NORMATIVA DEL TREBALL DE FI DE GRAU / MÀSTER
FACULTAT DE MEDICINA I CIÈNCIES DE LA SALUT URV

El treball de fi de grau i el de fi de màster —en el cas de Medicina— (en endavant TFG/TFM) permeten als estudiants acreditar, mitjançant el desenvolupament d'un treball, que han adquirit els coneixements i les competències associats al títol. Suposen finalitzar els estudis i permeten a l'estudiant mostrar, tot integrant coneixements, les competències adquirides al llarg de tota la formació, així com realitzar una reflexió i una anàlisi de la seva evolució personal, acadèmica i professional.

Per a l'organització i avaluació de les proves del TFG/TFM, es nomena una comissió (punt 5) que designa, al seu torn, els tribunals d'avaluació encarregats d'avaluar el TFG/TFM, d'acord amb les directrius que s'indiquen en aquest document.

Per matricular el TFG/TFM, cal tenir superats un mínim de 168 crèdits, en el cas de graus de 240 crèdits, o un mínim de 288 crèdits, en el cas de graus de 360 crèdits.

1. GUIA D'ELABORACIÓ DEL TFG/TFM

El TFG/TFM consta de tres blocs tal i com consta a la memòria de verificació (Annex 1): una avaluació clínica objectiva estructurada (ACOE), una defensa del Portafolis electrònic i un treball de pràctica o cas clínic complex o treball de recerca tutelats (anomenats en aquesta normativa com a *treball pràctic*, per facilitar la comprensió del text).

TFG/TFM	ACOE	
	PORTAFOLIS ELECTRÒNIC	
	TREBALL PRÀCTIC	1. Recerca 2. Cas clínic 3. Revisió bibliogràfica 4. Aprenentatge - servei

1.1. TIPOLOGIA I MEMÒRIA DE TREBALL PRÀCTIC

El treball pràctic ha de permetre, als estudiants, acreditar l'adquisició de coneixements i competències en l'àrea de la investigació bàsica, translacional o clínica, mitjançant l'elaboració, presentació i defensa oral d'un treball d'iniciació a la investigació. L'estudiant ha d'elaborar aquest treball sota la supervisió i direcció d'un professor que imparteixi docència del grau/màster.

El treball pràctic es pot realitzar de forma individual o en grup (4 estudiants com a màxim), però en tots els casos es presenta de forma individual (davant tribunals diferents) i cada alumne defensa la part que el tribunal d'avaluació li requereixi.

Es pot fer d'acord amb les següents modalitats:

1.1.1 Treball de recerca original

a) Realitzat durant els darrers cursos de grau i defensat al final del grau.

b) Realitzat al llarg de la titulació dins el programa Alumne intern (havent fet el programa complet durant dos anys) i defensat al final de l'últim curs.
En aquesta modalitat, hi ha de constar l'autoria, almenys, de dues comunicacions de treballs tutelats, presentades en dues edicions diferents dins del programa Alumne intern, amb la corresponent publicació al llibre de resums (AIFMCS). El nombre d'autors de cada comunicació no pot ser superior a quatre.

c) Realitzat al si d'un grup de recerca de la Facultat de Medicina i Ciències de la Salut, el treball ha de donar lloc a una publicació científica i l'alumne n'ha de ser coautor.

La presentació de la memòria és obligatòria per a qualsevol de les modalitats del treball de recerca exposades a l'apartat anterior. La memòria ha de tenir una extensió de 10 a 30 pàgines i ha d'incloure els apartats: títol, autor (o autors en el cas de les comunicacions col·lectives), nom del tutor - director del treball, resum (en català, castellà i anglès), introducció i justificació del treball, objectius, material i mètodes, resultats, discussió i bibliografia. Els apartats "resultats" i "discussió" es poden redactar conjuntament, si es considera oportú. S'hi pot adjuntar qualsevol documentació que sigui rellevant per comprendre i clarificar el treball desenvolupat.

1.1.2 Cas clínic complex

Descripció d'un o més casos clínics d'excepcional observació que suposin una aportació important al coneixement de la fisiopatologia, diagnòstic, tractament, etiologia o altres aspectes del procés.

En la modalitat 1.1.1 i 1.1.2, la memòria ha de seguir les pautes de les publicacions científiques de l'àrea de les ciències de la salut, com ara les normes publicades a la revista *Medicina Clínica* (annex 2) o els requisits d'uniformitat per a manuscrits establerts per a les revistes biomèdiques.
http://www.icmje.org/manuscript_1prepare.html

1.1.3 Treball de revisió bibliogràfica

S'han d'utilitzar articles de revistes dels darrers 5 anys i indexades, es deixa a criteri del tutor el número d'articles a revisar.
La memòria ha de seguir les pautes "Revisió sistemàtica d'estudis" (annex 3)
http://www.prisma-statement.org/statement.htm

1.1.4 Programa d'alumne mentor lligat a aprenentatge - servei de la URV

Planificació i aplicació de programes d'educació en salut dins dels programes d'agents promotors de salut i d'alumne mentor.

Aquests programes, que l'estudiant durà a terme dins de les assignatures del darrer any, s'indiquen a continuació segons l'ensenyament:

GRAU DE MEDICINA	
14204155	Àmbits d'actuació: atenció sociosanitària, urgències, intensius, investigació, medicina legal, gestió i altres
GRAU DE FISIOTERÀPIA	
14214130	Pràctiques tutelades opcionals: investigació, esport, hidroteràpia, externes
GRAU DE NUTRICIÓ HUMANA I DIETÈTICA	
14224401	Pràcticum opcional
14224402	Pràctiques comunitàries

L'alumne ha de presentar el projecte realitzat en dos centres com a mínim i els resultats obtinguts. El projecte es pot treballar com un estudi d'intervenció; per tant, es pot seguir la guia CONSORT (annex 4).
http://www.consort-statement.org/consort-statement/overview0/#checklist

L'elecció del tema del treball i la modalitat s'han d'acordar entre el tutor del treball pràctic i l'alumne (o els alumnes). A més, el treball s'ha d'adscriure a una de les línies de recerca que es desenvolupin en algun dels departaments amb docència en el grau.

El centre ha de publicar, d'acord amb els departaments implicats, una llista de professors i línies de recerca (annex 5).

1.2 TUTORITZACIÓ DEL TREBALL PRÀCTIC

1.2.1 TUTORS DE TREBALL PRÀCTIC

Podran exercir de tutors tots els professors doctors (grau/màster de Medicina) i no doctors[1] (grau de Fisioteràpia i Nutrició Humana i Dietètica) amb responsabilitats docents a la Facultat de Medicina i Ciències de la Salut. Es pot designar un únic cotutor per treball, que podrà ser qualsevol altre professor no doctor del centre, els tutors externs de Pràctiques Clíniques, altres metges dels hospitals universitaris adscrits a la nostra universitat (HUSJ, HUJXXIII i HUIPM) i personal investigador adscrit a les unitats d'investigació d'ambdós hospitals, instituts de recerca (IISPV, CTNS) i departaments de la facultat.

1.2.2 FUNCIONS DELS TUTORS

a) El tutor proposa a l'estudiant qüestions de rellevància que puguin ser objecte del treball pràctic i en què es puguin practicar les habilitats i competències de nivell de grau/màster.

b) El tutor orienta l'estudiant, li facilita l'accés als mitjans d'investigació que es considerin oportuns i realitza un seguiment continuat del progrés del treball durant la seva elaboració.

[1] S'aprova que els tutors de treball del grau de fisioteràpia i nutrició puguin ser doctors i no doctors de manera transitòria i en el curs 2017-2018 es revisarà.

c) Un cop finalitzat el treball, el tutor emet un informe que estableix si reuneix els requisits exigibles per ser presentat i defensat. L'informe del tutor s'ha d'afegir a la memòria que l'estudiant ha de lliurar abans defensar el treball. Només es defensaran públicament els TFG/TFM que tinguin un informe positiu del tutor.

1.2.3 PROPOSTA DE TREBALL PRÀCTIC

La proposta de treball pràctic la presentarà l'alumne a la Comissió, d'acord amb el professor tutor, en el termini que indiqui aquesta comissió.

Constarà de títol, tipus de treball pràctic, idioma de presentació i defensa (català, espanyol, anglès).

Anirà signada pel tutor i per l'estudiant.

S'accepta el tutor i el tema del treball, si la Comissió no indica el contrari en el termini de 10 dies hàbils.

En el cas que un estudiant no hagi sol·licitat tutor, La Comissió li assignarà un tutor que estigui disponible d'acord amb la càrrega docent del professorat.

I en el cas que hi hagi coincidència de sol·licituds de diversos estudiants, l'assignació es durà a terme mitjançant nota d'expedient.

1.3 COORDINACIÓ DE TREBALLS PRÀCTICS

Les coordinadores o la coordinadora elaboren la guia docent de l'assignatura i el pla de treball al Moodle. Hi indicaran la programació de totes les activitats: seminaris d'assessorament, tutories, lliurament d'activitats, avaluació, etc.
Aquesta tasca la realitzarà el responsable de la titulació o un professor doctor de la titulació en qui delegui.
Per aquesta feina, el coordinador tindrà 1 crèdit, que es restarà del còmput total assignat, aquell curs, als tutors pels treballs.

2. DEFENSA DEL TFG/TFM

La defensa oral i pública, davant els tribunals d'avaluació designats per la Comissió, és obligatòria per a les ACOE, el portafolis electrònic i totes les modalitats de treball pràctic descrites.

Per les ACOE, la defensa pública coincideix amb el desenvolupament de l'activitat mateixa i no caldrà realitzar-la de nou; però sí que s'hauran de lliurar els informes dels tutors de l'ACOE.

S'haurà de defensar el portafolis electrònic davant el tribunal d'avaluació, al qual, a més, caldrà lliurar la valoració final, feta pel tutor, del portafolis electrònic.

El treball pràctic haurà de defensar-se davant el tribunal d'avaluació, amb les següents excepcions:

a) A la modalitat de recerca original a l'AIFMCS (1.1.1.b), la defensa pública coincideix amb l'exposició del treball en l'activitat AIFMCS, i no caldrà realitzar-la de nou; però sí que s'haurà de lliurar la memòria completa, amb la referència als llibres de publicacions i els informes dels tutors.

b) En el cas de treballs de recerca original de modalitat "1.1.1.c" (coautoria de publicació), a la memòria s'hi han d'adjuntar una còpia del treball de recerca publicat (o la carta d'acceptació de la publicació) i un document, signat per l'investigador principal del treball, en què es facin constar les tasques realitzades per l'alumne en el treball en qüestió.

En qualsevol cas, els membres del tribunal d'avaluació han de disposar de la documentació necessària (memòria i altres documents, segons els casos) com a mínim 10 dies abans de l'acte de defensa oral.

La defensa del TFG/TFM es pot fer en llengua catalana, castellana o anglesa.

La durada màxima de l'exposició serà de 10 minuts i hi haurà 5 minuts de discussió amb els membres del tribunal d'avaluació.

3. AVALUACIÓ FINAL DEL TFG/TFM

La superació del TFG/TFM queda supeditada a l'assoliment de les competències específiques vinculades a aquesta assignatura. Per tant, és prerequisit, per superar l'assignatura, haver aprovat les matèries bàsiques i obligatòries de l'itinerari curricular.

El TFG/TFM s'avaluarà en una convocatòria única al mes de juny. L'alumnat podrà sol·licitar avançar aquesta convocatòria al gener, o endarrerir-la fins al setembre, presentant la corresponent instància a la Secretaria. (Consultar: Intranet>Gestió acadèmica>tràmits administratius).

El pes, a la nota final del TFG/TFM de cada bloc, és específic de cada titulació, d'acord amb la memòria verificada i ratificada per la Comissió del TFG/TFM.

3.1 AVALUACIÓ DE L'ACOE

L'ACOE es realitza al final de l'últim any de la titulació. La seva qualificació coincideix amb el percentatge descrit en la memòria verificada i especificat a la guia docent de l'assignatura.

Un tribunal d'avaluació de l'ACOE dóna a l'estudiant una qualificació. La qualificació de l'ACOE se suma a la nota final del TFG/TFM, sempre que s'aprovi.

3.2 AVALUACIÓ DEL PORTAFOLIS ELECTRÒNIC

El portafolis electrònic s'elabora des de 2n de carrera fins a l'últim any. El professor tutor del portafolis electrònic pot avaluar de 0 a 10 la qualitat de cada una de les evidències aportades per l'estudiant.

En finalitzar el portafolis electrònic, abans de la defensa del TFG/TFM, el professor tutor del portafolis electrònic valora, de manera global, tant les evidències de les assignatures com les dels alumnes, amb una nota de 5 a 10, sempre que l'estudiant hagi superat el 50% d'assoliment del portafolis electrònic i amb una qualificació superior a 5.

La qualificació que el tutor hagi posat al portafolis electrònic i la qualificació que li posi el tribunal d'avaluació (davant el qual l'estudiant defensa el seu portafolis electrònic) se sumen a la nota final del TFG/TFM.

3.3 AVALUACIÓ DEL TREBALL PRÀCTIC

El treball pràctic tutelat es recull en una memòria impresa, i es defensa oralment davant el tribunal d'avaluació.

El treball pràctic es qualifica com una altra assignatura del grau i aquesta qualificació també se sumarà a la nota final del TFG/TFM.

4. TRIBUNAL D'AVALUACIÓ

Els tribunals d'avaluació per a les ACOE, el portafolis electrònic i el treball pràctic, incloent-hi les comunicacions a l'AIFMCS, estan formades per tres membres (president, vocal i secretari) que han de pertànyer a la plantilla de professors de la Facultat de Medicina i Ciències de la Salut.

Si es considera oportú, podran, seguint els mateixos criteris, nomenar subcomissions de qualificació. Al final de cada sessió de defensa, se n'aixeca acta, on ha de constar la qualificació individual de cada alumne, d'acord amb els criteris esmentats anteriorment.

5. COMISSIÓ DE TFG/TFM

Les tasques de la revisió de la normativa i la designació dels membres dels tribunals d'avaluació dels treballs s'assumeixen a la Comissió Acadèmica delegada de la Junta de Centre de la Facultat de Medicina i Ciències de la Salut.

6. PROPIETAT INTEL·LECTUAL I DIPÒSIT DEL TREBALL

Els drets de propietat intel·lectual dels TFG/TFM es regulen en els termes i condicions previstos en la legislació vigent. En qualsevol cas, en qualsevol ús que es pugui fer dels TFG/TFM, sempre s'hi haurà de fer constar: l'autoria el tutor, la naturalesa del treball i la vinculació amb la URV.

Els TFG es dipositaran al repositori institucional de la URV. Els que obtinguin una qualificació igual o superior a 8 seran visibles i de lliure consulta per a usos docents, de recerca o d'estudi personal.

La Universitat Rovira i Virgili establirà els requisits formals dels TFG/TFM, a fi que després es puguin emmagatzemar al repositori institucional.

7. CONFIDENCIALITAT

En circumstàncies excepcionals determinades per la Comissió Acadèmica delegada de Centre, com pot ser, entre d'altres, la participació d'empreses en el programa, l'existència de convenis de confidencialitat amb empreses o la possibilitat de generació de patents que recaiguin sobre el contingut del TFG/M, es prendran les mesures oportunes per assegurar la no-publicitat d'aquests aspectes, sense que sigui en detriment de la presentació pública de l'aportació al coneixement del treball. Aquest fet s'ha d'informar a la Comissió Acadèmica delegada de la Junta de Centre.

Disposició addicional

Els alumnes que van presentar un treball de recerca a l'AIFMCS del curs 2010-11 i que van obtenir una qualificació major de 5 poden utilitzar aquestes presentacions per optar al model descrit en el punt 1.1.1.b.

ANNEX 1. ESPECIFICACIONS A LA MEMÒRIA VERIFICADA

1.1 GRAU DE FISIOTERÀPIA

COMPETÈNCIES I RESULTATS D'APRENENTATGE QUE L'ESTUDIANT ADQUIREIX AMB LA MATÈRIA:
3.2 Treball de fi de grau (6 ECTS)

Competències:

- **Competències específiques**: A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23, A24, A25, A26, A27, A28, A29, A30, A31, A32, A33, A34, A38, A39, A40, A41, A42
- **Competències transversals**: B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11
- **Competències nuclears**: C1, C2, C3, C4, C5, C6

Resultats d'aprenentatge:
Presentació i defensa davant del tribunal universitari d'un projecte de fi de grau que consisteixi en un exercici d'integració dels continguts formatius rebuts i les competències adquirides. (Ordre CIN/2135/2008, 3 de juliol)

Requisits:
Per poder realitzar les assignatures Fisioteràpia Hospitalària, Fisioteràpia en Atenció Primària, Fisioteràpia Geriàtrica i Fisioteràpia Domiciliària, l'alumne haurà d'haver superat les Pràctiques Clíniques I i les Pràctiques Clíniques II.
Per poder avaluar el treball de fi de grau, l'alumne haurà d'haver completat la llista competencial del portafolis electrònic. La universitat establirà els requisits que els estudiants hauran de posseir abans de poder-se matricular a l'assignatura Treball de Fi de Grau. Aquests requisits s'inclouran oportunament a la normativa universitària corresponent.

Treball de fi de grau. El treball de fi de grau ha de ser el resultat de la revisió i l'execució dels següents apartats:

Llista competencial: 20% de la nota final.
Elaboració i exposició de diferents casos clínics per àmbits d'actuació: 40% de la nota final.
ECOE: avaluació de la competència objectiva i estructurada, mitjançant una prova multiestació, amb la participació de pacients estandarditzats, maniquins, pictorials i casos problema: **40% de la nota final.**

En la següent taula s'hi resumeixen les metodologies d'avaluació, en quina activitat formativa s'utilitzarà l'avaluació i el pes que tindrà en la nota final. També s'hi especifiquen les metodologies d'avaluació dels tres apartats del treball de fi de grau i el seu pes en la nota final.

Metodologia	% Pràctiques clíniques	% Treball de fi de grau
Examen pràctic amb llista competencial		20
Laboratori d'habilitats		20
Atenció personalitzada	10	
Cas clínic	30	20
Aprenentatge basat en problemes (ABP)	15	20
Casos pràctics per avaluar	20	
Entrevistes amb pacients estandarditzats al laboratori d'habilitats		20
Observació estructurada de la pràctica clínica -	25	

Mini-CEX		

Matèria 3.2 Treball Fi de Grau
Avaluació de la competència objectiva estructurada (ECOE). Portafolis electrònic: registre d'evidències d'activitats formatives, experiències, exposicions, en què hagi participat l'alumne. Exposició d'un cas clínic complet o d'un treball d'investigació.

1.2 GRAU DE NUTRICIÓ HUMANA I DIETÈTICA

COMPETÈNCIES I RESULTATS D'APRENENTATGE QUE L'ESTUDIANT ADQUIREIX AMB LA MATÈRIA:
6.3 TREBALL DE FI DE GRAU (9 ECTS)

Competències:

- **Competències específiques**: A1,A2,A3,A4,A5,A6,A7,A8,A9,A10,A11,A12,A13,A14,A15,A16,A17, A18,A19,A20,A21,A22,A23,A24,A25,A26,A28,A29,A30,A31,A32,A33,A34,A35 A36,A37,A38,A39,A40,A41,A42,A43
- **Competències transversals**: B1,B2,B3,B4,B5,B6,B7,B8,B9,B10
- **Competències nuclears**: C1,C2,C3,C4,C5,C6

Resultats d'aprenentatge:

- Treball de fi grau: matèria transversal, el treball de la qual s'ha de realitzar associat a diverses matèries. (Ordre CIN/730/2009, 18 de març del 2009)

REQUISITS: Per poder avaluar l'assignatura Treball de Fi de Grau, l'alumne haurà d'haver superat la resta d'assignatures de la carrera.

Treball de fi de grau. El treball de fi de grau ha de ser el resultat de la revisió i l'execució dels següents apartats:

Llista competencial: 20% de la nota final.
Elaboració i exposició de diferents casos clínics per àmbits d'actuació: 40% de la nota final.
ECOE: avaluació de la competència objectiva i estructurada, mitjançant una prova multiestació, amb la participació de pacients estandarditzats, maniquins, pictorials i casos problema: **40% de la nota final.**

En la següent taula s'hi resumeixen les metodologies d'avaluació, en quina activitat formativa s'utilitzarà l'avaluació i el pes que tindrà en la nota final. També s'hi especifiquen les metodologies d'avaluació dels tres apartats del treball de fi de grau i el seu pes en la nota final.

Metodologia	% Pràctiques externes Pràctiques clíniques	% Treball de fi de grau
Examen pràctic amb llista competencial		20
Laboratori d'habilitats		20
Atenció personalitzada	10	
Cas clínic	15	20
Aprenentatge basat en problemes (ABP)	15	20
Casos pràctics per avaluar	10	
Entrevistes amb pacients estandarditzats al laboratori d'habilitats		20
Observació estructurada de la pràctica clínica - Mini-CEX	25	
Presentació de projectes	25	

Matèria 6.3 Treball de fi de grau

Avaluació de la competència objectiva estructurada (ECOE). Portafolis electrònic: registre d'evidències d'activitats formatives, experiències, exposicions, en què hagi participat l'alumne. Exposició d'un cas clínic complet o d'un treball d'investigació.

1.3. GRAU/MÀSTER DE MEDICINA

COMPETÈNCIES[2] I RESULTATS D'APRENENTATGE QUE L'ESTUDIANT ADQUIREIX AMB LA MATÈRIA: **5.2. Treball de fi de grau (6 ECTS)**
Competències: ☑ **Competències específiques**: Totes ☑ **Competències transversals**: Totes ☑ **Competències nuclears**: C1, C2, C3, C4, C5, C6 **Resultats d'aprenentatge:** ☑ Treball de fi grau: matèria transversal, el treball de la qual s'ha de realitzar associat a diverses matèries. (Ordre ECI/332/2008, 13 de febrer)
REQUISITS: Per poder avaluar les assignatures d'aquest mòdul, l'alumne ha de superar les assignatures Clínica I, Clínica II i Habilitats Diagnòstiques. Per poder realitzar les assignatures Clínica Mèdica i Clínica Quirúrgica, l'alumne ha de superar 50 ECTS de la matèria Formació Medicoquirúrgica. Per poder realitzar l'assignatura Clínica Obstètrica i Ginecològica, l'alumne ha de superar la matèria Obstetrícia i Ginecologia. Per poder realitzar l'assignatura Clínica Pediàtrica, l'alumne ha de superar la matèria Pediatria. Per poder realitzar l'assignatura Clínica Psiquiàtrica, l'alumne ha de superar la matèria Psicologia i Psiquiatria. Per poder avaluar el treball de fi de grau, l'alumne ha de superar la llista competencial del portafolis electrònic. La universitat establirà els requisits que els estudiants han de posseir abans de matricular-se a l'assignatura Treball de Fi de Grau. Aquests requisits s'inclouran oportunament a la normativa universitària corresponent.

Treball de fi de grau.
Llista competencial: 50%.
ECOE: 50%. Avaluació de la competència objectiva i estructurada, mitjançant una prova multiestació, amb la participació de pacients estandarditzats, maniquins, pictorials i casos problema.

5.2. Treball de fi de grau (6 ECTS)
Avaluació de la competència objectiva estructurada (ECOE). Dossier electrònic: registre d'evidències d'activitats formatives, experiències, exposicions, en què hagi participat l'alumne. Exposició d'un cas clínic complet o d'un treball d'investigació.

ANNEX 2 . NORMES PUBLICADES A LA REVISTA MEDICINA CLÍNICA

Información para los autores

Para una información más amplia consulten. Requisitos de uniformidad para manuscritos presentados para publicación en revistas biomédicas (http://www.icmje.org).

MEDICINA CLÍNICA considerará preferentemente para su publicación trabajos relacionados con la Medicina Interna y sus subespecialidades. Fundamentalmente, la Revista consta de las siguientes secciones:

Originales. Trabajos de investigación sobre etiología, fisiopatología, anatomía patológica, epidemiología, clínica, diagnóstico, pronóstico y tratamiento. Los diseños recomendados son de tipo analítico en forma de encuestas transversales, estudios de casos y controles, estudios de cohortes y ensayos controlados. La extensión recomendada del texto es de 12 páginas de 30 líneas a doble espacio (Times New Roman, punto 12; 4200 palabras, 25560 caracteres con espacios) y se admitirán hasta seis figuras y seis tablas. Es aconsejable que el número de firmantes no sea superior a seis. Se incluirán 30 referencias bibliográficas como máximo. Las unidades de medida en cualquier sección se expresarán en sistema internacional (SI). MEDICINA CLÍNICA podráconsiderar y publicar artículos originales en lengua inglesa. Excepcionalmente, los Artículos Especiales que tengan la estructura de un artículo original podrán también ser publicados en lengua inglesa.

Originales breves. En esta sección se considerarán los trabajos de investigación que por sus características especiales (series con número reducido de observaciones, trabajos de investigación con objetivo y resultados muy concretos, estudios epidemiológicos descriptivos, entre otros) pueden ser publicados en forma más abreviada y rápida. Estos trabajos deberán tener una extensión máxima de 120 líneas de texto, 1300 palabras, 8300 caracteres con espacios, hasta 10 referencias bibliográficas y no más de dos ilustraciones. El número máximo de firmantes será de seis. Cada trabajo deberá estructurarse como un artículo original. Los Originales y los Originales breves deberán seguir la guía de publicación (CONSORT, STROBE, PRISMA, STARD y otras) que aplique al tipo de estudio concreto. Pueden encontrarse en el sitio web de la red EQUATOR (www.español.equator-network.org).

Notas clínicas. Descripción de uno o más casos clínicos de excepcional observación que supongan una aportación importante al conocimiento de la fisiopatología o de otros aspectos del proceso.

La extensión máxima del texto será de 5 páginas de 30 líneas y a doble espacio (Times New Roman, punto 12; 1750 palabras, 10650 caracteres con espacios) se admitirán hasta dos figuras y dos tablas. Es aconsejable que el número de firmantes no sea superior a seis y que no haya más de 20 referencias bibliográficas.

Cartas al Editor. Tienen preferencia en esta Sección aquellas cartas que hagan referencia a trabajos publicados el último mes y aquellas que aporten opiniones, observaciones o experiencias que por sus características puedan ser resumidas en un breve texto. La sección se divide en "**Cartas Científicas**", es decir, trabajos que contienen nuevos estudios que pueden exponerse en forma abreviada y **"Cartas al Editor"** propiamente dichas.

La extensión máxima será de 60 líneas y se admitirán una figura o una tabla y diez referencias bibliográficas como máximo. El número de firmantes no debe exceder de cuatro.

Otras secciones. La Revista incluye otras secciones (Editoriales, Diagnóstico y Tratamiento, Revisiones y Artículos Especiales) cuyos artículos encarga el Comité de Redacción. Los autores que espontáneamente deseen colaborar en alguna de estas secciones deberán consultar previamente al Editor asociado de la Revista. El número máximo de autores será de dos para los Editoriales, de tres para las secciones de Revisiones y Diagnóstico y Tratamiento, y de cuatro para los Artículos Especiales. Se recomienda que los **Editoriales** tengan una extensión no superior a las 6 páginas a doble espacio (Times New Roman, punto 12; 2100 palabras, 12780 caracteres con espacios), sin tablas ni figuras, y un máximo de 30 citas bibliográ ficas. Los artículos para las secciones de **Revisiones, Diagnóstico** y **Tratamiento** y **Artículos Especiales** se presentarán con una extensión de 12 páginas a doble espacio (Times New Roman, punto 12; 4200 palabras, 25560 caracteres con espacios) y un máximo de 50-60 citas. Se admitirán como límite 4 figuras y 5 tablas que deberán contribuir de manera evidente a la mejor comprensión del texto. Las Revisiones se acompañarán de un Resumen en castellano e inglés y tendrán un último apartado de Conclusiones de aproximadamente un folio de extensión. Los artículos de la sección de Diagnóstico y Tratamiento y los Artículos Especiales no se acompañan de Resumen. MEDICINA podrá encargar artículos que aborden un tema que se considere de interés general para los lectores de la revista, y cuya publicación semanal o quincenal constituyan una Serie. Los trabajos de una **Serie** deberán ajustarse a las características editoriales de los Artículos Especiales. También se valorarán para su publicación como **Reportajes** las experiencias médicas de contenido sanitario o social. MEDICINA CLÍNICA publicará, asimismo, **Conferencias de Consenso** sobre diagnóstico o tratamiento de enfermedades siempre que estén promovidas por organismos oficiales de Sanidad o por sociedades científicas nacionales o internacionales. El número máximo de autores será de seis. Si fuera superior, se nombrará un Comité de Redacción hasta un máximo de seis firmantes y el resto figurarán como colaboradores. Debido a su extensión, estos artículos se publicarán en su totalidad en formato electrónico. En papel impreso se publicará el título, los autores, filiaciones y un resumen no estructurado en castellano e inglés de un máximo de 250 palabras. Medicina Clínica publicará **Conferencias Clinicopatológicas** y **Conferencias Clinicopatológicas-MIR,** con la siguiente estructura: presentación del caso, diagnóstico diferencial, diagnóstico clínico, discusión anatomopatológica y diagnóstico final. Por último, en la sección de **Imagen de la Semana** se publicarán imágenes de cualquier tipo (fotográfica, endoscópica, radiológica, microbiológica, anatomopatológica) que sean muy demostrativas y contengan por sí mismas un mensaje didáctico. Tendrán preferencia aquellas que combinen diversos aspectos de los mencionados. Deben acompa arse de un texto de menos de 10 líneas, el diagnóstico y un máximo de cuatro firmantes, con su centro de trabajo. Cuando haya posibilidad de que el paciente sea identificado en la imagen es imprescindible la autorización del mismo por escrito. Siempre que sea posible, la fotografía debe incluir recursos gráficos (flechas, asteriscos). Las imágenes se publicarán en papel o en formato electrónico (e-only) a criterio del Comité Editorial de la Revista.

