

CONSORT 2010 Spiegazione ed Elaborazione: linee guida aggiornate per il reporting di trial randomizzati a gruppi paralleli

David Moher^{1*}, Sally Hopewell², Kenneth F Schulz³, Victor Montori⁴, Peter C Gøtzsche⁵, P J Devereaux⁶, Diana Elbourne⁷, Matthias Egger⁸, Douglas G Altman²

¹Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa Hospital, Ottawa, Ontario, Canada, K1H 8L6, ²Centre for Statistics in Medicine, University of Oxford, Wolfson College, Oxford, ³Family Health International, Research Triangle Park, NC 27709, USA, ⁴UK Knowledge and Encounter Research Unit, Mayo Clinic, Rochester, MN, USA, ⁵The Nordic Cochrane Centre, Rigshospitalet, Blegdamsvej 9, Copenhagen, Denmark, ⁶McMaster University Health Sciences Centre, Hamilton, Canada, ⁷Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, ⁸Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland

ABSTRACT

Consistenti evidenze scientifiche dimostrano che la qualità del reporting dei trial controllati e randomizzati – randomized controlled trials (RCT) – non è ottimale. In assenza di un reporting chiaro e adeguato, i lettori non possono valutare l'affidabilità e la validità dei risultati di un trial, né sono in grado di identificare le informazioni necessarie alla produzione di revisioni sistematiche. Recenti studi metodologici dimostrano che un reporting e un disegno inadeguati si associano a stime distorte dell'efficacia del trattamento, compromettendo gravemente i RCT, considerati il gold standard per valutare l'efficacia degli interventi sanitari grazie alla loro capacità di minimizzare o evitare i bias.

Al fine di migliorare la qualità del reporting dei RCT un gruppo di ricercatori ed editori ha sviluppato il CONSORT (Consolidated Standards of Reporting Trials) Statement. Pubblicato nel 1996 e aggiornato nel 2001, il CONSORT Statement è costituito da una checklist e da un diagramma di flusso che gli autori possono utilizzare per il reporting di un RCT. Molte delle principali riviste mediche e i più importanti gruppi editoriali internazionali hanno adottato il CONSORT Statement che facilita l'approccio critico e l'interpretazione dei RCT.

Nel corso della revisione 2001 è emerso che una spiegazione ed elaborazione dei singoli item della checklist

del CONSORT avrebbe agevolato ricercatori e altri utenti nella redazione o nella valutazione del reporting dei trial. Di conseguenza, la versione 2001 del CONSORT è stata integrata con un articolo di spiegazione ed elaborazione.

Nel gennaio 2007, in occasione di un meeting di esperti, il CONSORT è stato sottoposto a revisione e pubblicato come CONSORT Statement 2010, che ha migliorato la formulazione e la chiarezza della checklist precedente e include raccomandazioni su argomenti – come il bias di reporting selettivo degli outcome – di cui solo recentemente è stata presa consapevolezza.

Anche il presente documento di spiegazione ed elaborazione, finalizzato a migliorare la comprensione, l'utilizzo e la diffusione del CONSORT Statement, è stato sottoposto ad accurata revisione. Il presente documento illustra il background e il rationale scientifico di ciascun item – nuovo o aggiornato – del CONSORT 2010, fornendo esempi di un adeguato reporting, riferimenti bibliografici a studi sperimentali rilevanti e numerosi esempi di diagrammi di flusso.

Il CONSORT Statement 2010, la presente revisione del documento di spiegazione ed elaborazione e il sito web dedicato (www.consort-statement.org) costituiscono risorse indispensabili per migliorare il reporting dei RCT.

Citazione. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Spiegazione ed Elaborazione: linee guida aggiornate per il reporting di trial randomizzati a gruppi paralleli. Evidence 2012;4(7): e1000024.

Pubblicato 27 Novembre 2012

Copyright. 2012 Moher et al. Questo è un articolo open-access, distribuito con licenza *Creative Commons Attribution*, che ne consente l'utilizzo, la distribuzione e la riproduzione su qualsiasi supporto esclusivamente per fini non commerciali, a condizione di riportare sempre autore e citazione originale.

Fonti di finanziamento. Per i finanziamenti ringraziamo: United Kingdom National Institute for Health Research; Canadian Institutes of Health Research; Presidents Fund, Canadian Institutes of Health Research; Johnson & Johnson; BMJ; American Society for Clinical Oncology.

Provenienza. Tradotto con permesso da: Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869.

* E-mail: dmoher@ohri.ca

“Tutta la medicina dipende dalla trasparenza del reporting dei trial clinici”¹. Infatti, i trial controllati e randomizzati – *randomized controlled trial* (RCT) – ben disegnati e adeguatamente condotti forniscono le migliori prove di efficacia degli interventi sanitari, mentre quelli con metodologia inadeguata sono associati a bias e, in particolare, tendono a sovrastimare l’efficacia dei trattamenti²⁻⁵. I risultati distorti di trial con disegno e reporting inadeguati possono determinare decisioni errate a tutti i livelli: dal trattamento del paziente individuale alle scelte nazionali di politica sanitaria.

Inoltre, la valutazione critica della qualità dei trial è possibile solo se il disegno, la conduzione e l’analisi dei RCT sono descritte in maniera completa e accurata. Lungi dall’essere trasparente, il reporting dei RCT è spesso incompleto⁶⁻⁹, e comporta problemi conseguenti a una metodologia inadeguata¹⁰⁻¹⁵.

Reporting incompleto e impreciso

Molte revisioni hanno documentato carenze nel reporting dei trial clinici: ad esempio, le informazioni sui metodi di assegnazione dei partecipanti a ciascun gruppo erano riportate solo nel 21% di 519 trial indicizzati in PubMed nel 2000¹⁶ e nel 34% dei 616 trial indicizzati nel 2006¹⁷. Analogamente, solo il 45% dei trial indicizzati in PubMed nel 2000¹⁶ e il 53% nel 2006¹⁷ definivano un end-point primario, e solo il 27% nel 2000 e il 45% nel 2006 riportavano le metodologie utilizzate per stimare la dimensione del campione. Il reporting dei trial non solo è spesso incompleto, ma a volte è anche impreciso. Dei 119 trial in cui tutti i partecipanti erano analizzati secondo i gruppi di assegnazione originari (intention-to-treat analysis), 15 (13%) escludevano pazienti dall’analisi o non analizzavano tutti i pazienti nel gruppo di assegna-

zione¹⁸. Numerose altre revisioni hanno riscontrato che un reporting inadeguato è più frequente nelle riviste specialistiche^{16,19} e in quelle pubblicate in lingua non inglese^{20,21}.

Un’adeguata randomizzazione, riducendo il bias di selezione all’inizio del trial, è la componente fondamentale dei RCT di elevata qualità²². Una randomizzazione corretta richiede due step: la generazione di una sequenza di allocazione casuale e l’occultamento di tale sequenza ai ricercatori che arruolano i partecipanti (box 1)^{2,23}.

Purtroppo, nonostante il loro ruolo chiave, la descrizione delle metodologie utilizzate per l’assegnazione dei partecipanti agli interventi è generalmente inadeguata. Ad esempio, il 5% di 206 trial pubblicati in riviste di ostetricia e ginecologia erano in realtà studi non randomizzati²³. Questa stima è da considerarsi conservativa, poiché la maggior parte dei report attuali non descrive adeguatamente le metodologie di allocazione^{20,23,30-33}.

Come migliorare il reporting dei RCT: il CONSORT statement

DerSimonian et coll. hanno suggerito che “gli editori potrebbero migliorare sensibilmente il reporting di trial clinici, fornendo agli autori un elenco di item da riportare rigorosamente”³⁴. All’inizio degli anni ‘90, due gruppi di editori, ricercatori e metodologi hanno pubblicato separatamente raccomandazioni per il reporting dei trial^{35,36}. A seguito di un invito di Rennie in un editoriale successivo, i due gruppi si sono incontrati sviluppando un set comune di indicazioni³⁷, dando vita al CONSORT (Consolidated Standards of Reporting Trials) Statement³⁸.

Il CONSORT Statement (o semplicemente CONSORT) comprende una checklist di item fondamentali che dovrebbero essere inclusi nel reporting dei RCT e un dia-

Box 1. Assegnazione dell’intervento: perché la randomizzazione è così importante?

La metodologia utilizzata per assegnare i partecipanti a ciascun gruppo di intervento è un aspetto cruciale del disegno di un trial. L’assegnazione casuale è il metodo ideale, impiegato regolarmente con successo nei trial per oltre 50 anni²⁴. La randomizzazione ha tre vantaggi principali²⁵: in primo luogo, se applicata correttamente, elimina i bias di selezione, assicurando che tutti i fattori prognostici – sia noti che sconosciuti – si distribuiscano omogeneamente nel gruppo sperimentale e in quello di controllo. In assenza di randomizzazione, i confronti tra gli interventi possono risultare alterati, consapevolmente o meno, per la presenza del bias di selezione. In secondo luogo, l’assegnazione casuale consente di utilizzare la teoria della probabilità per esprimere la possibilità che ciascuna differenza di esito tra i gruppi di studio sia da attribuire all’efficacia del trattamento in studio²⁶. Infine, l’assegnazione casuale può in alcuni casi facilitare il blinding di ricercatori, partecipanti e valutatori degli esiti, grazie all’impiego di un placebo per ridurre i bias dopo l’assegnazione degli interventi²⁷. Di questi tre vantaggi, la prevenzione del bias di selezione all’inizio del trial è, indubbiamente, il più importante²⁷.

Il successo della randomizzazione dipende da due aspetti strettamente correlati: un’adeguata generazione della sequenza di allocazione casuale e il suo occultamento sino all’assegnazione degli interventi^{2,23}. Un aspetto fondamentale è l’eventuale conoscenza o prevedibilità dell’assegnazione dei partecipanti ai gruppi di intervento da parte dei soggetti coinvolti nel trial²⁹. Il meccanismo di assegnazione degli interventi deve pertanto assicurare che chi arruola i partecipanti non conosca in anticipo a quale gruppo di intervento sarà assegnato il paziente successivo (occultamento della lista di randomizzazione)^{2,23}. Pertanto, se un’adeguata sequenza di allocazione consente di non prevedere le assegnazioni successive sulla base di quelle precedenti, un corretto occultamento della lista mantiene nascoste le assegnazioni successive.

gramma per documentare il flusso dei partecipanti nelle varie fasi di un trial. Pur essendo finalizzato a migliorare il reporting di trial a gruppi paralleli, numerosi item del CONSORT sono utili anche per altri disegni di trial: di non inferiorità, di equivalenza, fattoriali, cluster e crossover. Per migliorare il reporting dei trial con questi disegni – così come il reporting di specifiche tipologie di dati (effetti avversi⁴²), di interventi (trattamenti non farmacologici⁴⁴, fitoterapia⁴⁴) e degli abstract⁴⁵, sono state pubblicate specifiche estensioni del CONSORT³⁹⁻⁴¹.

Il CONSORT ha l'obiettivo di fornire agli autori una guida per migliorare il reporting dei loro trial, rendendolo chiaro, completo e trasparente. Lettori, revisori ed editori possono utilizzare il CONSORT anche per la valutazione critica dei trial, ma è opportuno ricordare che il CONSORT non è stato elaborato con questa finalità. Molti item che, pur non definiti esplicitamente nel CONSORT, dovrebbero essere inclusi nel reporting di un trial: ad esempio le informazioni riguardanti l'approvazione del comitato etico, l'ottenimento del consenso informato da parte dei partecipanti, e, se rilevante, l'esistenza del comitato per la sicurezza dei dati e il monitoraggio. Inoltre, tutti gli altri aspetti di un trial che vengono menzionati devono essere adeguatamente riportati, come ad esempio i risultati di analisi di costo/efficacia⁴⁶⁻⁴⁸.

Dalla sua prima pubblicazione nel 1996, il CONSORT è stato adottato da oltre 400 riviste (www.consort-statement.org) e promosso da diversi gruppi editoriali, come l'International Committee of Medical Journal Editors⁴⁹, determinando un miglioramento della qualità dei reporting dei trial^{17,50,51}. Tuttavia, il CONSORT è un'iniziativa in corso e il CONSORT Statement viene revisionato periodicamente³: l'ultima revisione risale al 2001⁵²⁻⁵⁴. Da allora le evidenze scientifiche per informare il CONSORT sono notevolmente aumentate e dati empirici hanno evidenziato l'importanza di nuove criticità metodologiche, come il reporting selettivo degli outcome⁵⁵⁻⁵⁷. Il gruppo CONSORT si è riunito in Canada nel gennaio 2007 per revisionare il CONSORT statement 2001 e il relativo documento di spiegazione ed elaborazione. La checklist revisionata è riportata nella tabella 1 e il diagramma di flusso, non revisionato, nella figura 1⁵²⁻⁵⁴.

Il CONSORT Statement 2010: spiegazione ed elaborazione

Durante la revisione del CONSORT del 2001 è emerso che una spiegazione ed elaborazione dei singoli item avrebbero agevolato ricercatori e altri utenti nel reporting dei trial. Pertanto, insieme alla versione 2001 del CONSORT, è stato pubblicato un articolo di spiegazione ed elaborazione⁵⁸ che illustrava il background e il razionale scientifico di ogni item, riportando esempi pubblicati di reporting adeguato. Il razionale per la revisione dell'articolo di spiegazione ed elaborazione è simile a quello del CONSORT Statement. Descriviamo brevemente di segui-

to le principali considerazioni aggiunte e quelle eliminate della precedente versione dell'articolo di spiegazione ed elaborazione.

Le variazioni del CONSORT 2010 Spiegazione ed Elaborazione

Abbiamo apportato diverse modifiche sostanziali e alcune modifiche formali a questa versione del CONSORT 2010 Spiegazione ed Elaborazione (per i dettagli si veda la versione 2010 del CONSORT Statement⁵⁹). Alcune sono costituite da variazioni della checklist del CONSORT: ci sono tre item nuovi nella checklist del CONSORT 2010 – come l'item 24, che invita gli autori a indicare dove reperire il protocollo completo del trial. Abbiamo aggiornato alcune informazioni esistenti, riportando evidenze metodologiche più recenti e migliorato alcuni esempi. Abbiamo eliminato il glossario, ora disponibile sul sito web del CONSORT (www.consort-statement.org). Dove possibile, riportiamo ulteriori evidenze da studi empirici rilevanti. Molti libri eccellenti sui trial clinici offrono una discussione più ampia degli aspetti metodologici⁶⁰⁻⁶². Infine, per comodità, a volte utilizziamo la definizione "trattamenti" e "pazienti", anche se non tutti gli interventi valutati nei RCT sono trattamenti e non tutti i partecipanti sono pazienti.

ITEM DELLA CHECKLIST

TITOLO E ABSTRACT

Item 1a. Riportare nel titolo che lo studio è un trial randomizzato.

Esempio. "Efficacia e sicurezza degli inalatori orali di nicotina per la disassuefazione al fumo: un trial clinico randomizzato in doppio cieco"⁶³.

Spiegazione. La capacità di identificare il report di un trial randomizzato in un database elettronico dipende in larga misura da come è stato indicizzato. È possibile che gli indicizzatori non classifichino un articolo come trial randomizzato se gli autori non riportano esplicitamente questa informazione⁶⁴. Per facilitare la corretta indicizzazione e l'identificazione di uno trial, gli autori dovrebbero utilizzare la parola "randomizzato" nel titolo, indicando che i partecipanti sono stati assegnati ai gruppi di confronto con modalità random.

Abstract strutturato: disegno, metodi, risultati, conclusioni (per indicazioni specifiche si veda l'estensione del CONSORT relativa agli abstract).

Per specifiche indicazioni si veda l'estensione del CONSORT relativa agli abstracts^{45,65}.

Spiegazione. È fondamentale che gli abstract siano chiari, trasparenti e sufficientemente dettagliati, perché spesso i lettori basano i loro giudizi esclusivamente su queste informazioni. Alcuni lettori utilizzano l'abstract come strumento di screening per decidere se leggere, o

Tabella 1. CONSORT 2010: checklist delle informazioni da includere nel report di un trial randomizzato*

Sezione/Topic	N° item	Item
TITOLO E ABSTRACT		
	1a	Riportare nel titolo che lo studio è un trial randomizzato
	1b	Abstract strutturato: disegno, metodi, risultati, conclusioni (per indicazioni specifiche si veda l'estensione del CONSORT relativa agli abstract)
INTRODUZIONE		
Background e obiettivi	2a	Background scientifico e spiegazione del razionale dello studio
	2b	Obiettivi o ipotesi specifici
METODI		
Disegno del trial	3a	Descrizione del disegno del trial (parallelo, fattoriale), incluso il rapporto di allocazione
	3b	Rilevanti variazioni metodologiche (es. criteri di eleggibilità) apportate dopo l'inizio del trial, con relative motivazioni
Partecipanti	4a	Criteri di eleggibilità dei partecipanti
	4b	Setting e aree geografiche in cui sono stati raccolti i dati
Interventi	5	Interventi somministrati a ciascun gruppo, riportando dettagli adeguati per renderli riproducibili, anche rispetto a modalità e tempi di somministrazione
Outcome	6a	Elenco completo degli outcome predefiniti, primari e secondari, inclusi tempi e modalità di rilevazione
	6b	Eventuali variazioni degli outcome dopo l'inizio del trial, con le relative motivazioni
Dimensione del campione	7a	Metodi utilizzati per stimare la dimensione del campione
	7b	Se applicabili, specificare le analisi intermedie (interim analysis) effettuate e riportare i criteri di interruzione del trial (stopping rules)
Randomizzazione		
Generazione della sequenza di allocazione	8a	Metodi utilizzati per generare la sequenza di allocazione random
	8b	Tipo di randomizzazione e dettagli di eventuali restrizioni (randomizzazione a blocchi e dimensione del blocco)
Meccanismo per occultare la lista di randomizzazione	9	Metodo per implementare la lista di randomizzazione (es. sequenza numerata di contenitori), descrivendo tutti gli step seguiti per mantenere nascosta la sequenza di allocazione sino all'assegnazione degli interventi.
Implementazione	10	Chi ha generato la sequenza di allocazione random, chi ha arruolato i partecipanti e chi li assegnati a ciascun gruppo di intervento
Cecità	11a	Se lo studio è in cieco, quali soggetti sono blinded dopo l'assegnazione al gruppo di intervento (es. partecipanti, professionisti che erogano l'assistenza, valutatori degli esiti) e con quali modalità
	11b	Se rilevante, descrizione dell'aspetto simile degli interventi
Metodi statistici	12a	Metodi statistici utilizzati per confrontare i gruppi di partecipanti per gli outcome primari e secondari
	12b	Metodi utilizzati per analisi statistiche aggiuntive come le analisi per sottogruppi e le analisi aggiustate
RISULTATI		
Flusso dei partecipanti (fortemente raccomandato l'uso di un diagramma di flusso)	13a	Per ciascun gruppo riportare i partecipanti randomizzati, quelli che hanno ricevuto il trattamento previsto e quelli inclusi nell'analisi per l'outcome primario
	13b	Per ciascun gruppo riportare i partecipanti persi al follow-up e quelli esclusi dopo la randomizzazione, con le relative motivazioni
Reclutamento	14a	Date relative ai periodi di reclutamento e di follow-up
	14b	Motivazioni per cui il trial si è concluso o è stato interrotto
Caratteristiche di base	15	Tabella che riporta le caratteristiche demografiche e cliniche di base per ciascun gruppo
Numeri analizzati	16	Per ciascun gruppo riportare i partecipanti (denominatore) inclusi in ciascuna analisi e se l'analisi è stata eseguita secondo i gruppi di assegnazione originari (intention-to-treat analysis)
Outcome e misure	17a	Per ciascun outcome primario e secondario, i risultati per ogni gruppo, la stima puntiforme dell'effetto e la sua precisione (limiti di confidenza al 95%)
	17b	Per gli outcome dicotomici è raccomandata la presentazione dell'efficacia dell'intervento sia con misure assolute, sia relative
Analisi ancillari	18	Risultati di tutte le altre analisi effettuate (analisi per sottogruppi, analisi aggiustate), distinguendo le analisi predefinite da quelle esplorative
Effetti avversi	19	Tutti i rilevanti effetti avversi o indesiderati in ciascun gruppo (per indicazioni specifiche consultare l'estensione del CONSORT sugli effetti avversi)
DISCUSSIONE		
Limiti	20	Limiti del trial, considerando le fonti di potenziali bias, l'imprecisione e – se rilevanti – le analisi multiple
Generalizzabilità	21	Generalizzabilità (validità esterna, applicabilità) dei risultati del trial
Interpretazione	22	Interpretazione coerente con i risultati, bilanciando benefici ed effetti avversi e tenendo conto di altre evidenze rilevanti
ALTRE INFORMAZIONI		
Registrazione	23	Numero di registrazione e nome del registro di trial
Protocollo	24	Dove è possibile reperire il protocollo completo del trial, se disponibile
Finanziamento	25	Fonti di finanziamento e altri supporti (es. fornitura dei farmaci), ruolo dei finanziatori

* È fortemente raccomandata la lettura del CONSORT 2010 Explanation and Elaboration per informazioni dettagliate su tutti gli item. Se necessario, è raccomandata la consultazione delle estensioni del CONSORT relative a: trial con randomizzazione cluster⁴⁰, trial di non-inferiorità e di equivalenza³⁹, trattamenti non farmacologici⁴³, interventi di fitoterapia⁴⁴, trial pragmatici⁴¹. Ulteriori estensioni saranno pubblicate a breve. Tutte le estensioni e i riferimenti bibliografici aggiornati relativi a questa checklist sono disponibili all'indirizzo: www.consort-statement.org

