

Valoración de la calidad metodológica y ética de los ensayos clínicos publicados en revistas de medicina de familia (2010-2013)

Alberto Castaño-García¹, Francisco Guillén-Grima² y Pilar León-Sanz³

¹Centro de Salud de Corella, Servicio Navarro de Salud, Corella, Navarra; ²Servicio de Medicina Preventiva, Clínica Universitaria de Navarra, Pamplona, Navarra; Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona, Navarra; ³Facultad de Medicina, Departamento de Humanidades Biomédicas, Universidad de Navarra, Pamplona, Navarra, España

Resumen

Objetivo: Evaluar algunas variables sobre la calidad metodológica y ética de los ensayos clínicos publicados en 10 revistas de medicina de familia. **Métodos:** Estudio descriptivo de calidad sobre 10 revistas de medicina de familia incluyendo ensayos clínicos en humanos publicados entre 2010 y 2013. Obtuvimos 141 ensayos clínicos y fueron excluidos 2447. **Resultados:** Ensayos clínicos controlados paralelos en el 92,9% (intervalo de confianza del 95% [IC 95%]: 92.0-93.9). Aleatorización enmascarada en el 72.3% (IC 95%: 71.7-73.1), descentralizada en el 51.8% (IC 95%: 51.4-52...4) y utilizando como control un tratamiento activo el 82.2% (IC 95%: 81.5-83.1). Consentimiento informado escrito en el 48.9% (IC 95%: 48.5-49.5) y no fue retirado en el 56.0% (IC 95%: 55.5-56.7). En 134 ensayos clínicos se contó con la aprobación por un comité ético de investigación clínica (CEIC), y en 117 no hubo conflicto de intereses. Se obtuvo un κ medio de 0,96 (IC 95%: 0.93-0.99). **Conclusiones:** Observamos, tras considerar las normas CONSORT, un aumento en algunas variables de calidad, como la aleatorización enmascarada (19.6%) y la aprobación por un CEIC (75%), en los ensayos clínicos publicados en 10 revistas de medicina de familia (2010-2013).

PALABRAS CLAVE: Lista de verificación. Ensayos clínicos como asunto. Medicina familiar y comunitaria. Análisis estadístico. Distribución aleatorizada. Comités de ética en investigación.

Abstract

Objective: To evaluate some methodological and ethical quality variables of clinical trials (CTs) published in 10 family medicine journals. **Methods:** Quality descriptive study of 10 family medicine journals including CTs in humans published since 2010 to 2013. We obtained 141 CT and 2447 were excluded. **Results:** CTs parallels controlled in 92.9% (95% confidence interval [95% CI]: 92.0-93.9). Masked randomization in 72.3% (95% CI: 71.7-73.1), decentralized in 51.8% (95% CI: 51.4-52.4) and using as control an active treatment in 82.2% (95% CI: 81.5-83.1). Wrote informed consent in 48.9% (95% CI: 48.5-49.5) and it was not withdrawn in 56.0% of cases (95% CI: 55.5-56.7). Approval by clinical research ethics committee (CREC) in 134, and there was no conflict of interest in 117 CTs. Average κ was 0.96 (95% CI: 0.93-0.99). **Conclusions:** We observe an increase in some quality variables like masked randomization (19.6%) and approval by CREC (75%) post CONSORT, in CTs published in 10 family medicine journals (2010-2013).

KEY WORDS: Checklist. Clinical trials as topic. Family practice. Statistical analysis. Random allocation. Research ethics committee.

Correspondencia:

Alberto Castaño-García
Avda. Pamplona, 19, Apto. 210
31010 Pamplona, España
E-mail: albertodoc@gmail.com

Fecha de recepción en versión modificada: 02-10-2016
Fecha de aceptación: 04-10-2016
DOI://dx.doi.org/10.24875/GMM.17002699

Gac Med Mex. 2018;154:92-104
Contents available at PubMed
www.gacetamedicademexico.com

Introducción

En la investigación actual, el ensayo clínico aleatorizado es el método de referencia para evaluar nuevas terapias o estrategias en medicina¹. Una baja calidad en su diseño o publicación² podría llevar a una minusvaloración del riesgo o incluso a falsedad, al utilizar tratamientos que podrían ser menos beneficiosos o incluso perjudiciales³, pudiendo tener consecuencias en la práctica clínica diaria.

Existen diversos modos de evaluar la calidad de las publicaciones de ensayos clínicos⁴. Dechartres, et al⁵ identificaron 26 escalas o listas-guía derivadas de CONSORT (CONsolidated Standards Of Reporting Trials). Encontraron un aumento de información sobre aspectos metodológicos, pero solo un 29% de asignación enmascarada adecuada, mostrando que la evaluación de la calidad todavía presenta problemas importantes. Revisiones sistemáticas actuales sugieren que la mejoría en la calidad de los ensayos clínicos aleatorizados está asociada con el uso de las normas CONSORT^{6,7}.

Los ensayos tipo *cluster* (aleatorización por grupos), comparados con otros individuales con el mismo número de sujetos, son poco eficientes y tienen un menor poder estadístico⁸.

Hasta la fecha, no hemos encontrado ningún estudio que utilice una lista-guía similar a la de CONSORT en revistas de la especialidad de medicina de familia⁹, por lo que hemos desarrollado una nueva herramienta más precisa para intentar profundizar en diferentes aspectos de la calidad metodológica y ética de los ensayos clínicos publicados en estas revistas.

El objetivo es analizar de una manera cuantitativa y cualitativa algunas variables que representan parte de la calidad metodológica y ética de los ensayos clínicos publicados entre 2010-2013 en las diez revistas de la especialidad de medicina de familia con mayor factor de impacto que publican ensayos clínicos.

Métodos

Se realizó un estudio descriptivo de calidad de los artículos publicados en las diez revistas de la especialidad de medicina de familia con mayor factor de impacto que publican ensayos clínicos.

Se realizó una búsqueda manual en las revistas *Annals of Family Medicine*, *Atención Primaria*, *Australian Family Physician*, *BioMed Central Family Practice*, *British Journal of General Practice*, *Canadian*

Family Physician, *European Journal of General Practice*, *Family Practice*, *Journal of American Board of Family Medicine* y *Scandinavian Journal of Primary Health Care*.

En esta revisión transversal, se tomó como base la declaración PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*)¹⁰, pero solo en las revistas señaladas al ser específicas de la especialidad de medicina de familia, publicar ensayos clínicos y tener el mayor factor de impacto (Journal Citation Report). La estrategia de búsqueda en PubMed se hizo con los términos “(((“family practice”) OR “general practice” OR “primary care” OR “family medicine”) AND “clinical trial” AND “quality”), clinical trial, humans”. Se encontraron 1096 artículos y además se revisaron los escogidos en *Related citations* de todos los artículos hasta dos niveles. En Web of Knowledge se utilizó el criterio de búsqueda “Topic=(“family practice”) OR Topic=(“general practice”) OR Topic=(“primary care”) OR Topic=(“family medicine”) AND Topic=(“clinical trial”) AND Topic=(“quality”), y se ajustaron los límites a “Document Types=(CLINICAL TRIAL) AND Research Domains=(SCIENCE TECHNOLOGY) AND Research Areas=(HEALTH CARE SCIENCES SERVICES OR GENERAL INTERNAL MEDICINE)”. Se encontraron 5749 estudios. En el buscador Google académico se siguió el criterio de búsqueda “(((“family practice”) OR “general practice” OR “primary care” OR “family medicine”) AND “clinical trial” AND “quality”)” y se encontraron 20.600 artículos. Finalmente, en Scopus se localizaron 1925 artículos con el criterio de búsqueda “(TITLE-ABS-KEY(“family practice”) OR TITLE-ABS-KEY(“general practice”) OR TITLE-ABS-KEY(“primary care”) OR TITLE-ABS-KEY(“family medicine”) AND TITLE-ABS-KEY(“clinical trial”) AND TITLE-ABS-KEY(“quality”)) AND DOCTYPE (ar) AND (LIMIT-TO (SUBJAREA,“MEDI”))”. Uniendo las tres fuentes bibliográficas se obtuvo un total de 2588 artículos originales en las diez revistas elegidas para la evaluación.

