Adequate reporting of randomized, controlled trials (RCTs) is necessary to allow accurate critical appraisal of the validity and applicability of the results. The CONSORT (Consolidated Standards of Reporting Trials) Statement, a 22-item checklist and flow diagram, is intended to address this problem by improving the reporting of RCTs. However, some specific issues that apply to trials of nonpharmacologic treatments (for example, surgery, technical interventions, devices, rehabilitation, psychotherapy, and behavioral intervention) are not specifically addressed in the CONSORT Statement. Furthermore, considerable evidence suggests that the reporting of nonpharmacologic trials still needs improvement. Therefore, the CONSORT group developed an extension of the CONSORT Statement for trials assessing nonpharmacologic treatments. A consensus meeting of 33 experts was organized in Paris, France, in February 2006, to develop an extension of the CONSORT Statement for trials of nonpharmacologic treatments. The participants extended 11 items from the CONSORT Statement, added 1 item, and developed a modified flow diagram.

To allow adequate understanding and implementation of the CONSORT extension, the CONSORT group developed this explanation and explanation document from a review of the literature to provide examples of adequate reporting. This extension, in conjunction with the main CONSORT Statement and other CONSORT extensions, should help to improve the reporting of RCTs performed in this field.

The CONSORT (Consolidated Standards of Reporting Trials) Statement, published in 1996 and revised in 2001, is a set of guidelines designed to improve the reporting of randomized, controlled trials (RCTs) (1, 2). Use of this evidence-based guideline is associated with improved quality of reporting in RCTs (3). The CONSORT Statement has been extended to cover different designs, such as noninferiority and equivalence trials (4); types of interventions, such as herbal therapies (5); and data, such as the reporting of harms (6). However, despite the wide dissemination of the CONSORT Statement, inadequate reporting remains common.

Nonpharmacologic treatments include surgery, technical procedures, devices, rehabilitation, psychotherapy, behavioral interventions, and complementary and alternative medicine. Of all RCTs published in 2000, RCTs of nonpharmacologic therapies account for 1 in 4 publications (7). However, the CONSORT Statement does not address some specific issues that apply to nonpharmacologic trials (8–12). For example, blinding is more difficult to achieve in nonpharmacologic trials (13) and, when feasible, relies on complex methods and specific design (14). Nonpharmacologic trials usually test complex interventions involving several components. Such treatments are consequently difficult to describe, standardize, reproduce, and administer consistently to all patients. All of these variations could have an important impact on the estimate of the treatment effect. In addition, care providers’ expertise and centers’ volume of care can also influence the estimate of the treatment effect (15).

Consequently, the CONSORT Group decided to develop an extension of the CONSORT Statement for nonpharmacologic treatments. The methods and processes leading up to these reporting guidelines are described in detail in an accompanying paper available online only at www.annals.org (16). A major element of the process was a meeting of 33 individuals in February 2006, at which consensus was achieved on guidance for reporting RCTs of nonpharmacologic treatments; this guidance consists of extensions to 11 checklist items, addition of 1 item, and modification of and the flow diagram (Table 1 and Figure 1).

To facilitate better understanding and dissemination of this CONSORT extension, the meeting participants recommended developing an explanation and elaboration document, similar to those developed for the revised CONSORT Statement (2), STARD (Standards for Reporting of Diagnostic Accuracy) (17), and STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) (18). As with those documents, this article uses a standard template: The modified checklist item is reported, along with the rationale, evidence base (whenever possible), and examples of good reporting provided in Table 2 (19–38). An example of reporting in the modified flow diagram is provided in Figure 2 (39). This document deals with only some of the CONSORT checklist items; it should thus be seen as an addendum to the main CONSORT explanatory paper (2) for trials of nonpharmacologic treatments. In this document, we have focused only on regular RCTs in which individual participants are


