Chapman & Hall/CRC Biostatistics Series

Randomized Clinical Trials of Nonpharmacological Treatments

Edited by
Isabelle Boutron
Philippe Ravaud
David Moher

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Randomized Clinical Trials of Nonpharmacological Treatments

Chapman & Hall/CRC Biostatistics Series

Chapman & Hall/CRC Biostatistics Series

Adaptive Design Theory and Implementation Using SAS and R
Mark Chang

Advanced Bayesian Methods for Medical Test Accuracy
Lyle D. Broemeling

Advances in Clinical Trial Biostatistics
Nancy L. Geller

Applied Statistical Design for the Researcher
Daryl S. Paulson

Basic Statistics and Pharmaceutical Statistical Applications, Second Edition
James E. De Muth

Bayesian Adaptive Methods for Clinical Trials
Scott M. Berry, Bradley P. Carlin, J. Jack Lee, and Peter Muller

Bayesian Analysis Made Simple: An Excel GUI for WinBUGS
Phil Woodward

Bayesian Methods for Measures of Agreement
Lyle D. Broemeling

Bayesian Missing Data Problems: EM, Data Augmentation and Noniterative Computation
Ming T. Tan, Guo-Liang Tian, and Kai Wang Ng

Bayesian Modeling in Bioinformatics
Dipak K. Dey, Samiran Ghosh, and Bani K. Mallick

Causal Analysis in Biomedicine and Epidemiology: Based on Minimal Sufficient Causation
Mikel Aickin

Clinical Trial Data Analysis using R
Ding-Geng (Din) Chen and Karl E. Peace

Clinical Trial Methodology
Karl E. Peace and Ding-Geng (Din) Chen

Computational Methods in Biomedical Research
Ravindra Khattree and Dayanand N. Naik

Computational Pharmacokinetics
Anders Källén

Controversial Statistical Issues in Clinical Trials
Shein-Chung Chow

Data and Safety Monitoring Committees in Clinical Trials
Jay Herson

Design and Analysis of Animal Studies in Pharmaceutical Development
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Bioavailability and Bioequivalence Studies, Third Edition
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Clinical Trials with Time-to-Event Endpoints
Karl E. Peace

Design and Analysis of Non-Inferiority Trials
Mark D. Rothmann, Brian L. Wiens, and Ivan S. F. Chan

Difference Equations with Public Health Applications
Lemuel A. Moyé and Asha Seth Kapadia

DNA Methylation Microarrays: Experimental Design and Statistical Analysis
Sun-Chong Wang and Arturas Petronis

DNA Microarrays and Related Genomics Techniques: Design, Analysis, and Interpretation of Experiments
David B. Allsion, Grier P. Page, T. Mark Beasley, and Jode W. Edwards

Dose Finding by the Continual Reassessment Method
Ying Kuen Cheung

Elementary Bayesian Biostatistics
Lemuel A. Moyé

Frailty Models in Survival Analysis
Andreas Wienke

Generalized Linear Models: A Bayesian Perspective
Dipak K. Dey, Sujit K. Ghosh, and Bani K. Mallick

Handbook of Regression and Modeling: Applications for the Clinical and Pharmaceutical Industries
Daryl S. Paulson

Measures of Interobserver Agreement and Reliability, Second Edition
Mohamed M. Shoukri

Medical Biostatistics, Second Edition
A. Indrayan

Meta-Analysis in Medicine and Health Policy
Dalene Stangl and Donal A. Berry

Monte Carlo Simulation for the Pharmaceutical Industry: Concepts, Algorithms, and Case Studies
Mark Chang

Multiple Testing Problems in Pharmaceutical Statistics
Alex Dmitrienko, Ajit C. Tamhane, and Frank Bretz

Randomized Clinical Trials of Nonpharmacological Treatments
Isabelle Boutron, Philippe Ravaud, and David Moher

Sample Size Calculations in Clinical Research, Second Edition
Shein-Chung Chow, Jun Shao and Hansheng Wang

Statistical Design and Analysis of Stability Studies
Shein-Chung Chow

Statistical Evaluation of Diagnostic Performance: Topics in ROC Analysis
Kelly H. Zou, Aiyi Liu, Andriy Bandos, Lucila Ohno-Machado, and Howard Rockette

Statistical Methods for Clinical Trials
Mark X. Norleans

Statistics in Drug Research: Methodologies and Recent Developments
Shein-Chung Chow and Jun Shao

Statistics in the Pharmaceutical Industry, Third Edition
Ralph Buncher and Jia-Yeong Tsay

Translational Medicine: Strategies and Statistical Methods
Dennis Cosmatos and Shein-Chung Chow

Chapman & Hall/CRC Biostatistics Series

Randomized Clinical Trials of Nonpharmacological Treatments

Edited by

Isabelle Boutron

Philippe Ravaud

David Moher

CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 2011922

International Standard Book Number-13: 978-1-4200-8802-1 (eBook - PDF)

Contents

Introduction

Isabelle Boutron and Philippe Ravaud
Paris Descartes University
Public Assistance—Hospitals of Paris

Nonpharmacological treatments represent a wide range of treatments proposed to patients. They could be defined as all interventions involving not just the administration of pharmacological treatments. Nonpharmacological treatments concern technical interventions such as surgical procedures; technical interventions such as joint lavage and angioplasty; implantable devices such as stents and arthroplasty; nonimplantable devices such as orthoses, ultrasound treatments, and laser treatments; and participative interventions such as rehabilitation, education, behavioral interventions, and psychotherapy.

The number of published randomized controlled trials assessing nonpharmacological treatments is increasing with time. A cross-sectional assessment of randomized trials published in 2000 identified 25% of such trials assessing nonpharmacological treatments (10% surgery or procedures, 11% counseling or lifestyle interventions, and 4% equipment) [1]. A similar recent study showed that randomized trials assessing nonpharmacological treatments concerned 42% of the trials published in 2006 (21% surgery or procedures, 18% counseling or lifestyle interventions, and 3% equipment).

Assessing nonpharmacological treatments raises specific issues. An important issue is the funding source. Most assessments of nonpharmacological treatments, except perhaps implantable and nonimplantable devices, rely on public funding, or more restricted amounts of money [2,3]. Further, the regulatory requirements for nonpharmacological treatments are less stringent than for pharmacological treatments. In most cases, the drug approval process of the U.S. Food and Drug Administration requires demonstrated treatment effectiveness from at least two adequate and well-controlled clinical trials.

In contrast, most nonpharmacological treatments such as surgical procedures or participative interventions have no specific requirements for approval. Consequently, they can be widely proposed in clinical practice but may not have been adequately evaluated. This situation is an important barrier for the evaluation of the beneficial effects of these treatments and the conduct of randomized controlled trials.

Finally, assessing nonpharmacological treatments raises specific methodological issues [3]. First, blinding of patients, care providers, and outcome assessors is frequently not feasible, particularly because of a lack of placebo for most nonpharmacological treatments [4]. Second, nonpharmacological treatments

are usually complex interventions made of several components that may all have an impact on the beneficial effect of the treatment [5]. These interventions are, consequently, difficult to describe, reproduce in the trial, and implement in clinical practice. Finally, care providers' expertise and centers' volume of care can have an important impact on the success of the interventions [6].

Nevertheless, it is essential to overcome these barriers and to adequately evaluate nonpharmacological treatments.

This book is divided in two parts. Part I is dedicated to specific issues when assessing nonpharmacological treatments. It highlights the difficulties of blinding and how these difficulties can be overcome. It discusses the placebos that can be used in such trials. It also addresses how the complexity of the intervention, the learning curve, and the clustering effect should be taken into account in trials. Issues of assessing harm and assessing the applicability of trials in this field are also raised. Different designs that are particularly useful in this context—cluster randomized controlled trials, expertise-based trials, pragmatic trials, and nonrandomized trials, as well as specific issues of systematic reviews in this field—are also presented.

