

# CAPÍTULO 34 - MÉTODOS DE INVESTIGACIÓN CLÍNICOS EN COT.

Autor: Francisco Forriol Campos

Facultad de Medicina Universidad San Pablo - CEU, Campus Montepríncipe (Madrid)

## 1.- INTRODUCCIÓN

La investigación para un cirujano es la mejor manera de evaluar su propia actividad. Una consecuencia de su práctica clínica y una necesidad para la docencia.

Investigar es una actividad bien vista socialmente aunque mal remunerada que empieza a ser una exigencia en el desarrollo profesional y requiere dedicación, continuidad con una metodología que se adapte a los principios del método científico. Salter (1), denomina al método científico como "ciclo de la naturaleza de la investigación médica" y lo divide en 16 fases (Tabla 1). Pero no hay un sólo método científico. Según el modo de razonar se puede seguir el método deductivo, del todo a las partes, de lo general a lo particular, o el método inductivo, de las partes al todo, de lo particular a lo universal.

	Salter		Janssen
1	Reconocer el problema clínico no resuelto	1	Identificar el problema
2	Pensar		
3	Revisar la literatura	2	Revisar la bibliografía
4	Plantear cuestiones inteligentes		
5	Formular hipótesis	3	Desarrollar la hipótesis
6	Plan de protocolo	4	Diseño experimental
7	Buscar colaboración		
8	Aplicar subvención		
9	Conducir la investigación	5	Realizar el experimento
10	Recogida y análisis de datos		
11	Interpretar datos		
12	Conclusiones válidas	6	Conclusiones
13	Responder a la cuestión original		
14	Presentación de los resultados		
15	Publicar los resultados		
16	Aplicar los nuevos conocimientos		

Existen dos tipos de investigación: la epidemiológico - clínica o "investigación clínica" y la básico - experimental o "de laboratorio". La investigación clínica se ha dividido, a veces de forma arbitraria, en dos categorías, básica y aplicada, que se rigen por los mismos principios aunque difieren en algunos aspectos. Ambas pueden dividirse en investigación experimental, efectuada en animales, o clínica, con pacientes.

La investigación clínica incluye cualquier investigación en la cual la unidad de análisis es la persona (2) y despierta cada vez mayor interés la **investigación translacional**, un intento de llevar la información del laboratorio a la práctica clínica. El nacimiento de la investigación translacional es una necesidad para trasladar la información obtenida en el laboratorio, desde los llamados sistemas no humanos al hombre, lo antes posible.

## 2.- EL DISEÑO EXPERIMENTAL

El investigador debe considerar los parámetros a comparar, los procedimientos a utilizar y los grupos a estudiar. Este proceso exige esfuerzo y meticulosidad, pues cualquier fallo puede invalidar los resultados, además precisa de creatividad, imaginación y conocimientos del método científico sin olvidar los conocimientos técnicos.

El modelo experimental es el diseño de un trabajo de investigación o el modo elegido para recoger datos de un hecho real. Se pueden adquirir de tres formas distintas, por la observación, por la medición o planteando una hipótesis (3). Según se manipulen los sujetos de estudio o no, los estudios serán experimentales u observacionales.

La investigación de calidad debe centrarse en proyectos concretos bien estructurados metodológicamente manteniendo el principio que los anglosajones han llamado KISS (beso) (Keep It Simple and Succint), es decir, medios simples, singulares en sus propósitos y sencillos en su diseño.

Un proyecto no puede encontrar todas las respuestas a todas las cuestiones; es más, un único proyecto de investigación no suele dar una respuesta completa a una cuestión.

Todo proyecto de investigación atraviesa por tres amplias fases diseñar el proyecto, realizar el diseño, y comunicar los resultados obtenidos.

En investigación lo primero es identificar el problema. Un problema mal planteado conlleva al fracaso de la línea de investigación. El segundo paso comprende una profunda revisión de la literatura. Los datos publicados por otros autores pueden variar las perspectivas de la línea de investigación y disminuir su coste, evitando repetir experiencias ya realizadas y ampliando el campo de visión. En tercer lugar, hay que desarrollar una hipótesis que se debe comprobar experimentalmente. El cuarto paso establece el modelo experimental, diseña el experimento para demostrar la hipótesis (Tabla 1).

El modelo experimental depende de la hipótesis y de los objetivos establecidos, del rigor científico deseado y, también, de los recursos disponibles. A la hora de diseñar un modelo hay que ser muy crítico, realista y práctico, pues

debe ser posible su realización (Tabla 2). El objetivo de cualquier modelo de investigación es planificar los medios que nos permitan recoger los datos de manera óptima y utilizarlos con la técnica analítica más apropiada.

