

Estudios de investigación: usos, validez y sesgos

J GARCÍA DÍAZ*, C. ECHEVARRÍA RUIZ DE VARGAS* y M.^a J ZARCO JMÉNEZ***

*Médico Rehabilitador. FREMAP. Sevilla. **Médico Rehabilitador. Hospital Virgen del Rocío. Sevilla. ***Médico Rehabilitador. Hospital Puerta del Mar. Cádiz.

Resumen.—Exponemos la importancia de desarrollar estudios en medicina de rehabilitación desde la perspectiva del modelo de enfermedad de la OMS, realizamos reseñas expresas a la aportación teórica dada por autores desde la Investigación de los Resultados (Outcomes Research) en Medicina de Rehabilitación. Subrayamos la importancia de incorporar a la investigación en nuestra especialidad tanto las tecnologías médicas que valoren los déficit provocados por la enfermedad, como el uso de escalas de valoración que objetiven variables subjetivas como puedan ser el dolor o el grado de discapacidad o de nivel de handicap. Igualmente destacamos la importancia de comprender y conocer bien el significado de parámetros estadísticos sencillos, en ocasiones «despreciados» o minusvalorados, como son el tamaño muestral, el error alfa, el error beta, y los errores metodológicos, dado que estos parámetros condicionan la calidad de los estudios.

Palabras clave: *Validez. Sesgos. Investigación de resultados.*

RESEARCH STUDIES: USES, VALIDITY AND BIASES

Summary.—We explain the importance of developing studies in rehabilitation medicine from the point of view of the WHO disease model, and we make an express review of the theoretical contribution contributed by the authors from Outcome Research in Rehabilitation Medicine. We stress the importance of incorporating both the medical technologies that assess the deficit provoked by the disease as well as the use of assessment scales that discover the subjective variables such as pain or the incapacity grade or the handicap level into the investigation in our specialty. We also stress the importance of having a good understanding and knowledge of the meaning of the simple statistical parameters which are sometimes overlooked or underevaluated, such as sample size, alpha errors, beta error, methodological errors, given that these parameters condition the quality of the studies.

Key words: *Accuracy. Bias. Outcome Research.*

«Si os pregunto por qué creéis en una determinada cuestión, debéis darme alguna razón; o debéis admitir que vuestra creencia carece por completo de fundamentos.»

DAVID HUME

INTRODUCCIÓN

Muir Gray, uno de los precursores de la Medicina Basada en la Evidencia, define la investigación en salud como «un proceso de indagación que produce conocimientos» (1). La definición es sencilla, y esta sencillez didáctica es igualmente utilizada a la hora de clasificar o sistematizar, distinguiéndose en la literatura diferentes tipos de investigación (básica, aplicada y dirigida / cualitativa o cuantitativa / horizontal o vertical) (2). Sin embargo, la investigación en Salud es un fenómeno complejo, Rebaggio et al (3) nos indican como en Salud se produce la convergencia de diversidad de teorías y metodologías (Medicina, Farmacología, Economía, Sociología, Bioética, Bioestadística,...); y subrayan el hecho de que estas disciplinas, más que comunicarse y complementarse, tienden a competir entre sí por alcanzar «la razón» y el reconocimiento de las otras.

Un acercamiento global a la investigación se da desde la epidemiología por Battista et al (4), que nos propugnan una clasificación de la investigación desde tres criterios:

1. *Objeto del estudio:* se distinguen dos áreas, los estados de Salud y las intervenciones en Salud.

2. *Nivel de estudio:* distinguiendo dentro de los estados de Salud los niveles de molecular, tisular, individual y poblacional. Dentro del estudio de las intervenciones se señalan la técnica, práctica, procesos, programas y política sanitaria.

3. *Propósito del estudio,* a saber: desarrollo, descripción, explicación o evaluación; tanto del estado como de las intervenciones en Salud.

De la intersección de estos ejes surgen las diferentes áreas de investigación en Salud, como serían la investigación básica, la epidemiología Clínica, la Investigación de los Servicios Sanitarios y la Promoción de la Salud.

EL MODELO DE ENFERMEDAD

Este enfoque global, muy descriptivo para las disciplinas de la «medicina clásica», no es tan útil en Rehabilitación. Asumimos, que en nuestro caso es necesario hacer referencia al modelo de enfermedad de la OMS y a su separación en niveles (Pathology / Impairment / Disability / Handicap) (5). En este sentido, Whyte (6) propone distinguir por un lado el nivel de enfermedad en el que el clínico actúa y por otro el nivel de resultado o de «outcome» en el que repercute esta intervención. Ambos, intervención y resultado, son clasificados y evaluados a nivel de: patología, déficit, discapacidad o handicap.

