

Juan O. Talavera,
Rodolfo Rivas-Ruiz

Investigación clínica VIII. Revisión estructurada de un artículo

Centro de Adiestramiento en Investigación Clínica,
Coordinación de Investigación en Salud,
Centro Médico Nacional Siglo XXI,
Instituto Mexicano del Seguro Social, Distrito Federal, México

Comunicación con: Juan O. Talavera.
Correo electrónico: jotalaverap@uaemex.mx

Resumen

Se han intentado varias estrategias para seleccionar un artículo bajo los supuestos de pertinencia y buena calidad, que en su mayoría dependen de la presencia o no de una serie de características y en otras ocasiones del juicio de quienes clasifican el artículo. Sin embargo, estas estrategias no nos permiten conocer la magnitud del error. Debido a que no hay artículo perfecto, resulta relevante identificar la magnitud del error y su impacto en el resultado final, para ello es necesario desarrollar destrezas que nos permitan revisar un artículo, identificar los posibles errores y generar una idea de su impacto sobre el resultado. Conforme la información contenida en las partes I a VII de esta serie de artículos sobre investigación clínica, hemos intentado mostrar su aplicación en una revisión estructurada de un artículo de causalidad, empezando por la revisión del estado basal, la maniobra y el resultado, con los errores sistemáticos (sesgos) que se generan en cada apartado, la pertinencia de la prueba, la oportunidad del tamaño de muestra y, finalmente, la relevancia clínica.

Palabras clave

artículo de revista
causalidad
estadística y datos numéricos
sesgos
tamaño de la muestra
medidas de asociación, exposición, riesgo o desenlace

Summary

There has been proposed several strategies to select an article under the assumption of quality and relevance. Mostly depend on the presence or absence of a number of features and others, depends of the judgment of other people. However, these strategies do not allow us to know the magnitude of the error contained within each article. And considering that there is not a perfect one, it is important to identify the magnitude of the error and the impact of this in the final result; consequently, it is necessary to develop skills that allow us to review a paper, identify potential errors, and generate a sense of the impact of it over the result. Using the information contained in Clinical Research Series I-VII, we have tried to show its application in a structured review of an article of causality beginning with the revision of the baseline, the maneuver and the outcome, with systematic errors (bias) generated in each section, followed by the appropriate of the statistical test and the sample size used, and finally, the clinical relevance.

Key words

journal article
causality
statistic and numerical data
bias
sample size
measures of association, exposure, risk or outcome

Introducción

Se han intentado varias estrategias para seleccionar un artículo bajo los supuestos de pertinencia y buena calidad, que en su mayoría dependen de la presencia o no de una serie de características y en otras ocasiones del juicio de quienes clasifican el artículo. Esto conlleva a una clasificación de adecuado o inadecuado, o en el mejor de los casos a una graduación de mayor a menor calidad o pertinencia. Sin embargo, estas estrategias

no nos permiten en realidad conocer la magnitud del error. Y debido a que no hay artículo perfecto, resulta relevante identificar la magnitud del error y el impacto que éste pudo haber tenido sobre el resultado final; para ello se vuelve necesario desarrollar destrezas que nos permitan revisar un artículo en forma estructurada, identificar los posibles errores y generar una idea del impacto de los mismos sobre el resultado. Es decir, no podemos depender de una clasificación ni del juicio de otros para decidir qué leer y qué no leer, o qué considerar ade-

cuado o no adecuado. Habrá que aprender la estructura base mínima que nos permita evaluar por nosotros mismos la pertinencia de cada artículo, sus errores y sus resultados.

De las partes I, III a VI de la presente serie sobre investigación clínica, hemos intentado mostrar las características que consideramos básicas para realizar una lectura y revisión estructurada de un artículo de causalidad (factor de riesgo o agente etiológico, pronóstico o tratamiento), una vez que se ha identificado el artículo a partir de una búsqueda sistemática (tema abordado en la parte VII). Comenzamos por un modelo integrado por el estado basal, la maniobra y el resultado (descrito en el artículo I), con los errores sistemáticos (sesgos) que se generan al definir y operar cada uno de estos apartados (artículo III). Y continuamos con la pertinencia de la prueba (parte IV), la pertinencia del tamaño de muestra (parte V) y, finalmente, la relevancia clínica (parte VI).