Tabla 2. Tipos de estudios de investigación, objetivos y diseño		
Investigación	Objetivo	Diseño ideal
Terapéutico	Eficacia nuevo tratamiento o cirugía	Ensayo controlado y aleatorizado
Diagnóstico	Probar si un nuevo diagnóstico o examen es válido y realizable	Estudio transversal
Rastreo (screening)	Probar el valor de un diagnóstico en un rastreo en una población grande	Estudio transversal
Pronóstico	Detectar pronto aspectos de una enfermedad	Estudio de cohortes longitudinal
Factores de riesgo	Determinar si un factor de riesgo está relacionado con el desarrollo de una enfermedad	Estudio de cohortes o caso-control

En el método científico el investigador parte de una idea. Después propone una hipótesis nula, una afirmación contraria a la idea que puede ser verdadera si no se observa ninguna diferencia entre los grupos experimentales. La hipótesis nula es, dentro de la investigación de la hipótesis, la hipótesis inicial que la prueba estadística debe considerar y, posiblemente, rechazar. El experimento, según los resultados obtenidos, puede refutar la hipótesis nula, si muestra que hay diferencia entre los grupos. Esto avala la idea original. De alguna manera investigar consiste en intentar rechazar una hipótesis nula.

### 3.- MEDIDAS Y VARIABLES. MUESTRAS Y SELECCIÓN

La mayoría de los estudios en medicina clínica involucran poblaciones que no pueden ser medidas en su totalidad, para lo cual se evalúan muestras parciales de esas poblaciones. La mejor medida para conocer si una muestra es representativa del grupo es su tamaño, pues las verdades estadísticas son más evidentes a medida que aumenta el número de observaciones. Cuando se investiga una hipótesis se introduce una incertidumbre porque nos tenemos que basar en una muestra y una muestra es una medición incompleta (4). Sólo podemos tener una cierta probabilidad (menor del 100%) de que la observación de nuestra muestra sea la misma de la que se obtendría de una medición de la población completa.

El método fundamental para establecer hechos biológicos es el experimento controlado. No olvidemos que un experimento es una situación en la que el examinador manipula el medio, aplicando una acción a un grupo y otra a un segundo grupo. Las mediciones de un experimento pueden obtenerse de forma directa sobre el espécimen o de manera indirecta con técnicas que no se integran en la pieza de estudio (radiografías, vídeo, modelos físicos, diseño por ordenador...) (5).

Cuando se efectúa investigación clínica, se deben especificar:

- los criterios seguidos en la admisión de los pacientes del estudio.
- conocer y explicar los pacientes que se pierden en el estudio.
- Cómo se evalúan los cambios radiográficos
- El control desarrollado en la población estudiada.
- Las estrategias seguidas para evitar la subjetividad.
- La exposición del análisis estadístico.
- Los criterios utilizados en la recogida de los datos.
- La evaluación correcta de su tratamiento. Tanto los buenos como los malos resultados son componentes de la salud del paciente (3)(6).

Como casi nunca se conoce la verdad con certeza absoluta, todos los resultados se deben expresar en términos de probabilidad. De hecho, los resultados se expresan con una  $p$ , que cuantifica la probabilidad de error. El criterio umbral de probabilidad es el nivel  $\alpha$ . Cuando  $p$  es inferior a  $\alpha$ , se considera que la hipótesis nula es falsa. En la mayoría de los estudios clínicos se establece  $\alpha$  en 0,05, es decir, se admite una probabilidad menor del 5% de que se haya cometido un error. Cuando el valor de  $p$  se encuentra por encima de este umbral se dice que los resultados obtenidos no son estadísticamente significativos.

Un estudio puede ser no significativo por que realmente no hay diferencia entre dos muestras o por que el estudio no tiene suficiente potencia para detectarla. La potencia estadística de un estudio viene determinado por el estadístico  $b$  que depende del tamaño de la muestra  $y$ , en menor grado, de la magnitud del efecto de interés y la variabilidad de los datos. También aumenta si se eleva el umbral  $\alpha$  (6).