Fast track. El procedimiento de revisión y publicación rápida está orientado a aquellos artículos de una importancia clínica excepcional o bien con un impacto directo sobre la práctica médica. Los trabajos seleccionados se publicarán al cabo de seis semanas de recibir el manuscrito en las versiones impresa y electrónica de MEDICINA CLÍNICA y se editarán con el símbolo de Pegasus en el sumario y al principio del artículo.

Los Fast track deben cumplir todos los requisitos de las "Normas para autores". Puede tratarse de artículos originales o de opinión, como pueden ser los editoriales. Debe hacerse constar claramente en la carta de presentación la solicitud de **Fast track** (revisión y publicación rápida).

Presentación y estructura de los trabajos

Todos los originales aceptados quedan como propiedad permanente de MEDICINA CLÍNICA, y no podrán ser reproducidos en parte o totalmente sin su permiso. El autor cede, en el supuesto de publicación de su trabajo, de forma exclusiva a ELSEVIER ESPAÑA, S.L. los derechos de reproducción, distribución, traducción y comunicación pública (por cualquier medio o soporte sonoro, audiovisual o electrónico) de su trabajo.

Los autores deben describir cualquier relación financiera que tengan y que pudiera dar lugar a un conflicto de intereses en relación con el artículo publicado.

Cuando se presenten estudios realizados en seres humanos debe indicarse si los métodos seguidos han cumplido las normas éticas del comité de investigación o de los ensayos clínicos correspondientes (del centro o regionales) y de la Declaración de Helsinki de 1975 (actualizaciones disponibles en: http://www.wma.net/s/policy/b3.htm).

Del mismo modo, los autores deberán declarar que se han seguido los protocolos establecidos por sus respectivos centros sanitarios para acceder a los datos de las historias clínicas a los fines de poder realizar este tipo de publicación con finalidad de investigación / divulgación para la comunidad científica.

No se aceptarán trabajos publicados o presentados al mismo tiempo en otra Revista.

Los trabajos se presentarán a doble espacio (30 líneas). Las hojas irán numeradas correlativamente en la parte inferior central. Cada parte del manuscrito empezará una página en el siguiente orden:

1. En la primera página del artículo se indicarán, en el orden que aquíse cita, los siguientes datos: título del artículo (en castellano y en inglés), nombre completo y uno o los dos apellidos de los autores, nombre completo y dirección del centro de trabajo, dirección postal, telefax, dirección de correo electrónico y otras especificaciones cuando se considere necesario.

En caso de autoría corporativa figurará un mínimo de un firmante y un máximo de seis. en nombre del grupo corporativo. El resto de participantes figurarán en un addendum.

2. *Texto.* Se recomienda la redacción del texto en impersonal. Conviene dividir claramente los trabajos en apartados, y es de desear que el esquema general sea el siguiente:

2.1. *Originales:* Introducción, Pacientes (Sujetos) o Material y Método, Resultados y Discusión. Resumen, en castellano e inglés (Abstract).

2.2. *Notas Clínicas:* Introducción, Observación clínica o Métodos, Resultados y Discusión. Resumen y Abstract.

a) *Introducción.* Será breve y debe proporcionar sólo la explicación necesaria para que el lector pueda comprender el texto que sigue a continuación. No

debe contener tablas ni figuras. Debe incluir un último párrafo en el que se exponga de forma clara el/los objetivo/s del trabajo. Siempre que se pretenda publicar una observación muy infrecuente debe precisarse en el texto el método de pesquisa bibliográfica, las palabras clave empleadas, los años de cobertura y la fecha de actualización.

b) *Pacientes (Sujetos) o Material y Métodos.* En este apartado se indican el centro donde se ha realizado el experimento o la investigación, el período de duración, las características de la serie estudiada, el criterio de selección empleado y las técnicas utilizadas, proporcionando los detalles suficientes para que una experiencia determinada pueda repetirse sobre la base de esta información. Se han de describir con detalle los métodos estadísticos.

c) *Resultados.* Relatan, no interpretan, las observaciones efectuadas con el método empleado. Estos datos se expondrán en el texto con el complemento de las tablas y figuras.

d) *Discusión.* Los autores tienen que exponer sus propias opiniones sobre el tema. Destacan aquí: 1) el significado y la aplicación práctica de los resultados; 2) las consideraciones sobre una posible inconsistencia de la metodología y las razones por las cuales pueden ser válidos los resultados; 3) la relación con publicaciones similares y comparación entre las áreas de acuerdo y desacuerdo, y 4) las indicaciones y directrices para futuras investigaciones. No deben efectuarse conclusiones. Por otra parte, debe evitarse que la discusión se convierta en una revisión del tema y que se repitan los conceptos que hayan aparecido en la introducción. Tampoco deben repetirse los resultados del trabajo.

e) *Agradecimiento.* Cuando se considere necesario se citaráa las personas, centros o entidades que hayan colaborado o apoyado la realización del trabajo. Si existen implicaciones comerciales también deben figurar en este apartado.

f) *Resumen/Abstract.* Debe adjuntarse en espaþal y en inglés bajo estos epígrafes. El resumen/abstract de las Revisiones debe tener una extensión de 150 palabras, aproximadamente, sin estructurarse. La extensión del resumen/abstract para los Originales debe ser como máximo de 250 palabras y para los Originales breves y las Notas Clínicas de 180 palabras aproximadamente. Su contenido debe estar estructurado y se divide en cuatro apartados: Fundamento y objetivo, Pacientes o Material y método, Resultados y Conclusiones. En cada uno de ellos se han de describir, respectivamente, el problema motivo de la investigación, la manera de llevarla a cabo, los resultados más destacados y las conclusiones que derivan de los resultados. Al final del resumen deben figurar las palabras clave de acuerdo con las incluidas en el Medical Subject Headings (MeSH) de Index Medicus/Medline, en inglés disponible en: http://www.nlm.nih. gov/mesh/meshhome.html y traducirlas al castellano.

3. Referencias bibliográficas. Se presentarán según el orden de aparición en el texto con la correspondiente numeración correlativa. En el artículo constará siempre la numeración de la cita en número volado, según los "Requisitos de uniformidad para manuscritos presentados para publicación en revistas biomédicas" elaborados por el Comité Internacional de Editores de Revistas Médicas (Med Clin (Barc). 1997;109:756-63). Actualizaciones disponibles en:http://www.icmje.org/ Los nombres de las revistas deben abreviarse de acuerdo con el estilo usado en el Index Medicus/Medline: "List of Journals Indexed" que se incluye todos los años en el número de enero del *Index Medicus*, también disponible en: http://www.ncbi.nlm.nih.gov/entrez/jrbrowser.cgi

Se evitará, en lo posible, la inclusión como referencias bibliográficas de libros de texto y actas de reuniones.

Es aconsejable evitar el uso de frases imprecisas como referencias bibliográficas y no pueden emplearse como tales "observaciones no publicadas" ni "comunicación personal", pero sípueden citarse entre paréntesis dentro del texto.

Las referencias bibliográficas deben comprobarse por comparación con los documentos originales, indicando siempre las páginas inicial y final de la cita. A continuación se dan unos ejemplos de formatos de citas bibliográficas:

Revista

1) Artículo ordinario.

Relacionar todos los autores si son seis o menos; si son siete o más, relacionar los seis primeros y añadir la expresión "et al" después de una coma.

Bonet J, Vicente A. Rigidez arterial, lesión subclínica de órganos y riesgo cardiovascular. Med Clin (Barc). 2009;133:137-8.

Forner A, Ayuso C, Isabel Real M, Sastre J, Robles R, Sangro B, et al. Diagnosis and treatment of hepatocellular carcinoma. Med Clin (Barc). 2009;132:272-87.

2) Autor corporativo

Expert Panel on Detection EaToHBCiA. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). JAMA. 2001;285:2486-97.

3) No se indica el nombre del autor

Las últimas transferencias sanitarias del INSALUD: una valoración de urgencia [editorial]. Medifam. 2002;12:11-3.

4) Suplemento de un volumen

Chouat G, Menu E, Delange G, Mareau JF, Khrishnan L, Hui L, et al. Immuno-endocrine interactions in early pregnancy. Human Reprod. 1995;10(Suppl. 2): 55-9.

5) Suplemento de un número

Boat TF. The future of pediatric research. J Pediatr. 2007;151(5 Suppl):21-7.

6) Número sin volumen

Kanis JA, McCloskey EV, Johansson H, Oden A, Melton III LJ, Khaltaev N. A reference standard for the description of osteoporosis. Bone. 2008;(3):467-75.

7) Indicación del tipo de artículo

Verdaguer JM. Alteraciones precoces en la producción vocal de los pacientes intervenidos de cirugía tiroidea [tesis doctoral]. Madrid: Universidad Autónoma de Madrid; 2007.

8) Trabajo en prensa

Bujanda L, Gil I, Sarasqueta C, Hijona E, Beraza M, Cosme A, et al. Características clinicopatológicas y supervivencia del cáncer de esófago. Resultados de 200 pacientes consecutivos. Med Clin (Barc). 2009. doi:10.1016/j.medcli.2009. 04. 049

Libros y otras monografías

9) Autores personales

Ware JE, Kosinski M, Dewey JE. How to score version 2 of the SF-36 Health Survey (standard & acute forms. Lincoln RI: Quality Metric Incorporated; 2000.

10) Directores o compiladores como autores

Charlton JE, editor. Core curriculum for professional education in pain. Seattle: IASP Press; 2005.

11) Capítulo de un libro

Greenland S, Lash TL. Bias analysis. En: Rothman KJ, Greenland S, Lash TL, editores. Modern Epidemiology, 3ª ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2008. p. 359.

12) Actas de reuniones

Aguillo IF, Granadino B, Ortega JL. Diseño, métodos y problemática documental en la construcción de un ranking web de hospitales del mundo [ponencia]. Actas de las X Jornadas Españolas de Documentación; 2007, mayo 9-11; Santiago de Compostela. Madrid: Federación Española de Sociedades de Archivística, Biblioteconomía, Documentación y Museística (FESABID); 2007.

Material electrónico

13) Artículo de revista en formato electrónico

Martínez A. Indicadores cibermétricos: nuevas propuestas para medir la información en el entorno digital. Acimed [revista electrónica]. 2006;14(4) [consultado 27 Feb 2008]. Disponible en: http://scielo.sld.cu/scielo.php?pid=S1024-94352006000400003&script= sci_arttext&tlng=es

14) Monografías en formato electrónico

Farreras/Rozman. Medicina Interna [edición en CDROM], 13ª ed. Barcelona: Ediciones Doyma; 1996.

4. Las imágenes se seleccionarán cuidadosamente, procurando que sean de buena calidad y omitiendo las que no contribuyan a una mejor comprensión del texto. Las imágenes se remitirán en archivos fotográficos electrónicos, con una resolución de 300 puntos pulgada. Siempre que se considere necesario se utilizarán recursos gráficos (flechas, asteriscos) para destacar la parte esencial de la imagen. Se procurará en lo posible evitar la identificación de los enfermos, en cualquier caso se deberá disponer de su permiso por escrito.

5. Las gráficas (hasta un máximo de seis). Se tendrán en cuenta las mismas normas del apartado 4 para las imágenes. Las imágenes y gráficas irán numeradas de manera correlativa y conjunta, como las figuras.

6. Las tablas se presentarán en hojas aparte que incluirán: a) numeración de la tabla con números arábigos; b) enunciado (título) correspondiente, y c) una sola tabla por hoja. Se procurará que sean claras y sin rectificaciones; las siglas y abreviaturas se acompaÞarán siempre de una nota explicativa al pie. Si una tabla ocupa más de una página se repetirán los encabezamientos en la hoja siguiente. La Revista admitirá tablas que ocupen hasta un máximo de una página impresa. Cuando se haya efectuado un estudio estadístico se indicará a pie de tabla la técnica empleada y el nivel de significación, si no se hubiera incluido en el texto de la tabla.

7. El Comité de Redacción acusará recibo electrónicamente de los trabajos enviados a la Revista e informará acerca de su aceptación. Todos los manuscritos originales se someterán a revisión por pares. Las Cartas al Editor pueden ser aceptadas directamente por el Comité de Redacción. Siempre que el Comité de Redacción sugiera efectuar modificaciones en los artículos, los autores deberán remitir, junto a la nueva versión del artículo, una relación de las modificaciones realizadas, tanto las sugeridas por el propio Comitéde Redacción como las que figuran en los informes de los expertos consultados.

8. Envío de manuscritos. Los manuscritos deben remitirse por vía web a través de http://ees.elsevier.com/medcli. Para enviar un artículo debe registrarse en la opción register del menúsuperior gris de la página y seguir las instrucciones de la pantalla.

La utilización de este método permite seguir el estado del artículo directamente a través de esta página web. El manuscrito se debe acompañar de una carta de presentación que incluya: 1) Sección de la revista donde se desea publicar el trabajo. 2) Declaración de que el manuscrito es original y no se encuentra en proceso de evaluación por ninguna otra revista científica. 3) Seguimiento de las responsabilidades éticas, incluyendo los requisitos de autoría y la declaración de la existencia o no de conflicto de intereses.

ANNEX 3. REVISIÓ SISTEMÀTICA D'ESTUDIS

Med Clin (Barc). 2010;**135(11)**:507–511

MEDICINA CLINICA

www.elsevier.es/medicinaclinica

Artículo especial

Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis

PRISMA declaration: A proposal to improve the publication of systematic reviews and meta-analyses

Gerard Urrútia [a,b,*,◆] y Xavier Bonfill [a,b,c]

[a] *Servei d'Epidemiologia Clínica i Salut Pública, Centro Cochrane Iberoamericano, Hospital de la Sant Creu i Sant Pau, Barcelona, España*
[b] *CIBER Epidemiología y Salud Pública, Barcelona, España*
[c] *Medicina Preventiva, Departamento de Pediatria, Obstetricia y Ginecologia, Universitat Autònoma de Barcelona, Barcelona, España*

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Recibido el 3 de diciembre de 2009
Aceptado el 12 de enero de 2010
On-line el 5 de marzo de 2010

Precedente de la declaración QUOROM

En el año 1999, después de 3 años de trabajo, se publicó la declaración QUOROM, cuyo objetivo era establecer unas normas para mejorar la calidad de la presentación de los metaanálisis de ensayos clínicos aleatorizados[1]. Se publicó un comentario de este documento en un monográfico sobre listas de comprobación para autores, revisores y editores de revistas médicas en Medicina Clínica[2].

En resumen, la declaración Quality Of Reporting Of Meta-analysis (QUOROM) incluye una lista de comprobación estructurada con 18 ítems que los autores de un metaanálisis, y también los editores de revistas, deberían considerar a la hora de publicar su trabajo en forma de artículo en una revista médica. Además, incluye un diagrama de flujo que describe todo el proceso, desde la identificación inicial de los estudios potencialmente relevantes hasta la selección definitiva de éstos. La finalidad de QUOROM era animar a los autores a que proporcionaran toda aquella información que resulta esencial para interpretar y utilizar adecuadamente los resultados de un metaanálisis.

Numerosos estudios realizados con posterioridad a la publicación de QUOROM han mostrado que la calidad de los metaanálisis publicados en revistas médicas todavía es deficiente. A pesar de esto, y a diferencia de otras iniciativas similares como CONSORT (dirigida a ensayos clínicos), la declaración QUOROM no parece haber logrado el mismo grado de aceptación por parte de los editores de revistas biomédicas, aun cuando el número de revisiones sistemáticas y metaanálisis que se publican anualmente es muy elevado y en número creciente (se estima en 2.500 solamente las revisiones publicadas en inglés)[3]. Una de las causas de este menor éxito de QUOROM podría ser que se han realizado pocos estudios que demuestren su impacto en la mejoría de la calidad de los metaanálisis después de su publicación o de su inclusión como criterio editorial de las revistas biomédicas[4].

En el momento de su publicación, el grupo de trabajo QUOROM estableció la necesidad de una revisión y una actualización periódica de las directrices conforme a la nueva evidencia publicada, y que podía comportar mantener, eliminar o añadir nuevos ítems a la lista de comprobación inicial. Como resultado, en julio de 2009 se publicó la declaración PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses), una actualización y ampliación de QUOROM.

Declaración PRISMA

A diferencia de QUOROM, la declaración PRISMA[5] viene acompañada de un extenso documento donde se detalla la explicación o la justificación de cada uno de los 27 ítems propuestos, así como el proceso de elaboración de estas directrices[6]. Reflejo del objetivo pedagógico de este documento, se incluyen 7 cuadros que proporcionan una explicación más detallada de ciertos aspectos clave sobre la metodología y la conducción de revisiones sistemáticas (terminología, formulación de la pregunta de investigación, identificación de los estudios y extracción de datos, calidad de los estudios y riesgo de sesgo,

* Autor para correspondencia.
Correo electrónico: gurrutia@santpau.cat (G. Urrútia).
◆ Este autor está adscrito al Programa de Doctorado en Salud Pública y Metodología de la Investigación Biomédica de la Universitat Autònoma de Barcelona.

0025-7753/$ - see front matter
doi:10.1016/j.medcli.2010.01.015

cuándo combinar datos, metaanálisis y análisis de la consistencia, y sesgo de publicación selectiva de estudios o resultados).

Como señalan sus autores, PRISMA incorpora varios aspectos conceptuales y metodológicos novedosos relacionados con la metodología de las revisiones sistemáticas que han emergido en los últimos años, período en el que ha habido una importante producción de revisiones y de investigación sobre éstas. Uno de ellos es el uso de la terminología utilizada para describir una revisión sistemática y un metaanálisis, hasta la fecha algo confusa e inconsistente. Los autores de PRISMA han adoptado las definiciones de la Colaboración Cochrane[7]. Según ésta, el metaanálisis (síntesis cuantitativa de resultados) sería solamente una parte, deseable pero no siempre posible, de un proceso más amplio —consistente en diversos pasos sucesivos que deben ser explícitos y reproducibles—, conocido como revisión sistemática. Por otro lado, PRISMA tiene una aplicabilidad más amplia que su predecesor QUOROM, ya que no se limita solamente a los metaanálisis de ensayos clínicos aleatorizados, sino que también es útil para las revisiones de otro tipo de estudios. La primera consecuencia de estos cambios se refleja en el nombre de PRISMA: «…directrices para la publicación de revisiones sistemáticas y metaanálisis de estudios que evalúan intervenciones sanitarias».

Cabe señalar que la elaboración y la publicación de PRISMA ha coincidido en el tiempo con la actualización y la modificación sustancial del Cochrane Handbook for Systematic Reviews of Interventions, versión 5 (Manual del Revisor Cochrane), cuya finalidad principal es ayudar a los autores de revisiones Cochrane a ser sistemáticos y explícitos en el desarrollo de éstas[8]. No en vano muchos de los 29 autores del grupo de trabajo PRISMA están también involucrados como asesores metodológicos de la Colaboración Cochrane. Por esta razón, el Manual del Revisor Cochrane incorpora muchos de los mismos cambios propuestos en PRISMA.

Aspectos conceptuales novedosos introducidos en PRISMA

Los autores de PRISMA identifican 4 aspectos conceptuales novedosos que conllevan la adición de nuevos ítems a la lista de comprobación:

1. El carácter iterativo del proceso de desarrollo de una revisión sistemática. La conducción de una revisión sistemática es un proceso complejo que implica numerosos juicios y decisiones por parte de los autores. Con el fin de minimizar el riesgo de sesgo en el proceso de la revisión, estos juicios y decisiones no deberían estar influidos por los resultados de los estudios incluidos en la revisión. El conocimiento anticipado que los autores puedan tener de los resultados de los estudios potencialmente elegibles podría, por ejemplo, influir en la pregunta que la revisión trata de responder, en los criterios de selección de los estudios, en la elección de las comparaciones que se van a analizar o en los resultados que se van a reportar en la revisión. Dada la naturaleza retrospectiva de las revisiones, es muy importante que los métodos que utilice la revisión se establezcan y se documenten a priori. La publicación previa del protocolo, tal como ocurre en las revisiones Cochrane, reduce el impacto de los sesgos inherentes al autor y promueve la transparencia acerca de los métodos y del proceso, además de evitar revisiones redundantes.
   La existencia de un protocolo no excluye que pueda haber razones justificadas para modificar el protocolo de revisión original. En este caso, es importante garantizar que los cambios (por ejemplo, la exclusión de estudios previamente seleccionados) no se han tomado a posteriori al saber cómo afectarán a los resultados. Tales decisiones son muy susceptibles de introducir sesgos y deben evitarse.
   Por todas las razones hasta aquí apuntadas, es importante que se detalle si existía un protocolo previo a la revisión y que se hagan explícitos los cambios introducidos, así como su justificación. Solo así es posible juzgar si tales cambios a posteriori fueron apropiados o no y si podrían introducir algún tipo de sesgo.
2. La conducción y la publicación de un estudio de investigación son conceptos distintos. Aunque esta distinción es menos evidente para una revisión sistemática que en el caso de un estudio primario, la publicación y la conducción de una revisión sistemática están muy entrelazadas. Por ejemplo, no comunicar si se ha realizado o no la evaluación del riesgo de sesgo de los estudios incluidos en una revisión es señal de una conducción deficiente dada la importancia de este aspecto en el proceso de la revisión.
3. Evaluación del riesgo de sesgo al nivel de los estudios o de los resultados. El grado en que una revisión puede arrojar conclusiones fiables sobre los efectos de una intervención depende de la validez de los datos y los resultados de los estudios incluidos en la revisión. Así, por ejemplo, un metaanálisis de estudios con poca validez interna producirá unos resultados engañosos. Por esto, la evaluación de la validez de los estudios incluidos es un componente esencial de una revisión y debe considerarse en los análisis, la interpretación y las conclusiones de ésta.
   La declaración QUOROM, y también las versiones previas del Manual del Revisor Cochrane, se referían a este importante aspecto como a la «evaluación de la calidad». No obstante, hoy se recomienda evitar el término «calidad» para hablar de «riesgo de sesgo», ya que el primero se refiere a la valoración de hasta qué punto el estudio se ha llevado a cabo según los mejores estándares posibles, lo que no implica necesariamente que se haya evitado el riesgo de sesgo en los resultados. Por ejemplo, el mejor estudio posible bajo determinadas circunstancias puede tener un importante riesgo de sesgo. El aspecto clave en una revisión es preguntarse hasta qué punto los resultados de los estudios incluidos pueden tener credibilidad, de modo que la evaluación del «riesgo de sesgo» aborda esta misma cuestión. Además, la valoración de la calidad del estudio está condicionada por la calidad del reporte y la disponibilidad de información detallada acerca de los métodos, información de la que no siempre se dispone.
   De entre las numerosas opciones existentes para valorar el riesgo de sesgo, los autores de PRISMA promueven un sistema basado en la evaluación de diversos componentes clave del diseño y la ejecución de los estudios para los que existen sólidas evidencias empíricas acerca de su relación con el sesgo, como es el caso del nuevo sistema adoptado por la Colaboración Cochrane[9].
   La evaluación rigurosa del riesgo de sesgo requiere tanto de una evaluación al nivel de los estudios (por ejemplo, evaluar aspectos genéricos del diseño, como la ocultación de la secuencia aleatorizada) como, en algunas ocasiones, también de los resultados (fiabilidad y validez de los datos para cada resultado específico a partir de los métodos utilizados para su medición en cada estudio individual). La calidad de la evidencia puede diferir entre resultados, aun dentro de un mismo estudio. Por ejemplo, algunas variables de resultado pueden haberse registrado de manera muy cuidadosa y exhaustiva, mientras que otras no tanto. Esta información debe referirse para permitir una evaluación explícita de hasta qué punto la estimación del efecto es correcta.
4. Importancia de los sesgos relacionados con la publicación. Las revisiones sistemáticas deben tratar de incorporar información de todos los estudios que sean relevantes para el tema de la revisión. No obstante, la ausencia de información de algunos estudios puede cuestionar la validez de la revisión. Este problema puede ocurrir cuando no se publican estudios completos (sesgo de publicación) o porque la información

publicada es incompleta o inexacta (por ejemplo, el informe selectivo de resultados). Tanto en un caso (omisión de estudios completos, habitualmente por causa de unos resultados no favorables) como en otro (omisión de resultados dentro de un estudio individual) existe evidencia empírica que demuestra su relación con el sesgo al estar basada la revisión en una muestra sesgada y no representativa de toda la información existente (publicada o no). Por todas estas razones, los autores de una revisión deben llevar a cabo estrategias para detectar estos posibles sesgos, así como investigar su posible relación con el efecto de la intervención y la precisión de la estimación.

Diferencias principales entre QUOROM y PRISMA

La nueva lista de comprobación PRISMA difiere en varios aspectos con respecto a QUOROM. En primer lugar, PRISMA presenta un total de 27 ítems (tabla 1) frente a los 18 ítems de QUOROM. Algunos ítems se han incorporado de nuevo a la lista, mientras que otros resultan del desdoblamiento de ítems ya presentes en QUOROM. Además, algunos ítems están interrelacionados para mejorar la consistencia en el informe de la revisión sistemática por parte de los autores.

Los cambios más relevantes introducidos en la lista de comprobación de PRISMA se presentan en la tabla 2.

