

REVIEW ARTICLE

THE REPORTING OF BLINDING IN PHYSICAL MEDICINE AND REHABILITATION RANDOMIZED CONTROLLED TRIALS: A SYSTEMATIC REVIEW

Mauricio F. Villamar, MD^{1,2}, Vanessa Suárez Contreras, MD¹, Richard E. Kuntz, MD, MSc³ and Felipe Fregni, MD, PhD, MPH¹

From the ¹Laboratory of Neuromodulation, Department of Physical Medicine & Rehabilitation, Spaulding Rehabilitation Hospital and Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA, ²School of Medicine, Pontifical Catholic University of Ecuador, Quito, Ecuador and ³Medtronic, Inc., Minneapolis, MN, USA

Objective: To conduct a systematic review evaluating the reporting of blinding in randomized controlled trials published in the field of Physical Medicine and Rehabilitation over two time periods.

Data sources: We searched MEDLINE via PubMed for all randomized controlled trials published in *American Journal of Physical Medicine and Rehabilitation*, *Archives of Physical Medicine and Rehabilitation*, *Clinical Rehabilitation*, *Disability and Rehabilitation* and *(Scandinavian) Journal of Rehabilitation Medicine* in the years 2000 and 2010.

Study selection: We initially identified 222 articles, and 139 (62.6%) met our selection criteria.

Data extraction: Two independent investigators collected data regarding study characteristics and blinding from each article. Consistency of data extraction was evaluated.

Data synthesis: When comparing articles from 2010 and 2000, the former showed significantly higher rates for reporting of blinding, explicitly describing key persons' blinding status, and discussing the absence of blinding as a study limitation. There was a trend for lower reporting among trials with positive outcomes. No improvement was observed in other CONSORT-enforced parameters.

Conclusions: Although the reporting of blinding in Physical Medicine and Rehabilitation randomized controlled trials shows some improvement over the past decade, it still does not fulfill current recommendations. Given its critical role in determining internal validity, stricter enforcement of CONSORT guidelines is needed.

Key words: blinding; clinical research; physical medicine; rehabilitation; randomized controlled trials as topic.

J Rehabil Med 2013; 45: 6–13

Correspondence address: Felipe Fregni, MD, PhD, MPH, Laboratory of Neuromodulation, Spaulding Rehabilitation Hospital, 125 Nashua Street #726, Boston, MA, USA 02114. E-mail: Fregni.Felipe@mgh.harvard.edu

Submitted March 30, 2012; accepted August 16, 2012

The paper was presented as an abstract in the III International Symposium on Neuromodulation, 17–19 October 2011 in Sao Paulo, Brazil.

INTRODUCTION

The term “blinding”, also referred to as “masking”, denotes the concealment of information about the assigned interventions or the true hypothesis of the study from key persons involved in a clinical trial, such as participants, healthcare providers, data collectors, outcome assessors and data analysts (1). Given that knowledge of a participant's treatment group can modify key persons' behavior and perceptions, thereby affecting outcomes, the main goal of blinding is to limit the possibility of introducing different sources of bias in a study (2–4). Effect estimates tend to be exaggerated when there is a lack of blinding (5–7), and this is particularly true when outcome measures involve some degree of subjectivity. In fact, a meta-epidemiological study by Wood et al. (8) covering 746 randomized controlled trials (RCTs) found that open-label trials tend to overestimate intervention effects by an average of 7% as compared with blind studies. Among those assessing subjective outcomes, treatment effects were found to be exaggerated by 25%.

Clinical trials in Physical Medicine and Rehabilitation (PM&R) pose some particular difficulties as compared to those conducted in other fields of medicine, since successful blinding is often hard to achieve (9, 10). Finding reliable “placebo” alternatives to physical therapy, exercise or certain devices may be challenging and sometimes even impossible (10, 11). Indeed, trials involving nonpharmacologic interventions report less blinding of healthcare providers, patients and outcome assessors as compared to pharmacologic studies (12, 13). Consequently, detection bias associated with a lack of adequate patient blinding tends to be most prominent for nonpharmacologic interventions (7).

Given the critical role of blinding in determining the likelihood of bias in a study, its characteristics must be thoroughly described in order to allow the reader to assess the internal validity of a RCT (14). However, despite its importance, reporting of blinding is not routinely provided in clinical trials (15, 16). Moreover, when blinding is reported, this is frequently done by using ambiguous terms such as “single-”, “double-” or “triple-blind” without further specifying which key persons were kept unaware of the assigned intervention (17–20). For

instance, in the field of general medicine, the blinding status of each of such groups was found to be explicitly reported in fewer than 25% of RCTs published in high-impact journals in 2000 (18). In a review that looked at RCTs assessing surgical interventions, only 8.2% were found to adequately report the blinding status of participants, 17.1% explained the blinding status of outcome assessors and none specified the blinding status of healthcare providers (13).