La magnitud del efecto de interés es un parámetro que se verifica simplemente, no se calcula. Por ejemplo, la división de las poblaciones en edades, alturas o pesos puede depender del interés del investigador. De haberlo hecho de otra forma a la elegida hubiera o no sido significativa. La variabilidad de los datos es un aspecto interesante pues también puede ser manipulado pero si los datos se agrupan alrededor de la media, se necesitan menos mediciones para determinar esa media. Cuanto mayor sea una muestra habrá más confianza de que la media calculada sea la media verdadera del grupo. No es bueno aumentar considerablemente una muestra pues se desperdician recursos y puede significar que algunos pacientes se someten sin necesidad a tratamientos placebos o inferiores. Para evitarlo se requieren cálculos de la muestra mínima necesaria. Por lo general y de forma aproximada, sin ser un principio estadístico, cuando se quiere detectar un efecto de magnitud moderada, asumiendo una variabilidad moderada y un error  $\alpha$  de 0,05, se requieren alrededor de 75 individuos (7).

Al efectuar un estudio muchas veces se precisa conocer cuál es el resultado con la evolución en el tiempo. Una rodilla intervenida puede moverse con normalidad o presentar unas imágenes normales pasado un tiempo pero, sin embargo, haber creado un trastorno en el tipo de vida del sujeto. Por eso incluir un protocolo de calidad de vida

relacionada con la salud en los estudios clínicos hace que se comprenda lo que se ganó o perdió como consecuencia del tratamiento. Se disponen de diferentes protocolos de calidad de vida, algunos generales (SF-36) y otros específicos (IKDC, para rodilla; Oswestry, para columna vertebral). Todos intentan valorar datos objetivos y subjetivos recogiendo la opinión del paciente, del médico y de las pruebas efectuadas.

Un protocolo de calidad de vida debe ser fiable, válido y con calidad de respuesta (8). La fiabilidad viene determinada por el grado en que un instrumento ofrece los mismos resultados cuando se aplica reiteradamente a una población. Por su parte, la validez del estudio de la muestra está avalado por la capacidad del instrumento para medir lo que se proponía. Además, es importante medir bien lo que es difícil de medir. Por ejemplo, la salud no se puede medir directamente pero se puede conocer el estado de salud por cosas que asociamos con la buena salud. La capacidad de respuesta es la capacidad de una variable para reflejar un cambio subyacente o una diferencia verdadera entre las puntuaciones antes y después de un tratamiento.

#### 4.- TIPOS DE ESTUDIOS - TIPO DE PUBLICACIONES

No todos los artículos publicados en una revista son iguales, cada tipo de estudio tiene una metodología y una forma de ser escritos diferentes.

Los estudios más complejos de diseñar son, sin lugar a dudas, los **ensayos clínicos**, siendo muchas veces irrealizables o no éticos, por lo que será necesario recurrir a la experimentación animal. En la investigación clínica hay que tener en cuenta dos aspectos, la manipulación y la aleatorización, para crear conjuntos iguales duplicados de individuos repartidos o asignados al azar; aleatorizando a cada sujeto al grupo control o al grupo donde se produce la variación. Si el tamaño de la muestra es grande el azar reparte homogéneamente todas las variables de estudio entre los grupos. Es recomendable utilizar grupos control pues aumentan el nivel de evidencia y mejoran la confianza y, por lo tanto, las posibilidades de su aplicación en clínica (3)(9).

Un diseño puede seguirse en el tiempo, será un estudio longitudinal, o analizarse en un momento determinado, en cuyo caso hablaremos de un estudio transversal o de corte. También los estudios se pueden centrar en individuos concretos o en un grupo de sujetos y, además, según la información disponible, pueden ser completos o incompletos.

El diseño de un estudio puede ser no experimental o experimental, que a su vez se dividen en estudios experimentales puros y los llamados cuasi-experimentales. Cada uno de ellos está indicado según las situaciones. Los estudios experimentales, son manipulados y prospectivos. Si bien hay que distinguir los estudios experimentales puros que están manipulados y son aleatorios, de los estudios cuasi-experimentales o comunitarios, manipulados pero no aleatorios. Además, los diseños pueden recoger una muestra con un final ya conocido, retrospectivos, o hacer un planteamiento para saber su final, en cuyo caso se denominan prospectivos.

Existen distintos tipos de investigación clínica, que podrían agruparse en dos grandes campos, la investigación sobre la eficacia y los estudios sobre la efectividad. La investigación sobre la eficacia es la investigación de técnicas y tratamientos, en su fase de desarrollo, que se lleva a cabo en instituciones que disponen de los medios necesarios para determinar la validez o el fracaso de un nuevo tratamiento. Por su parte, los estudios sobre la efectividad pretenden conocer una técnica o un fármaco ya adoptados y generalizados, para confirmar si los resultados de las investigaciones sobre eficacia son válidos en la práctica general. A este grupo pertenecen la mayoría de los trabajos clínicos. Pero la mayoría de los estudios clínicos no se puede olvidar que proceden de una investigación básica que tuvo lugar en un momento determinado, más o menos lejano y que sigue los procesos más variados.