### Table 1. Checklist of Items for Reporting Trials of Nonpharmacologic Treatments*

<table>
<thead>
<tr>
<th>Section</th>
<th>Item</th>
<th>Standard CONSORT Description</th>
<th>Extension for Nonpharmacologic Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td>1</td>
<td>How participants were allocated to interventions (e.g., “random allocation,” “randomized,” or “randomly assigned”))</td>
<td>In the abstract, description of the experimental treatment, comparator, care providers, centers, and blinding status</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>2</td>
<td>Scientific background and explanation of rationale</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>3</td>
<td>Eligibility criteria for participants and the settings and locations where the data were collected</td>
<td>When applicable, eligibility criteria for centers and those performing the interventions</td>
</tr>
<tr>
<td>Interventions</td>
<td>4</td>
<td>Precise details of the interventions intended for each group and how and when they were actually administered</td>
<td>When applicable, details of whether or not those administering co-interventions were blinded to group assignment</td>
</tr>
<tr>
<td></td>
<td>4A</td>
<td>Description of the different components of the interventions and, when applicable, descriptions of the procedure for tailoring the interventions to individual participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4B</td>
<td>Details of how the interventions were standardized</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4C</td>
<td>Details of how adherence of care providers with the protocol was assessed or enhanced</td>
<td></td>
</tr>
<tr>
<td><strong>Objectives</strong></td>
<td>5</td>
<td>Specific objectives and hypotheses</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td>6</td>
<td>Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors)</td>
<td>When applicable, details of whether and how the clustering by care providers or centers was addressed</td>
</tr>
<tr>
<td>Sample size</td>
<td>7</td>
<td>How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules</td>
<td>When applicable, details of whether and how the clustering by care providers or centers was addressed</td>
</tr>
<tr>
<td>Randomization-sequence generation</td>
<td>8</td>
<td>Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification)</td>
<td>When applicable, how care providers were allocated to each trial group</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>9</td>
<td>Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned</td>
<td></td>
</tr>
<tr>
<td><strong>Implementation</strong></td>
<td>10</td>
<td>Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups</td>
<td>When applicable, details of whether and how the clustering by care providers or centers was addressed</td>
</tr>
<tr>
<td>Blinding (masking)</td>
<td>11A</td>
<td>Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment</td>
<td>Whether or not those administering co-interventions were blinded to group assignment</td>
</tr>
<tr>
<td></td>
<td>11B†</td>
<td>If blinded, method of blinding and description of the similarity of interventionists</td>
<td>When applicable, details of whether and how the clustering by care providers or centers was addressed</td>
</tr>
<tr>
<td><strong>Statistical methods</strong></td>
<td>12</td>
<td>Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses</td>
<td>When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td>13</td>
<td>Flow of participants through each stage (a diagram is strongly recommended)—specifically, for each group, report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome; describe protocol deviations from study as planned, together with reasons</td>
<td>The number of care providers or centers performing the intervention in each group and the number of patients treated by each care provider or in each center</td>
</tr>
<tr>
<td>Participants flow</td>
<td>14</td>
<td>Dates defining the periods of recruitment and follow-up</td>
<td></td>
</tr>
<tr>
<td>Numbers analyzed</td>
<td>15</td>
<td>Baseline demographic and clinical characteristics of each group</td>
<td>When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group</td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>16</td>
<td>Number of participants (denominator) in each group included in each analysis and whether analysis was by “intention-to-treat”: state the results in absolute numbers when feasible (e.g., 10/20, not 50%)</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes and estimation</strong></td>
<td>17</td>
<td>For each primary and secondary outcome, a summary of results for each group and the estimated effect size and its precision (e.g., 95% confidence interval)</td>
<td></td>
</tr>
<tr>
<td>Ancillary analyses</td>
<td>18</td>
<td>Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those prespecified and those exploratory</td>
<td></td>
</tr>
<tr>
<td>Adverse events</td>
<td>19</td>
<td>All important adverse events or side effects in each intervention group</td>
<td></td>
</tr>
<tr>
<td><strong>Interpretation</strong></td>
<td>20</td>
<td>Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes</td>
<td>In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centers in each group</td>
</tr>
<tr>
<td>Generalizability</td>
<td>21</td>
<td>Generalizability (external validity) of the trial findings</td>
<td>Generalizability (external validity) of the trial findings according to the intervention, comparators, patients, and care providers and centers involved in the trial</td>
</tr>
<tr>
<td><strong>Overall evidence</strong></td>
<td>22</td>
<td>General interpretation of the results in the context of current evidence</td>
<td></td>
</tr>
</tbody>
</table>

* Additions or modifications to the CONSORT checklist. CONSORT = Consolidated Standards of Reporting Trials.
† This item anticipates a planned revision in the next version of the standard CONSORT checklist.
randomly assigned to groups. Nonpharmacologic treatments can also be evaluated in cluster RCTs, and in these cases, the CONSORT extension for cluster trials should also be consulted (40).

RECOMMENDATIONS FOR REPORTING TRIALS OF NONPHARMACOLOGIC TREATMENTS

Title and Abstract

Item 1

Standard CONSORT item: How participants were allocated to interventions (for example, “random allocation,” “randomized,” or “randomly assigned”).

In addition, for nonpharmacologic trials: In the abstract, description of the experimental treatment, comparator, care providers, centers, and blinding status.

The quality of reporting titles and abstracts is essential because it helps indexers, such as those compiling the National Library of Medicine’s MEDLINE database, to classify reports so they can be correctly identified electronically (41). Furthermore, abstracts are much more likely to be read than any other section of an article. Good evidence indicates that abstracts frequently underreport key features of trials assessing pharmacologic and nonpharmacologic treatments (41–49). The CONSORT guidelines for reporting journal and conference abstracts are forthcoming (50–52).

For nonpharmacologic trials, the abstract should also include data on centers or care providers, including, when applicable, details on the number of care providers participating in the trial and their expertise. The experimental and control procedures should also be clearly identified. Finally, authors should indicate who was blinded and, if blinding of participants and care providers was impossible, whether the outcome assessment was blinded. These details are necessary to allow an adequate appraisal of the internal and external validity of the trial.

Methods Section

Item 3: Participants

Standard CONSORT item: Eligibility criteria for participants and the settings and locations where the data were collected.

In addition, for nonpharmacologic trials: When applicable, eligibility criteria for centers and those performing the interventions.

Evidence suggests that patient outcome can be associated with hospital and care providers’ volume (15, 53–57). A systematic review of 135 trials (15) found that 71% observed a positive association between hospital volume and outcomes and 70% observed an association between care providers’ volume and outcomes. Differential expertise of care providers in each treatment group can bias treatment effect estimates (58). Furthermore, a nonpharmacologic treatment might be found to be safe and effective in an RCT performed in high-volume centers by high-volume care providers, but could have different results in low-volume centers. For example, the Asymptomatic Carotid Atherosclerosis Study investigators excluded 40% of all possible care providers, selecting only those with good safety records. This resulted in a postoperative mortality rate that was 8 times lower than in other trials with less stringent selection criteria (59–61). In most nonpharmacologic trials, care providers’ expertise and centers’ volume...
of care will influence the treatment effect (15, 53–57, 62–72).

Reporting of eligibility criteria for care providers and centers in nonpharmacologic trials is often poor. One study of surgical reports found that the setting and the center’s volume of activity was reported in only 7% and 3% of articles, respectively (79). Selection criteria were reported for care providers in 41% of the articles, and the number of care providers performing the intervention was reported in 32% (79).

A careful description of care providers involved in the trial, as well as details of the centers in which participants were treated, helps readers appraise the risk for bias and the applicability of the results. Selection criteria for centers typically relates to center volume for the procedure under investigation or similar procedures. Eligibility of care providers might include professional qualifications, years in practice, number of interventions performed, skill as assessed by level of complication when performing the intervention, and specific training before trial initiation. Eligibility criteria should be justified, because they will influence the applicability of the trial results (58, 73, 74).

Item 4: Interventions

Standard CONSORT item: Precise details of the interventions intended for each group and how and when they were actually administered.

