

Centro de Investigaciones y Referencias de Aterosclerosis de La Habana (CIRAH)

La ética y el método: los dos pilares básicos en la conducción de un ensayo clínico

Ethic and Method. Two basic miles stones in conduction of clinical assays

Jorge Bacallao Gallestey

Profesor titular de la Universidad de Ciencias Médicas de La Habana.
jbacallao@infomed.sld.cu

La mayor amenaza a la credibilidad y el prestigio de la ciencia moderna es el sesgo. Aunque normalmente -y con sobradas razones- se atribuye el sesgo a causas metodológicas, los componentes ético y axiológico también se hacen presentes en sus orígenes. Resulta hasta cierto punto paradójico que una actividad creativa organizada para el manejo racional del escepticismo, deba sostenerse sobre puntales basados en la confianza. Es por esa razón, que las violaciones a esa confianza con conductas impropias o con el fraude abierto, constituyan una grave amenaza a los fundamentos de la ciencia. Fuchs y Westervelt¹ aportan muy bien logrados argumentos sobre esta cuestión.

Muchas voces recientes dudan de la capacidad de la comunidad científica para garantizar la integridad en el desarrollo de la investigación. Los ensayos clínicos tampoco han quedado libres de un escrutinio suspicaz. Este giro drástico a partir de una cándida confianza, no en la infalibilidad de la actividad científica, pero sí en su desprejuicio y su integridad en la búsqueda de conclusiones fiables, es la consecuencia de visibles instancias de conductas impropias y la cambiante percepción pública sobre los procesos de investigación científica. A la par de su magnitud y de la influencia de la ciencia como empresa, han crecido las presiones económicas y políticas sobre los científicos, que ponen en tensión conflictiva, por un lado, el apego a la neutralidad (no en los fines, sino en el método) que forma parte de la sustancia misma de la ciencia, y, por la otra, el apremio por responder a encargos sociales cada vez más acuciantes.^{2,3} La urgencia por comercializar los resultados de la investigación (particularmente los de la industria farmacéutica), aun aquella que se genera en los ámbitos académicos, conlleva demandas cada vez más estrictas de revisión y escrutinio.⁴

Este estado de cosas no solo incrementa la conciencia pública, sino que contribuye a afinar los criterios y a rigORIZAR los estándares de calidad en los ensayos clínicos, tanto o más importante que en la ciencia en general, y ello por varias razones: (a) porque por su propia naturaleza hay más oportunidades de autocorrección en la investigación básica que en los ensayos clínicos, debido a que los hallazgos experimentales se replican de forma independiente; (b) porque los ensayos clínicos tienen mayores y más inmediatas consecuencias para la salud y el bienestar públicos y (c) porque los ensayos clínicos tienen una visibilidad pública mayor que la investigación básica.

El derecho a conducir un ensayo clínico es, en todo caso, un privilegio que conceden la sociedad y los pacientes participantes en el ensayo a un grupo de investigadores a partir de la confianza en su integridad y honestidad, por lo que la fabricación, el plagio, la adulteración o cualquiera de los estigmas y formas de inconducta son acciones francamente violatorias de esa confianza.

En relación con otras actividades creativas como el arte, la música o la literatura, la ciencia es singular en el sentido de que su esencia se agota en el método. El método científico *es* la integridad de la disciplina. El concepto de ética científica se ha visto distorsionado por la creencia de que la ciencia está compuesta por un canon fijo de reglas de calidad. La ética científica está mucho más determinada por la estructura del proceso que por sus resultados.⁵ En los ensayos clínicos, emerge además un componente ético en la relación del investigador con el paciente, que no siempre armoniza con los propósitos científicos o cognoscitivos.

Los ensayos clínicos son la regla de oro para establecer la efectividad de un tratamiento. El paradigma ideal de un ensayo clínico (EC) es el ensayo prospectivo, con asignación aleatoria, controlado, y doble ciego.

La necesidad de que sea prospectivo es obvia: puesto que el propósito del EC es conectar causalmente el tratamiento con ciertos efectos esperados, la precedencia temporal es una condición necesaria (aunque no suficiente).