Since physicians and textbooks have been shown to differ markedly in their definitions of terms such as “single-” or “double-blind” (21), the Consolidated Standards of Reporting Trials (CONSORT) Statement recommends abandoning them, and advocates for detailed reporting of blinding-related parameters. This includes defining who was blinded, the mechanism of blinding, and the similarity of characteristics of treatments. If any key trial persons are not blinded, authors should explain why this occurred (1).

To the best of our knowledge, no previous systematic reviews have been conducted in order to assess the reporting of blinding in PM&R RCTs and its changes over time. Thus, the purpose of this review was to examine this parameter among 5 prominent journals in the field (22), and to correlate it with RCT characteristics and results. Specifically, we aimed to determine: a) how frequently blinding is reported in studies, and the terms authors use to describe it; b) the factors related with a higher reporting of blinding; c) the extent to which authors explain which key persons were kept blinded; d) whether blind studies follow the recommendations of CONSORT in terms of describing the characteristics of blinding; e) if reporting rates for blinding differ between CONSORT-endorsing and non-endorsing journals; and f) whether there have been any changes in these parameters over the past decade. Our methodology follows the guidelines of the PRISMA Statement (23).

METHODS

Eligibility criteria

We reviewed all RCTs published either online or in print in *American Journal of Physical Medicine and Rehabilitation*, *Archives of Physical Medicine and Rehabilitation*, *Clinical Rehabilitation*, *Disability and Rehabilitation* and *Journal of Rehabilitation Medicine* (formerly *Scandinavian Journal of Rehabilitation Medicine*) in the years 2000 and 2010. Currently, all these journals have a 5-year impact factor greater than 2.0 according to Thomson-Reuters' Journal Citation Reports. RCTs were defined as “prospective studies assessing healthcare interventions in human participants who were randomly allocated to study groups” (19). We therefore excluded non-experimental trials, pilot studies, case reports and series, published reports of follow-up studies, secondary analyses of other trials, retrospective studies, systematic reviews and meta-analyses.

Our original goal was to analyze all RCTs published in the above-mentioned journals in the years 1990, 2000 and 2010. However, given the small number of publications from 1990 that fulfilled our selection criteria, inferential statistics were used for comparing the years 2000 and 2010 only.

Data sources and search strategy

We used the keywords shown in Fig. 1 to search MEDLINE via PubMed. We used the Medical Subject Headings (MeSH) terms “clinical trial*” or “random*” and other subheadings in order to find all the RCTs published

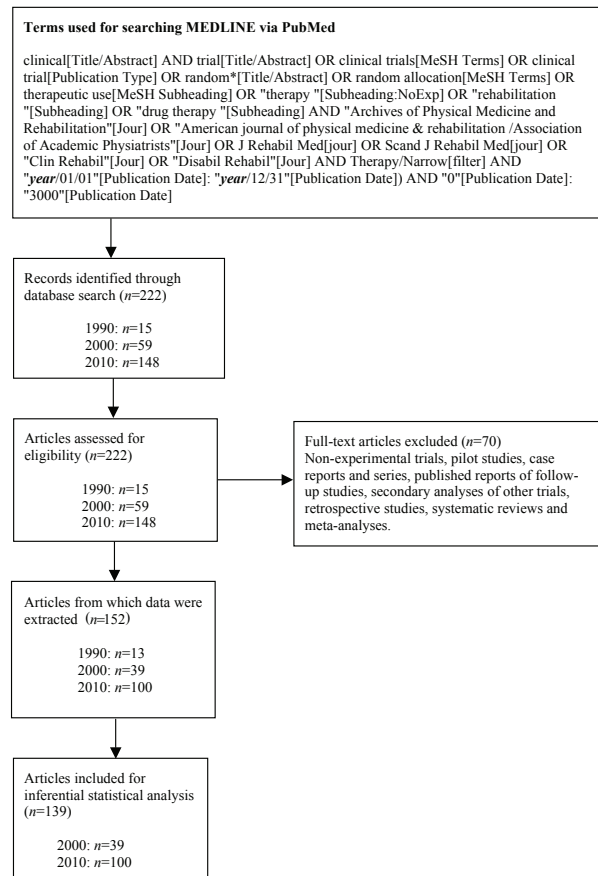


Fig 1. Selection process flow diagram.

in the aforementioned journals in 1990, 2000 and 2010. This search strategy was conducted in May 2011 and was modified from that used by Abdul Latif et al. (24) in their recent study on sample size calculation.