In addition, for nonpharmacologic trials: Precise details of both the experimental treatment and comparator.

Item 4A. Description of the different components of the interventions and, when applicable, description of the procedure for tailoring the interventions to individual participants.

It is important to provide a detailed description of nonpharmacologic treatments, which are usually complex interventions involving several components (75), each of which may influence the estimated treatment effect (27–32). For example, arterial endarterectomy and reconstruction during carotid endarterectomy can be performed in a variety of ways, some aspects of which may influence the treatment effect. For example, use of local anesthesia and patch closure, compared with other techniques, has been shown to reduce the risk for harms after carotid endarterectomy (76, 77). Therefore, authors should report all the different components of the treatment procedure. These descriptions will help introduce the safest and most effective treatments into clinical practice. They are also necessary to facilitate study comparison, reproducibility, and inclusion in systematic reviews (78).

In nonpharmacologic trials, the control treatment can be placebo, usual care, an active treatment, or a waiting list. If the control treatment is usual care, authors should report all the components received by the control group. This information will allow readers to compare the intensity of usual care with the experimental intervention and with what is usually provided to participants in their own setting.

Interventions in nonpharmacologic trials are often poorly described. A systematic review of reports of RCTs assessing surgical procedures highlighted the lack of reporting of other important components: only 35% of studies reported anesthesia management, 15% reported preoperative care, and 49% reported postoperative care (79). In a review of behavioral medicine interventions, insufficient intervention detail was a barrier to assessment of evidence and development of guidelines (80–82). A systematic review of articles published in 6 medical rehabilitation journals in 1997 to 1998 found that information about the timing of the intervention relative to the onset of the disorder was absent from 32% of the 171 reports. Descriptions of the interventions were either brief or absent in one half of the articles and lacked an operational definition in 9% of the articles (83).

The information that is required for a complete description of nonpharmacologic treatments depends on the type of intervention being tested. For surgery, technical procedure, or implantable devices, full details of preoperative care, intraoperative care, configuration of any device, and postoperative care are needed. For nonimplantable devices, the configuration of the device should be detailed and a user’s guide for the device should be prepared to enable reproducibility.

For rehabilitation, behavioral treatment, education, and psychotherapy, authors should report qualitative and quantitative data. Qualitative data describe the content of each session, how it is delivered (individual or group), whether the treatment is supervised, the content of the information exchanged with participants, and the instruments used to give information. Quantitative data describe the number of sessions, timing of each session, duration of each session, duration of each main component of each session, and overall duration of the intervention. It is also essential to report how the intervention was tailored to patients’ comorbid conditions, tolerance, and clinical course.

To aid the provision of a clear description of these complex interventions, Perera and colleagues proposed a graphical depiction of the experimental and control interventions (84).

Item 4B. Details of how the interventions were standardized.

Assessment of nonpharmacologic treatments in RCTs presents special difficulties because of the complexity of the treatment and the variability found across care providers and centers (23). The variety of settings that characterizes multicenter trials only exacerbates these problems (78). Authors should describe any method used to standardize the
Table 2. Checklist of Items for Reporting Trials of Nonpharmacologic Treatments, with Examples*

Section
Title and abstract

Item 1
Standard CONSORT Description
How participants were allocated to interventions (e.g., “random allocation,” “randomized,” or “randomly assigned”)

Extension for Nonpharmacologic Trials
In the abstract, description of the experimental treatment, comparator, care providers, centers, and blinding status

Examples of Good Reporting Based on Extension (Reference)
Objective: To compare the primary therapist model (PTM), provided by a single rheumatology-trained primary therapist, with the traditional treatment model (TTM), provided by a physical therapy (PT) and/or occupational therapy (OT) generalist, for treating patients with rheumatoid arthritis (RA).

Methods: Eligible patients were adults requiring rehabilitation treatment who had not received PT/OT in the past 2 years. Participants were randomized to the PTM or TTM group. The primary outcome was defined as the proportion of clinical responders who experienced a > or = 20% improvement in 2 of the following measures from baseline to 6 months: Health Assessment Questionnaire, pain visual analog scale, and Arthritis Community Research and Evaluation Unit RA Knowledge Questionnaire (19).

Participants

Item 3
Standard CONSORT Description
Eligibility criteria for participants and the settings and locations where the data were collected

Extension for Nonpharmacologic Trials
When applicable, eligibility criteria for centers and those performing the interventions

Examples of Good Reporting Based on Extension (Reference)
All participating centres . . . were major neurosurgical centres, treating large numbers of patients after aneurysmal subarachnoid haemorrhage (SAH), each centre treating between 60 and 200 cases annually. . . . Centres had to have expertise in both neurosurgical and endovascular management of ruptured aneurysms. Only accredited neurosurgeons with experience of aneurysm surgery were permitted to manage patients in the trial. Endovascular operators had to have done a minimum of 30 aneurysm treatment procedures, before they were permitted to treat patients in the trial (20).

Interventions

Item 4
Standard CONSORT Description
Precise details of the interventions intended for each group, and how and when they were actually administered

Extension for Nonpharmacologic Trials
Precise details of both the experimental treatment and comparator