El modelo de Whyte pone de manifiesto realidades conocidas por todos a la hora de enfrentarnos al problema de la investigación en medicina de rehabilitación:

1. Diversidad de niveles sobre los que puede actuar el médico rehabilitador, y por tanto, una multiplicidad de posibles variables Resultado de su práctica clínica.
2. No existencia de relaciones lineales entre el nivel de la intervención y el nivel de resultado en el que repercute la intervención. De modo, que aunque sólo se actúe sobre un nivel como la discapacidad, los resultados de la intervención pueden tener impacto en los cuatro niveles (patología, déficit, discapacidad y handicap).
3. Dificultad para crear condiciones «experimentales» o «controlables» en la investigación clínica, ya sea por la diversidad de variables que intervienen en el acto médico, o por la dificultad de crear grupos homogéneos de pacientes o por asignarlos aleatoriamente.

LA INVESTIGACIÓN DE LOS RESULTADOS

No obstante, se han realizado diversas propuestas a fin de intentar solventar los problemas de desarrollar estudios clínicos en nuestra especialidad. Estas van desde planteamientos sencillos, como son la creación de grandes bases de datos (p. e. CMBD) que incluyan valoraciones funcionales, las cuales serán básicas para la realización de estudios descriptivos e investigación de los «outcomes» (7), a desarrollos más complejos, como puedan ser los estudios «cuasi-experimentales», estudios clínicos basados en el «caso único», diseños basados en técnicas de análisis multivariante, algoritmos de «decisión clínica», estudios de individuos con déficit simples, o estudios basados en el tratamiento

específico de discapacidades o handicaps (6). Pero como base de estos estudios, hay un interés más enfocado en el análisis de la *efectividad* de los tratamientos e intervenciones, que en el simple estudio de la *eficacia* de los mismos (7); es decir, hay más interés por conocer y recoger que ocurre en la complejidad de nuestra práctica habitual, que en conseguir describir lo que ocurre en unas condiciones «ideales», propias de los ensayos clínicos, que aunque aportan un indudable grado de evidencia científica pueden estar alejados de la realidad de la práctica clínica.

Estos planteamientos que hemos apuntado brevemente aquí los hemos tomado de autores como M.V. Johnston, C. Granger, J Whyte, R.A. Keith... en una serie de artículos publicados en *American Journal of Physical Medicine and Rehabilitation*, denominados «outcomes research series». En ellos, estos autores exponen las bases de lo que ellos denominan la Investigación de los Resultados (outcomes), entendiendo aquí los resultados en un sentido amplio, es decir, como el estudio de la salud, la independencia funcional y calidad de vida tanto de los sujetos como de los grupos de individuos (6). Aunque este enfoque global de la investigación de los resultados no es «exclusivo» de este grupo de investigadores (8), ni mucho menos de los médicos rehabilitadores, nos llama la atención el esfuerzo que realizan por adaptar la investigación clínica y la gestión asistencial con un enfoque afín a nuestra especialidad: «Los estudios clínicos dentro de la metodología de la investigación de los resultados... no tendrán como objeto el resultado de una técnica sino la mejora del estado de salud y bienestar del paciente. Para cuantificar esta mejora tras nuestros tratamientos, se están desarrollando escalas para medir: calidad de vida, medidas de la funcionalidad, alivio del dolor y expectativas y satisfacción del paciente, y complicaciones» (9).

VARIABLES OBJETIVAS Y SUBJETIVAS

Argumentada ya por nuestra parte la necesidad de realizar diseños de investigación que tengan en cuenta la existencia de diferentes niveles o planos en la enfermedad, queremos ahora subrayar la necesidad de utilizar instrumentos adecuados para el tipo variable de resultado que se va registrar. Así como, remarcar la diferente metodología que requerirán estas variables a la hora de ser recogidas y manipuladas por el investigador. Básicamente se producen dos situaciones bien diferenciadas:

– En los problemas clínicos enfocados sobre el nivel de impairment (déficit y deformidades) las variables utilizadas pueden ser variables cuantitativas o categóricas, que pueden fácilmente objetivarse y manipularse de forma similar a otras variables biofísicas. Destaca aquí el papel que pueden aportar las «Nuevas Tecnologías» al

mejorar la precisión y calidad de las medidas objetivas ej. estudios isocinéticos, pruebas de neuroimagen, pruebas funcionales de esfuerzo, sistemas de valoración del equilibrio... (la valoración de estas nuevas tecnologías, es objeto de un artículo de este monográfico).