Part II provides several examples of the planning, conduct, analyses, and reporting of trials in different fields. It is obviously impossible to cover all the different clinical areas, but these examples in the field of surgery, technical interventions, devices, rehabilitation, psychotherapy, behavioral interventions, etc., should be very useful for readers to learn and grasp some ideas from various domains.

References

1. Chan, A.W., Altman, D.G. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet*, 365(9465), 1159–1162, 2005.
2. Balasubramanian, S.P., Wiener, M., Alshameeri, Z., Tiruvoipati, R., Elbourne, D., Reed, M.W. Standards of reporting of randomized controlled trials in general surgery: can we do better? *Ann Surg*, 244(5), 663–667, 2006.
3. Boutron, I., Tubach, F., Giraudeau, B., Ravaud, P. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *JAMA*, 290(8), 1062–1070, 2003.
4. 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*, 57(6), 543–550, 2004.
5. Glasziou, P., Meats, E., Heneghan, C., Shepperd, S. What is missing from descriptions of treatment in trials and reviews? *BMJ*, 336(7659), 1472–1474, 2008.
6. Halm, E.A., Lee, C., Chassin, M.R. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med*, 137(6), 511–520, 2002.

Contributors

Charles Abraham
Peninsula College of Medicine and Dentistry
University of Exeter
Exeter, United Kingdom

Hidefumi Aoyama
Department of Radiology
Niigata University Graduate School of Medical and Dental Sciences
Niigata, Japan

Vincenzo Berghella
Department of Obstetrics and Gynecology
Thomas Jefferson University
Philadelphia, Pennsylvania

David Biau
Department of Biostatistics and Medical Informatics
Public Assistance—Hospitals of Paris
Paris, France

Isabelle Boutron
INSERM U738
Paris Descartes University
and
Center for Clinical Epidemiology
Public Assistance—Hospitals of Paris
Paris, France

Jonathan A. Cook
Health Services Research Unit
University of Aberdeen
Aberdeen, United Kingdom

Nadine Elizabeth Foster
Arthritis Research UK Primary Care Centre
Keele University
Staffordshire, United Kingdom

Bruno Giraudeau
Center for Clinical Investigation (INSERM CIC 0202)
University Hospital of Tours
Tours, France

Paul Glasziou
Faculty of Health Sciences and Medicine
Bond University
Queensland, Australia

Asbjørn Hróbjartsson
The Nordic Cochrane Centre
Copenhagen, Denmark

John P.A. Ioannidis
Department of Hygiene and Epidemiology
University of Ioannina School of Medicine
and
Biomedical Research Institute
Foundation for Research and Technology–Hellas
Ioannina, Greece

and

Tufts University School of Medicine
Boston, Massachusetts

and

Stanford University School of Medicine
Stanford, California

Claire Jourdan
Cochin Hospital
Public Assistance—Hospitals of Paris
and
Paris Descartes University
and
Federative Research Institute on Disability
Paris, France

Arthur Kang'ombe
Department of Health Sciences
University of York
York, United Kingdom

Simon Lewin
Norwegian Knowledge Centre for the Health Services
Norway and Health Systems Research Unit
Medical Research Council of South Africa
Cape Town, South Africa

Isabelle Marc
Department of Pediatrics
Laval University
Quebec City, Quebec, Canada

Hugh MacPherson
Department of Health Sciences
University of York
York, United Kingdom

Franklin G. Miller
Department of Bioethics
National Institutes of Health
Bethesda, Maryland

David Moher
Clinical Epidemiology Program
Ottawa Hospital Research Institute
Ottawa, Ontario, Canada

Remy Nizard
Department of Orthopedic Surgery and Trauma
Lariboisière Hospital
Public Assistance—Hospitals of Paris
and
Faculty of Medicine
Paris Diderot University
Paris, France

Panagiotis N. Papanikolaou
Department of Hygiene and Epidemiology
University of Ioannina School of Medicine
Ioannina, Greece

Rafael Perera
Department of Primary Health Care
University of Oxford
Oxford, United Kingdom

Serge Poiraudeau
Cochin Hospital
Public Assistance—Hospitals of Paris
Paris Descartes University
and
Federative Research Institute on Disability
Paris, France

Craig R. Ramsay
Health Services Research Unit
University of Aberdeen
Aberdeen, United Kingdom

François Rannou
Cochin Hospital
Public Assistance—Hospitals of Paris
and
Paris Descartes University

and
Federative Research Institute
on Disability
Paris, France

Keith G. Rasmussen
Department of Psychiatry and
Psychology
Mayo Clinic
Rochester, Minnesota

Philippe Ravaud
INSERM U738
Paris Descartes University
and
Center for Clinical Epidemiology
Public Assistance—Hospitals
of Paris
Paris, France

Barnaby C. Reeves
Bristol Heart Institute
University of Bristol
Bristol, United Kingdom

Katherine Sanchez
Cochin Hospital
Public Assistance—Hospitals
of Paris
and
Paris Descartes University
and
Federative Research Institute
on Disability
Paris, France

Paula P. Schnurr
National Center for Posttraumatic
Stress Disorder
White River Junction, Vermont

and

Dartmouth Medical School
Hanover, New Hampshire

Larissa Shamseer
Clinical Epidemiology Program
Ottawa Hospital Research Institute
Ottawa, Ontario, Canada

Sasha Shepperd
Department of Public Health
University of Oxford
Oxford, United Kingdom

Michele Tansella
Department of Medicine and Public
Health
University of Verona
Verona, Italy

Graham Thornicroft
Health Service and Population
Research Department
King's College London
London, United Kingdom

Helen Tilbrook
Department of Health Sciences
University of York
York, United Kingdom

Jorge E. Tolosa
Department of Obstetrics and
Gynecology
Oregon Health Science University
Portland, Oregon

David Torgerson
Department of Health Sciences
University of York
York, United Kingdom

Tom Treasure
Department of Mathematics
University College London
London, United Kingdom

Alexander Tsertsvadze
Clinical Epidemiology Program
Ottawa Hospital Research Institute
Ottawa, Ontario, Canada

Martin Utley
Department of Mathematics
University College London
London, United Kingdom

Patricia Yudkin
Department of Primary Health Care
University of Oxford
Oxford, United Kingdom

Merrick Zwarenstein
Sunnybrook Research Institute
and
Institute for Clinical Evaluative Sciences
University of Toronto
Toronto, Ontario, Canada

and

Division of Global Health (IHCAR)
Karolinska Institute
Stockholm, Sweden

Part I

Assessing Nonpharmacological Treatments: Theoretical Framework

1

Blinding in Nonpharmacological Randomized Controlled Trials

Isabelle Boutron and Philippe Ravaud

Paris Descartes University and
Public Assistance—Hospitals of Paris

CONTENTS

1.1 General Framework on Blinding

Blinding is a cornerstone of unbiased therapeutic evaluation [1,2]. Blinding refers to keeping key people, such as participants, healthcare providers, and outcome assessors, unaware of the treatment administered or of the true hypothesis of the trial [3,4].

Blinding of participants and healthcare providers in a trial prevents performance bias, which occurs when knowledge of the treatment assignment may affect the willingness of healthcare providers to prescribe and participants to take co-interventions, participants to be compliant with the assigned treatment, and participants to cross over or withdraw from the trial [5–7]. For example, in a randomized controlled trial comparing surgery for lumbar intervertebral disk herniation with usual care, blinding of patients, care providers, and outcome assessors was not feasible [8]. Lack of blinding was responsible for an important contamination between the two groups, with 50% of patients assigned to surgery receiving surgery within 3 months of enrolment, and 30% of those assigned to nonoperative treatment receiving surgery in the same period. Blinding of outcome assessors also minimizes the risk of detection bias (i.e., observer, ascertainment, assessment bias). This type of bias occurs if participant assignment influences the process of outcome assessment [5–7]. For example, nonblinded neurologists assessing the outcome of a trial demonstrated an apparent treatment benefit, whereas

blinded neurologists did not [3]. Finally, blinding of data analysts can also prevent bias because knowledge of the intervention received may influence the choice of analytical strategies and methods [5].

