Published in final edited form as:
Psychosom Med. 2011 May ; 73(4): 323–335. doi:10.1097/PSY.0b013e318218e1fb.

Usual and Unusual Care: Existing Practice Control Groups In Randomized Controlled Trials of Behavioral Interventions

Kenneth E. Freedland, Ph.D., **David C. Mohr, Ph.D.**, **Karina W. Davidson, Ph.D.**, and **Joseph E. Schwartz, Ph.D.**
Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri (K.E.F.); Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois (D.C.M.); Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York (K.W.D.); and Department of Psychiatry and Behavioral Science, Stony Brook University School of Medicine, Stony Brook, New York (J.E.S.)

Abstract

Objective—To examine the use of existing practice control groups in randomized controlled trials of behavioral interventions, and the role of extrinsic healthcare services in the design and conduct of behavioral trials.

Method—Selective qualitative review.

Results—Extrinsic healthcare services, also known as nonstudy care, have important but under-recognized effects on the design and conduct of behavioral trials. Usual care, treatment as usual, standard of care, and other existing practice control groups pose a variety of methodological and ethical challenges, but they play a vital role in behavioral intervention research.

Conclusion—This review highlights the need for a scientific consensus statement on control groups in behavioral trials.

Keywords

control groups; delivery of health care; evidence-based medicine/methods; placebos; randomized controlled trials as topic/methods; research design

INTRODUCTION

This article focuses on control groups that are used to contrast behavioral interventions with existing treatments or healthcare practices. It also examines the effects of the extrinsic healthcare environment on other aspects of randomized controlled trials (RCTs) of behavioral interventions.

Historical Background

Across a vast assortment of disciplines and research domains, all clinical trialists are bound by the same core methodological and ethical principles. However, distinctive traditions and issues have emerged in different disciplines and research domains. Experimental psychology

Address correspondence to Kenneth E. Freedland, Ph.D., Department of Psychiatry, Washington University School of Medicine, 4320 Forest Park Avenue, Suite 301, St. Louis, Missouri 63108 USA. freedlak@bmc.wustl.edu.

provided the methodological foundation for many early trials of behavioral interventions, and it still exerts a profound influence on the design and conduct of behavioral RCTs. Rigorous behavioral experiments were supposed to be conducted in tightly controlled laboratories in which it could be guaranteed that an experimental manipulation was the only difference between groups of subjects formed by random allocation. When it was discovered that supposedly objective scientists could unintentionally bias the outcomes of their experiments (1), psychological laboratory research methods became even more fastidious.

Because they emulated the methods of experimental psychology, many early clinical behaviorists assumed that intervention research could and should be conducted in laboratory environments or institutional settings that were well insulated from external influences. This assumption pervaded many articles and textbooks on behavioral research methods (e.g., 2), and it provided the principal rationale for laboratory analogue studies of behavior therapy. (3) Early trials of systematic desensitization, contingency management, and other behavioral interventions adhered as closely as possible to the ideals of laboratory research.(e.g., 4, 5)

A gradual shift away from laboratory-based behavioral intervention research began in the late 1960's, for several reasons. First, the generalizability of laboratory analogue studies was questioned.(6) Second, Tharp and Wetzel's influential book (7) argued that behavioral interventions should be conducted in the recipient's natural environment, to make the best use of social resources and reinforcement contingencies. Finally, ecological validity became a more salient consideration when behavioral researchers emerged from their laboratories to investigate *in vivo* exposure and other treatments that *had* to be administered in the natural environment.(8)

Psychotherapy researchers followed a different path to the same destination. Early psychotherapy research consisted mostly of case studies, surveys, and observational studies. It focused on psychotherapy as practiced in actual clinical settings, but it had serious methodological limitations. This led some to conclude that laboratory analogue studies could provide more rigorous evidence (9), but the need for greater ecological validity ultimately prevailed. The most influential psychotherapy studies have been *bona fide* treatment trials conducted in clinical settings (e.g., 10), not analogue experiments conducted in laboratories.

However, the world outside the laboratory has turned out to be a rather complicated place in which to conduct intervention research. Contemporary behavioral RCTs are embedded in an elaborate matrix of healthcare environments and socioeconomic conditions. The medical and mental health services that exist in the settings in which we conduct our trials and in the communities from which we recruit our participants influence the kinds of control conditions we can or cannot employ, and they affect our trials in many other ways as well. Inescapable tradeoffs often have to be made between internal and external validity, and it is not always possible to exercise tight control over every conceivable threat to either kind of validity.(11) These challenges have grown along with the complexity of the healthcare environments in which contemporary trials are conducted, but awareness of this aspect of behavioral RCT methodology has not kept pace with these developments.

The Traditional Control Group Hierarchy

Many methodology texts (e.g., 12, 13, 14) present a hierarchy of traditional control conditions that differ in stringency, i.e., in the magnitude of the impact that exposure to the condition is designed or expected to have on the trial's outcome variable(s). The least stringent is a no-treatment control arm; the most is an active alternative treatment; and wait list, attention, placebo, and treatment component control groups are in between. This hierarchy has considerable heuristic value, but it is deficient in two respects.

First, it obscures the distinction between the *control* and *comparison* functions of control groups.(12) The primary purpose of most control groups is to control for threats to internal validity (15), but active comparators also make it possible to determine whether an experimental intervention is superior or inferior to some other form of treatment. Threats to internal validity can often be controlled with conditions that are less stringent than active comparators. All else being equal, it is more difficult to detect a treatment effect if an intervention is compared to an active comparator than to a less stringent control condition.(16) Because effect sizes are inversely related to control group stringency, trials with relatively stringent controls require larger samples than trials comparing the same intervention a less stringent control.(17) Consequently, researchers tend to use active comparators only when comparison is a primary aim or external validity is a salient consideration, such as in comparative effectiveness research.(18)

Second, the traditional hierarchy is based on the outmoded assumption that behavioral trials can be isolated from the healthcare environments in which they are conducted. For example, "no treatment" controls can only be used when it is possible to ensure that 1) the control group participants are not given the experimental treatment, 2) they cannot obtain it from any other provider, and 3) they cannot obtain any *other* treatment for the same problem. This seldom holds true in any RCT in which the target is a physical or mental health problem and the participants have access to nonstudy healthcare services, even if the services are rudimentary.

The availability of nonstudy care scrambles the traditional control group hierarchy by making it impossible to guarantee that ostensibly low stringency control conditions are actually and uniformly less stringent than more active or intensive ones. For example, some of the wait-listed participants in a behavioral intervention trial for insomnia may obtain nonstudy hypnotics, even if the research team tries to dissuade them from doing so. The higher the proportion of participants who take nonstudy medications, the less closely the control condition approximates the textbook ideal of a wait list and the more it resembles an active intervention.

Existing-Practice (EP) Comparison Groups

Existing practice control conditions are used to compare experimental interventions to existing treatments or clinical practices. These conditions play an important role in medical trials, and there is a growing literature on the challenges they pose.(19–25) They are also playing an increasingly important role in behavioral trials, but little is said about them in most behavioral research methodology textbooks. Nevertheless, there is growing interest in the utility and pitfalls of EP control groups in behavioral research.(16, 26–32) The characteristics of a variety of EP controls are summarized in Table 1.

