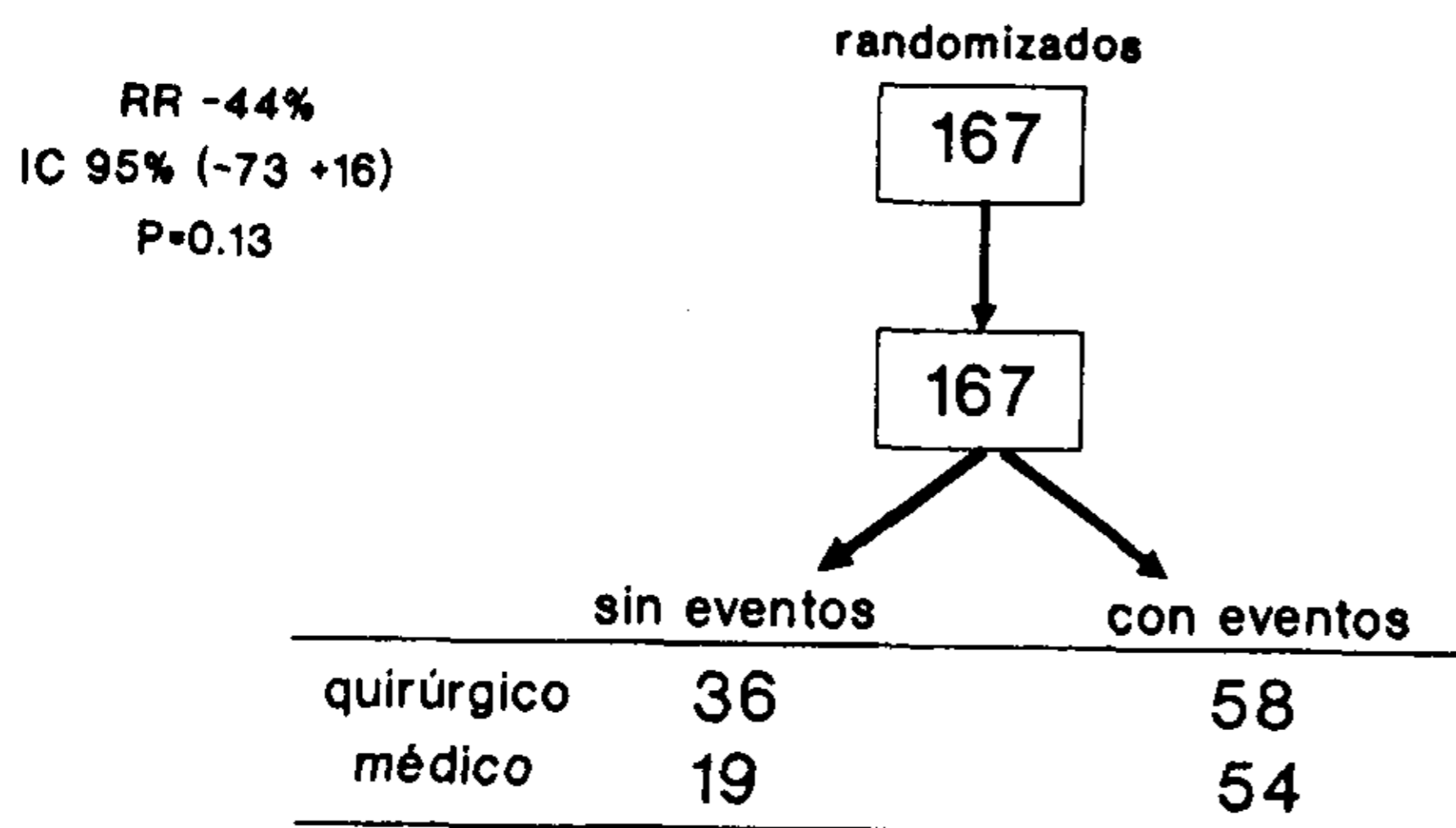


Fig. 7. Total de pacientes.

tratamiento. Si un paciente asignado a un grupo no recibe el tratamiento, el mismo es tenido en cuenta para el análisis de los resultados finales como si lo hubiese recibido. Esto no introduce desvíos porque el azar va a hacer que en el otro grupo también se den casos similares, con lo que al final el estudio se va a encontrar balanceado.

Si el porcentaje de pacientes que se cruzan de un grupo a otro es muy elevado, entonces no es posible valorar los resultados.