There is some empirical evidence of bias when blinding is lacking. Schulz et al. [1] evaluated the association of estimates of treatment effect and lack of double-blinding. Trials that were not double-blinded yielded larger effect estimates, with odds ratios exaggerated by 17%. Moher et al. [9] performed a meta-epidemiological study to estimate the effect of different quality indicators such as adequate randomization generation, allocation concealment, and reporting of double-blinding. The authors showed an overestimation of treatment effect estimates for randomization, generation, and allocation concealment but not for double-blinding. Recently, Wood et al. [10] showed that the impact of blinding depended highly on the type of outcome evaluated; in trials with subjective outcomes that lacked blinding, treatment effect estimates were exaggerated by 25%. In contrast, trials with objective outcomes showed no evidence of bias. Nevertheless, these meta-epidemiological studies raise some issues. They are indirect evidence susceptible to a considerable risk of confounding. In fact, double-blind randomized trials can differ from other trials in other important aspects such as the treatment evaluated (pharmacological or nonpharmacological), the randomization procedure, the funding source, or other unknown factors. Further, meta-epidemiological studies do not take into account: who was blinded, whether blinding was efficient, or the possible risk of bias.

Direct evidence of bias demonstrating the influence of lack of blinding is sparse. A systematic review of >20 randomized controlled trials with blinded and nonblinded outcome assessment showed a substantial impact of the blinding of outcome assessors, especially in trials with more subjective outcomes [11].

The reporting of blinding in published reports of randomized controlled trials is frequently inadequate. Most publications use a common terminology of single-blind, double-blind, or triple-blind study. However, this terminology should be used with caution. In fact, the use of the terms is confusing because it means different things to different people [12,13]. For example, a single-blind randomized trial could imply that patients are blinded or that outcome assessors are blinded. Further, many authors neglect to report whether their trial was blinded, who was blinded, and how blinding was achieved [14–16]. Haahr and Hrobjartsson evaluated how blinding was reported in 200 blinded randomized clinical trials with articles published in 2001; 78% of the articles described the trial as "double-blind," with only 2% explicitly reporting the blinding status for each trial participant (patients, care providers, outcome assessor). After contacting the authors of the studies, Haahr and Hrobjartsson [15] showed that about one-fifth of the "double-blind" trials did not blind patients, care providers, or outcome assessors. Hróbjartsson et al. [17] showed that the reporting of data related to blinding was better in protocols of studies than the published results, but a large proportion of protocols still report blinding unclearly.

To improve the quality of reporting of blinding, an international group, the CONSORT group, developed reporting guidelines, or statements, first published in 1996 [18] and updated in 2001 [19,20] and 2010 [21,22]. These guidelines are now endorsed and cited in the recommendations to authors of most peer-review journals. The guidelines clearly indicate that the authors should report "If [blinding was] done, who was blinded after assignment to interventions (e.g., participants, care providers, those assessing outcomes) and how" and provide "If relevant, a description of the similarity of interventions."

1.2 Blinding and Nonpharmacological Trials

Blinding is essential to limit the risk of bias; however, blinding is not always feasible. A study of a sample of randomized controlled trials assessing pharmacological and nonpharmacological treatments in the field of osteoarthritis showed that blinding was almost always feasible for patients, care providers, and outcome assessors in trials assessing pharmacological treatments. However, in trials assessing nonpharmacological treatments, blinding was considered feasible in only 42% of the trials for patients, 12% for care providers, and 34% for outcome assessors. When blinding was judged feasible, the perceived risk of nonblinding was more often considered moderate or important in trials assessing nonpharmacological treatments. When blinding was judged feasible, it was reported less often in nonpharmacological reports. These differences are linked to the difficulties of finding a placebo for nonpharmacological treatments. In fact, the procedures of blinding mainly rely on the use of a placebo defined as a control intervention with similar appearance as the experimental treatment but devoid of the components in the experimental intervention whose effects the trial is designed to evaluate. A placebo is usually feasible for pharmacological treatments but raises important issues for nonpharmacological treatments. For example, what is an appropriate placebo for a surgical procedure? for psychotherapy? These issues will be discussed in Chapter 2. Frequently, providing a completely similar placebo intervention in the control group is not possible, and the use of partial blinding could be proposed. Such a procedure should not be considered a panacea because complete blinding is not achieved. However, these procedures aim to limit the risk of bias. Patients could be blinded to the study hypothesis, that is, patients are aware that they will have a 50% chance of receiving one of the two interventions being evaluated, that they do not know which intervention is the most effective, and that for scientific reasons, they cannot be informed of all the hypotheses of the trial. In other situations, patients could be aware that one of the interventions is a placebo but will not be informed of the nature of the placebo.

To overcome the difficulties of blinding patients and care providers, a prospective randomized open, blinded end-point (PROBE) study could

be proposed [23]. Such a study limits the risk of detection bias because of blinded assessment of the outcome. This method has been proposed for different types of outcomes. It mainly relies on a centralized and blinded assessment of the outcome. For physician-driven data, the study could be a centralized assessment of clinical examinations through the use of photography, video, or audio of an interview. For example, in a trial evaluating the efficacy of multi-wavelength light therapy to treat pressure ulcers in subjects with disorders of the spinal cord [24], photographs of the ulcers were taken at the beginning and end of treatment and at 14 days after the last session. All evaluations were performed by a blinded outcome assessor. To assess the effects of a treatment for verbal communication in aphasia after stroke [25], patient responses were tape-recorded and scored by two independent blinded observers. When the outcome is a complementary test, a centralized assessment of the test will avoid bias. In a trial evaluating off-pump versus conventional coronary artery bypass grafting—early and 1 year graft patency—three cardiologists who were blinded to group assignment simultaneously reviewed angiograms [26].

For clinical events such as occurrence of myocardial infarction, a blinded adjudication committee is useful. However, this situation still entails a risk of bias, particularly if the adjudication committee evaluates and adjudicates only the events identified and transmitted by nonblinded investigators. Therefore, what was evaluated by the blinded adjudication committee must be considered: Did the adjudication committee evaluate all the patients included in the study (which is difficult to achieve because of time and cost)? Were patients systematically screened by a routine check of biochemical markers and electrocardiographic analyses by core laboratories? Were specific computer algorithms used to identify events? Mahaffey et al. [27] used a computer algorithm to show that 270 cases of myocardial infarction (5.2% of all patients enrolled) were not identified by site investigators and 134 cases (2.6%) identified by site investigators were not confirmed by the adjudication committee. Similarly, an independent review of case-report forms from the Rosiglitazone Evaluated for Cardiac Outcomes and Regulation of Glycemia in Diabetes (RECORD) trial [28] by the U.S. Food and Drug Administration (FDA) provided evidence of bias in that investigators were aware of the treatment allocated. In the RECORD study, the method for selecting cases to be assessed by the adjudication committee relied on whether the nonblinded investigators identified and reported the case. Only 12.5% of the case-report forms were reviewed. Errors such as a patient with an event not referred for adjudication were systematically bias with more errors in the experimental group and 81% of errors favoring the experimental group [27].