La aleatorización es el instrumento idóneo para minimizar el sesgo –posiblemente el único, puesto que opera igualmente sobre confusores observables y no observables. A menudo se cree que la aleatorización se realiza con el fin de balancear los brazos del ensayo con respecto a la presencia de confusores. El balance de los grupos es una consecuencia habitual de la aleatorización, pero no es su propósito. Es parte del lenguaje habitual del estadístico, no así del investigador biomédico, la referencia a la minimización del sesgo como propósito "a la larga": opera como parte de una racionalidad que garantiza el mejor ajuste posible entre la realidad y su apreciación a largo plazo, del mismo modo en que apostar porque en la próxima tirada de un dado equilibrado "no saldrá un 5" es una estrategia exitosa en un número grande de tiradas, pero poco o nada dice acerca del próximo lanzamiento.

La tabla 1 en la publicación de un ensayo clínico suele contener una comparación entre los dos grupos resultantes de la asignación aleatoria, para confirmar que los grupos están balanceados con respecto a características basales relevantes y que, por consiguiente, la aleatorización dio sus resultados esperados. Esta conducta es incorrecta. Cuando para alguno de los potenciales confusores los grupos no resultan balanceados, el investigador debe aplicar controles analíticos para remover su efecto, pero aun en tal eventualidad, la aleatorización cumplió cabalmente su función primordial.

Tabla 1. Frecuencias observadas de respuesta (Éxito=1; Fracaso=0) a un tratamiento dicotómico T, en presencia de un confusor C

Y	C = 1		C = 0	
	T = 1	T = 0	T = 1	T = 0
Y = 1	150	40	30	60
Y = 0	250	30	50	20
Total	400	70	80	80

Asignar un paciente a uno de los dos brazos del ensayo, no implica necesariamente su adherencia a esta asignación. La aleatorización puede considerarse una variable instrumental (en el caso más simple con dos resultados posibles) que tiene una correlación alta, pero no necesariamente igual a 1 con el grupo en el que finalmente se ubica el paciente.⁶ El 100% del efecto de la asignación sobre el paciente está mediado por su ubicación final en uno de los brazos del ensayo, y la asignación *per se*, debido a que es el resultado de un mecanismo aleatorio, no se asocia a ningún confusor, observable o no, salvo por puro azar. Estas dos condiciones definen a las variables instrumentales.

El uso de controles se explica modernamente bajo la lógica del paradigma contrafactual de causalidad. Sin pretensiones de rigor formal, el efecto de un tratamiento sobre un sujeto se define como la diferencia entre el resultado con y sin el tratamiento. Puesto que recibir el tratamiento y no recibirlo son sucesos mutuamente excluyentes, uno de los términos de la diferencia es inevitablemente un dato faltante, comúnmente denominado contrafactual (el que no ocurrió). Evaluar la efectividad de un tratamiento significa estimar la diferencia entre el resultado con y sin el tratamiento, lo cual implica elegir adecuadamente un sucedáneo del contrafactual, es decir, un grupo que se habría comportado como el contrafactual en ambas ramas del ensayo.

Por último, el ensayo debe ser doble ciego, para evitar el sesgo en la calificación de los resultados. Sin embargo, es relativamente frecuente encontrar en la literatura científica, informes técnicos de ensayos cuyo diseño y factura se aparta de este paradigma ideal.

Dentro de la vasta temática que desborda los propósitos y el formato de este texto, se hará breve referencia a dos aspectos del **ámbito metodológico** (los modelos estructurales marginales para el tratamiento analítico de los regímenes de tratamiento dinámico y el sesgo de medición en los ensayos clínicos) y a otros tantos aspectos en **la dimensión ética** (el sesgo del deseo y el dilema ético de los ensayos con potencia insuficiente).

La dimensión metodológica

Regímenes de tratamiento dinámico: Un caso bien interesante es el los llamados "regímenes de tratamiento dinámico", que aunque se presentan en el ámbito de los EC, también e incluso con mayor frecuencia emergen en los estudios observacionales. Como su nombre sugiere, en los regímenes de tratamiento dinámico (DTR, por sus siglas en inglés) el tratamiento cambia en el curso del ensayo. Una razón para ese cambio puede ser, por ejemplo, un deterioro en el estado del paciente.