Examples of Good Reporting Based on Extension (Reference)
The exercise training program . . . consisted of 2 approximately 3-month-long phases of exercise training. The initial phase of exercise was designed to prepare the participants for progressive resistance training and also to minimize injury. Exercises during the first 3-month phase (phase 1) were conducted by a physical therapist using a group format (2–5 participants/group) and were designed to enhance flexibility, balance, coordination, movement speed, and, to some extent, strength of all major muscle groups. Twenty-two exercises formed the basis of this program (protocol available from the authors). The exercises were made progressively more difficult by increasing the number of repetitions and/or by performing the exercises in more challenging ways. When safely able, participants also exercised on a stationary bicycle or treadmill. Participants attempted this exercise for a minimum of 5 minutes and progressed to a maximum of 15 minutes. The treadmill speed or bicycle resistance was set at the highest comfortable setting that was safe for the partici- pant’s ability and tolerance, which increased over the course of phase 1. During the second exercise phase (phase 2), progressive resistance training was added. One-repetition maximum (1-RM) voluntary strength was measured on each of 6 different exercises (knee extension, knee flexion, seated bench press, seated row, leg press, and biceps curl), which were performed bilaterally on a Hoist weightlifting machine (Hoist Fitness Systems, San Diego, Calif). Initially, the participants performed 1 to 2 sets of 6 to 8 repetitions of each exercise at 65% of their 1-RM. By the end of the first month of weight training, they progressed to 3 sets of 8 to 12 repetitions performed at 85% to 100% of the initial 1-RM. The 1-RM measurements were repeated at 6 weeks and used to progressively increase each individual’s exercise prescription. Participants continued to perform a shortened version of the phase 1 exercises and the treadmill or stationary bicycle warm-up exercise . . . (21).

The [control] treatment follows the same format [as experimental treatment], i.e., 10 weekly 90-min sessions. The therapist helps the patient identify daily stresses and discusses them in a supportive non-directive mode. No instructions for exposure are included. If the patient brings up trauma-related issues, the therapist gently redirects her to discuss other material (22).

Item 4A
Extension for Nonpharmacologic Trials
Description of the different components of the interventions and, when applicable, descriptions of the procedure for tailoring the interventions to individual participants

Examples of Good Reporting Based on Extension (Reference)
The usual practices of surgeons performing optic nerve decompression surgery were determined through literature review and through a survey of study surgeons. These practices were described in the protocol as a series of 31 steps, only six of which were required to be performed so as to ensure adequacy of the surgery as well as safety of the patient. The remaining steps could be performed according to surgeon preference as they did not directly affect either patient safety or adequacy of surgery. Each study surgeon signed a written commitment to adhere to the six required steps, which were: general anesthesia, medial approach, no mechanical static traction, subarachnoid dissection if no cerebrospinal fluid release was seen following fenestration of the optic nerve sheath, no more than 7 minutes of sustained traction on the globe at any one time and rest periods of at least 2 minutes following any 7-minute period of globe traction (23).
Table 2—Continued

Section

Item 4C

Extension for Nonpharmacologic Trials

Details of how adherence of care providers with the protocol was assessed or enhanced

Examples of Good Reporting Based on Extension (Reference)

All therapy sessions are videotaped. . . . A senior clinician who is independent of . . . treatment delivery will rate 10% of the videotapes using measures adapted from several randomized clinical trials of psychotherapy . . .; the 10% figure was chosen arbitrarily in an attempt to ensure an adequate sample of information from each treatment condition (22).

Sample size

Item 7

Standard CONSORT Description

How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules

Extension for Nonpharmacologic Trials

When applicable, details of whether and how the clustering by care providers or centers was addressed

Examples of Good Reporting Based on Extension (Reference)

The study was designed to enroll 8 participants for each of 4 therapists at the 12 participating sites. . . . Sample size was computed on the basis of our primary hypothesis, that PE [prolonged exposure] will be more effective than PCT [present-centered therapy] for the treatment of PTSD [posttraumatic stress disorder] due to military-related trauma in women as measured by the CAPS [clinician-administered PTSD scale] at 3 months posttreatment. Although treatment is delivered on an individual basis, each participant cannot be assumed to generate independent observations because participants are clustered within therapists. Thus, the computed sample size, based on the unpaired t-test statistic, was inflated by a factor, \( t = 1 + (m - 1)p \), to achieve the variance that one would have anticipated had there been no clustering. The cluster size \((m)\) is 8 (participants/therapist), and the intraclass correlation coefficient \((\rho)\) was estimated from prior studies to be in the range of .10 to .15, which in turn yields a sample size inflation factor of 1.7 to 2.05. With an estimated sample size of 384, this study has 85% to 90% statistical power to detect an effect of \( \delta = .50 \) at \( \alpha = .05 \), two-tailed . . . (22).

Randomization-sequence generation

Item 8

Standard CONSORT Description

Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification)

Extension for Nonpharmacologic Trials

When applicable, how care providers were allocated to each trial group

Examples of Good Reporting Based on Extension (Reference)

At each of the 12 sites, 4 female therapists were randomly assigned to deliver either PE [prolonged exposure] or PCT [present-centered therapy] \((n = 2\) per condition per site). . . . By design, each therapist treats 10 participants: 2 training cases during a 6-month run-up period, and 8 randomized cases during 2 years of recruitment (22).

Pancreatectoduodenectomy [in both experimental and control group] was performed by just 3 experienced surgeons who had done more than 40 pancreati-coduodenectomies with either the conventional [control treatment] or the binding pancreatectojunostomy [experimental treatment] (24).

The patients were randomly selected for one of two operative procedures: open reduction and internal fixation or external fixation and limited internal fixation. The six attending orthopaedic surgeons who performed the operations had been assigned to a treatment group according to their expertise or to their preference with regard to fixation. Each patient was managed by the one of the six surgeons who was on call when the patient was seen in the emergency room (25).

Blinding (masking)

Item 11A

Standard CONSORT Description

Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment

Extension for Nonpharmacologic Trials

Whether or not those administering co-interventions were blinded to group assignment

Examples of Good Reporting Based on Extension (Reference)

Patients were randomised [laparoscopic versus small-incision cholecystectomy] in the operating theatre and anaesthetic technique and pain-control methods were standardised. Four experienced surgeons did both types of procedure. Identical wound dressings were applied in both groups so that carers could be kept blind to the type of operation (26).