A frequent and difficult situation in assessing nonpharmacological treatments is the comparison of the treatment with usual care. In this situation, blinding of patients and care providers is not feasible. The risk of bias may be particularly important because of the deception of patients who will not receive any treatments, which is particularly problematic when the primary

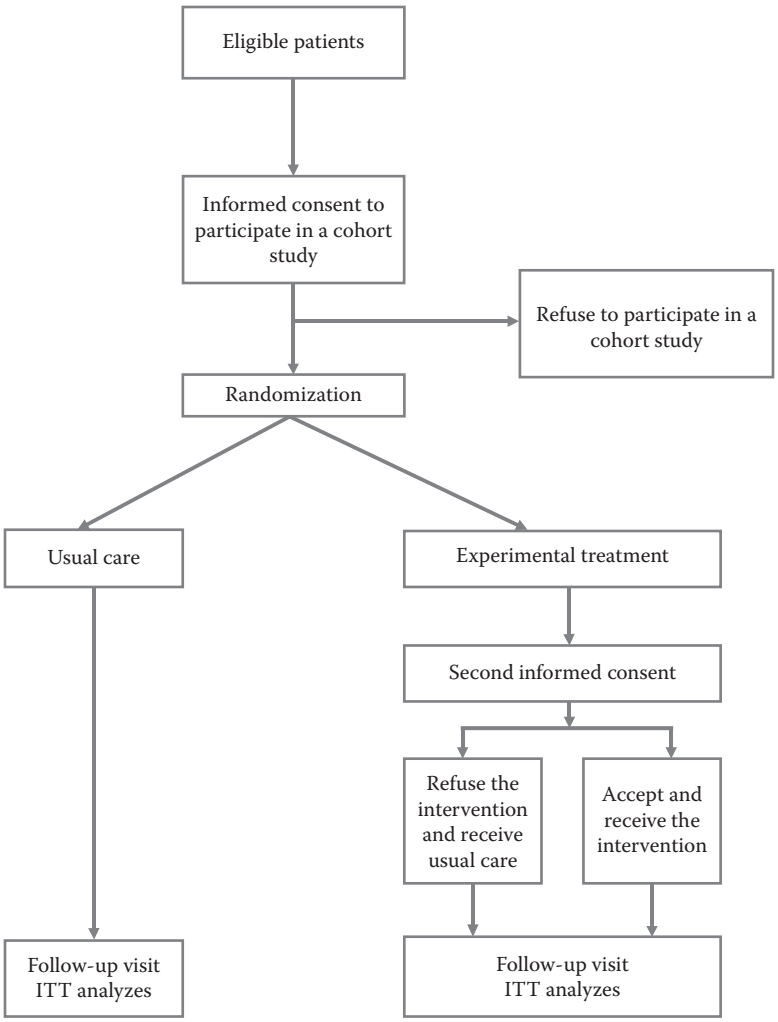

FIGURE 1.1
Modified Zelen design.

outcome is a patient-reported outcome (e.g., pain, quality of life). Some specific designs may be proposed in these situations. A modified Zelen design has been proposed, although it has been criticized for ethical issues [29,30] (see Figure 1.1). In a first step, patients are invited to participate in a cohort study. They are informed of the different follow-up visits and sign a consent form. In a second step, patients are randomized. Patients randomized to receive the experimental treatments are informed and sign a second consent form. Such a design avoids deceiving patients. However, this design raises some issues. An ethical issue is that some ethics committee will not agree to approve studies in which patients will be blinded to study hypotheses.

However, this issue has to be balanced with the ethical issue of conducting a trial knowing that there will be a high risk of bias. Other issues are related to logistics in terms of ensuring that patients from both groups do not meet. Finally, this design is not adequate in studies with a high risk that patients will refuse the experimental intervention because the analysis will have to be an intention-to-treat analysis and the high rate of refusal will decrease the power of the trial. Recently, Relton and colleagues proposed the cohort multiple randomized controlled trial design, which could also help address these methodological issues [31].

When blinding is not performed, the risk of bias must be evaluated. In fact, this approach is necessary for the critical appraisal of published results of randomized controlled trials but also when assessing the risk of bias in trial results included in systematic reviews and meta-analysis. The Cochrane collaboration has developed a specific tool to evaluate the risk of bias in randomized controlled trials, the Risk of Bias Tool (the RoB tool) [5]. The RoB tool recommends evaluating the risk of performance bias and the risk of detection bias. The risk of performance bias is high if patients and care providers are not blinded and if the outcome is likely to be influenced by the lack of blinding (e.g., crossover trials, differential co-interventions, and differential attrition). The risk of detection bias will be high if the outcome assessor is not blinded and the outcome measurement is likely to be influenced by lack of blinding (e.g., subjective outcomes).

1.3 Conclusion

Blinding is essential to limit the risk of bias. However, blinding is more difficult to achieve and maintain in trials assessing nonpharmacological treatments. There is a need to use creative methods of blinding, and in some situations to accept that the only way to limit the risk of bias is to attempt partial blinding. Nevertheless, frequently, blinding is not feasible, and evaluating the risk of bias in such studies is necessary.

References

1. Schulz, K.F., Chalmers, I., Hayes, R.J., Altman, D.G. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, 273(5), 408–412, 1995.
2. Schulz, K.F., Altman, D.G., Moher, D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*, 152(11), 726–732, 2010.

3. Noseworthy, J.H., Ebers, G.C., Vandervoort, M.K., Farquhar, R.E., Yetisir, E., Roberts, R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology*, 44(1), 16–20, 1994.
4. Day, S.J., Altman, D.G. Statistics notes: Blinding in clinical trials and other studies. *BMJ*, 321(7259), 504, 2000.
5. Julian, H., Sally, G. *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley-Blackwell, Hoboken, NJ, 2008.
6. Schulz, K.F., Grimes, D.A. Blinding in randomised trials: Hiding who got what. *Lancet*, 359(9307), 696–700, 2002.
7. Schulz, K.F., Grimes, D.A., Altman, D.G., Hayes, R.J. Blinding and exclusions after allocation in randomised controlled trials: Survey of published parallel group trials in obstetrics and gynaecology. *BMJ*, 312(7033), 742–744, 1996.
8. Weinstein, J.N., Tosteson, T.D., Lurie, J.D., et al. Surgical vs nonoperative treatment for lumbar disk herniation: The Spine Patient Outcomes Research Trial (SPORT): A randomized trial 10.1001/jama.296.20.2441. *JAMA*, 296(20), 2441–2450, 2006.
9. Moher, D., Pham, B., Jones, A., et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*, 352(9128), 609–613, 1998.
10. Wood, L., Egger, M., Gluud, L.L., et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: Meta-epidemiological study. *BMJ*, 336(7644), 601–605, 2008.
11. Hróbjartsson, A., Thomsen, A., Emanuelsson, F., et al. Empirical evidence of observer bias in randomised trials: A meta-analysis, 2011 (submitted).
12. Devereaux, P.J., Manns, B.J., Ghali, W.A., et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA*, 285(15), 2000–2003, 2001.
13. Devereaux, P.J., Bhandari, M., Montori, V.M., Manns, B.J., Ghali, W.A., Guyatt, G.H. Double blind, you are the weakest link—Good-bye! *ACP J Club*, 136(1), A11, 2002.
14. Montori, V., Bhandari, M., Devereaux, P., Manns, B., Ghali, W., Guyatt, G. In the dark: The reporting of blinding status in randomized controlled trials. *J Clin Epidemiol*, 55(8), 787, 2002.
15. Haahr, M.T., Hrobjartsson, A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clin Trials*, 3(4), 360–365, 2006.
16. Chan, A.W., Altman, D.G. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet*, 365(9465), 1159–1162, 2005.
17. Hróbjartsson, A., Pildal, J., Chan, A.W., Haahr, M.T., Altman, D.G., Gotzsche, P.C. Reporting on blinding in trial protocols and corresponding publications was often inadequate but rarely contradictory. *J Clin Epidemiol*, 62(9), 967–973, 2009.
18. Altman, D.G. Better reporting of randomised controlled trials: The CONSORT statement. *BMJ*, 313, 570–571, 1996.
19. Altman, D.G., Schulz, K.F., Moher, D., et al. The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Ann Intern Med*, 134(8), 663–694, 2001.
20. Moher, D., Schulz, K.F., Altman, D.G. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*, 357(9263), 1191–1194, 2001.