Varios estudios aleatorizados y observacionales han demostrado la efectividad de los antirretrovirales para incrementar el tiempo libre del Síndrome de Inmunodeficiencia Adquirida (SIDA) en pacientes portadores del virus de inmunodeficiencia humana (VIH).^{7,8} El tiempo óptimo para iniciar la terapia antirretroviral depende del conteo de CD4. Las normas vigentes aconsejan el inicio de la terapia cuando el conteo cae por debajo de 350 células/ μ l. Este es un caso típico de DTR en el que el tratamiento depende de una covariante asociada al tiempo que a su vez se asocia con la variable de respuesta que permite medir el éxito del tratamiento, lo cual la convierte en un confusor, es decir, una variable que se asocia tanto con el tratamiento como con la variable de respuesta. Lo característico de estos casos es que, por una parte, la covariante dependiente del tiempo predice la variable de respuesta y también el curso posterior del tratamiento y por la otra, que la trayectoria temporal del tratamiento predice el valor de la covariante. A las covariantes que satisfacen estas condiciones se

les denomina *confusores dependientes del tiempo*. Un error frecuente es aplicar el principio de la "intención de tratamiento" en presencia de este tipo de confusores.

El enfoque analítico moderno en este tipo de ensayos, cada vez más frecuentes, es el uso de modelos estructurales marginales.⁹ Se denominan estructurales los modelos que involucran una variable de respuesta que contiene un elemento contrafactual. Los modelos estructurales marginales (MEM) constituyen una nueva clase de modelos causales para la estimación del efecto causal de un tratamiento en presencia de covariantes dependientes del tiempo que pueden ser simultáneamente confusores y variables intermedias.^{10,11} El recurso de la ponderación por el inverso de la probabilidad produce estimaciones insesgadas, como se aprecia claramente en el ejemplo ficticio siguiente:

Sean T un tratamiento, Y una variable de respuesta y C un confusor, a todos los cuales consideraremos dicotómicos por razones de simplicidad. Supongamos que los resultados de un estudio son los que muestra la tabla 1.

Tabla 1. Frecuencias observadas de respuesta (Éxito=1; Fracaso=0) a un tratamiento dicotómico T, en presencia de un confusor C

Y	C = 1		C = 0	
	T = 1	T = 0	T = 1	T = 0
Y = 1	150	40	30	60
Y = 0	250	30	50	20
Total	400	70	80	80

La tabla 2 ilustra el cómputo de las ponderaciones por el inverso de la probabilidad y la consiguiente construcción de una pseudo-población en la que las estimaciones del efecto del tratamiento están libres de sesgo.

Tabla 2. Datos observados (N), probabilidades condicionales (P), ponderaciones (W) y valores ponderados de N (N1)

C	T	Y	N	P (T/C)	W = 1/p	N1
1	1	1	150	0,85	1,18	177
1	1	0	250	0,85	1,18	295
1	0	1	40	0,15	6,67	267
1	0	0	30	0,15	6,67	200
0	1	1	30	0,5	2	60
0	1	0	50	0,5	2	100
0	0	1	60	0,5	2	120
0	0	0	20	0,5	2	40

En la tabla 3, las observaciones de la Tabla 1 se han sustituido por observaciones ponderadas por el inverso de la probabilidad, que dan lugar a estimaciones libres del efecto confusor de C, como se advertirá más adelante.

Tabla 3. Frecuencias ponderadas de respuesta

Y	C = 1		C = 0	
	T = 1	T = 0	T = 1	T = 0
Y = 1	177	267	60	120
Y = 0	295	200	100	40
Total	472	467	160	160

Si ahora se colapsan las dos mitades de la Tabla 3, ignorando la existencia del confusor obtendríamos la siguiente tabla (tabla 4) que corresponde a una nueva población ficticia.