Treatment as usual (TAU) control groups are used to compare experimental interventions to treatments that are already used in clinical practice. The term "treatment as usual" seems to imply that most patients with the target problem ordinarily receive a particular treatment, but this is not always the case. "TAU" is used much more often to label EP control groups in psychotherapy studies (33), mental health services research (26, 34), and behavioral intervention trials for substance abuse (27) than in other areas of medical research.

Usual care (UC) is a roughly equivalent term that is used much more often than TAU in medical trials and in behavioral medicine.(e.g., 35, 36, 37, 38) The term "usual care" does not imply that patients ordinarily receive a specific treatment for the target problem. Instead, it suggests a broader range of possibilities. For example, usual care for hypertension may encompass a wide variety of drugs, nonpharmacological interventions, and clinical monitoring. Like TAU, however, the term "usual care" does not always align with its

broader connotations; sometimes, it actually does refer to a specific treatment.(19) Equally ambiguous synonyms such as standard care (SC)(39, 40) or routine care (RC)(41, 42) are occasionally used in trial reports instead of UC.

UC is a *laissez faire* condition in which healthcare providers who are independent of the research team determine the participants' care, but there are variants on UC that *are* influenced by the research team. In an enhanced usual care (EUC) condition, usual care is systematically improved by the research protocol to overcome ethical or methodological problems that would accompany ordinary UC. The degree of enhancement can range from minimal to extensive. The more extensively UC is enhanced, the less it approximates the routine care that the participants would ordinarily receive if they were not involved in an RCT, and the more it resembles an experimental intervention in its own right. EUC also stretches the meaning of "nonstudy" care. To the extent that a trial protocol modifies any aspect of the patient's healthcare other than the specific treatment(s) being tested, "nonstudy" care merges with "study" care. This applies not only to control groups, but to the intervention arms of RCTs as well.

The Enhancing Recovery In Coronary Heart Disease (ENRICHD) trial provides an example of a minimally enhanced usual care (mEUC) control group. All participants were given a patient education booklet, and their personal physicians were informed if major depression was diagnosed at baseline.(43) These enhancements addressed a potential confound (between-group differences in secondary prevention education) and an ethical problem (identifying but not treating major depression in the control group). The Pro-B Type Natriuretic Peptide Outpatient Tailored Chronic Heart Failure Therapy (PROTECT) Study provides an example an extensively enhanced usual care (eEUC) arm. All participants were given guideline-adherent heart failure care. In some cases, this differed from the care they received before enrollment in the trial.(44)

A second variant could be called constrained usual care (CUC), in which nonstudy care is restricted in some way. For example, benzodiazepines may interfere with behavioral interventions for anxiety disorders.(45, 46) Consequently, participants in a trial of a behavioral treatment for panic disorder may be required to refrain from nonstudy benzodiazepines.

A third variant is the standardized treatment regimen (STR), in which the same clinical care or treatment(s) are administered in the same way to all participants. Standardization does not necessarily mean that every patient receives identical treatment. Instead, each patient may be treated according to a standardized protocol or care path.(22, 47, 48) STR presumably differs from EUC and CUC in the extent to which treatment is homogenized. "Constraining" care suggests that a particular ingredient of usual care is withheld, whereas "standardizing" care implies more extensive calibration. Again, however, these sorts of terms are not always used consistently or precisely.

In standard of care (SOC) control groups, participants receive state-of-the-art, evidence-based, guideline-adherent clinical care. SOC is a naturalistic condition when patients are recruited from settings that provide it routinely (e.g., specialty clinics at leading academic medical centers). SOC may have to be imposed by enhancement of usual care when patients are recruited from less stellar settings. If care is extensively enhanced or standardized in order to provide the current standard of care, then eEUC, STR, and SOC may be indistinguishable.

A uniform or protocol-driven standard of care (uSOC) produces the best clinical outcomes for some conditions, but an individualized standard of care (iSOC) is best for others. Also, some best practices are supported by strong evidence but others are not, and thus leave more

room for expert judgment in clinical decision-making. Consequently, SOC control patients are treated in a very uniform fashion in some trials but not in others.(19, 22)

The most problematic EP control condition might be called inadequate care (IC), reflecting the inferior healthcare services to which underserved, uninsured, or captive patient populations may be relegated. Whether it is ethically justifiable to compare a behavioral intervention to IC depends on its translational potential.(49, 50) Hypothesized superiority to IC, by itself, is an insufficient justification. The intervention should also be designed to help close the gap between the target population's inadequate care and that which is available to more fortunate populations, and it should be potentially feasible to implement. Recent studies of diabetes self-management (51) and multifactorial cardiovascular disease risk reduction interventions (52) are examples of trials designed to improve upon IC for underserved populations..

Two or more kinds of EP control conditions may be combined to produce a hybrid. For example, UC may be enhanced in some ways but constrained in others. This could occur, for instance, in a weight management trial in which UC is enhanced with health education materials but constrained by prohibiting concurrent participation in other weight loss programs.

Stringency and Uniformity in Existing Practice Control Groups

Like the traditional control groups, a stringency dimension underlies EP control conditions. IC is the least stringent, and SOC is the most. To the extent that ordinary UC falls short of the current standard of care, it is also less stringent than SOC.

Traditional control conditions are usually designed to be as uniform as possible. In a typical placebo control group, for example, all participants are given the same placebo in exactly the same manner. Some EP control conditions, such as STR, are also very uniform, but others are not. There are several common sources of variability in EP controls. First, nonstudy care may differ across participants because they receive healthcare services from different providers, are covered by different insurance plans, or have different medical or psychiatric problems.

Second, the quality of nonstudy care may differ across recruitment sites. For example, a behavioral intervention trial for patients with heart disease might draw some participants from a renowned specialty clinic at a major academic medical center and others from the surrounding community. The quality of nonstudy care in these subgroups may be quite different.

Third, routine practices and standards of care can change over time, as when widely-used drugs fall into disfavor or new discoveries transform routine care in unanticipated ways. The longer it takes to complete an RCT, the more likely it is that such developments will introduce variability into EP control conditions.

Due to these sources of within-group variability, an RCT may be described as having a certain kind of EP control condition, such as "usual care", yet this descriptor can have very different meanings for different participants. Also, it is common for hybrid control groups, such as ones in which usual care is both constrained and enhanced, to simply be called UC or TAU. Consequently, the labels that are affixed to EP control groups often oversimplify the actual conditions to which the participants were exposed. In addition, if two trials have "usual care" control groups, there may be little resemblance between the groups other than their names.

Stringency and uniformity are nearly orthogonal dimensions of EP control groups. The least stringent IC control may be just as uniform in its paucity of care as the most stringent SOC control group is in its excellence. Similarly, highly personalized and highly standardized SOC control groups may be equally stringent. Whereas control group stringency tends to reduce between-group differences in outcomes, control group uniformity may decrease within-group variability and thereby increase between-group differences in outcomes. Consequently, both dimensions can affect statistical power.