Item 11BY

Extension for Nonpharmacologic Trials

If blinded, method of blinding and description of the similarity of interventions

Examples of Good Reporting Based on Extension (Reference)

Double blinding was achieved by shielding the subject’s view by a vertical drape and other means (described below) and by excluding the nurse assessor from the room until the procedure and clean-up were completed. . . . The subject’s contact with the investigator/procedurist was generally limited to the day of the procedure. . . . A 1-liter bag of sterile normal saline was hung at the edge of the drape within view of the subject. The knee was then draped with sterile towels, and the connecting tubing and 3-way stop-cocks were assembled and attached to an empty 1-liter waste bag and a 50-ml syringe, producing a closed system for fluid delivery, aspiration and ejection. To administer the SI [sham irrigation], the 14-gauge needle was advanced to, but not through, the joint capsule via the lateral suprapatellar port. Fresh saline was drawn from the supply bag in aliquots of 40-50 mL, and 3–5 mL of saline was expelled into the subcutaneous tissue with each mimicking “exchange” before the remainder of the saline was expelled into the waste bag. Positioning of the knee and manipulation were performed as for actual TI [tidal irrigation]. After passage of 1-liter of saline through the tubing, . . . the needle was removed. . . . All subjects and the nurse assessor remained blinded until the subject had completed study followup (27).

Statistical methods

Item 12

Standard CONSORT Description

Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses

Extension for Nonpharmacologic Trials

When applicable, details of whether and how the clustering by care providers or centers was addressed
Table 2—Continued

Section

Examples of Good Reporting Based on Extension (Reference)

Although the participants were individually randomised, a clustering of outcomes is potentially possible since a single therapist was treating several patients. If these clustering effects were strong, then this might alter the results. We therefore used multilevel modelling to check for any clustering effects by undertaking an analysis on the primary outcome (28).

Results

Participant flow

Item 13

Standard CONSORT Description

Flow of participants through each stage (a diagram is strongly recommended)—specifically, for each group, report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome; describe protocol deviations from study as planned, together with reasons

Extension for Nonpharmacologic Trials

The number of care providers or centers performing the intervention in each group and the number of patients treated by each care provider or in each center

Examples of Good Reporting Based on Extension (Reference)

See Figure 2, reconstructed with the authors’ permission (39).

Implementation of interventions

Item: New item

Extension for Nonpharmacologic Trials

Details of the experimental treatment and comparator as they were implemented

Examples of Good Reporting Based on Extension (Reference)

A single stent was implanted in 546 patients (40%), 2 stents in 487 (36%), 3 stents in 206 (15%), and 4 or more stents in 111 (8%) in both study groups (mean, 1.9 stents per patient and 1.4 stents per lesion). The mean stent diameter was 2.8 mm in both groups, and the mean length was 22.8 mm in the sirolimus-eluting stent group and 23.5 mm in the paclitaxel-eluting stent group. The maximum dilatation pressure during stent implantation was significantly lower in the paclitaxel-eluting stent group than in the sirolimus-eluting stent group (29).

On average, participants attended a mean of 9.4 exercise sessions (SD, 3.2) and 10.2 sham exercise sessions (SD, 3.0) of the planned 12 sessions. Participants attended a mean of 2.9 advice sessions (SD, 1.1) and 2.5 sham advice sessions (SD, 1.1) of the planned 3 sessions. The mean duration of exercise sessions was 54.0 minutes (SD, 16.3), of which 35.6 minutes (SD, 12.6) were spent with a physiotherapist. The mean duration of sham exercise sessions was 47.0 minutes (SD, 25.0), of which 22.9 minutes (SD, 8.4) were spent with a physiotherapist. Mean durations of advice and sham advice sessions were 20.0 minutes (SD, 4.9) and 19.0 minutes (SD, 5.3), respectively (30).

Baseline data

Item 15

Standard CONSORT Description

Baseline demographic and clinical characteristics of each group

Extension for Nonpharmacologic Trials

When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group

Examples of Good Reporting Based on Extension (Reference)

We dichotomized surgeon’s experience in laparoscopic repair into greater than 250 repairs (experienced) and less than 250 repairs (inexperienced). Surgeons participating in this trial ranged in age from 27 to 70 with a median of 42 years in the laparoscopic group (55 surgeons) and from 30 to 76 with a median of 42 in the open group (77 surgeons). In the laparoscopic group, 8 surgeons were classified as experienced and 47 as inexperienced (31, 32).

Discussion

Interpretation

Item 20

Standard CONSORT Description

Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes

Extension for Nonpharmacologic Trials

In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centers in each group

Examples of Good Reporting Based on Extension (Reference)

First, surgeons might not be proficient in one or both treatments. The difference in malunion rates between the two treatment groups was consistent across all four study sites, indicating the difference is due to the procedure and not technical proficiency. Staff from all four centres were experienced in both techniques and, therefore, the results are probably typical of other paediatric centers (35).

Generalizability

Item 21

Standard CONSORT Description

Generalizability (external validity) of the trial findings

Continued on following page
Assessing treatment adherence is essential to appraising the feasibility and reproducibility of the intervention in clinical practice. Several methods have been used to assess treatment adherence, such as review of case report forms, videotapes, and audiotapes (23, 73, 82, 85, 86). Authors should report the use of any adherence-improving strategies, such as decertifying and excluding surgeons who did not submit a videotape of the intervention rated as acceptable by an independent committee (23). Such strategies should enhance treatment adherence and may influence the treatment effect. Readers must be aware of these methods and strategies in order to accurately transpose the results of the trial into clinical practice and appraise the applicability of the trial’s results (82).

In a sample of 158 reports of surgical RCTs published in 2004 (79), only 4 articles reported care providers’ compliance with the planned procedure.

**Item 7: Sample Size**

Standard CONSORT item: How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.

In addition, for nonpharmacologic trials: When applicable, details of whether and how the clustering by care providers or centers was addressed.

Table 3 shows the rationale for the clustering effect (87–92). As with cluster randomized trials (40, 93), sample size estimates for individually randomized RCTs assessing nonpharmacologic treatments should ideally be adjusted for the clustering effect as estimated by the intraclass intraclass correlation coefficient. Authors should report whether and how they have incorporated these issues into the trial sample size calculations.
Item 8: Randomization (Sequence Generation)

Standard CONSORT item: Methods used to generate the random allocation sequence, including details of any restriction (for example, blocking, stratification).

