



Fundamentos de medicina basada en la evidencia

Algunas consideraciones en torno a la selección de la muestra y a la validez externa de los ensayos clínicos

Juan Gervas

Médico general, Equipo CESCO, Madrid (España). Correo electrónico: jgervasc@meditex.es

Términos clave en inglés: evidence-based medicine; reproducibility of results; clinical trial

Términos clave en español: medicina basada en pruebas; reproducibilidad de resultados; ensayo clínico

Fecha de recepción: 15 de mayo de 2007

Fecha de aceptación: 25 de mayo de 2007

Fecha de publicación: 1 de junio de 2007

Evid Pediatr. 2007; 3: 53 doi: [vol3/2007_numero_2/2007_vol3_numero2.24.htm](https://doi.org/10.1016/S1695-2688(07)70224-4)

Cómo citar este artículo

Gervás J. Algunas consideraciones en torno a la selección de la muestra y a la validez externa de los ensayos clínicos. Evid Pediatr 2007; 3: 53.

Para recibir Evidencias en Pediatría en su correo electrónico debe darse de alta en nuestro boletín por medio del ETOC <http://www.aepap.org/EvidPediatr/etoc.htm>

Este artículo está disponible en: http://www.aepap.org/EvidPediatr/numeros/vol3/2007_numero_2/2007_vol3_numero2.24.htm
EVIDENCIAS EN PEDIATRÍA es la revista oficial del Grupo de Pediatría Basada en la Evidencia de la Asociación Española de Pediatría de Atención Primaria. © 2005-06. Todos los derechos reservados

Algunas consideraciones en torno a la selección de la muestra y a la validez externa de los ensayos clínicos

Juan Gérvas

Médico general, Equipo CESCA, Madrid (España). Correo electrónico: jgervasc@meditex.es

Introducción (con algo de historia)

La Medicina Basada en Pruebas tiene herramientas varias para ayudar en la decisión médica, pero su más característico elemento es el ensayo clínico, cuyo diseño se considera como "patrón oro" en la evaluación de intervenciones terapéuticas.

El ensayo clínico aleatorizado doble enmascarado es el referente en cuanto a calidad científica en el apoyo al trabajo médico clínico, y en la organización de servicios sanitarios. De hecho, muchos ven el ensayo clínico doble enmascarado y aleatorizado como un estudio experimental, pese a que no es más que la adaptación al campo sanitario de un diseño de comienzos del siglo XX del área de la Pedagogía, un diseño de estudio "cuasi-experimental"¹.

En la historia hay eximios antecedentes de propuestas de ensayo clínico; el más conocido el del poeta italiano Petrarca², tan crítico con los médicos como otro poeta posterior, el español Quevedo. Escribió Petrarca a su amigo Boccaccio, en pleno siglo XIV: "si se tomasen a cien personas, o mil, de una misma edad y constitución general, habituadas también a una misma comida, que hubiesen caído todas víctimas de una enfermedad a la vez, y mitad de ellas siguiesen las prescripciones de nuestros médicos contemporáneos, y la otra mitad se guiase por su instinto natural y sentido común, pero sin ningún tipo de doctores, entonces, no tengo ninguna duda de qué grupo mejoraría". Como se ve, una propuesta sesgada en contra de los médicos, que tuvo poco eco. Resonancia que también faltó a los estudios del francés Pierre Louis, quien a comienzos del siglo XIX demostró la inutilidad y peligrosidad de la sangría en el tratamiento de la pulmonía (neumonía).

En 1948 se publicó el primer ensayo clínico propiamente dicho, acerca de la eficacia de la estreptomycin en la tuberculosis^{3,4}. Detrás de ese nuevo movimiento había mucho joven inglés crítico, y también un solterón y fumador empedernido, según su auto-necrológica, Archie Cochrane⁵. La beneficiosa influencia de sus ideas⁶ llega a nuestros días como bien demuestran los Centros Cochrane, y la Biblioteca Cochrane, aunque con alguna sombra⁷. En este trabajo revisaré varias cuestiones en torno a la selección de la muestra en los ensayos clínicos, una cuestión básica en la valoración de los resultados que luego se utilizan en la Medicina Basada en Pruebas.