Mediante una lista-guía se analizaron los ensayos clínicos publicados en las diez revistas de la especialidad de medicina de familia con mayor factor de impacto que publiquen ensayos clínicos. *Annals of Family Medicine* es la que tiene mayor factor de impacto (4.613)¹¹. Estas revistas siguen las normas CONSORT^{4,12} y las recomendaciones del International Committee of Medical Journal Editors (ICMJE)¹³. Se siguieron las declaraciones CONSORT⁴ y Helsinki¹⁴, que proporcionan listas-guía¹³ con una mayor calidad en la publicación de ensayos clínicos¹⁵.

Se incluyeron ensayos clínicos publicados en humanos, tanto pacientes como voluntarios sanos, entre enero de 2010 y diciembre del 2013. Puesto que las normas CONSORT se publicaron en 2001 por primera vez, se realizó este estudio posteriormente para aumentar la calidad^{7,16}.

Se actualizó una lista-guía¹⁷ basada en CONSORT¹⁸ con 80 ítems (ocho apartados), que se aumentaron a 133 (11 apartados). Siguiendo la Declaración de Helsinki¹⁹ y la gráfica PRECIS (*PRagmatic-Explanatory Continuum Indicator Summary*)²⁰, se modificaron o crearon ítems para adaptarlos a la especialidad de medicina de familia. Se crearon variables más explícitas que ayudasen a identificar de una manera más específica el uso del placebo en la comparación de los tratamientos. Se introdujeron los criterios de la Agency for Healthcare Research and Quality (AHRQ)²¹, la escala de Jadad. Se analizó si existían DSMB (Data and Safety Monitoring Board), reglas de parada (Haybittle-Peto, Pocock, O'Brien-Fleming)²², análisis de futilidad²³, ensayos tipo secuencial u optimizado, y comité ético de investigación clínica (CEIC); también si se indicaban la raza o la lengua de los sujetos participantes. Se investigaron los incentivos dados, la formación específica en bioética de los autores, el tipo de conflictos de intereses y la influencia del patrocinador en los ensayos clínicos.

Dos autores (FGG y ACG) revisaron independientemente los 141 artículos elegidos con la lista-guía. Utilizaron instrucciones detalladas para la cumplimentación de las variables (ítems) y las discrepancias fueron resueltas por consenso. Se calcularon la frecuencia, el porcentaje y el intervalo de confianza del 95% (IC 95%) de Clopper-Pearson²⁴. También se obtuvieron la media, la mediana, la moda, el mínimo, el máximo, la desviación típica y su IC 95% para las variables cuantitativas. Se obtuvieron el κ medio y su IC 95%²⁵. Se utilizó el paquete estadístico SPSS 20.0.

Resultados

Se encontraron 141 ensayos clínicos que cumplieran los criterios de inclusión (Fig. 1), ya que la mayoría fueron estudios observacionales. La revista *The Journal of Family Practice* no publicó ningún ensayo clínico durante el periodo de estudio.

En la tabla 1 se presentan las principales variables metodológicas analizadas en este estudio. Los resultados se muestran como porcentaje o en número absoluto, según la conveniencia para destacar el dato en cuestión.

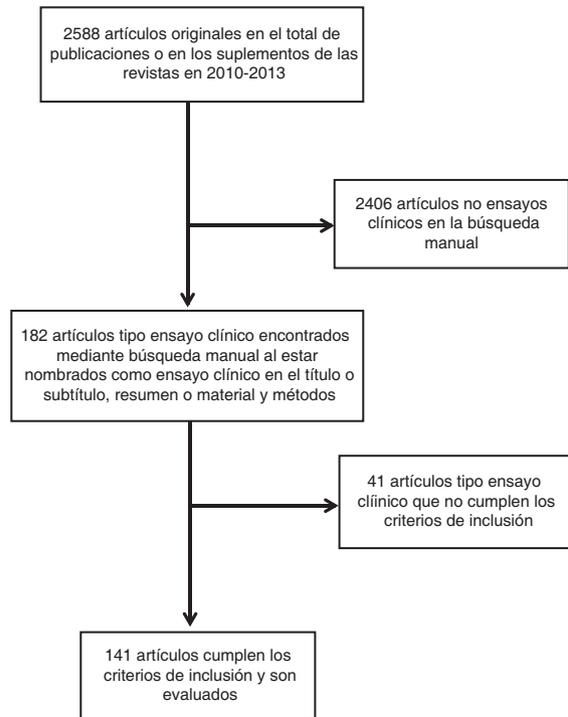


Figura 1. Diagrama de flujo simplificado de la revisión transversal.

Completando la tabla 1, en «País de los autores», dentro del grupo «Otros europeos» el que con más frecuencia se encontró fueron los Países Bajos (22 ensayos clínicos), y en el grupo «Otros» fue Canadá (14 ensayos clínicos). En cuanto al «Continente de estudio», en Europa el país más frecuente fueron los Países Bajos (22), seguido del Reino Unido (11), Suecia (9), España (8) y Dinamarca (6); en el grupo «Otros» destaca Nueva Zelanda (tres ensayos clínicos).

En el apartado de «Justificación y objetivos» de nuestra lista-guía se observó que aparece la denominación de ensayo clínico (*trial*) en el título o el subtítulo en 107, y se consideró dudoso (no lo nombra específicamente) en 33; los ensayos clínicos relacionados aparecieron en 135 y dudoso en dos; al menos una revisión sistemática citada en 128 ensayos clínicos y dudoso en tres; la referencia a una publicación anterior del mismo ensayo clínico en 44, y la hipótesis *a priori* aparece en todos.

Sobre el «Tipo de ensayo clínico», el 54.6% (IC 95%: 54.1-55.2) se realizó en más de cinco centros, el 19.1% (IC 95%: 19.0-19.4) en cinco o menos centros, y el 17.7% (IC 95%: 17.6-18.0) fueron unicéntricos. En un 71.6% (IC 95%: 71.0-72.4) algún autor era epidemiólogo o estadístico y actuó como colaborador en el 14.9% (IC 95%: 14.8-15.1).