Study selection

A total of 222 articles from these 3 time periods were found after the initial search. We retrieved full reports of all of them for detailed assessment by two reviewers, who scrutinized their titles, abstracts and methods sections in order to identify intervention studies. Articles were then selected based on our eligibility criteria. After this process, 139 were included for inferential statistical analysis, as shown in Fig. 1.

Data extraction

A preformatted database was used for data extraction. This database was generated based on a review of the literature and was previously discussed and tested by the researchers. From each article, information on variables of interest was independently extracted and entered into the database by two of the authors.

General characteristics of the studies were recorded. Articles were first examined to assess whether blinding was either mentioned or used. Next, we evaluated whether the authors explicitly reported which key persons were kept blinded, as opposed to the use of terms such as “single-” or “double-blind”. Additionally, the similarity of characteristics of treatments, the steps taken to maintain blinding and a description of the timing of unblinding were also analyzed. If the success of blinding was tested, we recorded the methods used by the authors. If blinding was not used, we assessed if an explanation was provided as to why this occurred, and whether lack of blinding was discussed as a study limitation. Table I lists the variables extracted from each article.

After examining the “Instructions for authors” section in each journal’s website, we found that, as of October 2011, *Archives of Physical Medicine and Rehabilitation*, *Clinical Rehabilitation* and *Journal of Rehabilitation Medicine* request that all RCTs submitted to them follow CONSORT guidelines. *American Journal of Physical Medicine and Rehabilitation* and *Disability and Rehabilitation* do not endorse them.

Agreement between reviewers

Each of the first two authors independently reviewed 50% of the 222 articles originally retrieved. During this process, 70 articles were not considered eligible for the aforementioned reasons, and information was extracted from the remaining 152 that met our eligibility criteria, including papers from 1990. Then, in order to evaluate the uniformity of data collection, the first author randomly examined 25% of the article reviewed by the second author, and vice versa. Any discrepancies between the two reviewers were resolved through consensus after re-checking the original manuscript and, if necessary, consulting the corresponding author. For each of the variables analyzed, interobserver agreement was not lower than 0.8, as measured by Cohen’s kappa coefficient.

Statistical analyses

Descriptive statistics were used for reporting characteristics of trials from which information was extracted. Given that all the independent variables analyzed were categorical, two-sided Fisher’s exact tests were conducted in order to compare differences between studies published in 2000 and 2010, and between trials that reported versus those that did not report blinding. For all analyses, two-tailed *p*-values < 0.05 were considered statistically significant.

Because we were limited by the number of articles published in these journals over these two time periods, we conducted a post-hoc power analysis and found that a sample of 139 articles (39 from 2000 and 100 from 2010) has a power of 92.5% to detect a 30.0% difference in reporting rates at an alpha value of 0.05. Statistical analyses were performed using STATA 10 (StataCorp. 2007. *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP).

RESULTS

Description of study characteristics

We extracted information from a total of 152 full reports of RCTs, including 13 articles from 1990, 39 from 2000 and 100

from 2010. A complete list of all the studies that were reviewed is shown in AppendixSI (available from <http://www.medical-journals.se/jrm/content/?doi=10.2340/16501977-1071>). Table II describes their general characteristics. In these studies, a total of 7,960 participants were randomly assigned to different intervention groups. The median sample size was 37.5 participants per trial. Notably, the number of RCTs published