Standards of Care

A number of issues concerning standards of care can impinge on behavioral RCTs. First, the treatment of a particular condition may be addressed in multiple clinical practice guidelines, issued by different disciplines involved in the treatment of the same condition, for care provided in different clinical settings or countries, or for care financed by different third-party payers.(e.g., 53, 54, 55, 56, 57, 58) Consequently, operative standards of care may depend on where a trial is conducted. In multicenter trials, standards of care may differ across centers.

Second, in jurisdictions that operate under what is known as the locality rule, physicians and other healthcare providers risk being held liable for malpractice if they adhere to national standards of care rather than to local community standards.(59, 60) Most states now regard the locality rule as anachronistic and have adopted national standards, but it is still in effect in some states. It may be difficult to conduct trials with SOC control groups in these states, with one notable exception: Federal facilities, such as Veterans Administration medical centers, are required to adhere to their own national standards regardless of local standards. Thus, behavioral trials conducted entirely within such facilities may be able to utilize true SOC control groups.

Third, controversies about standards of care have arisen when RCTs have been conducted in resource-poor countries. Some argue that it is unethical to conduct RCTs where local standards of care fall far short of that which is widely available in wealthier countries.(61) Others hold that prohibiting such trials would prevent tests of the very interventions that are needed to improve treatment and prevention services in resource-poor regions.(62) In practice, whether trials incorporate local levels of care or higher standards tends to differ across conditions. For example, RCTs conducted in sub-Saharan Africa to test treatments for tuberculosis tend to provide a high standard of tuberculosis care to control participants, while trials targeting AIDS usually do not provide state-of-the-art HIV care to control participants.(63)

Healthcare disparities within the United States (e.g., 64, 65) have raised concerns about racial and socioeconomic disparities in clinical trial enrollment (66, 67) and about the bioethics of recruiting medically underserved groups to participate in RCTs.(68) Much less attention has been paid to racial, ethnic, gender, and socioeconomic disparities in the nonstudy care that participants receive once they have been enrolled in clinical trials. Nevertheless, such disparities undoubtedly contribute to variability in EP control groups.

Ubiquity of Nonstudy Care in Behavioral Trials

Nonstudy care is ubiquitous in the clinical research environments in which contemporary behavioral trials are conducted. Its availability, quality, and effectiveness influence the self-selection of patients into clinical trials, as well as the adherence and retention of participants after trial enrollment.(69, 70) Nonstudy care is present in almost any control group that might be employed in a behavioral trial, including both EP and traditional control conditions. It is also present in most behavioral intervention arms as well. When participants

in the intervention arm of a trial receive nonstudy care, it is often referred to as concomitant therapy or co-intervention.(71, 72)

Nonstudy care is often overlooked when trial designs are described. For example, the design of a trial comparing cognitive-behavioral pain management (CBPM) to progressive muscle relaxation (PMR) would typically be presented as a comparison of CBPM vs. PMR. However, a more accurate description would be (CBPM + UC) vs. (PMR + UC), because all participants in the trial have access to nonstudy care. A report by Donta et al. (73) on a trial of CBT and aerobic exercise provides an example of this approach.

The presence of nonstudy care in both arms of a trial does not necessarily mean that their effects completely cancel each other out. Nonstudy alternatives may subtly influence the delivery or acceptability of an experimental intervention, or modify its effects. Hypothetically, if no other treatments for pain besides CBPM had ever been developed and patients had no better recourse other than to try this experimental intervention, they would be highly motivated to engage in the therapy and to adhere to the protocol. Of course, other treatments for pain *have* been developed, and their availability through nonstudy channels could affect the participants' motivation to engage in CBPM. If PMR is a less burdensome treatment, then it might be less permeable to this effect. This could create a subtle confound that would be difficult to prevent and to detect. On the other hand, this effect would strengthen the external validity of the findings, since CBPM and PMR can only be implemented in a world in which other treatments for pain already exist.

Unblinded practitioners can also influence behavioral trial outcomes.(74) For example, in a trial of cognitive behavior therapy (CBT) for depression, nonstudy physicians may be informed about the group to which their patients have been assigned. They might be more likely to prescribe nonstudy antidepressants to patients in the control group than to those in the CBT arm.

Behavioral trialists know that nonstudy treatments for the primary target of their experimental intervention can be problematic, both in intervention and in control arms. Thus, if testing an intervention for anxiety, most investigators would be vigilant about nonstudy treatment for anxiety. It is less widely understood that *any* form of nonstudy care that can affect any of a trial's outcomes, should be of concern. For example, a trial might be conducted to test a novel intervention for anxiety in patients with breast cancer. Obviously, nonstudy treatments for anxiety could affect the trial's anxiety outcomes, but the nonstudy treatment of cancer might also influence anxiety outcomes. Nonstudy treatments for medical and psychiatric comorbidities could also affect anxiety. Furthermore, all of these aspects of nonstudy care could affect secondary outcomes such as health-related quality of life.

If all aspects of a patient's nonstudy care are potentially relevant, then it is necessary to reconsider the meaning of "standard of care". If this anxiety trial were conducted at a leading center for cancer research, it would be safe to assume that the participants received state-of-the-art, nonstudy cancer care. It would not necessarily follow that all of them also received state-of-the-art, nonstudy care for anxiety, or for comorbidities such as diabetes that might affect their anxiety or their health-related quality of life. A behavioral trial probably should not be described as having an SOC control group unless the participants received state-of-the-art care for the primary target of the intervention, which in this case would be anxiety rather than cancer. However, if a medical variable such as time to cancer recurrence or death were either a primary or a major secondary outcome of the trial, the quality of nonstudy care both for anxiety and for cancer would be germane to the question of what to call the control group.

The broader relevance of nonstudy care also raises questions about how other EP control conditions are defined in behavioral trials. For example, a recent article (27) described four design options for effectiveness trials of behavioral interventions for substance abuse: 1) new intervention vs. TAU; 2) new intervention + TAU vs. TAU; 3) new intervention + TAU vs. alternative intervention + TAU; and 4) new intervention vs. standardized control treatment. This is a very useful framework but it assumes that an identifiable TAU for substance abuse is the only relevant aspect of nonstudy care. If one instead assumes that all aspects of nonstudy care are potentially relevant, design #1 expands to (new intervention + UC) vs. (TAU + UC), where TAU is whatever treatment is usually provided for substance abuse in the community setting of interest, and UC covers all other aspects of nonstudy care, such as for HIV, hepatitis C, smoking, or other health problems. The same logic applies to the other designs.