La investigación quirúrgica no es fácil ni sencilla pues hay muchos sesgos y las muestras son heterogéneas. Además, no siempre resulta factible programar grupos control ni controlar variables, como puede ser la habilidad o experiencia de los cirujanos. Cuando se comparan tratamientos quirúrgicos y conservadores hay una tendencia a que los pacientes más graves reciban un tratamiento quirúrgico y los pacientes prefieren los tratamientos más novedosos. No está justificado, profesional y éticamente, practicar cirugías placebo (9).

La mayoría de los trabajos aleatorizados y controlados en cirugía tienen un elevado número de exclusiones, influye la técnica de cada cirujano y la experiencia del centro, así como el volumen de pacientes intervenidos. Además, también se relaciona con la curva de aprendizaje. Por otro lado, si pensamos en los estudios comparativos con implantes es casi imposible que se pueda analizar el comportamiento de dos o más implantes que no estén aprobados y más difícil que se puedan hacer modificaciones de un implante después de un estudio (9).

Distinguimos entre estudios retrospectivos y prospectivos. Los estudios retrospectivos consisten en la recogida y elaboración de datos obtenidos antes de iniciar el estudio, y que por lo tanto no se obtuvieron con ese propósito. Entre los estudios retrospectivos encontramos los análisis de variación en pequeñas áreas, análisis de grandes bases de datos y el meta-análisis.

Los estudios prospectivos se realizan de forma planificada con unos protocolos establecidos previamente. Entre los estudios prospectivos tenemos los estudios de coste y eficacia, el análisis de decisión, las recomendaciones clínicas y los estudios clínicos.

Los estudios de coste y eficacia son cada vez más importantes y según el método de medición utilizado, las evaluaciones económicas pueden ser análisis de coste-eficacia; de coste-utilidad o de coste-beneficio.

Los ensayos clínicos longitudinales de asignación aleatoria son los estudios ideales para valorar un tratamiento. Es aconsejable disponer de un grupo control para contrastar los resultados del tratamiento y aumenta el valor del trabajo. En cirugía interesa efectuar estudios multicéntricos, donde cada uno realiza la técnica que domina con un equipo de personas que conocen el manejo

de esos pacientes, siguiendo unos protocolos analizados y evaluados y procurando disponer de muestras lo más grandes posibles (7).

Por otra parte, el número de publicaciones disponibles en la actualidad exige que el conocimiento sea gestionado de manera eficaz. Es decir, resulta interesante que alguien nos revise todo lo que está escrito sobre un tema. Cuando decimos todo se entiende que es todo lo que tiene visibilidad y calidad científica. Las revisiones narrativas están realizadas por un grupo de expertos por lo que suelen tener el peligro de intentar potenciar la idea del que escribe y no suele tener una metodología de búsqueda. Por eso las revisiones sistemáticas responden una pregunta buscando entre la evidencia existente con métodos específicos y sistemáticos para identificar, seleccionar y evaluar críticamente la investigación primaria y extraer y analizar los datos de los estudios que se incluyen en la revisión. Estas revisiones sistemáticas pueden utilizar o no técnicas de metaanálisis que no son otra cosa que la combinación por medio de técnicas de estadística para combinar los datos de los estudios publicados que han intentado responder a la misma pregunta. Aunque revisión sistemática y metaanálisis se utilizan, a veces, indistintamente, no son lo mismo. La revisión sistemática es el proceso para obtener conclusiones a partir de estudios originales, ya sea siguiendo técnicas estadísticas o no, el metaanálisis es la combinación matemática de al menos dos estudios para determinar una estimación del efecto de la intervención que se está evaluando.

En esta línea, la Colaboración Cochrane pretende generar revisiones sistemáticas de alta calidad sobre los efectos de las intervenciones médicas.

## 5.- MEDICINA BASADA EN EVIDENCIAS (MBE): UN MÉTODO DE PUBLICAR

Se puede pensar que todo estudio estadísticamente bien planteado es una forma de hacer MBE, pero no siempre es cierto. Muchas veces las mejores evidencias provienen de estudios no aleatorios que constituyen la base para desarrollar hipótesis de trabajo futuras (10).