Sobre la validez interna y externa, y el interés clínico

Un estudio tiene validez interna si lo que dice es cierto. Es decir, si se obtienen los mismos resultados cuando se repite en condiciones idénticas. Si un estudio carece de validez interna, sus resultados son falsos. Sin validez

interna un estudio no merece ser analizado, ni sus resultados aplicados.

Los resultados de un ensayo clínico que tiene validez interna nos dan idea de la eficacia, del máximo efecto. Es decir, de los resultados que consigue una intervención en condiciones ideales.

Cuestión distinta y clave para el médico clínico es la validez externa. Un estudio tiene validez externa si se obtienen los mismos resultados cuando se repite en condiciones similares. La validez externa permite extrapolar los resultados a otros ámbitos. Por ejemplo, con validez externa las conclusiones de un ensayo clínico hecho sobre niños escoceses hipertensos hospitalizados en Edimburgo, capital de Escocia, Reino Unido, son aplicables a los niños hipertensos de atención primaria de Santa Rosa, capital de la provincia de La Pampa, Argentina.

La validez externa es clave si se pretende deducir pautas más o menos universales de los ensayos clínicos, si se busca su aplicación en otros ámbitos. Lamentablemente, la validez externa de la mayoría de los ensayos clínicos es muy escasa o nula (dejando aparte las consideraciones filosóficas al respecto). Especialmente en lo que se refiere a llevar sus conclusiones del campo experimental al trabajo diario en atención primaria⁸.

Los pacientes, las instituciones y los profesionales que participan en los ensayos clínicos poco tienen que ver con los que llevan a cabo el trabajo clínico diario en los ambientes "reales". Por ello se han propuesto alternativas para obtener resultados de ensayos clínicos menos perfectos, pero más prácticos, los llamados "ensayos clínicos pragmáticos", en los que se trata de imitar la actividad clínica real, no ideal⁹. También se ha propuesto y demostrado la utilidad de los estudios observacionales, que carecen de aura experimental, pero con los que se obtienen resultados similares a los mejores ensayos clínicos¹⁰⁻¹². Son estudios en los que la "materia prima" es la práctica clínica diaria, con sus grandezas y miserias.

Los resultados de los trabajos citados, ensayos clínicos pragmáticos y estudios observacionales, nos dan idea de la efectividad, del máximo efecto que consigue una intervención en condiciones reales.

La metodología científica permite la fácil realización de ensayos clínicos aleatorizados doble enmascarados de validez interna incuestionable, al menos en teoría. De hecho, muchos ensayos clínicos tienen gran belleza en su diseño, y el mínimo, o nulo, interés clínico. Se responde elegantemente, con una perfección formal innegable, a cuestiones clínicamente irrelevantes¹³. A los médicos y a los pacientes lo que les interesa no es la elegancia interna de un estudio, ni su significación estadística, sino

su importancia clínica, su relevancia para el manejo de las situaciones en la práctica clínica habitual.

Más allá de las cuestiones sobre validez interna y externa, lo importante es lo que aporta el estudio, lo que significa para la mejora de la decisión clínica, para la salud del paciente. A los médicos y pacientes les interesa lo que se refiere a la calidad de vida, al alivio del sufrimiento y a la mejora del pronóstico.

En último término, la pregunta ante el paciente es: ¿hay algún estudio que aporte algo importante en la decisión a tomar en esta situación y problema concreto? O, en una formulación clásica: “dada la situación de este paciente, ¿existe una intervención que lleve el problema a un mejor estado, y en su caso, con qué beneficios y perjuicios?”¹⁴. Naturalmente, para responder a esta aparente sencilla pregunta lo primero es localizar algún estudio al respecto, para después valorar la oportunidad de la aplicación de sus resultados. Pero la cuestión básica es saber en qué grado nuestro paciente concreto se parece a los pacientes que participaron en dicho estudio, para poder dar el salto desde el conocimiento epidemiológico y estadístico a la atención personal individual¹⁵⁻¹⁷. Es un salto arriesgado sobre el abismo, una “tragedia”, según Feinstein¹⁸. El problema es hasta qué punto se puede deducir la validez externa de los ensayos clínicos a partir de las publicaciones de los mismos (a partir de la información publicada)¹⁹.