Confounding and Mediation By Nonstudy Care

Enhancement and other systematic modifications of a control group's nonstudy care can confound a trial's experimental variable unless comparable changes are made in the intervention arm. However, this problem can be prevented at the design stage of an RCT. For example, in the Heart Failure Adherence and Retention Trial (HART), there was identical educational enhancement of nonstudy care in the intervention and control arms. (75)

Behavioral trials can also be affected by post-randomization confounding and mediation. Different arms can have different effects on the nonstudy care environment, resulting in *differential intensification* or *differential abatement* of nonstudy care. If these are unintentional effects, they constitute a form of post-randomization confounding that can bias the trial's outcomes. If they are deliberate, they are potential mediators of the effects of the intervention.

Unintended differential intensification could occur, for example, in an RCT of a coping skills intervention vs. UC for cancer survivors. Intervention participants may tell their therapists about various signs, symptoms, and other health-related problems, whether or not they are asked to do so. When therapists hear about new or worsening medical or psychiatric problems, or ones that have not been adequately assessed or treated, they may advise participants to call their nonstudy physician or take other appropriate actions. In emergencies, they may take more active measures such as contacting the physician or calling 911. These actions may intensify nonstudy care by promoting nonstudy physician visits, hospital admissions, or more aggressive nonstudy medical treatment regimens, even though the intervention is not designed or intended to produce such effects. UC participants do not have therapists, so there are no comparable effects on their nonstudy care. If this pattern of differential intensification occurs, the apparent effects of the intervention could actually be due, at least in part, to nonstudy care.

The converse problem may occur if assessment staffers are not blinded to group assignments and are not in equipoise about the treatment and control conditions. In this circumstance, they may encourage more nonstudy care in the UC than in the intervention arm, out of concern that control group members are not getting any therapeutic help from the trial. This could work against the trial's hypothesis and increase the risk of a Type II error. Efforts to blind the research staff and to instill therapeutic equipoise can help to counteract this potential source of bias.

A recent trial of two different treatments for opioid addiction in adolescents provides an example of a study in which possible confounding by nonstudy care was examined. Woody and colleagues (76) found that significantly *fewer* participants in their extended

buprenorphine-nalaxone plus counseling arm obtained nonstudy addiction care during the trial than did the participants in the standard detoxification plus counseling arm. This effectively ruled out unintentional differential intensification of nonstudy addiction treatment as a rival explanation for the superior outcomes observed in the experimental treatment group.

This is a relatively rare example; few other behavioral RCT trial reports have included secondary analyses that might have revealed or ruled out post-randomization confounding by differential intensification of nonstudy care, even when differences in utilization of nonstudy care have been identified.(e.g., 77) Consequently, the prevalence of this potential threat to the internal validity of behavioral RCTs is unknown. It might be common, or it might not be. The only way to find out is for behavioral trialists to start collecting the requisite data.

Differential intensification of nonstudy care is a deliberate strategy in some behavioral trials. For example, in the Matching Alcoholism Treatments to Client Heterogeneity (Project MATCH) trial, participants were randomly assigned to CBT, motivational enhancement therapy (MET), or Twelve-Step Facilitation (TSF). The aim of TSF was to increase participation in Alcoholics Anonymous (AA), a type of non-study intervention. As intended, AA involvement was significantly higher in the TSF than in the other two arms. There were no significant between-group differences in any other form of nonstudy care.(78) In this example, nonstudy intervention (AA) was designed to *mediate* rather than confound the relationship between the TSF intervention and alcoholism outcomes.

Deliberate intensification resembles deliberate enhancement of nonstudy care, but they are not synonymous. Deliberate intensification is implemented only in the treatment arms of trials of disease management, collaborative care, and other interventions that work in conjunction with nonstudy care. In contrast, deliberate enhancements of nonstudy care are implemented in both arms of most EUC-controlled trials. Also, deliberate intensification is an experimental procedure, and the principle of equipoise requires uncertainty as to whether the intervention will actually succeed in intensifying nonstudy care or whether doing so will improve clinical outcomes. In contrast, deliberate enhancement is designed to improve nonstudy care in ways that will make it a more suitable comparator than ordinary UC or IC, and there should be little uncertainty as to whether the enhancements will be both feasible and beneficial.

Differential *abatement* of nonstudy care can also be a deliberate strategy. For example, Allen et al. (79) randomly assigned patients with somatization disorder to standard medical care augmented by a psychiatric consultation intervention or to a manualized cognitive-behavioral treatment program that was added to the psychiatric consultation intervention and designed to decrease somatization and healthcare utilization. Compared to the control group, the CBT group had significantly greater improvements in self-reported functioning and somatic symptoms and a greater decrease in nonstudy healthcare service utilization.

It may not be possible to completely prevent post-randomization confounding, and it would not be desirable to prevent intentional mediation by nonstudy care. Thus, it is necessary to document and analyze any elements of nonstudy care that could affect the trial's outcomes. Unfortunately, there is little guidance on this subject in the behavioral RCT literature, and few examples to follow. Much of the literature on treatment intensification has focused on clinical care, not on clinical trials. For example, there have been several studies of clinical inertia, which is defined as the failure of physicians or other healthcare providers to initiate or intensify therapy when indicated, e.g., when a diabetes patient's glycemic control starts to deteriorate, or when a first-line antidepressant fails to produce remission.(80–83) Other

studies have evaluated various methods for measuring treatment intensification in clinical practice.(84) A number of recent collaborative care studies (e.g., 36, 85) have included basic measures and analyses of nonstudy treatment intensification, because it is one of the explicit aims of the intervention.

In behavioral trials, any recommendations, referrals, or other actions taken by study staff that could impinge on nonstudy care should be systematically documented. This would include, for example, occasions when research staffers ask participants questions such as, "Have you told your doctor about that?", give more directive advice such as, "You should call your doctor right away about that," or intervene directly, e.g., "I'm calling 911." The documentation should classify the action (recommendation, referral, or direct intervention) and the agent (e.g., interventionist, research staff, or investigator). However, it can be hard to link such actions to actual changes in nonstudy care. For example, if a counselor advises a patient to talk to his doctor about a new symptom and the patient does so, an outpatient visit may or may not follow, and it might take place long after the advice was given. It might be difficult to connect these dots, but in some trials, it might be worth the effort to do so.

Differences between groups in the *quality* of nonstudy care can also pose a threat to internal validity. Published, evidence-based clinical guidelines can be converted into checklists and completed to assess key aspects of the quality of nonstudy care received by participants in both the intervention and control arms of an RCT. For example, a checklist based on the American Heart Association's guidelines for the management of chronic, stable angina (86) could be developed to assess the quality of nonstudy cardiac care in a trial of stress management for patients with angina pectoris. It is likely to be impractical to comprehensively assess every available quality indicator that might be relevant to a trial, so checklists should probably be limited to ones that are based on strong evidence from multiple RCTs or rigorous meta-analyses (e.g., Class Ia/Ib recommendations). The checklist data can be used to determine whether groups differ in the receipt of guideline-adherent nonstudy care during the trial. The data can also be used to characterize the quality of care received by the study sample as a whole. Consistent collection of such data would enable meta-analysts to evaluate whether the quality of nonstudy care moderates the effects of behavioral interventions.