Tabla 4. Frecuencias observadas de la Tabla 3 colapsada con respecto al confusor C

Y	T = 1	T = 0
Y = 1	237	387
Y = 0	395	240
Total	632	627

Si se calcula, por ejemplo, el "odds ratio" a partir de los datos de la Tabla 4 (que ignora la existencia del confusor C) se obtiene:

$$OR = \frac{237 \cdot 240}{395 \cdot 387} = 0,372$$

Este valor coincide con el que arroja el cálculo a partir de la Tabla no colapsada.

En ensayos clínicos DTR la probabilidad que se usa como base para la ponderación es la probabilidad de la trayectoria temporal específica de tratamientos que ha recibido el paciente en el curso del ensayo clínico.

El sesgo de medición en los ensayos clínicos

Uno de los problemas más críticos en la ciencia –los ensayos clínicos no son una excepción– es el sesgo de medición. En el campo de la inferencia causal, no medir adecuadamente la variable de respuesta, los factores de exposición o las covariantes, es letal para la estimación de los efectos causales. Los estadísticos suelen desentenderse de este serio problema, asumiendo que los datos que reciben son válidos y fiables, y que los investigadores son los responsables de garantizar su calidad. El espectro de errores posibles en la instancia de la medición es muy amplio para ser tratado en un texto con formato editorial, por lo que me limitaré a exponer unos pocos de los más frecuentes.

- Tratar datos ordinales como si fuesen nominales. Consideremos, por ejemplo, una variable de respuesta ordinal con cuatro valores posibles y un ensayo clínico con dos ramas. Si para la comparación de los grupos se empleara simplemente una tabla de contingencia de 2 x 4, o un modelo log-lineal, se

perdería mucha sensibilidad o potencia para detectar diferencias. Si la respuesta al tratamiento puede describirse como "baja", "moderada", "buena" o "excelente" y la expectativa natural es mover los pacientes de los niveles bajos a los niveles altos de la escala, disponer las observaciones en una tabla de contingencia puede arrojar $4! = 24$ permutaciones posibles de las columnas, todas las cuales arrojarían el mismo valor de un estadígrafo ji-Cuadrado de prueba, cuando en rigor solo unas pocas son estrictamente consistentes con el efecto esperado del tratamiento. Desde la publicación del trabajo de McCullagh¹² sobre el análisis de variables ordinales, se conoce que el modelo de los riesgos o de los "odds" proporcionales es la vía idónea para el tratamiento de estos datos. En estudios observacionales que suelen ser de carácter exploratorio y tienen el fin de generar o a lo sumo confirmar hipótesis, y no de verificarlas, el problema apenas se hace sentir, porque el análisis descansa en la descripción y la interpretación; no así en los ensayos clínicos, que tienen el propósito explícito de probar hipótesis de efectividad, superioridad o no inferioridad.^{13,14}

- La codificación de datos ordinales: la codificación "natural" de los datos ordinales no necesariamente coincide con la codificación óptima. Una terapia inmunológica contra el cáncer podría ser efectiva en pacientes en estadios 1 y 2, pero no serlo en pacientes en estadios superiores. Un procedimiento de prueba podría no detectar diferencias para una codificación politómica de la variable de respuesta y, por el contrario, sí detectarlas para otra codificación politómica o para una codificación binaria de la variable. Lo ideal sería poder construir el modelo para la prueba de hipótesis a partir de codificaciones formuladas y justificadas *a priori*, pero nada impide realizar estratificaciones *post hoc* exploratorias y utilizar procedimientos de escalamiento óptimo para maximizar el efecto del tratamiento sobre la variable de respuesta, siempre que se realice con una perspectiva exploratoria y no como parte de un proceso de manipulación *a posteriori* a la caza de significación estadística.¹⁵
- Dicotomización de datos continuos u ordinales: Una práctica generalizada entre los investigadores consiste en generar dicotomías en el recorrido de las variables ordinales a partir de puntos de corte arbitrarios,¹⁶ pese a que como regla general, esta práctica conlleva pérdida de información. Una suplementación a embarazadas puede incrementar el peso de los recién nacidos sin reducir la tasa de bajo peso al nacimiento, o un factor de exposición puede reducir el peso al nacimiento, sin llegar a incrementar la tasa de bajo peso.¹⁷
- El número necesario a tratar (NNT): el NNT es el recíproco de la reducción en el riesgo asociada a utilizar una de las ramas del ensayo en relación con la otra. Conceptualmente es inobjetable como medida de efecto cuando se aplica a verdaderas dicotomías y no a dicotomías arbitrarias impuestas a variables de respuesta ordinales o continuas; sin embargo, no es una buena medida descriptiva de los resultados de un ensayo clínico debido a que la población discrecional de la cual se obtienen los casos para el ensayo no tiene -o no hay cómo verificar que tenga- el mismo riesgo de base que la población virtual que es objeto de las inferencias. Este es un hecho clave porque la interpretación del NNT presupone que los riesgos en los grupos que se comparan -o al menos el riesgo basal- es conocido.¹⁸⁻²³ Aunque la aplicación del NNT es controversial y se puede encontrar tanto defensores como detractores, lo que parece inaceptable y fuera de toda discusión es la construcción artificial de dicotomías sólo con el fin de calcularlo.