21. Schulz, K.F., Altman, D.G., Moher, D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *PLoS Med*, 7(3), e1000251, 2010.
22. Schulz, K.F., Altman, D.G., Moher, D., Fergusson, D. CONSORT 2010 changes and testing blindness in RCTs. *Lancet*, 375(9721), 1144–1146, 2010.
23. Hansson, L., Hedner, T., Dahlof, B. Prospective randomized open blinded end-point (PROBE) study. A novel design for intervention trials. Prospective randomized open blinded end-point. *Blood Press*, 1(2), 113–119, 1992.
24. Taly, A.B., Sivaraman Nair, K.P., Murali, T., John, A. Efficacy of multiwavelength light therapy in the treatment of pressure ulcers in subjects with disorders of the spinal cord: A randomized double-blind controlled trial. *Arch Phys Med Rehabil*, 85(10), 1657–1661, 2004.
25. Doesborgh, S.J., van de Sandt-Koenderman, M.W., Dippel, D.W., van Harskamp, F., Koudstaal, P.J., Visch-Brink, E.G. Effects of semantic treatment on verbal communication and linguistic processing in aphasia after stroke: A randomized controlled trial. *Stroke*, 35(1), 141–146, 2004.
26. Puskas, J.D., Williams, W.H., Mahoney, E.M., et al. Off-pump vs conventional coronary artery bypass grafting: Early and 1-year graft patency, cost, and quality-of-life outcomes: A randomized trial. *JAMA*, 291(15), 1841–1849, 2004.
27. Mahaffey, K.W., Roe, M.T., Dyke, C.K., et al. Misreporting of myocardial infarction end points: Results of adjudication by a central clinical events committee in the PARAGON-B trial. Second platelet IIb/IIIa antagonist for the reduction of acute coronary syndrome events in a global organization network trial. *Am Heart J*, 143(2), 242–248, 2002.
28. Home, P.D., Pocock, S.J., Beck-Nielsen, H., et al. Rosiglitazone evaluated for cardiovascular outcomes in oral agent combination therapy for type 2 diabetes (RECORD): A multicentre, randomised, open-label trial. *Lancet*, 373(9681), 2125–2135, 2009.
29. Quilty, B., Tucker, M., Campbell, R., Dieppe, P. Physiotherapy, including quadriceps exercises and patellar taping, for knee osteoarthritis with predominant patello-femoral joint involvement: Randomized controlled trial. *J Rheumatol*, 30(6), 1311–1317, 2003.
30. Bravata, D.M., Smith-Spangler, C., Sundaram, V., et al. Using pedometers to increase physical activity and improve health: A systematic review. *JAMA*, 298(19), 2296–2304, 2007.
31. Relton, C., Torgerson, D., O'Cathain, A., Nicholl, J. Rethinking pragmatic randomised controlled trials: Introducing the "cohort multiple randomised controlled trial" design. *BMJ*, 340, c1066, 2010.

2

*Placebo in Nonpharmacological Randomized Trials**

Asbjørn Hróbjartsson
The Nordic Cochrane Centre

Franklin G. Miller
National Institutes of Health

CONTENTS

2.1 Introduction

One of the first clinical trials to explicitly use a placebo control group was published in 1938 by Diel et al. [1]. The trial compared patients treated with capsules containing vaccine for common cold with patients treated with placebo capsules containing lactose. In contrast to previous trials, Diel et al. found no effect of oral vaccine.

Since then, a vast multitude of placebo-controlled trials have been conducted. A search for "placebo*" in The Cochrane Central Register of Controlled Trials in April 2009 provided 111,592 references. The majority of hits are to pharmacological trials comparing a drug with a placebo, in some cases as an addition to a standard care regime, or as part of a "doubly dummy" design.

Placebo interventions are difficult to define unambiguously [2,3]. However, within a clinical trial, a placebo can be characterized as a control intervention

* The opinions expressed are the views of the author and do not necessarily reflect the policy of the National Institutes of Health, the Public Health Service, or the U.S. Department of Health and Human Services.

with similar appearance as the experimental treatment, but void of the components in the experimental intervention whose effects the trial is designed to evaluate.

Roughly 24% of contemporary randomized trials indexed in PubMed are nonpharmacological [4]. Nonpharmacological trials involve diverse types of interventions, and equally diverse types of placebo control groups. For the sake of clarity in the following, we will primarily address three types of nonpharmacological placebo control interventions: devices, surgical interventions, and psychological interventions. These three types of interventions exemplify the most typical challenges posed to nonpharmacological trials in general.

In this chapter, we will describe and discuss nonpharmacological placebo interventions as used in clinical trials. In general, device placebos and surgical placebos are conceptually similar to pharmacological placebos, though they pose some distinctive methodological and ethical issues. In both types of trials, there is a noticeable risk of unblinding, and surgical placebo trials carry risks to subjects from the placebo intervention itself. Psychological placebos are generally of a different type conceptually, as they are often designed to control for specific factors, such as expectancy, and not as a tool for blinding. When designing a nonpharmacological trial, or when interpreting the results of such a trial, the exact nature of the placebo intervention deserves considerable attention.

2.2 Placebo-Controlled Trials

The main aim of a placebo-controlled trial is to establish whether the components of the experimental intervention, hypothesized to be effective, in fact can produce clinically significant benefit in patients with a given medical condition. Accordingly, the experimental intervention is compared with a placebo control that appears indistinguishable and lacks the components of the intervention hypothesized to be responsible for its therapeutic efficacy.

There are two fundamental differences between a placebo control group and a no-treatment control group. First, a trial using a no-treatment control group tests whether an intervention as a whole has an effect. The design can say nothing clear about which component within a treatment package is the main causal factor. Second, the design is unreliable, especially when outcomes are subjective, because it does not permit masking of the study intervention and comparator.

The notion of a placebo control is historically linked with the idea that placebo interventions cause large effects. Especially after Beecher in 1955 published a review of the improvements reported in the placebo groups of 14 trials, it became a standard notion that placebo interventions had large

effects on many patients on both objective and subjective outcomes [5]. His assessment of "the powerful placebo" was based on a comparison between baseline and posttreatment in placebo control groups, and did not control for natural fluctuations in the patient's condition, including spontaneous remission, and regression to the mean. Nonetheless, this article was an important factor in persuading clinicians that randomized trials were necessary and ethical.

A recent update of a systematic review of 202 randomized trials with both placebo groups and no-treatment control groups found small to moderate differences between no treatment and placebo overall, but effects were more pronounced in trials with patient-reported outcomes [6]. For pain the mean effect corresponded to roughly 6 mm on a 100 mm visual analogue scale; however, in certain settings the effect was larger. Four well-performed German acupuncture trials reported an effect of placebo acupuncture of roughly 17 mm on a 100 mm visual analogue scale. It is noteworthy that the patients involved in the trials were falsely informed that the study involved a comparison between two types of acupuncture, and not between acupuncture and placebo acupuncture [7]. The effect of placebo was considerably smaller in other acupuncture trials [8]. When all trials with continuous outcomes were examined in a regression analysis, there was a clear tendency for larger effects in device placebos and psychological placebos as compared with pharmacological placebos. The general pattern of results from the review, when disregarding the obvious risk of reporting bias, is that placebo interventions can affect subjective outcomes, but that this effect is quite variable, and dependent on underlying causal factors, for example, patient information and type of placebo.