CONSORT 2010. Diagramma di flusso

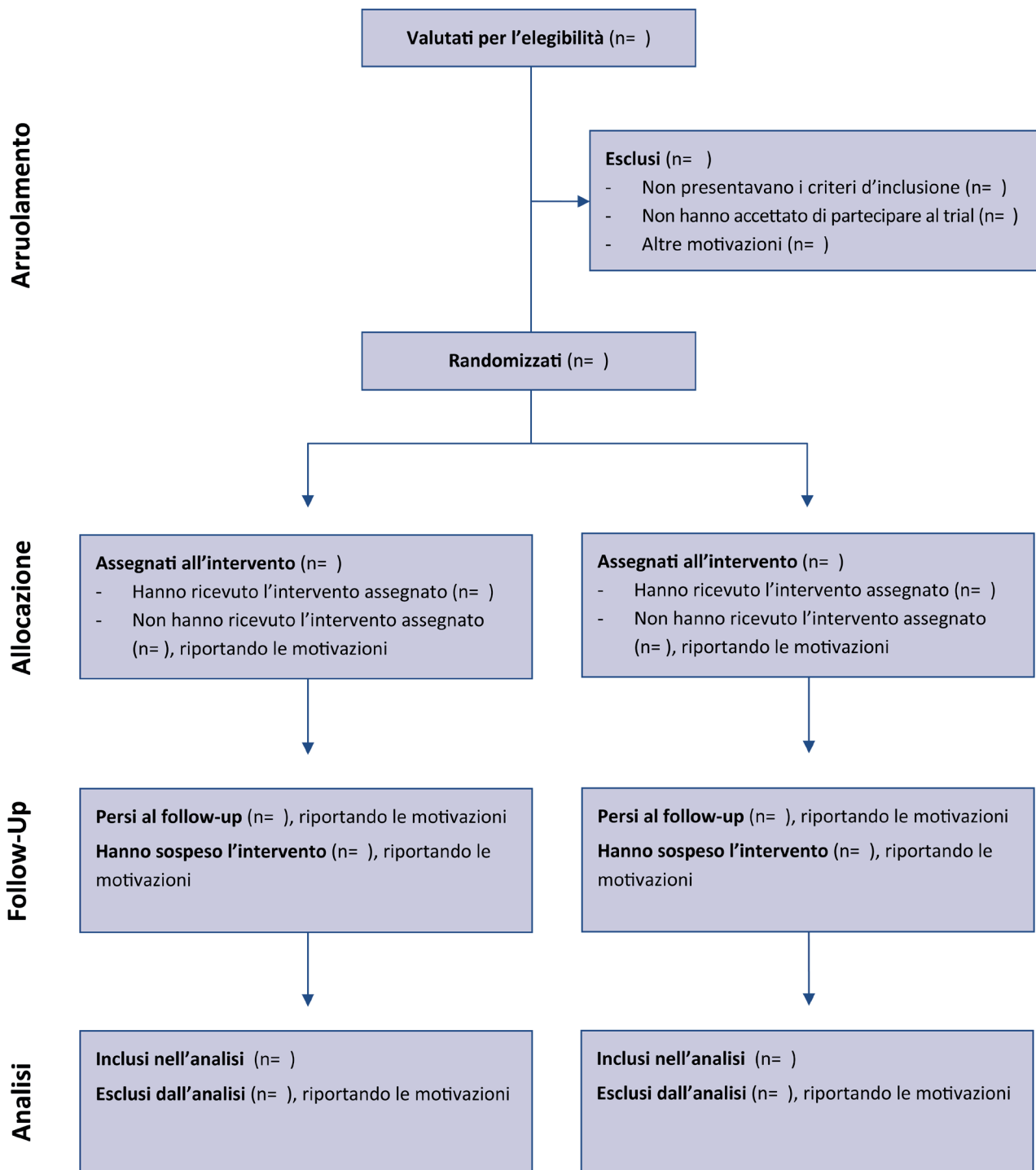


Figura 1. CONSORT 2010. Diagramma di flusso

meno, l'articolo integrale. Inoltre, poiché non tutti i trial sono disponibili gratuitamente e non tutti i professionisti sanitari hanno accesso ai report completi dei trial, le decisioni cliniche vengono talora formulate sulla base delle informazioni contenute negli abstract⁶⁶.

Pertanto, l'abstract pubblicato su una rivista dovrebbe contenere informazioni adeguate e sufficienti da rappresentare una sintesi accurata delle metodologie e dei risultati del trial, secondo i vincoli editoriali e il formato della rivista.

Un abstract adeguatamente strutturato e ben redatto permette ai lettori di valutare rapidamente la rilevanza dei risultati e facilita il reperimento del trial nei database elettronici⁶⁷. L'abstract dovrebbe riflettere accuratamente i contenuti dell'articolo, senza includere informazioni che non compaiono nel testo integrale. Studi comparativi sull'accuratezza delle informazioni contenute negli abstract rispetto a quanto riportato nelle corrispondenti pubblicazioni integrali hanno rilevato dati incongruenti o mancanti rispetto all'articolo completo⁶⁸⁻⁷¹. Al contrario, omettere dall'abstract eventuali effetti avversi può indurre il lettore a interpretare erroneamente i risultati del trial^{42,72}.

Una recente estensione del CONSORT statement fornisce un elenco di item essenziali che gli autori de-

vono includere pubblicando i risultati principali di un trial randomizzato in un abstract destinato ad una rivista o a un convegno (tabella 2)⁴⁵. Per il report di un trial randomizzato è fortemente raccomandato l'utilizzo di abstract strutturati, dove le informazioni sono organizzate in sezioni relative a disegno, conduzione, analisi e interpretazione⁷³. Alcuni studi hanno evidenziato che, rispetto agli abstract descrittivi, quelli strutturati hanno una qualità superiore^{74,75} e permettono al lettore di identificare più facilmente le informazioni⁷⁶. In ogni caso, poiché numerose riviste hanno un proprio layout e impongono un limite di battute per gli abstract, non suggeriamo di modificare questi formati, ma ci limitiamo a raccomandare le informazioni da riportare.

INTRODUZIONE

Item 2a. Background scientifico e spiegazione del razionale dello studio.

Esempio. "La chirurgia è il trattamento di prima scelta nei pazienti con carcinoma polmonare non a piccole cellule – *non-small cell lung cancer* (NSCLC) – di stadio I e II... Una meta-analisi sul NSCLC ha combinato i risultati di otto trial randomizzati che confrontavano l'intervento chirurgico rispetto alla combinazione chirurgia più chemioterapia adiuvante a base di cisplatino e ha mostrato

Tabella 2. Item da includere nell'abstract di un trial randomizzato

Item	Descrizione
Autori	Informazioni per la corrispondenza con gli autori
Disegno	Descrizione del disegno del trial (parallelo, cluster, di non-inferiorità)
Metodi	
Partecipanti	Criteri di eleggibilità dei partecipanti e setting in cui sono stati raccolti i dati
Interventi	Interventi somministrati a ciascun gruppo
Obiettivi	Obiettivi o ipotesi specifici
Outcome	Precisa definizione dell'outcome primario
Randomizzazione	Metodologie utilizzate per assegnare i partecipanti agli interventi
<i>Blinding</i>	Se lo studio è in cieco, quali soggetti (partecipanti, professionisti che erogano l'assistenza, valutatori degli esiti) sono <i>blinded</i> rispetto all'intervento assegnato
Risultati	
Numero di partecipanti randomizzati	Numero di partecipanti assegnati in maniera random a ciascun gruppo di intervento
Reclutamento	Status del trial
Numero di partecipanti analizzati	Numero di partecipanti inclusi nell'analisi per ciascun gruppo
Outcome	Per l'outcome primario: risultati per ciascun gruppo, dimensione dell'effetto e sua precisione
Effetti avversi	Rilevanti eventi avversi o indesiderati
Conclusioni	Interpretazione generale dei risultati
Registrazione	Numero di registrazione e nome del registro di trial
Finanziamento	Fonti di finanziamento

un piccolo, ma non significativo ($p = 0,08$), beneficio assoluto di sopravvivenza di circa il 5% a 5 anni (dal 50% al 55%). Al momento della pianificazione del presente trial (metà degli anni '90), la chemioterapia adiuvante non era ancora il trattamento clinico standard... Il razionale scientifico per la chemioterapia neo-adiuvante è triplice: la regressione della neoplasia primaria potrebbe essere ottenuta in modo da semplificare o ridurre il successivo intervento chirurgico; micro-metastasi non identificate potrebbero essere considerate all'inizio del trattamento; si potrebbe inibire lo stimolo a neoplasie residue da parte di fattori di crescita rilasciati durante l'intervento chirurgico e la guarigione della ferita... il presente trial è stato quindi disegnato per confrontare, in pazienti con NSCLC resecabile, l'intervento chirurgico con tre cicli di chemioterapia a base di platino seguita dalla chirurgia in termini di sopravvivenza, qualità della vita, stadio di malattia, tasso di resecabilità, estensione della chirurgia, tempistiche e sedi delle recidive⁷⁷."

Spiegazione. Di solito, l'introduzione è costituita da testo libero, dove gli autori spiegano il background e il razionale scientifico del trial e il suo schema generale. Può anche essere opportuno inserire nell'introduzione gli obiettivi del trial (item 2b). Il razionale può essere esplicativo (es., valutare la possibile influenza di un farmaco sulla funzione renale) o pragmatico (es., guidare la pratica clinica confrontando benefici ed effetti avversi di due interventi sanitari). Gli autori dovrebbero segnalare ogni evidenza di benefici ed effetti avversi degli interventi attivi inclusi in un trial e suggerire una spiegazione plausibile di come gli interventi dovrebbero funzionare, nel caso in cui non sia ovvio⁷⁸.

La Dichiarazione di Helsinki⁷⁹ sostiene che la ricerca biomedica che coinvolge soggetti umani dovrebbe essere basata su una conoscenza approfondita della letteratura scientifica, dal momento che non è etico esporre inutilmente gli esseri umani ai rischi della ricerca. Alcuni trial clinici sono risultati inutili perché il quesito valutato poteva essere risolto da una revisione sistematica della letteratura^{80,81}. Pertanto, la necessità di un nuovo trial deve essere giustificata nell'introduzione che, idealmente, dovrebbe includere un riferimento bibliografico a una revisione sistematica di trial simili già pubblicati, o una nota che specifichi che tali trial non sono ancora stati condotti⁸².

Item 2b. Obiettivi o ipotesi specifici.

Esempio. "In questo studio abbiamo testato l'ipotesi che una gestione attiva del travaglio in donne nullipare dovrebbe: 1. Ridurre l'incidenza di taglio cesareo, 2. Ridurre l'incidenza di travaglio prolungato, 3. Non influenzare la soddisfazione materna durante l'esperienza del parto⁸³."

Spiegazione. Gli obiettivi sono i quesiti a cui il trial intende dare una risposta, e spesso riguardano l'efficacia di uno specifico intervento terapeutico o preventivo. Le ipotesi sono i quesiti predefiniti che vengono valutati per poter raggiungere gli obiettivi. Le ipotesi sono più specifiche rispetto agli obiettivi e sottoposte a esplicita valutazione statistica; tuttavia, non sempre, obiettivi e ipotesi sono facilmente distinguibili. La maggior parte dei reporting di RCT fornisce adeguate informazioni relative a obiettivi e ipotesi del trial⁸⁴.

METODI

Item 3a. Descrizione del disegno del trial (parallelo, fattoriale), incluso il rapporto di allocazione.

Esempio. "Studio multicentrico, stratificato (da 6 a 11 anni e da 12 a 17 anni di età, con randomizzazione non bilanciata [2:1]), in doppio cieco, controllato verso placebo, a gruppi paralleli condotto negli Stati Uniti (41 centri)⁸⁵."

Spiegazione. Il termine "disegno" è spesso utilizzato per indicare tutti gli aspetti metodologici di pianificazione di un trial, ma può avere anche un'interpretazione più restrittiva. Molti aspetti specifici del disegno di un trial, come i dettagli di randomizzazione e cecità, sono trattati in altri item della checklist del CONSORT. In questa sezione sono presenti le informazioni sul tipo di trial, (es. a gruppi paralleli o fattoriale), il quadro concettuale (come la superiorità o la non-inferiorità), e altri aspetti correlati non analizzati in altri item della checklist.

Il CONSORT Statement si concentra principalmente sui trial con partecipanti randomizzati e assegnati individualmente a uno di due gruppi "paralleli". In realtà, poco più del 50% dei trial pubblicati ha questo tipo di disegno¹⁶. Tra i principali disegni alternativi ricordiamo i trial paralleli a bracci multipli, i trial con disegno crossover e fattoriali, i trial con randomizzazione cluster⁴⁰. Inoltre, se molti trial tendono, qualora esista, a dimostrare la superiorità di un nuovo intervento, altri ne valutano la non-inferiorità o l'equivalenza³⁹. È importante descrivere in modo chiaro questi aspetti dello studio, inclusa l'unità di randomizzazione (paziente, assistenza del medico di medicina generale, lesione), oltre a riportarli nell'abstract (item 1b).

Quando vengono utilizzati disegni di trial meno comuni, la scelta dovrebbe essere esplicitamente motivata, perché possono richiedere un campione di maggiori dimensioni, o analisi e interpretazioni più complesse.

Sebbene la maggior parte dei trial utilizzi la randomizzazione bilanciata (1:1 per due gruppi), si consiglia di esplicitare sempre il rapporto di allocazione. Per i trial farmacologici, può essere importante specificare anche la fase dello studio (I-IV).

Item 3b. Rilevanti variazioni metodologiche (es. criteri di eleggibilità) apportate dopo l'inizio del trial, con relative motivazioni.

Esempio. “I pazienti sono stati randomizzati e assegnati a uno dei sei gruppi paralleli, inizialmente secondo il rapporto 1:1:1:1:1:1, per ricevere uno dei cinque regimi di otamixaban [...] o un controllo attivo di eparina non frazionata [...] un comitato indipendente per il monitoraggio dei dati ha rivisto i dati relativi alla sicurezza del paziente non in cieco; non sono state eseguite analisi intermedie (*interim analysis*) per valutare l'efficacia. Durante il trial, il comitato ha raccomandato che il gruppo sottoposto alla dose più bassa di otamixaban (0,035 mg/kg/h) interrompesse il trattamento poiché la terapia anticoagulante risultava inadeguata. Sulla base di tale raccomandazione il protocollo è stato immediatamente modificato e i partecipanti sono stati successivamente randomizzati secondo il rapporto 2:2:2:2:1 rispettivamente ai 4 rimanenti gruppi di otamixaban e al controllo⁸⁶.”

Spiegazione. Se alcuni trial possono iniziare senza un protocollo stabilito (trial esplorativi), la maggior parte segue un protocollo che specifica nei dettagli le modalità di conduzione. Dal momento che è impossibile prevedere ogni eventuale cambiamento di situazioni, nel corso di un trial possono verificarsi deviazioni dal protocollo originale che richiederanno rilevanti variazioni metodologiche successive all'inizio del trial. Tali variazioni possono conseguire alla disponibilità di informazioni esterne derivanti da altri studi, a difficoltà finanziarie interne, oppure a un reclutamento inadeguato. Queste modifiche al protocollo dovrebbero essere apportate senza compromettere il *blinding* sugli outcome dei partecipanti. In alcuni trial, esiste un comitato indipendente di monitoraggio dei dati che può proporre variazioni del protocollo alla luce di dati non in cieco. Tali modifiche potrebbero influenzare i metodi di studio (es. variazioni di regimi di trattamento, criteri di eleggibilità, rapporto di randomizzazione o durata del follow-up) o la conduzione del trial (es. eliminazione di un centro con una scarsa qualità dei dati)⁸⁷.

Alcuni trial vengono progettati con un disegno “adattativo”. Non esiste una definizione universalmente accettata per questo disegno, ma in pratica lo si potrebbe definire “un disegno a più stadi che utilizza i dati raccolti per decidere come modificare gli aspetti dello studio, senza comprometterne validità e integrità”⁸⁸. Le modifiche, che solitamente riguardano la dimensione del campione e il numero dei bracci di trattamento, possono consentire decisioni più rapide e un più efficiente utilizzo delle risorse. Esistono, tuttavia, importanti aspetti etici, statistici e pratici da prendere in considerazione^{89,90}, e al fine di aiutare il lettore a interpretare i risultati, è essenziale riportare con estrema chiarezza

se le modifiche apportate fanno parte del disegno dello studio o conducono a mutate circostanze. Attualmente le variazioni metodologiche non vengono riportate in maniera adeguata: infatti, una revisione di confronti tra protocolli e rispettivi trial successivamente pubblicati ha evidenziato che circa la metà presenta inspiegabili discrepanze negli outcome primari⁵⁷, la randomizzazione, il *blinding*⁹¹ e le analisi statistiche⁹².

Item 4a. Criteri di eleggibilità dei partecipanti.

Esempio. “I partecipanti eleggibili erano gli adulti con infezione da HIV, di età ≥ 18 anni, che presentavano i criteri di eleggibilità per la terapia antiretrovirale secondo le linee guida nazionali del Malawi per il trattamento dell'HIV (stadio clinico III o IV OMS o qualsiasi stadio OMS con una conta di CD4 $<250/\text{mm}^3$) e che all'inizio del trattamento avevano un BMI $<18,5$. I criteri di esclusione erano la gravidanza e l'allattamento, o partecipazione a un altro programma di nutrizione supplementare⁹³.”

Spiegazione. Per aiutare i lettori a interpretare lo studio, è necessario descrivere dettagliatamente i criteri di eleggibilità utilizzati per selezionare i partecipanti del trial. In particolare, una comprensione chiara di tali criteri è uno degli elementi indispensabili per stabilire a chi possono essere applicati i risultati di uno studio – cioè, la generalizzabilità del trial (applicabilità) e la rilevanza per la pratica clinica o la salute pubblica (item 21)⁹⁴. Altrettanto importante in questo senso è la descrizione del metodo di reclutamento, come l'invio selettivo o la partecipazione volontaria (es. attraverso la pubblicità). Dal momento che vengono applicati prima della randomizzazione, i criteri di eleggibilità non influenzano la validità interna di un trial, ma condizionano la sua validità esterna.

I criteri di eleggibilità tipici e ampiamente accettati dipendono dalla natura e dallo stadio della malattia oggetto di studio, mentre quelli di esclusione evitano l'arruolamento di partecipanti a rischio di eventi avversi dell'intervento in studio o per il rispetto di norme legali ed etiche. Il consenso informato dei partecipanti, ad esempio, viene tipicamente richiesto negli studi sperimentali. La distinzione tra criteri d'inclusione ed esclusione non è necessaria; lo stesso criterio può essere formulato per includere o escludere i partecipanti⁹⁵.

Nonostante la loro importanza, i criteri di eleggibilità spesso non vengono adeguatamente riportati. Ad esempio, otto trial pubblicati che hanno portato a segnalazioni cliniche da parte dei *National Institutes of Health* hanno descritto, in media, 31 criteri di eleggibilità nei loro protocolli, ma solo il 63% di tali criteri era riportato negli articoli pubblicati, e solo il 19% nelle segnalazioni cliniche⁹⁶. Carenze simili sono state riscontrate per i trial sull'HIV⁹⁷. Tra i 364 report di RCT in chirurgia, il 25% non ha indicato alcun criterio di eleggibilità⁹⁸.

Item 4b. Setting e aree geografiche in cui sono stati raccolti i dati.

Esempio. “Lo studio è stato eseguito presso il dipartimento di terapia antiretrovirale del Queen Elizabeth Central Hospital a Blantyre, Malawi, dal gennaio 2006 all’aprile 2007. Blantyre è la principale città commerciale del Malawi, con un 1.000.000 di abitanti e una prevalenza stimata di HIV del 27% negli adulti nel 2004⁹³.”

Spiegazione. Oltre ai criteri di eleggibilità dei partecipanti (item 4a) e alla descrizione degli interventi (item 5), le informazioni sui setting e sulle aree geografiche sono fondamentali per valutare l’applicabilità di un trial. I partecipanti sono stati reclutati da setting di assistenza primaria, secondaria o terziaria o dalla comunità? Le istituzioni di assistenza sanitaria differiscono notevolmente in termini di organizzazione, esperienza e risorse e per quanto riguarda il rischio di base per la condizione in esame. Anche altri aspetti del setting (incluso il contesto sociale, economico e culturale e il clima) possono influenzare la validità esterna di uno studio.

Gli autori devono riportare il numero e la tipologia dei setting e descrivere i professionisti che erogano gli interventi sanitari. Devono inoltre riportare le aree geografiche in cui è stato condotto lo studio, inclusi il paese, se possibile la città, e l’ambiente circostante (es. la comunità, l’ambulatorio, l’ospedale o l’unità operativa). In particolare, dovrebbe essere chiaramente indicato se il trial è stato eseguito in uno o più centri (“trial multicentrici”). Questa descrizione dovrebbe fornire informazioni sufficienti per consentire ai lettori di valutare se i risultati dello studio possono essere rilevanti per il proprio setting assistenziale. Il contesto in cui viene condotto il trial può essere notevolmente diverso da quello in cui i suoi risultati saranno poi utilizzati per guidare la pratica clinica e le decisioni di politica sanitaria^{94,99}. Gli autori devono fornire anche tutti gli altri dettagli relativi a setting e aree geografiche che potrebbero aver influenzato i risultati (es. problemi di viabilità e trasporti possono condizionare la partecipazione del paziente o determinare ritardi nella somministrazione degli interventi).

Item 5. Interventi somministrati a ciascun gruppo, riportando dettagli adeguati per renderli riproducibili, anche rispetto a modalità e tempi di somministrazione.