La dimensión ética

El sesgo del deseo

El plagio, la adulteración o la fabricación de datos son formas descarnadas de fraude que violan los códigos éticos más elementales en los ensayos clínicos y la ciencia, en general. Sin embargo, hay otras formas "más inocentes" pero igualmente perniciosas y generadoras de ciertas modalidades de sesgo, que en la literatura se conocen como sesgo del deseo o sesgo interpretación. Es importante distinguir entre el sesgo intencional que adultera u oculta, y el sesgo no intencional que se deriva de los juicios *a priori*.²⁴⁻²⁶ En un trabajo reciente²⁷ en el que se reportan los resultados de un ensayo clínico para comparar dos analgésicos opioides para el manejo del dolor intenso de origen neuropático, sus autores incurren en varios de estos errores atribuibles a una falta de neutralidad metodológica que produce una falta de balance entre las ramas del ensayo. Entre esos errores se cuenta el uso de un diseño abierto (no doble ciego) combinado con el uso de un brazo de rescate asimétrico, que bajo ciertas condiciones asociadas a la evolución del dolor consiente el cambio en una dirección, pero no en la otra. El sesgo del deseo comporta por lo general la adopción de conductas no neutrales en la concepción y ejecución del diseño, como por ejemplo, no observar estrictamente las condiciones fármaco-dinámicas y fármaco-cinéticas, propias de los tratamientos en ambas ramas del ensayo. Pese a que la interpretación es una acción esencialmente subjetiva, el sesgo de interpretación implica una apreciación distorsionada de los datos del ensayo, a partir de juicios *a priori* sobredimensionados que se hacen prevalecer por encima de la evidencia objetiva.

El dilema ético de los ensayos con potencia insuficiente

La postura que, en general, han compartido especialistas dentro del campo de los ensayos clínicos y la Bioética es el rechazo a los ensayos con tamaños muestrales insuficientes para garantizar la debida sensibilidad en la detección de efectos relevantes.²⁸⁻²⁹ Sin embargo, la práctica no solo prevalece sino que ha sido apuntalada básicamente con los dos argumentos siguientes: (a) que la realización cada vez más extendida de los meta-análisis permite rescatar ensayos pequeños mediante recursos técnicos para combinar sus resultados con los de otros estudios en la estimación de la eficacia de un tratamiento y (b) que aunque un ensayo sea insuficiente para poner a prueba ciertas hipótesis, permite, sin embargo, calcular intervalos de confianza para la estimación de efectos, que luego pueden combinarse y arrojar intervalos más precisos, justamente mediante los recursos analíticos que aportan las técnicas de revisión sistemática y el meta-análisis. Bacchetti, *et al*,³⁰ a partir de consideraciones de riesgo-beneficio sostienen que los comités de ética tendrían que preguntarse si un ensayo es demasiado grande en lugar de si es demasiado pequeño. Es innegable que las transnacionales farmacéuticas suelen realizar mega-ensayos que abren el mercado a productos que no superan los estándares terapéuticos ya establecidos en magnitudes clínicamente relevantes, aunque estadísticamente significativas debido a la inflación artificial del tamaño de los ensayos.