Choosing Control Conditions for Behavioral Trials

Given the challenges that EP control groups can pose, one might wonder why they should ever be used in behavioral RCTs. However, they are indispensible in randomized effectiveness trials that evaluate whether new interventions can replace or augment existing practices.(27) They also have an important place in behavioral efficacy research. Indeed, the omnipresence of nonstudy care blurs the distinction between efficacy and effectiveness trials. Behavioral efficacy and effectiveness studies generally do not differ as to whether they are conducted in "real world" settings, but instead in the *kinds* of real world settings where they take place. Whether a clinical trial is conducted at a university teaching hospital or at a rural community health clinic, it involves real patients, real health problems, and real nonstudy healthcare services.

Efficacy and effectiveness trials do tend to differ as to whether the primary goal of the research is *explanatory* (e.g., to advance understanding about an intervention or to determine whether an experimental treatment and a comparator produce different outcomes) or *pragmatic* (e.g., to support policy decisions about which treatments to use in clinical practice).(87) EP control conditions play an important role not only in pragmatic research, but in explanatory trials as well. The pervasive presence of nonstudy care often makes it imperative to utilize EP controls in behavioral efficacy studies. It can also eliminate the option of using traditional control groups, at least in pure form. For example, it may not be

possible to expose control participants *only* to a placebo, to the exclusion of nonstudy care. Furthermore, existing practices are key determinants of whether it is considered ethical to employ traditional controls. Ethical concerns about placebo controls typically do not revolve around any harm that the placebo itself might cause, but focus instead on the risks of depriving patients of established therapies.(88)

As another example, a treatment development trial testing whether a novel intervention has any effect at all cannot employ a pure no-treatment control condition if nonstudy care is available. The closest approximation may be a UC control group in which participants receive no additional intervention besides their usual nonstudy care. Constrained usual care may be another alternative in some instances. The PREMIER trial, for example, tested the efficacy of lifestyle behavior and dietary interventions in individuals with prehypertension or stage 1 hypertension. The participants were not taking antihypertensive medications at enrollment and they were not given such medications during the intervention phase.(89, 90)

Control group stringency generally increases as research progresses from early treatment development work to definitive efficacy trials.(16) This holds not only for traditional control groups but for EP controls as well. For example, an intervention might be compared initially to ordinary UC. Even if the results are positive, they may not change clinical practice because the UC participants did not receive state-of-the-art therapy. If the next trial demonstrates superiority to a more rigorous SOC control, it could have a greater impact.

On the other hand, some trials are not undertaken with the goal of surpassing the current SOC or a well-established treatment. Their aim is instead to demonstrate the efficacy of a more practical or cost-effective alternative to an existing intervention or SOC that is too expensive, intensive, or inaccessible for the target population.(91, 92) For example, telephone-delivered interventions have been tested for depressed patients who are too medically ill, immobilized, time challenged, or distant from standard clinical services to be able to attend psychotherapy sessions in person. Similarly, internet-based lifestyle behavior interventions have been tested as alternatives to direct clinical services.(93, 94) Is the appropriate comparison for such interventions a nationally-established SOC to which the participants lack routine access? Or is it the local UC or IC that represents their *de facto* standard of care? The answer is debatable, like it has been in other controversies that have arisen where there are marked disparities in nonstudy healthcare services.

Explanatory trials designed to test biobehavioral mediation hypotheses are another exception to the rule that advanced stages of research should employ relatively stringent controls. Such trials aim to determine whether treatment of a behavioral or psychosocial problem improves medical outcomes. The goal is not to test the practical efficacy of the intervention but rather to test the hypothesis that the behavioral or psychosocial problem is a causal risk factor, not merely a risk marker, for the medical outcome. The success of the trial depends on producing the largest possible between-group difference in the behavioral or psychosocial outcome. This makes it highly desirable to choose an intervention with established efficacy for changing the behavioral or psychosocial outcome, along with the least stringent control that is possible and ethical to employ. Given the aims of these studies, it would be counterproductive to use more stringent control conditions or to design them as if they were pragmatic efficacy trials.

For example, ENRICHD was an explanatory trial designed to determine whether treatment of depression and inadequate social support decreases the risk of myocardial infarction or death in patients with coronary disease.(43) It was not a pragmatic trial of the efficacy of CBT for depression or inadequate social support. Its mEUC control group was the least stringent comparator that was both feasible and ethical to use. It would have been

counterproductive to employ a more stringent eEUC or SOC control, or an attention or placebo control group.

Equipoise and Existing Practice Controls

The bioethical principle of equipoise has been vigorously debated (95–97), but there is consensus that RCTs should be conducted only when there is significant uncertainty about what will be found. Researchers typically initiate a superiority trial because they believe that an experimental intervention *may* be superior to some alternative, but they should not do so if its superiority is almost inevitable, as this would violate the principle of equipoise. Efforts to adhere to this principle can also help to reduce experimenter bias in RCTs that cannot be double-blinded, in which there is always a risk that the deck will somehow be stacked, even if inadvertently, in favor of the experimental intervention. This is just as important a concern in EP-controlled trials as it is when traditional control conditions are employed in behavioral RCTs.

Research in England on community-based interventions for severely mentally ill individuals provides a cautionary lesson. The existing practice in the U.K. has long been for care to be provided by community mental health teams (CMHTs). An alternative approach known as Assertive Community Treatment (ACT) was tested in a series of studies, most recently by Killaspy et al.(98) ACT is more intensive and expensive than CMHT, but favorable findings in early trials led to substantial public investment in ACT implementation. Subsequent research found that ACT produces no better outcomes than CMHT in terms of the most important indicators, particularly hospitalization rates. After reviewing this line of research, Burns (28) argued that researchers and policy-makers might have been able to reach this conclusion many years sooner if CMHT had been viewed all along as an active comparator, rather than being framed as a presumptively inferior “treatment-as-usual” control condition. Labeling these trials as randomized comparisons of ACT vs. TAU, and failing to specify the latter condition as rigorously as the former, may have conveyed an assumption that ACT was the superior approach, thereby creating an implicit bias. Framing them instead as randomized trials of ACT vs. CMHT might have fostered greater neutrality and minimized inadvertent biases in favor of ACT.

Attention, Expectancy, and Existing Practice Controls

A common concern about EP control conditions is that they do not eliminate differential attention and expectancies as threats to internal validity.(99, 100) This issue deserves careful scrutiny. If the aim of a trial is to determine whether an intervention is superior to existing practices, then it has to be compared to those practices. Controlling for attention entails exposing control group participants to some sort of clinical experience that takes as much time and provides as much contact with a therapist as the intervention. Thus, a three-arm trial ([intervention + UC] vs. [attention + UC] vs. [UC alone]) would be needed to achieve the trial’s primary aim while also controlling for attention and expectancy. Three arms are more expensive than two, and they expose more patients to experimentation. The question is whether the attention arm is necessary.