In addition, for nonpharmacologic trials: When applicable, how care providers were allocated to each trial group.

In conventional RCTs, especially pharmacologic trials, participants are randomly assigned to 1 of 2 (or more) treatments. The treatments compared are usually administered by the same care providers. That approach is not desirable in many nonpharmacologic trials. First, the expertise of care providers for each procedure may differ, or one procedure may be more challenging than the other. This issue may result in differential expertise between interventions and may bias the treatment effect estimates, especially in surgery. Second, care providers are frequently unblinded in nonpharmacologic trials, and they can have preferences or differential expectations for one of the interventions. Thus, they may unconsciously bias the trial: for example, when prescribing a co-intervention or when proposing crossover between groups (that is, the experimental treatment is offered to participants randomly assigned to the control group). A survey of 139 surgeons participating in a large conventional RCT comparing 2 surgical procedures for treating a tibial shaft fracture showed that statistically significantly more surgeons had no or limited experience with the more technically challenging procedure (58). Furthermore, 87% of surgeons believed that the less-challenging procedure was superior, and differential crossover occurred: 8% of the patients assigned to the more-challenging procedure received the less-challenging procedure, whereas fewer than 1% of patients assigned to the less-challenging procedure received the more-challenging procedure.

To overcome these problems, care providers participating in a trial might perform interventions only in their preferred treatment group (expertise-based RCT) (58). This design is mandatory when comparing 2 different types of interventions, such as surgery versus physiotherapy for back problems. However, that design might limit the applicability of the trial results. In trials assessing behavioral intervention, rehabilitation, and psychotherapy, some researchers proposed selecting a random sample of care providers to avoid biased results and improve the applicability.

Consequently, so that others can evaluate the internal and external validity of a trial, authors should report on how care providers were allocated to each treatment group.

Item 11: Blinding (Masking)

11A. Standard CONSORT item: Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment.

This example was not reported in the article but was developed with the help of the authors (39). IQR = interquartile range; max = maximum; min = minimum.

In addition, for nonpharmacologic trials: Whether or not those administering co-interventions were blinded to group assignment.

Empirical evidence demonstrates that lack of reporting of blinding is associated with biased estimates of treatment effect (94–98). Blinding in trials is usually considered in relation to the participants, caregivers, and outcome assessors (14). In nonpharmacologic trials, the blinding status of other caregivers (for example, physicians administering co-interventions) should also be reported. In fact, other caregivers may have an important influence on the observed treatment effect. For example, in a trial assessing a surgical procedure, even if the surgeon cannot be blinded,
the health care professionals following the participants after the procedure might be blinded, and contact between other caregivers and the surgeon could be avoided, thus limiting the risk for performance bias.

Item 11B. If blinded, method of blinding and description of the similarity of interventions.

At the most recent CONSORT Group meeting (Montebello, Quebec, Canada, January 2007), the participants agreed to revise the 2001 CONSORT Statement. Item 11 of the checklist deals with reporting of blinding, and the wording of the item will be modified (Moher D. Personal communication.). Because blinding is an especially important issue for nonpharmacologic trials, we have used the wording of the revised checklist item on blinding for the nonpharmacologic extension. Part of the 2001 version states: “If done, how the success of blinding was evaluated.” This is now replaced by: “If blinded, method of blinding and description of the similarity of interventions.”

Blinding is often more difficult to carry out in trials assessing nonpharmacologic treatments (13), and the risk for unblinding is important (99, 100). A review of the methods of blinding in nonpharmacologic trials highlighted creative methods of blinding reported by some authors. Examples include use of sham procedures, such as simulation of surgical procedures, or partial blinding of participants, in which participants are blinded to the study hypothesis (14). The methods of blinding, as well as the similarity of treatments, should be highlighted.

Researchers are still working on how best to deal with some of these methodological challenges. In the meantime, authors should report how they have handled them in order to allow progress in understanding these potential biases.

Item 12: Statistical Methods

Compare CONSORT item: Statistical methods used to compare groups for primary outcome(s). Methods for additional analyses, such as subgroup analyses and adjusted analyses.

In addition, for nonpharmacologic trials: When applicable, details of whether and how the clustering by care providers or centers was addressed.

Table 3 shows the rationale for centers’ volume. In trials assessing nonpharmacologic treatments, the success of the intervention depends in part on the skill and training of care providers. As such, observations from participants treated by the same care provider or in the same center are correlated or clustered.

Standard methods of analysis that ignore clusters will result in incorrect estimation of treatment effect (87, 89, 101–104). Authors should use specific models that allow adjustment for participant characteristics while controlling for clustering effect (90, 91) in analyzing the results of this type of trial (91). Any allowance that was made in the analysis for the clustering of participants and care providers or care providers and center should be reported.

Results Section

Item 13: Participant Flow

Compare CONSORT item: Flow of participants through each stage (a diagram is strongly recommended)—specifically, for each group, report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome; describe protocol deviations from study as planned, together with reasons.

In addition, for nonpharmacologic trials: The number of care providers or centers performing the intervention in each group, and the number of patients treated by each care provider or in each center.

As outlined in the CONSORT Statement, the flow of individual participants through each stage of the trial should be reported: the number of persons evaluated for potential enrollment, randomly assigned to each group, who received treatment as allocated, who completed treatment as allocated, who completed follow-up as planned, and included in the main analyses in each group (2).

For trials assessing nonpharmacologic treatments, authors also should report information on the number of centers and care providers in each group and the distribution of participants treated by care providers or at each center. This information is crucial to allow others to criti-
cally appraise the applicability of the trial’s results. For instance, if 50 surgeons are treating participants, it is important to know whether most patients are being treated by only 1 surgeon or whether all surgeons treated similar numbers of patients. Authors should report the median (interquartile range, minimum and maximum value) of participants treated by each care provider or center. This information could be reported in a figure (Figure 1).