Tabla 1. Descripción de la muestra y variables metodológicas: frecuencias, porcentajes e intervalos de confianza del 95%

Variable	Valor	Frecuencia	%	IC 95%
País de los autores	EE.UU.	36	25.5	25.3 - 25.8
	Reino Unido	11	7.8	7.8 - 7.9
	España	10	7.1	7.1 - 7.2
	Escandinavia	18	12.8	12.7 - 13.0
	Otros europeos	31	22.0	21.8 - 22.3
	Otros	35	24.8	24.6 - 25.1
Continente donde se realiza el estudio	Norteamérica	50	35.5	35.2 - 35.9
	Latinoamérica	2	1.4	1.4 - 1.5
	Asia	4	2.8	2.8 - 2.9
	Australia	8	5.7	5.7 - 5.8
	Europa	71	50.4	50.0 - 51.0
	Otros	6	4.3	4.3 - 4.4
Tipo de financiación	Fines no lucrativos: pública	78	55.3	54.8 - 55.9
	Fines no lucrativos: privada*	71	50.4	50.0 - 51.0
	Industria farmacéutica*	17	11.9	11.8 - 12.1
	No se indica	8	5.7	5.7 - 5.8
	No existe	3	2.1	2.1 - 2.2
Sistema de asignación aleatoria	Aleatorización simple	49	34.8	34.5 - 35.2
	Por bloques	39	27.7	27.5 - 28.1
	Estratificada	23	16.3	16.2 - 16.5
	Apareamiento	5	3.6	3.6 - 3.7
	Minimización*	5	3.6	3.6 - 3.7
	Por grupos (<i>cluster</i>)*	46	32.5	32.3 - 32.9
	Otros*	2	1.4	1.4 - 1.5
	No aleatorizado: simple	4	2.8	2.8 - 2.9
¿Se siguió un método de aleatorización enmascarada?	Sí	102	72.3	71.7 - 73.1
	No	23	16.3	16.2 - 16.5
	Dudoso	11	7.8	7.8 - 7.9
	No existe aleatorización	5	3.6	3.6 - 3.7
¿Se especifica el método de aleatorización enmascarada empleado?	Descentralizada	73	51.8	51.4 - 52.4
	Sobres opacos con apertura secuencial*	21	14.9	14.8 - 15.1
	Otro*	18	12.8	12.7 - 13.0
	No se menciona	10	7.1	7.1 - 7.2
	No existe	29	20.6	20.5 - 20.9
Tipo de control	Placebo	13	9.2	9.1 - 9.4
	Tratamiento activo*	116	82.2	81.5 - 83.1
	No tratamiento*	12	8.5	8.5 - 8.6
	No se indica	1	0.7	0.7 - 0.7
	Otro	1	0.7	0.7 - 0.7
¿Se indica el grado de enmascaramiento de valoración?	Abierto o no ciego	17	11.9	11.8 - 12.1
	Simple ciego	18	12.8	12.7 - 13.0
	Doble ciego	10	7.1	7.1 - 7.2
	Evaluación ciega por terceros	26	18.4	18.3 - 18.7
	Triple ciego	6	4.3	4.3 - 4.4
	No se indica	64	45.4	45.0 - 45.9
¿Se explica por qué el ensayo no es ciego?	Sí	14	9.9	9.8 - 10.1
	No	58	41.1	40.8 - 41.6
	Dudoso	10	7.1	7.1 - 7.2
	Existe enmascaramiento	59	41.8	41.5 - 42.3
¿Se indica el método de enmascaramiento?	Sí	31	22.0	21.8 - 22.3
	No	22	15.6	15.5 - 15.8
	Dudoso	8	5.7	5.7 - 5.8
	No existe enmascaramiento	80	56.7	56.2 - 57.4

(Continúa)

Tabla 1. Descripción de la muestra y variables metodológicas: frecuencias, porcentajes e intervalos de confianza del 95% (Continuación)

Variable	Valor	Frecuencia	%	IC 95%
Tipo de tratamiento	Fármaco ya comercializado	10	7.1	7.1 - 7.2
	Dispositivo médico	5	3.6	3.6 - 3.7
	Método de intervención médica*	73	51.8	51.4 - 52.4
	Otro: acueroobic, yoga...*	53	37.6	37.3 - 38.1
Resultado del ensayo	Tratamiento significativamente mejor	50	35.5	35.2 - 35.9
	Diferencia por el/los autor/es, pero no es significativa	22	15.6	15.5 - 15.8
	Existe una tendencia hacia el tratamiento	12	8.5	8.5 - 8.6
	Dudoso	1	0.7	0.7 - 0.7
	No existe diferencia	44	31.2	31.0 - 31.6
	Existe una tendencia hacia el control	1	0.7	0.7 - 0.7
	Control es significativamente mejor	3	2.1	2.1 - 2.2
	Otro: mejoría significativa en una variable principal	8	5.7	5.7 - 5.8
Valoración de la respuesta (endpoint)	Muerte	3	2.1	2.1 - 2.2
	Diagnóstico de laboratorio	21	14.9	14.8 - 15.1
	Diagnóstico sintomático*	5	3.6	3.6 - 3.7
	Otro: manómetro de presión arterial, fichas médicas...*	55	39.0	38.7 - 39.5
	No existe	1	0.7	0.7 - 0.7
	Cuestionario*	79	55.9	55.5-56.6
	Escala de medición*	16	11.2	11.1-11.4

* Suma del mismo resultado en varias variables idénticas.

En cuanto al «Diseño del ensayo clínico», un 92.9% (IC 95%: 92.0-93.9) fueron ensayos clínicos controlados de grupos paralelos y seis fueron controlados factoriales. Según los criterios de la AHRQ, hubo mayor tendencia hacia el ensayo clínico pragmático que al explicativo.

En 132 ensayos clínicos se realizó una comparación basal en una tabla y apareció la *P* en un 57.4% (IC 95%: 56.9-58.1). No se encontraron ensayos clínicos no controlados. Solo el de Licciardone, et al.²⁶ fue secuencial y utilizó la regla de parada (O'Brien Fleming). En ninguno apareció la gráfica PRECIS²⁰ y tampoco hubo ensayos clínicos optimizados. La eficacia del enmascaramiento solo se observó en siete ensayos clínicos. Las retiradas debidas a efectos adversos se encontraron en 11 ensayos clínicos y fueron dudosas en 26 estudios.

En cuanto a «Selección de los sujetos», se incluyeron sujetos blancos en un 27.0% (IC 95%: 26.8-27.4), afroamericanos en un 18.4% (IC 95%: 18.3-18.7), hispanos en un 14.2% (IC 95%: 14.1-14.4) y asiáticos en 11,3% IC 95% (11,2-11,5) en 16 estudios. No se indicó la raza en un 71.6% (IC 95%: 71.0-72.4), no se especificó en 25 ensayos clínicos (17,7% IC 95% (17,6-18,0), y en nueve los sujetos fueron minorías étnicas, como los indios americanos y los vietnamitas. La inclusión de población emigrante o con escasos recursos económicos ha aumentado en los últimos 3 años (20 ensayos clínicos) y se observaron personas «vulnerables»²⁷.

En «Tamaño de la muestra», las variables cuantitativas metodológicas se indican en la tabla 2. Sus estadísticos descriptivos y el IC 95% de la desviación típica²⁸ se muestran en la tabla 3.

El error alfa fue del 87.9% (IC 95%: 87.1-88.8), el beta fue del 83.7% (IC 95%: 82.9-84.6), la magnitud de la diferencia mínima a encontrar fue del 80.9% (IC 95%: 80.2-81.8) y el diagrama de flujo se observó en el 86.5% (IC 95%: 85.7-87.4).