Un aspecto filosófico que subyace detrás de esta forma de análisis es que con los ECR no se estudian fármacos sino problemas en el manejo de los pacientes. En todos los casos se está tratando de resolver una situación para la que no existe una terapéutica adecuada y no preocupándose de demostrar que un fármaco es bueno o no. La intención de tratar permite analizar los resultados tomando una población que no va a diferir demasiado con lo que sucede en la práctica: hay pacientes que cumplen con el tratamiento, otros que lo hacen en menor grado y otros que no lo hacen.²

CONCLUSION

En las principales revistas es cada vez mayor el número de ECR que se publican. Son también cada vez más frecuentes los estudios multicéntricos en los que participan diversos centros de uno o varios países, los que representan realidades de grupos más extensos de poblaciones. Pero no todos son de buena calidad, por lo que el porcentaje de artículos que satisfagan tanto los criterios de consistencia científica y de aplicabilidad clínica como los descritos en el presente trabajo no van a ser muchos. Sin embargo, los artículos que sobrevivan a la aplicación de estas guías van a ser los más válidos, relevantes y aplicables en su propia práctica. De todas formas,

los ECR correctos no dan una respuesta definitiva para todo momento. Tampoco invalidan la "cascada" anterior: el laboratorio, los estudios en voluntarios sanos, la observación individual, la pequeña casuística y el criterio clínico, pues éstos son los generadores de las hipótesis que luego serán evaluadas en los estudios randomizados.

Estos últimos pueden contribuir —en alguna medida— a mejorar la capacidad para manejar los problemas de los pacientes hasta que nueva evidencia obligue a adoptar otras conductas.

Esperamos que estas guías puedan ser de utilidad para que —tomando una frase del grupo canadiense— los médicos sigamos haciendo más bien que mal a nuestros pacientes.

AGRADECIMIENTOS

Deseo agradecer los valiosos comentarios y sugerencias realizados por los Dres. Carlos Bertolasi, Joaquín Cobas, Ricardo Lorenzo, Jorge Pouso, Hugo Senra y Gianni Tognoni.

BIBLIOGRAFIA

1. Fletcher RH, Fletcher SN: Clinical epidemiology: a new discipline for an old art. *Ann Intern Med* 99: 401-403, 1983.
2. Tognoni G, Maggioni AP, Franzosi MG: Curso de Metodología Científica en Cardiología. Praxis Médica. Buenos Aires, 1989.
3. Department of Clinical Epidemiology and Biostatistics Mac Master University: How to read clinicals journals I: Why to read them and how to star reading them critically. *Can Med Assoc J* 124: 555-558, 1981.
4. Department of Clinical Epidemiology and Biostatistics Mac Master University: How to read clinicals journals V: To distinguish useful from useless or even harmful therapy. *Can Med Assoc J* 124: 1156-1162, 1981.
5. Elwood JM: Interpreting clinical trials: seven steps to understanding. *Can Med Assoc J* 123: 343-346, 1980.
6. Temple E: Government viewpoint of Clinical Trials of Cardiovascular Drugs. *Med Con N Am* 73: 495-509, 1989.
7. Stinson ER, Mueller DA: Survey of health professionals' information habits and needs: conducted through personal interviews. *JAMA* 243: 140-143, 1980.
8. Curry L, Putnam RW: Continuing Medical Education in Maritime Canada: the methods physicians use, would prefer and find most effective. *Can Med Assoc J* 124: 563-566, 1981.
9. Covell DG, Uman GC, Manning PR: Information needs in office practice: are they being met? *Ann Intern Med* 103: 596-599, 1985.
10. Belitzky R: ¿Qué conviene leer? Montevideo: Publicación CLAP 843, 1979.
11. Preto R, Pike MC, Armitage P et al: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I: Introduction and design. *Br J Cancer* 24: 585-612, 1976.
12. Yusuf S, Wittes J, Friedman L: Overview of results of randomized clinical trials in heart disease. I: Treatment following myocardial infarction. *JAMA* 260: 2088-2093, 1988.
13. GISSI (Gruppo Italiano per lo Studio Della Streptochinasi Nell'Infarto Miocardico): Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. *Lancet* 1: 397-401, 1986.