Esempi. “Nel trial POISE, i pazienti hanno ricevuto la prima dose del farmaco in studio (100 mg di metoprololo per via orale a rilascio prolungato) o del placebo 2-4 ore prima dell’intervento chirurgico. La somministrazione del farmaco richiedeva una frequenza cardiaca ≥ 50 bpm e una pressione sistolica ≥ 100 mmHg; questi parametri emodinamici sono stati monitorati prima di ogni somministrazione. Nelle prime 6 ore dopo l’intervento chirurgico, se in qualsiasi momento la frequenza cardiaca era ≥ 80 bpm e la pressione arteriosa sistolica

≥ 100 mmHg, i pazienti ricevevano la prima dose post-operatoria per via orale (100 mg di metoprololo a rilascio prolungato o placebo). Se il farmaco non era somministrato durante le prime 6 ore, i pazienti ricevevano la prima dose post-operatoria a 6 ore dall’intervento chirurgico. 12 ore dopo la prima dose post-operatoria, i pazienti iniziavano ad assumere 200 mg di metoprololo a rilascio prolungato per via orale o placebo, 1 volta/die per 30 giorni. Se la frequenza cardiaca del paziente era costantemente < 45 bpm o la pressione arteriosa sistolica scendeva sotto i 100 mmHg, il farmaco veniva sospeso sino ad un aumento della frequenza cardiaca o della pressione sistolica; il farmaco veniva, quindi, somministrato nuovamente al dosaggio di 100 mg/die. I pazienti in cui la frequenza cardiaca era costantemente 45-49 bpm e la pressione sistolica < 100 mmHg ritardavano l’assunzione del farmaco di 12 ore¹⁰⁰.”

“I pazienti erano randomizzati e assegnati a ricevere un tutore in neoprene su misura da indossare durante la notte oppure all’assistenza tradizionale. Il tutore, un’ortesi rigida raccomandata per uso esclusivamente notturno, ricopriva la base del pollice e il palmo, ma non il polso. I tutori erano stati realizzati da tre terapisti professionali esperti, che regolavano il tutore per ogni paziente in modo che il primo fotorecettore potesse essere aperto e il pollice posizionato dal lato opposto al primo dito lungo. I pazienti erano invitati a contattare il terapeuta se ritenevano necessaria una regolazione del tutore, in caso di aumento del dolore mentre indossavano il tutore, o di comparsa di effetti collaterali (es. erosione della pelle). Poiché in questa condizione non esiste nessun trattamento standard, i pazienti dei due gruppi (intervento e controllo) ricevevano cure tradizionali a discrezione del proprio medico di medicina generale o del reumatologo. Si è deciso di non utilizzare un placebo perché, secondo la nostra esperienza, nessun placebo per splintaggio è in grado di ottenere un *blinding* efficace dei pazienti¹⁰¹.”

Spiegazione. Considerato che il medico che intende utilizzare l’intervento deve conoscerne esattamente le sue modalità di somministrazione nel trial, gli autori dovrebbero descrivere accuratamente ciascun intervento, inclusi quelli di controllo¹⁰². Per un intervento farmacologico, le informazioni dovrebbero includere il nome del farmaco, la dose, le modalità di somministrazione (es. per via orale, endovenosa), le tempistiche e la durata della somministrazione, le condizioni in cui non è possibile erogare gli interventi e il regime di titolazione, se applicabile. Se il gruppo di controllo riceve “l’assistenza convenzionale” – *usual care* – è importante descrivere accuratamente da cosa è costituita. Se il gruppo sperimentale o quello di controllo riceve un intervento combinato, gli autori dovrebbero fornire sia una descrizione dettagliata di ciascun intervento, sia una spiegazione

delle modalità con cui la combinazione degli interventi viene somministrata o interrotta, oltre ai fattori trigger per la loro introduzione, se applicabile.

Estensioni specifiche del CONSORT Statement riguardano il reporting di interventi non-farmacologici e fitoterapici ed evidenziano i dettagli specifici da prevedere^{43,44}, come il livello di competenza professionale o i dettagli relativi alle modalità di standardizzazione degli interventi. Si raccomanda ai lettori di consultare le indicazioni per gli interventi non-farmacologici e fitoterapici, quando necessario.

Item 6a. Elenco completo degli outcome predefiniti, primari e secondari, inclusi tempi e modalità di rilevazione.

Esempio. “L’end-point primario relativo all’efficacia dell’intervento nella psoriasi era la percentuale di pazienti che raggiungeva un miglioramento del 75% dell’attività di malattia rispetto al valore di base dopo 12 settimane, secondo la misura del *Psoriasis Area and Severity Index* (PASI). Ulteriori analisi valutavano la variazione percentuale del punteggio PASI e il miglioramento delle lesioni psoriasiche¹⁰³.”

Spiegazione. Tutti i RCT valutano gli outcome (o end-point) per confrontare i risultati dei due gruppi di partecipanti. La maggior parte dei trial prevede diversi outcome, di cui alcuni hanno una maggiore rilevanza di altri. L’outcome primario è l’esito pre-definito considerato molto rilevante dai principali *stakeholders* (pazienti, *policy-makers*, medici, finanziatori) e viene utilizzato per stimare la dimensione del campione (item 7). La definizione di diversi outcome primari, anche se possibile, non è raccomandata perché può determinare problemi di interpretazione conseguenti alle analisi multiple (item 18 e 20). Gli outcome primari dovrebbero essere esplicitamente indicati come tali nel reporting di un RCT. Altri outcome di interesse sono quelli secondari (o accessori): gli outcome secondari possono essere diversi e spesso includono gli effetti avversi dell’intervento (item 19), che in realtà dovrebbero essere sempre considerati importanti, indipendentemente dal fatto che siano definiti primari o secondari.

Tutti gli outcome, primari e secondari, devono essere identificati e ben definiti nei dettagli per consentire l’utilizzo ad altri ricercatori¹⁰². Quando gli outcome vengono valutati in periodi diversi dopo la randomizzazione, gli autori dovrebbero indicare il *time point* predefinito dell’outcome primario. Per molti interventi non farmacologici è utile specificare chi ha valutato gli esiti (es. se sono necessarie particolari competenze per farlo) e quanti erano i valutatori⁴³.

Qualora siano disponibili adeguate scale o linee guida, il loro utilizzo deve essere riportato^{104,105}, sia per migliorare la qualità delle misure, sia per consentire il confronto con studi simili¹⁰⁶. Ad esempio, la valutazione della

qualità della vita può migliorare attraverso l’utilizzo di uno strumento validato¹⁰⁷. Gli autori dovrebbero sempre indicare la provenienza e le caratteristiche delle scale.

Oltre 70 outcome sono stati utilizzati in 196 RCT su farmaci anti-infiammatori non steroidei per l’artrite reumatoide¹⁰⁸, e 640 strumenti diversi sono stati utilizzati in 2.000 trial sulla schizofrenia, di cui 369 impiegati una sola volta³³. La valutazione di 149 di questi 2.000 trial ha dimostrato che l’utilizzo di scale non pubblicate in letteratura costituisce fonte di bias. Nei trial non farmacologici, un terzo delle definizioni di superiorità del trattamento, basato su scale non pubblicate, non sarebbe stato formulato se fosse stata utilizzata una scala pubblicata¹⁰⁹. Risultati analoghi sono stati riportati in altri campi^{110,111}. Solo il 45% di un gruppo di 519 RCT pubblicati nel 2000 specificava l’outcome primario¹⁶, analogamente al 53% di un gruppo simile di 614 RCT pubblicati nel 2006¹⁷.

Item 6b. Eventuali variazioni degli outcome dopo l’inizio del trial, con le relative motivazioni.

Esempio. “L’end-point primario originario era la mortalità per tutte le cause, ma, nel corso di un’analisi in cieco, il comitato di monitoraggio dei dati e della sicurezza ha riscontrato che la mortalità totale era inferiore a quella stimata e che lo studio non poteva essere completato con la dimensione del campione e la potenza definiti inizialmente. Il comitato direttivo ha pertanto deciso di utilizzare un duplice end-point primario: oltre alla mortalità per tutte le cause (end-point primario originario), la mortalità per tutte le cause insieme ai ricoveri per cause cardiovascolari (il primo end-point secondario predefinito)¹¹².”

Spiegazione. Esistono numerose motivazioni che richiedono di apportare variazioni al protocollo originale dello studio (item 24). Gli autori devono descrivere tutte le principali modifiche al protocollo, incluse quelle non previste ai criteri di inclusione, agli interventi, alle valutazioni, alla raccolta dei dati, ai metodi di analisi e agli outcome, informazioni che non sempre vengono riportate.

Come indicato in precedenza (item 6a), la maggior parte dei trial prende in considerazione diversi outcome, con il rischio di riportare i risultati solo per un sottogruppo selezionato (item 17); tale rischio dovrebbe essere prevenuto con la definizione preliminare e il reporting di outcome primari e secondari (item 6a). In alcuni trial, tuttavia, le circostanze richiedono una modifica delle modalità di valutazione di un outcome o addirittura, come nel precedente esempio, la modifica dell’outcome primario. Ad esempio, evidenze derivanti da altri trial o revisioni sistematiche possono suggerire che l’end-point potrebbe non essere appropriato; oppure il reclutamento dei partecipanti o l’incidenza dell’evento nel trial potrebbero essere inferiori al previsto¹¹². La variazione di

un end-point sulla base di dati non in cieco è molto più problematica, anche se può essere specificata nel contesto di un trial con disegno adattativo⁸⁸. Gli autori devono identificare e motivare tali modifiche, oltre che così riportare e spiegare eventuali variazioni agli outcome apportate dopo l'inizio del trial.

Un confronto tra protocolli e pubblicazioni di 102 trial randomizzati ha riscontrato che il 62% dei trial aveva almeno un outcome primario modificato, introdotto, oppure omesso rispetto al protocollo originale⁵⁵. Differenze negli outcome primari erano anche evidenti tra protocolli e pubblicazioni nel 40% di 48 trial finanziati dal *Canadian Institutes of Health Research*¹¹³. Nessuno dei 150 report di trial successivamente pubblicati riportava, e tanto meno spiegava, le modifiche apportate al protocollo. Analoghi risultati provenienti da altri studi sono stati recentemente menzionati da una revisione sistematica di studi empirici che analizza i bias nel reporting degli outcome⁵⁷.

Item 7a. Metodi utilizzati per stimare la dimensione del campione.

Esempi. “Per rilevare una riduzione della degenza post-operatoria di 3 giorni (deviazione standard 5 giorni), in accordo con lo studio di Lobo et al.¹⁷ con un livello di significatività del 5% e una potenza dell'80%, è stato stimato un campione di 50 pazienti per gruppo, considerando un'incidenza di abbandono del 10%. Per reclutare questo numero di pazienti è stato previsto un periodo di 12 mesi¹¹⁴.”

“Sulla base di un'incidenza attesa dell'end-point primario composito del 11% dopo 2,25 anni nel gruppo placebo, abbiamo stimato necessari 950 eventi dell'end-point primario e un campione di 9.650 pazienti, con una potenza di 90% per rilevare una differenza significativa tra ivabradina e placebo, corrispondente a una riduzione relativa del 19% del rischio (errore α pari al 5%). Inizialmente abbiamo disegnato un trial *event-driven*, stabilendo di interromperlo una volta rilevati 950 end-point primari. Tuttavia, l'incidenza dell'end-point primario è stata superiore al previsto, forse per le caratteristiche di base dei pazienti reclutati, che presentavano un rischio più elevato rispetto a quello atteso (es. minore percentuale di pazienti classe NYHA I e incidenza più elevata di diabete e ipertensione). Abbiamo calcolato che nel momento in cui si fossero verificati 950 end-point primari, i pazienti inclusi più recentemente sarebbero stati trattati solo per circa 3 mesi. Di conseguenza, nel gennaio 2007, il comitato esecutivo ha deciso di modificare lo studio da *event-driven* a *time-driven* e di proseguire sino a 12 mesi il follow-up degli ultimi pazienti randomizzati. Questa variazione non ha modificato la durata di 3 anni prevista dello studio¹¹⁵.”

Spiegazione. Per ragioni scientifiche ed etiche, la di-

mensione del campione di un trial deve essere accuratamente stimata, bilanciando valutazioni cliniche e statistiche. Idealmente, uno studio deve essere sufficientemente numeroso da avere un'elevata probabilità (potenza) di rilevare come statisticamente significativa una differenza clinicamente importante, se questa esiste realmente. La dimensione dell'effetto considerato importante è inversamente proporzionale alla dimensione del campione necessario per rilevarla: in altre parole per rilevare piccole differenze sono necessari campioni molto numerosi. Gli elementi per calcolare la dimensione del campione sono (1) l'incidenza degli outcome stimati in ciascun gruppo (che implica la differenza target clinicamente rilevante tra i gruppi di intervento); (2) l'errore α (tipo I); (3) la potenza statistica, oppure l'errore β (tipo II); (4) per gli outcome continui, la deviazione standard¹¹⁶. L'interazione di questi elementi e il loro reporting saranno diversi per i trial con randomizzazione cluster⁴⁰ e per i trial di non-inferiorità e di equivalenza³⁹.

Gli autori dovrebbero dettagliare come è stata stimata la dimensione del campione; se è stata utilizzata una stima formale della potenza, dovrebbero definire l'outcome primario su cui è basata la stima (item 6a), tutte le misure utilizzate per la stima, e la conseguente dimensione del campione per ciascun gruppo dello studio. È preferibile citare l'incidenza dell'evento attesa nel gruppo di controllo e la differenza tra i gruppi che non si intende tralasciare. In alternativa, gli autori potrebbero presentare la percentuale dell'evento o la media per ogni gruppo utilizzata nelle loro stime. Dovrebbero essere forniti i dettagli relativi ai margini previsti in caso di attrito o non-compliance durante lo studio.

Secondo alcuni metodologi, i cosiddetti trial sottodimensionati (*underpowered*) sono comunque accettabili perché potrebbero essere combinati in una revisione sistematica con meta-analisi¹¹⁷⁻¹¹⁹, e perché è sempre meglio disporre di alcune evidenze, piuttosto che di nessuna. Occorre comunque che il trial sia imparziale, che sia riportato correttamente, pubblicato indipendentemente dai risultati e pertanto disponibile per le meta-analisi¹¹⁸. D'altra parte, molti ricercatori temono che i trial sottodimensionati con risultati indeterminati non verranno mai pubblicati e insistono sul fatto che ciascun trial deve avere una potenza adeguata. Questo dibattito, destinato a continuare, con punti di vista differenti anche da parte degli stessi membri del gruppo CONSORT è assolutamente irrilevante per il reporting di un trial. Infatti, qualunque sia la potenza di un trial, gli autori devono riportare correttamente la sua dimensione con tutte le metodologie e i presupposti¹¹⁸. In questo modo la potenza dello studio viene indicata in maniera trasparente ai lettori, fornendo loro gli strumenti per valutare se il trial ha raggiunto o meno la dimensione prevista.

In alcuni trial, vengono utilizzate analisi intermedie (*interim*) per decidere se è opportuno interrompere precocemente o continuare il reclutamento oltre la fine definita del trial (item 7b). Se la dimensione reale del campione differisce da quella inizialmente stimata per altre motivazioni (es. scarso reclutamento o modifica della dimensione del campione), queste dovrebbero essere esplicitamente dichiarate.

Trial di piccole dimensioni spesso concludono erroneamente che non esistono differenze tra i gruppi di intervento, quando in realtà il numero di pazienti studiati è troppo esiguo per giustificare una simile affermazione¹²⁰. Alcune revisioni di trial pubblicati hanno rilevato che una elevata percentuale di trial ha una potenza inadeguata per rilevare effetti clinicamente significativi dell'intervento¹²¹⁻¹²³. In realtà, è molto più probabile che esistano piccole differenze clinicamente significative rispetto a grandi differenze rilevabili solo con trial di grandi dimensioni¹²⁴.

In generale, le dimensioni del campione riportate nei trial sono piccole: la dimensione media del campione era di 54 pazienti in 196 trial sull'artrite¹⁰⁸, di 46 pazienti in 73 trial in dermatologia⁸, e di 65 pazienti in 2.000 trial sulla schizofrenia³³. Questi campioni di piccole dimensioni sono coerenti con quelli di uno studio condotto su 519 trial indicizzati in PubMed nel dicembre 2000¹⁶ e su un altro relativo a 616 di trial indicizzati nel 2006¹⁷: il numero medio di pazienti reclutati per trial a gruppi paralleli era di 80 in entrambi i casi. Inoltre, molte revisioni hanno riscontrato che pochi autori riportano le metodologie utilizzate per stimare la dimensione del campione^{8,14,32,33,123}.

La stima *post hoc* della potenza statistica utilizzando i risultati di un trial è di limitato valore perché viene poi correttamente indicata dai limiti di confidenza (item 17)¹²⁵.

Item 7b. Se applicabili, specificare le analisi intermedie (*interim analysis*) effettuate e riportare i criteri di interruzione del trial (*stopping rules*).

Esempi. "Durante il trial sono state eseguite due analisi intermedie: i livelli di significatività hanno mantenuto un *P-value* pari a 0,05 e sono stati calcolati secondo i limiti pre-definiti di interruzione di O'Brien-Fleming. Quest'ultima analisi, ha utilizzato un punteggio Z di 1,985, con un *P-value* associato di 0,0471¹²⁶."

"Un comitato indipendente di monitoraggio dei dati valutava periodicamente i risultati di efficacia e sicurezza. I criteri di interruzione si basavano sui limiti modificati di Haybittle-Peto di 4 deviazioni standard (DS) nella prima metà dello studio e di 3 DS nella seconda metà per i dati di efficacia; di 3 DS nella prima metà dello studio e di 2 DS nella seconda metà per i dati di sicurezza. Sono state eseguite due analisi intermedie di efficacia

quando si è verificato il 50% e il 75% del numero atteso di outcome primari; per le analisi intermedie non è stata effettuata alcuna correzione del *P-value*¹²⁷".

Spiegazione. Considerato che molti trial reclutano i partecipanti durante un lungo periodo di tempo, se un intervento ha effetti particolarmente favorevoli o sfavorevoli, lo studio dovrà verosimilmente essere interrotto in anticipo per motivi etici. Questo aspetto può essere affrontato esaminando i risultati man mano che i dati vengono raccolti, preferibilmente da parte di un comitato indipendente di monitoraggio dei dati. Tuttavia, eseguire analisi statistiche multiple dei dati raccolti senza un'adeguata correzione può portare a risultati e interpretazioni errati¹²⁸. Ad esempio, se vengono effettuate cinque analisi intermedie che utilizzano un *P-value* di 0,05, l'incidenza complessiva di risultati falsamente positivi sarà prossima al 19%.

Esistono diversi metodi statistici sequenziali per correggere le analisi multiple¹²⁹⁻¹³¹, e il loro utilizzo dovrebbe essere definito nel protocollo del trial. Con questi metodi, i dati vengono confrontati a ogni analisi intermedia, e un *P-value* inferiore al valore soglia, definito dal metodo sequenziale, indica significatività statistica. Alcuni ricercatori utilizzano metodi sequenziali come supporto al processo decisionale, mentre altri¹³² li considerano criteri d'interruzione formali con l'intento di interrompere il trial se il *P-value* osservato è inferiore al valore soglia.

Gli autori dovrebbero specificare se la periodica verifica dei dati è stata effettuata da loro in prima persona oppure da un comitato di monitoraggio: dovrebbero essere quindi riportati quanti erano gli esaminatori, le loro motivazioni iniziali, i metodi statistici utilizzati (inclusi gli eventuali criteri d'interruzione formali), e se erano stati previsti prima dell'inizio del trial, prima che il comitato di monitoraggio dei dati avesse esaminato tutti i dati intermedi per l'allocazione, o qualche tempo dopo. Queste informazioni spesso non vengono incluse nelle pubblicazioni di trial¹³³, anche in quelli che sono stati interrotti precocemente¹³⁴.

Item 8a. Metodi utilizzati per generare la sequenza di allocazione random.

Esempi. "Farmacisti indipendenti hanno erogato inattori attivi o contenenti placebo secondo una lista di randomizzazione generata da un software⁶³."

"Per l'assegnazione dei partecipanti, è stata utilizzata una lista di numeri casuali generata da un software¹³⁵."

Spiegazione. I partecipanti del trial devono essere assegnati ai gruppi di confronto attraverso un metodo casuale (random) e imprevedibile (box 1). Gli autori dovrebbero fornire informazioni sufficienti per consentire al lettore di valutare le metodologie utilizzate per generare la sequenza di allocazione random e il rischio di

bias nell'assegnazione ai gruppi di intervento. Le informazioni sul processo di randomizzazione devono essere incluse nel corpo dell'articolo principale e non riportate in un documento separato che potrebbe sfuggire all'attenzione del lettore.

Il termine random ha un significato metodologico ben preciso: con l'allocazione random, ciascun partecipante ha una probabilità nota di ricevere ciascun intervento prima dell'assegnazione dell'intervento, determinata da un processo casuale e imprevedibile. Tuttavia, spesso il termine random viene utilizzato impropriamente per descrivere trial in cui sono stati utilizzati metodi di assegnazione deterministici e non casuali: assegnazione alternata, numero di ospedali, data di nascita. Quando i ricercatori utilizzano questi metodi non casuali dovrebbero descriverli con precisione, evitando di utilizzare il termine random o qualsiasi sua variante, incluso il termine "quasi-random". I trial che utilizzano

metodologie non-random producono in genere risultati affetti da bias^{2-4,136}, che derivano presumibilmente da metodi inadeguati per occultare la lista di randomizzazione (item 9).

Esistono diversi metodi adeguati per generare la sequenza di allocazione, ma in assenza di dettagli, difficilmente i lettori sono in grado di valutarne l'appropriatezza sulla base di semplici termini quali "assegnazione random", "randomizzazione" o "random". Pertanto, gli autori dovrebbero sempre specificare il metodo di generazione della sequenza: ad esempio una tabella di numeri random o un software per generare numeri random. La sequenza può anche essere generata attraverso il processo di minimizzazione, un metodo non random generalmente accettabile (box 2).