Besides controlling for placebo effects, there are additional compelling reasons for implementing masked placebo control groups. Reporting bias occurs when patients report their symptoms more favorably than they otherwise would have, for example, because of courtesy to the doctor who offers them treatment. In the case of placebo-controlled trials of invasive treatments, patients may be disposed to perceive or report benefit as a result of having undergone a burdensome or seemingly powerful intervention. This type of bias is much more likely to occur when an intervention is compared with a no-treatment control group instead of a placebo group. Similarly, attrition bias occurs when patients stop the trial, or do not adhere to the treatment, because they wanted to be in the other treatment group. This type of behavior is also likely to be more pronounced in no-treatment group as compared with a placebo group.

It is a common misunderstanding that placebo control interventions have to be "inert" [9]. Strictly speaking, classic placebos such as sugar pills and saline infusions are not inert, since they contain biologically active ingredients. They are "inert" only in the relative sense that there is no scientific reason to think that the sugar or salt in the placebo intervention will have an effect on the outcomes of interest in a clinical trial. Similarly, in sham surgery trials, comparing a real to a fake surgical procedure, the invasive

placebo control obviously is not inert. But as long as the placebo surgical intervention does not include the surgical manipulation hypothesized to be responsible for the outcomes under investigation, it counts as a valid control. Likewise, a sham acupuncture intervention (whether superficial needling at non-acupuncture points or a retractable device) constitutes a valid control for detecting whether the needling techniques characteristic of traditional acupuncture are responsible for clinical effects, regardless of the possibility that the physical stimulus provided by the sham acupuncture intervention might itself have an effect.

The ethics of placebo-controlled trial has been debated intensely [10]. There is no doubt that placebo-controlled trials are used in many situations where there is an established treatment. For example, it is routine to use of placebo controls in many psychiatric conditions and conditions in which pain is the outcome, despite proven effective treatment [11]. The fifth revision (1996) of the Helsinki Declaration stated that: "The benefits, risks, burdens and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods. This does not exclude the use of placebo, or of no-treatment, in studies where no proven prophylactic, diagnostic or therapeutic method exist." [12]. Taken at face value, this means that a large number of placebo-controlled trials would be in violation with the fifth revision of the declaration.

However, the sixth revision (2008) is importantly different: "The benefits, risks, burdens and effectiveness of a new intervention must be tested against those of the best current proven intervention, except in the following circumstances: The use of placebo, or no treatment, is acceptable in studies where no current proven intervention exists; or where for compelling and scientifically sound methodological reasons the use of placebo is necessary to determine the efficacy or safety of an intervention and the patients who receive placebo or no treatment will not be subject to any risk of serious or irreversible harm. Extreme care must be taken to avoid abuse of this option." [13].

This shift toward endorsing placebo-controlled trials finds support from three strange bedfellows. First, the pharmaceutical industry is generally very interested in establishing effect beyond placebo, and much less interested in a trial that risk showing a drug to be less effective than a standard therapy. Second, the U.S. Food and Drug Administration (FDA) are concerned with the so-called "assay sensitivity," by which they mean the ability of trials to detect an effect compared with an ineffective therapy, and normally require new drugs to show superiority to placebo in at least two trials, before they approve new drugs [14]. Third, researchers writing from an evidence-based medicine perspective have pointed out that "proven" effective may quite often be illusive [15]. What clinicians have thought to be an effective therapy has later often been shown to have no effect, or that the harmful effects outweighed the beneficial effects. If an intervention falsely is regarded effective, and new interventions are compared with this false positive yardstick, we risk introducing a number of equally ineffective interventions.

2.3 Nonpharmacological Placebo Interventions: Device Placebos

To illustrate the nature of contemporary nonpharmacological placebo interventions, a search on PubMed from November 2008 to February 2009 for publications containing the terms placebo* or sham* and indexed as "randomized controlled trial" provided 21 references (Table 2.1). Twelve trials used various forms of placebo devices: two trials using ineffective ultrasound machines, two trials using ineffective lasers, three trials using ineffective magnetic or electric stimulation, and five trials using other forms of sham devices, for example, paper filters in an air cleaner. There were six trials with placebo acupuncture or acupressure procedures. One trial used a manual placebo procedure, and there were two psychological trials.

The typical nonpharmacological placebo intervention thus seems to be a device. An illustrative example is the trial by Sulser et al., investigating the effect of high-efficiency particulate arresting (HEPA) air cleaner filters on the symptoms of asthma in children and adolescents sensitized to cat and dog allergens. The machines containing the filters were identical, and the only difference between the "active" and the "placebo" machine was that the active air cleaners contained HEPA filters, and that the placebo machines contained paper filters.

The trial is very similar to the standard pharmacological placebo-controlled trial. It is fairly easy to construct two machines that appear to be identical, one with a true HEPA filter and one with a paper filter. The active component in the trial is clearly defined and delineated. There are similarly no conceptual challenges in constructing placebo devices for ultrasound, or magnetic/electronic devices, that appear identical to the real devices, but without their magnetic or electronic property.

The difference between a pharmacological and a device placebo is most often of a practical kind, often concerning the risk of unblinding. Patients may try to check whether their intervention is placebo or not and this may be easier when treated with a device placebo than a pharmacological placebo. For example, in Chen et al.'s trial of magnetic knee wrappers (Table 2.1), it would be easy for patients to test whether their wrappers attracted metal sometimes during the 12-week period of the trial.

Another difference appears when it is difficult to construct a device placebo void of the active component tested in the experimental group. For example, in Chermont et al.'s trial of the effect of continuous positive airway pressure (CPAP) for chronic heart failure, a CPAP placebo was used with a low air pressure (0–1 mm H_2O) as compared with the higher pressure in the real CPAP group (3 mm). This is different from the classic pharmacological placebo trial, in that placebos differ from the active treatment in dose, not in nature. The procedure is meaningful only as long as the assumption of subtherapeutic dose is correct.

TABLE 2.1

Nonpharmacological Randomized Trials with Placebo Groups[a]

Trial	Clinical Problem	Experimental Procedure	Placebo Procedure
Stowell et al. [1]	Phlebotomy pain	Ultrasound	Unclear
Özgönenel et al. [2]	Knee osteoarthritis	Ultrasound	The applicator was disconnected to the ultrasound machine
Deng et al. [3]	Pain after thoracotomy	Acupuncture	Sham studs not penetrating the skin, and placed at sites not … true acupuncture sites
Elden et al. [4]	Pelvic girdle pain in pregnancy	Acupuncture	Nonpenetrating needles and no attempt to evoke "Qi"
Gaudet et al. [5]	Labor initiation	Acupuncture	Needling in sites not known to have an effect on the initiation of labor
Nordio and Romanelli [6]	Insomnia	Acupressure	Application of wrist pressure at a site different from the true HT 7 Shenmen acupuncture point
Sima and Wang [7]	Cisplatin-induced nausea	Acupuncture	Needling at points not regarded effective for nausea and vomiting
Desantana et al. [8]	Postoperative pain	Transcutaneous electrical nerve stimulation	No electronic stimulation, but the machine displayed an active indicator light
Schutter et al. [9]	Mood in healthy subjects	Transcranial magnetic stimulation	The device mimics the sound click … but the brain is shielded from actual stimulation [with an aluminum plate]
Lisanby et al. [10]	Major depression	Transcranial magnetic stimulation	A magnetic shield "limited the magnetic energy reaching the cortex to 10% …." Active and sham coils had "similar appearance, placement, and acoustic properties"
Koenigs et al. [11]	Emotional function in healthy subjects	Transcranial direct current stimulation	Stimulation for only 30 s
Chen et al. [12]	Knee osteoarthritis	Magnetic knee wrap	No magnetic activity
Kuhn et al. [13]	Oral mucositis	Low-level infrared laser therapy	The laser was turned off, but patients were blindfolded

TABLE 2.1 (continued)

Nonpharmacological Randomized Trials with Placebo Groups[a]