La selección de la muestra

La muestra que participa en un ensayo clínico se elige a través de una serie de seis pasos sucesivos, que parte de lo más general y llega hasta los pacientes que finalizan el estudio:

1.- Población con una característica de interés

Casi nunca se estudia toda una población, todo el conjunto de elementos que tienen una característica común que nos interesa. Lo habitual es analizar una muestra. Una muestra “representativa” de la población. Pero lo básico es tener claro qué caracteriza a la población que nos interesa, y no tanto la representatividad de la muestra, pues siempre cabe señalar las limitaciones en su selección. Así, si nos interesa la prevención de las complicaciones y la atención a los niños recién nacidos con displasia de desarrollo de caderas, lo importante es definir con precisión que significa “displasia de desarrollo de caderas en el recién nacido” (¿qué caracteriza a la población de la que extraeremos la muestra?). En este caso, son pacientes desde con cadera luxada a cadera luxable, pues incluye todas las situaciones congénitas de relación anormal entre la cabeza femoral y el acetábulo (también la cadera displásica, subluxada y otras). Como se ve por el ejemplo, hay que afinar en la definición de la población, lo que muchas veces no es fácil, y no se hace en muchas publicaciones, ni siquiera en los artículos publicados en las mejores revistas del mundo. Cuando no se definen con precisión las características básicas de la población de interés se impide la reproducción y la comparación (y análisis conjunto) del ensayo o estudio considerado.

2.- Población diana

El siguiente paso es definir la población diana. Es decir, aquella población con la característica que nos interesa y de la que vamos a extraer la muestra. Supongamos que en el ejemplo previo, de displasia de desarrollo de caderas en el recién nacido, hemos decidido considerar clave el factor de riesgo de “parto en posición podálica”. No es lo mismo partir de la población con estas características encontradas en un estudio en la comunidad, de niños aparentemente sanos, que partir de los niños nacidos en un hospital universitario, o de los niños que acuden a su médico de atención primaria por distintos motivos. La definición precisa de la población diana permite la posterior extrapolación de los resultados. Naturalmente, al definir la población diana habría que especificar también las características de donde se obtiene (la comunidad, las consultas hospitalarias, los centros de salud, sus profesionales y demás), pues la variabilidad a este respecto puede ser muy importante, tanto para la validez interna del estudio (la eficacia de la intervención), como para la validez externa del mismo (la posible generalización de los resultados).

3.- Población diana seleccionada (participantes potenciales)

De toda la población diana se selecciona una muestra, bien por día, bien por época del año, bien por medio (rural, urbano), por organización sanitaria (pública, privada), por país, o por otras cuestiones prácticas y convenientes. Así, lo normal es acotar la población diana para hacer factible el estudio, ya que suele ser raro poder acceder a toda la población diana. En esta acotación se pueden introducir sesgos importantes, por lo que conviene explicar los criterios de selección de la población diana. En muchos casos esta cuestión se considera irrelevante, y apenas merece un párrafo, o ninguno¹⁵. En esta población diana seleccionada se considerará la posible participación de los centros, profesionales, y pacientes. Toda la población diana seleccionada puede participar en el estudio (participantes potenciales), pero sólo algunos lo harán: los que cumplan los requisitos planteados.