In alcuni trial, i partecipanti possono essere intenzionalmente assegnati in maniera non bilanciata ai diversi gruppi di intervento: ad esempio per acquisire

Box 2. Randomizzazione e minimizzazione

- **Randomizzazione semplice.** La randomizzazione basata su un singolo rapporto di allocazione è nota come randomizzazione semplice. La randomizzazione semplice con un rapporto di allocazione 1:1 è analoga al lancio di una moneta, meccanismo non accettabile per un RCT. Tuttavia, "semplice" è un termine, in qualche modo, improprio: sebbene altri metodi di randomizzazione sembrano più complessi e sofisticati, la randomizzazione semplice è in realtà più imprevedibile e consente una maggiore prevenzione dei bias rispetto a tutti gli altri metodi.
- **Randomizzazione con restrizioni.** Viene così definito qualunque metodo di assegnazione diverso dalla randomizzazione semplice. Il metodo più comune è la randomizzazione a blocchi; altri metodi di randomizzazione con restrizioni, meno frequenti, sono la randomizzazione con sostituzione, con lancio della moneta, e con urna¹⁴¹.
- **Randomizzazione a blocchi.** Viene utilizzata per garantire che i gruppi di confronto siano generati in base a un rapporto predeterminato, di solito 1:1, oppure che i gruppi abbiano approssimativamente le stesse dimensioni. Il blocco può essere utilizzato per garantire un numero analogo di pazienti in qualsiasi momento dell'arruolamento: per ogni blocco di otto partecipanti, ad esempio, quattro dovrebbero essere assegnati a ciascun braccio del trial¹⁴². Un bilanciamento migliore si ottiene a scapito di una riduzione dell'imprevedibilità della sequenza: infatti, anche se l'ordine degli interventi varia in maniera casuale all'interno di ciascun blocco, i ricercatori potrebbero dedurre alcune assegnazioni successive dell'intervento nel caso in cui sia loro nota la dimensione del blocco. Il *blinding*¹⁴³ e l'utilizzo di blocchi permutati (variabili) di maggiori dimensioni possono limitare questo inconveniente.
- **Randomizzazione stratificata.** La stratificazione viene utilizzata per garantire un bilanciamento adeguato dei fattori prognostici in ciascun gruppo. Per effetto del caso, specialmente nei trial di piccole dimensioni, i due gruppi possono essere sbilanciati per i fattori prognostici, come l'età e lo stadio della malattia, riducendo la credibilità del trial¹⁴⁴. È possibile evitare questi sbilanciamenti senza rinunciare ai vantaggi della randomizzazione: infatti, la stratificazione garantisce che il numero di partecipanti assegnati a un intervento sia bilanciato all'interno di ogni strato. La randomizzazione stratificata è realizzata attraverso una procedura di randomizzazione separata per ciascuno di due o più sottoinsiemi (strati) di partecipanti, relativi ad esempio al centro dello studio, all'età, o alla severità della malattia. La stratificazione per centro è comune nei trial multicentrici. La stratificazione richiede una qualche forma di restrizione (come i blocchi all'interno degli strati) perché la stratificazione senza blocchi è inefficace.
- **Minimizzazione.** La minimizzazione assicura il bilanciamento tra i gruppi d'intervento per i diversi fattori prognostici (es. l'età)^{22,60}. Il primo paziente viene assegnato in maniera casuale e i successivi allocati in relazione alla distribuzione dei principali fattori prognostici con l'obiettivo di minimizzare lo sbilanciamento tra i due gruppi. Con il progredire dell'arruolamento, se i principali fattori prognostici risultano bilanciati, il paziente successivo viene assegnato in maniera casuale. La minimizzazione ha il vantaggio di rendere piccoli gruppi molto simili in termini di fattori prognostici durante tutto il periodo dell'arruolamento. La minimizzazione rappresenta l'unica alternativa accettabile alla randomizzazione: anche se alcuni la considerano addirittura superiore¹⁴⁵, occorre tenere presente che la minimizzazione manca dei principi teorici per prevenire i bias su tutti i fattori prognostici noti e sconosciuti. In generale, i trial che utilizzano la minimizzazione sono considerati metodologicamente equivalenti ai trial randomizzati, anche quando non è stato incluso un elemento random.

maggior esperienza con una nuova procedura o per limitare i costi del trial. In questi casi, gli autori dovrebbero specificare il rapporto di allocazione (es. 2:1, oppure due partecipanti sottoposti all'intervento per ciascun controllo) (item 3a).

In un campione rappresentativo di trial indicizzati in PubMed nel 2000, solo il 21% riporta un approccio adeguato per la generazione di una sequenza random¹⁶, valore che aumenta al 34% per un gruppo simile di trial indicizzati nel 2006¹⁷. In oltre il 90% di questi casi, i ricercatori hanno utilizzato un software per generare numeri random o una tabella di numeri random.

Item 8b. Tipo di randomizzazione e dettagli di eventuali restrizioni (randomizzazione a blocchi e dimensione del blocco)

Esempi. "La sequenza di randomizzazione è stata generata utilizzando il software statistico Stata 9.0 (StataCorp, College Station, TX) ed è stata stratificata per centro, con un rapporto di allocazione 1:1 utilizzando blocchi casuali di 2, 4 e 6¹³⁷."

"I partecipanti sono stati assegnati mediante randomizzazione semplice (numeri random generati dal computer) a ciascuno di due gruppi di trattamento¹³⁸."

Spiegazione. Nei trial che arruolano numerose centinaia, o addirittura migliaia di partecipanti, la randomizzazione semplice di solito garantisce numeri simili nei due gruppi¹³⁹, solitamente confrontabili rispetto alle variabili prognostiche sia note, sia sconosciute¹⁴⁰. Per i trial di dimensioni più piccole (item 7 b), così come per quelli che possono essere interrotti prima di arruolare la dimensione del campione stimata, possono risultare utili alcune randomizzazioni con restrizioni, ovvero procedure che consentono di raggiungere un bilanciamento tra i gruppi in termini di dimensioni o fattori prognostici (item 2).

È importante specificare se non è stata utilizzata alcuna restrizione, dichiarandolo esplicitamente o indicando che è stata eseguita una "randomizzazione semplice". In caso contrario, accanto ai metodi di selezione random, dovrebbero essere specificati quelli utilizzati per la randomizzazione con restrizioni. Per la randomizzazione a blocchi, gli autori dovrebbero indicare i metodi di generazione dei blocchi (es. utilizzando un disegno a blocchi permutati con un software che genera numeri random), la/e dimensione/i del blocco e se la dimensione del blocco era fissa o variava in modo casuale. Se i ricercatori sono venuti a conoscenza della dimensione del blocco, questo dovrebbe essere riportato, poiché potrebbe violare il codice di assegnazione. Gli autori dovrebbero specificare se è stata utilizzata una randomizzazione stratificata e, in tal caso, riportare i fattori considerati (es. centri di arruolamento, genere, stadio della malattia), i valori cut-off di categorizzazione all'interno

degli strati, e il metodo utilizzato per la restrizione. Nonostante la randomizzazione stratificata sia una tecnica utile, specialmente per i trial di piccole dimensioni, è complessa da attuare nella pratica, o addirittura impossibile se vengono previsti molti strati. Occorre specificare se è stata utilizzata la minimizzazione (box 2), così come le variabili incorporate nel processo. Se è stato impiegato un elemento casuale, deve essere indicato.

Solo il 9% di 206 reporting di trial pubblicati su riviste specialistiche²³ e il 39% di 80 trial su riviste mediche di carattere generale hanno riportato l'utilizzo della randomizzazione stratificata³² e solo la metà dei report ha specificato l'utilizzo della randomizzazione con restrizioni. Tuttavia, questi studi e quello di Adetugbo e Williams⁸ hanno riscontrato che le dimensioni dei gruppi di intervento in molti trial erano uguali o molto simili, nonostante la mancata segnalazione di blocchi e strati. Pertanto, una possibile spiegazione per il bilanciamento dei numeri è proprio la mancata segnalazione dell'utilizzo della randomizzazione con restrizioni.

Item 9. Metodo per implementare la lista di randomizzazione (es. sequenza numerata di contenitori), descrivendo tutti gli step seguiti per mantenere nascosta la sequenza di allocazione sino all'assegnazione degli interventi.

Esempi. "La doxiciclina e il placebo, in forma di capsule di aspetto identico, sono stati pre-confezionati in contenitori numerati progressivamente per ciascuna donna secondo la lista di randomizzazione. A ciascuna donna è stato assegnato un numero e le capsule nel corrispondente contenitore pre-confezionato¹⁴⁶."

"La sequenza di allocazione è stata mantenuta nascosta al ricercatore (JR) che arruolava e valutava i partecipanti utilizzando buste numerate progressivamente, opache, sigillate e pinzate. Al fine di renderla impermeabile alla luce intensa, la busta è stata rivestita internamente da un foglio di alluminio. Per evitare il sovrvertimento della sequenza di allocazione, la busta – su cui erano riportati nome e data di nascita del partecipante – è stata videoregistrata dopo essere stata sigillata con i dettagli del partecipante ben visibili. Con un foglio di carta carbone all'interno della busta sono stati trasferiti i dati sulla scheda di allocazione all'interno della busta e un secondo ricercatore (CC) ha esaminato successivamente le registrazioni per assicurarsi che le buste fossero ancora sigillate quando sono stati scritti i nomi dei partecipanti. Le buste corrispondenti sono state aperte solo dopo che i partecipanti arruolati avevano completato tutte le valutazioni basali ed era il momento di assegnare l'intervento¹⁴⁷."

Spiegazione. L'item 8a ha discusso la necessità di generare una sequenza di allocazione imprevedibile. Altrettanta importanza rivestono le metodologie utilizzate

per applicare la sequenza di allocazione ai partecipanti arruolati nel trial (box 1). Una lista di allocazione, una volta generata, dovrebbe essere implementata previo occultamento²³, meccanismo critico che impedisce ai professionisti che arruolano i pazienti di conoscere a quale gruppo sarà assegnato il paziente successivo e, quindi, di essere influenzati da questa consapevolezza. La decisione di accettare o rifiutare un partecipante deve sempre essere presa – e il consenso informato del partecipante ottenuto – senza conoscere l'assegnazione successiva¹⁴⁸.

L'occultamento della lista di randomizzazione non deve essere confuso con la cecità o *blinding* (item 11). L'occultamento della lista di randomizzazione contribuisce a prevenire il bias di selezione, protegge la sequenza di allocazione sino all'assegnazione e può sempre essere implementata con successo². L'obiettivo del *blinding*, che non sempre può essere attuato è, invece, di prevenire i bias di performance e di accertamento dell'esito (*detection bias*) e di proteggere la sequenza dopo l'assegnazione²³. In assenza di occultamento adeguato della lista di randomizzazione, anche sequenze di allocazione casuali e imprevedibili possono essere sovvertite^{2,149}.

È auspicabile un'assegnazione centralizzata o da parte di terzi: adeguati meccanismi di occultamento della lista di randomizzazione prevedono infatti il coinvolgimento di un centro esterno. Due tecniche comuni sono l'uso di una farmacia o di sistema telefonico centralizzato di randomizzazione¹⁵⁰. Quando non è possibile ricorrere a un centro esterno, un'eccellente metodologia di occultamento della lista di randomizzazione è costituita dall'impiego di contenitori numerati. Gli interventi (spesso farmaci) sono sigillati in contenitori identici, progressivamente numerati, secondo la sequenza di allocazione¹⁵¹. La predisposizione di una sequenza numerata di buste opache e sigillate può essere un adeguato meccanismo di occultamento dell'assegnazione se attuato e monitorato con attenzione. Questo metodo, tuttavia, se non correttamente eseguito, può essere manipolato. I ricercatori dovrebbero garantire che le buste siano opache anche se esposte alla luce, e aperte in sequenza solo dopo che il nome del partecipante e altri dettagli sono stati riportati sulla relativa busta¹⁴³.

Alcuni studi metodologici forniscono evidenze empiriche a sostegno di tali precauzioni^{152,153}: i trial in cui la sequenza di allocazione è stata occultata in modo inadeguato o non chiaro sovrastimano l'efficacia dell'intervento, rispetto ai trial in cui gli autori hanno riportato un adeguato occultamento della lista di randomizzazione. Questi risultati forniscono chiare evidenze che un inadeguato occultamento della lista di randomizzazione distorce la stima dell'efficacia dell'intervento. Nonostante l'importanza del meccanismo di occultamento della lista di randomizzazione, i trial pubblicati spesso

omettono tali dettagli. Il meccanismo utilizzato per l'assegnazione degli interventi è stato omesso nel reporting dell'89% dei trial sull'artrite reumatoide¹⁰⁸, nel 48% dei trial nelle riviste di ostetricia e ginecologia²³, e nel 44% dei trial in riviste mediche di carattere generale³². In un campione più rappresentativo di tutti i trial randomizzati indicizzati su PubMed, solo il 18% ha riportato una strategia, non sempre adeguata, per occultare la lista di randomizzazione¹⁶.

Item 10. Chi ha generato la sequenza di allocazione random, chi ha arruolato i partecipanti e chi li assegna a ciascun gruppo di intervento.

Esempi. “Per determinare se un paziente sarebbe stato trattato con streptomicina e riposo a letto (caso S) o solo con riposo a letto (caso C) si è fatto riferimento a una serie di metodi statistici basati su campionamento random di numeri elaborati dal professor Bradford Hill per ciascun sesso presso ogni centro; i dettagli della lista erano sconosciuti ai ricercatori o al coordinatore [...] Dopo l'accettazione di un paziente e prima dell'ammissione al centro di somministrazione della streptomicina, la relativa busta numerata è stata aperta presso l'ufficio centrale; la scheda all'interno ha indicato se il paziente doveva essere un caso S oppure C, informazione che è stata poi fornita al medico del centro²⁴.”

“I dettagli relativi al gruppo assegnato sono stati riportati su schede colorate contenute in buste numerate progressivamente, opache, sigillate, preparate presso la *National Perinatal Epidemiology Unit* (NPEU) e conservate in un luogo definito in ciascun dipartimento. La randomizzazione è stata effettuata alla fine della seconda fase del travaglio quando l'ostetrica ha ritenuto che il parto vaginale fosse imminente. Per inserire una donna nello studio, l'ostetrica ha aperto la busta successiva seguendo la numerazione progressiva¹⁵⁴.”

“La randomizzazione a blocchi ha previsto una lista di numeri random realizzata con un software da un ricercatore non coinvolto nel trial dal punto di vista clinico. Abbiamo stratificato per l'ammissione a una procedura oncologica. Dopo avere ottenuto il consenso del paziente, l'infermiera ha contattato telefonicamente un soggetto indipendente dall'arruolamento per effettuare l'assegnazione¹⁵⁵.”

Spiegazione. Come indicato nell'item 9, l'occultamento dell'intervento al momento dell'arruolamento è particolarmente importante. Oltre alle metodologie utilizzate, è importante conoscere anche come è stata implementata la sequenza random, in particolare chi ha generato la lista di randomizzazione, chi ha arruolato i partecipanti e chi li ha assegnati a ciascun gruppo di intervento.

Il processo di randomizzazione dei partecipanti in un trial ha tre diverse fasi: generazione della sequenza, oc-

cultamento della lista di randomizzazione e implementazione (box 3). Anche se gli stessi soggetti possono essere coinvolte in più fasi, i ricercatori dovrebbero effettuare una netta distinzione tra quelli coinvolti nella generazione della lista di randomizzazione e del suo occultamento, da quelli che assegnano gli interventi. Anche in presenza di una generazione della sequenza di allocazione e di un occultamento della lista di randomizzazione ineccepibili, non separarli dall'implementazione della lista ai gruppi di intervento può essere fonte di bias. Ad esempio, il soggetto che ha generato una sequenza di allocazione potrebbe conservare una copia e consultarla durante l'intervista ai potenziali partecipanti di un trial, influenzando l'arruolamento o l'assegnazione, a prescindere dall'imprevedibilità della sequenza di allocazione. I ricercatori devono, quindi, garantire che il programma di assegnazione sia imprevedibile e protetto anche dalla persona che lo ha generato (come ad esempio in una cassaforte di un edificio inaccessibile dal luogo dell'arruolamento). Nel report del trial deve essere specificato dove i ricercatori hanno conservato la lista di randomizzazione.

Box 3. Fasi standard del processo di randomizzazione

Generazione della sequenza di allocazione

- Generare la sequenza di allocazione in maniera casuale e imprevedibile

Occultamento della lista di randomizzazione

- Sviluppare una strategia di occultamento della lista di randomizzazione (es. bottiglie numerate identiche, o buste progressivamente numerate, sigillate, opache)
- Preparare il meccanismo di occultamento della lista di randomizzazione utilizzando la sequenza di allocazione dal momento di generazione della sequenza

Implementazione

- Arruolare i partecipanti
 - Verificarne i criteri di eleggibilità
 - Fornire spiegazioni sugli obiettivi del trial
 - Ottenere il consenso informato
 - Arruolare i partecipanti
- Assegnare l'intervento (es. aprendo la busta successiva)
- Somministrare l'intervento

Item 11a. Se lo studio è in cieco, quali soggetti sono *blinded* dopo l'assegnazione al gruppo di intervento (es. partecipanti, professionisti che erogano l'assistenza, valutatori degli esiti) e con quali modalità.

Esempi. "Mentre pazienti e medici assegnati al gruppo d'intervento erano a conoscenza del braccio assegnato, i valutatori degli esiti e gli statistici che analizzavano i dati erano mantenuti all'oscuro (*blinded*)¹⁵⁶."

"*Blinding* ed equipollenza erano rispettati rigorosamente, sottolineando al personale che somministrava l'intervento e ai partecipanti che ciascuna dieta aderisce

a principi sani ed è raccomandata da alcuni esperti poiché è più efficace per la perdita di peso a lungo termine. Fatta eccezione per i professionisti che erogavano l'intervento (dietologi e psicologi comportamentali), i ricercatori e il personale non erano a conoscenza della dieta assegnata ai partecipanti. Il trial aderiva alle procedure previste per mantenere separati i professionisti che valutavano gli esiti da quelli che erogavano l'intervento: i primi non erano a conoscenza della dieta assegnata, mentre i dietologi e gli psicologi comportamentali che erogavano l'intervento non effettuavano la misurazione degli outcome. Tutti i ricercatori, i professionisti sanitari e i partecipanti erano mantenuti all'oscuro (*blinded*) delle misurazioni degli esiti e dei risultati del trial¹⁵⁷."

Spiegazione. Il termine *blinding* o *masking* indica che le informazioni sugli interventi assegnati sono tenute nascoste ai soggetti coinvolti nel trial che possono potenzialmente essere influenzati da questa consapevolezza. Il *blinding* è un importante metodo di prevenzione di bias, in particolare per la valutazione di esiti soggettivi¹⁵³.

Benjamin Franklin è stato il primo a utilizzare il *blinding* in un esperimento scientifico¹⁵⁸: mantenendo ignari i pazienti per non far sapere loro quando stava applicando il mesmerismo (popolare "fluido per la guarigione" utilizzato nel 18° secolo), ha dimostrato che il mesmerismo era un imbroglio. Grazie a questo esperimento, la comunità scientifica ha riconosciuto che il *blinding* può ridurre i bias ed è rimasto una strategia comunemente utilizzata nelle sperimentazioni cliniche.

Il box 4, relativo al *blinding*, definisce i gruppi di individui (partecipanti, professionisti che erogano l'assistenza, soggetti che raccolgono i dati, valutatori degli esiti, statistici che analizzano i dati) che possono potenzialmente introdurre bias in un trial per la conoscenza degli interventi assegnati. I partecipanti possono avere risposte terapeutiche differenti se consapevoli dell'intervento assegnato: ad esempio rispondere in modo più favorevole quando ricevono il trattamento innovativo.¹⁵³ La mancanza di *blinding* può anche influenzare l'aderenza all'intervento, l'utilizzo di co-interventi, e il rischio di abbandono del trial.

I professionisti che erogano l'assistenza in maniera non *blinded* possono introdurre bias analoghi e i soggetti non *blinded* che raccolgono i dati possono valutare in maniera differente gli esiti (come la frequenza o la tempistica), ripetere la misura di risultati anomali o incoraggiare i pazienti durante i test di performance. I valutatori degli esiti non *blinded* possono valutare in maniera differente gli esiti soggettivi, e gli statistici non *blinded* che analizzano i dati possono introdurre bias utilizzando specifiche strategie di analisi, quali la selezione di punti temporali o esiti favorevoli, e attraverso l'esclusione di alcuni pazienti dalle analisi. Tutti questi bias sono ben documentati in letteratura^{71,153,159-162}.

Box 4. Terminologia relativa al “blinding”

Affinché un termine tecnico risulti utile deve essere coerente con il suo utilizzo e la sua interpretazione. Gli autori dei trial comunemente usano il termine “doppio cieco” e, meno comunemente, i termini “singolo cieco” e “triplo cieco”. Il problema di questa terminologia è che esiste una grande variabilità di definizioni epidemiologiche e interpretazioni cliniche di questi termini¹⁶⁹. Inoltre, uno studio di 200 RCT in doppio cieco ha riscontrato 18 diverse combinazioni di gruppi realmente blinded, quando gli autori di questi trial sono stati intervistati, e circa uno su cinque di questi trial in doppio cieco non aveva attuato il blinding di partecipanti, professionisti che erogavano l'intervento o che raccoglievano i dati¹⁷⁰.

Poiché questo studio ha dimostrato che i termini singolo, doppio e triplo cieco sono ambigui, autori ed editori dovrebbero abbandonarne l'utilizzo, riportando esplicitamente le categorie di soggetti che vengono mantenuti blinded.

I professionisti che erogano l'intervento includono tutto il personale che assiste i partecipanti durante il trial (es. medici, fisioterapisti, infermieri, chiropratici). I data collectors sono i soggetti che raccolgono i dati relativi agli outcome del trial. I valutatori degli esiti (outcome assessors) stabiliscono se il partecipante ha raggiunto, o meno, l'outcome di interesse.

Anche se alcuni ricercatori hanno anche sostenuto il blinding e il reporting dello stato di blinding del comitato di monitoraggio dei dati e degli autori del manoscritto¹⁶⁰, il blinding di questi gruppi è raro e il suo valore è controverso¹⁷¹.