Están de moda los niveles de evidencia, un método o mejor una llamada de atención a tener en cuenta la metodología a la hora de realizar una publicación y plantear un trabajo científico. El *Journal of Bone and Joint Surgery americano* clasifica sus trabajos siguiendo la clasificación del *Oxford Centre for Evidence-based Medicine*, en cinco niveles de evidencia según el diseño del artículo y englobado en uno de los diferentes apartados: terapéutico, pronóstico, diagnóstico y económico o análisis de decisión (Tabla 3).

En el nivel I, están los ensayos clínicos controlados de alta calidad aleatorios con o sin diferencias estadísticamente significativas pero, siempre con estrechos intervalos de confianza. Al nivel II pertenecen ensayos controlados aleatorios de menor calidad, como un estudio comparativo o una revisión sistemática de estudios de nivel II. En el nivel III se encuentran los estudios de caso-control, comparaciones retrospectivas y revisiones de artículos de este tipo. El nivel IV lo ocupan los casos clínicos y, por último, el nivel V corresponde a la opinión de los expertos

en un tema. En principio, los niveles más bajos muestran los estudios con mayor garantía para la aplicación en el tratamiento de los pacientes.

La jerarquía establecida para los niveles de evidencia es muy rígida ya que defiende que los estudios aleatorios y controlados son los mejores y que lo demás es una metodología de peor calidad. Es una mentalidad muy simple pues cada estudio tiene valor si está bien hecho.

Los niveles de evidencia son difíciles de valorar en el mundo de la cirugía pues no resulta fácil realizar estudios aleatorios o controlados de alta calidad de los diferentes tipos de tratamientos quirúrgicos ya que participan consideraciones éticas, dificultades para adherirse a los protocolos de cada cirujano, organización del estudio con suficiente poder y diferencias en las técnicas individuales. Como señala el actual editor del *Journal of Bone and Joint Surgery británico*, los mayores avances en cirugía ortopédica y cirugía general son ensayos controlados; la mayoría de los avances estarían en el nivel III y muy pocos en el nivel II. En muchas ocasiones, algunas cuestiones planteadas en clínica no pueden ser desarrolladas según la MBE con niveles I o II por lo infrecuente de la enfermedad o patología dejando como única opción la recogida de casos clínicos.

## 6.- ALGUNOS ERRORES ELEMENTALES EN LOS DISEÑOS DE INVESTIGACIÓN

En los diseños de investigación se pueden producir errores sistemáticos o sesgos que impiden detectar los efectos buscados. Pueden ser predecibles y, a veces, se pueden evitar con un diseño correcto y el análisis adecuado de los datos. Hay sesgos de selección, de clasificación, de confusión y de interpretación.

Los sesgos de selección se producen cuando el procedimiento selecciona individuos que no representan a la población a la que se extrapolan los resultados. El sesgo de mayor interés es el error en la estimación del tamaño de la muestra.

En ocasiones el entrevistador fuerza situaciones que tienden a apoyar la hipótesis de partida del propio investigador. Por eso es bueno que los entrevistadores que participan en un estudio desconozcan las hipótesis del proyecto. Para evitar los sesgos del observador se utilizan las técnicas de enmascaramiento o de ciego muy frecuentes en los ensayos clínicos. El ciego simple es cuando el participante en el estudio no sabe si pertenece al grupo experimental o al control; doble ciego cuando ni el participante ni el que recoge la información saben a que grupo pertenece y triple ciego cuando ni el participante, ni el que recoge la información ni tampoco el analizador de los datos conocen el grupo al que está adscrito.

Los sesgos de confusión producen una distorsión en la estimación del efecto pues se sobrestima, se subestima o se anula, por la presencia de una tercera variable que se conoce como factor de confusión. Para evitar los sesgos de confusión en la etapa de diseño del estudio hay que admitir