Respecto a «Descripción del tratamiento», con un tratamiento estándar eficaz no se utilizó placebo o no intervención como control en un 79.4% (IC 95%: 78.7-79.3). Una razón metodológica sólida y convincente para la utilización de placebo o no intervención se dio en 11 ensayos clínicos, y la determinación de eficacia y seguridad en 13. No se encontró ningún artículo en el que el uso del placebo o no intervención produjese un riesgo adicional de daño grave e irreversible. El ensayo clínico fue reproducible en todos los casos salvo en dos, y uno de ellos fue dudoso. En 48 estudios hubo tratamientos concomitantes (permitidos o prohibidos) y en 68 la intervención fue inocua. En un 31.2% (IC 95%: 31.0-31.6) se observaron incidentes adversos graves en general (incluso fallecimientos) y fueron dudosos en 21 estudios. El análisis por intención de tratar se realizó en 64 ensayos clínicos, no se indicó (se realizó, pero no se nombró) en siete y fue dudoso (se nombró, pero no se realizó adecuadamente) en cinco. Según la puntuación Jadad, 70

Tabla 2. Calidad de la información en variables cuantitativas: frecuencias, porcentajes e intervalos de confianza del 95%

Variable: calidad de la información	Frecuencia	%	IC 95%
Sujetos evaluados			
Mencionado en el artículo	114	80.9	80.2 - 81.8
Calculado o sobreentendido por el lector	8	5.7	5.7 - 5.8
Omitido e incalculable	17	11.9	11.8 - 12.1
No se muestran sujetos elegibles	2	1.4	1.4 - 1.5
Sujetos evaluados (<i>cluster</i>)			
Mencionado en el artículo	35	24.8	24.6 - 25.1
Calculado o sobreentendido por el lector	5	3.6	3.6 - 3.7
Omitido e incalculable	4	2.8	2.8 - 2.9
No es un ensayo clínico tipo <i>cluster</i>	97	68.8	68.2 - 69.6
Sujetos que rechazan participar			
Mencionado en el artículo	48	34.1	33.8 - 34.6
Calculado o sobreentendido por el lector	72	51.1	50.7 - 51.7
Omitido e incalculable	19	13.5	13.4 - 13.7
No mostrados	2	1.4	1.4 - 1.5
Sujetos que rechazan participar (<i>cluster</i>)			
Mencionado en el artículo	16	11.3	11.2 - 11.5
Calculado o sobreentendido por el lector	24	17.0	16.9 - 17.2
Omitido e incalculable	4	2.8	2.8 - 2.9
No es un ensayo clínico tipo <i>cluster</i>	97	68.8	68.2 - 69.6
Sujetos aleatorizados			
Mencionado en el artículo	116	82.3	81.6 - 83.2
Calculado o sobreentendido por el lector	19	13.5	13.4 - 13.7
Omitido e incalculable	3	2.1	2.1 - 2.2
No se muestran	3	2.1	2.1 - 2.2
Sujetos aleatorizados (<i>cluster</i>)			
Mencionado en el artículo	33	23.4	23.2 - 23.7
Calculado o sobreentendido por el lector	10	7.1	7.1 - 7.2
Omitido e incalculable	1	0.7	0.7 - 0.7
No es un ensayo clínico tipo <i>cluster</i>	97	68.8	68.2 - 69.6
Recibieron tratamiento asignado			
Mencionado en el artículo	71	50.4	50.0 - 51.0
Calculado o sobreentendido por el lector	64	45.4	45.0 - 45.9
Omitido e incalculable	6	4.3	4.3 - 4.4
Recibieron tratamiento asignado (<i>cluster</i>)			
Mencionado en el artículo	20	14.2	14.1 - 14.4
Calculado o sobreentendido por el lector	23	16.3	16.2 - 16.5
Omitido e incalculable	1	0.7	0.7 - 0.7
No es un ensayo clínico tipo <i>cluster</i>	97	68.8	68.2 - 69.6
Retiraron su consentimiento informado			
Mencionado en el artículo	22	15.6	15.5 - 15.8
Calculado o sobreentendido por el lector	78	55.3	54.8 - 55.9
Omitido e incalculable	16	11.3	11.2 - 11.5
No existe consentimiento informado	25	17.7	17.6 - 18.0
Retiraron su consentimiento informado (<i>cluster</i>)			
Mencionado en el artículo	6	4.3	4.3 - 4.4
Calculado o sobreentendido por el lector	25	17.7	17.6 - 18.0
Omitido e incalculable	4	2.8	2.8 - 2.9
No existe consentimiento informado o no ensayo clínico tipo <i>cluster</i>	106	75.2	74.5 - 76.0
Pérdidas durante el seguimiento			
Mencionado en el artículo	24	17.0	16.9 - 17.2
Calculado o sobreentendido por el lector	92	65.2	64.6 - 65.9
Omitido e incalculable	25	17.7	17.6 - 18.0

(Continúa)

Tabla 2. Calidad de la información en variables cuantitativas: frecuencias, porcentajes e intervalos de confianza del 95% (Continuación)

Variable: calidad de la información	Frecuencia	%	IC 95%
Pérdidas durante el seguimiento (<i>cluster</i>)			
Mencionado en el artículo	13	9.2	9.1 - 9.4
Calculado o sobreentendido por el lector	24	17.0	16.9 - 17.2
Omitido e incalculable	6	4.3	4.3 - 4.4
No se muestra o no es un ensayo clínico tipo <i>cluster</i>	98	69.5	68.9 - 70.3
Análisis principal sobre un total			
Mencionado en el artículo	62	44.0	43.6 - 44.5
Calculado o sobreentendido por el lector	75	53.2	52.8 - 53.8
Omitido e incalculable	4	2.8	2.8 - 2.9
Análisis principal sobre un total (<i>cluster</i>)			
Mencionado en el artículo	12	8.5	8.5 - 8.6
Calculado o sobreentendido por el lector	28	19.9	19.8 - 20.2
Omitido e incalculable	3	2.1	2.1 - 2.2
No se muestra o no es un ensayo clínico tipo <i>cluster</i>	98	69.5	68.9 - 70.3
Casos <i>endpoint</i> deseaban evitar			
Mencionado en el artículo	8	5.7	5.7 - 5.8
Calculado o sobreentendido por el lector	40	28.4	28.2 - 28.8
Omitido e incalculable	92	65.2	64.6 - 65.9
No existe <i>endpoint</i>	1	0.7	0.7 - 0.7
Casos <i>endpoint</i> deseaban evitar (<i>cluster</i>)			
Calculado o sobreentendido por el lector	5	3.6	3.6 - 3.7
Omitido e incalculable	37	26.2	26.0 - 26.5
No se muestra o no es un ensayo clínico tipo <i>cluster</i>	99	70.2	69.6 - 71.0