El diagrama de flujo también se ha modificado, y este nuevo diagrama de PRISMA es más detallado e informativo (fig. 1). El diagrama de QUOROM partía de los ensayos clínicos aleatorizados potencialmente relevantes sometidos a cribado y terminaba con los ensayos que proporcionan información útil para cada resultado en el metaanálisis, mientras que el de PRISMA parte desde el inicio mismo del proceso (los registros o las citas identificados en las búsquedas realizadas en cada una de las diferentes bases de datos u otras fuentes utilizadas), continúa por el número total de registros o citas únicas una vez eliminados los duplicados y termina con los estudios individuales incluidos

Tabla 1
Lista de comprobación de los ítems para incluir en la publicación de una revisión sistemática (con o sin metaanálisis). La declaración PRISMA

Sección/tema	*Número*	*Ítem*
Título		
Título	1	Identificar la publicación como revisión sistemática, metaanálisis o ambos
Resumen		
Resumen estructurado	2	Facilitar un resumen estructurado que incluya, según corresponda: antecedentes; objetivos; fuente de los datos; criterios de elegibilidad de los estudios, participantes e intervenciones; evaluación de los estudios y métodos de síntesis; resultados; limitaciones; conclusiones e implicaciones de los hallazgos principales; número de registro de la revisión sistemática
Introducción		
Justificación	3	Describir la justificación de la revisión en el contexto de lo que ya se conoce sobre el tema
Objetivos	4	Plantear de forma explícita las preguntas que se desea contestar en relación con los participantes, las intervenciones, las comparaciones, los resultados y el diseño de los estudios (PICOS)*
Métodos		
Protocolo y registro	5	Indicar si existe un protocolo de revisión al que se pueda acceder (por ej., dirección web) y, si está disponible, la información sobre el registro, incluyendo su número de registro
Criterios de elegibilidad	6	Especificar las características de los estudios (por ej., PICOS, duración del seguimiento) y de las características (por ej., años abarcados, idiomas o estatus de publicación) utilizadas como criterios de elegibilidad y su justificación
Fuentes de información	7	Describir todas las fuentes de información (por ej., bases de datos y períodos de búsqueda, contacto con los autores para identificar estudios adicionales, etc.) en la búsqueda y la fecha de la última búsqueda realizada
Búsqueda	8	Presentar la estrategia completa de búsqueda electrónica en, al menos, una base de datos, incluyendo los límites utilizados, de tal forma que pueda ser reproducible
Selección de los estudios	9	Especificar el proceso de selección de los estudios (por ej., el cribado y la elegibilidad incluidos en la revisión sistemática y, cuando sea pertinente, incluidos en el metaanálisis)
Proceso de extracción de datos	10	Describir los métodos para la extracción de datos de las publicaciones (por ej., formularios pilotado, por duplicado y de forma independiente) y cualquier proceso para obtener y confirmar datos por parte de los investigadores
Lista de datos	11	Listar y definir todas las variables para las que se buscaron datos (por ej., PICOS, fuente de financiación) y cualquier asunción y simplificación que se hayan hecho
Riesgo de sesgo en los estudios individuales	12	Describir los métodos utilizados para evaluar el riesgo de sesgo en los estudios individuales (especificar si se realizó al nivel de los estudios o de los resultados) y cómo esta información se ha utilizado en la síntesis de datos
Medidas de resumen	13	Especificar las principales medidas de resumen (por ej., razón de riesgos o diferencia de medias)
Síntesis de resultados	14	Describir los métodos para manejar los datos y combinar resultados de los estudios, cuando esto es posible, incluyendo medidas de consistencia (por ej., ítem 2) para cada metaanálisis
Riesgo de sesgo entre los estudios	15	Especificar cualquier evaluación del riesgo de sesgo que pueda afectar la evidencia acumulativa (por ej., sesgo de publicación o comunicación selectiva)
Análisis adicionales	16	Describir los métodos adicionales de análisis (por ej., análisis de sensibilidad o de subgrupos, metarregresión), en el caso de que se hiciera, indicar cuáles fueron preespecificados
Resultados		
Selección de estudios	17	Facilitar el número de estudios cribados, evaluados para su elegibilidad e incluidos en la revisión, y detallar las razones para su exclusión en cada etapa, idealmente mediante un diagrama de flujo
Características de los estudios	18	Para cada estudio presentar las características para las que se extrajeron los datos (por ej., tamaño, PICOS y duración del seguimiento) y proporcionar las citas bibliográficas

Tabla 1 (*continuación*)

Sección/tema	*Número*	*Ítem*
Riesgo de sesgo en los estudios	19	Presentar datos sobre el riesgo de sesgo en cada estudio y, si está disponible, cualquier evaluación del sesgo en los resultados (ver ítem 12)
Resultados de los estudios individuales	20	Para cada resultado considerado en cada estudio (beneficios o daños), presentar: a) el dato resumen para cada grupo de intervención y b) la estimación del efecto con su intervalo de confianza, idealmente de forma gráfica mediante un diagrama de bosque (*forest plot*)
Síntesis de los resultados	21	Presentar los resultados de todos los metaanálisis realizados, incluyendo los intervalos de confianza y las medidas de consistencia
Riesgo de sesgo entre los estudios	22	Presentar los resultados de cualquier evaluación del riesgo de sesgo entre los estudios (ver ítem 15)
Análisis adicionales	23	Facilitar los resultados de cualquier análisis adicional, en el caso de que se hayan realizado (por ej., análisis de sensibilidad o de subgrupos, metarregresión [ver ítem 16])
Discusión		
Resumen de la evidencia	24	Resumir los hallazgos principales, incluyendo la fortaleza de las evidencias para cada resultado principal; considerar su relevancia para grupos clave (por ej., proveedores de cuidados, usuarios y decisores en salud)
Limitaciones	25	Discutir las limitaciones de los estudios y de los resultados (por ej., riesgo de sesgo) y de la revisión (por ej., obtención incompleta de los estudios identificados o comunicación selectiva)
Conclusiones	26	Proporcionar una interpretación general de los resultados en el contexto de otras evidencias, así como las implicaciones para la futura investigación
Financiación		
Financiación	27	Describir las fuentes de financiación de la revisión sistemática y otro tipo de apoyos (por ej., aporte de los datos), así como el rol de los financiadores en la revisión sistemática

* PICOS: se trata de un acrónimo formado por: P: participants; I: interventions; C: comparisons; O: outcomes; S: study design.

Tabla 2
Cambios más relevantes introducidos en la lista de comprobación de PRISMA

Sección/tema	**Ítem**	**Comentario**
Título		PRISMA solicita la identificación de la publicación como revisión sistemática, metaanálisis o ambos, mientras que QUOROM se refería exclusivamente a metaanálisis de ensayos clínicos
Resumen		Tanto QUOROM como PRISMA solicitan a los autores un resumen estructurado, pero PRISMA no especifica su formato, aunque sí hace recomendaciones
Introducción	Objetivo	Este nuevo ítem (4) se focaliza en la pregunta explícita que aborda la revisión utilizando el formato PICO (descripción de los participantes, las intervenciones, las comparaciones y las medidas de resultado de la revisión sistemática), así como el tipo de estudio (diseño);este ítem está interrelacionado con los ítems 6, 11 y 18
Métodos	Protocolo	Este nuevo ítem (5) solicita al autor que explique si un protocolo precedió la revisión y, en tal caso, cómo puede accederse a éste
Métodos	Búsqueda	PRISMA desdobla el ítem sobre la «búsqueda» de QUOROM en 2 ítems: a) fuentes de información (ítem 7) y b) búsqueda (ítem 8). Aunque la estrategia de búsqueda se reporta tanto en QUOROM como en PRISMA, éste solicita a los autores que proporcionen la descripción completa de, al menos, una estrategia de búsqueda electrónica (ítem 8), sin ésta no es posible reproducir la búsqueda
Métodos	Evaluación del riesgo desesgo en los estudios incluidos	En QUOROM este ítem aparecía como «evaluación de la calidad». Ahora, este ítem (12) se focaliza en la evaluación del riesgo de sesgo dentro de cada estudio incluido en la revisión. Además, este ítem está interrelacionado con otro nuevo ítem incorporado en PRISMA: la comunicación de esta información en los resultados (ítem 19). También se introduce el nuevo concepto de evaluación del sesgo al nivel de los resultados

PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses.

en la síntesis cualitativa (revisión sistemática) y cuantitativa (metaanálisis). PRISMA también establece la diferencia en cada etapa del proceso entre los registros o las referencias bibliográficas (resultado de aplicar las estrategias de búsqueda electrónica en las bases bibliográficas), los artículos a texto completo (artículos que deben obtenerse a texto completo para decidir con seguridad acerca de su elegibilidad o no) y los estudios individuales (estudios que cumplen los criterios de elegibilidad de la revisión y que pueden corresponderse con una o más publicaciones o artículos).

En cuanto a las expectativas para el futuro, PRISMA se ha concebido como una herramienta para contribuir a mejorar la claridad y la transparencia en la publicación de revisiones sistemáticas. Por el contrario, PRISMA no se ha formulado como un instrumento para valorar la calidad de las revisiones y no debería utilizarse como tal.

Como los autores señalan, seguir las recomendaciones de la lista de comprobación PRISMA puede suponer aumentar la extensión de la publicación de una revisión sistemática, lo que puede colisionar con las normas editoriales de algunas revistas. No obstante, las ventajas de proporcionar a los lectores información completa, clara y transparente superan los inconvenientes de tener que leer un texto algo más extenso.

Si las revistas biomédicas van a adoptar la declaración PRISMA, tal como ya lo fueron otras propuestas como CONSORT, debería existir evidencia que demuestre que PRISMA realmente mejora la claridad y la transparencia de las revisiones sistemáticas publicadas. Los autores de PRISMA tienen la intención de evaluar

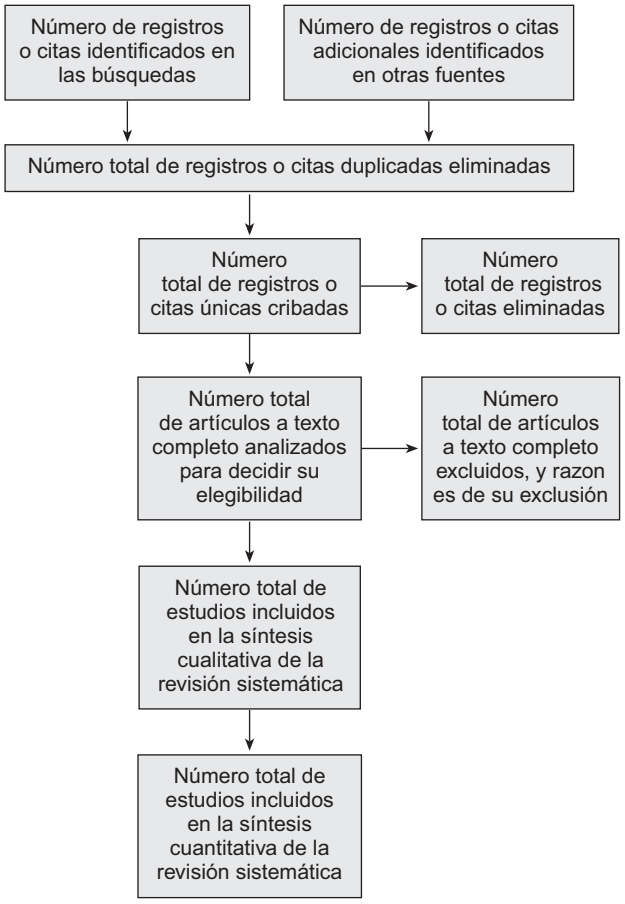

Figura 1. Diagrama de flujo de la información a través de las diferentes fases de una revisión sistemática.

los beneficios (calidad mejorada de la publicación) y los posibles efectos adversos (aumento de la extensión del texto) de PRISMA y animan a otros a hacer lo mismo.

Una de las limitaciones apuntadas por los autores de PRISMA es que no se realizó una revisión sistemática previa para la confección de la lista de comprobación. No obstante, PRISMA se desarrolló sobre la base de las evidencias disponibles siempre que fue posible. Solo se incluyeron en el listado ítems para los que hubiera evidencia de una relación entre su omisión y el riesgo de sesgo, o bien que existiera amplio consenso acerca de su necesidad para evaluar la fiabilidad de una revisión.

Con el fin de mantener PRISMA actualizado y basado en la evidencia, los autores realizarán una revisión periódica de la literatura médica metodológica. Por ejemplo, no existen evidencias publicadas para apoyar algunos ítems, como por ejemplo, si el entrenamiento previo contribuye a mejorar la precisión y la fiabilidad en la extracción de los datos. Por esto, se espera que PRISMA actúe de catalizador de futuros estudios que generen la evidencia que pueda incorporarse en futuras actualizaciones del listado de comprobación.

Los autores confían en que PRISMA tendrá una revisión y una actualización más frecuentes, así como también una mejor implementación que la que tuvo en su momento QUOROM. Este optimismo se basa en el número creciente de revisiones sistemáticas publicadas y su utilización, también creciente, por parte de los proveedores, los decisores y los gestores en salud para fundamentar las decisiones clínicas, sanitarias y de investigación. También basan su optimismo en los beneficios esperados del desarrollo de la Red EQUATOR (Enhancing the QUAlity and Transparency Of health Research), cuyo objetivo es ayudar a individuos y a grupos interesados en el desarrollo de futuras directrices (o su traducción) dirigidas a mejorar la calidad de las publicaciones en ciencias de la salud[10,11]. La red dispone de una web donde se facilita el acceso a los recursos (http://www.equator-network.org/).

El éxito de PRISMA está por verse, y dependerá de la medida en que las revistas biomédicas y los grupos editoriales la adopten y la incluyan en las «Instrucciones para autores», donde se establezca un *link* con la web de PRISMA, y realicen acciones editoriales para llamar la atención de los lectores y los potenciales autores de revisiones acerca de su importancia.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. Lancet. 1999;354:1896–900.
2. Urrutia G, Tort S, Bonfill X. Metaanálisis (QUOROM). Med Clin (Barc). 2005;125:32–7.
3. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PLoS Med. 2007;4:e78, doi:10.1371/journal.pmed.0040078.
4. Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. Crit Care. 2005;9:R575e82.
5. Site web de PRISMA. [consultado 23/02/2010] Disponible en: http://www.prisma-statement.org.
6. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. BMJ. 2009;339:b2700, doi:10.1136/bmj.b2700.
7. Green S, Higgins JPT, Alderson P, Clarke M, Mulrow CD. Chapter 1: What is a systematic review?. [actualizado 2/2008]. En: Higgins JPT, Green S, editores. Cochrane handbook for systematic reviews of interventions version 5.0.0. The Cochrane Collaboration; 2008 [consultado 26/11/2009]. Disponible en: http://www.cochrane-handbook.org/.
8. Higgins JPT, Green S, editores. Cochrane handbook for systematic reviews of interventions version 5.0.0. The Cochrane Collaboration; 2008 [consultado 26/11/2009]. Disponible en: http://www.cochrane-handbook.org/.
9. Higgins JPT, Altman DG. Chapter 8: Assessing risk of bias in included studies. [actualizado 2/2008]. En: Higgins JPT, Green S, editores. Cochrane handbook for systematic reviews of interventions version 5.0.0. The Cochrane Collaboration; 2008 [consultado 26/11/2009]. Disponible en: http://www.cochrane-handbook.org/.
10. Altman DG, Simera I, Hoey J, Moher D, Schulz K. EQUATOR: Reporting guidelines for health research. Lancet. 2008;371:1149e50.
11. Simera I, Altman DG, Moher D, Schulz KF, Hoey J. Guidelines for reporting health research: The EQUATOR network's survey of guideline authors. PLoS Med. 2008;5:e139, doi:10.1371/journal.pmed.0050139.

ANNEX 4. GUIA CONSORT

ELSEVIER

Journal of Clinical Epidemiology 63 (2010) e1−e37

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials

David Moher[a,*], Sally Hopewell[b], Kenneth F. Schulz[c], Victor Montori[d], Peter C. Gøtzsche[e], P.J. Devereaux[f], Diana Elbourne[g], Matthias Egger[h], Douglas G. Altman[b]

[a]*Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa Hospital, Ottawa, Ontario, Canada, K1H 8L6*
[b]*Centre for Statistics in Medicine, University of Oxford, Wolfson College, Oxford*
[c]*Family Health International, Research Triangle Park, NC 27709, USA*
[d]*UK Knowledge and Encounter Research Unit, Mayo Clinic, Rochester, MN, USA*
[e]*The Nordic Cochrane Centre, Rigshospitalet, Blegdamsvej 9, Copenhagen, Denmark*
[f]*McMaster University Health Sciences Centre, Hamilton, Canada*
[g]*Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London*
[h]*Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland*

Accepted 8 February 2010

Abstract

Overwhelming evidence shows the quality of reporting of randomised controlled trials (RCTs) is not optimal. Without transparent reporting, readers cannot judge the reliability and validity of trial findings nor extract information for systematic reviews. Recent methodological analyses indicate that inadequate reporting and design are associated with biased estimates of treatment effects. Such systematic error is seriously damaging to RCTs, which are considered the gold standard for evaluating interventions because of their ability to minimise or avoid bias.

A group of scientists and editors developed the CONSORT (Consolidated Standards of Reporting Trials) statement to improve the quality of reporting of RCTs. It was first published in 1996 and updated in 2001. The statement consists of a checklist and flow diagram that authors can use for reporting an RCT. Many leading medical journals and major international editorial groups have endorsed the CONSORT statement. The statement facilitates critical appraisal and interpretation of RCTs.

During the 2001 CONSORT revision, it became clear that explanation and elaboration of the principles underlying the CONSORT statement would help investigators and others to write or appraise trial reports. A CONSORT explanation and elaboration article was published in 2001 alongside the 2001 version of the CONSORT statement.

After an expert meeting in January 2007, the CONSORT statement has been further revised and is published as the CONSORT 2010 Statement. This update improves the wording and clarity of the previous checklist and incorporates recommendations related to topics that have only recently received recognition, such as selective outcome reporting bias.

This explanatory and elaboration document—intended to enhance the use, understanding, and dissemination of the CONSORT statement—has also been extensively revised. It presents the meaning and rationale for each new and updated checklist item providing examples of good reporting and, where possible, references to relevant empirical studies. Several examples of flow diagrams are included.

The CONSORT 2010 Statement, this revised explanatory and elaboration document, and the associated website (www.consort-statement.org) should be helpful resources to improve reporting of randomised trials.

In order to encourage dissemination of the CONSORT 2010 Statement, this article is freely accessible on bmj.com and will also be published in the Journal of Clinical Epidemiology. The authors jointly hold the copyright for this article. For details on further use, see the CONSORt website (www.consort-statement.org).

* Correspondence to: D Moher.
E-mail address: dmoher@ohri.ca (D. Moher).

> ''The whole of medicine depends on the transparent reporting of clinical trials''[1].

Well designed and properly executed randomised controlled trials (RCTs) provide the most reliable evidence on the efficacy of healthcare interventions, but trials with inadequate methods are associated with bias, especially exaggerated treatment effects [2−5]. Biased results from poorly designed and reported trials can mislead decision making in health care at all levels, from treatment decisions for a patient to formulation of national public health policies.

0895-4356/$ − see front matter
doi: 10.1016/j.jclinepi.2010.03.004

Critical appraisal of the quality of clinical trials is possible only if the design, conduct, and analysis of RCTs are thoroughly and accurately described in the report. Far from being transparent, the reporting of RCTs is often incomplete [6−9], compounding problems arising from poor methodology [10−15].

1. Incomplete and inaccurate reporting

Many reviews have documented deficiencies in reports of clinical trials. For example, information on the method used in a trial to assign participants to comparison groups was reported in only 21% of 519 trial reports indexed in PubMed in 2000 [16], and only 34% of 616 reports indexed in 2006 [17]. Similarly, only 45% of trial reports indexed in PubMed in 2000 [16] and 53% in 2006 [17] defined a primary end point, and only 27% in 2000 and 45% in 2006 reported a sample size calculation. Reporting is not only often incomplete but also sometimes inaccurate. Of 119 reports stating that all participants were included in the analysis in the groups to which they were originally assigned (intention-to-treat analysis), 15 (13%) excluded patients or did not analyse all patients as allocated [18]. Many other reviews have found that inadequate reporting is common in specialty journals [16,19] and journals published in languages other than English [20,21].

Proper randomisation reduces selection bias at trial entry and is the crucial component of high quality RCTs [22]. Successful randomisation hinges on two steps: generation of an unpredictable allocation sequence and concealment of this sequence from the investigators enrolling participants (see Box 1) [2,23].

Unfortunately, despite that central role, reporting of the methods used for allocation of participants to interventions is also generally inadequate. For example, 5% of 206 reports of supposed RCTs in obstetrics and gynaecology journals described studies that were not truly randomised [23]. This estimate is conservative, as most reports do not at present provide adequate information about the method of allocation [20,23,30−33].

2. Improving the reporting of RCTs: the CONSORT statement

DerSimonian and colleagues suggested that ‘‘editors could greatly improve the reporting of clinical trials by providing authors with a list of items that they expected to be strictly reported’’[34]. Early in the 1990s, two groups of journal editors, trialists, and methodologists independently published recommendations on the reporting of trials [35,36]. In a subsequent editorial, Rennie urged the two groups to meet and develop a common set of recommendations [37]; the outcome was the CONSORT statement (Consolidated Standards of Reporting Trials) [38].

The CONSORT statement (or simply CONSORT) comprises a checklist of essential items that should be included in reports of RCTs and a diagram for documenting the flow of participants through a trial. It is aimed at primary reports of RCTs with two group, parallel designs. Most of CONSORT is also relevant to a wider class of trial designs, such as noninferiority, equivalence, factorial, cluster, and crossover trials. Extensions to the CONSORT checklist for reporting trials with some of these designs have been published [39−41], as have those for reporting certain types of data (harms [42]), types of interventions (non-pharmacological treatments [43], herbal interventions [44]), and abstracts [45].

The objective of CONSORT is to provide guidance to authors about how to improve the reporting of their trials.

Box 1. Treatment allocation. What’s so special about randomisation?

The method used to assign interventions to trial participants is a crucial aspect of clinical trial design. Random assignment is the preferred method; it has been successfully used regularly in trials for more than 50 years [24]. Randomisation has three major advantages [25]. First, when properly implemented, it eliminates selection bias, balancing both known and unknown prognostic factors, in the assignment of treatments. Without randomisation, treatment comparisons may be prejudiced, whether consciously or not, by selection of participants of a particular kind to receive a particular treatment. Second, random assignment permits the use of probability theory to express the likelihood that any difference in outcome between intervention groups merely reflects chance [26]. Third, random allocation, in some situations, facilitates blinding the identity of treatments to the investigators, participants, and evaluators, possibly by use of a placebo, which reduces bias after assignment of treatments [27]. Of these three advantages, reducing selection bias at trial entry is usually the most important [28].

Successful randomisation in practice depends on two interrelated aspects—adequate generation of an unpredictable allocation sequence and concealment of that sequence until assignment occurs [2,23]. A key issue is whether the schedule is known or predictable by the people involved in allocating participants to the comparison groups [29]. The treatment allocation system should thus be set up so that the person enrolling participants does not know in advance which treatment the next person will get, a process termed allocation concealment [2,23]. Proper allocation concealment shields knowledge of forthcoming assignments, whereas proper random sequences prevent correct anticipation of future assignments based on knowledge of past assignments.

Trial reports need be clear, complete, and transparent. Readers, peer reviewers, and editors can also use CONSORT to help them critically appraise and interpret reports of RCTs. However, CONSORT was not meant to be used as a quality assessment instrument. Rather, the content of CONSORT focuses on items related to the internal and external validity of trials. Many items not explicitly mentioned in CONSORT should also be included in a report, such as information about approval by an ethics committee, obtaining informed consent from participants, and, where relevant, existence of a data safety and monitoring committee. In addition, any other aspects of a trial that are mentioned should be properly reported, such as information pertinent to cost effectiveness analysis [46−48].

Since its publication in 1996, CONSORT has been supported by more than 400 journals (www.consort-statement.org) and several editorial groups, such as the International Committee of Medical Journal Editors [49]. The introduction of CONSORT within journals is associated with improved quality of reports of RCTs [17,50,51]. However, CONSORT is an ongoing initiative, and the CONSORT statement is revised periodically [3]. CONSORT was last revised nine years ago, in 2001 [52−54]. Since then the evidence base to inform CONSORT has grown considerably; empirical data have highlighted new concerns regarding the reporting of RCTs, such as selective outcome reporting [55−57]. A CONSORT Group meeting was therefore convened in January 2007, in Canada, to revise the 2001 CONSORT statement and its accompanying explanation and elaboration document. The revised checklist is shown in Table 1 and the flow diagram, not revised, in Fig 1 [52−54].

3. The CONSORT 2010 Statement: explanation and elaboration

During the 2001 CONSORT revision, it became clear that explanation and elaboration of the principles underlying the CONSORT statement would help investigators and others to write or appraise trial reports. The CONSORT explanation and elaboration article [58] was published in 2001 alongside the 2001 version of the CONSORT statement. It discussed the rationale and scientific background for each item and provided published examples of good reporting. The rationale for revising that article is similar to that for revising the statement, described above. We briefly describe below the main additions and deletions to this version of the explanation and elaboration article.

4. The CONSORT 2010 Explanation and Elaboration: changes

We have made several substantive and some cosmetic changes to this version of the CONSORT explanatory document (full details are highlighted in the 2010 version of the CONSORT statement [59]). Some reflect changes to the CONSORT checklist; there are three new checklist items in the CONSORT 2010 checklist—such as item 24, which asks authors to report where their trial protocol can be accessed. We have also updated some existing explanations, including adding more recent references to methodological evidence, and used some better examples. We have removed the glossary, which is now available on the CONSORT website (www.consort-statement.org). Where possible, we describe the findings of relevant empirical studies. Many excellent books on clinical trials offer fuller discussion of methodological issues [60−62]. Finally, for convenience, we sometimes refer to ‘‘treatments’’ and ‘‘patients,’’ although we recognise that not all interventions evaluated in RCTs are treatments and not all participants are patients.

5. Checklist items

5.1. Title and abstract

5.1.1. Item 1a. Identification as a randomised trial in the title

Example—‘‘Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety’’[63].

Explanation—The ability to identify a report of a randomised trial in an electronic database depends to a large extent on how it was indexed. Indexers may not classify a report as a randomised trial if the authors do not explicitly report this information [64]. To help ensure that a study is appropriately indexed and easily identified, authors should use the word ‘‘randomised’’ in the title to indicate that the participants were randomly assigned to their comparison groups.

5.1.2. Item 1b. Structured summary of trial design, methods, results, and conclusions

For specific guidance see CONSORT for abstracts [45,65].

Explanation—Clear, transparent, and sufficiently detailed abstracts are important because readers often base their assessment of a trial on such information. Some readers use an abstract as a screening tool to decide whether to read the full article. However, as not all trials are freely available and some health professionals do not have access to the full trial reports, healthcare decisions are sometimes made on the basis of abstracts of randomised trials [66].

A journal abstract should contain sufficient information about a trial to serve as an accurate record of its conduct and findings, providing optimal information about the trial within the space constraints and format of a journal. A properly constructed and written abstract helps individuals to assess quickly the relevance of the findings and aids the retrieval of relevant reports from electronic databases [67].

Table 1
CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts [45,65])	
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	
Recruitment	14a	Dates defining the periods of recruitment and follow-up	
	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms [42])	
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	

(*Continued*)

Table 1
Continued

Section/Topic	Item No	Checklist item	Reported on page No
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

* We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials [40], non-inferiority and equivalence trials [39], non-pharmacological treatments [43], herbal interventions [44], and pragmatic trials [41]. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

The abstract should accurately reflect what is included in the full journal article and should not include information that does not appear in the body of the paper. Studies comparing the accuracy of information reported in a journal abstract with that reported in the text of the full publication have found claims that are inconsistent with, or missing from, the body of the full article [68−71]. Conversely, omitting important harms from the abstract could seriously mislead someone's interpretation of the trial findings [42,72].

A recent extension to the CONSORT statement provides a list of essential items that authors should include when reporting the main results of a randomised trial in a journal (or conference) abstract (see Table 2) [45]. We strongly recommend the use of structured abstracts for reporting randomised trials. They provide readers with information about the trial under a series of headings pertaining to the design, conduct, analysis, and interpretation [73]. Some studies have found that structured abstracts are of higher quality than the more traditional descriptive abstracts [74,75] and that they allow readers to find information more easily [76]. We recognise that many journals have developed their own structure and word limit for reporting abstracts. It is not our intention to suggest changes to these formats, but to recommend what information should be reported.

5.2. Introduction

5.2.1. Item 2a. Scientific background and explanation of rationale

Example—"Surgery is the treatment of choice for patients with disease stage I and II non-small cell lung cancer (NSCLC) … An NSCLC meta-analysis combined the results from eight randomised trials of surgery versus surgery plus adjuvant cisplatin-based chemotherapy and showed a small, but not significant (p=0.08), absolute survival benefit of around 5% at 5 years (from 50% to 55%). At the time the current trial was designed (mid-1990s), adjuvant chemotherapy had not become standard clinical practice … The clinical rationale for neo-adjuvant chemotherapy is three-fold: regression of the primary cancer could be achieved thereby facilitating and simplifying or reducing subsequent surgery; undetected micro-metastases could be dealt with at the start of treatment; and there might be inhibition of the putative stimulus to residual cancer by growth factors released by surgery and by subsequent wound healing … The current trial was therefore set up to compare, in patients with resectable NSCLC, surgery alone versus three cycles of platinum-based chemotherapy followed by surgery in terms of overall survival, quality of life, pathological staging, resectability rates, extent of surgery, and time to and site of relapse"[77].

Explanation—Typically, the introduction consists of free flowing text, in which authors explain the scientific background and rationale for their trial, and its general outline. It may also be appropriate to include here the objectives of the trial (see item 2b). The rationale may be explanatory (for example, to assess the possible influence of a drug on renal function) or pragmatic (for example, to guide practice by comparing the benefits and harms of two treatments). Authors should report any evidence of the benefits and harms of active interventions included in a trial and should suggest a plausible explanation for how the interventions might work, if this is not obvious [78].

The Declaration of Helsinki states that biomedical research involving people should be based on a thorough knowledge of the scientific literature [79]. That is, it is unethical to expose humans unnecessarily to the risks of research. Some clinical trials have been shown to have been unnecessary because the question they addressed had been or could have been answered by a systematic review of the existing literature [80,81]. Thus, the need for a new trial should be justified in the introduction. Ideally, it should include a reference to a systematic review of previous similar trials or a note of the absence of such trials [82].