Trial	Clinical Problem	Experimental Procedure	Placebo Procedure
Teggi et al. [14]	Chronic tinnitus	Low-level laser	The [laser] device was pointed into the ear canal but the laser remained inactive
Azarpazhooh et al. [15]	Dentin hypersensitivity	Ozone machine	Machine delivered air
Sulser et al. [16]	Asthmatic children	Air cleaners with HEPA filters	Air cleaners with paper filters
Chermont et al. [17]	Chronic heart failure	Continuous positive airway pressure (3 mm H_2O)	Continuous positive airway pressure with low pressure (0–1 mm H_2O)
Lettieri and Eliasson [18]	Restless legs syndrome	Pneumatic compression (40 mm H_2O)	Pneumatic compression with low air pressure (3–4 mm H_2O)
Perry and Green [19]	Nervous activity in the lower limbs	Lumbar oscillatory mobilization	Same hand positioning but without any oscillatory movements
Edinger et al. [20]	Insomnia	Cognitive behavioral therapy and relaxation	Quasi-desensitization procedure
Walkup et al. [21]	Childhood anxiety	Cognitive behavioral therapy and sertraline	Placebo pill

References

1. Stowell, C.P., Trieu, M.Q., Chuang, H., Katz, N., Quarrington, C. Ultrasound-enabled topical anesthesia for pain reduction of phlebotomy for whole blood donation. *Transfusion*, 49(1), 146–153, 2009.
2. Ozgönenel, L., Aytekin, E., Durmuşoglu, G. A double-blind trial of clinical effects of therapeutic ultrasound in knee osteoarthritis. *Ultrasound Med Biol*, 35(1), 44–49, 2009.
3. Deng, G., Rusch, V., Vickers, A., et al. Randomized controlled trial of a special acupuncture technique for pain after thoracotomy. *J Thorac Cardiovasc Surg*, 136(6), 1464–1469, 2008.
4. Elden, H., Fagevik-Olsen, M., Ostgaard, H.C., Stener-Victorin, E., Hagberg, H. Acupuncture as an adjunct to standard treatment for pelvic girdle pain in pregnant women: Randomised double-blinded controlled trial comparing acupuncture with non-penetrating sham acupuncture. *BJOG*, 115(13), 1655–1668, 2008.
5. Gaudet, L.M., Dyzak, R., Aung, S.K., Smith, G.N. Effectiveness of acupuncture for the initiation of labour at term: A pilot randomized controlled trial. *J Obstet Gynaecol Can*, 30(12), 1118–1123, 2008.
6. Nordio, M., Romanelli, F. Efficacy of wrists overnight compression (HT 7 point) on insomniacs: Possible role of melatonin? *Minerva Med*, 99(6), 539–547, 2008.

(*continued*)

TABLE 2.1 (continued)

Nonpharmacological Randomized Trials with Placebo Groups[a]

7. Sima, L., Wang, X. Therapeutic effect of acupuncture on cisplatin-induced nausea and vomiting. *Zhongguo Zhen Jiu*, 29(1), 3–6, 2009.
8. Desantana, J.M., Sluka, K.A., Lauretti, G.R. High and low frequency TENS reduce postoperative pain intensity after laparoscopic tubal ligation: A randomized controlled trial. *Clin J Pain* 25(1), 12–19, 2009.
9. Schutter, D.J., Enter, D., Hoppenbrouwers, S.S. High-frequency repetitive transcranial magnetic stimulation to the cerebellum and implicit processing of happy facial expressions. *J Psychiatry Neurosci*, 34(1), 60–65, 2009.
10. Lisanby, S.H., Husain, M.M., Rosenquist, P.B., et al. Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: Clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology*, 34(2), 522–534, 2009.
11. Koenigs, M., Ukueberuwa, D., Campion, P., Grafman, J., Wassermann, E. Bilateral frontal transcranial direct current stimulation: Failure to replicate classic findings in healthy subjects. *Clin Neurophysiol*, 120(1), 80–84, 2009.
12. Chen, C.Y., Chen, C.L., Hsu, S.C., Chou, S.W., Wang, K.C. Effect of magnetic knee wrap on quadriceps strength in patients with symptomatic knee osteoarthritis. *Arch Phys Med Rehabil*, 89(12), 2258–2264, 2008.
13. Kuhn, A., Porto, F.A., Miraglia, P., Brunetto, A.L. Low-level infrared laser therapy in chemotherapy-induced oral mucositis: A randomized placebo-controlled trial in children. *J Pediatr Hematol Oncol*, 31(1), 33–37, 2009.
14. Teggi, R., Bellini, C., Piccioni, L.O., Palonta, F., Bussi, M. Transmeatal low-level laser therapy for chronic tinnitus with cochlear dysfunction. *Audiol Neurotol*, 14(2), 115–120, 2009.
15. Azarpazhooh, A., Limeback, H., Lawrence, H.P., Fillery, E.D. Evaluating the effect of an ozone delivery system on the reversal of dentin hypersensitivity: A randomized, double-blinded clinical trial. *J Endod*, 35(1), 1–9, 2009.
16. Sulser, C., Schulz, G., Wagner, P., et al. Can the use of HEPA cleaners in homes of asthmatic children and adolescents sensitized to cat and dog allergens decrease bronchial hyperresponsiveness and allergen contents in solid dust? *Int Arch Allergy Immunol*, 148(1), 23–30, 2009.
17. Chermont, S., Quintão, M.M., Mesquita, E.T., Rocha, N.N., Nóbrega, A.C. Noninvasive ventilation with continuous positive airway pressure acutely improves 6-minute walk distance in chronic heart failure. *J Cardiopulm Rehabil Prev*, 29(1), 44–48, 2009.
18. Lettieri, C.J., Eliasson, A.H. Pneumatic compression devices are an effective therapy for restless legs syndrome: A prospective, randomized, double-blinded, sham-controlled trial. *Chest*, 135(1), 74–80, 2009.
19. Perry, J., Green, A. An investigation into the effects of a unilaterally applied lumbar mobilisation technique on peripheral sympathetic nervous system activity in the lower limbs. *Man Ther*, 13(6), 492–499, 2008.
20. Edinger, J.D., Carney, C.E., Wohlgemuth, W.K. Pretherapy cognitive dispositions and treatment outcome in cognitive behavior therapy for insomnia. *Behav Ther*, 39(4), 406–416, 2008.
21. Walkup, J.T., Albano, A.M., Piacentini, J., et al. Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *N Engl J Med*, 359, 2753–2766, 2008.

[a] Based on a PubMed search for placebo* or sham*, restricted to "randomized clinical trials" and publicized from November 2008 to February 2009.

A somewhat special type of device trials is sham-acupuncture studies. There were 5 placebo acupuncture or acupressure trials among the 12 device trials listed in Table 2.1, reflecting that acupuncture placebo trials are common. The type of placebo control (usually called sham control within acupuncture research) varied. In one trial, the placebo intervention was penetrative ("Needling at points not regarded effective for nausea and vomiting"), whereas that was not the case for another trial ("Sham studs not penetrating the skin, and placed at sites not … true acupuncture sites"). Placebo acupuncture procedures involve manual manipulation, and typically an intense doctor–patient interaction and they therefore differ from most other device placebos. The risk of unblinding the patient through subtle cues in the acupuncturist's behavior is considerable, and the effects of needling in non-acupuncture points cannot be ruled out [8,16,17].

Placebo-controlled trials of device interventions tend to be conceptually similar to trials of pharmacological trials, though in general the practical construction of the placebo devices may be more challenging, and the risk of unblinding is higher. Acupuncture trials involve additional challenges, for example, to deal with intense patient–provider interaction.

2.4 Nonpharmacological Placebo Interventions: Surgical Placebos

In the 1950s, ligation of the internal mammary arteries for angina pectoris became popular in the United States. However, two small trials comparing the effect of ligation of the internal mammary arteries with the effect of placebo operation only (skin incision only) concluded that the procedure had no effect, and the operation became unfashionable. In 1961, Beecher energetically described the classic story emphasizing the need for rigorous randomized trials of surgical procedures [18].