Table II. Characteristics of articles included for data extraction

Characteristics	Year		
	1990 (n=13)	2000 (n=39)	2010 (n=100)
Sample size, median (IQR)	24 (25)	32 (51)	40 (36.5)
Journal, n (%)			
American Journal of Physical Medicine & Rehabilitation	2 (15.4)	4 (10.3)	15 (15.0)
Archives of Physical Medicine and Rehabilitation	9 (69.2)	22 (56.4)	23 (23.0)
Clinical Rehabilitation	0 (0.0)	10 (25.6)	41 (41.0)
Disability and Rehabilitation	0 (0.0)	3 (7.7)	10 (10.0)
(Scandinavian) Journal of Rehabilitation Medicine	2 (15.4)	0 (0.0)	11 (11.0)
Origin, n (%)			
USA/Canada	8 (61.5)	9 (23.1)	11 (11.0)
Europe	5 (38.5)	20 (51.3)	50 (50.0)
Asia	0 (0.0)	6 (15.4)	28 (28.0)
Oceania	0 (0.0)	4 (10.2)	6 (6.0)
Latin America	0 (0.0)	0 (0.0)	4 (4.0)
Africa	0 (0.0)	0 (0.0)	1 (1.0)
Rehabilitation area, n (%)			
Pain/musculoskeletal	4 (30.7)	16 (41.0)	38 (38.0)
Neurorehabilitation	3 (23.1)	10 (25.7)	32 (32.0)
Spinal	3 (23.1)	2 (5.1)	7 (7.0)
Geriatric	0 (0.0)	3 (7.7)	6 (6.0)
Cardiac	1 (7.7)	2 (5.1)	5 (5.0)
Pulmonary	1 (7.7)	3 (7.7)	3 (3.0)
Other	1 (7.7)	3 (7.7)	9 (9.0)
Type of intervention, n (%)			
Pharmacologic	1 (7.7)	4 (10.3)	8 (8.0)
Non-pharmacologic	12 (92.3)	35 (89.7)	92 (92.0)
Physical therapy	5 (41.7)	17 (48.6)	53 (57.6)
Devices	6 (50.0)	9 (25.7)	26 (28.3)
Other	1 (8.3)	9 (25.7)	13 (14.1)
Type of control, n (%)			
Active intervention	6 (46.1)	24 (61.6)	60 (60.0)
No intervention	3 (23.1)	10 (25.6)	24 (24.0)
Placebo/Sham	4 (30.8)	5 (12.8)	16 (16.0)
Study sites, n (%)			
Single center	11 (84.6)	31 (79.5)	92 (92.0)
Multicenter	2 (15.4)	8 (20.5)	8 (8.0)
Outcome, n (%)			
Primary	1 (7.7)	11 (28.2)	28 (28.0)
Multiple	12 (92.3)	28 (71.8)	72 (72.0)
Result, n (%)			
Positive	8 (61.5)	30 (76.9)	74 (74.0)
Negative	5 (38.5)	9 (23.1)	26 (26.0)
Any reporting of blinding in the study, n (%)			
No	4 (30.8)	17 (43.6)	15 (15.0)
Yes	9 (69.2)	22 (56.4)	85 (85.0)
Specific reporting of key persons who were blinded, n (%)			
Absent	6 (46.2)	17 (43.6)	21 (21.0)
Present	7 (53.8)	22 (56.4)	79 (79.0)

IQR: Interquartile range.

Table I. Variables extracted from each article

Journal
CONSORT Endorser vs. Non-endorser
Year
Origin
Number of study sites
Sample size
Rehabilitation area
Type of intervention
Type of control
Outcome (primary vs. multiple)
Result (positive vs. negative)
Any reporting of blinding
Terms used by authors to describe blinding
Specific reporting of key trial persons who were blinded
Actions taken to ensure similarity of characteristics of interventions
Steps taken to maintain blinding
Description of the timing of unblinding
Assessment of blinding success
In case that blinding was not used
Do authors explain the reason?
Is it discussed as a study limitation?

Table III. Differences between trials reporting and not reporting blinding (years 2000 and 2010)

Parameters	Blinding not reported n=32 n (%)	Blinding reported n=107 n (%)	P
Year			
2000	17 (43.6)	22 (56.4)	0.001
2010	15 (15.0)	85 (85.0)	
Type of intervention			
Pharmacologic	1 (9.1)	10 (90.9)	0.46
Non-pharmacologic	31 (24.2)	97 (75.8)	
Study sites			
Single center	27 (22.0)	96 (78.0)	0.53
Multicenter	5 (31.3)	11 (68.7)	
Type of control			
Active intervention	24 (28.6)	60 (71.4)	0.01
No intervention	8 (23.5)	26 (76.5)	
Placebo/sham	0 (0.0)	21 (100.0)	
Sample size			
≤50	23 (25.6)	67 (74.4)	0.40
>50	9 (18.4)	40 (81.6)	
Outcome			
Primary	10 (25.6)	29 (74.4)	0.66
Multiple	22 (22.0)	78 (78.0)	
Result			
Positive	28 (26.9)	76 (73.1)	0.07
Negative	4 (11.4)	31 (88.6)	

in the 5 journals almost tripled every 10 years, with the highest number coming from *Archives of Physical Medicine and Rehabilitation* (n=54, 35.5%) and *Clinical Rehabilitation* (n=51, 33.6%) over the 3 time periods. The number of trials from the United States and Canada has remained relatively stable over the years, while those from Europe and Asia have increased markedly, accounting for most of the publications from 2010.

In all time periods, the main area of research was pain/musculoskeletal, followed by neurorehabilitation. However,

in recent years there has been a growing interest in areas such as spinal, geriatric, and urinary rehabilitation. Approximately 90% of the studies performed in the field of PM&R comprise nonpharmacologic interventions, a percentage that has remained constant over the years. Among them, those involving physical therapy are most common. Most published trials used active interventions as controls, and only a minority used placebo/sham-based controls. Single-center studies, those assessing multiple outcomes and those with positive results were more frequently published in all time periods.