14. The CONSENSUS Trial Study Group: Effects of enalapril on mortality in severe congestive heart failure. *N Engl J Med* 316: 1429-1435, 1987.
15. Sandoya E: ¿Por qué se necesitan grandes estudios randomizados? *Rev Urug Card* 4 (Suppl 1): 38A-41A.
16. Medical Research Council: Streptomycin treatment of tuberculous meningitis. *Lancet* 1: 582-586, 1948.
17. Abraham EP: Further observations on penicillin. *Lancet* 2: 177-180, 1941.
18. Haynes RB, McKibbon A, Fitzgerald D et al: How to keep up with the medical literature. I: Why try to keep up and how to get started. *Ann Inter Med* 105: 149-153, 1986.
19. TIMI Study Group: The thrombolysis in myocardial infarction (TIMI) trial. *N Engl J Med* 312: 932-936, 1985.
20. Six and twelve-month follow-up of the phase I TIMI trial. *Am J Cardiol* 62: 179-185, 1988.
21. Wilcox RG, von der Lippe G, Olsson CG et al: Angl-Scandinavian study of early thrombolysis (ASSET). *Lancet* 2: 525-530, 1988.
22. ISIS-2 (Second International Study on Infarct Survival) Collaborative Group: Randomized trial of intravenous streptokinase, oral aspirin, both or neither among 17.189 cases of suspected acute myocardial infarction. *Lancet* 2: 349-360, 1988.
23. Collen D, Topol E, Tiefenbrunn AJ et al: Analysis of coagulation and fibrinolysis during intravenous infusion of recombinant human tissue-type plasminogen activator in patients with acute myocardial infarction. *Circulation* 73: 511-517, 1986.
24. Medical Research Council Working Party: MRC trial of treatment of mild hypertension: principal results. *Br Med J* 291: 97-102, 1985.
25. Gifford RW: Review of the long-term controlled trials of usefulness of therapy or systemic hypertension. *Am J Cardiol* 63: 8B-16B, 1989.
26. Pardell H: Manual de Bioestadística. Masson, Barcelona, 1986.
27. Evans SJ, Mills P, Dawson J: The end of p value? *Br Heart J* 60: 177-180, 1988.
28. Norwegian Multicenter Study Group: Timolol-induced reduction in mortality and reinfarction in patients surviving acute myocardial infarction. *N Engl J Med* 304: 801-807, 1981.
29. ISIS-3 (Third International Study Infarction Survival): Protocol. ISIS Coordination Center, Oxford, 1989.
30. Coronary Artery Surgery Study (CASS): A randomized trial of coronary artery bypass surgery. *Circulation* 68: 939-951, 1983.
31. Maggioni AP, Franziosi MG, Tognoni G: Los estudios GISSI sobre infarto agudo de miocardio. Su lógica, diseño y resultados. *Rev Urug Card* 4: 83-91, 1989.
32. Cardiac Arrhythmia Suppression Trial (CAST): Preliminary report. *N Engl J Med* 321: 406-412, 1989.
33. Comité Organizador EMERA (Estudio Multicéntrico Estreptoquinasa República Argentina). *Rev Fed Argent Cardiol* 16: 238-240, 1987.
34. Fields WS, Maslenikov V, Meyer JS et al: Joint study of extracranial occlusion V. Progress report of prognosis following surgery or nonsurgical treatment for transient ischemic attacks and cervical carotid lesions. *JAMA* 211: 1993-2003, 1970.

Propuesta operativa

GIANNI TOGNONI

Director del Laboratorio de Farmacología Clínica, Instituto de Investigación Farmacológica "Mario Negri",
Miembro del Comité Organizador de los estudios GISSI

Dirección para separatas: Dr. Gianni Tognoni, Via Eritrea 62 - 20157 Milan - Fax (39-2) 33200049, Italia

Principios relativamente innovadores de actualización profesional y de metodología del juicio clínico como los que han sido discutidos en el trabajo que antecede pueden aparecer muy atractivos pero un poco abstractos y lejos de las exigencias y de las posibilidades de aplicación en la práctica diaria. La única manera de que eso no suceda es tomar este artículo, no como una revisión más, sino como una *invitación formal* a entrar en un ejercicio colectivo de reconocimiento experimental de la practicabilidad de los principios y de su relevancia.

Prácticamente se propone lo siguiente:

1) Quienes estén interesados en el problema pueden escribir y entrar en un ejercicio de evaluación, que consiste en el recibir una muestra de trabajos-modelos elegidos entre los más significativos de los últimos años en cardiología, a los que cada uno podrá evaluar, entregando después su respuesta en una ficha predetermi-

nada. Se puede interpretar esto como una "escuela pública abierta", la que permitiría aumentar dramáticamente la masa crítica de personas que pueden orientarse en el laberinto de la literatura con una autonomía cultural real, que es también el prerrequisito para crear un grupo creativo y crítico en la recepción de las novedades y en la planificación de investigaciones clínicamente importantes.

2) Establecer un "observatorio" de la calidad de la literatura cardiológica publicada en revistas latinoamericanas e internacionales durante un período experimental de un año, como trabajo colectivo que permita tener una idea directa de la calidad de la información disponible. Es posible imaginar una publicación periódica de este observatorio que permita a otros individuos repetir el ejercicio en lo que se refiere a la realidad que conocen y a la literatura que llega a sus manos.