Implementation of Interventions

New item, for nonpharmacologic trials: Details of the experimental treatment and comparator as they were implemented.

Although a nonpharmacologic intervention can be standardized (see item 4B), there may be differences between how it was intended to be administered and how it actually was administered—for example, because of lack of the reproducibility of the treatment (82). Furthermore, because participants and care providers are frequently not blinded to treatment assignment, a risk for unequal administration of additional treatments (co-intervention) and consequent “contamination” (that is, administration of the experimental treatment to the control group) might influence the estimates of treatment effect. Care providers may introduce part or all of the experimental intervention into the control program if they are convinced of its superiority. Participants in the control group may also treat themselves with the experimental intervention if they believe in its efficacy. Reporting how the intervention was actually administered is thus crucial to accurate evaluation of the results (22, 23).

Item 15: Baseline Data

Standard CONSORT item: Baseline demographic and clinical characteristics of each group.

In addition, for nonpharmacologic trials: When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group.

Although the eligibility criteria (item 3) provide some information on care providers and centers participating in an RCT, further details of the characteristics of the care providers and the centers that recruited and treated participants are important to know.

A table can efficiently present this baseline information. The mean and SD can be used to summarize quantitative data for each group. When quantitative data have an asymmetrical distribution, a preferable approach may be to give the median and percentile range (perhaps the 25th and 75th percentiles). Authors should report numbers and proportions for categorical and qualitative data (2). These data are essential to appraise the risk for bias linked to care providers’ expertise and the external validity of the results.

Discussion Section

Item 20: Interpretation

Standard CONSORT item: Interpretation of the results, taking into account the study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes.

In addition, for nonpharmacologic trials: In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centers in each group.

Three aspects specific to nonpharmacologic trials should be addressed in the Discussion section. First, the choice of the comparator is important and will influence the observed treatment effect of the intervention (75, 105). In particular, debate surrounds the use of placebo interventions in trials assessing nonpharmacologic treatments, because these treatments may have a specific therapeutic effect associated with the relationship between participants and care providers. Consequently, trials with placebos might underestimate the treatment effect (14, 106). Some placebos are also questionable from an ethical perspective, such as the use of simulated or sham surgery (27, 37, 107).

Second, blinding issues associated with the feasibility of blinding, risk for blinding failure (13, 14, 108), and risk for bias when blinding is not feasible should be discussed (109, 110). When participants and care providers are not blinded, performance bias (that is, unequal provision of care according to the treatment administered) could occur; a discussion of co-interventions, contamination, and the rate of follow-up in each group is therefore useful. Lack of blinding of outcome assessors could be responsible for ascertainment bias. Any methods used to reduce bias should be discussed. For situations in which outcome assessors cannot be blinded, an objective primary outcome, such as mortality or assessment by an independent end point committee, could limit the risk for bias.

Finally, authors should discuss the possibility of differential expertise bias (58) linked to unequal expertise of care providers in each group (item 3).

Item 21: Generalizability

Standard CONSORT item: Generalizability (external validity) of the trial findings.

In addition, for nonpharmacologic trials: Generalizability (external validity) of the trial findings according to the intervention, comparators, patients, care providers, and centers involved in the trial.

To be clinically useful, the results of RCTs should provide data on the external validity, also called generalizability and applicability. Lack of external validity is frequently cited as a reason why interventions found to be
effective in clinical trials are underutilized in clinical practice (111). In trials comparing pharmacologic with nonpharmacologic treatment, the characteristics of the patients included, the trial setting, the treatment regimens, and the outcomes assessed should all be reported and discussed (2, 111). In nonpharmacologic trials, the health care system (112), selection of participating centers and care providers (59), and intervention actually administered are also essential to evaluate the external validity. For example, differences between health care systems affected the external validity in the European Carotid Surgery Trial (112), an RCT of endarterectomy for recently symptomatic carotid stenosis. In this international trial, countries differed in the speed with which patients were investigated. These differences were not mentioned in any of the publications for the European Carotid Surgery Trial, yet they probably had an important impact on the outcomes (111). Similarly, differences between countries in methods of diagnosis and management can affect the external validity of the trial results. Finally, the volume of centers and care providers can influence the treatment effect estimates, and exclusive participation of high-volume centers has obvious implications for external validity. Authors should clearly indicate whether the intervention evaluated could be performed in all settings by all centers or should be reserved for high-volume centers.

**DISCUSSION**

We developed this CONSORT extension to help improve the reporting of RCTs investigating nonpharmacologic treatments. This document provides explanation of and elaboration on the CONSORT checklist items specific to nonpharmacologic treatments. Authors should use this document in conjunction with the main CONSORT guidelines (2) when addressing all 22 items on the checklist. Depending on the type of trial conducted, authors may also find it useful to consult the CONSORT extensions for cluster trials (40) and noninferiority trials (4), and the detailed guidelines for reporting harms associated with interventions (6). All CONSORT guidelines can be found on the CONSORT Web site (www.consort-statement.org).

We hope that journals endorsing and enforcing CONSORT for reporting nonpharmacologic RCTs will recommend that authors also review this explanatory document. We believe that the promotion of this extension will improve the quality of reporting RCTs of nonpharmacologic treatments.

From INSERM U738, AP-HP, Hôpital Bichat-Claude Bernard, and Université Paris 7 Denis Diderot, Paris, France; Chalmers Research Group, Children’s Hospital of Eastern Ontario Research Unit, University of Ottawa, Ottawa, Ontario, Canada; University of Oxford, Oxford, United Kingdom; and Family Health International, Research Triangle Park, North Carolina.

**Acknowledgment:** The authors thank Lola Fourcade, who was responsible for preparing the minutes of the CONSORT meeting, and Lucie Ronceray, who developed the Web-based interface used for the preliminary survey. They also thank Drs. Homs and Siersma (Departments of Gastroenterology and Hepatology, Erasmus MC/University Medical Centre Rotterdam, Rotterdam, the Netherlands), who helped reconstruct Figure 2.