4.- Participantes elegibles

La muestra de participantes elegibles (centros, profesionales, y pacientes) se obtiene en relación con la población diana seleccionada, y generalmente se hacen salvedades. Por ejemplo, en el ejemplo citado de la displasia de desarrollo de caderas del recién nacido se pueden excluir a los que nazcan con otras alteraciones congénitas macroscópicas (como espina bífida evidente), y considerar sólo a los bebés de sexo femenino. Los criterios de inclusión pueden llegar a ser tan rígidos que excluyan al 90%, y más, de la población diana seleccionada. Así, en un trabajo sobre cribado mediante ecografía de displasia de cadera en el recién nacido se seleccionó al 10% del total de la muestra con dos criterios: existencia de factores de riesgo (aquí, antecedentes familiares, presentación podálica y chasquido presente) y cadera normal en el examen clínico a los dos días de nacimiento

en el hospital^{20,21}. En otro ejemplo, en el estudio SPAF, sobre anticoagulación y fibrilación auricular, se excluyó al 97% de los 18.376 participantes potenciales²². En un trabajo sobre eficacia de la endarterectomía en pacientes con estenosis asintomática de la carótida se excluyó inicialmente al 40% de los cirujanos, por baja calidad en su práctica, y no se renovó la participación de los que, ya en el estudio, tuvieron malos resultados¹⁵. Cuando las exclusiones son muchas, se asegura la validez interna, pero casi se afirma la ausencia de validez externa, o al menos se dificulta la generalización de los resultados^{16,23}. En las publicaciones se hacen constar raramente todas las exclusiones, y frecuentemente no se hace constar ninguna, como si no las hubiera. Se ha demostrado que esta aparente ausencia de criterios de exclusión se asocia a peor calidad del ensayo clínico²⁴. En cualquier caso, tras considerar la posible participación de la población diana seleccionada, y una vez tamizada por los criterios de selección, se genera un subconjunto al que se le ofrece participar en el estudio.

5.- Participantes que aceptan las condiciones del estudio

Aceptar la participación en un ensayo clínico no es cuestión baladí, pues se trata de someterse a una situación experimental por "el bien de la ciencia", lo que significa aceptar ciertos riesgos y, al tiempo, no tener una clara preferencia por las alternativas en juego^{15,23,25}. No todos los pacientes están dispuestos a aceptar los riesgos y costes de la participación, y son pocos los neutrales acerca de las preferencias entre alternativas. Estas cuestiones, por cierto, también afectan a la participación de centros y profesionales en los ensayos clínicos. Por ello se suele remunerar la participación en los ensayos clínicos, lo que puede llevar a un abuso en el "trabajo" de reclutar pacientes. Dejando aparte estos problemas éticos, y centrales, las preferencias de los pacientes son difíciles de controlar cuando se valora la validez externa de los ensayos clínicos, pues en muchos casos la decisión final es el resultado de una compleja interacción entre las creencias del paciente, su relación con el profesional, y los conocimientos, creencias e intereses de éste. De hecho, se habla del "efecto terapéutico de la preferencia del paciente", en cierto grado similar al efecto placebo y nocebo²³, lo que hace difícil medir las consecuencias de la neutralidad al respecto^{26,27}. Sirva de ejemplo de la importancia de las preferencias del participante, la situación que se les creó a las mujeres a las que se les ofreció participar en un ensayo clínico sobre la mastectomía en el cáncer de mama, radical o parcial. Sólo el 10% aceptó participar, pues la gran mayoría tenía una clara preferencia entre las alternativas²⁸. Hay que tener en cuenta que muchas veces los clínicos descartan la participación de pacientes por motivos nunca claros, que no constan en los protocolos de investigación, y sobre los que no se publica nada¹⁵. Sería muy útil que en los ensayos clínicos se siguieran hasta el final del estudio los dos subconjuntos en que se divide el conjunto de pacientes elegibles: el de los que aceptan participar y el de los que lo rechazan²³. Con ello

tendríamos un segundo grupo control, claro, el de los pacientes elegibles que rechazaron participar.