Las antinomias que enfrentan en el terreno de los valores a una ética proyectada hacia el individuo y otra proyectada hacia las poblaciones, y que emerge en el diseño y ejecución de los ensayos clínicos en cualquiera de sus fases, se han examinado más extensamente en un artículo anterior.³¹

Reflexiones finales

La ciencia médica contemporánea discurre por dos caminos cuyas directrices axiológicas, derroteros metodológicos y valores epistemológicos parecen cada vez

más distinguibles. Uno de esos caminos es el preferido de los investigadores que se empeñan en el descubrimiento y la explicación de las causas de las enfermedades. Los descubrimientos son a menudo el resultado de la conjunción de eventos que parecen guiados por el azar: el curso inusitado de una enfermedad en un paciente, los resultados inesperados de un experimento, el comportamiento atípico de algunos grupos poblacionales o la convergencia de varios artículos en escenarios afines de la literatura científica. Los investigadores formulan conjeturas y tratan de confirmarlas en su compatibilidad con la observación sistemática. Cuando los datos se alinean con las conjeturas aparece una nueva publicación. La siguiente oleada de investigadores impone controles más rigurosos a los procedimientos de prueba: estratifican a los pacientes según diferentes criterios, hacen variar en términos cuantitativos y cualitativos los factores de exposición, examinan la influencia de potenciales confusores o varían las condiciones del laboratorio. Aplican, en otras palabras, el paradigma epistemológico de las conjeturas y refutaciones que conducen a la extinción o al fortalecimiento de las primeras y al diseño de nuevos estudios con controles analíticos cada vez más rigurosos.

El otro camino es el que siguen los investigadores cuyo fin es diseñar e implementar estudios para evaluar si una terapéutica o un medio diagnóstico tiene la efectividad que se les atribuye. Se trata de los cultivadores de la investigación evaluativa cuyo ejemplar más paradigmático son los ensayos clínicos aleatorizados. La principal condición para la credibilidad de estos ensayos es la planeación rigurosa y explícita de cada etapa y cada aspecto del ensayo.³² En ese proceso, emergen aspectos éticos y metodológicos que este texto ha examinado muy someramente, a partir de una selección personal del autor. En el plano metodológico, se ha llamado la atención sobre los regímenes de tratamiento dinámico que hasta hace muy poco se enfrentaban con herramientas analíticas muy insuficientes que hoy se han enriquecido con nuevos modelos y procedimientos más refinados de estimación; también se ha puesto énfasis en algunos puntos críticos relacionados con los sesgos de medición, que no son privativos de los ensayos clínicos, pero que emergen allí con lamentable frecuencia, debido a la importancia secundaria que se les atribuye en comparación con otras aristas metodológicas de los ensayos. En el plano ético, se han considerado brevemente dos temas críticos: el tamaño de los ensayos que remite a antinomias que se presentan en función de la importancia relativa de una ética basada en el individuo y otra basada en la población, y el sesgo del deseo que distrae, a veces de forma inconsciente a los investigadores, de los cánones de una buena práctica científica en los que la neutralidad y la objetividad son dos de los valores supremos.

REFERENCIAS BIBLIOGRÁFICAS

1. Fuchs S, Westervelt SD. Fraud and trust in science. *PerspBiolMed*.1995; 39: 248-269.
2. Bacallao J. Neutralidad y compromiso: la presencia de la dimensión ética en el trabajo científico. En: Acosta JR (Ed). *Bioética para la sustentabilidad*. La Habana: Publicaciones Acuario; 2002.
3. Bacallao J. Filosofía de la práctica científica. En: Tealdi, JC (Ed). *Diccionario Latinoamericano de Bioética*. Bogotá. Colombia: UNESCO; 2008.