In a drug efficacy trial, the aim is to determine the effects of a specific chemical compound. The most important rival explanations are the placebo effect and clinical attention. Neither is plausible for some drugs (e.g., general anesthetics), but both are quite plausible for others (e.g., antidepressants).(101) To control for these twin threats to internal validity, patients are randomly assigned in double-blind fashion to drug or placebo. The groups are given equal attention, typically in the form of clinical management.(102) Thus, clinical management is co-administered with medication or placebo. The Sertraline Antidepressant Heart Attack Randomized Trial (SADHART) is an example of this type of trial.(103)

In many psychotherapy efficacy studies, the primary aim is to determine whether an intervention's putative active ingredients yield incremental benefits over those of its nonspecific ingredients, which are also known as common factors. Attention is a nonspecific ingredient of any psychotherapeutic or behavioral intervention that involves human therapists, but it cannot be distilled and delivered in pure form by having therapists sit and stare at control group participants for the same amount of time that treatment group therapists spend with theirs. Clinical attention inevitably must be wrapped in a credible intervention. Thus, attention control groups in psychotherapy trials require two ingredients, i.e., clinical attention and an intervention that induces an expectation of therapeutic benefits, much like drug trial control groups combine clinical attention with pill placebos.

For example, a psychotherapy trial comparing two interventions to a supportive therapy control group was cited in a recent methodology textbook as an excellent example of an attention/placebo control design.(13) The supportive therapy did control for common factors, but it also helped patients express their feelings and identify options for addressing personal problems.(104) Clearly, it provided something more than content-free clinical attention.

The aim of some behavioral trials is to compare the efficacy of different dosages of the same therapy. For example, Edinger et al. (105) conducted an RCT to compare CBT for insomnia, delivered in 1, 2, 4, or 8 sessions over an 8-week period. In such trials, the amount of clinical attention is the experimental variable and it differs between groups by design. In other trials, the aim to compare the efficacy or effectiveness of two (or more) different treatments for the same problem. There is often an effort in such trials to hold attention constant across groups, and to statistically adjust if necessary for differential attention, so that differences in outcomes can be attributed to differences in the putative active ingredients of the interventions. In a recent trial, for example, Safran et al. (106) compared CBT to relaxation with educational support for residual symptoms of attention-deficit/ hyperactivity disorder (ADHD) in medication-treated adults. Both groups received 12 weekly, 50-minute sessions.

In many trials, however, differentiating between specific and nonspecific ingredients is not a priority. If the goal is to push the envelope of efficacy and achieve better outcomes by whatever means seem most promising, rather than to disaggregate a particular intervention, it may be unnecessary to control for attention but a UC or SOC control group may be essential. The same intervention can be studied in different trials for different reasons, and control conditions that are optimal for some of these trials may be unsuitable for others. The paramount consideration in choosing a control condition is how well it fits the purpose of the trial.(107)

The assumption that it is *always* necessary to control for attention in psychotherapy or behavioral intervention trials rests on the questionable premise that attention is always a threat to internal validity. It is a genuine threat to the internal validity of standard drug trials because they are designed to answer questions about chemical compounds, not about their clinical delivery. Psychotherapy differs from pharmacotherapy in that clinical attention is an integral component of psychotherapy, not something that is co-administered with it. Attention is not a rival explanation for the effects of psychotherapy; it is an indispensible ingredient of psychotherapy. Since attention can only be delivered in the guise of a credible intervention, *comparison* rather than control is the primary function of an attention control group in a behavioral RCT.

If an RCT shows that a novel intervention is superior to UC, critics might dismiss the findings because they could be partially due to an imbalance in the amount of clinical

attention that was given to the two groups. In response, one might ask, “So what? Why does that matter?” The intervention cannot be delivered without attention, so what difference does it make how much of its effect is due to attention? A more important question is whether comparable clinical outcomes can be achieved with a less expensive or intensive intervention, like the kind that might be employed in an “attention control” group. That, however, is a comparative effectiveness study for another day, and there may never be a need for it unless the intervention is first shown to be superior to usual care.

A landmark depression trial by DeRubeis and colleagues (108, 109) provides an example of a well-designed RCT in which controlling for attention was not a high priority. The primary purpose of the trial was to compare cognitive therapy to antidepressant therapy. A placebo control group was included because the comparison between the two active arms would have been difficult to interpret unless it was confirmed that the antidepressant was superior to placebo.(110) Both the placebo and the active medication participants received less clinical attention than the cognitive therapy participants. This was an appropriate design decision, because the medication and placebo arms received as much clinical attention as was consistent with the current “best practices” for antidepressant pharmacotherapy, and because the main purpose of the study was to compare the efficacy of the active interventions, not to parse their specific and nonspecific ingredients.

Participants’ outcome expectations play an important role in clinical trials. If two groups are assigned to different treatments but have identical outcome expectations, then any differences in outcomes are not attributable to differential expectations. This is the foundation of the double-blind, placebo-controlled RCT design. With rare exceptions, behavioral trials cannot be double-blinded, so they have to be conducted without the protection against differential expectancies that double blinding is designed to provide. If two groups in a behavioral RCT differ both with respect to a treatment variable (e.g., one group receives counseling and the other receives usual care) *and* to outcome expectations, then differential expectations confound the treatment variable and can complicate the interpretation of the trial’s results.

Few behavioral RCT reports have included data on participant expectations, and even fewer have systematically investigated differential expectations as a rival explanation for the trial’s outcome. There has been more research on the effects of treatment preferences in behavioral RCTs. Treatment preferences are strongly influenced by expectations of benefit, although there are other determinants as well, such as concerns about potential risks, side effects, and exposure to novel or experimental interventions.(111–113) Some studies have found little or no difference in outcomes between groups that are randomly assigned to treatments and ones that are assigned according to patient preference, or between groups in which treatment assignments are matched or mismatched to patient preferences.(112, 114–117) However, others have shown that outcome expectations and treatment preferences can have significant effects on trial outcomes. In one study, for example, primary care patients with major depression who chose counseling had better outcomes than those who were randomized to it (118, 119), and the outcomes in several RCTs have differed between participants who were matched to their preferred treatment and those who were mismatched.(119–122) Despite these inconsistent findings, it is safe to conclude that expectancy and preference effects should be taken into account in planning and conducting behavioral RCTs, including but not limited to ones that employ EP control conditions.

However, controlling patient expectations is easier said than done. This is evident, for example, in the highly variable placebo response rates that dozens of antidepressant trials have yielded among patients with major depression (101), in controversies that have arisen concerning the magnitude of placebo effects in many other trials (123–127), and in the

surprising efficacy of "open label" placebos.(128) Participant expectations are influenced by multiple inputs, only some of which are determined by the study's design or procedures. For example, expectations of benefit depend on the number of active and/or placebo arms in double-blinded RCTs.(129–131) A variety of other inputs, such as the patient's past experiences with similar treatments, are not determined by anything that the trialist does or fails to do.(132)

Participants have idiosyncratic expectations about their nonstudy care, just as they do about whatever treatment (if any) a trial exposes them to. Individual differences in expectancies can arise even under highly standardized conditions in double-blinded comparisons of equally credible treatments (133), and variability in nonstudy care can widen these differences. No matter which type of control group one chooses, and no matter how diligently one tries to equilibrate expectancies, they can still differ within and between groups. Thus, although equivalent expectancies are desirable, they cannot be guaranteed and should not be required when planning behavioral trials.