A volte un gruppo di individui (es. i professionisti che erogano l'intervento) è costituito dagli stessi soggetti che svolgono un altro ruolo in un trial (es. i professionisti che raccolgono i dati). Anche in questo caso, gli autori dovrebbero esplicitamente indicare lo stato di blinding di questi gruppi per consentire ai lettori di giudicare la validità del trial.

Il *blinding*, a differenza dell'occultamento della lista di randomizzazione (item 10), non è una strategia sempre praticabile (es. in un trial che confronta i livelli di dolore associato al prelievo di sangue dall'orecchio o dal pollice¹⁶³). Il *blinding* è particolarmente importante quando la valutazione degli outcome richiede una certa soggettività, come la valutazione del dolore. Mantenere *blinded* i soggetti che raccolgono i dati e quelli che valutano gli esiti avrà poche conseguenze sugli outcome oggettivi, come la mortalità per tutte le cause. Anche in quel caso, però, la mancanza di *blinding* dei partecipanti o dei professionisti che erogano l'assistenza può portare ad altri problemi, come l'attrito differenziale¹⁶⁴. In alcuni trial, in particolare in chirurgia, il *blinding* di partecipanti e chirurghi è spesso difficile o impossibile, ma quello dei soggetti che raccolgono i dati e valutano gli esiti è generalmente fattibile. Ad esempio, le lesioni possono essere fotografate prima e dopo l'intervento e valutate da un osservatore esterno¹⁶⁵. Indipendentemente dalla fattibilità del *blinding*, gli autori dovrebbero sempre riportare quali soggetti sono *blinded*: partecipanti, professionisti che erogano l'assistenza, soggetti che raccolgono i dati e che valutano gli esiti.

Purtroppo, gli autori spesso non riportano l'eventuale utilizzo del *blinding*. Per esempio, il 51% di 506 trial sulla fibrosi cistica¹⁶⁷, il 33% di 196 trial sull'artrite reumatoide¹⁰⁸, e il 38% di 68 trial in dermatologia⁸ non riportano se è stato utilizzato il *blinding*. Sino a quando gli autori dei trial non miglioreranno la loro segnalazione del *blinding*, i lettori avranno difficoltà a giudicare la validità dei trial che vorrebbero eventualmente utilizzare per guidare la loro pratica clinica.

Il termine *masking* è talvolta preferito a *blinding* per evitare confusione con la condizione medica del non vedente. Tuttavia, *blinding* nel suo significato metodologi-

co sembra essere ormai accettato universalmente nel reporting dei trial.

Item 11b. Se rilevante, descrizione dell'aspetto simile degli interventi.

Esempio. “La Jamieson Laboratories Inc ha fornito 500 mg di niacina a rilascio immediato in una compressa bianca, oblunga, divisa in due parti. Abbiamo confermato in modo indipendente il contenuto della compressa utilizzando la cromatografia liquida ad elevate prestazioni [...] Il placebo, reso identico al farmaco oggetto di studio per gusto, colore e dimensione, conteneva cellulosa microcristallina, biossido di silicio, fosfato bicalcico, stearato di magnesio e acido stearico.”

Spiegazione. Come è importante disporre dell'evidenza dell'occultamento per essere certi che l'assegnazione sia avvenuta effettivamente in maniera casuale, allo stesso modo è fondamentale conoscere la metodologia di *blinding*. Nei trial con *blinding* di partecipanti o professionisti che erogano l'assistenza, gli autori dovrebbero indicare chiaramente la similitudine degli interventi rispetto all'aspetto, il gusto, l'odore e le modalità di somministrazione.

E' stato proposto di verificare il *blinding* chiedendo, al termine di un trial, a partecipanti e professionisti coinvolti se ritengono che il partecipante abbia ricevuto l'intervento sperimentale o di controllo. Tuttavia, poiché partecipanti e professionisti conoscono in genere se il partecipante ha sperimentato o meno l'outcome primario, questo rende difficile stabilire se le loro risposte riflettono il fallimento del *blinding* o una precisa assunzione sull'efficacia dell'intervento. Considerata l'incertezza che scaturisce da queste informazioni, abbiamo scartato la proposta di riportare questo tipo di test per il *blinding* dal CONSORT 2010. Raccomandiamo, tuttavia,

agli autori di riportare eventuali variazioni note del *blinding*. Per esempio, gli autori dovrebbero segnalare se è stato necessario smascherare eventuali partecipanti in qualsiasi momento durante lo svolgimento del trial.

Item 12a. Metodi statistici utilizzati per confrontare i gruppi di partecipanti per gli outcome primari e secondari.

Esempio. “L’end-point primario era costituito dalle variazioni del peso corporeo nella popolazione in studio durante le 20 settimane dello studio, secondo il principio dell’*intention-to-treat* [...] Gli end-point secondari includevano le variazioni della circonferenza addominale e della pressione arteriosa sistolica e diastolica, la prevalenza della sindrome metabolica [...] Per l’end-point primario è stata utilizzata l’analisi di covarianza (ANCOVA), mentre per gli end-point secondari le misure di circonferenza della vita, di pressione arteriosa, e altri outcome riportati dal paziente, il tutto completato da un’analisi delle misure ripetute. Il modello ANCOVA includeva il trattamento, il paese e il genere come effetti fissi; il peso corporeo al momento della randomizzazione come co-variata. Abbiamo valutato se i dati fornivano evidenze di superiorità di ciascuna dose di liraglutide rispetto al placebo (obiettivo primario) e all’orlistat (obiettivo secondario)¹⁷⁶.”

Spiegazione. I dati possono essere analizzati con diversi metodi statistici, alcuni dei quali inadeguati in particolari situazioni. È indispensabile specificare il metodo statistico utilizzato per ciascuna analisi; ulteriori dettagli possono essere necessari nella sezione dei risultati. Il principio da seguire è: “Descrivere i metodi statistici con sufficienti dettagli per consentire al lettore esperto l’accesso ai dati originali per verificare i risultati riportati” (www.icmje.org). È anche importante specificare i dettagli dell’analisi statistica come l’analisi *intention-to-treat* (box 6).

Quasi tutti i metodi di analisi consentono una stima dell’efficacia del trattamento, che risulta da un confronto degli outcome tra i gruppi di partecipanti. Gli autori dovrebbero completare la stima puntiforme dell’efficacia con i limiti di confidenza, che indicano il range centrale di incertezza rispetto al reale effetto del trattamento. I limiti di confidenza possono essere interpretati come un range di valori per l’effetto dell’intervento compatibile con i dati osservati. È consuetudine riportare i limiti di confidenza al 95%, definiti come il range di valori entro cui è contenuto, con una probabilità del 95%, il valore reale, valido per l’intera popolazione.

I risultati dello studio possono essere valutati anche in termini di significatività statistica. Il *P-value* rappresenta la probabilità che i dati osservati (o un risultato più estremo) siano dovuti al caso quando gli interventi in realtà non differiscono tra loro. I *P-value* puntiformi

(es. $p = 0,003$) sono da preferire a valori soglia imprecisi come $p < 0,05$ ^{48,177}.

Tutti i metodi standard di analisi assumono che i dati siano “indipendenti”: per i trial controllati, questo significa che è stata effettuata una sola osservazione per ciascun partecipante. Considerare come dati indipendenti osservazioni multiple effettuate sul singolo partecipante è un grave errore: questo accade quando gli outcome possono essere misurati in diverse parti del corpo, come in odontoiatria o reumatologia. In questo caso l’analisi dei dati dovrebbe conteggiare ciascun partecipante una sola volta^{178,179} o dovrebbe essere eseguita utilizzando metodi statistici più complessi¹⁸⁰. Analisi multiple non aggiustate per singolo partecipante sono state identificate in 123 su 196 (63%) trial condotti in pazienti con artrite reumatoide¹⁰⁸.

Item 12b. Metodi utilizzati per analisi statistiche aggiuntive come le analisi per sottogruppi e le analisi aggiustate.

Esempio. “Le percentuali di pazienti che hanno risposto al trattamento sono state confrontate tra i due gruppi con il test χ^2 Mantel-Haenszel, aggiustate per la variabile di stratificazione, ossia l’impiego di metotrexate¹⁰³.”

“Le analisi per sottogruppi pre-definite in base all’assegnazione del trattamento antiossidante, alla presenza/assenza di precedenti cardiovascolari, all’assunzione di acido folico con la dieta, all’abitudine al fumo, al diabete, all’uso di aspirina, alla terapia ormonale e all’utilizzo di complessi multivitaminici sono state effettuate utilizzando i modelli di Cox di regressione proporzionale dei rischi. Queste analisi hanno utilizzato le valutazioni delle esposizioni di base e sono state limitate ai partecipanti per i quali erano disponibili i dati basali sottogruppo¹⁸¹.”

Spiegazione. Analogamente alle analisi primarie, anche quelle per sottogruppi devono essere chiaramente specificate. Le analisi più affidabili sono, in particolare, quelle che vogliono rilevare una differenza dell’efficacia del trattamento in sottogruppi complementari (es. partecipanti anziani vs giovani), un confronto noto come test di interazione^{182,183}. Un approccio comune, ma fuorviante, è confrontare i *P-value* per analisi di efficacia separate in ciascun gruppo. Dedurre un effetto sottogruppo (interazione) da un *P-value* significativo e uno non significativo non è corretto¹⁸⁴ e determina un’elevata incidenza di falsi positivi.

Per il rischio elevato di risultati falsamente positivi, le analisi per sottogruppi sono spesso scoraggiate^{14,185}. Confronti per sottogruppi *post-hoc* (eseguiti dopo aver esaminato i dati) difficilmente vengono confermati da studi ulteriori e non sono molto attendibili.

In alcuni studi, lo sbilanciamento dei fattori prognostici viene corretto attraverso alcune tipologie di analisi di regressione multipla. Sebbene la necessità di

“aggiustamento” nei RCT è inferiore rispetto agli studi epidemiologici, un’analisi “aggiustata” può essere ragionevole, soprattutto se una o più variabili sono considerate prognostiche¹⁸⁶. Idealmente, le analisi “aggiustate” dovrebbero essere specificate nel protocollo dello studio (item 24): ad esempio, l’aggiustamento è spesso raccomandato per ciascuna variabile di stratificazione (item 8b), seguendo il principio che la strategia di analisi dovrebbe essere coerente con il disegno dello studio. Nei trial la decisione di effettuare analisi aggiustate non dovrebbe essere determinata dal fatto che esistono caratteristiche di base statisticamente significative (item 16)^{183,187}. Il rationale per ogni analisi aggiustata e i metodi statistici utilizzati dovrebbero essere chiaramente specificati: infatti, gli autori dovrebbero indicare la scelta delle variabili che sono state aggiustate, dichiarare come le variabili continue sono state manipolate e specificare se l’analisi era stata pianificata (*pre-planned*) o suggerita dai dati (*post hoc*)¹⁸⁸. Alcune revisioni di trial pubblicati indicano che il reporting delle analisi aggiustate è inadeguato per tutti questi aspetti¹⁸⁸⁻¹⁹¹.

RISULTATI

Item 13. Flusso dei partecipanti (fortemente raccomandato l’uso di un diagramma di flusso).

Item 13a. Per ciascun gruppo riportare i partecipanti randomizzati, quelli che hanno ricevuto il trattamento previsto e quelli inclusi nell’analisi per l’outcome primario.

Esempi. Figura 2 e figura 3.

Spiegazione. Il disegno e la conduzione di alcuni RCT è semplice, e il relativo flusso dei partecipanti, in particolare in assenza di persi al follow-up e/o di partecipanti esclusi dall’analisi, può essere descritto adeguatamente in poche frasi per ciascuna fase dello studio. Nel caso di trial più complessi, può essere difficile per i lettori distinguere se e perché alcuni partecipanti non hanno ricevuto il trattamento assegnato, e/o sono stati persi al follow-up, e/o sono stati esclusi dall’analisi⁵¹. Queste informazioni sono fondamentali per vari motivi. In particolare, è improbabile che i partecipanti esclusi dopo l’assegnazione siano rappresentativi dell’intero campione: ad esempio, i pazienti potrebbero non essere disponibili per le valutazioni di follow-up perché hanno avuto una riacutizzazione della malattia o sperimentato effetti avversi del trattamento^{22,192}.

L’attrito conseguente ai persi al follow-up, spesso inevitabile, deve essere distinto dall’esclusione determinata dai ricercatori per motivi quali ineleggibilità, la sospensione del trattamento, la scarsa aderenza al protocollo del trial. Se i partecipanti vengono infatti esclusi dall’analisi si può giungere a conclusioni errate e gli sbilanciamenti in tali omissioni tra i gruppi possono essere particolarmente indicativi di bias¹⁹²⁻¹⁹⁴. È quindi molto importante riportare se i ricercatori hanno incluso nell’analisi tutti i partecipanti randomizzati, ciascuno nel gruppo originale di allocazione, [analisi intention-to-treat (item 16 e box 6)]. Conoscere il numero di partecipanti che non hanno ricevuto l’intervento, rispetto al gruppo di allocazione, o che non hanno completato il trattamento consente al lettore di valutare in che misura l’efficacia reale dell’intervento potrebbe essere sotto-stimata rispetto alle condizioni ideali.

Se disponibile, dovrebbe anche essere riportato il numero di partecipanti eleggibili: anche se questo numero è rilevante solo per la validità esterna ed è probabilmente meno importante rispetto ad altri fattori¹⁹⁵, indica che i partecipanti al trial sono, o meno, rappresentativi di tutti i partecipanti eleggibili.

Una revisione di RCT pubblicati nel 1998 in cinque importanti riviste di medicina ha riportato che il reporting del flusso dei partecipanti è spesso incompleto, in particolare per il numero di partecipanti che ricevono l’intervento assegnato e per il numero dei persi al follow-up⁵¹. Anche informazioni di base come il numero di partecipanti assegnati in maniera casuale e quelli esclusi dalle analisi non era disponibile nel 20% degli studi⁵¹.

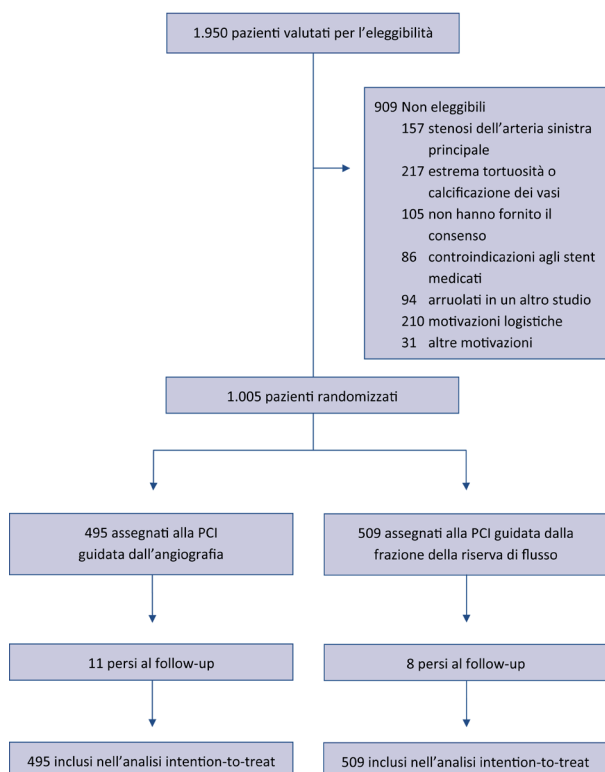


Figura 2. Diagramma di flusso di un trial multicentrico che confronta la frazione della riserva di flusso rispetto all’angiografia per guidare l’angioplastica coronarica percutanea – percutaneous coronary intervention (PCI). Il diagramma include informazioni dettagliate sui partecipanti esclusi (adattato da Tonino et al.³¹³)

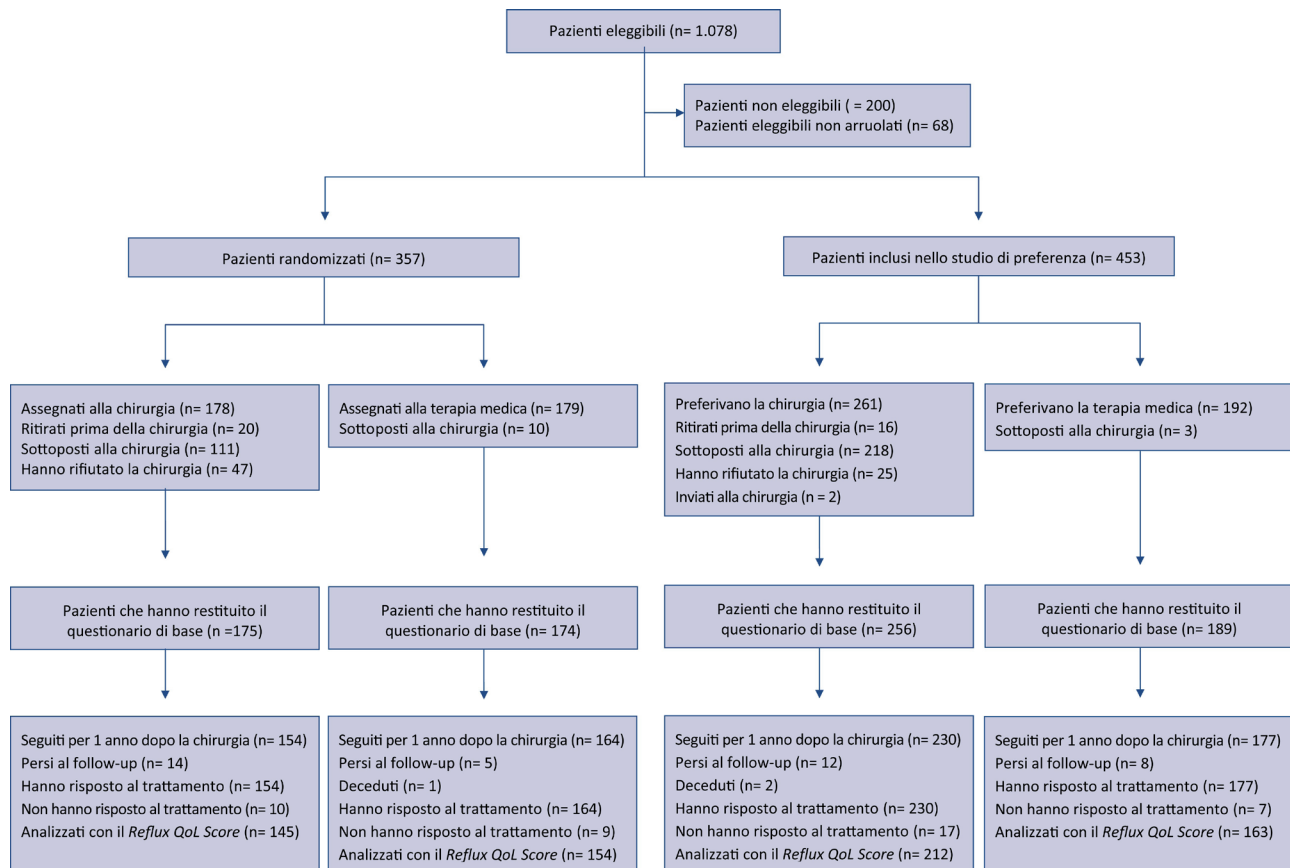


Figura 3. Diagramma di flusso di un trial che confronta l'efficacia della chirurgia minore con la terapia medica nei pazienti con reflusso gastro-esofageo cronico. Il diagramma mostra un trial multicentrico con un gruppo parallelo di preferenza non randomizzato (adattato da Grant et al.¹⁹⁶)

Il reporting era notevolmente più accurato nei trial che utilizzavano un diagramma del flusso dei partecipanti, come raccomandato dal CONSORT. Questo studio è stato utilizzato per effettuare una revisione del diagramma di flusso del CONSORT Statement⁵²⁻⁵⁴ (figura 1); i numeri richiesti sono descritti in dettaglio nella tabella 3.

Alcune informazioni (es. numero di individui eleggibili) non sempre sono disponibili¹⁴ e, in relazione alla natura di un trial, alcuni numeri possono essere più importanti di altri, quindi può essere utile o necessario adattare la struttura del diagramma di flusso a un determinato trial. In altri casi può essere utile prevedere informazioni aggiuntive: ad esempio, il diagramma di flusso di un trial a gruppi paralleli sulla chirurgia minore rispetto alla terapia medica per il reflusso gastro-esofageo cronico comprendeva anche un gruppo parallelo di preferenza non randomizzato (figura 3)¹⁹⁶.

La struttura e il contenuto del diagramma di flusso possono pertanto essere modificati in base alle caratteristiche specifiche del trial: ad esempio, molti trial che valutano l'efficacia di un intervento chirurgico o di una vaccinazione non prevedono la possibilità di interruzione.

Anche se il CONSORT raccomanda fortemente di utilizzare questo approccio grafico per riportare il flusso dei partecipanti nel corso dell'intero studio, non esiste un formato specifico suggerito.

Item 13b. Per ciascun gruppo riportare i partecipanti persi al follow-up e quelli esclusi dopo la randomizzazione, con le relative motivazioni.

Esempi. "C'è stata una sola violazione di protocollo: una donna nel gruppo di studio, che presentava una misura anormale del bacino, alla quale è stato programmato il taglio cesareo. Tuttavia, l'ostetrica ha ritenuto eseguibile una prova di travaglio: il taglio cesareo è stato praticato quando non c'era alcuna progressione della prima fase di travaglio¹⁹⁷".

"Il monitoraggio ha determinato il ritiro di nove centri, dove non si è potuta dimostrare la presenza di alcuni pazienti o dove si erano verificate altre gravi violazioni al protocollo¹⁹⁸".