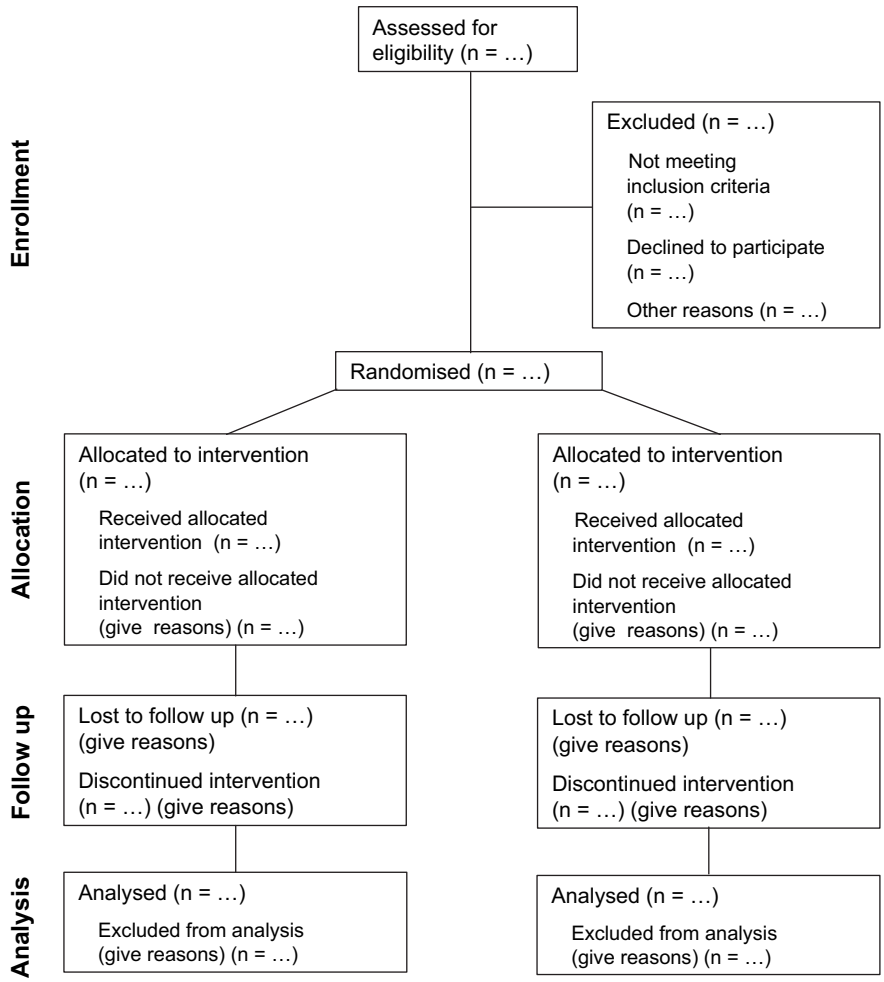

Fig. 1. Flow diagram of the progress through the phases of a parallel randomised trial of two groups (that is, enrolment, intervention allocation, follow-up, and data analysis) [52–54].

5.2.2. *Item 2b. Specific objectives or hypotheses*

Example—"In the current study we tested the hypothesis that a policy of active management of nulliparous labour would: 1. reduce the rate of caesarean section, 2. reduce the rate of prolonged labour; 3. not influence maternal satisfaction with the birth experience"[83].

Explanation—Objectives are the questions that the trial was designed to answer. They often relate to the efficacy of a particular therapeutic or preventive intervention. Hypotheses are pre-specified questions being tested to help meet the objectives. Hypotheses are more specific than objectives and are amenable to explicit statistical evaluation. In practice, objectives and hypotheses are not always easily differentiated. Most reports of RCTs provide adequate information about trial objectives and hypotheses [84].

5.3. *Methods*

5.3.1. *Item 3a. Description of trial design (such as parallel, factorial) including allocation ratio*

Example—"This was a multicenter, stratified (6 to 11 years and 12 to 17 years of age, with imbalanced randomisation [2:1]), double-blind, placebo-controlled, parallel-group study conducted in the United States (41 sites)"[85].

Explanation—The word "design" is often used to refer to all aspects of how a trial is set up, but it also has a narrower interpretation. Many specific aspects of the broader trial design, including details of randomisation and blinding, are addressed elsewhere in the CONSORT checklist. Here we seek information on the type of trial, such as parallel group or factorial, and the conceptual framework, such as superiority or non-inferiority, and other related issues not addressed elsewhere in the checklist.

The CONSORT statement focuses mainly on trials with participants individually randomised to one of two "parallel" groups. In fact, little more than half of published trials have such a design [16]. The main alternative designs are multi-arm parallel, crossover, cluster [40], and factorial designs. Also, most trials are set to identify the superiority of a new intervention, if it exists, but others are designed to assess non-inferiority or equivalence [39]. It is important that researchers clearly describe these aspects of their trial, including the unit of randomisation (such as patient, GP

Table 2
Items to include when reporting a randomised trial in a journal abstract

Item	Description
Authors	Contact details for the corresponding author
Trial design	Description of the trial design (such as parallel, cluster, non-inferiority)
Methods:	
Participants	Eligibility criteria for participants and the settings where the data were collected
Interventions	Interventions intended for each group
Objective	Specific objective or hypothesis
Outcome	Clearly defined primary outcome for this report
Randomisation	How participants were allocated to interventions
Blinding (masking)	Whether participants, care givers, and those assessing the outcomes were blinded to group assignment
Results:	
Numbers randomised	Number of participants randomised to each group
Recruitment	Trial status
Numbers analysed	Number of participants analysed in each group
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision
Harms	Important adverse events or side effects
Conclusions	General interpretation of the results
Trial registration	Registration number and name of trial register
Funding	Source of funding

practice, lesion). It is desirable also to include these details in the abstract (see item 1b).

If a less common design is employed, authors are encouraged to explain their choice, especially as such designs may imply the need for a larger sample size or more complex analysis and interpretation.

Although most trials use equal randomisation (such as 1:1 for two groups), it is helpful to provide the allocation ratio explicitly. For drug trials, specifying the phase of the trial (I-IV) may also be relevant.

5.3.2. *Item 3b. Important changes to methods after trial commencement (such as eligibility criteria), with reasons*

Example—"Patients were randomly assigned to one of six parallel groups, initially in 1:1:1:1:1:1 ratio, to receive either one of five otamixaban … regimens … or an active control of unfractionated heparin … an independent Data Monitoring Committee reviewed unblinded data for patient safety; no interim analyses for efficacy or futility were done. During the trial, this committee recommended that the group receiving the lowest dose of otamixaban (0·035 mg/kg/h) be discontinued because of clinical evidence of inadequate anticoagulation. The protocol was immediately amended in accordance with that recommendation, and participants were subsequently randomly assigned in 2:2:2:2:1 ratio to the remaining otamixaban and control groups, respectively"[86].

Explanation—A few trials may start without any fixed plan (that is, are entirely exploratory), but the most will have a protocol that specifies in great detail how the trial will be conducted. There may be deviations from the original protocol, as it is impossible to predict every possible change in circumstances during the course of a trial. Some trials will therefore have important changes to the methods after trial commencement.

Changes could be due to external information becoming available from other studies, or internal financial difficulties, or could be due to a disappointing recruitment rate. Such protocol changes should be made without breaking the blinding on the accumulating data on participants' outcomes. In some trials, an independent data monitoring committee will have as part of its remit the possibility of recommending protocol changes based on seeing unblinded data. Such changes might affect the study methods (such as changes to treatment regimens, eligibility criteria, randomisation ratio, or duration of follow-up) or trial conduct (such as dropping a centre with poor data quality) [87].

Some trials are set up with a formal "adaptive" design. There is no universally accepted definition of these designs, but a working definition might be "a multistage study design that uses accumulating data to decide how to modify aspects of the study without undermining the validity and integrity of the trial"[88]. The modifications are usually to the sample sizes and the number of treatment arms and can lead to decisions being made more quickly and with more efficient use of resources. There are, however, important ethical, statistical, and practical issues in considering such a design [89,90].

Whether the modifications are explicitly part of the trial design or in response to changing circumstances, it is essential that they are fully reported to help the reader interpret the results. Changes from protocols are not currently well reported. A review of comparisons with protocols showed that about half of journal articles describing RCTs had an unexplained discrepancy in the primary outcomes [57]. Frequent unexplained discrepancies have also been observed for details of randomisation, blinding [91], and statistical analyses [92].

5.3.3. Item 4a. Eligibility criteria for participants

Example—"Eligible participants were all adults aged 18 or over with HIV who met the eligibility criteria for antiretroviral therapy according to the Malawian national HIV treatment guidelines (WHO clinical stage III or IV or any WHO stage with a CD4 count $<250/mm^3$) and who were starting treatment with a BMI <18.5. Exclusion criteria were pregnancy and lactation or participation in another supplementary feeding programme"[93].

Explanation—A comprehensive description of the eligibility criteria used to select the trial participants is needed to help readers interpret the study. In particular, a clear understanding of these criteria is one of several elements required to judge to whom the results of a trial apply—that is, the trial's generalisability (applicability) and relevance to clinical or public health practice (see item 21) [94]. A description of the method of recruitment, such as by referral or self selection (for example, through advertisements), is also important in this context. Because they are applied before randomisation, eligibility criteria do not affect the internal validity of a trial, but they are central to its external validity.

Typical and widely accepted selection criteria relate to the nature and stage of the disease being studied, the exclusion of persons thought to be particularly vulnerable to harm from the study intervention, and to issues required to ensure that the study satisfies legal and ethical norms. Informed consent by study participants, for example, is typically required in intervention studies. The common distinction between inclusion and exclusion criteria is unnecessary; the same criterion can be phrased to include or exclude participants [95].

Despite their importance, eligibility criteria are often not reported adequately. For example, eight published trials leading to clinical alerts by the National Institutes of Health specified an average of 31 eligibility criteria in their protocols, but only 63% of the criteria were mentioned in the journal articles, and only 19% were mentioned in the clinical alerts [96]. Similar deficiencies were found for HIV clinical trials [97]. Among 364 reports of RCTs in surgery, 25% did not specify any eligibility criteria [98].

5.3.4. Item 4b. Settings and locations where the data were collected

Example—"The study took place at the antiretroviral therapy clinic of Queen Elizabeth Central Hospital in Blantyre, Malawi, from January 2006 to April 2007. Blantyre is the major commercial city of Malawi, with a population of 1 000 000 and an estimated HIV prevalence of 27% in adults in 2004"[93].

Explanation—Along with the eligibility criteria for participants (see item 4a) and the description of the interventions (see item 5), information on the settings and locations is crucial to judge the applicability and generalisability of a trial. Were participants recruited from primary, secondary, or tertiary health care or from the community? Healthcare institutions vary greatly in their organisation, experience, and resources and the baseline risk for the condition under investigation. Other aspects of the setting (including the social, economic, and cultural environment and the climate) may also affect a study's external validity.

Authors should report the number and type of settings and describe the care providers involved. They should report the locations in which the study was carried out, including the country, city if applicable, and immediate environment (for example, community, office practice, hospital clinic, or inpatient unit). In particular, it should be clear whether the trial was carried out in one or several centres ("multicentre trials"). This description should provide enough information so that readers can judge whether the results of the trial could be relevant to their own setting. The environment in which the trial is conducted may differ considerably from the setting in which the trial's results are later used to guide practice and policy [94,99]. Authors should also report any other information about the settings and locations that could have influenced the observed results, such as problems with transportation that might have affected patient participation or delays in administering interventions.

5.3.5. Item 5. The interventions for each group with sufficient details to allow replication, including how and when they were actually administered

Examples—"In POISE, patients received the first dose of the study drug (ie, oral extended-release metoprolol 100 mg or matching placebo) 2–4 h before surgery. Study drug administration required a heart rate of 50 bpm or more and a systolic blood pressure of 100 mm Hg or greater; these haemodynamics were checked before each administration. If, at any time during the first 6 h after surgery, heart rate was 80 bpm or more and systolic blood pressure was 100 mm Hg or higher, patients received their first postoperative dose (extended-release metoprolol 100 mg or matched placebo) orally. If the study drug was not given during the first 6 h, patients received their first postoperative dose at 6 h after surgery. 12 h after the first postoperative dose, patients started taking oral extended-release metoprolol 200 mg or placebo every day for 30 days. If a patient's heart rate was consistently below 45 bpm or their systolic blood pressure dropped below 100 mm Hg, study drug was withheld until their heart rate or systolic blood pressure recovered; the study drug was then restarted at 100 mg once daily. Patients whose heart rate was consistently 45–49 bpm and systolic blood pressure exceeded 100 mm Hg delayed taking the study drug for 12 h"[100].

"Patients were randomly assigned to receive a custom-made neoprene splint to be worn at night or to usual care. The splint was a rigid rest orthosis recommended for use only at night. It covered the base of the thumb and the thenar eminence but not the wrist (Figure 1). Splints were made by 3 trained occupational therapists, who adjusted the splint for each patient so that the first web could be

opened and the thumb placed in opposition with the first long finger. Patients were encouraged to contact the occupational therapist if they felt that the splint needed adjustment, pain increased while wearing the splint, or they had adverse effects (such as skin erosion). Because no treatment can be considered the gold standard in this situation, patients in the control and intervention groups received usual care at the discretion of their physician (general practitioner or rheumatologist). We decided not to use a placebo because, to our knowledge, no placebo for splinting has achieved successful blinding of patients, as recommended''[101].

Explanation—Authors should describe each intervention thoroughly, including control interventions. The description should allow a clinician wanting to use the intervention to know exactly how to administer the intervention that was evaluated in the trial [102]. For a drug intervention, information would include the drug name, dose, method of administration (such as oral, intravenous), timing and duration of administration, conditions under which interventions are withheld, and titration regimen if applicable. If the control group is to receive ''usual care'' it is important to describe thoroughly what that constitutes. If the control group or intervention group is to receive a combination of interventions the authors should provide a thorough description of each intervention, an explanation of the order in which the combination of interventions are introduced or withdrawn, and the triggers for their introduction if applicable.

Specific extensions of the CONSORT statement address the reporting of non-pharmacologic and herbal interventions and their particular reporting requirements (such as expertise, details of how the interventions were standardised) [43,44]. We recommend readers consult the statements for non-pharmacologic and herbal interventions as appropriate.

5.3.6. Item 6a. Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed

Example—''The primary endpoint with respect to efficacy in psoriasis was the proportion of patients achieving a 75% improvement in psoriasis activity from baseline to 12 weeks as measured by the PASI [psoriasis area and severity index] Additional analyses were done on the percentage change in PASI scores and improvement in target psoriasis lesions''[103].

Explanation—All RCTs assess response variables, or outcomes (end points), for which the groups are compared. Most trials have several outcomes, some of which are of more interest than others. The primary outcome measure is the pre-specified outcome considered to be of greatest importance to relevant stakeholders (such a patients, policy makers, clinicians, funders) and is usually the one used in the sample size calculation (see item 7). Some trials may have more than one primary outcome. Having several primary outcomes, however, incurs the problems of interpretation associated with multiplicity of analyses (see items 18 and 20) and is not recommended. Primary outcomes should be explicitly indicated as such in the report of an RCT. Other outcomes of interest are secondary outcomes (additional outcomes). There may be several secondary outcomes, which often include unanticipated or unintended effects of the intervention (see item 19), although harms should always be viewed as important whether they are labelled primary or secondary.

All outcome measures, whether primary or secondary, should be identified and completely defined. The principle here is that the information provided should be sufficient to allow others to use the same outcomes [102]. When outcomes are assessed at several time points after randomisation, authors should also indicate the pre-specified time point of primary interest. For many non-pharmacological interventions it is helpful to specify who assessed outcomes (for example, if special skills are required to do so) and how many assessors there were [43].

Where available and appropriate, the use of previously developed and validated scales or consensus guidelines should be reported [104,105], both to enhance quality of measurement and to assist in comparison with similar studies [106]. For example, assessment of quality of life is likely to be improved by using a validated instrument [107]. Authors should indicate the provenance and properties of scales.

More than 70 outcomes were used in 196 RCTs of non-steroidal anti-inflammatory drugs for rheumatoid arthritis [108], and 640 different instruments had been used in 2000 trials in schizophrenia, of which 369 had been used only once [33]. Investigation of 149 of those 2000 trials showed that unpublished scales were a source of bias. In non-pharmacological trials, a third of the claims of treatment superiority based on unpublished scales would not have been made if a published scale had been used [109]. Similar data have been reported elsewhere [110,111]. Only 45% of a cohort of 519 RCTs published in 2000 specified the primary outcome [16]; this compares with 53% for a similar cohort of 614 RCTs published in 2006 [17].

5.3.7. Item 6b. Any changes to trial outcomes after the trial commenced, with reasons

Example—''The original primary endpoint was all-cause mortality, but, during a masked analysis, the data and safety monitoring board noted that overall mortality was lower than had been predicted and that the study could not be completed with the sample size and power originally planned. The steering committee therefore decided to adopt co-primary endpoints of all-cause mortality (the original primary endpoint), together with all-cause mortality or cardiovascular hospital admissions (the first prespecified secondary endpoint)''[112].

Explanation—There are many reasons for departures from the initial study protocol (see item 24). Authors

should report all major changes to the protocol, including unplanned changes to eligibility criteria, interventions, examinations, data collection, methods of analysis, and outcomes. Such information is not always reported.

As indicated earlier (see item 6a), most trials record multiple outcomes, with the risk that results will be reported for only a selected subset (see item 17). Prespecification and reporting of primary and secondary outcomes (see item 6a) should remove such a risk. In some trials, however, circumstances require a change in the way an outcome is assessed or even, as in the example above, a switch to a different outcome. For example, there may be external evidence from other trials or systematic reviews suggesting the end point might not be appropriate, or recruitment or the overall event rate in the trial may be lower than expected [112]. Changing an end point based on unblinded data is much more problematic, although it may be specified in the context of an adaptive trial design [88]. Authors should identify and explain any such changes. Likewise, any changes after the trial began of the designation of outcomes as primary or secondary should be reported and explained.

A comparison of protocols and publications of 102 randomised trials found that 62% of trials reports had at least one primary outcome that was changed, introduced, or omitted compared with the protocol [55]. Primary outcomes also differed between protocols and publications for 40% of a cohort of 48 trials funded by the Canadian Institutes of Health Research [113]. Not one of the subsequent 150 trial reports mentioned, let alone explained, changes from the protocol. Similar results from other studies have been reported recently in a systematic review of empirical studies examining outcome reporting bias [57].

5.3.8. Item 7a. How sample size was determined

Examples—"To detect a reduction in PHS (postoperative hospital stay) of 3 days (SD 5 days), which is in agreement with the study of Lobo et al [17] with a two-sided 5% significance level and a power of 80%, a sample size of 50 patients per group was necessary, given an anticipated dropout rate of 10%. To recruit this number of patients a 12-month inclusion period was anticipated"[114].

"Based on an expected incidence of the primary composite endpoint of 11% at 2.25 years in the placebo group, we calculated that we would need 950 primary endpoint events and a sample size of 9650 patients to give 90% power to detect a significant difference between ivabradine and placebo, corresponding to a 19% reduction of relative risk (with a two-sided type 1 error of 5%). We initially designed an event-driven trial, and planned to stop when 950 primary endpoint events had occurred. However, the incidence of the primary endpoint was higher than predicted, perhaps because of baseline characteristics of the recruited patients, who had higher risk than expected (e.g., lower proportion of NYHA class I and higher rates of diabetes and hypertension). We calculated that when 950 primary endpoint events had occurred, the most recently included patients would only have been treated for about 3 months. Therefore, in January 2007, the executive committee decided to change the study from being event-driven to time-driven, and to continue the study until the patients who were randomised last had been followed up for 12 months. This change did not alter the planned study duration of 3 years"[115].

Explanation—For scientific and ethical reasons, the sample size for a trial needs to be planned carefully, with a balance between medical and statistical considerations. Ideally, a study should be large enough to have a high probability (power) of detecting as statistically significant a clinically important difference of a given size if such a difference exists. The size of effect deemed important is inversely related to the sample size necessary to detect it; that is, large samples are necessary to detect small differences. Elements of the sample size calculation are (1) the estimated outcomes in each group (which implies the clinically important target difference between the intervention groups); (2) the α (type I) error level; (3) the statistical power (or the β (type II) error level); and (4), for continuous outcomes, the standard deviation of the measurements [116]. The interplay of these elements and their reporting will differ for cluster trials [40] and non-inferiority and equivalence trials [39].

Authors should indicate how the sample size was determined. If a formal power calculation was used, the authors should identify the primary outcome on which the calculation was based (see item 6a), all the quantities used in the calculation, and the resulting target sample size per study group. It is preferable to quote the expected result in the control group and the difference between the groups one would not like to overlook. Alternatively, authors could present the percentage with the event or mean for each group used in their calculations. Details should be given of any allowance made for attrition or non-compliance during the study.

Some methodologists have written that so called underpowered trials may be acceptable because they could ultimately be combined in a systematic review and meta-analysis [117−119], and because some information is better than no information. Of note, important caveats apply—such as the trial should be unbiased, reported properly, and published irrespective of the results, thereby becoming available for meta-analysis [118]. On the other hand, many medical researchers worry that underpowered trials with indeterminate results will remain unpublished and insist that all trials should individually have "sufficient power." This debate will continue, and members of the CONSORT Group have varying views. Critically however, the debate and those views are immaterial to reporting a trial. Whatever the power of a trial, authors need to properly report their intended size with all their methods and assumptions [118]. That transparently reveals the power of

the trial to readers and gives them a measure by which to assess whether the trial attained its planned size.

In some trials, interim analyses are used to help decide whether to stop early or to continue recruiting sometimes beyond the planned trial end (see item 7b). If the actual sample size differed from the originally intended sample size for some other reason (for example, because of poor recruitment or revision of the target sample size), the explanation should be given.

Reports of studies with small samples frequently include the erroneous conclusion that the intervention groups do not differ, when in fact too few patients were studied to make such a claim [120]. Reviews of published trials have consistently found that a high proportion of trials have low power to detect clinically meaningful treatment effects [121−123]. In reality, small but clinically meaningful true differences are much more likely than large differences to exist, but large trials are required to detect them [124].

In general, the reported sample sizes in trials seem small. The median sample size was 54 patients in 196 trials in arthritis [108], 46 patients in 73 trials in dermatology [8], and 65 patients in 2000 trials in schizophrenia [33]. These small sample sizes are consistent with those of a study of 519 trials indexed in PubMed in December 2000 [16] and a similar cohort of trials (n=616) indexed in PubMed in 2006 [17], where the median number of patients recruited for parallel group trials was 80 across both years. Moreover, many reviews have found that few authors report how they determined the sample size [8,14,32,33,123].

There is little merit in a post hoc calculation of statistical power using the results of a trial; the power is then appropriately indicated by confidence intervals (see item 17) [125].

5.3.9. Item 7b. When applicable, explanation of any interim analyses and stopping guidelines

Examples—"Two interim analyses were performed during the trial. The levels of significance maintained an overall P value of 0.05 and were calculated according to the O'Brien-Fleming stopping boundaries. This final analysis used a Z score of 1.985 with an associated P value of 0.0471"[126].

"An independent data and safety monitoring board periodically reviewed the efficacy and safety data. Stopping rules were based on modified Haybittle-Peto boundaries of 4 SD in the first half of the study and 3 SD in the second half for efficacy data, and 3 SD in the first half of the study and 2 SD in the second half for safety data. Two formal interim analyses of efficacy were performed when 50% and 75% of the expected number of primary events had accrued; no correction of the reported P value for these interim tests was performed"[127].

Explanation—Many trials recruit participants over a long period. If an intervention is working particularly well or badly, the study may need to be ended early for ethical reasons. This concern can be addressed by examining results as the data accumulate, preferably by an independent data monitoring committee. However, performing multiple statistical examinations of accumulating data without appropriate correction can lead to erroneous results and interpretations [128]. If the accumulating data from a trial are examined at five interim analyses that use a P value of 0.05, the overall false positive rate is nearer to 19% than to the nominal 5%.

Several group sequential statistical methods are available to adjust for multiple analyses [129−131], and their use should be pre-specified in the trial protocol. With these methods, data are compared at each interim analysis, and a P value less than the critical value specified by the group sequential method indicates statistical significance. Some trialists use group sequential methods as an aid to decision making [132], whereas others treat them as a formal stopping rule (with the intention that the trial will cease if the observed P value is smaller than the critical value).

Authors should report whether they or a data monitoring committee took multiple "looks" at the data and, if so, how many there were, what triggered them, the statistical methods used (including any formal stopping rule), and whether they were planned before the start of the trial, before the data monitoring committee saw any interim data by allocation, or some time thereafter. This information is often not included in published trial reports [133], even in trials that report stopping earlier than planned [134].

5.3.10. Item 8a. Method used to generate the random allocation sequence

Examples—"Independent pharmacists dispensed either active or placebo inhalers according to a computer generated randomisation list"[63].

"For allocation of the participants, a computer-generated list of random numbers was used"[135].

Explanation—Participants should be assigned to comparison groups in the trial on the basis of a chance (random) process characterised by unpredictability (see Box 1). Authors should provide sufficient information that the reader can assess the methods used to generate the random allocation sequence and the likelihood of bias in group assignment. It is important that information on the process of randomisation is included in the body of the main article and not as a separate supplementary file; where it can be missed by the reader.

The term "random" has a precise technical meaning. With random allocation, each participant has a known probability of receiving each intervention before one is assigned, but the assigned intervention is determined by a chance process and cannot be predicted. However, "random" is often used inappropriately in the literature to describe trials in which non-random, deterministic allocation methods were used, such as alternation, hospital numbers, or date of birth. When investigators use such non-random methods, they should describe them precisely and should not use the term "random" or any variation of it.

Even the term ''quasi-random'' is unacceptable for describing such trials. Trials based on non-random methods generally yield biased results [2–4,136] Bias presumably arises from the inability to conceal these allocation systems adequately (see item 9).

Many methods of sequence generation are adequate. However, readers cannot judge adequacy from such terms as ''random allocation,'' ''randomisation,'' or ''random'' without further elaboration. Authors should specify the method of sequence generation, such as a random-number table or a computerised random number generator. The sequence may be generated by the process of minimisation, a non-random but generally acceptable method (see Box 2).

In some trials, participants are intentionally allocated in unequal numbers to each intervention: for example, to gain more experience with a new procedure or to limit costs of the trial. In such cases, authors should report the randomisation ratio (for example, 2:1 or two treatment participants per each control participant) (see item 3a).

In a representative sample of PubMed indexed trials in 2000, only 21% reported an adequate approach to random sequence generation [16]; this increased to 34% for a similar cohort of PubMed indexed trials in 2006 [17]. In more than 90% of these cases, researchers used a random number generator on a computer or a random number table.

5.3.11. Item 8b. Type of randomisation; details of any restriction (such as blocking and block size)

Examples—''Randomization sequence was created using Stata 9.0 (StataCorp, College Station, TX) statistical software and was stratified by center with a 1:1 allocation using random block sizes of 2, 4, and 6''[137].

''Participants were randomly assigned following simple randomization procedures (computerized random numbers) to 1 of 2 treatment groups''[138].

Explanation—In trials of several hundred participants or more simple randomisation can usually be trusted to generate similar numbers in the two trial groups [139] and to generate groups that are roughly comparable in terms of known and unknown prognostic variables [140]. For smaller trials (see item 7a)—and even for trials that are not intended to be small, as they may stop before reaching their target size—some restricted randomisation (procedures to help achieve balance between groups in size or characteristics) may be useful (see Box 2).

It is important to indicate whether no restriction was used, by stating such or by stating that ''simple randomisation'' was done. Otherwise, the methods used to restrict the randomisation, along with the method used for random selection, should be specified. For block randomisation, authors should provide details on how the blocks were generated (for example, by using a permuted block design with a computer random number generator), the block size or sizes, and whether the block size was fixed or randomly varied. If the trialists became aware of the block size(s), that information should also be reported as such knowledge could lead to code breaking. Authors should specify whether stratification was used, and if so, which factors were involved (such as recruitment site, sex, disease stage), the categorisation cut-off values within strata, and the method used for restriction. Although stratification is a useful technique, especially for smaller trials, it is complicated to implement and may be impossible if many stratifying factors are used. If minimisation (see Box 2) was used, it should be explicitly identified, as should the variables incorporated into the scheme. If used, a random element should be indicated.

Only 9% of 206 reports of trials in specialty journals [23] and 39% of 80 trials in general medical journals reported use of stratification [32]. In each case, only about half of the reports mentioned the use of restricted randomisation. However, these studies and that of Adetugbo and Williams [8] found that the sizes of the treatment groups in many trials were the same or quite similar, yet blocking or stratification had not been mentioned. One possible explanation for the close balance in numbers is underreporting of the use of restricted randomisation.

5.3.12. Item 9. Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned

Examples—''The doxycycline and placebo were in capsule form and identical in appearance. They were prepacked in bottles and consecutively numbered for each woman according to the randomisation schedule. Each woman was assigned an order number and received the capsules in the corresponding prepacked bottle''[146].

''The allocation sequence was concealed from the researcher (JR) enrolling and assessing participants in sequentially numbered, opaque, sealed and stapled envelopes. Aluminium foil inside the envelope was used to render the envelope impermeable to intense light. To prevent subversion of the allocation sequence, the name and date of birth of the participant was written on the envelope and a video tape made of the sealed envelope with participant details visible. Carbon paper inside the envelope transferred the information onto the allocation card inside the envelope and a second researcher (CC) later viewed video tapes to ensure envelopes were still sealed when participants' names were written on them. Corresponding envelopes were opened only after the enrolled participants completed all baseline assessments and it was time to allocate the intervention''[147].