A continuación haremos un ejercicio sobre el uso de dicha información bajo una propuesta de revisión estructurada, para ello utilizaremos un artículo de nuestra propia autoría: *Reduction*

in the incidence of poststroke nosocomial pneumonia by using the "Turn-Mob" Program, publicado en *Journal of Stroke and Cerebrovascular Diseases* 2010;19:23-28. El objetivo del estudio fue demostrar la eficacia de un programa de movilización en cama llamado "turn-mob" para disminuir la incidencia de neumonía nosocomial en pacientes con evento vascular cerebral isquémico.

En la figura 1 encontramos las características del estado basal como la forma de selección de la prueba y la demarcación pronóstica; se puede apreciar que la aleatorización logró balancear las características de los grupos, con excepción de la enfermedad pulmonar obstructiva crónica, discretamente mayor en el grupo b (14 % *adversus* 7 %, $p = 0.088$), que pudo impactar en el resultado final. Dado que no se realizó un análisis estratificado, no es posible observar el efecto de cada maniobra según distintos factores de riesgo, con lo que el resultado se puede atribuir principalmente a las características promedio de la población estudiada.

En la figura 2 hay que considerar la calidad de la aplicación de la maniobra principal (programa *turn-mob* contra cam-

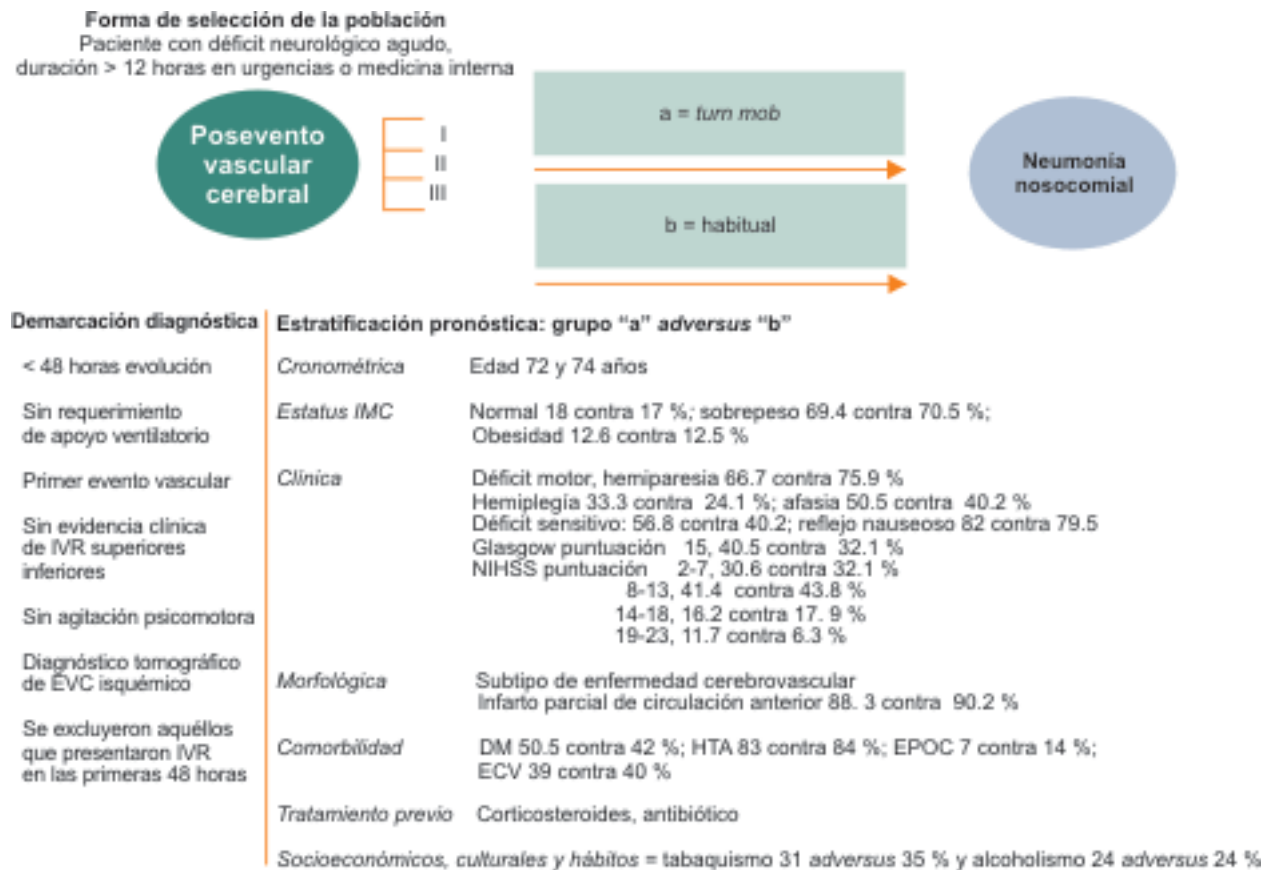


Figura 1 Características del estado basal: demarcación diagnóstica (criterios de selección) y estratificación (demarcación) pronóstica (variables que impactan en el desenlace independiente de la maniobra)
IVR = infección de vías respiratorias, EVC = evento vascular cerebral, IMC = índice de masa corporal, DM = diabetes mellitus, HTA = hipertensión arterial, EPOC = enfermedad pulmonar obstructiva crónica