4. Schwartz RP. Maintaining integrity and credibility in industry-sponsored clinical research. *Controlled Clin Trials*. 1991; 12: 753-760.
5. Friedlander MW. *At the Fringes of Science*. Boulder: Westview Press; 1995.
6. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000; 29: 722-729.
7. Hammer SM, Squires KE, Hughes MD, Grimes JM, *et al*. A controlled trial of two nucleoside analogues plus zidovudine in persons with HIV infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *New Engl J Med*. 1997; 337: 725-733.
8. Sterne JA, Hernán MA, Lederberger B, Tilling C, *et al*. Long term effectiveness of potent antiretroviral therapy in preventing AIDS and death: A prospective cohort study. *Lancet*. 2005; 366: 378-384.
9. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology*. 2003; 14: 680-686.
10. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Halloran E, Berry D. (Eds). *Statistical Models in Epidemiology: The Environment and Clinical Trials*. New York: Springer-Verlag; 1999, p.95-134.
11. Robins JM. Correction for non-compliance in equivalence trials. *Stat Med*. 1998; 17: 269-302.
12. McCullagh P. Regression models for ordinal data. *J Royal Stat Soc B*. 1980; 42: 109-142.
13. Vandembroucke JP. Observational Research, Randomised Trials, and Two Views of Medical Science. *PLoS Med* 2008; 5(3): e67.
14. Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: Importance, indications, and interpretation. *Lancet*. 2005; 365: 176-186.
15. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using optimal cut-points in the evaluation of prognostic factors. *J Nat Inst Cancer*. 1994; 86: 829-835.
16. Julious SA, George S, Machin D, Stephens RJ. Sample sizes for randomized trials measuring quality of life in cancer patients. *Quality of Life Research*. 1997; 6: 109-117.
17. Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binaru, ordered categorical and continuous outcomes in 2 groups comparisons. *British Med Journal*. 1995; 311: 1145-1148.
18. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *British Med Journal*. 1995; 310: 452-454.
19. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *New Eng J of Med*. 1988; 318: 1728-1733.
20. Grieve AP. The number needed to treat: a useful clinical measure or a case of the Emperor's new clothes? *Pharmac Stat*. 2003; 2: 87-102.

21. Hutton JL. Numbers needed to treat: properties and problems (with comments). *J Royal Stat Society A*. 2000; 163: 403-419.
22. Smeeth L, Haines A, Ebrahim S. Numbers needed to treat derived from meta-analyses—sometimes informative, usually misleading. *British Med Journal*. 1999; 318: 1548-1551.
23. Julious SA. Issues with number needed to treat. *Stat in Med* 2005; 24: 3233-3235.
24. Wynder EL, Higgins IT, Harris RE. The wish bias. *J ClinEpidemiol*. 1990;43:619-621.
25. Vandembroucke JP. Medical journals and the shaping of medical knowledge. *The Lancet*. 1998; 352: 2001-2006.
26. Kaptchuk TJ. Effect of interpretive bias on research evidence. *BMJ*. 2003; 28;326:1453-1455.
27. Baron R, Likar R, Martín-Mola E, *et al*. Effectiveness of Tapentadol Prolonged Release (PR) compared with Oxycodone / Naloxone PR for the management of severe chronic low back pain with a neuropathic component: A randomized, controlled, open-label, Phase 3b/4 study. *Pain Practice*. 2015.
28. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA*. 2000; 283: 2701-2711.
29. Halpern SD, Karlawish JHT, Berlin JA. The continuing unethical conduct of underpowered clinical trials. *JAMA*. 2002; 288: 358-362.
30. Bacchetti P, Wolf LE, Segal MR, *et al*. Ethics and sample size. *Am J Epidemiol*. 2005; 161: 105-110.
31. Bacallao J. Bases éticas de algunas dicotomías en el contexto de los ensayos clínicos. En: Acosta JR. (Ed). *Bioética desde una perspectiva cubana*. La Habana: Centro Félix Varela; 1997, p. 89-91.
32. Laine C, Horton R, DeAngelis CD, Drazen JM, Frizelle FA, *et al*. Clinical trial registration—Looking back and moving ahead. *N Engl J Med*. 2007; 356: 2734-2736.

Recibido: 25 de octubre de 2015.

Aprobado: 12 de noviembre de 2015.