**Tabla 3. Niveles de evidencia para cuestiones de investigación primarias**

	Estudios terapéuticos - Investigación de los resultados de tratamiento	Estudios pronósticos Análisis prospectivo de enfermedades	Estudios diagnósticos - Investigación pruebas de diagnóstico	Análisis económicos y de decisión - Desarrollo de un modelo económico o de decisión
<b>Nivel I</b>	1. ensayos controlados y aleatorizados. a. con diferencias significativas b. sin diferencias significativas pero con intervalos de confianza estrechos  2. revisión sistemática (2) de ensayos controlados aleatorizados (estudios homogéneos)	a. Estudios prospectivos (1)  b. Revisión sistemática (2) de estudios nivel I	1. Pruebas de criterios de diagnóstico desarrollados previamente en series de pacientes consecutivos (con una referencia universal "gold standard")  2. Revisión sistemática (2) de estudios de nivel I	1. Costes clínicos y alternativas; valores obtenidos de muchos estudios; análisis de sensibilidad por diferentes vías  2. Revisión sistemática (2) de estudios nivel I
<b>Nivel II</b>	1. Estudios prospectivos de cohortes (3)  2. Ensayos aleatorizados de pobre calidad (p. Ej., seguimiento <80%)  3. Revisiones sistematicas (2) a. estudios nivel II b. estudios no homogéneos nivel I	1. Estudio retrospectivo (4)  2. Estudio de controles no tratados de un ensayo clínico aleatorizado previo  3. Revisión sistemática (2) de estudios de nivel II	1. Desarrollo de criterios de diagnóstico en base de pacientes consecutivos (con una referencia universal "gold standard")  2. Revisión sistemática (2) de estudios de nivel II	1. Costes clínicos y alternativas; valores obtenidos de estudios limitados; análisis de sensibilidad por diferentes vías  2. Revisión sistemática (2) de estudios de nivel II
<b>Nivel III</b>	1. Estudios caso - control (5)  2. Estudio retrospectivo de cohortes  3. Revisión sistemática (2) de estudios nivel III		1. Estudio de pacientes no consecutivos (sin referencia a un "gold standard")  2. Revisión sistemática (2) de estudios nivel III	1. Alternativas limitadas y costes; estimación pobre  2. Revisión sistemática (2) de estudios nivel III
<b>Nivel IV</b>	Serie de casos (no, o históricas, grupo control)	Serie de casos	1. Estudio caso - control  2. Referencias pobres	Análisis sin sensibilidad
<b>Nivel V</b>	Opinión de un experto	Opinión de un experto	Opinión de un experto	Opinión de un experto

1. todos los pacientes están incluidos en el mismo punto durante su enfermedad (inception cohorte) con un seguimiento  $\geq 80\%$  de los pacientes incluidos
2. un estudio de resultados de dos o más estudios
3. pacientes son comparados con un grupo control de pacientes tratados en el mismo tiempo e institución
4. el estudio se inició después de aplicar el tratamiento
5. pacientes con un seguimiento particular ("casos" con, p. Ej., una prótesis de cadera de revisión) son comparados con aquellos que no tienen seguimiento ("controles" con, por ejemplo, una prótesis de cadera que no hay que revisar)

únicamente a los sujetos que no presentan factores de confusión; realizando en los estudios experimentales puros, una asignación aleatoria o un emparejamiento (matching), procurando que los grupos sean homogéneos. También se pueden evitar los sesgos de confusión con la estratificación, es decir, dividiendo la muestra global en estratos, según los posibles valores de la potencial variable de confusión.

Por último, entendemos por sesgos de interpretación cuando se comenten errores por la interpretación de los resultados estadísticos.

## BIBLIOGRAFÍA

1. Salter RB. Trastornos y lesiones del sistema músculoesquelético. Barcelona, Salvat, 2ª ed, 1990
2. Schuster DP, Powers WJ. Translational and experimental clinical research. Lipincott Williams&Wilkins, Philadelphia, PA, 2005.
3. Bhandari M, Schemitsch EH. Planning a randomized clinical trial. An overview. Techniques in orthopaedics. 2004; 19:66-71.
4. Jacobs WCH. Methodology and reporting of randomised controlled trials. The Knee 2004; 11:423-5.
5. Marx RG, Dunn WR. Randomized controlled trials in knee surgery. J Knee Surg 2005; 18:65-8.
6. Boutron I, Ravaud P, Nizard R. The design and assessment of prospective randomised, controlled trials in orthopaedic surgery. J Bone Joint Surg (Br) 2007; 89-B:858-63.
7. Freedman KB, Bernstein J. Sample size and statistical power in clinical orthopaedic research. J Bone Joint Surg (Am) 1999; 81-A:1454-60.
8. Lang T. Twenty statistical errors even YOU can find in biomedical research articles. CMJ 2004; 45:361-70.
9. Wright JG, Gebhardt MC. Multicenter clinical trials in orthopaedics: time for musculoskeletal specialty societies to take action. J Bone Joint Surg (Am) 2005; 87-A:214-7.

10. Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the Journal. *J Bone Joint Surg (Am)* 2003; 85-A:1-3.