Reporting of blinding

Articles were first reviewed in order to determine whether blinding was reported, either as part of the study design or by stating its absence, as opposed to failure to allude to it. The reporting rate for this parameter was found to be significantly higher among papers from 2010 than among those from 2000 (85.0% vs. 56.4%, $p=0.001$), as depicted in Table III. In terms of trial design, significantly higher reporting rates were seen in placebo/sham-controlled trials as compared to those using active interventions or no intervention as controls (100.0% vs. 71.4% and 76.5%, respectively; $p=0.01$). When comparing by main trial outcome (positive vs. negative results), data suggest a trend for lower reporting of blinding among the former ($p=0.07$). Factors such as the type of intervention ($p=0.46$), number of study sites ($p=0.53$), sample size ($p=0.40$) and number of outcome variables ($p=0.66$) were not significantly related to the reporting of blinding.

We found that a significantly higher percentage of articles from 2010 explicitly describe which key trial persons were kept blinded as compared to those from 2000 (79.0% vs. 56.4%, $p=0.01$). Among studies described as "single-blind" by their authors, outcome assessors (41.9%) were the key persons most frequently blinded. However, 5 papers (16.0%) listed more than one group as being blinded, and one study (3.2%) did not

Table IV. Reporting of blinding status of key trial persons in studies reporting blinding (years 2000 and 2010)

Key trial persons blinded	Blinding according to authors n=97		
	"Single-blind" n=31 (32.0 %) n (%)	"Double-blind" n=19 (19.6 %) n (%)	Not specified n=47 (48.4 %) n (%)
None/not specified	1 (3.2)	0 (0.0)	6 (12.8)
Participants	1 (3.2)	1 (5.3)	1 (2.1)
Data collectors	11 (35.6)	1 (5.3)	12 (25.5)
Outcome assessors	13 (41.9)	1 (5.3)	14 (29.8)
Data analysts	0 (0.0)	0 (0.0)	1 (2.1)
Participants, data collectors	0 (0.0)	3 (15.8)	1 (2.1)
Participants, outcome assessors	1 (3.2)	3 (15.8)	4 (8.5)
Participants, data analysts	0 (0.0)	1 (5.3)	0 (0.0)
Data collectors, outcome assessors	0 (0.0)	0 (0.0)	1 (2.1)
Data collectors, healthcare providers	1 (3.2)	1 (5.3)	3 (6.4)
Data collectors, data analysts	1 (3.2)	0 (0.0)	0 (0.0)
Outcome assessors, healthcare providers	1 (3.2)	0 (0.0)	0 (0.0)
Outcome assessors, data analysts	1 (3.2)	0 (0.0)	0 (0.0)
Participants, data collectors, healthcare providers	0 (0.0)	3 (15.8)	1 (2.1)
Participants, outcome assessors, healthcare providers	0 (0.0)	4 (21.1)	3 (6.4)
Participants, outcome assessors, healthcare providers, data analysts	0 (0.0)	1 (5.3)	0 (0.0)

provide any clarification as of who was blinded. When RCTs described as “double-blind” by their authors were assessed, we found that 3 of them (15.9%) specifically reported the blinding status of one key person only, while 8 (42.2%) listed 4 unique combinations of two groups who were kept unaware of the assigned intervention. Three or more key persons were blinded in 8 “double-blind” studies (42.2%). Table IV lists the main findings.

Table V compares several blinding-related characteristics for articles published in 2000 and 2010. Papers from 2010 where blinding was not used tend to discuss its absence as a study limitation more frequently than those from 2000 ($p=0.004$). However, no significant differences were found in terms of describing other parameters, such as the actions taken to ensure the similarity of characteristics of interventions ($p=0.81$), the steps taken to maintain blinding ($p=0.41$), a description of the timing of unblinding ($p=0.45$) or a justification in case that blinding was not used ($p=0.14$).

We assessed whether there were any disparities in the reporting of blinding between CONSORT-endorsing and non-endorsing journals. No significant differences were found in relation to any of the parameters listed in the previous paragraph, either when articles published in 2000 or in 2010 were analyzed (data not shown). Lastly, we found that none of the RCTs published in 2000 or 2010 reported having assessed blinding success. In fact, of 152 articles from which information was extracted, only one, a 1990 study specifically focused on the feasibility of blinding in the field of PM&R, tested for the success of blinding.

DISCUSSION

Over the past decade, some significant improvement has been demonstrated in terms of reporting and discussing certain

blinding-related characteristics in RCTs in the field of PM&R. Articles from 2010 showed significantly higher rates for reporting the presence or absence of blinding, specifically reporting which groups were kept blinded and discussing the absence of blinding as a study limitation, as compared to those from 2000. However, despite CONSORT recommendations, reporting of blinding did not fully improve as evidenced by deficiency in the following areas: a) the actions taken to ensure the similarity of characteristics of interventions, b) the steps taken to maintain blinding, c) a description of the time of unblinding and d) a justification in case that blinding was not used. Altogether, though some improvement in the reporting of blinding in PM&R RCTs has been observed, it remains suboptimal.