Spiegazione. Alcune variazioni al protocollo possono essere riportate nel diagramma di flusso (item 13a), come ad esempio i partecipanti che non hanno ricevuto

Tabella 3. Informazioni richieste per riportare il flusso dei partecipanti nelle varie fasi di un trial randomizzato

Fase	Numero di partecipanti inclusi	Numero di partecipanti non inclusi o esclusi	Razionale
Arruolamento	Partecipanti valutati per il potenziale arruolamento	Partecipanti che non presentavano i criteri di inclusione o che, pur presentandoli, hanno rifiutato l'arruolamento	Questi numeri indicano se i partecipanti arruolati erano rappresentativi di tutti quelli osservati. Fondamentali per la valutare la validità esterna, ma spesso non disponibili
Randomizzazione	Partecipanti randomizzati	-	Numeri fondamentali per definire la dimensione del trial e valutare se un trial è stato analizzato secondo l'analisi intention-to-treat
Assegnazione dell'intervento	Partecipanti che hanno ricevuto l'intervento assegnato, secondo il gruppo originario di assegnazione	Partecipanti che non hanno ricevuto l'intervento assegnato, secondo il gruppo originario di assegnazione	Numeri importanti per valutare la validità interna e interpretare i risultati; dovrebbero essere specificate le motivazioni per cui i partecipanti non hanno ricevuto l'intervento assegnato
Follow-up	Partecipanti che hanno completato il trattamento, secondo il gruppo originario di assegnazione	Partecipanti che non hanno completato il trattamento, secondo il gruppo originario di assegnazione	Numeri importanti per valutare la validità interna e l'interpretazione dei risultati; dovrebbero essere specificate le motivazioni per cui i partecipanti non hanno completato il trattamento o il follow-up
	Partecipanti che hanno completato il follow-up previsto, secondo il gruppo originario di assegnazione	Partecipanti che non hanno completato il follow-up previsto, secondo il gruppo originario di assegnazione	
Analisi	Partecipanti inclusi nell'analisi principale, secondo il gruppo originario di assegnazione	Partecipanti esclusi dall'analisi principale, secondo il gruppo originario di assegnazione	Numeri importanti per valutare se un trial è stato analizzato mediante analisi intention-to-treat: dovrebbero essere specificate le motivazioni per l'esclusione di partecipanti dall'analisi

l'intervento previsto. Se i partecipanti sono stati esclusi dopo la randomizzazione (contrariamente ai principi dell'analisi per *intention-to-treat*) perché non presentavano i criteri di eleggibilità (item 16), dovrebbero essere inclusi nel diagramma di flusso. L'utilizzo del termine "variazione al protocollo" negli articoli pubblicati non è sufficiente a giustificare l'esclusione dei partecipanti dopo la randomizzazione: la natura della variazione al protocollo e l'esatta motivazione dell'esclusione dopo la randomizzazione dovrebbero essere sempre riportati.

Item 14a. Date relative ai periodi di reclutamento e di follow-up.

Esempio. "L'arruolamento dei partecipanti eleggibili per età si è svolto [...] dal febbraio 1993 al settembre 1994 [...] I partecipanti sono stati sottoposti alle visite cliniche al momento della randomizzazione (basale) e a intervalli di 6 mesi per i successivi 3 anni¹⁹⁹."

Spiegazione. Conoscere quando uno studio è stato condotto, oltre che i periodi di reclutamento dei parte-

cipanti, permette di collocare il trial in un preciso contesto storico. Terapie mediche e chirurgiche, inclusi i trattamenti concomitanti, si evolvono continuamente e possono influenzare l'assistenza convenzionale offerta ai partecipanti nel corso di un trial. Specialmente per altri ricercatori può anche risultare utile conoscere la velocità di reclutamento dei partecipanti.

La durata del follow-up non è sempre un periodo fisso successivo alla randomizzazione: infatti, in molti trial in cui l'end-point è costituito dal "tempo a un evento", il follow-up di tutti i partecipanti si conclude in una data specifica, che dovrebbe essere fornita insieme alla durata minima, massima e mediana del follow-up^{200,201}.

Una revisione di studi pubblicati in riviste di oncologia – la maggior parte dei quali non erano RCT²⁰¹ – che utilizzavano l'analisi di sopravvivenza ha rilevato che quasi l'80% (104/132 studi) includeva le date di inizio e di fine del reclutamento dei pazienti, ma solo il 24% (32/132 studi) riportava anche la data di conclusione del follow-up.

Item 14b. Motivazioni per cui il trial si è concluso o è stato interrotto.

Esempi. “Quando è stata eseguita l’analisi intermedia, il follow-up totale includeva il 63% stimato del numero totale di anni-paziente che sarebbero stati reclutati alla fine dello studio, determinando un valore soglia di 0,0095, in base al metodo di funzione d’uso (*spending function*) di α di Lan-DeMets [...] Al momento dell’analisi intermedia, il rischio relativo era di 0,37 nel gruppo di intervento, rispetto al gruppo di controllo, con un *P-value* di 0,00073, inferiore al valore soglia. Il board di monitoraggio dei dati e della sicurezza ha raccomandato ai ricercatori di interrompere il trial. Al gruppo di controllo, invitato a raggiungere il centro di ricerca, è stata consigliata e proposta la circoncisione medica [...] A causa dell’interruzione dello studio, alcuni partecipanti non avevano completato il follow-up, e le loro visite non ancora effettuate sono descritte in questo articolo come “programmate²⁰²”.

“Nel gennaio 2000, i problemi di fornitura del vaccino hanno determinato la sostituzione temporanea nazionale della componente cellulare del vaccino combinato DPT/Hib con il vaccino acellulare anti-pertosse. Dal momento che questo vaccino ha un diverso profilo di reattività locale, abbiamo deciso di interrompere precocemente il trial²⁰³.”

Spiegazione. I ricercatori che eseguono arbitrariamente analisi intermedie non programmate dopo aver “accumulato” pochi eventi, senza utilizzare criteri statistici, rischiano di “raccolgere” dati casuali estremi, che spesso

sovrastimano l’efficacia dell’intervento²⁰⁴.

I lettori probabilmente trarranno conclusioni più deboli da un trial interrotto in modo *data-driven* rispetto a uno che riporta le sue conclusioni dopo aver raggiunto un obiettivo indipendente dai risultati. Per questo i trial dovrebbero specificare le motivazioni per cui il trial si è concluso (box 5): il reporting dovrebbe includere i fattori estrinseci al trial che hanno contribuito alla decisione di interrompere il trial, indicando sia chi ha preso tale decisione, sia il ruolo dell’ente finanziatore in questa decisione¹³⁴.

Da una revisione sistematica di 143 trial interrotti precocemente per benefici del trattamento si evince che la loro interruzione è avvenuta dopo aver “accumulato” in media 66 eventi, riscontrando un rischio relativo medio di 0,47 e una forte relazione tra il numero di eventi e le dimensioni dell’effetto. I trial di dimensioni minori con un numero di eventi inferiore hanno prodotto i maggiori effetti del trattamento (OR 31, IC 95% da 12 a 82)¹³⁴. Sebbene un numero sempre più elevato di trial pubblicati in riviste biomediche a elevato impact factor riporti l’interruzione precoce, solo lo 0,1% riporta l’interruzione precoce per benefici, un dato che contrasta con le stime derivanti da studi di simulazione²⁰⁵ e con le indagini dei comitati di monitoraggio dei dati²⁰⁶. Di conseguenza, è verosimile che molti piccoli trial che riportano una notevole efficacia del trattamento siano stati interrotti precocemente, ma non lo hanno riportato.

Box 5. Interruzione precoce del trial

I trial possono essere interrotti quando raggiungono la dimensione del campione stimata, il numero di eventi atteso, la durata del follow-up definito o la data prevista di completamento. In questi casi il trial si concluderà indipendentemente dai suoi risultati, rendendo improbabile l’introduzione di bias. In alternativa, i trial possono essere interrotti precocemente perché il risultato di un’analisi intermedia mostra benefici o effetti avversi superiori a quelli attesi per l’intervento sperimentale, oppure quando i ricercatori non riscontrano alcuna differenza importante tra gli interventi (interruzione per inutilità). Inoltre, i trial possono essere interrotti precocemente per altre motivazioni: ad esempio viene meno il finanziamento, oppure i ricercatori non possono accedere ai pazienti arruolati o agli interventi in studio, o ancora quando i risultati di altri studi rendono irrilevante il quesito di ricerca.

Il reporting completo delle motivazioni di conclusione di un trial è importante per un processo decisionale *evidence-based* (item 14b). Su 143 studi interrotti prematuramente per benefici, molti non hanno riportato informazioni metodologiche fondamentali su come è stata presa questa decisione: in particolare, la dimensione del campione stimata ($n = 28$), le analisi intermedie che hanno determinato l’interruzione del trial ($n = 45$) o se uno specifico criterio di interruzione (*stopping rule*) ha determinato la decisione ($n = 48$)¹³⁴. L’item 7b della checklist prevede il reporting delle tempistiche delle analisi intermedie, i fattori che le hanno determinate, quante ne sono state effettuate, se erano pianificate o effettuate ad hoc e se erano stati stabiliti a priori criteri statistici di interruzione del trial. Inoltre, è utile sapere se un comitato di monitoraggio dei dati indipendente ha partecipato alle analisi (e da chi era composto, con particolare attenzione alle fonti di finanziamento) e chi ha deciso di interrompere il trial. Spesso il comitato di monitoraggio dei dati formula le raccomandazioni e i finanziatori (sponsor) o i ricercatori prendono la decisione di interrompere il trial.

I trial che vengono interrotti precocemente per ragioni indipendenti dai risultati e quelli che raggiungono il termine previsto, non rischiano di introdurre bias²⁰⁷. In questi casi, gli autori dovrebbero specificare se sono state effettuate analisi intermedie e se i loro risultati erano a disposizione del finanziatore.

I trial in cui vengono apportate variazioni all’intervento in risposta a risultati intermedi, consentendo così una valutazione più rapida di interventi sanitari promettenti per malattie mortali e/o con rapida evoluzione, richiederanno un reporting ancora più accurato del processo e della decisione di interrompere precocemente il trial²⁰⁸.

Item 15. Tabella che riporta le caratteristiche demografiche e cliniche di base per ciascun gruppo.

Esempio. Tabella 4

Spiegazione. Anche se i criteri di inclusione (item 4a) specificano quali soggetti erano eleggibili per il trial, è importante conoscere le caratteristiche dei partecipanti realmente arruolati. Questa informazione consente ai lettori, in particolare ai professionisti sanitari, di valutare la rilevanza e l'applicabilità dei risultati di un trial per i loro pazienti.

I RCT hanno l'obiettivo di confrontare gruppi di partecipanti nei quali l'unica differenza è l'intervento (trattamento) in studio. Una corretta assegnazione casuale previene il bias di selezione, ma non garantisce sempre che i gruppi abbiano caratteristiche di base omogenee: eventuali differenze sono, tuttavia, dovute al caso piuttosto

che a bias³². I gruppi di studio dovrebbero essere confrontati per le caratteristiche demografiche e cliniche di base in modo che i lettori possano valutarne la comparabilità. I dati di base sono particolarmente importanti per gli esiti che possono essere misurati anche all'inizio del trial (es. la pressione arteriosa).

Il modo migliore per presentare le caratteristiche di base è una tabella (tabella 4): per variabili continue, come il peso o la pressione arteriosa, dovrebbe essere riportata la variabilità dei dati, con i valori medi. Le variabili continue possono essere riassunte per ciascun gruppo con la media e la deviazione standard. Quando i dati continui hanno una distribuzione asimmetrica, l'approccio preferibile potrebbe essere quello di citare la media e un range di centili (come 25° e 75° centile)¹⁷⁷.

Tabella 4. Esempio di reporting delle caratteristiche demografiche e cliniche di base* (adattato dalla tabella 1 di Yusuf et al.²⁰⁹)

	Telmisartan (N= 2.954)	Placebo (N= 2.972)
Età (anni)	66,9 (± 7,3)	66,9 (± 7,4)
Genere (femmina)	1.280 (43,3%)	1.267 (42,6%)
Abitudine al fumo		
Attuale	293 (9,9%)	289 (9,7%)
Precedente	1.273 (43,1%)	1.283 (43,2%)
Etnia		
Asiatica	637 (21,6%)	624 (21,0%)
Araba	37 (1,3%)	40 (1,3%)
Africana	51 (1,7%)	55 (1,9%)
Europea	1.801 (61,0%)	1820 (61,2%)
Nativa o aborigena	390 (13,2%)	393 (13,2%)
Altra	38 (1,3%)	40 (1,3%)
Pressione arteriosa (mmHg)		
Sistolica	140,7 (± 16,8)	141,3 (± 16,4)
Diastolica	81,8 (± 10,1)	82,0 (± 10,2)
Frequenza cardiaca (bpm)	68,8 (± 11,5)	68,8 (± 12,1)
Colesterolo (mmol/l)		
Totale	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)
Malattia coronarica	2.211 (74,8%)	2.207 (74,3%)
Infarto del miocardio	1.381 (46,8%)	1.360 (45,8%)
Angina pectoris	1.412 (47,8%)	1.412 (47,5%)
Arteriopatia periferica	349 (11,8%)	323 (10,9%)
Ipertensione	2.259 (76,5%)	2.269 (76,3%)
Diabete	1.059 (35,8%)	1.059 (35,6%)

* I dati sono espressi come medie (± DS) o numeri (%)

Errori standard e limiti di confidenza non sono appropriati per descrivere la variabilità, perché misure di statistica inferenziale e non descrittiva. Le variabili con un piccolo numero di categorie ordinate (come gli stadi da I a IV della malattia) non dovrebbero essere considerate variabili continue, ma per ciascuna categoria dovrebbero essere riportati numeri assoluti e percentuali^{48,177}.

Purtroppo i test di significatività delle differenze di base sono ancora comuni^{23,32,210}: sono stati segnalati nella metà dei 50 RCT pubblicati nelle principali riviste generaliste nel 1997¹⁸³. Tali test di significatività valutano la probabilità che le differenze di base siano imputabili al caso, cosa che in realtà è già nota. Di conseguenza, questi test non sono errati, ma superflui, privi di logica²¹¹, e possono indurre in errore ricercatori e lettori. Piuttosto, i confronti di base dovrebbero considerare la potenza prognostica delle variabili misurate e l'entità degli sbilanciamenti casuali che si sono verificati²¹¹.

Item 16. Per ciascun gruppo riportare il numero dei partecipanti (denominatore) inclusi in ciascuna analisi e se l'analisi è stata eseguita secondo i gruppi di assegnazione originari (*intention-to-treat analysis*).

Esempi. "L'analisi primaria era un'analisi *intention-to-treat* e ha coinvolto tutti i pazienti randomizzati²¹²."

"Un paziente nel gruppo trattato con alendronato è stato perso al follow-up; di conseguenza, erano disponibili i dati di 31 pazienti per l'analisi *intention-to-treat*. Considerato che cinque pazienti costituivano violazioni di protocollo [...] solo 26 pazienti sono stati inclusi nella *per-protocol analysis*²¹³."

Spiegazione. Il numero di partecipanti in ciascun gruppo è un elemento fondamentale per le analisi dei dati. Sebbene il diagramma di flusso (item 13a) riporti il numero dei partecipanti inclusi nell'analisi, spesso i numeri variano in relazione ad outcome differenti. Il numero di partecipanti per gruppo dovrebbe essere fornito per tutte le analisi. Per gli outcome binari (come il rapporto di rischio e la differenza di rischio) devono essere riportati anche i denominatori o l'incidenza dell'evento, perché esprimere i risultati come frazioni aiuta il lettore a valutare se alcuni dei partecipanti randomizzati sono stati esclusi dall'analisi; pertanto, i risultati non dovrebbero mai essere presentati solo come misure aggregate (rischio relativo, riduzione del rischio relativo, etc.).

A volte alcuni pazienti non eleggibili vengono randomizzati per errore, oppure possono ricevere l'intervento solo in parte. Per gestire tali situazioni si raccomanda di analizzare tutti i partecipanti secondo i gruppi di assegnazione originari, indipendentemente da quello che si è verificato successivamente (box 6). Questa tipologia di analisi, definita *intention-to-treat* (letteralmente, secondo l'intenzione al trattamento) non è sempre facile da attuare. Considerato che è abbastanza comune che alcu-

ni pazienti non completino lo studio – persi al follow-up, sospensione del trattamento – tali pazienti non vengono valutati alla conclusione del trial. Se l'outcome è la mortalità, tali pazienti possono essere inclusi nell'analisi sulla base delle informazioni registrate, mentre può essere necessario utilizzare le metodologie di assunzione se mancano altri dati di outcome. Il termine *intention-to-treat* è spesso utilizzato impropriamente, ad esempio, quando coloro che non hanno ricevuto la prima dose di un farmaco sono esclusi dalle analisi¹⁸.

Al contrario, l'analisi può essere limitata solo ai partecipanti che soddisfano il protocollo in termini di eleggibilità, somministrazione degli interventi in studio e valutazione degli outcome. Questa tipologia di analisi è nota come *per-protocol analysis* o *on-treatment analysis*. Escludere partecipanti dalle analisi può portare a conclusioni errate: ad esempio, in un trial che ha confrontato nei pazienti con stenosi carotidea la terapia medica con quella chirurgica, l'analisi limitata ai partecipanti che hanno completato per il follow-up (*per-protocol*) ha mostrato che la chirurgia riduce il rischio di attacco ischemico transitorio, di stroke e la mortalità. Tuttavia, l'analisi *intention-to-treat* eseguita includendo tutti i partecipanti, secondo i gruppi originali di assegnazione, non ha dimostrato una maggiore efficacia della terapia chirurgica²¹⁴.

L'analisi *intention-to-treat* è generalmente preferita perché previene i bias conseguenti alla perdita non casuale di partecipanti²¹⁵⁻²¹⁷. Indipendentemente dal fatto che gli autori utilizzino il termine *intention-to-treat*, occorrerebbe sempre specificare quali e quanti partecipanti sono inclusi in ogni analisi (item 13). L'assenza di compliance agli interventi assegnati può far sì che l'analisi *intention-to-treat* sottovaluti i potenziali benefici dell'intervento, rendendo quindi opportuno eseguire analisi accessorie, come l'analisi per protocollo^{218,219}, tuttavia spesso notevolmente imprecise²²⁰.

In una revisione di 403 trial pubblicati nel 2002 su 10 importanti riviste mediche, 249 (62%) hanno riportato l'impiego dell'analisi secondo *intention-to-treat* come analisi primaria. Questa percentuale è stata più elevata per le riviste aderenti al CONSORT Statement (70% vs 48%). Tra gli articoli che hanno riportato l'utilizzo dell'analisi *intention-to-treat*, solo il 39% ha realmente incluso nell'analisi tutti i partecipanti randomizzati, mentre oltre il 60% dei trial presentava dati mancanti nell'analisi primaria²²¹. Altri studi mostrano simili risultati^{18,222,223}. I trial che non riportano i pazienti esclusi dalle analisi, rispetto a quelli che li riportano¹⁷³, hanno una minore qualità metodologica anche in altri item, indicando chiaramente che alcuni ricercatori non segnalano volutamente l'esclusione dei partecipanti dalle analisi. Un altro studio ha riscontrato che il reporting dell'analisi *intention-to-treat* si associa ad altri adeguati aspetti

Box 6. Analisi *intention-to-treat*

Un vantaggio importante dei RCT è quello di evitare bias durante l'assegnazione dei partecipanti a ciascun gruppo di intervento (box 1), arrivando a conclusioni certe su causa ed effetto, non possibili con altri disegni di studio. Per preservare integralmente i benefici della randomizzazione bisognerebbe includere nell'analisi tutti i partecipanti randomizzati, ovvero tutti quelli originariamente presenti nei gruppi di assegnazione. Queste due condizioni definiscono l'analisi *intention-to-treat*, fortemente raccomandata dal punto di vista metodologico^{18,223}. L'analisi *intention-to-treat*, che consente di analizzare i gruppi secondo la randomizzazione, è spesso difficile da effettuare per due principali ragioni: la mancanza di dati di outcome per alcuni partecipanti e la mancata aderenza al protocollo del trial.

Mancanza di outcome. Molti ricercatori escludono i pazienti per i quali non sono disponibili dati di esito: anche se spesso questo appare ragionevole, occorre tenere presente che se qualche partecipante randomizzato viene escluso l'analisi non può essere considerata *intention-to-treat*. In realtà, la maggior parte dei trial randomizzati è carente di alcune osservazioni: i ricercatori devono effettivamente scegliere tra escludere i partecipanti senza dati di esiti finali o formulare assunzioni sui dati mancanti di outcome²²⁵. Un'analisi secondo "caso completo" (o "caso disponibile") include solo i partecipanti il cui outcome è noto. Se è vero che la mancanza di risultati per un paio di partecipanti non causa alcun problema, uno studio documenta che nella metà dei trial esaminati mancano i dati in oltre il 10% dei pazienti randomizzati²²⁶. Questo frequente approccio genera una perdita di potenza riducendo le dimensioni del campione e può introdurre bias se la perdita al follow-up è legata alla risposta del paziente al trattamento, in particolare quando la frequenza o le cause di abbandono differiscono tra i gruppi di intervento.

I partecipanti per i quali mancano risultati possono essere inclusi nell'analisi solo se i loro risultati sono dedotti (ovvero stimati da altre informazioni raccolte). La deduzione dei dati mancanti consente un'analisi conforme ai criteri *intention-to-treat*, ma richiede assunzioni forti, talora difficili da giustificare²²⁷. Metodi di assunzione semplici, seppure attraenti, sono da sconsigliare: in particolare, un metodo ampiamente utilizzato è quello "dell'ultima osservazione effettuata", in cui i valori mancanti di outcome finali sono sostituiti dall'ultimo valore noto prima che il partecipante fosse perso al follow-up. Questo metodo, non tenendo conto dell'incertezza dell'assunzione²², può introdurre bias²²⁸ ed è stato severamente criticato da molti autori²²⁹⁻²³¹.