Tabla 3. Estadísticos más representativos de las variables cuantitativas

Variable	Media	Mediana	Moda	Mínimo	Máximo	Desviación típica (IC 95%)
Sujetos evaluados	9098.15	641.00	91 (varias)	32	572546	53871.94 (48233.54-61015.48)
Sujetos evaluados (<i>cluster</i>)	108.68	69.50	40	9	373	95.02 (85.07-107.62)
Sujetos rechazan participar	363.08	3.00	0	0	18882	1936.66 (1733.96-2193.46)
Sujetos rechazan participar (<i>cluster</i>)	25.34	0.00	0	0	203	53.44 (47.85-60.53)
Sujetos aleatorizados	1150.84	280.00	125 (varias)	11	38147	3936.06 (3524.10-4457.99)
Sujetos aleatorizados (<i>cluster</i>)	107.58	51.00	45 (varias)	9	1361	211.22 (189.11-239.23)
Recibieron tratamiento asignado	1161.22	289.00	237	11	38147	4057.90 (3633.19-4595.99)
Recibieron tratamiento asignado (<i>cluster</i>)	101.35	51.00	41	9	1361	208.97 (187.10-236.68)
Retiraron su consentimiento informado	4.22	0.00	0	0	97	14.20 (12.71-16.08)
Retiraron su consentimiento (<i>cluster</i>)	0.74	0.00	0	0	16	3.10 (2.77-3.51)
Pérdidas durante el seguimiento	74.78	13.00	0	0	1612	233.09 (208.69-264.00)
Pérdidas durante el seguimiento (<i>cluster</i>)	4.11	0.00	0	0	54	9.92 (8.88-11.23)
Análisis principal sobre un total	1062.51	234.00	164	11	36092	3864.55 (3460.07-4377.00)
Análisis principal sobre un total (<i>cluster</i>)	100.95	47.00	40	9	1361	216.22 (193.59-244.89)
Casos <i>endpoint</i> deseaban evitar	181.46	73.00	7 (varias)	7	1912	348.60 (312.11-394.82)
Casos <i>endpoint</i> deseaban evitar (<i>cluster</i>)	3.20	0.00	0	0	16	7.15 (6.40-8.10)

estudios tuvieron buena calidad (3 o 4), 54 baja calidad (1 o 2) y 11 excelente (5).

Al analizar el apartado «Desarrollo del ensayo y evaluación de la respuesta», la variable de estudio principal (*endpoint*) no se describió en tres ensayos clínicos y fue dudosa en el 9.2% (IC 95%: 9.1-9.4). En 139 ensayos clínicos se especificaron los métodos estadísticos para evaluar los resultados. Se aplicó análisis multivariante en un 61.7% (IC 95%: 61.2-62.4) y el número necesario a tratar solo en cuatro. Se encontró un DSMB en un 5.7% (IC 95%: 5.7-5.8). Se realizaron cambios en el protocolo durante el ensayo clínico en nueve, y en tres además se informó de ello al CEIC. Finalmente, se observó que un 29.1% (IC 95%: 28.9-29.5) indican que siguen las normas CONSORT.

En la tabla 4 se presentan las variables éticas más importantes. En siete estudios se tuvo en cuenta solo la Declaración de Helsinki (menor calidad que la aprobación por un CEIC), y se publicó un estudio en el que no estaba indicado que la siguiese (muy baja calidad ética). En nueve ensayos clínicos se observó que se había dado una información parcial a individuos que podían ser incluidos sobre objetivos, financiación, etc., del ensayo clínico, y en ocho ensayos clínicos con una comprobación dudosa de que se entendió dicha información¹⁹.

Se encontraron sujetos participantes menores de edad en 14 ensayos clínicos. Los padres o representantes cumplimentaron el consentimiento informado en cuatro ensayos clínicos y dudosamente lo cumplimentaron en 10. En un 2.8% (IC 95%: 2.8-2.9) los sujetos menores participaron como característica necesaria de estudio, y dudosa en el 7.1% (IC 95%: 7.1-7.2).

Se observó que se guardaba la confidencialidad de la información personal de los sujetos participantes en un 44.7% (IC 95%: 44.3-45.2), y en ningún estudio se incluyó en el consentimiento informado la posibilidad de guardar esos datos para utilizarlos en el futuro. Un 4.3% (IC 95%: 4.3-4.4) de los estudios mostraron un daño en los sujetos participantes, pero no se indicó que existiese compensación. En un ensayo clínico, el protocolo evidenció el arreglo del daño potencial sufrido para no salir con secuelas, y en dos ensayos clínicos, los arreglos para que al final se equiparase el potencial resultado beneficioso de todos los sujetos participantes. Además, en un 32.6% (IC 95%: 32.4-33.0) no se nombró un protocolo de estudio. Cinco ensayos clínicos proporcionaron la intervención más beneficiosa al resto de los sujetos participantes tras su finalización, pero solo en un caso

se informó de esa posibilidad en el consentimiento informado. Se mostraron los resultados generales del estudio al finalizarlo a todos los sujetos participantes en un 1.4% (IC 95%: 1.4-1.5) de los estudios.

Ningún estudio evidenció una formación específica en bioética de los autores. En todos los ensayos clínicos se informó de posibles conflictos de intereses. No se describió la influencia del patrocinador en un 57.4% (IC 95%: 56.9-58.1). Los estudios analizados aparecieron registrados (ClinicalTrials.gov o similar, según el país de estudio) en un 55.3% (IC 95%: 54.8-55.9).

Se obtuvo entre 112 variables comparables un κ medio de 0.96 (IC 95%: 0,93-0,99).

Discusión

Es posible afirmar que el tipo de ensayo clínico más frecuente publicado en las 10 revistas revisadas es multicéntrico, con financiación pública, controlado, paralelo, aleatorizado simple, con una aleatorización enmascarada descentralizada, usando un tratamiento activo como control, con una duración de 12 meses, en pacientes blancos y con un método de intervención médica, con un resultado positivo en un 50% (medido mediante un cuestionario). En el 45% se realiza el análisis por intención de tratar. Se muestran los efectos adversos graves y la buena calidad (Jadad 3). Además, el consentimiento informado (oral o no mencionado), la no incentivación y la no existencia de conflictos de intereses aparecen con mayor frecuencia que en otras publicaciones citadas.

Por otra parte, la lista-guía utilizada en este análisis tiene más ítems que la lista CONSORT para considerar aspectos específicos en los cuales CONSORT no incide lo suficiente según nuestro criterio. Además, al basarnos en la última Declaración de Helsinki profundizamos aún más en diversos aspectos éticos que las normas CONSORT, a nuestro juicio, no consideran lo suficiente.

Si realizamos un análisis comparativo con los ensayos clínicos del estudio realizado sobre las revistas *New England Journal of Medicine*, *British Medical Journal*, *Journal of the American Medical Association* y *Lancet*¹⁷, encontramos que ha aumentado la aleatorización enmascarada progresivamente: del 33.1% en 1993-1995 al 56.7% en 2004¹⁷, y aún mayor en nuestro estudio. Podría explicarse porque tras CONSORT se insistió en que esta era fundamental para disminuir el sesgo de selección. Por el contrario, el doble ciego es menos frecuente que en revisiones anteriores (43.9% en 1993-1995 y 43.3% en 2004)¹⁷,

Tabla 4. Variables éticas: frecuencias, porcentajes e intervalos de confianza del 95%