6.- Participantes que llegan (o deberían haber llegado) al final del estudio

Habitualmente, entre los pacientes que aceptan participar se establece una "lotería" o sistema de distribución al azar, que tiene por objeto producir dos grupos estadísticamente comparables. En el ensayo clínico aleatorizado doble enmascarado, la distribución al azar en un componente básico y crítico del mismo (el otro componente fundamental es lograr que ni clínicos ni pacientes sepan al grupo a que están asignados, lo que resulta más difícil de lo que se imagina, pero no es cuestión relevante en este trabajo). Sin embargo, en muchas publicaciones de ensayos clínicos no se alude a la forma de aleatorización. Por ejemplo, en un trabajo sobre 50 ensayos clínicos publicados en las mejores revistas, en 27 no se describió el método estadístico de distribución al azar de los participantes²⁹. No se olvide que también se deberían distribuir al azar los centros y profesionales participantes, por la heterogeneidad de los mismos³⁰. Estos defectos, y otros que atañen a la validez interna, más que a la externa, se han tratado de obviar con normas tipo CONSORT, QUOROM y otras, que obligan a hacer declaraciones explícitas en las publicaciones sobre ensayos clínicos y metanálisis, pero con poco éxito en teoría y en la práctica¹⁵. Así, por ejemplo, se recomienda que no haya pérdida ninguna en los dos grupos (intervención y control), y que en todo caso el análisis de haga por "intención de tratar"; es decir, según la asignación inicial (aunque se haya "perdido" o muerto el paciente, incluso antes de empezar el tratamiento, o haya cambiado de grupo, por causas clínicas o por otras). La mayoría de las publicaciones de los ensayos clínicos reclaman para sí el haber realizado un análisis según intención de tratar, pero cuando se estudian en detalle los resultados, se trata más bien de una frase hecha, como se ha demostrado con el estudio de 249 ensayos clínicos publicados en las mejores revistas del mundo, en los que más de la mitad reclamaban el análisis por intención de tratar, pero el 75% de estos tenían defectos que hacían imposible el cumplimiento de tal premisa³¹. Es defecto frecuente el excluir a los participantes no cumplidores, tanto en la rama control como en la experimental. Resulta absurdo, pues sabemos que el simple cumplimiento modifica el resultado, como se demostró al comparar el grupo cumplidor de tratamiento con clofibrato, que tenía menos mortalidad que el no cumplidor con dicho tratamiento. La comparación en el grupo placebo entre cumplidores y no cumplidores demostró el mismo efecto sobre la mortalidad³¹.

Conclusión

Los ensayos clínicos no son experimentos, aunque intentan cumplir con rigor una serie de normas que los transforman en semi-experimentos¹. Con todo, los resultados de los ensayos clínicos deberían aplicarse

con mucha precaución al paciente individual, pues de lo que se trata es de utilizarlos de forma tal que produzcan beneficios sin perjuicios, algo difícil incluso cuando se parte de estudios bien diseñados, bien realizados, y correctamente publicados¹⁴⁻¹⁹. La situación es peculiar en atención primaria, pues muchos ensayos clínicos se publican a partir de pacientes, centros y profesionales hospitalarios, y sus resultados son ajenos a las situaciones habituales en la práctica de la Medicina General/ de Familia y de la Pediatría^{8,31}. Las críticas a la Medicina Basada en Pruebas tienen frentes varios, que en este trabajo no he analizado, y van desde los problemas conceptuales a la defectuosísima aplicación práctica de sus principios^{7,32-34}. No es extraño, pues, que al evaluar 61 guías de práctica clínica muy difundidas en España, se llegase a la conclusión de que sólo cuatro eran "muy recomendables" (contra 25 "muy poco recomendables")³⁴.

En este trabajo he considerado los problemas en torno a la elección de la muestra, que obligan a poner en duda gran parte de los resultados de los ensayos clínicos (por cuestiones en torno a la validez interna y a la validez externa). Estos problemas han llevado a decir que los ensayos clínicos son mejores experimentos que encuestas³⁵. Por conveniencia, las muestras de pacientes y profesionales que participan en los ensayos clínicos no suelen ser representativas de las poblaciones con la característica de interés analizada³⁶, y ello provoca graves problemas cuando se generalizan los hallazgos y se ignora la heterogeneidad de los resultados^{37,38}.