Mancata aderenza al protocollo. Diverso è il caso in cui il protocollo del trial non è stato interamente seguito da alcuni partecipanti: ad esempio quelli che non presentavano i criteri d'inclusione (es. diagnosi errata o età troppo giovane), che hanno ricevuto un co-intervento non consentito, che non hanno assunto tutti i trattamenti stabiliti, o che hanno ricevuto un trattamento diverso o nessun intervento. Il modo più semplice per far fronte a eventuali deviazioni dal protocollo è di ignorarle: tutti i partecipanti possono essere inclusi nell'analisi, indipendentemente dall'aderenza al protocollo (*intention-to-treat*). Si definisce "analisi *intention-to-treat* modificata" o "analisi per protocollo" un'analisi che esclude i partecipanti che non hanno aderito adeguatamente al protocollo, in particolare quelli che non hanno ricevuto una quantità dell'intervento minima predefinita²³². Sebbene l'analisi per protocollo possa essere appropriata in alcuni setting, dovrebbe essere correttamente definita come un confronto osservazionale non randomizzato. L'eventuale esclusione dei pazienti dall'analisi compromette la randomizzazione e può introdurre bias nei risultati.

Come "*intention-to-treat*", nessuna di queste altre definizioni chiarisce esattamente in modo affidabile quali pazienti sono stati inclusi. Per questo, nella checklist del CONSORT abbiamo mantenuto, per l'analisi *intention-to-treat*, la richiesta specifica di descrivere esattamente i partecipanti inclusi in ciascuna analisi.

di disegno e reporting dello studio, come la descrizione delle metodologie utilizzate per stimare la dimensione del campione²²⁴.

Item 17a. Per ciascun outcome primario e secondario, i risultati per ogni gruppo, la stima puntiforme dell'effetto e la sua precisione (limiti di confidenza al 95%).

Esempi. Tabella 5 e tabella 6.

Spiegazione. Per ciascun outcome, i risultati dello studio devono essere riportati come una sintesi degli outcome di ogni gruppo (es. il numero di partecipanti, con o senza evento e i denominatori o la media e la deviazione standard delle misure), insieme al confronto tra i gruppi, conosciuto come dimensione dell'effetto. Per gli

outcome binari (dicotomici), la dimensione dell'effetto potrebbe essere il rapporto di rischio (rischio relativo), l'*odds ratio*, o la differenza di rischio; per i dati di sopravvivenza, il rapporto di rischio (*hazard ratio*) o la differenza nella sopravvivenza mediana, mentre per i dati continui di solito la differenza delle medie. È necessario indicare i limiti di confidenza per il confronto tra i gruppi: è un errore comune presentare limiti di confidenza separati per l'esito di ciascun gruppo, piuttosto che per l'efficacia del trattamento²³³. Per la presentazione dei risultati è preferibile utilizzare una tabella piuttosto che descriverli nel testo (tabella 5, tabella 6).

Per tutti gli outcome, gli autori dovrebbero fornire i limiti di confidenza per indicare la precisione (incertez-

Tabella 5. Esempio di reporting di outcome dicotomici per ciascun gruppo* (adattato dalla tabella 2 di Mease et al¹⁰³)

Endpoint	Etanercept (n=30)	Placebo (n=30)	Riduzione del rischio assoluto (IC 95%)
End-point primario			
PsARC raggiunto dopo 12 settimane	26 (87%)	7 (23%)	63% (44 – 83)
Endpoint secondario			
Percentuale di pazienti che soddisfano i criteri ACR			
ACR20	22 (73%)	4 (13%)	60% (40 – 80)
ACR50	15 (50%)	1 (3%)	47% (28 – 66)
ACR70	4 (13%)	0 (0%)	13% (1 – 26)

* Vedi anche l'esempio dell'item 6a

PsARC= criteri di risposta al trattamento dell'artrite psoriasica (*Psoriatic Arthritis Response Criteria*)

ACR=American College of Rheumatology

Tabella 6. Esempio di reporting di outcome continui per ciascun gruppo (adattato dalla tabella 3 di van Linschoten²³⁴)

	Terapia con esercizio fisico (n=65)		Controllo (n=66)		Differenza aggiustata* a 12 mesi (IC 95%)
	Basale (media ± DS)	12 mesi (media ± DS)	Basale (media ± DS)	12 mesi (media ± DS)	
Score di funzionalità (0-100)	64,4 ± 13,9	83,2 ± 14,8	65,9 ± 15,2	79,8 ± 17,5	4,52 (da -0,73 a 9,76)
Dolore a riposo (0-100)	4,14 ± 2,3	1,43 ± 2,2	4,03 ± 2,3	2,61 ± 2,9	-1,29 (da -2,16 a -0,42)
Dolore durante l'attività (0-100)	6,32 ± 2,2	2,57 ± 2,9	5,97 ± 2,3	3,54 ± 3,38	-1,19 (da -2,22 a -0,16)

*Score di funzionalità aggiustato per valori di base, età e durata dei sintomi.

za) della stima^{48,235}. I limiti di confidenza convenzionali sono al 95%, ma a volte vengono utilizzati altri valori. Molte riviste incoraggiano l'uso dei limiti di confidenza²³⁶, particolarmente utili per le differenze tra i gruppi in studio che non raggiungono la significatività statistica, perché indicano se il risultato non esclude un'importante differenza clinica. L'utilizzo dei limiti di confidenza è notevolmente aumentato negli ultimi anni, anche se non in tutte le specialità mediche²³³. I *P-value* possono essere riportati insieme ai limiti di confidenza, ma i risultati non devono essere mai essere riportati solo come *P-value*^{237,238}. I risultati dovrebbero essere riportati per tutti gli end-point primari e secondari predefiniti, non solo per le analisi statisticamente significative o ritenute "interessanti": infatti, il reporting selettivo degli outcome di un trial è una problematica grave e molto diffusa^{55,57}. Nei trial in cui sono state effettuate analisi intermedie (ad-interim), l'interpretazione dovrebbe concentrarsi solo sui dati finali²³⁹. Sia per gli outcome dicotomici, sia per il tempo di sopravvivenza, può essere utile esprimere i risultati anche come numero necessa-

rio da trattare – *number-needed-to-treat/harm* – per ottenere un beneficio o per osservare un effetto avverso del trattamento (item 21)^{240,241}.

Item 17b. Per gli outcome dicotomici è raccomandata la presentazione dell'efficacia dell'intervento sia con misure assolute, sia relative.

Esempio. "Il rischio di dipendenza dall'ossigeno o di morte era ridotto del 16% (IC 95% da 7 a 25). La differenza assoluta era -6,3% (IC 95% da 9,9% a -2,7%); la somministrazione precoce a 16 bambini evita quindi che 1 bambino muoia o rimanga dipendente a lungo termine dall'ossigeno (tabella 7)".

Spiegazione. Se l'outcome primario è dicotomico, devono essere riportati sia l'efficacia relativa (rischio relativo, odds ratio) sia quella assoluta (differenza di rischio), con i relativi limiti di confidenza, in quanto né la misura relativa né la misura assoluta offrono da sole un quadro completo dell'efficacia del trattamento e delle sue implicazioni. Differenti tipologie di lettori possono preferire il rischio relativo o quello assoluto, ma sia i medici sia

Tabella 7. Esempio di reporting delle dimensioni dell'effetto assolute e relative (adattato dalla tabella 3 dell'OSIRIS Collaborative Group²⁴²)

Outcome primario	Numero (%)		Rischio relativo (IC 95%)	Riduzione del rischio assoluto (IC 95%)
	Somministrazione precoce (n= 1.344)	Somministrazione selettiva ritardata (n=1.346)		
Morte o ossigeno-dipendenza alla data presunta del parto	429 (31,9)	514 (38,2)	0,84 (0,75 – 0,93)	6,3% (9,9 – 2,7)

i cittadini tendono a sovrastimare l'effetto quando viene presentato in termini di rischio relativo. La differenza di rischio è meno generalizzabile ad altre popolazioni rispetto al rischio relativo in quanto dipende dal rischio di base nel gruppo di controllo, che tende a variare tra le popolazioni. Per le malattie in cui l'outcome è comune, un rischio relativo vicino all'unità potrebbe indicare differenze clinicamente importanti in termini di salute pubblica. Al contrario, quando l'outcome è raro un elevato rischio relativo potrebbe non essere rilevante per la salute pubblica, ma solo per soggetti a rischio elevato.

Item 18. Risultati di tutte le altre analisi effettuate (analisi per sottogruppi, analisi aggiustate), distinguendo le analisi predefinite da quelle esplorative.

Esempio. “Sulla base di uno studio che ha suggerito che l'efficacia della somministrazione peri-operatoria di β -bloccanti è influenzata dal rischio coronarico, abbiamo predefinito la nostra analisi primaria per sottogruppi in relazione alle carte del rischio cardiovascolare. Abbiamo predefinito anche analisi per sottogruppi secondarie in relazione al genere, al tipo di chirurgia, al ricorso ad un'anestesia spinale o epidurale. Per tutte le analisi per sottogruppi abbiamo utilizzato il modello di Cox a rischi proporzionali che include il test di interazione, con un livello di significatività fissato a $p < 0,05$... La figura 3 mostra i risultati delle nostre analisi per sottogruppi predefinite e indica la consistenza degli effetti ... Le nostre analisi per sottogruppi non avevano una potenza adeguata per rilevare differenze modeste negli effetti di sottogruppo che ci si potrebbe aspettare in presenza di un reale effetto di sottogruppo¹⁰⁰.”

Spiegazione. Considerato che analisi multiple condotte sugli stessi dati rischiano di generare risultati falsamente positivi²⁴⁶, gli autori non dovrebbero cedere alla tentazione di effettuare un numero eccessivo di analisi per sottogruppi^{183,185,247}. Considerato che le analisi predefinite (*pre-planned*) nel protocollo del trial (item 24) sono molto più affidabili rispetto a quelle suggerite dai risultati (*post-hoc*), gli autori dovrebbero riportare quali analisi erano predefinite. Inoltre, poiché il reporting selettivo di analisi per sottogruppi potrebbe essere fonte di bias²⁴⁸, gli autori dovrebbero specificare se sono sta-

te effettuate analisi per sottogruppi, quali sottogruppi sono stati presi in esame con le relative motivazioni, e quanti sottogruppi erano predefiniti.

Quando si valuta un sottogruppo non si intende verificare se nel sottogruppo esiste un risultato statisticamente significativo, ma se gli effetti del trattamento sono significativamente diversi dagli altri sottogruppi. A tal fine è utile un test di interazione, tenendo conto che la loro potenza è generalmente bassa. Quando vengono condotti test formali di interazione (item 12b), i risultati dovrebbero essere riportati come differenza stimata nell'effetto dell'intervento in ciascun sottogruppo (con relativi limiti di confidenza) e non solo come *P-value*.

In una survey, 35/50 trial includevano analisi per sottogruppi, di cui solo il 42% utilizzava test di interazione¹⁸³: è stato spesso difficile determinare se si trattava di analisi per sottogruppi predefinite nel protocollo. In un'altra survey condotta su trial chirurgici pubblicati su riviste ad elevato impact factor, 27/72 trial hanno riportato 54 analisi per sottogruppi, di cui 91% erano *post-hoc* e solo il 6% prevedeva un test di interazione per valutare se esisteva un effetto di sottogruppo²⁴⁹.

Analoghe raccomandazioni sono valide per le analisi che hanno effettuato un aggiustamento” per le variabili di base (*adjusted analyses*). Qualora vengano condotte analisi aggiustate e non aggiustate, dovrebbero essere riportate entrambe, indicando se le analisi aggiustate, inclusa la scelta delle variabili da aggiustare, sono state predefinite. Idealmente, il protocollo del trial dovrebbe dichiarare se l'aggiustamento è effettuato per le variabili di base utilizzando l'analisi di covarianza¹⁸⁷. L'aggiustamento delle variabili motivato da significative differenze di base verosimilmente introdurrà bias nella stima dell'efficacia del trattamento¹⁸⁷. Una survey ha rilevato che le discrepanze non riconosciute tra protocolli e pubblicazioni erano presenti in 25 trial che riportavano analisi per sottogruppi e in 23/28 trial che effettuavano analisi aggiustate⁹².

Item 19. Tutti i rilevanti effetti avversi o indesiderati in ogni gruppo.

(per indicazioni specifiche consultare l'estensione del CONSORT sugli effetti avversi⁴²)

Esempio. “La percentuale di pazienti che ha sperimentato eventi avversi era simile tra il gruppo trattato con rBPI21 [proteina battericida ricombinante aumentante la permeabilità] e il gruppo placebo – rispettivamente 88,4% (168/190) e 88,7% (180/203) – ed era più bassa nei pazienti trattati con rBPI21 rispetto a quelli trattati con placebo per 11 di 12 sistemi corporei [...] la percentuale di pazienti che ha sperimentato un grave evento avverso era inferiore nel gruppo con rBPI21 rispetto al gruppo placebo: rispettivamente 27,9% (53/190) vs 36,5% (74/203). Si sono verificati solo tre gravi eventi avversi farmaco-correlati, tutti nel gruppo placebo²⁵⁰.”

Spiegazione. I lettori necessitano di informazioni su benefici ed effetti avversi degli interventi sanitari al fine di prendere decisioni razionali ed equilibrate. L'esistenza e la natura degli effetti avversi possono contribuire a stabilire se un particolare intervento è accettabile e utile. Non tutti gli eventi avversi osservati nel corso di un trial sono necessariamente una conseguenza dell'intervento: alcuni possono infatti derivare dalla condizione trattata. I RCT costituiscono l'approccio migliore per fornire le evidenze di efficacia e sicurezza, ma non sono in grado di rilevare effetti collaterali rari.

Inoltre, molti trial forniscono informazioni inadeguate sugli eventi avversi: una survey condotta su 192 trial di interventi farmacologici, pubblicati dal 1967 al 1999, ha mostrato che solo il 39% segnalava adeguatamente gli eventi avversi e il 29% la tossicità definita di laboratorio⁷². Più recentemente, un confronto tra i dati di eventi avversi presentati al registro di trial del *National Cancer Institute*, che aveva sponsorizzato i relativi trial, e le informazioni incluse nei report hanno permesso di riscontrare che gli eventi avversi di bassa severità erano sotto-riportati nelle riviste. Eventi molto severi (grado di severità da 3 a 5 secondo i *Common Toxicity Criteria*) sono stati riportati negli articoli in modo inconsistente, così come incompleta risultava la loro attribuzione ai farmaci sperimentali²⁵¹.

Infine, una revisione di trial pubblicati in sei riviste mediche generali dal 2006 al 2007 ha evidenziato che, sebbene l'89% di 133 studi riportasse gli eventi avversi, non veniva fornita alcuna informazione sui gravi eventi avversi (27% degli articoli), né sulla sospensione del trattamento a causa di un evento avverso (48% degli articoli)²⁵².

Una estensione del CONSORT Statement è stata sviluppata per fornire raccomandazioni dettagliate sul reporting di effetti avversi nei trial randomizzati⁴². Raccomandazioni ed esempi di reporting adeguati sono disponibili gratuitamente sul sito web del CONSORT (www.consort-statement.org). Esse completano il CONSORT Statement 2010 e dovrebbero essere consultate, in particolare se lo studio degli effetti avversi costituisce un obiettivo chiave del trial. Sostanzialmente, se i dati sugli

eventi avversi sono disponibili, gli eventi dovrebbero essere elencati e specificati, preferibilmente con riferimenti bibliografici ai criteri standardizzati. Occorre descrivere i metodi utilizzati per la raccolta di dati e l'assegnazione degli eventi. Per ogni braccio dello studio, devono essere specificati sia il rischio assoluto di ciascun evento avverso, utilizzando misure appropriate per eventi ricorrenti, sia il numero dei partecipanti che hanno sospeso il trattamento a causa di effetti avversi. Infine, gli autori dovrebbero fornire una discussione equilibrata su benefici e rischi dei trattamenti in studio⁴².

DISCUSSIONE

Item 20. Limiti del trial, considerando le fonti di potenziali bias, l'imprecisione e – se rilevanti – le analisi multiple.

Esempio. “La prevalenza di pazienti di sesso maschile (85%) rappresenta un limite del nostro studio [...] Abbiamo utilizzato stent di metallo nudo, poiché quelli medicati non erano disponibili sino ad arruolamento inoltrato. Sebbene questo aspetto possa essere percepito come un limite dello studio, i dati pubblicati non indicano alcun beneficio (a breve e a lungo termine) in termini di mortalità e di infarto del miocardio nei pazienti con malattia coronarica stabile nei quali vengono impiantati stent medicati, rispetto a quelli di metallo nudo²⁵³.”

Spiegazione. La sezione della discussione degli articoli scientifici è spesso piena di retorica per supportare i risultati ottenuti⁵⁴, ma riserva poco spazio ai pro e ai contro dello studio e ai suoi risultati. Alcune riviste hanno tentato di risolvere questo problema incoraggiando una discussione strutturata^{255, 256}. Ad esempio, *Annals of Internal Medicine* suggerisce agli autori di organizzare la sezione della discussione in: (1) breve sintesi dei principali risultati; (2) considerazioni su possibili meccanismi e spiegazioni; (3) confronto con i risultati rilevanti di altri studi pubblicati, possibilmente attraverso una revisione sistematica che combini i risultati dello studio con quelli di tutti gli studi rilevanti già pubblicati; (4) limiti dello studio e le strategie utilizzate per ridurli e compensarli; (5) breve sezione per riassumere le implicazioni cliniche e di ricerca dello studio²⁵⁵. Gli autori sono invitati a seguire questi suggerimenti, eventualmente suddividendo in paragrafi la sezione della discussione.

Sebbene la discussione dei limiti sia spesso omessa dagli articoli scientifici²⁵⁷, la loro identificazione ed approfondimento è in realtà molto importante²⁵⁸. Ad esempio, un team di chirurghi ha riportato che nei pazienti con colecistite acuta, la colecistectomia laparoscopica, procedura tecnicamente difficile, ha determinato un'incidenza di complicanze significativamente inferiore rispetto alla tradizionale colecistectomia in aperto²⁵⁹. Tuttavia, gli autori hanno omesso di presentare un evidente bias dello studio: infatti tutte le cole-

cistectomie laparoscopiche erano state effettuate dai ricercatori, mentre l'80% delle colecistectomie in aperto veniva effettuato da specialisti in formazione.

Gli autori dovrebbero anche discutere di tutte le imprecisioni dei risultati relative a diversi aspetti dello studio, inclusa la misurazione dell'outcome primario (item 6a) o la diagnosi (item 4a). Ad esempio, una scala validata negli adulti potrebbe essere utilizzata in una popolazione pediatrica, oppure il valutatore potrebbe non essere stato istruito sulle modalità di utilizzo.

La differenza tra la significatività statistica e la rilevanza clinica dovrebbe sempre essere tenuta presente: in particolare, gli autori dovrebbero evitare l'errore molto frequente di interpretare un risultato statisticamente non significativo come indicativo dell'equivalenza degli interventi a confronto. I limiti di confidenza (item 17a) forniscono informazioni preziose sul fatto che il risultato di un trial sia o meno compatibile con un effetto clinicamente rilevante, indipendentemente dal *P-value*¹²⁰.

Gli autori dovrebbero prestare particolare attenzione quando valutano i risultati del trial attraverso confronti multipli, perché tale molteplicità deriva da vari interventi, misure di outcome, time point, analisi per sottogruppi e altri fattori. In queste circostanze, alcuni risultati statisticamente significativi potrebbero essere dovuti solo al caso.

Item 21. Generalizzabilità (validità esterna, applicabilità) dei risultati del trial.

Esempi. "Poiché l'intervento è stato erogato ad uomini e donne di tutte le età, per tutti i tipi di sport e a diversi livelli agonistici, i risultati indicano che l'intera gamma di atleti, da giovani altamente selezionati ad atleti senior di ceto intermedio, potrebbe trarre benefici dal programma di allenamento per la prevenzione delle recidive della distorsione alla caviglia. Includendo nello studio gli atleti sottoposti a terapia non farmacologica e farmacologica, abbiamo coperto un ampio spettro di severità della lesione, concludendo che il programma di allenamento può essere attuato per il trattamento di tutti gli atleti. Inoltre, assumendo ragionevolmente che le distorsioni alla caviglia non correlate allo sport siano simili a quelle conseguenti all'attività sportiva, il programma potrebbe offrire benefici alla popolazione generale²⁶⁰."

"Questo riproduce ed estende il lavoro di Clarke et al. e dimostra che il programma cognitivo-comportamentale di prevenzione può essere erogato in maniera affidabile ed efficace in altri setting, anche da medici non appartenenti al gruppo che ha sviluppato originariamente l'intervento. La dimensione dell'effetto era consistente con quella già riportata in letteratura negli studi di prevenzione della depressione, eseguiti in singoli centri e tra centri diversi rispetto sia ai disturbi

depressivi sia ai sintomi [...] In questo trial di generalizzabilità abbiamo scelto un confronto rilevante per la salute pubblica: la terapia convenzionale [...] Il campione comprendeva prevalentemente individui dalla classe operaia alla classe media che usufruivano di assicurazioni sanitarie. Considerando che la terapia cognitivo-comportamentale può essere più efficace negli adolescenti provenienti da famiglie con redditi più elevati, è importante valutare l'efficacia di questo programma di prevenzione in campioni di etnia e condizioni socio-economiche differenti²⁶¹."