Variable	Valor	Frecuencia	%	IC 95%
¿Se menciona si se obtuvo el consentimiento informado?	Sí	111	78.7	78.0 - 79.6
	No	30	21.3	21.2 - 21.6
El consentimiento informado, ¿se obtuvo de manera voluntaria?	Sí	10	7.1	7.1 - 7.2
	Dudoso	101	71.6	71.0 - 72.4
	No se obtuvo consentimiento informado	30	21.3	21.2 - 21.6
¿Se especifica el tipo de consentimiento informado, y si es atestiguado y documentado?	Oral (no se menciona testigo)	1	0.7	0.7 - 0.7
	Escrito	69	48.9	48.5 - 49.5
	No se menciona (si es escrito u oral)	41	29.1	28.9 - 29.5
	No se obtuvo consentimiento informado	30	21.3	21.2 - 21.6
¿Recibió información el paciente al obtener el consentimiento informado?	Sí	43	30.5	30.3 - 30.9
	No	68	48.2	47.8 - 48.8
	No se obtuvo consentimiento informado	30	21.3	21.2 - 21.6
¿Se obtuvo el consentimiento informado bajo una relación de dependencia o bajo presión?	Relación de dependencia	44	31.2	31.0 - 31.6
	No	14	9.9	9.8 - 10.1
	Dudoso	53	37.6	37.3 - 38.1
	No se obtuvo consentimiento informado	30	21.3	21.2 - 21.6
¿Designada una persona calificada sin relación para la obtención del consentimiento informado?	Sí	2	1.4	1.4 - 1.5
	No	35	24.8	24.6 - 25.1
	Dudoso	60	42.6	42.3 - 43.1
	No consentimiento informado o no bajo presión o dependencia	44	31.2	31.0 - 31.6
¿El consentimiento informado del sujeto participante fue retirado en algún momento del ensayo clínico?	Sí	21	14.9	14.8 - 15.1
	No	79	56.0	55.5 - 56.7
	Dudoso	11	7.8	7.8 - 7.9
	No se obtuvo consentimiento informado	30	21.3	21.2 - 21.6
¿Algún incentivo a investigadores, centros participantes o sujetos?	No	96	68.1	67.5 - 68.9
	A investigadores	26	18.4	18.3 - 18.7
	A centros participantes	4	2.8	2.8 - 2.9
	A sujetos participantes*	16	11.2	11.1 - 11.4
	A ayudantes de investigadores*	6	4.3	4.3 - 4.4
	A médicos participantes (<i>cluster</i>)*	7	5.0	5.0 - 5.1
Si la respuesta es afirmativa, especificar el tipo y la cuantía:	Prima a investigadores	7	5.0	5.0 - 5.1
	Subvención o beca a investigadores*	23	16.3	16.2 - 16.5
	Prima a centros participantes	4	2.8	2.8 - 2.9
	Prima a los sujetos por participar*	16	11.2	11.1 - 11.4
	Prima a los ayudantes*	6	4.3	4.3 - 4.4
	Prima a médicos (<i>cluster</i>) por participar*	7	5.0	5.0 - 5.1
	No se dieron incentivos	96	68.1	67.5 - 68.9
El incentivo dado, ¿ha sido proporcionado?	Sí	41	29.1	28.9 - 29.5
	No	4	2.8	2.8 - 2.9
	No se dieron incentivos	96	68.1	67.5 - 68.9
Información proporcionada en relación con conflictos de intereses	No hay conflicto de intereses	117	83.0	82.2 - 83.9
	Remuneración o beca de compañía	12	8.5	8.5 - 8.6
	Algún autor trabaja para compañía*	14	9.9	9.8 - 10.1
	Otra: autor fue asesor o propietario*	7	5.0	5.0 - 5.1
	Algún autor posee acciones de la compañía	1	0.7	0.7 - 0.7

* Suma del mismo resultado en varias variables idénticas.

y también ha ido disminuyendo el placebo como control (37.0% en 1993-1995 y 16.7% en 2004)¹⁷, y aún menor en nuestro análisis; probablemente porque con los años hay más tratamientos estándar para comparar y una mayor exigencia de los organismos controladores para plantearse la salida al mercado de nuevos tratamientos, teniendo que demostrar su eficacia con tratamientos ya estandarizados.

Respecto a los indicadores de la calidad ética, ha aumentado el número de ensayos clínicos en que aparece una referencia a un CEIC y más en los últimos años: 25.5% en 1993-1995 y 20% en 2004¹⁷. Los ensayos clínicos con niños o discapacitados son muy escasos, posiblemente por ser estudios sobre ensayos clínicos publicados en revistas de atención primaria y de medicina interna.

Sinha, et al.²⁹, sobre 42 ensayos clínicos, obtuvieron una puntuación Jadad ≥ 3 en un 40%²⁹, resultado inferior al que encontramos nosotros.

Agha, et al.³⁰ coinciden con nosotros, en su muestra de 57 ensayos clínicos, en que los EE.UU. (con 15 estudios) es el país donde con más frecuencia se realizan ensayos clínicos.

La financiación pública en los 74 ensayos clínicos de Lai, et al.³¹ es la más frecuente, al igual que en nuestro estudio, probablemente porque los ensayos clínicos son muy caros. Además, los efectos adversos importantes son más frecuentes (64.9%; IC 95%: 52.9-75.6) que en nuestro estudio, posiblemente al no estudiar intervenciones médicas inocuas sino tratamientos más agresivos. También la calidad metodológica es mayor en nuestro estudio, al solo aparecer en su revisión³¹ en un 45.9% (IC 95%: 34.3-57.9) la variable principal, el cálculo del tamaño muestral en el 39.2% (IC 95%: 28.0-51.2), la aleatorización enmascarada en el 29.7% (IC 95%: 19.7-41.5), el enmascaramiento en el 2.2% (IC 95%: 4.5-19.8) y el análisis por intención de tratar en el 27% (IC 95%: 16.7-37.4).

Comparando diversos aspectos de calidad metodológica de nuestro estudio con otros realizados recientemente en otras especialidades, podemos destacar lo siguiente. En una revisión de 300 ensayos clínicos tipo *cluster*³², el 55% escribieron sobre el cálculo del tamaño muestral (dato inferior a nuestra revisión). En otra revisión³³ de 89 ensayos clínicos aleatorizados sobre endodoncias, se encontró en menor porcentaje, comparando con nuestro estudio, ya que solo el 45% generaron una secuencia de aleatorización para disminuir el riesgo de sesgos, en un 75% no dieron información sobre el método de asignación y un 19.1% tenían un diseño no ciego. Además, solo el 55% habían realizado

de manera adecuada la estimación del cálculo del tamaño muestral, únicamente el 16% obtuvieron intervalos de confianza y el 25% no proporcionaron el valor de p. Finalmente, el 2.2% no realizaron pruebas estadísticas. En el campo de la ortopedia, en una revisión sobre 222 estudios³⁴, en ninguno realizaron un cálculo del tamaño muestral adecuado, en el 33.4% generaron una secuencia de aleatorización adecuada, en solo el 3.7% realizaron una asignación aleatorizada correcta y en el 8.1% hubo un enmascaramiento adecuado (datos inferiores que los de nuestra revisión). En concreto, en los ocho artículos escogidos de la revista *Orthopaedics Traumatology Surgery & Research* encontraron que en el 12.5% no mencionaron la realización de una secuencia de aleatorización adecuada, en el 50% no aportaron una asignación aleatoria correcta y en el 25% hubo un enmascaramiento adecuado (datos también inferiores a los nuestros). En el campo del aparato respiratorio, en una revisión sobre 176 ensayos³⁵ se encontró que el 53% mencionaron una generación adecuada de la secuencia de aleatorización, en el 38% era correcta la asignación aleatorizada, el 45% eran doble ciego, en el 70% el cálculo del tamaño muestral era adecuado, en el 50% realizaron un análisis por intención de tratar y el 69% incluyeron un diagrama de flujo. Salvo el porcentaje del doble ciego, el resto son bastante inferiores a los encontrados en nuestra revisión. En una revisión³⁶ de 305 estudios publicados en revistas sobre la diabetes, los autores encontraron que en un 35.4% los ensayos mostraron una adecuada generación de secuencia de aleatorización, en el 28.5% la asignación aleatorizada fue correcta y en el 23.8% realizaron un análisis por intención de tratar (porcentajes menores que los encontrados en nuestro estudio). Sobre una revisión³⁷ de 321 artículos sobre anestesia epidural, los autores encontraron un valor medio en la escala de Jadad de 2 (inferior al observado en nuestra revisión) y un incremento general en la calidad metodológica en los publicados después de la aparición de las normas CONSORT, con lo cual coincidimos en nuestra revisión. En el campo de los trasplantes, los autores de una revisión³⁸ de 290 estudios encontraron que un 43% tenían buena calidad según la escala de Jadad (dato un poco inferior al nuestro) y en un tercio de estos la asignación aleatoria y el análisis de los datos fueron satisfactorios (bastante más bajo que lo observado en nuestro estudio). Además, la calidad metodológica y la adhesión a las normas CONSORT fue mayor en los ensayos de aquellas revistas que estaban adheridas a CONSORT³⁸. En el tratamiento del cáncer colorrectal, los autores de una revisión³⁹ encontraron