Explanation—Item 8a discussed generation of an unpredictable sequence of assignments. Of considerable importance is how this sequence is applied when participants are enrolled into the trial (see Box 1). A generated allocation schedule should be implemented by using allocation concealment [23], a critical mechanism that prevents foreknowledge of treatment assignment and thus shields those who enroll participants from being influenced by this

Box 2. Randomisation and minimisation

Simple randomisation—Pure randomisation based on a single allocation ratio is known as simple randomisation. Simple randomisation with a 1:1 allocation ratio is analogous to a coin toss, although we do not advocate coin tossing for randomisation in an RCT. ''Simple'' is somewhat of a misnomer. While other randomisation schemes sound complex and more sophisticated, in reality, simple randomisation is elegantly sophisticated in that it is more unpredictable and surpasses the bias prevention levels of all other alternatives.

Restricted randomisation—Any randomised approach that is not simple randomisation. Blocked randomisation is the most common form. Other means of restricted randomisation include replacement, biased coin, and urn randomisation, although these are used much less frequently [141].

Blocked randomisation—Blocking is used to ensure that comparison groups will be generated according to a predetermined ratio, usually 1:1 or groups of approximately the same size. Blocking can be used to ensure close balance of the numbers in each group at any time during the trial. For every block of eight participants, for example, four would be allocated to each arm of the trial [142]. Improved balance comes at the cost of reducing the unpredictability of the sequence. Although the order of interventions varies randomly within each block, a person running the trial could deduce some of the next treatment allocations if he or she knew the block size [143]. Blinding the interventions, using larger block sizes, and randomly varying the block size can ameliorate this problem.

Stratified randomisation—Stratification is used to ensure good balance of participant characteristics in each group. By chance, particularly in small trials, study groups may not be well matched for baseline characteristics, such as age and stage of disease. This weakens the trial's credibility [144]. Such imbalances can be avoided without sacrificing the advantages of randomisation. Stratification ensures that the numbers of participants receiving each intervention are closely balanced within each stratum. Stratified randomisation is achieved by performing a separate randomisation procedure within each of two or more subsets of participants (for example, those defining each study centre, age, or disease severity). Stratification by centre is common in multicentre trials. Stratification requires some form of restriction (such as blocking within strata). Stratification without blocking is ineffective.

Minimisation—Minimisation ensures balance between intervention groups for several selected patient factors (such as age) [22,60]. The first patient is truly randomly allocated; for each subsequent participant, the treatment allocation that minimises the imbalance on the selected factors between groups at that time is identified. That allocation may then be used, or a choice may be made at random with a heavy weighting in favour of the intervention that would minimise imbalance (for example, with a probability of 0.8). The use of a random component is generally preferable. Minimisation has the advantage of making small groups closely similar in terms of participant characteristics at all stages of the trial. Minimisation offers the only acceptable alternative to randomisation, and some have argued that it is superior [145]. On the other hand, minimisation lacks the theoretical basis for eliminating bias on all known and unknown factors. Nevertheless, in general, trials that use minimisation are considered methodologically equivalent to randomised trials, even when a random element is not incorporated.

knowledge. The decision to accept or reject a participant should be made, and informed consent should be obtained from the participant, in ignorance of the next assignment in the sequence [148].

The allocation concealment should not be confused with blinding (see item 11). Allocation concealment seeks to prevent selection bias, protects the assignment sequence until allocation, and can always be successfully implemented [2]. In contrast, blinding seeks to prevent performance and ascertainment bias, protects the sequence after allocation, and cannot always be implemented [23]. Without adequate allocation concealment, however, even random, unpredictable assignment sequences can be subverted [2,149].

Centralised or ''third-party'' assignment is especially desirable. Many good allocation concealment mechanisms incorporate external involvement. Use of a pharmacy or central telephone randomisation system are two common techniques. Automated assignment systems are likely to become more common [150]. When external involvement is not feasible, an excellent method of allocation concealment is the use of numbered containers. The interventions (often drugs) are sealed in sequentially numbered identical containers according to the allocation sequence [151]. Enclosing assignments in sequentially numbered, opaque, sealed envelopes can be a good allocation concealment mechanism if it is developed and monitored diligently. This method can be corrupted, however, particularly if it is poorly executed. Investigators should ensure that the envelopes are opaque when held to the light, and opened sequentially and only after the participant's name and other details are written on the appropriate envelope [143].

A number of methodological studies provide empirical evidence to support these precautions [152,153]. Trials in

which the allocation sequence had been inadequately or unclearly concealed yielded larger estimates of treatment effects than did trials in which authors reported adequate allocation concealment. These findings provide strong empirical evidence that inadequate allocation concealment contributes to bias in estimating treatment effects.

Despite the importance of the mechanism of allocation concealment, published reports often omit such details. The mechanism used to allocate interventions was omitted in reports of 89% of trials in rheumatoid arthritis [108], 48% of trials in obstetrics and gynaecology journals [23], and 44% of trials in general medical journals [32]. In a more broadly representative sample of all randomised trials indexed on PubMed, only 18% reported any allocation concealment mechanism, but some of those reported mechanisms were inadequate [16].

5.3.13. Item 10. Who generated the allocation sequence, who enrolled participants, and who assigned participants to interventions

Examples—"Determination of whether a patient would be treated by streptomycin and bed-rest (S case) or by bed-rest alone (C case) was made by reference to a statistical series based on random sampling numbers drawn up for each sex at each centre by Professor Bradford Hill; the details of the series were unknown to any of the investigators or to the co-ordinator … After acceptance of a patient by the panel, and before admission to the streptomycin centre, the appropriate numbered envelope was opened at the central office; the card inside told if the patient was to be an S or a C case, and this information was then given to the medical officer of the centre"[24].

"Details of the allocated group were given on coloured cards contained in sequentially numbered, opaque, sealed envelopes. These were prepared at the NPEU and kept in an agreed location on each ward. Randomisation took place at the end of the 2nd stage of labour when the midwife considered a vaginal birth was imminent. To enter a women into the study, the midwife opened the next consecutively numbered envelope"[154].

"Block randomisation was by a computer generated random number list prepared by an investigator with no clinical involvement in the trial. We stratified by admission for an oncology related procedure. After the research nurse had obtained the patient's consent, she telephoned a contact who was independent of the recruitment process for allocation consignment"[155].

Explanation—As noted in item 9, concealment of the allocated intervention at the time of enrolment is especially important. Thus, in addition to knowing the methods used, it is also important to understand how the random sequence was implemented—specifically, who generated the allocation sequence, who enrolled participants, and who assigned participants to trial groups.

The process of randomising participants into a trial has three different steps: sequence generation, allocation concealment, and implementation (see Box 3). Although the same people may carry out more than one process under each heading, investigators should strive for complete separation of the people involved with generation and allocation concealment from the people involved in the implementation of assignments. Thus, if someone is involved in the sequence generation or allocation concealment steps, ideally they should not be involved in the implementation step.

Even with flawless sequence generation and allocation concealment, failure to separate creation and concealment of the allocation sequence from assignment to study group may introduce bias. For example, the person who generated an allocation sequence could retain a copy and consult it when interviewing potential participants for a trial. Thus, that person could bias the enrolment or assignment process, regardless of the unpredictability of the assignment sequence. Investigators must then ensure that the assignment schedule is unpredictable and locked away (such as in a safe deposit box in a building rather inaccessible to the enrolment location) from even the person who generated it. The report of the trial should specify where the investigators stored the allocation list.

5.3.14. Item 11a. If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how

Examples—"Whereas patients and physicians allocated to the intervention group were aware of the allocated arm, outcome assessors and data analysts were kept blinded to the allocation"[156].

"Blinding and equipoise were strictly maintained by emphasising to intervention staff and participants that each diet adheres to healthy principles, and each is advocated by certain experts to be superior for long-term weight-loss. Except for the interventionists (dieticians and behavioural psychologists), investigators and staff were kept blind to diet assignment of the participants. The trial adhered to established procedures to maintain separation between staff that take outcome measurements and staff that deliver the intervention. Staff members who obtained outcome measurements were not informed of the diet group assignment. Intervention staff, dieticians and behavioural psychologists who delivered the intervention did not take outcome measurements. All investigators, staff, and participants were kept masked to outcome measurements and trial results"[157].

Explanation—The term "blinding" or "masking" refers to withholding information about the assigned interventions from people involved in the trial who may potentially be influenced by this knowledge. Blinding is an important safeguard against bias, particularly when assessing subjective outcomes [153].

Benjamin Franklin has been credited as being the first to use blinding in a scientific experiment [158]. He blindfolded participants so they would not know when he was

Box 3. Steps in a typical randomisation process

Sequence generation
- Generate allocation sequence by some random procedure

Allocation concealment
- Develop allocation concealment mechanism (such as numbered, identical bottles or sequentially numbered, sealed, opaque envelopes)
- Prepare the allocation concealment mechanism using the allocation sequence from the sequence generation step

Implementation
- Enrol participants:

  Assess eligibility

  Discuss the trial

  Obtain informed consent

  Enrol participant in trial
- Ascertain intervention assignment (such as opening next envelope)
- Administer intervention

applying mesmerism (a popular "healing fluid" of the 18th century) and in so doing showed that mesmerism was a sham. Based on this experiment, the scientific community recognised the power of blinding to reduce bias, and it has remained a commonly used strategy in scientific experiments.

Box 4, on blinding terminology, defines the groups of individuals (that is, participants, healthcare providers, data collectors, outcome adjudicators, and data analysts) who can potentially introduce bias into a trial through knowledge of the treatment assignments. Participants may respond differently if they are aware of their treatment assignment (such as responding more favourably when they receive the new treatment) [153]. Lack of blinding may also influence compliance with the intervention, use of co-interventions, and risk of dropping out of the trial.

Unblinded healthcare providers may introduce similar biases, and unblinded data collectors may differentially assess outcomes (such as frequency or timing), repeat measurements of abnormal findings, or provide encouragement during performance testing. Unblinded outcome adjudicators may differentially assess subjective outcomes, and unblinded data analysts may introduce bias through the choice of analytical strategies, such as the selection of favourable time points or outcomes, and by decisions to remove patients from the analyses. These biases have been well documented [71,153,159–162].

Blinding, unlike allocation concealment (see item 10), may not always be appropriate or possible. An example is a trial comparing levels of pain associated with sampling blood from the ear or thumb [163]. Blinding is particularly important when outcome measures involve some subjectivity, such as assessment of pain. Blinding of data collectors and outcome adjudicators is unlikely to matter for objective outcomes, such as death from any cause. Even then, however, lack of participant or healthcare provider blinding can lead to other problems, such as differential attrition [164]. In certain trials, especially surgical trials, blinding of participants and surgeons is often difficult or impossible, but blinding of data collectors and outcome adjudicators is often achievable. For example, lesions can be photographed before and after treatment and assessed by an external observer [165]. Regardless of whether blinding is possible, authors can and should always state who was blinded (that is, participants, healthcare providers, data collectors, and outcome adjudicators).

Unfortunately, authors often do not report whether blinding was used [166]. For example, reports of 51% of 506 trials in cystic fibrosis [167], 33% of 196 trials in rheumatoid arthritis [108], and 38% of 68 trials in dermatology [8] did not state whether blinding was used. Until authors of trials improve their reporting of blinding, readers will have difficulty in judging the validity of the trials that they may wish to use to guide their clinical practice.

The term masking is sometimes used in preference to blinding to avoid confusion with the medical condition of being without sight. However, "blinding" in its methodological sense seems to be understood worldwide and is acceptable for reporting clinical trials [165,168]

Box 4. Blinding terminology

In order for a technical term to have utility it must have consistency in its use and interpretation. Authors of trials commonly use the term ''double blind'' and, less commonly, the terms ''single blind''or ''triple blind.'' A problem with this lexicon is that there is great variability in clinician interpretations and epidemiological textbook definitions of these terms [169]. Moreover, a study of 200 RCTs reported as double blind found 18 different combinations of groups actually blinded when the authors of these trials were surveyed, and about one in every five of these trials—reported as double blind—did not blind participants, healthcare providers, or data collectors [170].

This research shows that terms are ambiguous and, as such, authors and editors should abandon their use. Authors should instead explicitly report the blinding status of the people involved for whom blinding may influence the validity of a trial.

Healthcare providers include all personnel (for example, physicians, chiropractors, physiotherapists, nurses) who care for the participants during the trial. Data collectors are the individuals who collect data on the trial outcomes. Outcome adjudicators are the individuals who determine whether a participant did experience the outcomes of interest.

Some researchers have also advocated blinding and reporting the blinding status of the data monitoring committee and the manuscript writers [160]. Blinding of these groups is uncommon, and the value of blinding them is debated [171].

Sometimes one group of individuals (such as the healthcare providers) are the same individuals fulfilling another role in a trial (such as data collectors). Even if this is the case, the authors should explicitly state the blinding status of these groups to allow readers to judge the validity of the trial.

5.3.15. Item 11b. If relevant, description of the similarity of interventions

Example—''Jamieson Laboratories Inc provided 500-mg immediate release niacin in a white, oblong, bisect caplet. We independently confirmed caplet content using high performance liquid chromatography … The placebo was matched to the study drug for taste, color, and size, and contained microcrystalline cellulose, silicon dioxide, dicalcium phosphate, magnesium stearate, and stearic acid''[172].

Explanation—Just as we seek evidence of concealment to assure us that assignment was truly random, we seek evidence of the method of blinding. In trials with blinding of participants or healthcare providers, authors should state the similarity of the characteristics of the interventions (such as appearance, taste, smell, and method of administration) [35,173]

Some people have advocated testing for blinding by asking participants or healthcare providers at the end of a trial whether they think the participant received the experimental or control intervention [174]. Because participants and healthcare providers will usually know whether the participant has experienced the primary outcome, this makes it difficult to determine if their responses reflect failure of blinding or accurate assumptions about the efficacy of the intervention [175]. Given the uncertainty this type of information provides, we have removed advocating reporting this type of testing for blinding from the CONSORT 2010 Statement. We do, however, advocate that the authors report any known compromises in blinding. For example, authors should report if it was necessary to unblind any participants at any point during the conduct of a trial.

5.3.16. Item 12a. Statistical methods used to compare groups for primary and secondary outcomes

Example—''The primary endpoint was change in bodyweight during the 20 weeks of the study in the intention-to-treat population … Secondary efficacy endpoints included change in waist circumference, systolic and diastolic blood pressure, prevalence of metabolic syndrome … We used an analysis of covariance (ANCOVA) for the primary endpoint and for secondary endpoints waist circumference, blood pressure, and patient-reported outcome scores; this was supplemented by a repeated measures analysis. The ANCOVA model included treatment, country, and sex as fixed effects, and bodyweight at randomisation as covariate. We aimed to assess whether data provided evidence of superiority of each liraglutide dose to placebo (primary objective) and to orlistat (secondary objective)''[176].

Explanation—Data can be analysed in many ways, some of which may not be strictly appropriate in a particular situation. It is essential to specify which statistical procedure was used for each analysis, and further clarification may be necessary in the results section of the report. The principle to follow is to, ''Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results'' (www.icmje.org). It is also important to describe details of the statistical analysis such as intention-to-treat analysis (see Box 6).

Almost all methods of analysis yield an estimate of the treatment effect, which is a contrast between the outcomes in the comparison groups. Authors should accompany this by a confidence interval for the estimated effect, which indicates a central range of uncertainty for the true treatment effect. The confidence interval may be interpreted as the

range of values for the treatment effect that is compatible with the observed data. It is customary to present a 95% confidence interval, which gives the range expected to include the true value in 95 of 100 similar studies.

Study findings can also be assessed in terms of their statistical significance. The P value represents the probability that the observed data (or a more extreme result) could have arisen by chance when the interventions did not truly differ. Actual P values (for example, P=0.003) are strongly preferable to imprecise threshold reports such as P<0.05 [48,177].

Standard methods of analysis assume that the data are "independent." For controlled trials, this usually means that there is one observation per participant. Treating multiple observations from one participant as independent data is a serious error; such data are produced when outcomes can be measured on different parts of the body, as in dentistry or rheumatology. Data analysis should be based on counting each participant once [178,179] or should be done by using more complex statistical procedures [180]. Incorrect analysis of multiple observations per individual was seen in 123 (63%) of 196 trials in rheumatoid arthritis [108].

5.3.17. Item 12b. Methods for additional analyses, such as subgroup analyses and adjusted analyses

Examples—"Proportions of patients responding were compared between treatment groups with the Mantel-Haenszel χ^2 test, adjusted for the stratification variable, methotrexate use"[103].

"Pre-specified subgroup analyses according to antioxidant treatment assignment(s), presence or absence of prior CVD, dietary folic acid intake, smoking, diabetes, aspirin, hormone therapy, and multivitamin use were performed using stratified Cox proportional hazards models. These analyses used baseline exposure assessments and were restricted to participants with nonmissing subgroup data at baseline"[181].

Explanation—As is the case for primary analyses, the method of subgroup analysis should be clearly specified. The strongest analyses are those that look for evidence of a difference in treatment effect in complementary subgroups (for example, older and younger participants), a comparison known as a test of interaction [182,183]. A common but misleading approach is to compare P values for separate analyses of the treatment effect in each group. It is incorrect to infer a subgroup effect (interaction) from one significant and one non-significant P value [184]. Such inferences have a high false positive rate.

Because of the high risk for spurious findings, subgroup analyses are often discouraged [14,185]. Post hoc subgroup comparisons (analyses done after looking at the data) are especially likely not to be confirmed by further studies. Such analyses do not have great credibility.

In some studies, imbalances in participant characteristics are adjusted for by using some form of multiple regression analysis. Although the need for adjustment is much less in RCTs than in epidemiological studies, an adjusted analysis may be sensible, especially if one or more variables is thought to be prognostic [186]. Ideally, adjusted analyses should be specified in the study protocol (see item 24). For example, adjustment is often recommended for any stratification variables (see item 8b) on the principle that the analysis strategy should follow the design. In RCTs, the decision to adjust should not be determined by whether baseline differences are statistically significant (see item 16) [183,187]. The rationale for any adjusted analyses and the statistical methods used should be specified.

Authors should clarify the choice of variables that were adjusted for, indicate how continuous variables were handled, and specify whether the analysis was planned or suggested by the data [188]. Reviews of published studies show that reporting of adjusted analyses is inadequate with regard to all of these aspects [188−191].

5.4. Results

5.4.1. Item 13. Participant flow (a diagram is strongly recommended)

5.4.1.1. Item 13a. For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome.

Examples—See Figs 2 and 3.

Explanation—The design and conduct of some RCTs is straightforward, and the flow of participants, particularly were there are no losses to follow-up or exclusions, through each phase of the study can be described adequately in a few sentences. In more complex studies, it may be difficult for readers to discern whether and why some participants did not receive the treatment as allocated, were lost to follow-up, or were excluded from the analysis [51]. This information is crucial for several reasons. Participants who were excluded after allocation are unlikely to be representative of all participants in the study. For example, patients may not be available for follow-up evaluation because they experienced an acute exacerbation of their illness or harms of treatment [22,192].

Attrition as a result of loss to follow up, which is often unavoidable, needs to be distinguished from investigator-determined exclusion for such reasons as ineligibility, withdrawal from treatment, and poor adherence to the trial protocol. Erroneous conclusions can be reached if participants are excluded from analysis, and imbalances in such omissions between groups may be especially indicative of bias [192−194]. Information about whether the investigators included in the analysis all participants who underwent randomisation, in the groups to which they were originally allocated (intention-to-treat analysis (see item 16 and Box 6)), is therefore of particular importance. Knowing the number of participants who did not receive the intervention as allocated or did not complete treatment permits the reader to assess to what extent the estimated efficacy of

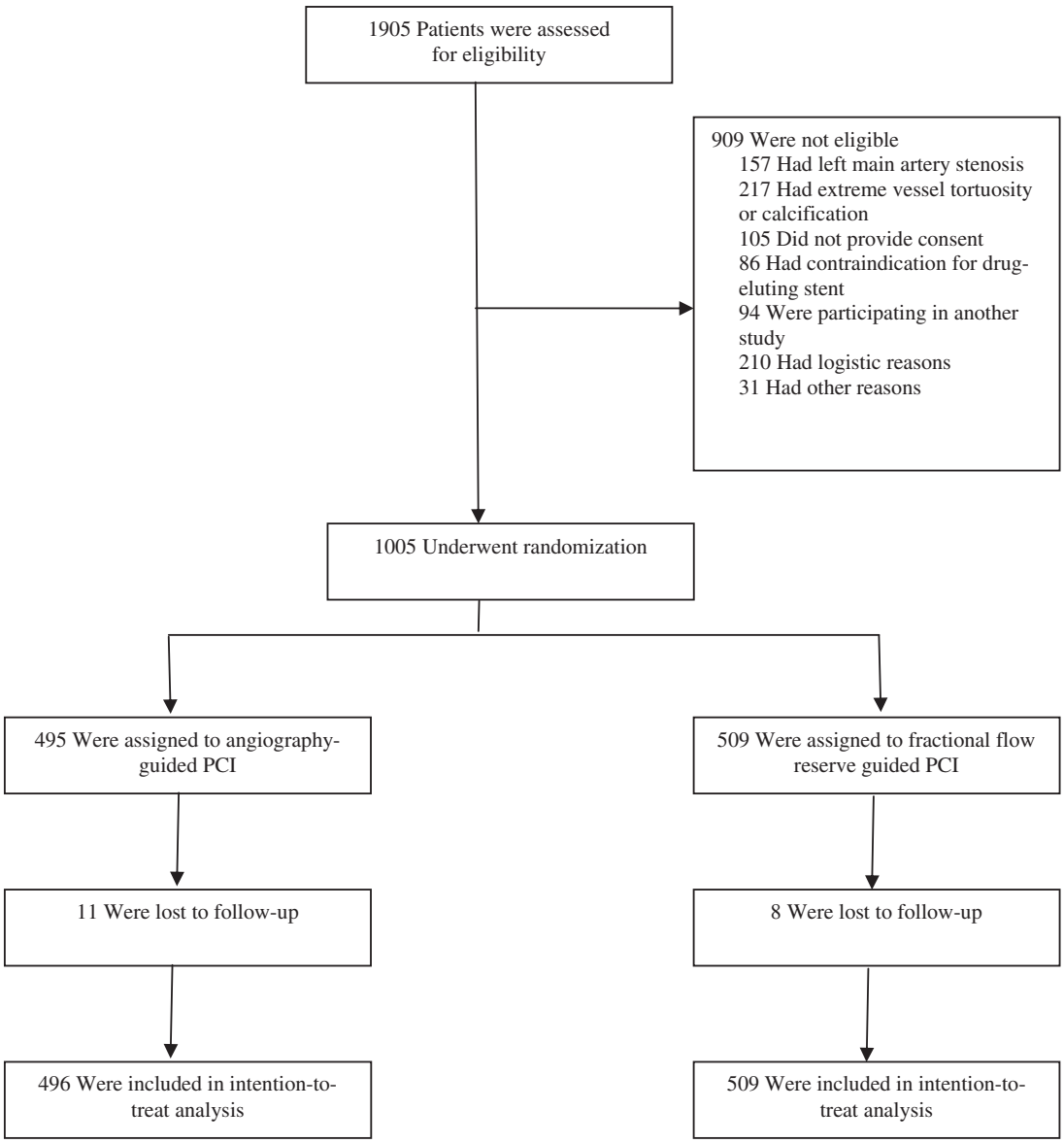

Fig. 2. Flow diagram of a multicentre trial of fractional flow reserve versus angiography for guiding percutaneous coronary intervention (PCI) (adapted from Tonino et al [313]). The diagram includes detailed information on the excluded participants.

therapy might be underestimated in comparison with ideal circumstances.

If available, the number of people assessed for eligibility should also be reported. Although this number is relevant to external validity only and is arguably less important than the other counts [195], it is a useful indicator of whether trial participants were likely to be representative of all eligible participants.

A review of RCTs published in five leading general and internal medicine journals in 1998 found that reporting of the flow of participants was often incomplete, particularly with regard to the number of participants receiving the allocated intervention and the number lost to follow-up [51]. Even information as basic as the number of participants who underwent randomisation and the number excluded from analyses was not available in up to 20% of articles [51]. Reporting was considerably more thorough in articles that included a diagram of the flow of participants through a trial, as recommended by CONSORT. This study informed the design of the revised flow diagram in the revised CONSORT statement [52−54]. The suggested template is shown in Fig 1, and the counts required are described in detail in Table 3.

Some information, such as the number of individuals assessed for eligibility, may not always be known [14], and, depending on the nature of a trial, some counts may be more relevant than others. It will sometimes be useful or necessary to adapt the structure of the flow diagram to a particular trial. In some situations, other information may usefully be added. For example, the flow diagram of a parallel

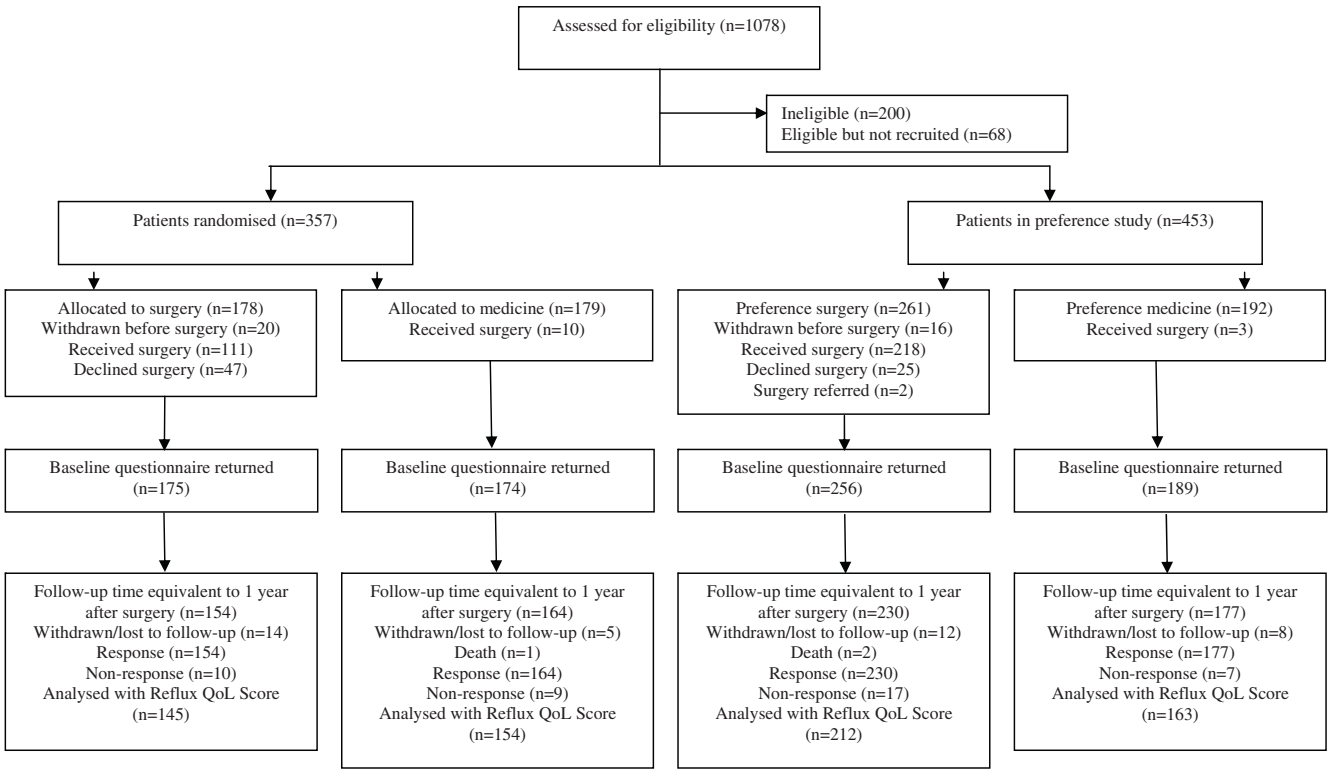

Fig. 3. Flow diagram of minimal surgery compared with medical management for chronic gastro-oesophageal reflux disease (adapted from Grant et al [196]). The diagram shows a multicentre trial with a parallel non-randomised preference group.

group trial of minimal surgery compared with medical management for chronic gastro-oesophageal reflux also included a parallel non-randomised preference group (see Fig 3) [196].

The exact form and content of the flow diagram may be varied according to specific features of a trial. For example, many trials of surgery or vaccination do not include the possibility of discontinuation. Although CONSORT strongly recommends using this graphical device to communicate participant flow throughout the study, there is no specific, prescribed format.

5.4.1.2. Item 13b. For each group, losses and exclusions after randomisation, together with reasons.

Examples—"There was only one protocol deviation, in a woman in the study group. She had an abnormal pelvic measurement and was scheduled for elective caesarean section. However, the attending obstetrician judged a trial of labour acceptable; caesarean section was done when there was no progress in the first stage of labour"[197].

"The monitoring led to withdrawal of nine centres, in which existence of some patients could not be proved, or other serious violations of good clinical practice had occurred"[198].

Explanation—Some protocol deviations may be reported in the flow diagram (see item 13a)—for example, participants who did not receive the intended intervention. If participants were excluded after randomisation (contrary to the intention-to-treat principle) because they were found not to meet eligibility criteria (see item 16), they should be included in the flow diagram. Use of the term "protocol deviation" in published articles is not sufficient to justify exclusion of participants after randomisation. The nature of the protocol deviation and the exact reason for excluding participants after randomisation should always be reported.