Spiegazione. La validità esterna, definita anche generalizzabilità o applicabilità, è il grado con cui i risultati di uno studio possono essere generalizzati ad altre popolazioni²⁶². La validità interna, strettamente legata alla capacità dello studio di prevenire i bias nelle fasi di pianificazione, conduzione e analisi, è un prerequisito della validità esterna: infatti, i risultati di un trial metodologicamente inadeguato non sono validi e la questione della sua validità esterna diventa irrilevante. Non esiste una validità esterna assoluta: il termine è significativo soltanto per le condizioni chiaramente specificate che non sono state valutate direttamente nel trial. I suoi risultati possono essere generalizzati a un singolo partecipante o a popolazioni differenti da quella arruolata nel trial rispetto all'età, al genere, alla gravità della malattia e alle comorbidità? I risultati sono applicabili ad altri farmaci della stessa classe, a dosi, tempi, vie di somministrazione e terapie concomitanti differenti? E' possibile attendersi risultati simili in setting di assistenza primaria, secondaria e terziaria? Quali effetti sugli esiti correlati non sono stati valutati nel trial? Quale rilevanza per la durata del follow-up e del trattamento, soprattutto per quanto riguarda l'insorgenza di effetti avversi?²⁶³

La validità esterna è una questione di giudizio e dipende dalle caratteristiche dei partecipanti inclusi nel trial, dal setting assistenziale, dai regimi dei trattamenti valutati e dagli outcome misurati^{5,136}. Pertanto, è fondamentale che lo studio fornisca informazioni adeguate su: criteri di inclusione, setting e area geografica (item 4b), interventi a confronto e relative modalità di somministrazione (item 5), definizione degli outcome (item 6), date relative ai periodi di reclutamento e di follow-up (item 14). È importante indicare la percentuale dei partecipanti del gruppo di controllo in cui si verifica l'esito (rischio del gruppo di controllo). La percentuale di partecipanti eleggibili che si rifiutano di partecipare al trial, come riportato nel diagramma di flusso (item 13), è rilevante per la generalizzabilità del trial, in quanto può indicare preferenze e accettabilità di un intervento. Simili considerazioni possono essere applicate alle preferenze dei professionisti sanitari^{264,265}.

Molti aspetti sono rilevanti nell'applicazione dei risultati di un trial al paziente individuale²⁶⁶⁻²⁶⁸: nonostan-

te la diversa risposta al trattamento tra il singolo paziente e quelli arruolati in un trial o inclusi in una revisione sistematica, le differenze tendono a riguardare l'entità e non la direzione dell'effetto del trattamento.

Nonostante rilevanti eccezioni²⁶⁸, i trattamenti (in particolare farmacologici²⁶⁹) di documentata efficacia in una popolazione ristretta di pazienti hanno in genere un'applicazione più ampia nella pratica. Sono stati proposti specifici *framework* per valutare la validità esterna dei trial: studi qualitativi, come le "valutazioni del processo" integrali²⁷⁰ e checklist²⁷¹. Le misure di efficacia terapeutica che incorporano il rischio basale, come il *number-needed-to-treat* (NNT) per ottenere un successo terapeutico e il *number-needed-to-harm* (NNH) per causare un effetto avverso, sono utili per valutare il rapporto benefici/rischi in pazienti con caratteristiche diverse da quelle dei partecipanti del trial^{268,272,273}. Infine, dopo aver ottenuto le stime centrate sul paziente relative ai potenziali benefici ed effetti avversi di un intervento, il medico deve integrarle con i valori e le preferenze del paziente. Analoghe considerazioni si applicano per valutare la generalizzabilità dei risultati di un trial in setting e con interventi differenti.

Item 22. Interpretazione coerente con i risultati, bilanciando benefici ed effetti avversi e tenendo conto di altre evidenze rilevanti.

Esempio. "Gli studi pubblicati prima del 1990 hanno suggerito che l'immunoterapia profilattica riduce anche le infezioni nosocomiali in neonati con peso alla nascita molto basso. Tuttavia, questi studi hanno arruolato piccoli numeri di pazienti, utilizzando disegni, dosi e preparazioni diversi e includendo popolazioni differenti. In questo ampio trial controllato e randomizzato multicentrico, la somministrazione profilattica ripetuta di immunoglobuline per via endovenosa non ha ridotto significativamente l'incidenza di infezioni nosocomiali in neonati prematuri con peso alla nascita compreso tra 501 e 1.500 grammi²⁷⁴."

Spiegazione. Per consentire ai lettori di conoscere come i risultati dello studio si integrano con quelli di altri trial, la strategia migliore è includere una formale revisione sistematica nella sezione dei risultati o della discussione^{83,275-277}. Se tale obiettivo può essere poco praticabile per gli autori di un trial, spesso è possibile citare una revisione sistematica di studi simili. Una revisione sistematica può aiutare i lettori a valutare se i partecipanti arruolati e i risultati del trial sono simili a quelli di trial analoghi. Il reporting dei RCT spesso non affronta questo aspetto in maniera adeguata²⁷⁷. Il metodo bayesiano può essere utilizzato per combinare statisticamente i dati di un trial con evidenze precedenti²⁷⁸.

Quale requisito minimo raccomandiamo che la discussione sia il più sistematica e ampia possibile, piuttosto

che limitata solo agli studi che supportano i risultati dello studio²⁷⁹.

ALTRE INFORMAZIONI

Item 23. Numero di registrazione e nome del registro del trial.

Esempio. "Il trial è registrato in ClinicalTrials.gov, numero NCT00244842²⁸⁰."

Spiegazione. Le conseguenze della mancata pubblicazione dei trial per esteso^{281,282}, del reporting selettivo degli outcome all'interno dei trial, e di analisi "per protocollo", piuttosto che "intention-to-treat", sono ben documentate^{55,56,283}. Anche le pubblicazioni ridondanti di trial possono rappresentare un problema, in particolare per gli autori di revisioni sistematiche che, inavvertitamente, includono più volte i risultati dello stesso trial²⁸⁴.

Per minimizzare o evitare questi problemi, negli ultimi 25 anni molti sforzi sono stati compiuti per registrare i trial al loro avvio con numeri unici di identificazione, includendo altre informazioni di base per rendere disponibili alla comunità scientifica i dettagli fondamentali dello studio²⁸⁵⁻²⁸⁸. Per gravi problemi dovuti alla mancata comunicazione di dati²⁸⁹, l'Organizzazione Mondiale della Sanità ha affermato che "la registrazione di tutti i trial è una responsabilità scientifica, etica e morale" (www.who.int/ictip/en). Registrando un trial randomizzato, gli autori riportano un set minimo di informazioni e ottengono un numero unico di registrazione del trial.

Nel settembre 2004 l'*International Committee of Medical Journal Editors* (ICMJE) ha modificato la sua policy, dichiarando di prendere in considerazione per la pubblicazione solo i trial preventivamente registrati prima dell'arruolamento del primo partecipante²⁹⁰, determinando uno straordinario aumento delle registrazioni di trial²⁹¹. L'ICMJE fornisce indicazioni sui registri validi (www.icmje.org/faq.pdf).

In una recente survey delle istruzioni per gli autori di 165 riviste ad elevato impact factor, 44 riviste hanno specificato che requisito fondamentale per sottomettere un trial alla pubblicazione è la sua registrazione²⁹².

Gli autori dovrebbero fornire il nome del registro e il numero unico di registrazione del trial: se gli autori non hanno registrato il trial dovrebbero specificarlo, riportandone le motivazioni.

Item 24. Dove è possibile reperire il protocollo completo del trial, se disponibile.

Esempio. "Il protocollo completo del trial è disponibile nell'appendice supplementare del testo integrale dell'articolo all'indirizzo: www.nejm.org"²⁹³.

Spiegazione. Il protocollo completo del trial (o quello relativo a specifiche procedure) è di fondamentale importanza perché predefinisce numerosi aspetti metodologici del RCT, tra cui l'outcome primario (item 6a). La

disponibilità del protocollo può aiutare a limitare la probabilità di variazioni post-hoc non dichiarate dai metodi del trial e il reporting selettivo degli outcome (item 6b). Gli elementi importanti da includere nel protocollo di un trial randomizzato sono descritti altrove²⁹⁴.

Gli autori hanno a disposizione varie opzioni per rendere accessibile il protocollo del trial ai lettori interessati. Come nell'esempio sopra riportato, le riviste che pubblicano i risultati di un trial possono rendere disponibile il protocollo attraverso il loro sito web. L'accessibilità a risultati e protocollo dei trial è maggiore nelle riviste open-access. Alcune riviste (come *Trials*) pubblicano protocolli dei trial che possono costituire il riferimento bibliografico nel reporting dei risultati principali dei trial. Inoltre, la registrazione del trial (item 23) garantisce la disponibilità di molti dettagli del protocollo, visto che le caratteristiche minime del trial incluse in un registro di trial comprendono vari dettagli del protocollo e dei risultati (www.who.int/ictrp/en). I ricercatori possono anche rendere disponibile il protocollo del trial attraverso il sito web della loro organizzazione. Qualunque sia il metodo utilizzato, incoraggiamo tutti i ricercatori a rendere il protocollo del loro trial facilmente accessibile ai lettori interessati.

Item 25. Fonti di finanziamento e altri supporti (es. fornitura dei farmaci), ruolo dei finanziatori.

Esempi. "È stato ricevuto un supporto finanziario da Plan International per l'intervento e da Wellcome Trust e Joint United Nations Programme on HIV/AIDS (UNAIDS) per la ricerca. I finanziatori non hanno avuto alcun ruolo nel disegno dello studio, nella raccolta e analisi dei dati, nella decisione di pubblicare e nella preparazione del manoscritto²⁹⁵."

"Questo studio è stato finanziato dalla GlaxoSmithKline Pharmaceuticals. La GlaxoSmithKline è stata coinvolta nel disegno e nella conduzione dello studio e ha fornito supporto logistico durante il trial. Dipendenti dello sponsor hanno collaborato con i ricercatori per pianificare le analisi statistiche eseguite presso l'Università dello Utah. Il manoscritto è stato redatto dal dottor Shaddy e dai membri dello steering committee. Alla GlaxoSmithKline è stato permesso di rivedere il manoscritto e di suggerire modifiche, ma la decisione finale sui contenuti è stata esclusivamente degli autori²⁹⁶."

Spiegazione. Gli autori dovrebbero specificare le fonti di finanziamento dello studio perché rappresentano un'informazione fondamentale per chi valuta un trial. Alcuni studi hanno dimostrato che la ricerca sponsorizzata dall'industria farmaceutica ha maggiori probabilità di produrre risultati a favore del farmaco prodotto dall'azienda sponsor rispetto agli studi finanziati da altre organizzazioni²⁹⁷⁻³⁰⁰. Una revisione sistematica di 30 studi sulle fonti di finanziamento ha dimostrato che la ricerca

finanziata dall'industria farmaceutica ha una probabilità quattro volte superiore di avere outcome favorevoli allo sponsor rispetto alla ricerca finanziata da altre organizzazioni (*odds ratio* 4.05; intervallo di confidenza al 95% 2.98 – 5.51)²⁹⁷. Un'ampia percentuale dei trial pubblicati non riporta ancora le fonti di finanziamento, ma l'entità di questo reporting inadeguato è difficile da quantificare. Una survey condotta su 370 trial farmacologici ha evidenziato che il 29% non riporta le fonti di finanziamento³⁰¹. In un'altra indagine su trial randomizzati indicizzati in PubMed pubblicati nel dicembre 2000, le fonti di finanziamento sono state riportate nel 66% di 519 trial¹⁶.

Il livello di coinvolgimento di chi finanzia un trial e l'influenza nel disegno, conduzione, analisi e reporting dello studio possono variare: di conseguenza, è importante che gli autori descrivano dettagliatamente il ruolo di chi ha finanziato lo studio, inclusa anche l'assenza di coinvolgimento. Allo stesso modo, gli autori dovrebbero riportare qualunque altro supporto, quali la fornitura e la preparazione di farmaci o apparecchiature, oppure durante l'analisi dei dati e la stesura del manoscritto³⁰.

Reporting di trial che non hanno un disegno a gruppi paralleli

Il target principale delle raccomandazioni del CONSORT sono i trial a disegno parallelo con due gruppi di intervento, che rappresentano la maggioranza. Tuttavia una percentuale non trascurabile di RCT non ha questo disegno: il 45% (233/519) di RCT pubblicati nel dicembre 2000¹⁶ e il 39% (242/616) nel dicembre 2006¹⁷.

La maggior parte delle raccomandazioni del CONSORT Statement può essere applicata anche ad altri varianti di trial che, tuttavia, presentano ulteriori complessità da prendere in considerazione. Prima della pubblicazione della revisione del CONSORT Statement nel 2001, il gruppo CONSORT ha deciso di sviluppare alcune estensioni relative a specifici disegni di trial: sono già state pubblicate quelle per i trial con randomizzazione cluster⁴⁰ e i trial di non-inferiorità e di equivalenza³⁹. Per la mancanza di risorse non sono ancora state completate le estensioni previste per i trial con disegno parallelo a più bracci, con disegno fattoriale, disegno crossover e i trial *within-person*.

Gli autori di trial con randomizzazione cluster o di trial di non-inferiorità o di equivalenza dovrebbero consultare le relative estensioni del CONSORT, oltre al presente documento. Qui ci occuperemo degli altri disegni, ricordando che il disegno dello studio deve essere ben specificato sia nel testo dell'articolo, sia nell'abstract.

Trial paralleli a più bracci (> 2 gruppi) necessitano quantomeno della modifica delle linee guida standard del CONSORT. Il diagramma di flusso può essere facilmente esteso: le differenze principali, rispetto ai trial a

due bracci, riguardano il fatto che le ipotesi dello studio e le conseguenti metodologie di analisi e interpretazione dei dati sono riferite a più di due gruppi. Per i trial con disegno fattoriale, occorre in genere considerare le potenziali interazioni tra gli interventi; oltre al confronto generale dei partecipanti che hanno ricevuto, o meno, ciascun intervento in studio, i ricercatori dovrebbero anche riportare i risultati per ciascuna combinazione degli interventi³⁰³.

Negli studi con disegno *cross-over*, ciascun partecipante riceve due (o più) trattamenti secondo una sequenza casuale: i principali aspetti da affrontare riguardano la natura doppia dei dati che influenza il disegno e l'analisi³⁰⁴. Considerazioni simili riguardano i confronti *within-person*, in cui i partecipanti ricevono contemporaneamente due trattamenti (spesso agli organi doppi). In entrambi i casi, a causa del rischio di effetti di "trascinamento" (*carry over*) temporali e sistemici, la scelta del disegno deve essere giustificata.

Il gruppo CONSORT intende pubblicare estensioni relative a tutti questi disegni. È inoltre prevista la pubblicazione di aggiornamenti alle raccomandazioni esistenti (trial con randomizzazione cluster e i trial di non-inferiorità e di equivalenza) per uniformarle a questo rilevante aggiornamento delle raccomandazioni standard del CONSORT.

Discussione

La valutazione degli interventi sanitari può essere adeguata solo se i ricercatori garantiscono confronti imparziali. In particolare, l'assegnazione casuale ai gruppi di studio rimane l'unica strategia per eliminare i bias di selezione e di confondimento: infatti, rispetto ai trial randomizzati, quelli non randomizzati tendono a sovrastimare l'efficacia degli interventi sanitari^{305,306}.

I bias, tuttavia, sono possibili anche nei RCT se i ricercatori non li conducono in maniera metodologicamente adeguata³⁰⁷. Una recente revisione sistematica, combinando i risultati di diversi studi metodologici, ha evidenziato che, per gli outcome soggettivi, i trial che utilizzavano un occultamento della lista di randomizzazione inadeguato o poco chiaro producevano stime dell'efficacia superiori del 31% rispetto a quelli che impiegavano un occultamento adeguato; nei trial senza *blinding* le stime erano superiori del 25%¹⁵³. Come è ragionevole aspettarsi, esiste una forte associazione tra i due aspetti metodologici.

Il disegno e la conduzione di un RCT richiedono competenze metodologiche e cliniche, impegno scrupoloso^{143,308}, e un elevato livello di attenzione per gestire le difficoltà impreviste. Allo stesso modo, al fine di limitare i bias, la stessa attenzione deve esser posta nel reporting dei trial. I lettori non dovrebbero avere alcun dubbio: le metodologie utilizzate devono essere ben specificate e

trasparenti, per consentire di distinguere facilmente i trial con risultati non distorti da quelli con risultati discutibili. Adeguate evidenze scientifiche prevedono un reporting adeguato e la conduzione di trial eticamente accettabili si basa su consistenti evidenze scientifiche³⁰⁹.

Ci auguriamo che questo aggiornamento dell'articolo esplicativo del CONSORT sia utile agli autori per utilizzare la versione 2010 del CONSORT e sottolinei l'importanza di un adeguato reporting dei trial. Il CONSORT Statement può aiutare i ricercatori a progettare trial futuri¹⁰ e può guidare i revisori nella valutazione dei manoscritti, semplicemente verificando se gli autori hanno o meno riportato gli item della checklist del CONSORT. Tali valutazioni probabilmente miglioreranno la chiarezza e la trasparenza dei trial pubblicati. Poiché il CONSORT è un documento in continua evoluzione, richiede un processo dinamico di valutazione e revisione continue e – se necessario – di modifica, da cui è scaturita la necessità di aggiornare la checklist e l'articolo di spiegazione ed elaborazione. Considerato che emergono continuamente nuove evidenze e commenti critici, valuteremo la necessità di futuri aggiornamenti.

Dal 1996 la prima versione del CONSORT Statement sembra aver determinato un miglioramento della qualità del reporting di RCT nelle riviste che lo hanno adottato⁵⁰⁻⁵⁴. Altri gruppi stanno utilizzando il modello del CONSORT per migliorare il reporting di altri disegni di studio, come i test diagnostici³¹¹ e gli studi osservazionali³¹².

Il sito web del CONSORT (www.consort-statement.org), creato per fornire materiale didattico e un database di informazioni rilevanti per il reporting di RCT, contiene molti esempi di trial pubblicati tra cui tutti quelli riportati in questo articolo. Continueremo ad aggiungere esempi validi e non validi di reporting al database, e invitiamo i lettori a inviare ulteriori suggerimenti, contattandoci attraverso il sito web. Il gruppo CONSORT continuerà a effettuare una revisione della letteratura per reperire articoli rilevanti su aspetti fondamentali del reporting di RCT: gli autori di eventuali articoli sono invitati a segnalarceli. Tutte queste informazioni saranno disponibili sul sito web del CONSORT, che viene regolarmente aggiornato.

Più di 400 riviste mediche generali e specialistiche e gruppi editoriali biomedici – inclusi l'*International Committee of Medical Journals Editors* (ICMJE), la *World Association of Medical Journal Editors* (WAME) e il *Council of Science Editors* (CSE) – hanno ufficialmente supportato il CONSORT. Invitiamo altre riviste attente alla qualità dei reporting dei trial ad approvare il documento CONSORT e a contattarci attraverso il nostro sito web per notificarci il loro supporto. I beneficiari finali di questi sforzi collettivi saranno tutte le persone che, per qualsiasi motivo, necessitano di un intervento da parte della comunità sanitaria.

Ringraziamenti

Ringraziamo Frank Davidoff e Tom Lang per la loro partecipazione al CONSORT 2001 Explanation and Elaboration. Un ringraziamento speciale a Maria Ocampo, coordinatore del CONSORT di Ottawa, che ha contribuito a diffondere il CONSORT 2001 Explanation and Elaboration e il CONSORT Statement.

Membri del CONSORT group che hanno contribuito al 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.

Al fine di incoraggiare la diffusione del CONSORT Statement 2010, questo articolo, liberamente accessibile sul sito www.bmj.com, sarà pubblicato anche sul *Journal of Clinical Epidemiology*. Gli autori detengono congiuntamente i diritti d'autore di questo articolo. Per ulteriori dettagli sull'utilizzo, consultare il sito web del CONSORT (www.consort-statement.org).

MATERIALE SUPPLEMENTARE

Diagramma di flusso.
Checklist.

NOTE ALLA VERSIONE ITALIANA

La Fondazione GIMBE ha sostenuto la traduzione italiana dell'articolo senza alcun supporto istituzionale o commerciale.

TEAM CHE HA REALIZZATO LA VERSIONE ITALIANA

Responsabile scientifico

Dott. Antonino Cartabellotta

Coordinamento editoriale

Dott. Marco Mosti

Traduzione

Dott.ssa Luigia Atorino, Dott.ssa Elena Cottafava

Revisione editoriale

Dott. Roberto Luceri, Dott.ssa Francesca Torre

BIBLIOGRAFIA

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. 7th Cochrane Colloquium, Rome, Italy 1999.
5. Jüni 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, Jüni 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, Gøtzsche PC, Kunz R, Oxman AD, Chalmers I. So what's so special about randomisation. In: Maynard A, Chalmers I, eds. *Non-random reflections on health services research*. BMJ Books, 1997:93-106.
 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, Doré 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, McPeck 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, Gøtzsche 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, Jüni 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, Hróbjartsson A, Haahr MT, Gøtzsche 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 D, for 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. Gøtzsche 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-e776.
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, Sooriyachchi 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, Hróbjartsson A, Forfang E, Altman DG, Gøtzsche PC. Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study. *BMJ* 2005;330:1049.
92. Chan AW, Hróbjartsson A, Jørgensen KJ, Gøtzsche 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, Macé 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. Gøtzsche 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-analysis. Prepared for the Therapeutics Initiative, 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, Kroleza-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. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. In: Egger M, Davey Smith G, Altman DG, eds. *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 cooperative 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, Gülmezoglu 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, Hróbjartsson A, Jørgensen KJ, Hilden J, Altman DG, Gøtzsche 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, Jüni 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. Gøtzsche 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, Hróbjartsson 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, Rössner 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, eds. *Statistics with confidence: confidence intervals and statistical guidelines*. 2nd ed. *BMJ Books*, 2000:171-90.
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, eds. *Encyclopedia of biostatistics*. John Wiley, 1998:1000-5.
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, Mancica 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, Stepniowska 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 non-compliance. *Stat Med* 2000;19:1849-64.
220. Melander H, Ahlqvist-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, Martínez-González MA, de Irala-Estévez 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, eds. *Statistics with confidence*. 2nd ed. BMJ Books, 2000:6-14.
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, eds. *Statistics with confidence*. 2nd ed. BMJ Books, 2000:120-38.

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, Nexøe 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, Schünemann 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, Sirén 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, Lacey 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, eds. *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. Bornhöft 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. Gøtzsche PC, Gjørup I, Bonnén 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. Gøtzsche 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. Tramèr 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, Hébert 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, Gøtzsche PC, Hróbjartsson 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, Als-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. Als-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, Gøtzsche 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, van't Veer M, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. *N Engl J Med* 2009;360:213-24.