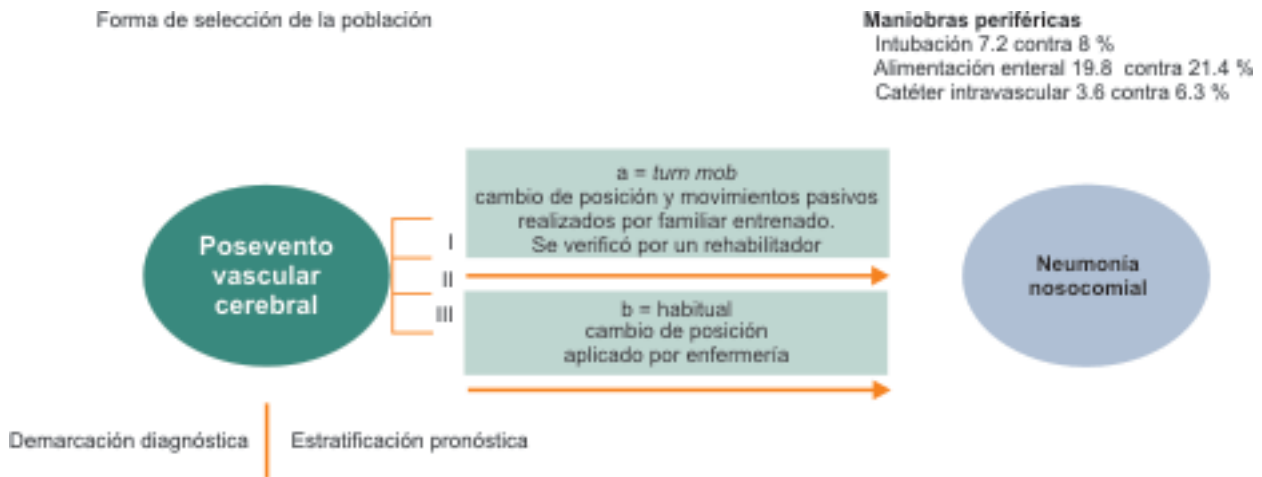


Figura 2 | Características por considerar durante la aplicación de la maniobra

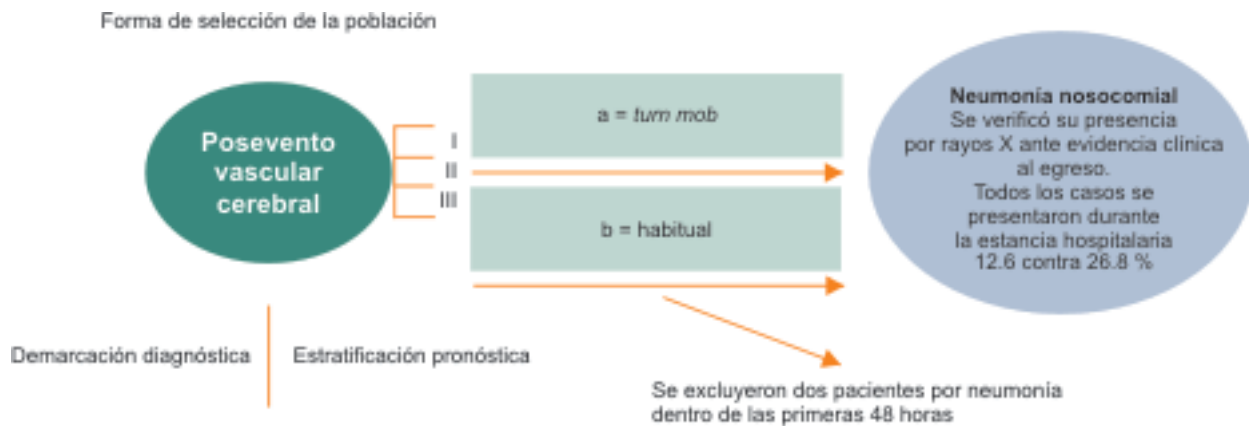


Figura 3 | Características por considerar en el resultado

bios de posición habitual) y verificar que se apliquen las maniobras periféricas en forma similar en ambos grupos.

Si bien no existieron diferencias en las maniobras periféricas, la aplicación del programa *turn-mob* fue estandarizada inicialmente y verificada día a día; por su parte, nunca se estandarizó ni se verificó la aplicación del tratamiento habitual, por lo cual no hay garantía de que se llevara a cabo; más aún, al egresar el paciente del hospital, en casa no contó con el apoyo de enfermería. Esto pudo representar más que una superioridad del programa *turn-mob* sobre el tratamiento habitual: el resultado de la aplicación del programa *turn-mob* contra nada.

En lo referente al desenlace, no existió la posibilidad de haber detectado en forma diferencial la presencia de neumonía nosocomial, dado que a todos los pacientes se les realizó rayos X de tórax al egreso o ante la mínima sospecha clínica. De igual forma no existió problema por pérdida de pacientes (sesgo de transferencia), ya que tan solo se excluyeron dos

pacientes de un total de 225 por la presencia de neumonía en las primeras 48 horas del ingreso hospitalario (figura 3).

Comentarios generales

Como comentario global del diseño metodológico y desarrollo del proyecto podemos decir que la selección de la población fue adecuada (ensamble adecuado), al considerar sujetos con alta probabilidad de desarrollar una neumonía nosocomial y en quienes era factible la aplicación del programa *turn-mob*. Se demostró que la distribución de distintos factores pronósticos fue similar entre los grupos, con lo que se evitó parcialmente el sesgo de susceptibilidad, dado que no se realizó un análisis estratificado que permitiera evaluar la maniobra en distintos grupos de riesgo (susceptibilidad pronóstica). En cuanto a la maniobra no se cuidó la adecuada ejecución de la maniobra habitual con lo que no podemos garantizar que no hubiese

existido un sesgo de ejecución. En la medición del desenlace fue igual para ambos grupos, lo que evita el sesgo de detección. Finalmente, no observamos pérdidas que pudiesen haber invertido la diferencia observada del desenlace entre los grupos (no sesgo de transferencia).