5.4.2. Item 14a. Dates defining the periods of recruitment and follow-up

Example—"Age-eligible participants were recruited … from February 1993 to September 1994 … Participants attended clinic visits at the time of randomisation (baseline) and at 6-month intervals for 3 years"[199].

Explanation—Knowing when a study took place and over what period participants were recruited places the study in historical context. Medical and surgical therapies, including concurrent therapies, evolve continuously and may affect the routine care given to participants during a trial. Knowing the rate at which participants were recruited may also be useful, especially to other investigators.

The length of follow-up is not always a fixed period after randomisation. In many RCTs in which the outcome is time to an event, follow-up of all participants is ended on a specific date. This date should be given, and it is also useful to report the minimum, maximum, and median duration of follow-up [200,201].

A review of reports in oncology journals that used survival analysis, most of which were not RCTs [201], found that nearly 80% (104 of 132 reports) included the starting

Table 3
Information required to document the flow of participants through each stage of a randomised trial

Stage	Number of people included	Number of people not included or excluded	Rationale
Enrolment	People evaluated for potential enrolment	People who did not meet the inclusion criteria or met the inclusion criteria but declined to be enrolled	These counts indicate whether trial participants were likely to be representative of all patients seen; they are relevant to assessment of external validity only, and they are often not available.
Randomisation	Participants randomly assigned		Crucial count for defining trial size and assessing whether a trial has been analysed by intention to treat
Treatment allocation	Participants who completed treatment as allocated, by study group	Participants who did not complete treatment as allocated, by study group	Important counts for assessment of internal validity and interpretation of results; reasons for not receiving treatment as allocated should be given.
Follow-up	Participants who completed treatment as allocated, by study group Participants who completed follow-up as planned, by study group	Participants who did not complete treatment as allocated, by study group Participants who did not complete follow-up as planned, by study group	Important counts for assessment of internal validity and interpretation of results; reasons for not completing treatment or follow-up should be given.
Analysis	Participants included in main analysis, by study group	Participants excluded from main analysis, by study group	Crucial count for assessing whether a trial has been analysed by intention to treat; reasons for excluding participants should be given.

and ending dates for accrual of patients, but only 24% (32 of 132 reports) also reported the date on which follow-up ended.

5.4.3. Item 14b. Why the trial ended or was stopped

Examples—"At the time of the interim analysis, the total follow-up included an estimated 63% of the total number of patient-years that would have been collected at the end of the study, leading to a threshold value of 0.0095, as determined by the Lan-DeMets alpha-spending function method … At the interim analysis, the RR was 0.37 in the intervention group, as compared with the control group, with a p value of 0.00073, below the threshold value. The Data and Safety Monitoring Board advised the investigators to interrupt the trial and offer circumcision to the control group, who were then asked to come to the investigation centre, where MC (medical circumcision) was advised and proposed … Because the study was interrupted, some participants did not have a full follow-up on that date, and their visits that were not yet completed are described as "planned" in this article"[202].

"In January 2000, problems with vaccine supply necessitated the temporary nationwide replacement of the whole cell component of the combined DPT/Hib vaccine with acellular pertussis vaccine. As this vaccine has a different local reactogenicity profile, we decided to stop the trial early"[203].

Explanation—Arguably, trialists who arbitrarily conduct unplanned interim analyses after very few events accrue using no statistical guidelines run a high risk of "catching" the data at a random extreme, which likely represents a large overestimate of treatment benefit [204].

Readers will likely draw weaker inferences from a trial that was truncated in a data-driven manner versus one that reports its findings after reaching a goal independent of results. Thus, RCTs should indicate why the trial came to an end (see Box 5). The report should also disclose factors extrinsic to the trial that affected the decision to stop the trial, and who made the decision to stop the trial, including reporting the role the funding agency played in the deliberations and in the decision to stop the trial [134].

A systematic review of 143 RCTs stopped earlier than planned for benefit found that these trials reported stopping after accruing a median of 66 events, estimated a median relative risk of 0.47 and a strong relation between the number of events accrued and the size of the effect, with smaller trials with fewer events yielding the largest treatment effects (odds ratio 31, 95% confidence interval 12 to 82) [134]. While an increasing number of trials published in high impact medical journals report stopping early, only 0.1% of trials reported stopping early for benefit, which contrasts with estimates arising from simulation studies [205] and surveys of data safety and monitoring committees [206]. Thus, many trials accruing

few participants and reporting large treatment effects may have been stopped earlier than planned but failed to report this action.

5.4.4. Item 15. A table showing baseline demographic and clinical characteristics for each group

Example—See Table 4

Explanation—Although the eligibility criteria (see item 4a) indicate who was eligible for the trial, it is also important to know the characteristics of the participants who were actually included. This information allows readers, especially clinicians, to judge how relevant the results of a trial might be to an individual patient.

Randomised trials aim to compare groups of participants that differ only with respect to the intervention (treatment). Although proper random assignment prevents selection bias, it does not guarantee that the groups are equivalent at baseline. Any differences in baseline characteristics are, however, the result of chance rather than bias [32]. The study groups should be compared at baseline for important demographic and clinical characteristics so that readers can assess how similar they were. Baseline data are especially valuable for outcomes that can also be measured at the start of the trial (such as blood pressure).

Baseline information is most efficiently presented in a table (see Table 4). For continuous variables, such as weight or blood pressure, the variability of the data should be reported, along with average values. Continuous variables can be summarised for each group by the mean and standard deviation. When continuous data have an asymmetrical distribution, a preferable approach may be to quote the median and a centile range (such as the 25th and 75th centiles) [177]. Standard errors and confidence intervals are not appropriate for describing variability—they are inferential rather than descriptive statistics. Variables with a small number of ordered categories (such as stages of disease I to IV) should not be treated as continuous variables; instead, numbers and proportions should be reported for each category [48,177].

Unfortunately significance tests of baseline differences are still common [23,32,210]; they were reported in half of 50 RCTs trials published in leading general journals in 1997 [183]. Such significance tests assess the probability that observed baseline differences could have occurred by chance; however, we already know that any differences are caused by chance. Tests of baseline differences are not necessarily wrong, just illogical [211]. Such hypothesis testing is superfluous and can mislead investigators and their readers. Rather, comparisons at baseline should be based on consideration of the prognostic strength of the variables measured and the size of any chance imbalances that have occurred [211].

Box 5. Early stopping

RCTs can end when they reach their sample size goal, their event count goal, their length of follow-up goal, or when they reach their scheduled date of closure. In these situations the trial will stop in a manner independent of its results, and stopping is unlikely to introduce bias in the results. Alternatively, RCTs can stop earlier than planned because of the result of an interim analysis showing larger than expected benefit or harm on the experimental intervention. Also RCTs can stop earlier than planned when investigators find evidence of no important difference between experimental and control interventions (that is, stopping for futility). In addition, trials may stop early because the trial becomes unviable: funding vanishes, researchers cannot access eligible patients or study interventions, or the results of other studies make the research question irrelevant.

Full reporting of why a trial ended is important for evidence based decision making (see item 14b). Researchers examining why 143 trials stopped early for benefit found that many failed to report key methodological information regarding how the decision to stop was reached—the planned sample size (n=28), interim analysis after which the trial was stopped (n=45), or whether a stopping rule informed the decision (n=48) [134]. Item 7b of the checklist requires the reporting of timing of interim analyses, what triggered them, how many took place, whether these were planned or ad hoc, and whether there were statistical guidelines and stopping rules in place a priori. Furthermore, it is helpful to know whether an independent data monitoring committee participated in the analyses (and who composed it, with particular attention to the role of the funding source) and who made the decision to stop. Often the data safety and monitoring committee makes recommendations and the funders (sponsors) or the investigators make the decision to stop.

Trials that stop early for reasons apparently independent of trial findings, and trials that reach their planned termination, are unlikely to introduce bias by stopping [207]. In these cases, the authors should report whether interim analyses took place and whether these results were available to the funder.

The push for trials that change the intervention in response to interim results, thus enabling a faster evaluation of promising interventions for rapidly evolving and fatal conditions, will require even more careful reporting of the process and decision to stop trials early [208].

Table 4
Example of reporting baseline demographic and clinical characteristics*

	Telmisartan (N=2954)	Placebo (N=2972)
Age (years)	66.9 (7.3)	66.9 (7.4)
Sex (female)	1280 (43.3%)	1267 (42.6%)
Smoking status:		
Current	293 (9.9%)	289 (9.7%)
Past	1273 (43.1%)	1283 (43.2%)
Ethnic origin:		
Asian	637 (21.6%)	624 (21.0%)
Arab	37 (1.3%)	40 (1.3%)
African	51 (1.7%)	55 (1.9%)
European	1801 (61.0%)	1820 (61.2%)
Native or Aboriginal	390 (13.2%)	393 (13.2%)
Other	38 (1.3%)	40 (1.3%)
Blood pressure (mm Hg)	140.7 (16.8/81.8) (10.1)	141.3 (16.4/82.0) (10.2)
Heart rate (beats per min)	68.8 (11.5)	68.8 (12.1)
Cholesterol (mmol/l):		
Total	5.09 (1.18)	5.08 (1.15)
LDL	3.02 (1.01)	3.03 (1.02)
HDL	1.27 (0.37)	1.28 (0.41)
Coronary artery disease	2211 (74.8%)	2207 (74.3%)
Myocardial infarction	1381 (46.8%)	1360 (45.8%)
Angina pectoris	1412 (47.8%)	1412 (47.5%)
Peripheral artery disease	349 (11.8%)	323 (10.9%)
Hypertension	2259 (76.5%)	2269 (76.3%)
Diabetes	1059 (35.8%)	1059 (35.6%)

* Data are means (SD) or numbers (%).
Adapted from Table 1 of Yusuf et al [209].

5.4.5. Item 16. For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups

Examples—"The primary analysis was intention-to-treat and involved all patients who were randomly assigned"[212].

"One patient in the alendronate group was lost to follow up; thus data from 31 patients were available for the intention-to-treat analysis. Five patients were considered protocol violators … consequently 26 patients remained for the per-protocol analyses"[213].

Explanation—The number of participants in each group is an essential element of the analyses. Although the flow diagram (see item 13a) may indicate the numbers of participants analysed, these numbers often vary for different outcome measures. The number of participants per group should be given for all analyses. For binary outcomes, (such as risk ratio and risk difference) the denominators or event rates should also be reported. Expressing results as fractions also aids the reader in assessing whether some of the randomly assigned participants were excluded from the analysis. It follows that results should not be presented solely as summary measures, such as relative risks.

Participants may sometimes not receive the full intervention, or some ineligible patients may have been randomly allocated in error. One widely recommended way to handle such issues is to analyse all participants according to their original group assignment, regardless of what subsequently occurred (see Box 6). This "intention-to-treat" strategy is not always straightforward to implement. It is common for some patients not to complete a study—they may drop out or be withdrawn from active treatment—and thus are not assessed at the end. If the outcome is mortality, such patients may be included in the analysis based on register information, whereas imputation techniques may need to be used if other outcome data are missing. The term "intention-to-treat analysis" is often inappropriately used—for example, when those who did not receive the first dose of a trial drug are excluded from the analyses [18].

Conversely, analysis can be restricted to only participants who fulfil the protocol in terms of eligibility, interventions, and outcome assessment. This analysis is known as an "on-treatment" or "per protocol" analysis. Excluding participants from the analysis can lead to erroneous conclusions. For example, in a trial that compared medical with surgical therapy for carotid stenosis, analysis limited to participants who were available for follow-up showed that surgery reduced the risk for transient ischaemic attack, stroke, and death. However, intention-to-treat analysis based on all participants as originally assigned did not show a superior effect of surgery [214].

Intention-to-treat analysis is generally favoured because it avoids bias associated with non-random loss of participants [215–217]. Regardless of whether authors use the term "intention-to-treat," they should make clear which and how many participants are included in each analysis (see item 13). Non-compliance with assigned therapy may mean that the intention-to-treat analysis underestimates the potential benefit of the treatment, and additional analyses, such as a per protocol analysis, may therefore be considered [218,219]. It should be noted, however, that such analyses are often considerably flawed [220].

In a review of 403 RCTs published in 10 leading medical journals in 2002, 249 (62%) reported the use of intention-to-treat analysis for their primary analysis. This proportion was higher for journals adhering to the CONSORT statement (70% *v* 48%). Among articles that reported the use of intention-to-treat analysis, only 39% actually analysed all participants as randomised, with more than 60% of articles having missing data in their primary analysis [221]. Other studies show similar findings [18,222,223]. Trials with no reported exclusions are methodologically weaker in other respects than those that report on some excluded participants [173], strongly indicating that at least some researchers who have excluded participants do not report it. Another study found that reporting an intention-to-treat analysis was associated with other aspects of good study design and reporting, such as describing a sample size calculation [224].

Box 6. Intention-to-treat analysis

The special strength of the RCT is the avoidance of bias when allocating interventions to trial participants (see Box 1). That strength allows strong inferences about cause and effect that are not justified with other study designs. In order to preserve fully the huge benefit of randomisation we should include all randomised participants in the analysis, all retained in the group to which they were allocated. Those two conditions define an ''intention-to-treat'' analysis, which is widely recommended as the preferred analysis strategy [18,223]. Intention-to-treat analysis corresponds to analysing the groups exactly as randomised. Strict intention-to-treat analysis is often hard to achieve for two main reasons—missing outcomes for some participants and non-adherence to the trial protocol.

Missing outcomes

Many trialists exclude patients without an observed outcome. Often this is reasonable, but once any randomised participants are excluded the analysis is not strictly an intention-to-treat analysis. Indeed, most randomised trials have some missing observations. Trialists effectively must choose between omitting the participants without final outcome data or imputing their missing outcome data [225]. A ''complete case'' (or ''available case'') analysis includes only those whose outcome is known. While a few missing outcomes will not cause a problem, in half of trials more than 10% of randomised patients may have missing outcomes [226]. This common approach will lose power by reducing the sample size, and bias may well be introduced if being lost to follow-up is related to a patient's response to treatment. There should be concern when the frequency or the causes of dropping out differ between the intervention groups.

Participants with missing outcomes can be included in the analysis only if their outcomes are imputed (that is, their outcomes are estimated from other information that was collected). Imputation of the missing data allows the analysis to conform to intention-to-treat analysis but requires strong assumptions, which may be hard to justify [227]. Simple imputation methods are appealing, but their use may be inadvisable. In particular, a widely used method is ''last observation carried forward'' in which missing final values of the outcome variable are replaced by the last known value before the participant was lost to follow up. This is appealing through its simplicity, but the method may introduce bias [228], and no allowance is made for the uncertainty of imputation [229]. Many authors have severely criticised last observation carried forward [229−231].

Non-adherence to the protocol

A separate issue is that the trial protocol may not have been followed fully for some trial participants. Common examples are participants who did not meet the inclusion criteria (such as wrong diagnosis, too young), received a proscribed co-intervention, did not take all the intended treatment, or received a different treatment or no intervention. The simple way to deal with any protocol deviations is to ignore them: all participants can be included in the analysis regardless of adherence to the protocol, and this is the intention-to-treat approach. Thus, exclusion of any participants for such reasons is incompatible with intention-to-treat analysis.

The term ''modified intention-to-treat'' is quite widely used to describe an analysis that excludes participants who did not adequately adhere to the protocol, in particular those who did not receive a defined minimum amount of the intervention [232]. An alternative term is ''per protocol.'' Though a per protocol analysis may be appropriate in some settings, it should be properly labelled as a non-randomised, observational comparison. Any exclusion of patients from the analysis compromises the randomisation and may lead to bias in the results.

Like ''intention-to-treat,'' none of these other labels reliably clarifies exactly which patients were included. Thus, in the CONSORT checklist we have dropped the specific request for intention-to-treat analysis in favour of a clear description of exactly who was included in each analysis.

5.4.6. Item 17a. For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)

Examples—See Tables 5 and 6.

Explanation—For each outcome, study results should be reported as a summary of the outcome in each group (for example, the number of participants with or without the event and the denominators, or the mean and standard deviation of measurements), together with the contrast between the groups, known as the effect size. For binary outcomes, the effect size could be the risk ratio (relative risk), odds ratio, or risk difference; for survival time data, it could be the hazard ratio or difference in median survival time; and for continuous data, it is

Table 5
Example of reporting of summary results for each study group (binary outcomes)*

Endpoint	Number (%)		Risk difference (95% CI)
	Etanercept (n=30)	Placebo (n=30)	
Primary endpoint			
Achieved PsARC at 12 weeks	26 (87)	7 (23)	63% (44 to 83)
Secondary endpoint			
Proportion of patients meeting ACR criteria:			
ACR20	22 (73)	4 (13)	60% (40 to 80)
ACR50	15 (50)	1 (3)	47% (28 to 66)
ACR70	4 (13)	0 (0)	13% (1 to 26)

PsARC=psoriatic arthritis response criteria. ACR=American College of Rheumatology.

* See also example for item 6a.

Adapted from table 2 of Mease et al [103].

usually the difference in means. Confidence intervals should be presented for the contrast between groups. A common error is the presentation of separate confidence intervals for the outcome in each group rather than for the treatment effect [233]. Trial results are often more clearly displayed in a table rather than in the text, as shown in Tables 5 and 6.

For all outcomes, authors should provide a confidence interval to indicate the precision (uncertainty) of the estimate [48,235]. A 95% confidence interval is conventional, but occasionally other levels are used. Many journals require or strongly encourage the use of confidence intervals [236]. They are especially valuable in relation to differences that do not meet conventional statistical significance, for which they often indicate that the result does not rule out an important clinical difference. The use of confidence intervals has increased markedly in recent years, although not in all medical specialties [233]. Although P values may be provided in addition to confidence intervals, results should not be reported solely as P values [237,238]. Results should be reported for all planned primary and secondary end points, not just for analyses that were statistically significant or "interesting." Selective reporting within a study is a widespread and serious problem [55,57]. In trials in which interim analyses were performed, interpretation should focus on the final results at the close of the trial, not the interim results [239].

For both binary and survival time data, expressing the results also as the number needed to treat for benefit or harm can be helpful (see item 21) [240,241].

5.4.7. Item 17b. For binary outcomes, presentation of both absolute and relative effect sizes is recommended

Example—"The risk of oxygen dependence or death was reduced by 16% (95% CI 25% to 7%). The absolute difference was −6.3% (95% CI −9.9% to −2.7%); early administration to an estimated 16 babies would therefore prevent 1 baby dying or being long-term dependent on oxygen" (also see Table 7) [242].

Explanation—When the primary outcome is binary, both the relative effect (risk ratio (relative risk) or odds ratio) and the absolute effect (risk difference) should be reported (with confidence intervals), as neither the relative measure nor the absolute measure alone gives a complete picture of the effect and its implications. Different audiences may prefer either relative or absolute risk, but both doctors and lay people tend to overestimate the effect when it is presented in terms of relative risk [243–245]. The size of the risk difference is less generalisable to other populations than the relative risk since it depends on the baseline risk in the unexposed group, which tends to vary across populations. For diseases where the outcome is common, a relative risk near unity might indicate clinically important differences in public health terms. In contrast, a large relative risk when the outcome is rare may not be so important for public health (although it may be important to an individual in a high risk category).

5.4.8. Item 18. Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory

Example—"On the basis of a study that suggested perioperative β-blocker efficacy might vary across baseline risk, we prespecified our primary subgroup analysis on the basis of the revised cardiac risk index scoring system. We also did prespecified secondary subgroup analyses based on sex, type of surgery, and use of an epidural or spinal anaesthetic. For all subgroup analyses, we used Cox proportional hazard models that incorporated tests for

Table 6
Example of reporting of summary results for each study group (continuous outcomes)

	Exercise therapy (n=65)		Control (n=66)		Adjusted difference* (95% CI) at 12 months
	Baseline (mean (SD))	12 months (mean (SD))	Baseline (mean (SD))	12 months (mean (SD))	
Function score (0–100)	64.4 (13.9)	83.2 (14.8)	65.9 (15.2)	79.8 (17.5)	4.52 (−0.73 to 9.76)
Pain at rest (0–100)	4.14 (2.3)	1.43 (2.2)	4.03 (2.3)	2.61 (2.9)	−1.29 (−2.16 to −0.42)
Pain on activity (0–100)	6.32 (2.2)	2.57 (2.9)	5.97 (2.3)	3.54 (3.38)	−1.19 (−2.22 to −0.16)

* Function score adjusted for baseline, age, and duration of symptoms.

Adapted from table 3 of van Linschoten [234].

interactions, designated to be significant at $p<0.05$ … Figure 3 shows the results of our prespecified subgroup analyses and indicates consistency of effects … Our subgroup analyses were underpowered to detect the modest differences in subgroup effects that one might expect to detect if there was a true subgroup effect''[100].

Explanation—Multiple analyses of the same data create a risk for false positive findings [246]. Authors should resist the temptation to perform many subgroup analyses [183,185,247]. Analyses that were prespecified in the trial protocol (see item 24) are much more reliable than those suggested by the data, and therefore authors should report which analyses were prespecified. If subgroup analyses were undertaken, authors should report which subgroups were examined, why, if they were prespecified, and how many were prespecified. Selective reporting of subgroup analyses could lead to bias [248]. When evaluating a subgroup the question is not whether the subgroup shows a statistically significant result but whether the subgroup treatment effects are significantly different from each other. To determine this, a test of interaction is helpful, although the power for such tests is typically low. If formal evaluations of interaction are undertaken (see item 12b) they should be reported as the estimated difference in the intervention effect in each subgroup (with a confidence interval), not just as P values.

In one survey, 35 of 50 trial reports included subgroup analyses, of which only 42% used tests of interaction [183]. It was often difficult to determine whether subgroup analyses had been specified in the protocol. In another survey of surgical trials published in high impact journals, 27 of 72 trials reported 54 subgroup analyses, of which 91% were post hoc and only 6% of subgroup analyses used a test of interaction to assess whether a subgroup effect existed [249].

Similar recommendations apply to analyses in which adjustment was made for baseline variables. If done, both unadjusted and adjusted analyses should be reported. Authors should indicate whether adjusted analyses, including the choice of variables to adjust for, were planned. Ideally, the trial protocol should state whether adjustment is made for nominated baseline variables by using analysis of covariance [187]. Adjustment for variables because they differ significantly at baseline is likely to bias the estimated treatment effect [187]. A survey found that unacknowledged discrepancies between protocols and publications were found for all 25 trials reporting subgroup analyses and for 23 of 28 trials reporting adjusted analyses [92].

5.4.9. Item 19. All important harms or unintended effects in each group

For specific guidance see CONSORT for harms [42].

Example—''The proportion of patients experiencing any adverse event was similar between the rBPI21 [recombinant bactericidal/permeability-increasing protein] and placebo groups: 168 (88.4%) of 190 and 180 (88.7%) of 203, respectively, and it was lower in patients treated with rBPI21 than in those treated with placebo for 11 of 12 body systems … the proportion of patients experiencing a severe adverse event, as judged by the investigators, was numerically lower in the rBPI21 group than the placebo group: 53 (27.9%) of 190 versus 74 (36.5%) of 203 patients, respectively. There were only three serious adverse events reported as drug-related and they all occurred in the placebo group''[250].

Explanation—Readers need information about the harms as well as the benefits of interventions to make rational and balanced decisions. The existence and nature of adverse effects can have a major impact on whether a particular intervention will be deemed acceptable and useful. Not all reported adverse events observed during a trial are necessarily a consequence of the intervention; some may be a consequence of the condition being treated. Randomised trials offer the best approach for providing safety data as well as efficacy data, although they cannot detect rare harms.

Many reports of RCTs provide inadequate information on adverse events. A survey of 192 drug trials published from 1967 to 1999 showed that only 39% had adequate reporting of clinical adverse events and 29% had adequate reporting of laboratory defined toxicity [72]. More recently, a comparison between the adverse event data submitted to the trials database of the National Cancer Institute, which sponsored the trials, and the information reported in journal articles found that low grade adverse events were underreported in journal articles. High grade events (Common Toxicity Criteria grades 3 to 5) were reported inconsistently in the articles, and the information regarding attribution to investigational drugs was incomplete [251]. Moreover, a review of trials published in six general medical journals in 2006 to 2007 found that, although 89% of 133 reports mentioned adverse events, no information on severe adverse events and withdrawal of patients due to an adverse event was given on 27% and 48% of articles, respectively [252].

An extension of the CONSORT statement has been developed to provide detailed recommendations on the reporting of harms in randomised trials [42]. Recommendations and examples of appropriate reporting are freely available from the CONSORT website (www.consort-statement.org). They complement the CONSORT 2010 Statement and should be consulted, particularly if the study of harms was a key objective. Briefly, if data on adverse events were collected, events should be listed and defined, with reference to standardised criteria where appropriate. The methods used for data collection and attribution of events should be described. For each study arm the absolute risk of each adverse event, using appropriate metrics for recurrent events, and the number of participants withdrawn due to harms should be presented. Finally, authors should provide a balanced discussion of benefits and harms [42].

Table 7
Example of reporting both absolute and relative effect sizes

	Percentage (No)			
Primary outcome	Early administration (n=1344)	Delayed selective administration (n=1346)	Risk ratio (95% CI)	Risk difference (95% CI)
Death or oxygen dependence at "expected date of delivery"	31.9 (429)	38.2 (514)	0.84 (0.75 to 0.93)	−6.3 (−9.9 to −2.7)

Adapted from table 3 of The OSIRIS Collaborative Group [242].

5.5. Discussion

5.5.1. Item 20. Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses

Example—"The preponderance of male patients (85%) is a limitation of our study … We used bare-metal stents, since drug-eluting stents were not available until late during accrual. Although the latter factor may be perceived as a limitation, published data indicate no benefit (either short-term or long-term) with respect to death and myocardial infarction in patients with stable coronary artery disease who receive drug-eluting stents, as compared with those who receive bare-metal stents"[253].

Explanation—The discussion sections of scientific reports are often filled with rhetoric supporting the authors' findings [254] and provide little measured argument of the pros and cons of the study and its results. Some journals have attempted to remedy this problem by encouraging more structure to authors' discussion of their results [255,256]. For example, *Annals of Internal Medicine* recommends that authors structure the discussion section by presenting (1) a brief synopsis of the key findings, (2) consideration of possible mechanisms and explanations, (3) comparison with relevant findings from other published studies (whenever possible including a systematic review combining the results of the current study with the results of all previous relevant studies), (4) limitations of the present study (and methods used to minimise and compensate for those limitations), and (5) a brief section that summarises the clinical and research implications of the work, as appropriate [255]. We recommend that authors follow these sensible suggestions, perhaps also using suitable subheadings in the discussion section.

Although discussion of limitations is frequently omitted from research reports [257], identification and discussion of the weaknesses of a study have particular importance [258]. For example, a surgical group reported that laparoscopic cholecystectomy, a technically difficult procedure, had significantly lower rates of complications than the more traditional open cholecystectomy for management of acute cholecystitis [259]. However, the authors failed to discuss an obvious bias in their results. The study investigators had completed all the laparoscopic cholecystectomies, whereas 80% of the open cholecystectomies had been completed by trainees.

Authors should also discuss any imprecision of the results. Imprecision may arise in connection with several aspects of a study, including measurement of a primary outcome (see item 6a) or diagnosis (see item 4a). Perhaps the scale used was validated on an adult population but used in a paediatric one, or the assessor was not trained in how to administer the instrument.

The difference between statistical significance and clinical importance should always be borne in mind. Authors should particularly avoid the common error of interpreting a non-significant result as indicating equivalence of interventions. The confidence interval (see item 17a) provides valuable insight into whether the trial result is compatible with a clinically important effect, regardless of the P value [120].

Authors should exercise special care when evaluating the results of trials with multiple comparisons. Such multiplicity arises from several interventions, outcome measures, time points, subgroup analyses, and other factors. In such circumstances, some statistically significant findings are likely to result from chance alone.

5.5.2. Item 21. Generalisability (external validity, applicability) of the trial findings

Examples—"As the intervention was implemented for both sexes, all ages, all types of sports, and at different levels of sports, the results indicate that the entire range of athletes, from young elite to intermediate and recreational senior athletes, would benefit from using the presented training programme for the prevention of recurrences of ankle sprain. By including non-medically treated and medically treated athletes, we covered a broad spectrum of injury severity. This suggests that the present training programme can be implemented in the treatment of all athletes. Furthermore, as it is reasonable to assume that ankle sprains not related to sports are comparable with those in sports, the programme could benefit the general population"[260].

"This replicates and extends the work of Clarke and colleagues and demonstrates that this CB (cognitive behavioural) prevention program can be reliably and effectively delivered in different settings by clinicians outside of the group who originally developed the intervention. The effect size was consistent with those of previously reported,

single-site, indicated depression prevention studies and was robust across sites with respect to both depressive disorders and symptoms … In this generalisability trial, we chose a comparison condition that is relevant to public health—usual care … The sample also was predominantly working class to middle class with access to health insurance. Given evidence that CB therapy can be more efficacious for adolescents from homes with higher incomes, it will be important to test the effects of this prevention program with more economically and ethnically diverse samples''[261].