en 40 ensayos clínicos aleatorizados unos datos muy inferiores a los nuestros, ya que solo en el 7.5% aparecía identificado en el título como ensayo clínico aleatorizado, también en solo el 7.5% aplicaron un método de asignación aleatoria de las secuencias, en un 5% mencionaron el tipo de aleatorización y solo en el 2.5% aplicaron un método de enmascaramiento o calcularon el tamaño muestral. Únicamente el 2.5% describieron el diagrama de flujo, las variables de estudio primarias y secundarias con efecto estimado de la muestra. Por último, el 32.5% mostraron las características demográficas y clínicas basales, y el 25% realizaron un análisis por intención de tratar. En una revisión⁴⁰ de 53 ensayos clínicos sobre esclerosis múltiple, los autores encontraron que el 79% fueron patrocinados por una compañía farmacéutica, el 13% fue con financiamiento mixto y solo el 6% por instituciones sin ánimo de lucro. Esto podría comprometer la validez de este estudio, pues es un financiamiento claramente sesgado, justo lo opuesto a lo encontrado en nuestra revisión. La media de pacientes evaluados fue de 297 (rango: 30-2244). Un 83% fueron controlados con placebo y no mostraron sistemas de asignación aleatoria. El método estadístico usado para generar la lista aleatorizada lo especificaron el 81.1%, siendo por bloques el más utilizado (81.4%). El doble ciego lo usaron el 90.6% y solo uno no fue ciego. Entre los ensayos con enmascaramiento simple ciego, el 94.2% mostraron detalles de quién lo había realizado y el 59.6% de cómo se había realizado. Definieron la variable de respuesta principal todos los estudios excepto dos. Un 64.2% fueron ensayos significativamente positivos. En el 79.2% se utilizó el análisis por intención de tratar. El 20.8% no mostraron ningún diagrama de flujo según CONSORT. En 52 estudios se informó sobre el número de pacientes perdidos en el seguimiento. Todos los ensayos mostraron las características basales en una tabla. Al igual que en estudios anteriores, los datos son menos frecuentes, salvo el doble ciego, los ensayos con significación positiva y la muestra de las características basales en una tabla. Los autores de una revisión⁴¹ de 73 ensayos clínicos tipo *cluster* observaron que el 50% se habían realizado en los EE.UU. o en el Reino Unido (porcentaje superior al hallado por nosotros). El cálculo del tamaño muestral lo realizaron el 60%. La información sobre pérdidas en el seguimiento individual apareció en 53 estudios, mientras que sobre los grupos solo en 28. La evaluación ciega por terceros apareció en el 44% (dato superior al de nuestra revisión). En 51 ensayos aparece como coautor un estadístico (menos

que en nuestro estudio) y esta aparición, junto al periodo de publicación, tiene mayor influencia en la calidad que la fuerza de la unión de la revista a las normas CONSORT⁴¹. En una revisión⁴² sobre inmunosupresión en el trasplante de riñón, los autores observaron que, en 179 estudios, todos tuvieron un diseño paralelo. Un 35% no declararon ninguna fuente de financiación o no la tuvieron, un 57% fueron financiados por la industria farmacéutica y un 14% por una entidad pública (datos muy diferentes a los encontrados en nuestra revisión). El 17% fueron multinacionales, reclutando pacientes desde más de un continente: el 26% de Norteamérica, el 44% de Europa, el 7% de Asia y el 11% de otras zonas (porcentajes similares a los nuestros). En el 67%, los estudios fueron publicados en revistas unidas a las normas CONSORT (porcentaje inferior al de nuestro estudio). El número de sujetos participantes fue desde 15 hasta 1645. El 2% de los estudios revisados mostraron resultados en el seguimiento ≤ 1 mes y el 7% en el seguimiento ≤ 6 meses, con un rango de 0.25 a 240 meses (datos también inferiores a los encontrados en nuestro estudio).

Finalmente, si comparamos nuestro estudio con el de Hopewell, et al.⁷, observamos que en los estudios revisados del año 2006 aparece lo siguiente: la variable de respuesta principal se define en el 53%, el cálculo del tamaño muestral en el 45%, el método de aleatorización enmascarada empleado en el 34%, el método de asignación aleatoria en el 25%, el enmascaramiento aparece en el 59%, el diagrama de flujo en el 28%, las pérdidas durante el seguimiento en el 69%, el análisis por intención de tratar en el 31%, la financiación farmacéutica en el 17% y el registro del ensayo clínico en el 9%.

Como Baskerville, et al.¹⁵, solo hemos podido acceder a los datos publicados de los ensayos clínicos y no a los originales, que podrían ser diferentes.

En nuestro análisis existe una variabilidad importante entre los estudios respecto a la medición de las variables principales, debido a los muchos tipos de variables encontradas según el tratamiento estudiado.

Como en el estudio de Hopewell⁷, podemos haber perdido artículos en la búsqueda, ya que la fecha de indización en PubMed y en los otros buscadores lleva un cierto retraso respecto a la publicación del artículo.

Además, aunque hemos realizado una evaluación pareada consensuada con un coeficiente κ medio óptimo, siempre existe la posibilidad del error en la toma

de datos al señalar las respuestas de los ítems en la lista-guía utilizada.

La creación de una lista-guía con tantos ítems ha sido con el fin de analizar las características metodológicas y éticas, generales y de calidad, y no tanto como herramienta habitual para la revisión de ensayos clínicos.

Por último, por limitaciones de espacio no ponemos la lista-guía utilizada ni sus instrucciones para completarla, las cuales podrán estar disponibles para los lectores si así lo desean.

Conclusiones

Tras la publicación de las normas CONSORT ha aumentado de manera considerable la calidad general de los ensayos clínicos publicados, pero existen algunos aspectos que podrían mejorarse. Han aumentado la aleatorización enmascarada en un 19,6%¹⁷ y la aprobación por un CEIC en un 75%¹⁷, pero han disminuido el doble ciego en un 36,2%¹⁶ y el uso de placebo para comparar un nuevo tratamiento en un 7,5%¹⁷ como criterios de calidad metodológica y ética, respectivamente, en los ensayos clínicos publicados en las 10 revistas revisadas de la especialidad de medicina de familia (2010-2013).