Although blinding leads to clear methodological advantages, its absence does not necessarily invalidate the results of a given study. In fact, some reviews have found no association between presence or absence of “double-blinding” and effect sizes in placebo-controlled trials (25), though trial characteristics such as type of outcomes can influence the likelihood that lack of blinding would cause biased results. Similarly, a trial does not need to be “double-blind” to be judged as the best quality, since many trials where blinding of one key person only is achievable may be equally or more rigorous in terms of their methodology (26). However, in order to allow the reader to adequately appraise the internal validity of a RCT, the characteristics of blinding must be thoroughly described (14).

Previous studies have suggested that implementation of CONSORT has resulted in more detailed reporting over time in certain areas of RCTs (27, 28). Among them, reporting of blinding is one of the parameters that have most consistently demonstrated a greater improvement, as described in RCTs in palliative care (29) and in cluster randomized trials (30). In line with these results, we found that 85.0% of PM&R trials published in 2010 mention the presence or absence of blinding, as compared to 56.4% in 2000. Given its methodological importance, these findings are certainly auspicious. However, when we compare these numbers with what has been found in other areas of medicine, it becomes clear that reporting in the field of PM&R was considerably below average a decade ago and that it is still deficient. Meinert et al. (31) showed that 86.0% of trials published in MEDLINE in 1980, encompassing a wide variety of medical specialties, reported the presence or absence of blinding. Likewise, Chan & Altman (19) found this percentage to be of 92.0% in papers from 2000. Therefore, despite the fact that reporting of blinding in PM&R trials has shown notable progress, further efforts are required in order to equal the rates seen in other fields.

The low rate for reporting of blinding among PM&R trials may be due in part to the difficulty of blinding key persons in this area and/or because of a lack of familiarity with potential alternatives. The latter may be favored by the fact that no significant differences related to reporting of blinding were seen in our study between RCTs assessing pharmacologic and nonpharmacologic interventions.

Although blinding of participants and personnel who administer the interventions may not always be feasible in PM&R

Table V. Description of blinding-related characteristics among articles from 2000 and 2010

Parameters	Year		p
	2000 n (%)	2010 n (%)	
Specific reporting of key trial persons who were blinded			
Absent	17 (43.6)	21 (21.0)	0.01
Present	22 (56.4)	79 (79.0)	
Actions taken to ensure similarity of characteristics of interventions			
Not specified	13 (59.1)	44 (55.7)	0.81
Specified	9 (40.9)	35 (44.3)	
Steps taken to maintain blinding			
Not specified	15 (68.2)	61 (77.2)	0.41
Specified	7 (31.8)	18 (22.8)	
Description of the timing of unblinding			
Not specified	21 (95.5)	69 (87.3)	0.45
Specified	1 (4.5)	10 (12.7)	
Justification in case that blinding was not used			
Absent	18 (100.0)	23 (82.1)	0.14
Present	0 (0.0)	5 (17.9)	
Discussion of absence of blinding as a study limitation			
Absent	18 (100.0)	18 (64.3)	0.004
Present	0 (0.0)	10 (35.7)	

RCTs, it should always be possible to blind data collectors and outcome assessors, as well as data analysts. This is particularly important due to the subjective nature of many outcome measurements in this field. In addition to participants, those administering interventions and outcome assessors, trials assessing nonpharmacologic treatments should report whether or not those administering co-interventions were blind to group assignment (32). In this way, even if the interventionist cannot be blinded to the procedure or therapy, blinding key persons responsible for the follow-up of participants may decrease the likelihood of performance bias (14).

We considered it helpful to provide a few examples of proper reporting of blinding found among the studies included in our review. A comprehensive description of the blinding status of study personnel involved in a trial (with blinding of participants described elsewhere in the same manuscript), would be the following: “A second individual, a physical therapist with extensive experience using the study measures, was not involved in treating the patient but performed the assessment measures... Both the treating and assessing therapists were blinded to the treatment of group assignment. A study coordinator applied the stimulator electrodes and turned on and off the stimulation or sham stimulation, as assigned. The coordinator was aware of the group assignment.” (33).