En relación con la prueba utilizada (tema desarrollado en la parte IV de esta serie sobre investigación clínica), la χ^2 muestra la comparación entre dos grupos de una variable de desenlace nominal como la presencia o no de neumonía nosocomial.

Por otro lado, si bien se demostró que no había diferencia entre la presencia de las distintas características y el grupo de tratamiento (prueba χ^2), hubiese resultado atractivo un ajuste multivariado del efecto del programa *turn-mob*, por las distintas características del estado basal y de las comaniobras que pudiesen haber impactado en el desenlace, en cuyo caso dado que es nominal, la prueba de regresión logística múltiple hubiese resultado adecuada.

En cuanto al tamaño de muestra (abordado en la parte V), no se menciona cómo se calculó, sin embargo, recordemos que este cálculo se lleva a cabo para obtener el número necesario de pacientes para demostrar que una diferencia esperada entre dos grupos es real y no al azar. En este caso, la diferencia clínica observada de 12.6 % *adversus* 26.8 % resultó estadísticamente significativa, con lo cual podemos asumir que es real dado que la probabilidad de que se deba al azar es menor a 5 % ($p < 0.05$). Y si bien no se describen los cálculos, con la incidencia mencionada en la introducción de 2 a 23 %, podemos estimar que se utilizó el valor mayor y se consideró una reducción directa de alrededor de 15 %, con lo que da un tamaño de muestra entre 90 y 103 sujetos por gru-

po (fórmula Fleiss-Kelsey) y si a esto agregamos 20 % obtenemos un valor alrededor de los 225 sujetos incluidos en el estudio (estimación de tamaño de muestra para diferencia de proporciones).

Finalmente, en general se consideraron como clínicamente relevante las diferencias directas mayores a 10 % o un NNT ≈ 10 (IC-VI). En este caso la diferencia fue de 14.2 y el NNT de 7.04 pacientes (que redondeado equivale a 8) para ver el beneficio en uno. Con estos resultados claramente podemos concluir que es clínicamente relevante.

Conclusiones

No podemos descartar la presencia de un sesgo de ejecución donde el tratamiento habitual no se hubiese realizado, en cuyo caso la conclusión no sería que el programa *turn-mob* es mejor que la movilización habitual realizada por enfermería, sino que se tendría que concluir que el programa *turn-mob* en un paciente posevto vascular cerebral isquémico es mejor que la no rotación ni movilización. Por otro lado, no podemos identificar si el programa *turn-mob* conserva su beneficio en distintos estratos de severidad del paciente dado que no se realizó un análisis estratificado, ni se ajustó a través de un análisis multivariado; es probable que no se hayan realizado dichos análisis por el tamaño de muestra, ya que 44 casos de neumonía nosocomial resultan insuficientes al estratificar o ajustar. Como podemos ver, todo estudio tiene errores, no obstante existe información valiosa pero para ponderarla resulta indispensable tener un conocimiento mínimo en investigación clínica.

Bibliografía

1. Talavera JO. Clinical research I. The importance of the research design. *Rev Med Inst Mex Seguro Soc* 2011;49(1): 53-58.
2. Talavera JO, Wachter-Rodarte NH, Rivas-Ruiz R. Clinical research II. Studying the process (the diagnosis test). *Rev Med Inst Mex Seguro Soc* 2011;49(2):163-170.
3. Talavera JO, Wachter-Rodarte NH, Rivas-Ruiz R. Clinical research III. The causality studies. *Rev Med Inst Mex Seguro Soc* 2011;49(3):289-294.
4. Talavera JO, Rivas-Ruiz R. Clinical research IV. Relevancy of the statistical test chosen. *Rev Med Inst Mex Seguro Soc* 2011;49(4):401-405.
5. Talavera JO, Rivas-Ruiz R. Clinical research V. Sample size. *Rev Med Inst Mex Seguro Soc* 2011;49(5):517-522.
6. Talavera JO, Rivas-Ruiz R. Clinical research VI. Clinical relevance. *Rev Med Inst Mex Seguro Soc* 2011;49(6):631-635.