Los sesgos que conllevan los problemas en la selección de la muestra tienen que ver con la heterogeneidad de los resultados, y se deben a:

- 1.- La distinta probabilidad inicial de que los pacientes presenten un evento indeseable relacionado con su enfermedad (susceptibilidad, o riesgo sin tratamiento).
- 2.- La variabilidad en la respuesta a la intervención.
- 3.- La probabilidad de presentar un efecto adverso como consecuencia de la intervención.
- 4.- La distinta utilidad de los pacientes (la valoración y preferencias por los resultados esperables con y sin intervención)^{37,38}.

Para terminar, sirva de ejemplo extremo el ensayo clínico PROACT, de utilización de la trombolisis en el ictus isquémico agudo. Se partió de una población total de 12.323 pacientes, de los que hubo que excluir por diversas razones a 12.143 (98,5%). Es decir, los resultados se obtuvieron en un ensayo clínico con 180 pacientes repartidos en dos grupos (control y experimental)³⁹. Evidentemente, esa muestra no es representativa de la población de pacientes con ictus agudo isquémico a la que se le aplicarían los resultados del PROACT¹⁹. Contra lo que pueda parecer es muy frecuente el obtener resultados de poblaciones fuertemente seleccionadas, no representativas^{19,40}. La moraleja es simple: detrás de la etiqueta "Medicina Basada en Pruebas/ la Evidencia" hay a veces validez interna, y raramente validez externa^{16,32,33,38,40}. Por ello es clave en los análisis de ensayos clínicos el

apartado de "aplicabilidad en la práctica clínica", que facilita la toma de decisiones basadas en las mejores pruebas científicas.

Nota

Parte de lo expuesto procede del material de un libro en elaboración por el firmante, con Mercedes Pérez Fernández, que se titulará "Reivindicación de una medicina cercana, científica y humana".

Bibliografía:

- 1.- Gervas J, Pérez Fernández M. El diseño semi-experimental y la investigación evaluativa. *Aten Primaria*. 1991;8:567-76.
- 2.- Editorial. El referente: el ensayo clínico. *Informatiu Agencia d'Avaluació de Tecnologia y Recerca Mèdiques*. 2003;31:1-2.
- 3.- Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis. *BMJ*. 1948;30:769-82.
- 4.- Crofton J. The MRC randomized trial of streptomycin and its legacy: a view from the clinical front line. *J R Soc Med*. 2006;99:531-4.
- 5.- Hill GB. Archie Cochrane and his legacy. An internal challenge to physicians's autonomy. *J Clin Epidemiol*. 2000;53:1193-9.
- 6.- Cochrane A. Eficacia y eficiencia. Relexiones al azar sobre los servicios sanitarios. Barcelona: Salvat; 1985.
- 7.- Links M. Analogies between reading of medical and religious texts. *BMJ*. 2006;333:1068-70.
- 8.- Graham RP, James PA, Cown TM. Are clinical practice guidelines valid for primary care?. *J Clin Epidemiol*. 2000;53:949-54.
- 9.- Vallvé C. Revisión crítica del ensayo clínico pragmático. *Med Clí (Barc)*. 2003;121: 384-8.
- 10.- Barton S. Which clinical studies provide the best evidence. The best RCT still trumps the best observational study. *BMJ*. 2000;321:255-66.
- 11.- Benson K, Hartz AJ. A comparison of observational studies and randomized controlled trials. *N Engl J Med*. 2000;342:86.
- 12.- Concato J, Horwitz RI. Beyond randomised versus observational studies. *Lancet*. 2004;363:1660-1.
- 13.- Starfield B. Quality of care research. Internal elegance and external relevance. *JAMA*. 1998;280:1006-8.
- 14.- Fletcher RH. Evaluation of interventions. *J Clin Epidemiol*. 2002;55:1183-90.
- 15.- Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply". *Lancet*. 2005;365:82-93.
- 16.- Pereira D. Evidence-based medicine and patient-centred medicine: the need to harmonize. *J Health Serv Res Policy*. 2005;10:66-8.
- 17.- Gervas J, Pérez Fernández M. Uso apropiado de la medicina basada en pruebas: revisión de diez artículos recientes. *AMF*. 2005;1:46-6.
- 18.- Feinstein AR. The problem of cogent subgroups: a clinico-statistical tragedy. *J Clin Epidemiol*. 1998;51:297-9.
- 19.- Gross CP, Mallory R, Helat A, Krumholz HM. Reporting the recruitment process in clinical trials: who are these patients and how did they get there?. *Ann Intern Med*. 2002;137:10-6.
- 20.- Roovers EA, Boere-Boonekamp MM, Castelein RM, Zielhuis GA, Kerkhoff TH. Effectiveness of ultrasound screening for devel-