Explanation—External validity, also called generalisability or applicability, is the extent to which the results of a study can be generalised to other circumstances [262]. Internal validity, the extent to which the design and conduct of the trial eliminate the possibility of bias, is a prerequisite for external validity: the results of a flawed trial are invalid and the question of its external validity becomes irrelevant. There is no absolute external validity; the term is meaningful only with regard to clearly specified conditions that were not directly examined in the trial. Can results be generalised to an individual participant or groups that differ from those enrolled in the trial with regard to age, sex, severity of disease, and comorbid conditions? Are the results applicable to other drugs within a class of similar drugs, to a different dose, timing, and route of administration, and to different concomitant therapies? Can similar results be expected at the primary, secondary, and tertiary levels of care? What about the effect on related outcomes that were not assessed in the trial, and the importance of length of follow-up and duration of treatment, especially with respect to harms? [263].

External validity is a matter of judgment and depends on the characteristics of the participants included in the trial, the trial setting, the treatment regimens tested, and the outcomes assessed [5,136]. It is therefore crucial that adequate information be described about eligibility criteria and the setting and location (see item 4b), the interventions and how they were administered (see item 5), the definition of outcomes (see item 6), and the period of recruitment and follow-up (see item 14). The proportion of control group participants in whom the outcome develops (control group risk) is also important. The proportion of eligible participants who refuse to enter the trial as indicated on the flow-chart (see item 13) is relevant for the generalisability of the trial, as it may indicate preferences for or acceptability of an intervention. Similar considerations may apply to clinician preferences [264,265].

Several issues are important when results of a trial are applied to an individual patient [266−268]. Although some variation in treatment response between an individual patient and the patients in a trial or systematic review is to be expected, the differences tend to be in magnitude rather than direction.

Although there are important exceptions [268], therapies (especially drugs [269]) found to be beneficial in a narrow range of patients generally have broader application in actual practice. Frameworks for the evaluation of external validity have been proposed, including qualitative studies, such as in integral ''process evaluations''[270] and checklists [271]. Measures that incorporate baseline risk when calculating therapeutic effects, such as the number needed to treat to obtain one additional favourable outcome and the number needed to treat to produce one adverse effect, are helpful in assessing the benefit-to-risk balance in an individual patient or group with characteristics that differ from the typical trial participant [268,272,273]. Finally, after deriving patient centred estimates for the potential benefit and harm from an intervention, the clinician must integrate them with the patient's values and preferences for therapy. Similar considerations apply when assessing the generalisability of results to different settings and interventions.

5.5.3. Item 22. Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence

Example—''Studies published before 1990 suggested that prophylactic immunotherapy also reduced nosocomial infections in very-low-birth-weight infants. However, these studies enrolled small numbers of patients; employed varied designs, preparations, and doses; and included diverse study populations. In this large multicenter, randomised controlled trial, the repeated prophylactic administration of intravenous immune globulin failed to reduce the incidence of nosocomial infections significantly in premature infants weighing 501 to 1500 g at birth''[274].

Explanation—Readers will want to know how the present trial's results relate to those of other RCTs. This can best be achieved by including a formal systematic review in the results or discussion section of the report [83,275−277]. Such synthesis may be impractical for trial authors, but it is often possible to quote a systematic review of similar trials. A systematic review may help readers assess whether the results of the RCT are similar to those of other trials in the same topic area and whether participants are similar across studies. Reports of RCTs have often not dealt adequately with these points [277]. Bayesian methods can be used to statistically combine the trial data with previous evidence [278].

We recommend that, at a minimum, the discussion should be as systematic as possible and be based on a comprehensive search, rather than being limited to studies that support the results of the current trial [279].

5.6. Other information

5.6.1. Item 23. Registration number and name of trial registry

Example—''The trial is registered at ClinicalTrials.gov, number NCT00244842''[280].

Explanation—The consequences of non-publication of entire trials [281,282], selective reporting of outcomes

within trials, and of per protocol rather than intention-to-treat analysis have been well documented [55,56,283]. Covert redundant publication of clinical trials can also cause problems, particularly for authors of systematic reviews when results from the same trial are inadvertently included more than once [284].

To minimise or avoid these problems there have been repeated calls over the past 25 years to register clinical trials at their inception, to assign unique trial identification numbers, and to record other basic information about the trial so that essential details are made publicly available [285−288]. Provoked by recent serious problems of withholding data [289], there has been a renewed effort to register randomised trials. Indeed, the World Health Organisation states that ''the registration of all interventional trials is a scientific, ethical and moral responsibility'' (www.who.int/ictrp/en). By registering a randomised trial, authors typically report a minimal set of information and obtain a unique trial registration number.

In September 2004 the International Committee of Medical Journal Editors (ICMJE) changed their policy, saying that they would consider trials for publication only if they had been registered before the enrolment of the first participant [290]. This resulted in a dramatic increase in the number of trials being registered [291]. The ICMJE gives guidance on acceptable registries (www.icmje.org/faq.pdf).

In a recent survey of 165 high impact factor medical journals' instructions to authors, 44 journals specifically stated that all recent clinical trials must be registered as a requirement of submission to that journal [292].

Authors should provide the name of the register and the trial's unique registration number. If authors had not registered their trial they should explicitly state this and give the reason.

5.6.2. Item 24. Where the full trial protocol can be accessed, if available

Example—''Full details of the trial protocol can be found in the Supplementary Appendix, available with the full text of this article at www.nejm.org''[293].

Explanation—A protocol for the complete trial (rather than a protocol of a specific procedure within a trial) is important because it pre-specifies the methods of the randomised trial, such as the primary outcome (see item 6a). Having a protocol can help to restrict the likelihood of undeclared post hoc changes to the trial methods and selective outcome reporting (see item 6b). Elements that may be important for inclusion in the protocol for a randomised trial are described elsewhere [294].

There are several options for authors to consider ensuring their trial protocol is accessible to interested readers. As described in the example above, journals reporting a trial's primary results can make the trial protocol available on their web site. Accessibility to the trial results and protocol is enhanced when the journal is open access. Some journals (such as *Trials*) publish trial protocols, and such a publication can be referenced when reporting the trial's principal results. Trial registration (see item 23) will also ensure that many trial protocol details are available, as the minimum trial characteristics included in an approved trial registration database includes several protocol items and results (www.who.int/ictrp/en). Trial investigators may also be able to post their trial protocol on a website through their employer. Whatever mechanism is used, we encourage all trial investigators to make their protocol easily accessible to interested readers.

5.6.3. Item 25. Sources of funding and other support (such as supply of drugs), role of funders

Examples—''Grant support was received for the intervention from Plan International and for the research from the Wellcome Trust and Joint United Nations Programme on HIV/AIDS (UNAIDS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript''[295].

''This study was funded by GlaxoSmithKline Pharmaceuticals. GlaxoSmithKline was involved in the design and conduct of the study and provided logistical support during the trial. Employees of the sponsor worked with the investigators to prepare the statistical analysis plan, but the analyses were performed by the University of Utah. The manuscript was prepared by Dr Shaddy and the steering committee members. GlaxoSmithKline was permitted to review the manuscript and suggest changes, but the final decision on content was exclusively retained by the authors''[296].

Explanation—Authors should report the sources of funding for the trial, as this is important information for readers assessing a trial. Studies have showed that research sponsored by the pharmaceutical industry are more likely to produce results favouring the product made by the company sponsoring the research than studies funded by other sources [297−300]. A systematic review of 30 studies on funding found that research funded by the pharmaceutical industry had four times the odds of having outcomes favouring the sponsor than research funded by other sources (odds ratio 4.05, 95% confidence interval 2.98 to 5.51) [297]. A large proportion of trial publications do not currently report sources of funding. The degree of underreporting is difficult to quantify. A survey of 370 drug trials found that 29% failed to report sources of funding [301]. In another survey, of PubMed indexed randomised trials published in December 2000, source of funding was reported for 66% of the 519 trials [16].

The level of involvement by a funder and their influence on the design, conduct, analysis, and reporting of a trial varies. It is therefore important that authors describe in detail the role of the funders. If the funder had no such involvement, the authors should state so. Similarly, authors should report any other sources of support, such as supply and preparation of drugs or equipment, or in the analysis of data and writing of the manuscript [302].

6. Reporting RCTs that did not have a two group parallel design

The primary focus of the CONSORT recommendations is RCTs with a parallel design and two treatment groups. Most RCTs have that design, but a substantial minority do not: 45% (233/519) of RCTs published in December 2000 [16], and 39% (242/616) in December 2006 [17].

Most of the CONSORT statement applies equally to all trial designs, but there are a few additional issues to address for each design. Before the publication of the revised CONSORT statement in 2001, the CONSORT Group decided to develop extensions to the main CONSORT statement relevant to specific trial designs. Extensions have been published relating to reporting of cluster randomised trials [40] and non-inferiority and equivalence trials [39]. Lack of resources has meant that other planned extensions have not been completed; they will cover trials with the following designs: multiarm parallel, factorial, crossover, within-person.

Authors reporting trials with a cluster design or using a non-inferiority or equivalence framework should consult the CONSORT recommendations in addition to those in this document. Here we make a few interim comments about the other designs. In each case, the trial design should be made clear in both the main text and the article's abstract.

Multiarm (>2 group) parallel group trials need the least modification of the standard CONSORT guidance. The flow diagram can be extended easily. The main differences from trials with two groups relate to clarification of how the study hypotheses relate to the multiple groups, and the consequent methods of data analysis and interpretation. For factorial trials, the possibility of interaction between the interventions generally needs to be considered. In addition to overall comparisons of participants who did or did not receive each intervention under study, investigators should consider also reporting results for each treatment combination [303].

In crossover trials, each participant receives two (or more) treatments in a random order. The main additional issues to address relate to the paired nature of the data, which affect design and analysis [304]. Similar issues affect within-person comparisons, in which participants receive two treatments simultaneously (often to paired organs). Also, because of the risk of temporal or systemic carryover effects, respectively, in both cases the choice of design needs justification.

The CONSORT Group intends to publish extensions to CONSORT to cover all these designs. In addition, we will publish updates to existing guidance for cluster randomised trials and non-inferiority and equivalence trials to take account of this major update of the generic CONSORT guidance.

7. Discussion

Assessment of healthcare interventions can be misleading unless investigators ensure unbiased comparisons. Random allocation to study groups remains the only method that eliminates selection and confounding biases. Non-randomised trials tend to result in larger estimated treatment effects than randomised trials [305,306].

Bias jeopardises even RCTs, however, if investigators carry out such trials improperly [307]. A recent systematic review, aggregating the results of several methodological investigations, found that, for subjective outcomes, trials that used inadequate or unclear allocation concealment yielded 31% larger estimates of effect than those that used adequate concealment, and trials that were not blinded yielded 25% larger estimates [153]. As might be expected, there was a strong association between the two.

The design and implementation of an RCT require methodological as well as clinical expertise, meticulous effort [143,308], and a high level of alertness for unanticipated difficulties. Reports of RCTs should be written with similarly close attention to reducing bias. Readers should not have to speculate; the methods used should be complete and transparent so that readers can readily differentiate trials with unbiased results from those with questionable results. Sound science encompasses adequate reporting, and the conduct of ethical trials rests on the footing of sound science [309].

We hope this update of the CONSORT explanatory article will assist authors in using the 2010 version of CONSORT and explain in general terms the importance of adequately reporting of trials. The CONSORT statement can help researchers designing trials in future [310] and can guide peer reviewers and editors in their evaluation of manuscripts. Indeed, we encourage peer reviewers and editors to use the CONSORT checklist to assess whether authors have reported on these items. Such assessments will likely improve the clarity and transparency of published trials. Because CONSORT is an evolving document, it requires a dynamic process of continual assessment, refinement, and, if necessary, change, which is why we have this update of the checklist and explanatory article. As new evidence and critical comments accumulate, we will evaluate the need for future updates.

The first version of the CONSORT statement, from 1996, seems to have led to improvement in the quality of reporting of RCTs in the journals that have adopted it [50−54]. Other groups are using the CONSORT template to improve the reporting of other research designs, such as diagnostic tests [311] and observational studies [312].

The CONSORT website (www.consort-statement.org) has been established to provide educational material and a repository database of materials relevant to the reporting of RCTs. The site includes many examples from real trials, including all of the examples included in this article. We will continue to add good and bad examples of reporting to the database, and we invite readers to submit further suggestions by contacting us through the website. The CONSORT Group will continue to survey the literature to find relevant articles that address issues relevant to the reporting of RCTs, and we invite authors of any such articles

to notify us about them. All of this information will be made accessible through the CONSORT website, which is updated regularly.

More than 400 leading general and specialty journals and biomedical editorial groups, including the ICMJE, World Association of Medical Journal Editors, and the Council of Science Editors, have given their official support to CONSORT. We invite other journals concerned about the quality of reporting of clinical trials to endorse the CONSORT statement and contact us through our website to let us know of their support. The ultimate benefactors of these collective efforts should be people who, for whatever reason, require intervention from the healthcare community.

Acknowledgments

We are grateful to Frank Davidoff and Tom Lang for their involvement in the 2001 version of CONSORT explanation and elaboration document. A special thanks to Mary Ocampo, the Ottawa CONSORT coordinator, who helped deliver this explanation and elaboration paper and the CONSORT statement.

The CONSORT Group contributors to CONSORT 2010: Douglas G Altman, Centre for Statistics in Medicine, University of Oxford, UK; Virginia Barbour, *PLoS Medicine*, UK; Jesse A Berlin, Johnson & Johnson Pharmaceutical Research and Development, USA; Isabelle Boutron, University Paris 7 Denis Diderot, Assistance Publique des Hôpitaux de Paris, INSERM, France; PJ Devereaux, McMaster University, Canada; Kay Dickersin, Johns Hopkins Bloomberg School of Public Health, USA; Diana Elbourne, London School of Hygiene & Tropical Medicine, UK; Susan Ellenberg, University of Pennsylvania School of Medicine, USA; Val Gebski, University of Sydney, Australia; Steven Goodman, *Journal of the Society for Clinical Trials*, USA; Peter C Gøtzsche, Nordic Cochrane Centre, Denmark; Trish Groves, *BMJ*, UK; Steven Grunberg, American Society of Clinical Oncology, USA; Brian Haynes, McMaster University, Canada; Sally Hopewell, Centre for Statistics in Medicine, University of Oxford, UK; Astrid James, *Lancet*; Peter Juhn, Johnson & Johnson, USA; Philippa Middleton, University of Adelaide, Australia; Don Minckler, University of California Irvine, USA; David Moher, Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Canada; Victor M Montori, Knowledge and Encounter Research Unit, Mayo Clinic College of Medicine, USA; Cynthia Mulrow, *Annals of Internal Medicine*, USA; Stuart Pocock, London School of Hygiene & Tropical Medicine, UK; Drummond Rennie, *JAMA*, USA; David L Schriger, *Annals of Emergency Medicine*, USA; Kenneth F Schulz, Family Health International, USA; Iveta Simera, EQUATOR Network, UK; Elizabeth Wager, Sideview, UK.

Funding: We gratefully acknowledge financial support from United Kingdom National Institute for Health Research; Canadian Institutes of Health Research; Presidents Fund, Canadian Institutes of Health Research; Johnson & Johnson; *BMJ*; and the American Society for Clinical Oncology.

References

[1] Rennie D. CONSORT revised—improving the reporting of randomized trials. JAMA 2001;285:2006−7.
[2] Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273: 408−12.
[3] Moher D. CONSORT: an evolving tool to help improve the quality of reports of randomized controlled trials. Consolidated Standards of Reporting Trials. JAMA 1998;279:1489−91.
[4] Kjaergard LL, Villumsen J, Gluud C. Quality of randomised clinical trials affects estimates of intervention efficacy. Rome, Italy: 7th Cochrane Colloquium; 1999.
[5] Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ 2001;323:42−6.
[6] Veldhuyzen van Zanten SJ, Cleary C, Talley NJ, Peterson TC, Nyren O, Bradley LA, et al. Drug treatment of functional dyspepsia: a systematic analysis of trial methodology with recommendations for design of future trials. Am J Gastroenterol 1996;91:660−73.
[7] Talley NJ, Owen BK, Boyce P, Paterson K. Psychological treatments for irritable bowel syndrome: a critique of controlled treatment trials. Am J Gastroenterol 1996;91:277−83.
[8] Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? Arch Dermatol 2000;136: 381−5.
[9] Kjaergard LL, Nikolova D, Gluud C. Randomized clinical trials in HEPATOLOGY: predictors of quality. Hepatology 1999;30:1134−8.
[10] Schor S, Karten I. Statistical evaluation of medical journal manuscripts. JAMA 1966;195:1123−8.
[11] Gore SM, Jones IG, Rytter EC. Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976. BMJ 1977;1:85−7.
[12] Hall JC, Hill D, Watts JM. Misuse of statistical methods in the Australasian surgical literature. Aust N Z J Surg 1982;52:541−3.
[13] Altman DG. Statistics in medical journals. Stat Med 1982;1:59−71.
[14] Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N Engl J Med 1987;317:426−32.
[15] Altman DG. The scandal of poor medical research. BMJ 1994;308: 283−4.
[16] Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. Lancet 2005;365:1159−62.
[17] Hopewell S, Dutton S, Yu LM, Chan AW, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. BMJ 2010;340. c723.
[18] Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;319: 670−4.
[19] Lai TY, Wong VW, Lam RF, Cheng AC, Lam DS, Leung GM. Quality of reporting of key methodological items of randomized controlled trials in clinical ophthalmic journals. Ophthalmic Epidemiol 2007;14:390−8.
[20] Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le LJ, et al. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. Lancet 1996;347:363−6.
[21] Junker CA. Adherence to published standards of reporting: a comparison of placebo-controlled trials published in English or German. JAMA 1998;280:247−9.

[22] Altman DG. Randomisation. BMJ 1991;302:1481−2.
[23] Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. JAMA 1994;272:125−8.
[24] Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. BMJ 1948;2:769−82.
[25] Schulz KF. Randomized controlled trials. Clin Obstet Gynecol 1998;41:245−56.
[26] Greenland S. Randomization, statistics, and causal inference. Epidemiology 1990;1:421−9.
[27] Armitage P. The role of randomization in clinical trials. Stat Med 1982;1:345−52.
[28] Kleijnen J, Gotzsche PC, Kunz R, Oxman AD, Chalmers I. So what's so special about randomisation. In: Maynard A, Chalmers I, editors. Non-random reflections on health services research. BMJ Books; 1997. p. 93−106.
[29] Chalmers I. Assembling comparison groups to assess the effects of health care. J R Soc Med 1997;90:379−86.
[30] Nicolucci A, Grilli R, Alexanian AA, Apolone G, Torri V, Liberati A. Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer. A lost opportunity for meta-analysis. JAMA 1989;262:2101−7.
[31] Ah-See KW, Molony NC. A qualitative assessment of randomized controlled trials in otolaryngology. J Laryngol Otol 1998;112:460−3.
[32] Altman DG, Dore CJ. Randomisation and baseline comparisons in clinical trials. Lancet 1990;335:149−53.
[33] Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. BMJ 1998;317:1181−4.
[34] DerSimonian R, Charette LJ, McPeek B, Mosteller F. Reporting on methods in clinical trials. N Engl J Med 1982;306:1332−7.
[35] A proposal for structured reporting of randomized controlled trials. The Standards of Reporting Trials Group. JAMA 1994;272:1926−31.
[36] Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature. Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Ann Intern Med 1994;121:894−5.
[37] Rennie D. Reporting randomized controlled trials. An experiment and a call for responses from readers. JAMA 1995;273:1054−5.
[38] Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. JAMA 1996;276:637−9.
[39] Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. JAMA 2006;295:1152−60.
[40] Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. BMJ 2004;328:702−8.
[41] Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337. a2390.
[42] Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004;141: 781−8.
[43] Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 2008;148: 295−309.
[44] Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. Ann Intern Med 2006;144:364−7.
[45] Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med 2008;5:e20.
[46] Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. JAMA 1996;276:1339−41.
[47] Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ 1996;313:275−83.
[48] Lang TA, Secic M. How to report statistics in medicine. Annotated guidelines for authors, editors, and reviewers. ACP 1997.
[49] Davidoff F. News from the International Committee of Medical Journal Editors. Ann Intern Med 2000;133:229−31.
[50] Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. Med J Aust 2006;185:263−7.
[51] Egger M, Juni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. JAMA 2001;285:1996−9.
[52] Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann Intern Med 2001;134:657−62.
[53] Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 2001;285:1987−91.
[54] Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001;357:1191−4.
[55] Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457−65.
[56] Al-Marzouki S, Roberts I, Evans S, Marshall T. Selective reporting in clinical trials: analysis of trial protocols accepted by the Lancet. Lancet 2008;372:201.
[57] Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS ONE 2008;3:e3081.
[58] Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001;134:663−94.
[59] Schulz KF, Altman DG, Moher Dfor the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.
[60] Pocock SJ. Clinical trials: a practical approach. John Wiley; 1983.
[61] Meinert CL. Clinical trials: design, conduct and analysis. Oxford University Press; 1986.
[62] Friedman LM, Furberg CD, DeMets DL. Fundamentals of clinical trials. 3rd ed. Springer; 1998.
[63] Bolliger CT, Zellweger JP, Danielsson T, van Biljon X, Robidou A, Westin A, et al. Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety. BMJ 2000;321:329−33.
[64] Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. Eval Health Prof 2002;25:38−64.
[65] Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. Lancet 2008;371:281−3.
[66] The impact of open access upon public health. PLoS Med 2006;3: e252.
[67] Harbourt AM, Knecht LS, Humphreys BL. Structured abstracts in MEDLINE, 1989-1991. Bull Med Libr Assoc 1995;83:190−5.
[68] Harris AH, Standard S, Brunning JL, Casey SL, Goldberg JH, Oliver L, et al. The accuracy of abstracts in psychology journals. J Psychol 2002;136:141−8.
[69] Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. JAMA 1999;281:1110−1.
[70] Ward LG, Kendrach MG, Price SO. Accuracy of abstracts for original research articles in pharmacy journals. Ann Pharmacother 2004;38:1173−7.

[71] Gotzsche PC. Believability of relative risks and odds ratios in abstracts: cross sectional study. BMJ 2006;333:231−4.
[72] Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. JAMA 2001;285:437−43.
[73] Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. Ann Intern Med 1990;113: 69−76.
[74] Taddio A, Pain T, Fassos FF, Boon H, Ilersich AL, Einarson TR. Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association. CMAJ 1994;150:1611−5.
[75] Wager E, Middleton P. Technical editing of research reports in biomedical journals. Cochrane Database Syst Rev 2008;. MR000002.
[76] Hartley J, Sydes M, Blurton A. Obtaining information accurately and quickly: Are structured abstracts more efficient? J Inform Sci 1996;22:349−56.
[77] Gilligan D, Nicolson M, Smith I, Groen H, Dalesio O, Goldstraw P, et al. Preoperative chemotherapy in patients with resectable non-small cell lung cancer: results of the MRC LU22/NVALT 2/EORTC 08012 multicentre randomised trial and update of systematic review. Lancet 2007;369:1929−37.
[78] Sandler AD, Sutton KA, DeWeese J, Girardi MA, Sheppard V, Bodfish JW. Lack of benefit of a single dose of synthetic human secretin in the treatment of autism and pervasive developmental disorder. N Engl J Med 1999;341:1801−6.
[79] World Medical Association. Declaration of Helsinki: ethical principle for medical research involving human subjects. 59th WMA General Assembly, Seoul 2008; www.wma.net/e/policy/b3.htm (accessed 2 June 2009).
[80] Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. N Engl J Med 1992;327:248−54.
[81] Fergusson D, Glass KC, Hutton B, Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? Clin Trials 2005;2:218−29.
[82] Savulescu J, Chalmers I, Blunt J. Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. BMJ 1996;313:1390−3.
[83] Sadler LC, Davison T, McCowan LM. A randomised controlled trial and meta-analysis of active management of labour. BJOG 2000;107: 909−15.
[84] Bath FJ, Owen VE, Bath PM. Quality of full and final publications reporting acute stroke trials: a systematic review. Stroke 1998;29: 2203−10.
[85] Blumer JL, Findling RL, Shih WJ, Soubrane C, Reed MD. Controlled clinical trial of zolpidem for the treatment of insomnia associated with attention-deficit/hyperactivity disorder in children 6 to 17 years of age. Pediatrics 2009;123:e770−6.
[86] Sabatine MS, Antman EM, Widimsky P, Ebrahim IO, Kiss RG, Saaiman A, et al. Otamixaban for the treatment of patients with non-ST-elevation acute coronary syndromes (SEPIA-ACS1 TIMI 42): a randomised, double-blind, active-controlled, phase 2 trial. Lancet 2009;374:787−95.
[87] Grant AM, Altman DG, Babiker AB, Campbell MK, Clemens FJ, Darbyshire JH, et al. Issues in data monitoring and interim analysis of trials. Health Technol Assess 2005;9:1-iv.
[88] Gallo P, Krams M. PhRMA Working Group on adaptive designs, ''White Paper.'' Drug Information Journal 2006;40:421−82.
[89] Brown CH, Ten Have TR, Jo B, Dagne G, Wyman PA, Muthen B, et al. Adaptive designs for randomized trials in public health. Annu Rev Public Health 2009;30:1−25.
[90] Kelly PJ, Sooriyarachchi MR, Stallard N, Todd S. A practical comparison of group-sequential and adaptive designs. J Biopharm Stat 2005;15:719−38.
[91] Pildal J, Chan AW, Hrobjartsson A, Forfang E, Altman DG, Gotzsche PC. Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study. BMJ 2005;330:1049.
[92] Chan AW, Hrobjartsson A, Jorgensen KJ, Gotzsche PC, Altman DG. Discrepancies in sample size calculations and data analyses reported in randomised trials: comparison of publications with protocols. BMJ 2008;337:a2299.
[93] Ndekha MJ, van Oosterhout JJ, Zijlstra EE, Manary M, Saloojee H, Manary MJ. Supplementary feeding with either ready-to-use fortified spread or corn-soy blend in wasted adults starting antiretroviral therapy in Malawi: randomised, investigator blinded, controlled trial. BMJ 2009;338:1867−75.
[94] Rothwell PM. External validity of randomised controlled trials: ''to whom do the results of this trial apply?'' Lancet 2005;365: 82−93.
[95] Fuks A, Weijer C, Freedman B, Shapiro S, Skrutkowska M, Riaz A. A study in contrasts: eligibility criteria in a twenty-year sample of NSABP and POG clinical trials. National Surgical Adjuvant Breast and Bowel Program. Pediatric Oncology Group. J Clin Epidemiol 1998;51:69−79.
[96] Shapiro SH, Weijer C, Freedman B. Reporting the study populations of clinical trials. Clear transmission or static on the line? J Clin Epidemiol 2000;53:973−9.
[97] Gandhi M, Ameli N, Bacchetti P, Sharp GB, French AL, Young M, et al. Eligibility criteria for HIV clinical trials and generalizability of results: the gap between published reports and study protocols. AIDS 2005;19:1885−96.
[98] Hall JC, Mills B, Nguyen H, Hall JL. Methodologic standards in surgical trials. Surgery 1996;119:466−72.
[99] Weiss NS, Koepsell TD, Psaty BM. Generalizability of the results of randomized trials. Arch Intern Med 2008;168:133−5.
[100] Devereaux PJ, Yang H, Yusuf S, Guyatt G, Leslie K, Villar JC, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. Lancet 2008;371:1839−47.
[101] Rannou F, Dimet J, Boutron I, Baron G, Fayad F, Mace Y, et al. Splint for base-of-thumb osteoarthritis: a randomized trial. Ann Intern Med 2009;150:661−9.
[102] Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? BMJ 2008;336:1472−4.
[103] Mease PJ, Goffe BS, Metz J, VanderStoep A, Finck B, Burge DJ. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. Lancet 2000;356:385−90.
[104] McDowell I, Newell C. Measuring health: a guide to rating scales and questionnaires. 3rd ed. New York: Oxford University Press; 2006.
[105] Streiner D, Norman C. Health measurement scales: a practical guide to their development and use. 3rd ed. Oxford: Oxford University Press; 2003.
[106] Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials 2007;8:39.
[107] Sanders C, Egger M, Donovan J, Tallon D, Frankel S. Reporting on quality of life in randomised controlled trials: bibliographic study. BMJ 1998;317:1191−4.
[108] Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. Control Clin Trials 1989;10:31−56.
[109] Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. Br J Psychiatry 2000;176:249−52.
[110] Jadad AR, Boyle M, Cunnigham C, Kim M, Schachar R. Treatment of Attention-Deficit/Hyperactivity Disorder. Evidence Report/Technology Assessment No. 11. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Healthcare Research and Quality. AHQR publication no. 00-E005; 1999.