Similarly, a detailed description of the actions taken to ensure the similarity of characteristics of the interventions would be, for example: “Identical nonfunctioning units were provided by the manufacturer. Both real and sham units had ‘on’ lights that flashed at the stimulation frequency set by the control knobs, and a battery indicator light that flashed when batteries were low.” (11) However, given that providing an exhaustive description of methods for facilitating blinding in the setting of PM&R trials is beyond the scope of this review, the reader is prompted to refer to Boutron et al. (34) for thorough recommendations focused on different types of outcomes as well as on participative and device-based interventions, two of the most common types of interventions studied in PM&R trials. Additionally, Lowe et al. (35) describe a successful implementation of blind outcome assessment in a physiotherapy RCT evaluating recovery after knee arthroplasty.

Even though a significantly higher proportion of articles from 2010 do provide information specifying which key persons were kept blinded, the use of terms such as “single-” or “double-blind” is still common in these publications despite CONSORT recommendation to abandon them. Given their confusing nature, these terminologies were used for describing the blinding status of several possible groups or their combinations in our sample. For this reason, it becomes evident that they are hindering effective and objective research communication, complicating the evaluation of trial validity, the interpretation of their findings and their application into clinical practice. Therefore, description of the blinding status of each group taking part in a clinical trial might be the optimal method for reporting blinding characteristics.

Our results suggest that trials with positive results may have lower reporting rates for blinding, although this finding was

not statistically significant. The fact that unblinded trials produce positive results more often as compared to blind studies has long been recognized in the field of PM&R (36), and may be explained by the influence of observer bias on outcome measurements. Therefore, results from open-label trials should be viewed very carefully while acknowledging this potential limitation, though the extent of potential bias may depend on factors such as nature of outcomes. Some progress has been made in this regard over the years. While no open-label trials from 2000 discussed the absence of blinding as a study limitation and its potential implication for study results, 35.7% of those from 2010 acknowledged it. Nevertheless, if blinding were not used in a trial, it would also be desirable that authors provide an explanation of why this occurred, a characteristic that has shown no improvement over the past decade.

Due to limitations in the interpretation of tests for the success of blinding, CONSORT 2010 no longer encourages performing them (37). However, we decided to assess this parameter since 3 of the journals included in our search strategy still require authors to discuss it, and because trials published in 2010 were either designed or carried out when the previous guidelines were in effect. The fact that only one study from 1990 fulfilled this requisite is certainly worrisome and represents yet another call for a more detailed reporting of blinding in this field.

We did not find any significant differences related to reporting of any blinding-related parameters between CONSORT-endorsing and non-endorsing journals in our sample. Similarly, a systematic review by Plint et al. (28) concluded that the CONSORT-adopting status of a journal appears to have little impact on reporting of blinding. This could suggest that implementation of CONSORT is not the sole responsible for the higher reporting rate for blinding that was found in our study. A tendency toward higher methodological rigor over time, determined by more careful planning and analysis by authors or by stricter submission guidelines and review processes by journals, may play an important role. Therefore, and quoting Montori et al. (18), “for the revised CONSORT statement to effectively improve the quality of reporting, it seems likely that journals adopting the CONSORT statement will have to move from endorsement to enforcement of the checklist”.

Study limitations

Our study has some limitations. First, we were not able to include articles from 1990 for inferential statistical analysis due to the reduced number of RCTs published in that year. This would have been useful to determine reporting of blinding before and after publication of the first CONSORT Statement in 1996. Nevertheless, our study was considered to be adequately powered to detect differences and assess changes between the two other time periods.

Secondly, it was not feasible for us to include all trials published in all PM&R journals over the 3 time periods. If RCTs published in other journals were systematically different with respect to their methodological rigor and their reporting of blinding, this could reduce the external generalizability of our results. However, because of the thoroughness of the editorial

review of the journals chosen for this study, it is likely that this sample would represent at least the articles with the most meticulous reporting in the field of PM&R.

Given the fact that we used a search strategy, there is also a chance that we might have missed some eligible studies classified under keywords or MeSH terms that we did not cover. Nonetheless, we used a search strategy that aimed to capture all of the possible terms for their categorization. Finally, in order to have a homogeneous sample, our study only included randomized trials. However, blinding can also be implemented in quasi-experimental and observational studies, and it would have been interesting to determine the reporting of blinding in these studies as well.

In conclusion, stricter enforcement of CONSORT guidelines is needed in order to guarantee clear and detailed reporting of blinding in PM&R RCTs, and its importance must be acknowledged by authors, journal editors and reviewers. This is a necessary step as to facilitate the interpretation and judging of trial methodology, enabling the judicious implementation of their findings into clinical practice.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers of this manuscript for their insightful feedback and comments. MF Villamar is grateful to Ana Cristina Albuja, M.D. for constant support.

Sources of financial support: Departmental funds.

Authors' financial disclosure: The authors have no proprietary or commercial interest in any materials discussed in this article.

The authors declare no conflicts of interest.