— Por otro lado cuantificar los logros obtenidos en los terrenos del dolor, la discapacidad y el handicap requieren de la aplicación de técnicas «cualitativas» mediante cuestionarios o escalas de valoración. Su manejo necesita de metodologías más complejas, desarrolladas en las áreas de la psicometría y la educación (10, 11). Pero dado el uso y sistematización con que estos instrumentos comienzan a utilizarse, es necesario por parte del médico cierto conocimiento sobre la calidad de medida que le aportan estas herramientas y del modo en que ha de utilizar los resultados en su práctica clínica (12-16).

VALIDEZ, PRECISIÓN Y SESGOS

Con independencia del tipo de variable de nuestro estudio, es necesario que su diseño nos permita reflejar el problema que nos ocupa de un modo efectivo, es decir, necesitamos que nuestro estudio posea *Validez Interna*. Si además, los resultados son reproducibles y extensibles para poblaciones distintas a la estudiada estamos diciendo que posee también *Validez Externa* (17).

Asociados a la validez interna tenemos otros dos conceptos importantes. El primero de ellos, es el de

error aleatorio o *precisión* del estudio y el segundo, es el concepto de error sistemático o *sesgo*.

Los errores aleatorios son debidos al azar y estarán presentes en cualquier estudio que realicemos sobre una muestra, y no lo estarán por definición, cuando utilizemos a toda la población en estudio. No son errores predecibles. No pueden ser eliminados, pero sí pueden ser estimados estadísticamente (18). Es importante señalar también, que los *errores aleatorios* pueden controlarse y por lo tanto tenemos una vía para mejorar la precisión de los estudios (12, 18). Son cuatro los factores a tener en cuenta como datos importantes para definir el grado de precisión que podemos alcanzar en nuestro estudio: tamaño muestral, error tipo I o alfa, errores tipo II o beta y potencia estadística del test.

Ottenbacher (19), ha revisado el papel de estos parámetros en los estudios de nuestra especialidad aportando las siguientes conclusiones: 1) la significación estadística es confundida frecuentemente con la significación clínica. 2) Los estudios actuales aportan un buen nivel de error tipo I, pero descuidan los errores tipo II, y con ello, los tests tienen una potencia muy baja. 3) La baja potencia de los tests suele asociarse a una dispersión en los resultados y por lo tanto la reproducibilidad de los estudios es baja (validez externa). La consecuencia es un grado de evidencia en los trabajos de investigación bajo, y por tanto es difícil alcanzar un consenso entre los autores, se dificulta por tanto la elaboración de protocolos y guías de práctica clínica sobre una evidencia científica, que por lo comentado, o no existe o es escasa.

TABLA 1. Clasificación de Sesgos (17, 18).

Sesgos		
Sesgos de selección	Selección inadecuada de la muestra. Si utilizamos otra muestra de la misma población las estimaciones obtenidas serían diferentes.	Sesgo del voluntario o de auto selección. Sesgo de «no respuesta». Sesgo del «trabajador sano». Sesgo de sospecha diagnóstica o «sesgo de detección». Sesgo de Berkson. Falacia de Neyman.
Sesgos de información o medida	Las medidas no son válidas, por defecto de los instrumentos de medida o la fiabilidad inter-observador es baja. Otro efectos pueden ser causados por la «memoria», o por los «diagnósticos» previos.	Mala clasificación no diferencial. Mala clasificación diferencial. — Sesgos amnésico o de recuerdo. — Sesgo de atención o efecto «Hawthorne». — Sesgo del entrevistador. Regresión a la media. Falacia ecológica.
Sesgo de confusión	Se atribuye un efecto a una variable, factor de confusión, cuando en realidad no es debido a ella. Se produce cuando la variable que denominamos f. de confusión se asocia de modo independiente con un factor de riesgo de la enfermedad o del efecto que se busca. Además, este factor de confusión no es un eslabón intermedio en la vía causal entre los factores de riesgo y la enfermedad o efecto. Ej. el alcohol con f. confusión con el tabaco en el cáncer de lengua-boca.	

Por otro lado, tenemos los denominados errores metodológicos o sesgos (12, 13, 30). Este tipo de error cuestiona la validez de los estudios realizados, y suele agruparse en tres categorías: a) sesgos de selección, b) sesgos de clasificación o de información y c) sesgos de confusión. Los hemos resumido en la tabla 1, más allá de la clasificación no existen reglas, son como hemos comentado errores metodológicos recopilados sobre el trabajo de otros autores.