**Grant Support:** By the Département de la Recherche Clinique et du Développement, Assistance Publique des Hôpitaux de Paris; Département d’Épidémiologie, Biostatistique et Recherche Clinique, Hôpital Bichat-Claude Bernard; INSERM U738; and the Eli Lily Institute. Dr. Moher is funded in part by a University of Ottawa Research Chair. Dr. Altman is supported by Cancer Research UK. Dr. Schulz is supported by Family Health International.

**Potential Financial Conflicts of Interest:** None disclosed.

**Requests for Single Reprints:** Isabelle Boutron, MD, PhD, Département d’Épidémiologie Biostatistique et Recherche Clinique, INSERM U738, AP-HP, Hôpital Bichat-Claude Bernard, Université Paris 7 Denis Diderot, 46 Rue Henri Huchard, 75018 Paris, France; e-mail, isabelle.boutron@bch.aphp.fr.

Current author addresses are available at www.annals.org.

**References**

13. Boutron I, Tubach F, Giraudreau B, Ravaud P. Blinding was judged more...


106. Paterson C, Dieppe P. Characteristic and incidental (placebo) effects in complex interventions such as acupuncture. BMJ. 2005;330:1202-5. [PMID: 15905259]


111. Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?”. Lancet. 2005;365:82-93. [PMID: 15639683]


---

**ANNALS AUDIO SUMMARIES**

Author interviews and audio summaries of each *Annals* issue are now available at www.annals.org. For transcripts, or for instructions on how to download to a portable MP3 player, go to www.annals.org/podcast. You may also subscribe directly through iTunes.

After listening, send feedback to podcast@acponline.org.
Current Author Addresses: Drs. Boutron and Ravaud: Département d’Épidémiologie Biostatistique et Recherche Clinique, INSERM U738, AP-HP, Hôpital Bichat-Claude Bernard, Université Paris 7 Denis Diderot, 46 Rue Henri Huchard, 75018 Paris, France. Dr. Moher: Chalmers Research Group, Children’s Hospital of Eastern Ontario Research Institute, Faculty of Medicine, University of Ottawa, 401 Smyth Road, Room 210, Ottawa, Ontario, Canada. Dr. Altman: Centre for Statistics in Medicine, University of Oxford, Wolfson College Annexe, Linton Road, Oxford OX2 6UD, United Kingdom. Dr. Schulz: Quantitative Sciences, Family Health International, PO Box 13950, Research Triangle Park, NC 27709.

Appendix: Contributors to the CONSORT Extension for Nonpharmacologic Treatment Interventions

Douglas G. Altman (University of Oxford, Oxford, United Kingdom); Mohit Bhandari (McMaster University, Hamilton General Hospital, McMaster Clinic, Hamilton, Ontario, Canada); Isabelle Boutron (INSERM U738, AP-HP, Hôpital Bichat-Claude Bernard, and Université Paris 7 Denis Diderot, Paris, France); Marion Campbell (Health Services Research Unit, Aberdeen, United Kingdom); Philip Devereaux (McMaster University Health Sciences Centre, Hamilton, Ontario, Canada); Peter C. Gøtzsche (The Nordic Cochrane Centre, Copenhagen, Denmark); Teodor P. Grantcharov (Saint Michael’s Hospital, Toronto, Ontario, Canada); Jeremy Grimshaw (Ottawa Health Research Institute, Ottawa, Ontario, Canada); Ethan A. Halm (Mount Sinai School of Medicine, New York, New York); Erik Hendriks (Maastricht University, Maastricht, the Netherlands); Ashbjorn Hróbjartsson (The Nordic Cochrane Centre, Copenhagen, Denmark); John Ioannidis (University of Ioannina School of Medicine and Biomedical Research Institute, Foundation for Research and Technology–Hellas, Ioannina, Greece); Astrid James (The Lancet, London, United Kingdom); Giselle Jones (BMJ Editorial, London, United Kingdom); Richard J. Lilford (University of Birmingham, Edgbaston, Birmingham, United Kingdom); Robin McLeod (Annals of Surgery and Mount Sinai Hospital, Toronto, Ontario, Canada); David Moher (Chalmers Research Group, Children’s Hospital of Eastern Ontario Research Institute, and University of Ottawa, Ottawa, Ontario, Canada); Andrew J. Molyneux (Neurovascular Research Unit, Radcliffe Infirmary, Oxford, United Kingdom); Victor M. Montori (Mayo Clinic College of Medicine, Rochester, Minnesota); Cynthia Mulrow (Annals of Internal Medicine, American College of Physicians, Philadelphia, Pennsylvania); Amy Plint (Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada); Philippe Ravaud (INSERM U738, AP-HP, Hôpital Bichat-Claude Bernard, and Université Paris 7 Denis Diderot, Paris, France); Drummond Rennie (JAMA, Chicago, Illinois, and University of California, San Francisco, San Francisco, California); Peter M. Rothwell (Radcliffe Infirmary, Oxford, United Kingdom); Paula P. Schnurr (Veterans Affairs Medical and Regional Office Center, White River Junction, Vermont); Kenneth F. Schulz (Quantitative Sciences, Family Health International, Research Triangle Park, North Carolina); Christoph M. Seiler (University of Heidelberg Medical School, Heidelberg, Germany); Judith Stephen­son (Centre for Sexual Health and HIV Research, Research Unit, London, United Kingdom); Simon G. Thompson (Medical Research Council, Biostatistics Unit, Institute of Public Health, Cambridge, United Kingdom); Graham Thornicroft (Institute of Psychiatry, King’s College London, London, United Kingdom); David Torgerson (University of York, York, United Kingdom); Tom Treasure (Guy’s and St Thomas’ Hospital, London, United Kingdom); Peter Tugwell (Journal of Clinical Epidemiology and University of Ottawa, Ottawa, Ontario, Canada).