[111] Schachter HM, Pham B, King J, Langford S, Moher D. The efficacy and safety of methylphenidate in attention deficit disorder: A systematic review and meta-analyis. Prepared for the Therapeutics Inititiative, Vancouver, B.C., and the British Columbia Ministry for Children and Families, 2000.
[112] Dargie HJ. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. Lancet 2001;357:1385−90.
[113] Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ 2004;171:735−40.
[114] Vermeulen H, Hofland J, Legemate DA, Ubbink DT. Intravenous fluid restriction after major abdominal surgery: a randomized blinded clinical trial. Trials 2009;10:50.
[115] Fox K, Ford I, Steg PG, Tendera M, Ferrari R. Ivabradine for patients with stable coronary artery disease and left-ventricular systolic dysfunction (BEAUTIFUL): a randomised, double-blind, placebo-controlled trial. Lancet 2008;372:807−16.
[116] Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. BMJ 1995;311:1145−8.
[117] Guyatt GH, Mills EJ, Elbourne D. In the era of systematic reviews, does the size of an individual trial still matter. PLoS Med 2008;5:e4.
[118] Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. Lancet 2005;365:1348−53.
[119] Halpern SD, Karlawish JH, Berlin JA. The continuing unethical conduct of underpowered clinical trials. JAMA 2002;288:358−62.
[120] Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ 1995;311:485.
[121] Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994;272:122−4.
[122] Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. N Engl J Med 1978;299:690−4.
[123] Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials: review. BMJ 2009;338:b1732.
[124] Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? Stat Med 1984;3:409−22.
[125] Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. Ann Intern Med 1994;121:200−6.
[126] Galgiani JN, Catanzaro A, Cloud GA, Johnson RH, Williams PL, Mirels LF, et al. Comparison of oral fluconazole and itraconazole for progressive, nonmeningeal coccidioidomycosis. A randomized, double-blind trial. Mycoses Study Group. Ann Intern Med 2000;133:676−86.
[127] Connolly SJ, Pogue J, Hart RG, Hohnloser SH, Pfeffer M, Chrolavicius S, et al. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. N Engl J Med 2009;360:2066−78.
[128] Geller NL, Pocock SJ. Interim analyses in randomized clinical trials: ramifications and guidelines for practitioners. Biometrics 1987;43:213−23.
[129] Berry DA. Interim analyses in clinical trials: classical vs. Bayesian approaches. Stat Med 1985;4:521−6.
[130] Pocock SJ. When to stop a clinical trial. BMJ 1992;305:235−40.
[131] DeMets DL, Pocock SJ, Julian DG. The agonising negative trend in monitoring of clinical trials. Lancet 1999;354:1983−8.
[132] Buyse M. Interim analyses, stopping rules and data monitoring in clinical trials in Europe. Stat Med 1993;12:509−20.
[133] Sydes MR, Altman DG, Babiker AB, Parmar MK, Spiegelhalter DJ. Reported use of data monitoring committees in the main published reports of randomized controlled trials: a cross-sectional study. Clin Trials 2004;1:48−59.
[134] Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, et al. Randomized trials stopped early for benefit: a systematic review. JAMA 2005;294:2203−9.
[135] Coutinho IC, Ramos de Amorim MM, Katz L, Bandeira de Ferraz AA. Uterine exteriorization compared with in situ repair at cesarean delivery: a randomized controlled trial. Obstet Gynecol 2008;111:639−47.
[136] Juni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. In: Egger M, Davey Smith G, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. BMJ Books; 2001.
[137] Creinin MD, Meyn LA, Borgatta L, Barnhart K, Jensen J, Burke AE, et al. Multicenter comparison of the contraceptive ring and patch: a randomized controlled trial. Obstet Gynecol 2008;111:267−77.
[138] Tate DF, Jackvony EH, Wing RR. Effects of internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. JAMA 2003;289:1833−6.
[139] Lachin JM. Properties of simple randomization in clinical trials. Control Clin Trials 1988;9:312−26.
[140] Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 1976;34:585−612.
[141] Schulz KF, Grimes DA. The Lancet handbook of essential concepts in clinical research. Elsevier; 2006.
[142] Altman DG, Bland JM. How to randomise. BMJ 1999;319:703−4.
[143] Schulz KF. Subverting randomization in controlled trials. JAMA 1995;274:1456−8.
[144] Enas GG, Enas NH, Spradlin CT, Wilson MG, Wiltse CG. Baseline comparability in clinical trials: prevention of poststudy anxiety. Drug Information Journal 1990;24:541−8.
[145] Treasure T, MacRae KD. Minimisation: the platinum standard for trials? Randomisation doesn't guarantee similarity of groups; minimisation does. BMJ 1998;317:362−3.
[146] Sinei SK, Schulz KF, Lamptey PR, Grimes DA, Mati JK, Rosenthal SM, et al. Preventing IUCD-related pelvic infection: the efficacy of prophylactic doxycycline at insertion. Br J Obstet Gynaecol 1990;97:412−9.
[147] Radford JA, Landorf KB, Buchbinder R, Cook C. Effectiveness of low-Dye taping for the short-term treatment of plantar heel pain: a randomised trial. BMC Musculoskelet Disord 2006;7:64.
[148] Chalmers TC, Levin H, Sacks HS, Reitman D, Berrier J, Nagalingam R. Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials. Stat Med 1987;6:315−28.
[149] Pocock SJ. Statistical aspects of clinical trial design. Statistician 1982;31:1−18.
[150] Haag U. Technologies for automating randomized treatment assignment in clinical trials. Drug Information Journal 1998;32:11.
[151] Piaggio G, Elbourne D, Schulz KF, Villar J, Pinol AP, Gulmezoglu AM. The reporting of methods for reducing and detecting bias: an example from the WHO Misoprostol Third Stage of Labour equivalence randomised controlled trial. BMC Med Res Methodol 2003;3:19.
[152] Pildal J, Hrobjartsson A, Jorgensen KJ, Hilden J, Altman DG, Gotzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. Int J Epidemiol 2007;36:847−57.
[153] Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008;336:601−5.
[154] McCandlish R, Bowler U, van Asten H, Berridge G, Winter C, Sames L, et al. A randomised controlled trial of care of the perineum during second stage of normal labour. Br J Obstet Gynaecol 1998;105:1262−72.

[155] Webster J, Clarke S, Paterson D, Hutton A, van Dyk S, Gale C, et al. Routine care of peripheral intravenous catheters versus clinically indicated replacement: randomised controlled trial. BMJ 2008;337: a339.

[156] Smith SA, Shah ND, Bryant SC, Christianson TJ, Bjornsen SS, Giesler PD, et al. Chronic care model and shared care in diabetes: randomized trial of an electronic decision support system. Mayo Clin Proc 2008;83:747−57.

[157] Sacks FM, Bray GA, Carey VJ, Smith SR, Ryan DH, Anton SD, et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. N Engl J Med 2009;360: 859−73.

[158] Kaptchuk TJ. Intentional ignorance: a history of blind assessment and placebo controls in medicine. Bull Hist Med 1998;72:389−433.

[159] Guyatt GH, Pugsley SO, Sullivan MJ, Thompson PJ, Berman L, Jones NL, et al. Effect of encouragement on walking test performance. Thorax 1984;39:818−22.

[160] Gotzsche PC. Blinding during data analysis and writing of manuscripts. Control Clin Trials 1996;17:285−90.

[161] Karlowski TR, Chalmers TC, Frenkel LD, Kapikian AZ, Lewis TL, Lynch JM. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. JAMA 1975;231:1038−42.

[162] Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. Neurology 1994;44:16−20.

[163] Carley SD, Libetta C, Flavin B, Butler J, Tong N, Sammy I. An open prospective randomised trial to reduce the pain of blood glucose testing: ear versus thumb. BMJ 2000;321:20.

[164] Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. Ann Intern Med 2002;136:254−9.

[165] Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. BMJ 2000;321:504.

[166] Montori VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH. In the dark: the reporting of blinding status in randomized controlled trials. J Clin Epidemiol 2002;55:787−90.

[167] Cheng K, Smyth RL, Motley J, O'Hea U, Ashby D. Randomized controlled trials in cystic fibrosis (1966-1997) categorized by time, design, and intervention. Pediatr Pulmonol 2000;29:1−7.

[168] Lang T. Masking or blinding? An unscientific survey of mostly medical journal editors on the great debate. Med Gen Med 2000;2:E25.

[169] Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. JAMA 2001;285:2000−3.

[170] Haahr MT, Hrobjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. Clin Trials 2006;3:360−5.

[171] Meinert CL. Masked monitoring in clinical trials—blind stupidity? N Engl J Med 1998;338:1381−2.

[172] Mills E, Prousky J, Raskin G, Gagnier J, Rachlis B, Montori VM, et al. The safety of over-the-counter niacin. A randomized placebo-controlled trial [[ISRCTN18054903]]. BMC Clin Pharmacol 2003;3:4.

[173] Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. BMJ 1996;312:742−4.

[174] Fergusson D, Glass KC, Waring D, Shapiro S. Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. BMJ 2004;328:432.

[175] Sackett DL. Turning a blind eye: why we don't test for blindness at the end of our trials. BMJ 2004;328:1136.

[176] Astrup A, Rossner S, Van Gaal L, Rissanen A, Niskanen L, Al HM, et al. Effects of liraglutide in the treatment of obesity: a randomised, double-blind, placebo-controlled study. Lancet 2009;374:1606−16.

[177] Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. In: Altman DG, Machin D, Bryant TN, Gardner MJ, editors. Statistics with confidence: confidence intervals and statistical guidelines. 2nd ed. BMJ Books; 2000. p. 171−90.

[178] Altman DG, Bland JM. Statistics notes. Units of analysis. BMJ 1997;314:1874.

[179] Bolton S. Independence and statistical inference in clinical trial designs: a tutorial review. J Clin Pharmacol 1998;38:408−12.

[180] Greenland S. Principles of multilevel modelling. Int J Epidemiol 2000;29:158−67.

[181] Albert CM, Cook NR, Gaziano JM, Zaharris E, MacFadyen J, Danielson E, et al. Effect of folic acid and B vitamins on risk of cardiovascular events and total mortality among women at high risk for cardiovascular disease: a randomized trial. JAMA 2008;299: 2027−36.

[182] Matthews JN, Altman DG. Interaction 3: How to examine heterogeneity. BMJ 1996;313:862.

[183] Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet 2000;355:1064−9.

[184] Matthews JN, Altman DG. Statistics notes. Interaction 2: Compare effect sizes not P values. BMJ 1996;313:808.

[185] Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med 1992;116:78−84.

[186] Steyerberg EW, Bossuyt PM, Lee KL. Clinical trials in acute myocardial infarction: should we adjust for baseline characteristics? Am Heart J 2000;139:745−51.

[187] Altman DG. Adjustment for covariate imbalance. In: Armitage P, Colton T, editors. Encyclopedia of biostatistics. John Wiley; 1998. p. 1000−5.

[188] Mullner M, Matthews H, Altman DG. Reporting on statistical methods to adjust for confounding: a cross-sectional survey. Ann Intern Med 2002;136:122−6.

[189] Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. Ann Intern Med 1993;118:201−10.

[190] Bender R, Grouven U. Logistic regression models used in medical research are poorly presented. BMJ 1996;313:628.

[191] Khan KS, Chien PF, Dwarakanath LS. Logistic regression models in obstetrics and gynecology literature. Obstet Gynecol 1999;93: 1014−20.

[192] Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. N Engl J Med 1979;301:1410−2.

[193] May GS, DeMets DL, Friedman LM, Furberg C, Passamani E. The randomized clinical trial: bias in analysis. Circulation 1981;64: 669−73.

[194] Altman DG, Cuzick J, Peto J. More on zidovudine in asymptomatic HIV infection. N Engl J Med 1994;330:1758−9.

[195] Meinert CL. Beyond CONSORT: need for improved reporting standards for clinical trials. Consolidated Standards of Reporting Trials. JAMA 1998;279:1487−9.

[196] Grant AM, Wileman SM, Ramsay CR, Mowat NA, Krukowski ZH, Heading RC, et al. Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomised trial. BMJ 2008;337:a2664.

[197] van Loon AJ, Mantingh A, Serlier EK, Kroon G, Mooyaart EL, Huisjes HJ. Randomised controlled trial of magnetic-resonance pelvimetry in breech presentation at term. Lancet 1997;350: 1799−804.

[198] Brown MJ, Palmer CR, Castaigne A, de Leeuw PW, Mancia G, Rosenthal T, et al. Morbidity and mortality in patients randomised to double-blind treatment with a long-acting calcium-channel blocker or diuretic in the International Nifedipine GITS study: Intervention as a Goal in Hypertension Treatment (INSIGHT). Lancet 2000;356:366−72.

[199] LaCroix AZ, Ott SM, Ichikawa L, Scholes D, Barlow WE. Low--dose hydrochlorothiazide and preservation of bone mineral density

in older adults. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2000;133:516−26.
[200] Shuster JJ. Median follow-up in clinical trials. J Clin Oncol 1991;9: 191−2.
[201] Altman DG, de Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. Br J Cancer 1995;72: 511−8.
[202] Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, Puren A. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. PLoS Med 2005;2:e298.
[203] Diggle L, Deeks J. Effect of needle length on incidence of local reactions to routine immunisation in infants aged 4 months: randomised controlled trial. BMJ 2000;321:931−3.
[204] Pocock S, White I. Trials stopped early: too good to be true? Lancet 1999;353:943−4.
[205] Hughes MD, Pocock SJ. Stopping rules and estimation problems in clinical trials. Stat Med 1988;7:1231−42.
[206] Kiri A, Tonascia S, Meinert CL. Treatment effects monitoring committees and early stopping in large clinical trials. Clin Trials 2004;1: 40−7.
[207] Psaty BM, Rennie D. Stopping medical research to save money: a broken pact with researchers and patients. JAMA 2003;289: 2128−31.
[208] Temple R. FDA perspective on trials with interim efficacy evaluations. Stat Med 2006;25:3245−9.
[209] Yusuf S, Teo K, Anderson C, Pogue J, Dyal L, Copland I, et al. Effects of the angiotensin-receptor blocker telmisartan on cardiovascular events in high-risk patients intolerant to angiotensin-converting enzyme inhibitors: a randomised controlled trial. Lancet 2008;372:1174−83.
[210] Senn S. Base logic: tests of baseline balance in randomized clinical trials. Clin Res Regulatory Affairs 1995;12:171−82.
[211] Altman DG. Comparability of randomised groups. Statistician 1985;34:125−36.
[212] Heit JA, Elliott CG, Trowbridge AA, Morrey BF, Gent M, Hirsh J. Ardeparin sodium for extended out-of-hospital prophylaxis against venous thromboembolism after total hip or knee replacement. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2000;132:853−61.
[213] Haderslev KV, Tjellesen L, Sorensen HA, Staun M. Alendronate increases lumbar spine bone mineral density in patients with Crohn's disease. Gastroenterology 2000;119:639−46.
[214] Fields WS, Maslenikov V, Meyer JS, Hass WK, Remington RD, Macdonald M. Joint study of extracranial arterial occlusion. V. Progress report of prognosis following surgery or nonsurgical treatment for transient cerebral ischemic attacks and cervical carotid artery lesions. JAMA 1970;211:1993−2003.
[215] Lee YJ, Ellenberg JH, Hirtz DG, Nelson KB. Analysis of clinical trials by treatment actually received: is it really an option? Stat Med 1991;10:1595−605.
[216] Lewis JA, Machin D. Intention to treat—who should use ITT? Br J Cancer 1993;68:647−50.
[217] Lachin JL. Statistical considerations in the intent-to-treat principle. Control Clin Trials 2000;21:526.
[218] Sheiner LB, Rubin DB. Intention-to-treat analysis and the goals of clinical trials. Clin Pharmacol Ther 1995;57:6−15.
[219] Nagelkerke N, Fidler V, Bernsen R, Borgdorff M. Estimating treatment effects in randomized clinical trials in the presence of noncompliance. Stat Med 2000;19:1849−64.
[220] Melander H, hlqvist-Rastad J, Meijer G, Beermann B. Evidence b(i)ased medicine−selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. BMJ 2003;326:1171−3.
[221] Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? Clin Trials 2007;4:350−6.
[222] Kruse RL, Alper BS, Reust C, Stevermer JJ, Shannon S, Williams RH. Intention-to-treat analysis: who is in? Who is out? J Fam Pract 2002;51:969−71.
[223] Herman A, Botser IB, Tenenbaum S, Chechick A. Intention-to-treat analysis and accounting for missing data in orthopaedic randomized clinical trials. J Bone Joint Surg Am 2009;91:2137−43.
[224] Ruiz-Canela M, Martinez-Gonzalez MA, de Irala-Estevez J. Intention to treat analysis is related to methodological quality. BMJ 2000;320:1007−8.
[225] Altman DG. Missing outcomes in randomised trials: addressing the dilemma. Open Med 2009;3:e21−3.
[226] Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clin Trials 2004;1: 368−76.
[227] Streiner DL. Missing data and the trouble with LOCF. Evid Based Ment Health 2008;11:3−5.
[228] Molnar FJ, Hutton B, Fergusson D. Does analysis using "last observation carried forward" introduce bias in dementia research? CMAJ 2008;179:751−3.
[229] Ware JH. Interpreting incomplete data in studies of diet and weight loss. N Engl J Med 2003;348:2136−7.
[230] Streiner DL. The case of the missing data: methods of dealing with dropouts and other research vagaries. Can J Psychiatry 2002;47: 68−75.
[231] Lane P. Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. Pharm Stat 2008;7: 93−106.
[232] Abraha I, Montedori A, Romagnoli C. Modified intention to treat: frequency, definition and implication for clinical trials [abstract]. Sao Paulo, Brazil: XV Cochrane Colloquium; 2007:86-7.
[233] Altman DG. Confidence intervals in practice. In: Altman DG, Machin D, Bryant TN, Gardner MJ, editors. Statistics with confidence. 2nd ed. BMJ Books; 2000. p. 6−14.
[234] van Linschoten R, van Middelkoop M, Berger MY, Heintjes EM, Verhaar JA, Willemsen SP, et al. Supervised exercise therapy versus usual care for patellofemoral pain syndrome: an open label randomised controlled trial. BMJ 2009;339:b4074.
[235] Altman DG. Clinical trials and meta-analyses. In: Altman DG, Machin D, Bryant TN, Gardner MJ, editors. Statistics with confidence. 2nd ed. BMJ Books; 2000. p. 120−38.
[236] Uniform requirements for manuscripts submitted to biomedical journals. International Committee of Medical Journal Editors. Ann Intern Med 1997;126:36−47.
[237] Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. BMJ 1986;292:746−50.
[238] Bailar JC III, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. Ann Intern Med 1988;108:266−73.
[239] Bland JM. Quoting intermediate analyses can only mislead. BMJ 1997;314:1907−8.
[240] Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ 1995;310:452−4.
[241] Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. BMJ 1999;319: 1492−5.
[242] The OSIRIS Collaborative Group. Early versus delayed neonatal administration of a synthetic surfactant—the judgment of OSIRIS (open study of infants at high risk of or with respiratory insufficiency—the role of surfactant). Lancet 1992;340:1363−9.
[243] Sorensen L, Gyrd-Hansen D, Kristiansen IS, Nexoe J, Nielsen JB. Laypersons' understanding of relative risk reductions: randomised cross-sectional study. BMC Med Inform Decis Mak 2008;8:31.
[244] Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians' willingness to prescribe. Lancet 1994;343:1209−11.

[245] Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? Ann Intern Med 1992;117:916−21.
[246] Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. Science 1977;198:679−84.
[247] Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. JAMA 1991;266:93−8.
[248] Hahn S, Williamson PR, Hutton JL, Garner P, Flynn EV. Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. Stat Med 2000;19:3325−36.
[249] Bhandari M, Devereaux PJ, Li P, Mah D, Lim K, Schunemann HJ, et al. Misuse of baseline comparison tests and subgroup analyses in surgical trials. Clin Orthop Relat Res 2006;447:247−51.
[250] Levin M, Quint PA, Goldstein B, Barton P, Bradley JS, Shemie SD, et al. Recombinant bactericidal/permeability-increasing protein (rBPI21) as adjunctive treatment for children with severe meningococcal sepsis: a randomised trial. rBPI21 Meningococcal Sepsis Study Group. Lancet 2000;356:961−7.
[251] Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. J Clin Oncol 2006;24:3933−8.
[252] Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. Arch Intern Med 2009;169:1756−61.
[253] Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, Kostuk WJ, et al. Optimal medical therapy with or without PCI for stable coronary disease. N Engl J Med 2007;356:1503−16.
[254] Horton R. The rhetoric of research. BMJ 1995;310:985−7.
[255] Annals of Internal Medicine. Information for authors. Available at www.annals.org (accessed 15 Jan 2008).
[256] Docherty M, Smith R. The case for structuring the discussion of scientific papers. BMJ 1999;318:1224−5.
[257] Purcell GP, Donovan SL, Davidoff F. Changes to manuscripts during the editorial process: characterizing the evolution of a clinical paper. JAMA 1998;280:227−8.
[258] Ioannidis JP. Limitations are not properly acknowledged in the scientific literature. J Clin Epidemiol 2007;60:324−9.
[259] Kiviluoto T, Siren J, Luukkonen P, Kivilaakso E. Randomised trial of laparoscopic versus open cholecystectomy for acute and gangrenous cholecystitis. Lancet 1998;351:321−5.
[260] Hupperets MD, Verhagen EA, van Mechelen W. Effect of unsupervised home based proprioceptive training on recurrences of ankle sprain: randomised controlled trial. BMJ 2009;339. b2684.
[261] Garber J, Clarke GN, Weersing VR, Beardslee WR, Brent DA, Gladstone TR, et al. Prevention of depression in at-risk adolescents: a randomized controlled trial. JAMA 2009;301:2215−24.
[262] Campbell D. Factors relevant to the validity of experiments in social settings. Psychol Bull 1957;54:297−312.
[263] Rothwell PM. Factors that can affect the external validity of randomised controlled trials. PLoS Clin Trials 2006;1:e9.
[264] King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, et al. Conceptual framework and systematic review of the effects of participants' and professionals' preferences in randomised controlled trials. Health Technol Assess 2005;9:1-iv.
[265] Djulbegovic B, Lacevic M, Cantor A, Fields KK, Bennett CL, Adams JR, et al. The uncertainty principle and industry-sponsored research. Lancet 2000;356:635−8.
[266] Dans AL, Dans LF, Guyatt GH, Richardson S. Users' guides to the medical literature: XIV. How to decide on the applicability of clinical trial results to your patient. Evidence-Based Medicine Working Group. JAMA 1998;279:545−9.
[267] Smith GD, Egger M. Who benefits from medical interventions? BMJ 1994;308:72−4.
[268] McAlister FA. Applying the results of systematic reviews at the bedside. In: Egger M, Davey Smith G, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. BMJ Books; 2001.
[269] Bartlett C, Doyal L, Ebrahim S, Davey P, Bachmann M, Egger M, et al. The causes and effects of socio-demographic exclusions from clinical trials. Health Technol Assess 2005;9:iii-x, 1.
[270] Bonell C, Oakley A, Hargreaves J, Strange V, Rees R. Assessment of generalisability in trials of health interventions: suggested framework and systematic review. BMJ 2006;333:346−9.
[271] Bornhoft G, Maxion-Bergemann S, Wolf U, Kienle GS, Michalsen A, Vollmar HC, et al. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. BMC Med Res Methodol 2006;6:56.
[272] Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med 1988;318:1728−33.
[273] Altman DG. Confidence intervals for the number needed to treat. BMJ 1998;317:1309−12.
[274] Fanaroff AA, Korones SB, Wright LL, Wright EC, Poland RL, Bauer CB, et al. A controlled trial of intravenous immune globulin to reduce nosocomial infections in very-low-birth-weight infants. National Institute of Child Health and Human Development Neonatal Research Network. N Engl J Med 1994;330:1107−13.
[275] Randomised trial of intravenous atenolol among 16 027 cases of suspected acute myocardial infarction: ISIS-1. First International Study of Infarct Survival Collaborative Group. Lancet 1986;2:57−66.
[276] Gotzsche PC, Gjorup I, Bonnen H, Brahe NE, Becker U, Burcharth F. Somatostatin v placebo in bleeding oesophageal varices: randomised trial and meta-analysis. BMJ 1995;310:1495−8.
[277] Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. J R Soc Med 2007;100:187−90.
[278] Goodman SN. Toward evidence-based medical statistics. 1: The P value fallacy. Ann Intern Med 1999;130:995−1004.
[279] Gotzsche PC. Reference bias in reports of drug trials. BMJ 1987;295:654−6.
[280] Papp K, Bissonnette R, Rosoph L, Wasel N, Lynde CW, Searles G, et al. Efficacy of ISA247 in plaque psoriasis: a randomised, multicentre, double-blind, placebo-controlled phase III study. Lancet 2008;371:1337−42.
[281] Dickersin K. How important is publication bias? A synthesis of available data. AIDS Educ Prev 1997;9:15−21.
[282] Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. Health Technol Assess 2000;4:1−115.
[283] Williamson PR, Gamble C. Identification and impact of outcome selection bias in meta-analysis. Stat Med 2005;24:1547−61.
[284] Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. BMJ 1997;315:635−40.
[285] Simes RJ. Publication bias: the case for an international registry of clinical trials. J Clin Oncol 1986;4:1529−41.
[286] Chalmers I. From optimism to disillusion about commitment to transparency in the medico-industrial complex. J R Soc Med 2006;99:337−41.
[287] Tonks A. A clinical trials register for Europe. BMJ 2002;325:1314−5.
[288] Dickersin K, Rennie D. Registering clinical trials. JAMA 2003;290:516−23.
[289] Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. Lancet 2004;363:1341−5.
[290] De Angelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Is this clinical trial fully registered? A statement from the International Committee of Medical Journal Editors. Lancet 2005;365:1827−9.
[291] Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DA. Issues in the registration of clinical trials. JAMA 2007;297:2112−20.
[292] Hopewell S, Altman DG, Moher D, Schulz KF. Endorsement of the CONSORT Statement by high impact factor medical journals:

a survey of journal editors and journal 'instructions to authors'. Trials 2008;9:20.
[293] Russell JA, Walley KR, Singer J, Gordon AC, Hebert PC, Cooper DJ, et al. Vasopressin versus norepinephrine infusion in patients with septic shock. N Engl J Med 2008;358:877−87.
[294] Chan AW, Tetzlaff J, Altman D, Gotzsche PC, Hrobjartsson A, Krleza-Jeric K, et al. The SPIRIT initiative: defining standard protocol items for randomised trials. Oral presentation at the 16th Cochrane Colloquium: Evidence in the era of globalisation; 2008 Oct 3-7; Freiburg, Germany [[abstract]]. Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen 2008;102:27.
[295] Gregson S, Adamson S, Papaya S, Mundondo J, Nyamukapa CA, Mason PR, et al. Impact and process evaluation of integrated community and clinic-based HIV-1 control: a cluster-randomised trial in eastern Zimbabwe. PLoS Med 2007;4:e102.
[296] Shaddy RE, Boucek MM, Hsu DT, Boucek RJ, Canter CE, Mahony L, et al. Carvedilol for children and adolescents with heart failure: a randomized controlled trial. JAMA 2007;298:1171−9.
[297] Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ 2003;326:1167−70.
[298] Kjaergard LL, ls-Nielsen B. Association between competing interests and authors' conclusions: epidemiological study of randomised clinical trials published in the BMJ. BMJ 2002;325:249.
[299] Bero L, Oostvogel F, Bacchetti P, Lee K. Factors associated with findings of published trials of drug-drug comparisons: why some statins appear more efficacious than others. PLoS Med 2007;4:e184.
[300] Sismondo S. Pharmaceutical company funding and its consequences: a qualitative systematic review. Contemp Clin Trials 2008;29:109−13.
[301] ls-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003;290:921−8.
[302] Ross JS, Hill KP, Egilman DS, Krumholz HM. Guest authorship and ghostwriting in publications related to rofecoxib: a case study of industry documents from rofecoxib litigation. JAMA 2008;299:1800−12.
[303] McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials: a systematic review. JAMA 2003;289:2545−53.
[304] Senn S. Crossover trials in clinical research. 2nd ed. Wiley; 2002.
[305] Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, et al. Evaluating non-randomised intervention studies. Health Technol Assess 2003;7. iii-173.
[306] Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials. Cochrane Database Syst Rev 2007;. MR000012.
[307] Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. Lancet 2001;357:373−80.
[308] Schulz KF. Randomised trials, human nature, and reporting guidelines. Lancet 1996;348:596−8.
[309] Murray GD. Promoting good research practice. Stat Methods Med Res 2000;9:17−24.
[310] Narahari SR, Ryan TJ, Aggithaya MG, Bose KS, Prasanna KS. Evidence-based approaches for the Ayurvedic traditional herbal formulations: toward an Ayurvedic CONSORT model. J Altern Complement Med 2008;14:769−76.
[311] Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003;138:40−4.
[312] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007;147:573−7.
[313] Tonino PA, De Bruyne B, Pijls NH, Siebert U, Ikeno F, Veer M, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. N Engl J Med 2009;360:213−24.
Cite this as: *BMJ* 2010;340:c869.