Pero, ¿pueden controlarse los sesgos? Básicamente, sólo pueden evitarse o prevenirse durante la fase de diseño, de ahí la necesidad de conocerlos, porque su estudio se basa más en el conocimiento de «errores» que otros han cometido antes, que en reglas para evitarlos. No obstante, los *sesgos de selección y clasificación* pueden disminuirse mediante técnicas de enmascaramiento o de «ciego». Para el control de los *sesgos de confusión* pueden emplearse: a) en la fase de diseño técnicas de restricción y técnicas de emparejamiento. b) En la fase de análisis de los datos, pueden utilizarse técnicas de estratificación y de análisis multivariante. Pero se requiere, siempre, una sospecha previa sobre la variable que puede ejercer el efecto de confusión.

La palabra sesgo, es una palabra ampliamente utilizada en la literatura, y en este sentido es una palabra que ha hecho «fortuna» y por tanto pierde significación. Así por ejemplo, en un contexto diferente al comentado podemos encontrar el denominado *sesgo fraudulento*, cuando el autor modifica voluntariamente los datos para obtener un resultado, o el *sesgo de publicación*, condicionado por la tendencia en la literatura a publicar los estudios que encuentran significación estadística en la hipótesis de estudio.

A modo de conclusión comentar que el diseño de los estudios en nuestra especialidad debe de plantearse sobre la existencia de diferentes planos en el estatus de enfermedad del paciente. Indicar que los instrumentos para objetivar y medir las variables en cada nivel son diferentes, lo que obliga en ocasiones al manejo de técnicas estadísticas complejas en el análisis, lo que no ha de ser excusa para cometer «errores» con parámetros estadísticos más sencillos, pero que pueden definir la calidad del estudio.

BIBLIOGRAFÍA

- Muir Gray JA. Atención Sanitaria Basada en la Evidencia. Madrid: Churchill Livingstone; 1997.
- Peña A, Lillo MJ, Lafuente G, Vara MT, Rodríguez I. Importancia de la investigación en Rehabilitación. Rehabilitación (Madr) 1997;31:5-11.
- Rebagliato M, Ruiz I, Arranz M. Metodología de investigación en epidemiología. Madrid: Díaz de Santos; 1996.
- Battista RN, Contrandriopoulos AP, Champagne F, Williams J, Pineault R, Boyle P. An integrative framework for health-related research. Comentado en Rebagliato M, Ruiz I, Arranz M. Metodología de investigación en epidemiología. Madrid: Díaz de Santos; 1996.
- SERMEFF. Estado Actual de la Rehabilitación en España. Formación de especialistas. Madrid: Garsi, Masson; 1996.
- Whyte J. Toward a methodology for Rehabilitation research. Am J Phys Med Rehabil 1994;73:428-35.
- Johnston MV, Granger CV. Outcomes research in medical rehabilitation. Am J Phys Med Rehabil 1994;73:296-303.
- Peiró Moreno S. Medidas de actividad y producto sanitario. En: Del Llano Señaris J, Ortún Rubio V, Martín Moreno JM, Millán Núñez-Cortés J, Badía Gené J, directores. Gestión Sanitaria. Innovaciones y desafíos. Barcelona: Masson; 1998.
- Vaquero Martín J. Investigación clínica y estudio de resultados en cirugía y ortopédica. MAPFRE Med 1996;7:53-7.
- Streiner DL, Norman G. Health measurement scales. A practical guide to their development and use. 2nd ed. New York: Oxford Medical Publications; 1995.
- Granger CV, Gresham GE. New developments in functional assessment. Phys Med Rehab Clin North Am 1993;4:417-609.
- Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of Misinference. Arch Phys Med Rehabil 1989;70:308-12.
- Wright BD, Linacre JM. Observation are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857-60.
- Silverstein B, Kilgore KM, Ficher WP, Harley JP, Harvey RF. Applying psychometric criteria to functional assessment in medical rehabilitation: I. Exploring unidimensionality. Arch Phys Med Rehabil 1991;72:631-7.
- Silverstein B, Kilgore KM, Ficher WP, Harley JP, Harvey RF. Applying psychometric criteria to functional assessment in medical rehabilitation: II. Defining interval measures. Arch Phys Med Rehabil 1992;73:507-18.
- Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000;81(12 Supl 2):S15-20.
- Álvarez Cáceres R. Validez de los estudios científicos. Precisión y Sesgos. En: Álvarez Cáceres R. El Método Científico en las Ciencias de la Salud. Madrid: Díaz de Santos; 1996.
- Martínez González MA. Sesgos y error aleatorio. MAPFRE Med 1996;7(Supl IV):69-77.
- Ottenbacher KJ. Why Rehabilitation research does not work (as well as we think it should). Arch Phys Med Rehabil 1976;121-9.
- Silva Ayçaguer LC. Cultura estadística e investigación científica en el campo de la salud: una mirada crítica. Madrid: Díaz de Santos; 1997.

Correspondencia:

Jesús García Díaz
Trastámara nº 19; 2º B
41001 Sevilla