REFERENCES

- Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010; 63: e1–e37.
- Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. *Ann Intern Med* 2002; 136: 254–259.
- Hrobjartsson A, Boutron I. Blinding in randomized clinical trials: imposed impartiality. *Clin Pharmacol Ther* 2011; 90: 732–736.
- Polit DF, Gillespie BM, Griffin R. Deliberate ignorance: a systematic review of blinding in nursing clinical trials. *Nurs Res* 2011; 60: 9–16.
- 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.
- 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: 408–412.
- Nuesch E, Reichenbach S, Trelle S, Rutjes AW, Liewald K, Sterchi R, et al. The importance of allocation concealment and patient blinding in osteoarthritis trials: a meta-epidemiologic study. *Arthritis Rheum* 2009; 61: 1633–1641.
- Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008; 336: 601–605.
- Boutron I, Tubach F, Giraudeau B, Ravaud P. Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. *J Clin Epidemiol* 2004; 57: 543–550.
- Fregni F, Imamura M, Chien HF, Lew HL, Boggio P, Kaptchuk TJ, et al. Challenges and recommendations for placebo controls in randomized trials in physical and rehabilitation medicine: a report of the international placebo symposium working group. *Am J Phys Med Rehabil* 2010; 89: 160–172.
- Deyo RA, Walsh NE, Schoenfeld LS, Ramamurthy S. Can trials of physical treatments be blinded? The example of transcutaneous electrical nerve stimulation for chronic pain. *Am J Phys Med Rehabil* 1990; 69: 6–10.
- Boutron I, Tubach F, Giraudeau B, Ravaud P. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *JAMA* 2003; 290: 1062–1070.
- Jacquier I, Boutron I, Moher D, Roy C, Ravaud P. The reporting of randomized clinical trials using a surgical intervention is in need of immediate improvement: a systematic review. *Ann Surg* 2006; 244: 677–683.
- 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.
- Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989; 10: 31–56.
- 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.
- Devereaux PJ, Bhandari M, Montori VM, Manns BJ, Ghali WA, Guyatt GH. Double blind, you are the weakest link – good-bye! *ACP J Club* 2002; 136: A11.
- 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: 787–790.
- Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005; 365: 1159–1162.
- Haahr MT, Hrobjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clin Trials* 2006; 3: 360–365.
- 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: 2000–2003.
- Frontera WR, Grimby G, Basford J, Muller D, Ring H. Publishing in physical and rehabilitation medicine. *Am J Phys Med Rehabil* 2008; 87: 215–220.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; 151: 264–269, W64.
- Abdul Latif L, Daud Amadera JE, Pimentel D, Pimentel T, Fregni F. Sample size calculation in physical medicine and rehabilitation: a systematic review of reporting, characteristics, and results in randomized controlled trials. *Arch Phys Med Rehabil* 2011; 92: 306–315.
- Balk EM, Bonis PA, Moskowitz H, Schmid CH, Ioannidis JP, Wang C, et al. Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *JAMA* 2002; 287: 2973–2982.
- Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002; 359: 696–700.
- Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001; 285: 1992–1995.
- 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: 263–267.

29. Piggott M, McGee H, Feuer D. Has CONSORT improved the reporting of randomized controlled trials in the palliative care literature? A systematic review. *Palliat Med* 2004; 18: 32–38.
30. Ivers NM, Taljaard M, Dixon S, Bennett C, McRae A, Taleban J, et al. Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000–2008. *BMJ* 2011; 343: d5886.
31. Meinert CL, Tonascia S, Higgins K. Content of reports on clinical trials: a critical review. *Control Clin Trials* 1984; 5: 328–347.
32. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Methods and processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments. *Ann Intern Med* 2008; 148: W60–W60.
33. Stein J, Hughes R, D’Andrea S, Therrien B, Niemi J, Krebs K, et al. Stochastic resonance stimulation for upper limb rehabilitation poststroke. *Am J Phys Med Rehabil* 2010; 89: 697–705.
34. Boutron I, Guittet L, Estellat C, Moher D, Hrobjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing non-pharmacological treatments. *PLoS Med* 2007; 4: e61.
35. Lowe CJ, Wilson MS, Sackley CM, Barker KL. Blind outcome assessment: the development and use of procedures to maintain and describe blinding in a pragmatic physiotherapy rehabilitation trial. *Clin Rehabil* 2011; 25: 264–674.
36. Falconer J, Hayes KW, Chang RW. Therapeutic ultrasound in the treatment of musculoskeletal conditions. *Arthritis Care Res* 1990; 3: 85–91.
37. Schulz KF, Altman DG, Moher D, Fergusson D. CONSORT 2010 changes and testing blindness in RCTs. *Lancet* 2010; 375: 1144–1146.