Differential expectancies may arise when behavioral interventions are compared to existing practices, but this should not preclude the use of EP controls. Whatever steps can be taken to minimize these differences should be taken, as long as they are compatible with the trial's aims. Investigators should not assume that participants will necessarily prefer or expect greater benefit from an experimental intervention than they do from usual care. There are a variety of reasons why some participants' expectancies and preferences may not favor the treatment arm in a behavioral RCT. For example, some patients may have more confidence in their own personal physician's advice about managing behavioral or emotional problems than they do in an unfamiliar study therapist into whose care they are randomly assigned. Others may consent to participate in an RCT for altruistic reasons, or for remuneration; in either case, they may prefer usual care because it is less burdensome than participation in the treatment arm.

Questionnaires should be used in behavioral trials to systematically assess expectancies and treatment preferences. The resulting data should be used in secondary analyses to evaluate the role of expectancies and of matched or mismatched preferences in the trial's outcomes. (100, 112, 123) Routine inclusion of these analyses in behavioral RCT reports would strengthen the quality of the evidence that we generate.

Conclusions and Recommendations

Behavioral trialists have to weigh a variety of scientific, ethical, practical, and other considerations in choosing control groups for their RCTs. If nonstudy care is present in the clinical trial environment, it may not be possible to control for every conceivable threat to internal validity, at least not in a single trial. Consequently, it may be necessary to control for whichever threats are judged to be the most important, at the expense of remaining vulnerable to less important ones. It is not always possible to give equal weight to internal and external validity in designing a trial when nonstudy care is available. It is not always ethical to employ a traditional control group if doing so would deprive participants of treatments that have already been incorporated into routine clinical care. In short, difficult tradeoffs often have to be made in choosing a control or comparison group for a behavioral trial. Existing practices and nonstudy care are what make many of these tradeoffs so difficult. On the other hand, existing practice comparison groups such as UC or SOC play invaluable roles in behavioral trials. Some of the most important goals of behavioral intervention research could not be achieved without them.

This article has highlighted a number of issues that have received relatively little attention in the behavioral trial literature, but that are sources of uncertainty and disagreement among

investigators, reviewers, bioethicists, and behavioral research methodologists. The best way to address these issues would be to develop an evidence-based scientific consensus statement or guideline on control groups in behavioral RCTs. We recommend that a consensus task force be created that would adhere to current standards for evidence-based guidelines.(134) This would help to accelerate progress in numerous areas of behavioral research.

Acknowledgments

This work was supported, in part, by Grant 5R01HL091918 from the National Heart, Lung, and Blood Institute, National Institutes of Health (K.E.F.).

We thank Patrick J. Lustman, Ph.D and Robert M. Carney, PhD, Washington University School of Medicine, St. Louis, Missouri, and Nathan Cobb, M.D., Schroeder Institute, Washington, D.C., for their very helpful comments on preliminary versions of this work.

Glossary

CUC	constrained usual care
EP	existing practice
EUC	enhanced usual care
eEUC	extensively enhanced usual care
HIV	human immunodeficiency virus
IC	inadequate care
iSOC	individualized standard of care
mEUC	minimally enhanced usual care
PMR	progressive muscle relaxation
RC	routine care
RCT	randomized controlled trial
SC	standard care
SOC	standard of care
uSOC	uniform standard of care
STR	standardized treatment regimen
TAU	treatment as usual
UC	usual care

REFERENCES

1. Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton-Century-Crofts; 1966.
2. Paul GL. Strategy of outcome research in psychotherapy. J Consult Psychol. 1967; 31(2):109–118. [PubMed: 5342732]
3. Kazdin, AE. Therapy analogues and clinical trials in psychotherapy research. In: Hersen, M.; Michelson, L.; Bellack, AS., editors. Issues in psychotherapy research. New York: Plenum Press; 1984. p. 227-250.
4. Lang P, Lazovik A. Experimental desensitization of phobia. J Abn Soc Psychol. 1963; 66(6):519–525.

5. Ayllon T, Azrin NH. The measurement and reinforcement of behavior of psychotics. J Exp Anal Behav. 1965; 8(6):357–383. [PubMed: 5851397]
6. Bernstein DA, Paul GL. Some comments on therapy analogue research with small animal 'phobias.'. J Behav Ther Exp Psychiatry. 1971; 2(4):225–237.
7. Tharp, RG.; Wetzel, RJ. Behavior modification in the natural environment. New York: Academic Press; 1969.
8. Rachman S, Marks I, Hodgson R. The treatment of obsessive-compulsive neurotics by modelling and flooding in vivo. Behav Res Ther. 1973; 11(4):463–471. [PubMed: 4777642]
9. Hersen, M.; Michelson, L.; Bellack, AS. Historical overview. In: Hersen, M.; Michelson, L.; Bellack, AS., editors. Issues in psychotherapy research. New York: Plenum Press; 1984. p. 33-250.
10. Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, Glass DR, Pilkonis PA, Leber WR, Docherty JP. National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. Arch Gen Psychiatry. 1989; 46(11):971–982. [PubMed: 2684085]
11. Cook, TD.; Campbell, DT. Quasi-experimentation: design & analysis issues for field settings. Boston: Houghton Mifflin; 1979.
12. Kazdin, AE. Research design in clinical psychology. 4th ed.. Boston: Allyn and Bacon; 2003.
13. Nock, MK.; Janis, IB.; Wedig, MM. Research designs. In: Nezu, AM.; Nezu, CM., editors. Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychosocial interventions. New York, NY US: Oxford University Press; 2008. p. 201-218.
14. Kendall, PC.; Holmbeck, G.; Verduin, T. Methodology, design, and evaluation in psychotherapy research. In: Lambert, MJ.; Lambert, Michael J., editors. Bergin and Garfield's handbook of psychotherapy and behavior change. 5th ed.. New York, NY: Wiley; 2004. p. 16-43.
15. Campbell, DT.; Stanley, JC. Experimental and quasi-experimental designs for research. Chicago: R. McNally; 1966.
16. Mohr DC, Spring B, Freedland KE, Beckner V, Arean P, Hollon SD, Ockene J, Kaplan R. The selection and design of control conditions for randomized controlled trials of psychological interventions. Psychother Psychosom. 2009; 78(5):275–284. [PubMed: 19602916]
17. Leon AC, Solomon DA. Toward rapproachment in the placebo control debate. A calculated compromise of power. Eval Health Prof. 2003; 26(4):404–414. [PubMed: 14631611]
18. Hochman M, McCormick D. Characteristics of published comparative effectiveness studies of medications. JAMA. 2010; 303(10):951–958. [PubMed: 20215609]
19. Thompson BT, Schoenfeld D. Usual care as the control group in clinical trials of nonpharmacologic interventions. Proc Am Thorac Soc. 2007; 4(7):577–582. [PubMed: 17878473]
20. van der Graaf R, van Delden JJ. What is the best standard for the standard of care in clinical research? Am J Bioeth. 2009; 9(3):35–43. [PubMed: 19247887]
21. Au DH, Castro M, Krishnan JA. Selection of controls in clinical trials: introduction and conference summary. Proc Am Thorac Soc. 2007; 4(7):567–569. [PubMed: 17878470]
22. Dawson L, Zarin DA, Emanuel EJ, Friedman LM, Chaudhari B, Goodman SN. Considering usual medical care in clinical trial design. PLoS Med. 2009; 6(9):e1000111. [PubMed: 19787044]
23. Minneci PC, Eichacker PQ, Danner RL, Banks SM, Natanson C, Deans KJ. The importance of usual care control groups for safety monitoring and validity during critical care research. Intensive Care Med. 2008; 34(5):942–947. [PubMed: 18214424]
24. Silverman HJ, Miller FG. Control group selection in critical care randomized controlled trials evaluating interventional strategies: An ethical assessment. Crit Care Med. 2004; 32(3):852–857. [PubMed: 15090973]
25. Takala J. Better conduct of clinical trials: the control group in critical care trials. Crit Care Med. 2009; 37(1 Suppl):S80–S90. [PubMed: 19104230]
26. Arean PA, Alvidrez J. Ethical considerations in psychotherapy effectiveness research: choosing the comparison group. Ethics Behav. 2002; 12(1):63–73. [PubMed: 12171083]
27. Brigham GS, Feaster DJ, Wakim PG, Dempsey CL. Choosing a control group in effectiveness trials of behavioral drug abuse treatments. J Subst Abuse Treat. 2009; 37(4):388–397. [PubMed: 19553062]