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

- Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924-6.
- Pildal J, Chan AW, Hrobjartsson A, et al. Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study. *BMJ*. 2005;330:1049-52.
- Moher D, Altman DG, Schulz KF, et al. Opportunities and challenges for improving the quality of reporting clinical research: CONSORT and beyond. *CMAJ*. 2004;171:349-50.
- Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
- Dechartres A, Charles P, Hopewell S, Ravaut P, Altman DG. Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. *J Clin Epidemiol*. 2011;64:136-44.
- Cui Q, Tian J, Song X, et al. Does the CONSORT checklist for abstracts improve the quality of reports of randomized controlled trials on clinical pathways? *J Eval Clin Pract*. 2014;20:827-33.
- Hopewell S, Dutton S, Yu LM, et al. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ*. 2010;340:c723.
- Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ*. 2004;328:702-8.
- Ahmed Ali U, van der Sluis PC, Issa Y, et al. Trends in worldwide volume and methodological quality of surgical randomized controlled trials. *Ann Surg*. 2013;258:199-207.
- Moher D, Liberati A, Tetzlaff J, et al.; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6:e1000097.
- ISI Web of Knowledge. Journal Citations Reports. Thomson Reuters; 2014. (Consultado el 4 de febrero de 2014.) Disponible en: <http://admin-apps.webofknowledge.com/JCR/JCR>
- Mills E, Wu P, Gagnier J, et al. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemporary Clin Trials*. 2005;26:480-7.
- Hoof L, Korevaar DA, Molenaar N, et al. Endorsement of ICMJE's clinical trial registration policy: a survey among journal editors. *Neth J Med*. 2014;72:349-55.
- Millum J, Wendler D, Emanuel EJ. The 50th anniversary of the Declaration of Helsinki. Progress but many remaining challenges. *JAMA*. 2013;310:2143-4.
- Baskerville NB, Liddy C, Hogg W. Systematic review and meta-analysis of practice facilitation within primary care settings. *Ann Fam Med*. 2012;10:63-74.
- Bai Y, Gao J, Zou DW, Li ZS. Methodological reporting of randomized clinical trials in major gastroenterology and hepatology journals in 2006. *Hepatology*. 2009;49:2108-12.
- Castaño García A. Evolución de la calidad de los ensayos clínicos publicados en 4 revistas generales de medicina (1993-5 y 2004). Un estudio preliminar. [Diploma de Estudios Avanzados]. Pamplona: Universidad de Navarra; 2005.
- Moher D, Schulz KF, Altman D; CONSORT Group (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001;285:1987-91.
- WMA.net. Declaración de Helsinki de la AMM. Principios éticos para las investigaciones médicas en seres humanos. Asociación Médica Mundial; 2014. (Consultado el 3 de abril de 2014.) Disponible en: <http://www.wma.net/es/30publications/10policies/b3>
- Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *CMAJ*. 2009;180:E47-57.
- Garthner G, Hansen RA, Nissman D, et al. Criteria for distinguishing effectiveness from efficacy trials in systematic reviews. *AHRQ Publication No. 06-0046*; abril 2006. (Consultado el 12 de enero de 2014.) Disponible en: <http://archive.ahrq.gov/downloads/pub/evidence/pdf/efftrials/efftrials.pdf>
- Grant AM, Altman DG, Babiker AB, et al. Issues in data monitoring and interim analysis of trials. *Health Technol Assess*. 2005;9 (7).
- Goldman B, LeBlanc M, Crowley J. Interim futility analysis with intermediate endpoints. *J Clinical Trials*. 2008;5:14-22.
- AusVet Animal Health Services. Calculate confidence limits for a sample proportion. (Consultado el 16 de septiembre de 2014.) Disponible en: <http://epitools.ausvet.com.au/content.php?page=CIPProportion&SampleSize=141&Positive=1&Conf=0.95&Digits=1>
- Stats Calculator. Confidence interval calculator for means. McCallum Layton; 2014. (Consultado el 16 de septiembre de 2014.) Disponible en: <https://www.allto.co.uk/tools/statistic-calculators/confidence-interval-for-mean-calculator/>
- Licciardone JC, Minotti DE, Gatchel RJ, et al. Osteopathic manual treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. *Ann Fam Med*. 2013;11:122-129.
- Yank V, Rennie D. Reporting of informed consent and ethics committee approval in clinical trials. *JAMA*. 2002;287: 2835-8.
- GraphPad Software. QuickCalcs. Confidence interval of a SD. (Consultado el 6 de diciembre de 2015.) Disponible en: <http://graphpad.com/quickcalcs/CISD1>
- Sinha S, Sinha S, Ashby E, et al. Quality of reporting in randomised trials published in high-quality surgical journals. *J Am Coll Surg*. 2009;209:565-71.e1.
- Agha RA, Camm CF, Doganay E, et al. Randomised controlled trials in plastic surgery: a systematic review of reporting quality. *Eur J Plast Surg*. 2014;37:55-62.
- Lai R, Chu R, Fraumeni M, et al. Quality of randomized controlled trials reporting in the primary treatment of brain tumors. *J Clin Oncol*. 2006;24:1136-44.
- Rutterford C, Taljaard M, Dixon S, et al. Reporting and methodological quality of simple size calculations in cluster randomized trials could be improved: a review. *J Clin Epidemiol*. 2015;68:716-23.
- Lucena C, Souza EM, Voinea GC, et al. A quality assessment of randomized controlled trials reports in endodontics. *Int Endod J*. 2017;50: 237-50.
- Zhang J, Chen X, Zhu Q, et al. Methodological reporting quality of randomized controlled trials: a survey of seven core journals of orthopaedics from Mainland China over 5 years following the CONSORT statement. *Orthop Traumatol Surg Res*. 2016;102:933-8.
- Lu Y, Yao Q, Gu J, et al. Methodological reporting of randomized clinical trials in respiratory research in 2010. *Respir Care*. 2013;58:1546-51.

36. Zhai X, Wang Y, Mu Q, et al. Methodological reporting quality of randomized controlled trials in 3 leading diabetes journals from 2011 to 2013 following CONSORT statement: a system review. *Medicine (Balt)*. 2015;94:e1083.
37. Liu SS, Togioka BM, Hurley RW, et al. Methodological quality of randomized controlled trials of postoperative epidural analgesia: validation of the Epidural Analgesia Trial Checklist as a specific instrument to evaluate methodology. *Reg Anesth Pain Med*. 2010;35:549-55.
38. Liu LQ, Morris PJ, Pengel LH. Compliance to the CONSORT statement of randomized controlled trials in solid organ transplantation: a 3-year overview. *Transpl Int*. 2013;26:300-6.
39. Huang D, Jin X, Gao J, et al. Quality evaluation of randomized controlled trials reports of laparoscopy compared with open colorectal resection for colorectal cancer. *Expert Rev Anticancer Ther*. 2015;15:727-32.
40. Signori A, Baccino A, Sormani MP. The quality of reports of randomized trials in multiple sclerosis: a review. *Mult Scler*. 2012;18:776-81.
41. Diaz-Ordaz K, Froud R, Sheehan B, et al. A systematic review of cluster randomised trials in residential facilities for older people suggests how to improve quality. *BMC Med Res Methodol*. 2013;13:127.
42. Masson P, Duthie FA, Ruster LP, et al. Consistency and completeness of reported outcomes in randomized trials of primary immunosuppression in kidney transplantation. *Am J Transplant*. 2013;13:2892-901.