His call for surgical trials is equally relevant today. The number of randomized trials in surgery is still very low compared with medicine in general. Only 10% of trials indexed in PubMed in December 2000 were Surgical/procedure trials, whereas 76% of were pharmacological trials [4]. Though the scarcity of surgical trials may be explained for reasons of practicality, history, lack of regulatory oversight requiring clinical trials before new surgical procedures are introduced into practice, and sparse funding sources [19], it remains a public health scandal that numerous surgical procedures are not reliably evaluated.

In the comparatively few surgical trials conducted, placebo-controlled trials are rare, though they tend to be highly publicized [20]. An illustrative example of a surgical placebo trial is Moseley et al.'s trial of arthroscopic lavage vs. arthroscopic debridement vs. placebo surgery in patients with

osteoarthritis of the knee [21]. The placebo surgery consisted in three 1 cm incisions in the skin after having received a short-acting intravenous tranquillizer and an opioid, and spontaneously breathed oxygen-enriched air. Patients were unaware of which treatment they had received. The main outcome was pain after 2 years, and the trial found no difference between the three groups.

In the trial by Moseley et al., both active interventions involved instrumentation within the knee joint, not performed on the patients in the placebo group. It is meaningful and comparatively easy to clearly define and delineate the active surgical components. From a conceptual point of view, many surgical placebos are similar to the trial by Moseley et al. It seems meaningful to construct a placebo surgery procedure as long as the active part of the surgical procedure is conducted on an anatomical entity that is either covered by skin, for example, a bone, or confined within an anatomical space, for example, a ligament within a joint, or procedures within the abdominal or thoracic cavities. However, some surgical procedures involve procedures that are impossible to mimic with a placebo surgery control group, for example, amputation of the lower limb.

The main problem with surgical placebo interventions is that, unlike pharmacological placebos, they can harm patients directly. The patients in the placebo surgical group may have to undergo a skin incision, have pain medication or anesthesia, with its potential harmful effects, and risk postoperative infection. In that sense, surgical placebo trials involve a more direct ethical challenge than pharmacological or device trials, where patients may be harmed, but more indirectly only if their participation precludes them from access to alternative effective treatment.

The result of the trial by Moseley et al. is much more reliable due to its placebo procedure than it would have been with a control group receiving usual medical therapy to treat pain or a nonblinded no-treatment control group. Still, Moseley et al. must have struggled somewhat with the unavoidable dilemma of when the added risk to the included patients was outweighed by the benefit to future patients. The authors implemented a quite strict informed consent procedure, stating that "placebo surgery will not benefit my knee arthritis" (44% of screened patients declined). Furthermore, the trialists made an effort to minimize the risk to placebo patients by not giving standard general anesthesia.

Surgical placebo trials are similar to the device placebo trials in that there often are potential problems of nonblinding. For example, it seems possible for some of the patients in the placebo knee surgery group to realize that their anesthesia procedure was different from standard procedures, and from there deduce that they had received placebo.

Another high-profile surgical placebo-controlled trial evaluated the effect of surgical implantation of fetal tissue to patients with Parkinson's disease [22]. The placebo procedure included "the placement of a stereotactic frame, target localization on magnetic resonance imaging,

the administration of general anesthesia with a laryngeal-mask airway, and a skin incision with a partial burr hole that does not penetrate the inner cortex of the skull." Whether the added risk to the included patients was outweighed by the benefit to future patients in this trial has been discussed stormily [23,24].

The additional risk involved in surgical trials is not necessarily as dramatic as a scull burr hole. The risk involved in some surgical trials is similar to that of other generally accepted procedures, for example, "muscle biopsy, bronchoscopy, and phase 1 testing of experimental drugs in healthy volunteers, which do not offer participants a prospect of direct benefit" [25]. Providing that the trial authors minimize the risk of the included patients, justify the remaining risk necessary to produce scientifically valid results, and carefully think through the informed consent procedure, a surgical placebo control group is not unethical per se.

Surgery is always harmful and sometimes is also beneficial. Placebo-controlled surgery trials are needed to reliably assess this balance, especially when outcomes are subjective. The ethical considerations involved in any placebo surgery control group needs to be delicately and cautiously analyzed, and the conclusion will probably differ according to the clinical scenario and the type of placebo surgery involved. However, in general, a surgical placebo control group seems attractive when outcomes are subjective, and the risks of the surgical placebo procedure are minor.

2.5 Nonpharmacological Placebo Interventions: Psychological Placebos

Effects of placebo interventions and effects of psychological interventions are both psychologically mediated. The discussion of placebo control groups in psychological trials is closely linked to the discussion of what exactly distinguishes psychological placebo interventions from psychological "verum" interventions. Both issues have been hotly debated in psychology for years, from Frank's classic characterization of placebo as a form of psychotherapy [26] to a more recent theme issue in *Journal of Clinical Psychology* [27].

One of the psychological trials listed in Table 2.1 is illustrative: Edinger et al. compared cognitive behavioral therapy with a placebo intervention for insomnia (there was also a third arm of relaxation training). The cognitive behavior intervention consisted of sessions providing practical information on sleep and stimulus control, and instructions to establish standard wake-up times, to get out of bed during extended awakenings, to avoid sleep-incompatible behavior in the bedroom, and to eliminate daytime napping. Furthermore, patients were given an initial time in bed prescription, which was modified during the sessions.

The placebo intervention ("quasi-desensitization") was presented to patients as a method to overcome "conditioned arousal." Therapists helped each patient to develop a hierarchy of common activities he/she did on awakening at night (e.g., opening eyes, clock watching). Therapists also helped them develop scenes of themselves engaged in neutral activities (e.g., reading the newspaper). In each session, patients were taught to pair neutral scenes with items on the hierarchy. The exercise was tape-recorded and the patient was given this tape locked in a player. The patients were told to practice their exercises at home once each day, but to avoid using the tape or exercise during sleep periods.

It is clear that the cognitive behavioral therapy intervention and the placebo intervention were not identical in appearance. In fact, the two treatments though vaguely similar were quite diverse, and the trial is more like a trial of two different, and differently appearing, interventions.

Psychological placebo control groups differ, but will rarely appear similar to the experimental intervention. Thus, though they are called "placebo groups," and have some similarities with standard placebo control groups, they are dissimilar, and in this title we use the term "placebo analogue control group."

In its most pragmatic, and primitive form, a placebo analogue control group consists of a pill placebo. For example, Walkup et al. compared the effect of cognitive behavioral therapy, sertraline, and combined therapy, with pill placebo (Table 2.1). Other trials compare psychological interventions with attention placebos, a kind of basic psychotherapy, often described as "neutral nondirective" void of any "specific" content. For example, in one trial of the effect of cognitive behavioral therapy for depression after stroke, the "attention placebo" was described as "a conversation that focused on day-to-day occurrences and discussions regarding the physical effects of stroke and life changes" [28].

Finally, a third group of placebo control groups have been laboriously developed as to be as similar as possible to an experimental psychological treatment. Such "authoritative" placebos, after having been successful as comparators to a highly specialized psychological intervention, are then sometimes used as controls for other, quite diverse, interventions. Examples are "quasi-desensitization" described earlier, and "pseudo-meditation" [29].

The basic idea behind such psychological placebo analogues is that patients are presented to a treatment with equal credibility as the experimental treatment, and often also equal patient–provider time, and thus positively control for these two important factors. Placebo analogue control groups are fundamentally different from the classic placebo control groups in many pharmacological or device trials, because they aim to control for specifically defined factors (typically patient–provider time and treatment credibility, or sometimes what is called "common factors"), whereas classic placebos aim to control for all known and unknown factors.

The major challenge is that placebo analogues cannot control for unknown factors. However, as long as patients can perceive the differences between the