opmental dysplasia of the hip. Arch Dis Child Fetal Neonatal Ed. 2005;90:F25-30.

21.- Perdikidis L, González de Dios J. El cribado ecográfico de displasia de desarrollo de caderas sólo se justifica en recién nacidos con determinados factores de riesgo. Evid Pediatr. 2005;1:4.

22.- Stroke Prevention in Atrial Fibrillation Investigators. Stroke prevention in atrial fibrillation study: final results. Circulation. 1991;184:527-39.

23.- Mant D. Can randomised trials inform clinical decisions about individual patients?. Lancet. 1999;353:743-6.

24.- Schulz KF, Grimes DA, Altman DG, Hayes JR. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynecology. BMJ. 1996;312:742-4.

25.- Sackett DL. Participants in research. BMJ. 2005;330:1164.

26.- McPherson K, Britton A, Wennberg J. Are randomised controlled trials controlled? Patients preferences and unblind trials. J R Soc Med. 1997;90:652-6.

27.- Flum DR. Interpreting surgical trials with subjective outcomes: avoiding UnSPORTsmanlike conduct. JAMA. 2006;296:2483-5.

28.- Olschewski M, Schumacher M, Davis KB. Analysis of randomised and non-randomised patients in clinical trials using the comprehensive cohort follow-up study design. Control Clin Trials. 1992;13:226-39.

29.- Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trial. Lancet. 2000;355:1064-9.

30.- Lee KJ, Thompson SG. Clustering by health professional in individually randomised trials. BMJ. 2005;330:142-4.

31.- Hollis S, Campbell F. What is meant by intention to treat analysis?. Survey of published randomised controlled trials. BMJ. 1999;319:670-4.

32.- Gervás J. Experiencia, ciencia y "medicina basada en pruebas" en atención primaria. SEMERGEN. 2002;28:302-4.

33.- Alonso-Coello P, García JM, Solá I. Limitaciones y subterfugios de la crítica a la medicina basada en la evidencia. Med Clí (Barc). 2005;124:237-8.

34.- Navarro MA, Ruiz F, Reyes A, Gutiérrez I, Hermosilla T, Alonso C, et al. ¿Las guías que nos guían son fiables?. Evaluación de las guías de práctica clínica españolas. Rev Clí Esp. 2005;205:533-40.

35.- Longford NT, Nelder JA. Statistics versus statistical science in the regulatory process. Statistics Med. 1999;18:2311-20.

36.- Schmoor C, Olschewski M, Schumacher M. Randomized and non-randomized patients in clinical trials: experiences with comprehensive cohort studies. Statistics Med. 1996;15:263-71.

37.- Longford NT. Selection bias and treatment heterogeneity in clinical trials. Statistics Med. 1999;18:1467-74.

38.- Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q. 2004;4:661-87.

39.- Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, Kase C, et al. Intra-arterial prourokinase for acute ischemic stroke. The PROACT II Study: a randomized controlled trial. Prolyse in Acute Cerebral Thromboembolism. JAMA. 1999;282:2003-11.

40.- Steinberg EP, Luce BR. Evidence based?. Caveat emptor!. Health Affairs. 2005;24:80-92.