28. Burns T. End of the road for treatment-as-usual studies? Br J Psychiatry. 2009; 195(1):5–6. [PubMed: 19567887]
29. Degenholtz HB, Parker LS, Reynolds CF. Trial design and informed consent for a clinic-based study with a treatment as usual control arm. Ethics Behav. 2002; 12(1):43–62. [PubMed: 12171082]
30. Reynolds CF III, Degenholtz H, Parker LS, Schulberg HC, Mulsant BH, Post E, Rollman B. Treatment as usual (TAU) control practices in the PROSPECT Study: managing the interaction and tension between research design and ethics. Int J Geriatr Psychiatry. 2001; 16(6):602–608. [PubMed: 11424169]
31. de Bruin M, Viechtbauer W, Hospers HJ, Schaalma HP, Kok G. Standard care quality determines treatment outcomes in control groups of HAART-adherence intervention studies: implications for the interpretation and comparison of intervention effects. Health Psychol. 2009; 28(6):668–674. [PubMed: 19916634]
32. de Bruin M, Viechtbauer W, Schaalma HP, Kok G, Abraham C, Hospers HJ. Standard care impact on effects of highly active antiretroviral therapy adherence interventions: A meta-analysis of randomized controlled trials. Arch Intern Med. 2010; 170(3):240–250. [PubMed: 20142568]
33. Gava I, Barbui C, Aguglia E, Carlino D, Churchill R, De Vanna M, McGuire HF. Psychological treatments versus treatment as usual for obsessive compulsive disorder (OCD). Cochrane Database Syst Rev. 2007; (2):CD005333. [PubMed: 17443583]
34. Staines GL, McKendrick K, Perlis T, Sacks S, De Leon G. Sequential assignment and treatment-as-usual. Alternatives to standard experimental designs in field studies of treatment efficacy. Eval Rev. 1999; 23(1):47–76. [PubMed: 10346073]
35. Davidson KW, Rieckmann N, Clemow L, Schwartz JE, Shimbo D, Medina V, Albanese G, Kronish I, Hegel M, Burg MM. Enhanced depression care for patients with acute coronary syndrome and persistent depressive symptoms: coronary psychosocial evaluation studies randomized controlled trial. Arch Intern Med. 2010; 170(7):600–608. [PubMed: 20386003]
36. Rollman BL, Belnap BH, LeMenager MS, Mazumdar S, Houck PR, Counihan PJ, Kapoor WN, Schulberg HC, Reynolds CF III. Telephone-delivered collaborative care for treating post-CABG depression: a randomized controlled trial. JAMA. 2009; 302(19):2095–2103. [PubMed: 19918088]
37. Alexopoulos GS, Reynolds CF III, Bruce ML, Katz IR, Raue PJ, Mulsant BH, Oslin DW, Ten HT. Reducing suicidal ideation and depression in older primary care patients: 24-month outcomes of the PROSPECT study. Am J Psychiatry. 2009; 166(8):882–890. [PubMed: 19528195]
38. Grossman P, Kappos L, Gensicke H, D'Souza M, Mohr DC, Penner IK, Steiner C. MS quality of life, depression, and fatigue improve after mindfulness training: a randomized trial. Neurology. 2010; 75(13):1141–1149. [PubMed: 20876468]
39. Barratt R, Frost G, Millward DJ, Truby H. A randomised controlled trial investigating the effect of an intensive lifestyle intervention v. standard care in adults with type 2 diabetes immediately after initiating insulin therapy. Br J Nutr. 2008; 99(5):1025–1031. [PubMed: 18197995]
40. Byrne NM, Meerkin JD, Laukkanen R, Ross R, Fogelholm M, Hills AP. Weight loss strategies for obese adults: personalized weight management program vs. standard care. Obesity (Silver Spring). 2006; 14(10):1777–1788. [PubMed: 17062808]
41. Raskin J, Smith TR, Wong K, Pritchett YL, D'Souza DN, Iyengar S, Wernicke JF. Duloxetine versus routine care in the long-term management of diabetic peripheral neuropathic pain. J Palliat Med. 2006; 9(1):29–40. [PubMed: 16430342]
42. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Kerstens PJ, Nielen MM, Vos K, van Schaardenburg D, Speyer I, Seys PE, Breedveld FC, Allaart CF, Dijkmans BA. DAS-driven therapy versus routine care in patients with recent-onset active rheumatoid arthritis. Ann Rheum Dis. 2010; 69(1):65–69. [PubMed: 19155234]
43. Berkman LF, Blumenthal J, Burg M, Carney RM, Catellier D, Cowan MJ, Czajkowski SM, DeBusk R, Hosking J, Jaffe A, Kaufmann PG, Mitchell P, Norman J, Powell LH, Raczynski JM, Schneiderman N. Effects of treating depression and low perceived social support on clinical events after myocardial infarction: the Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) Randomized Trial. JAMA. 2003; 289(23):3106–3116. [PubMed: 12813116]

44. Bhardwaj A, Rehman SU, Mohammed A, Baggish AL, Moore SA, Januzzi JL Jr. Design and methods of the Pro-B Type Natriuretic Peptide Outpatient Tailored Chronic Heart Failure Therapy (PROTECT) Study. Am Heart J. 2010; 159(4):532–538. [PubMed: 20362709]
45. Spiegel DA, Bruce TJ. Benzodiazepines and exposure-based cognitive behavior therapies for panic disorder: conclusions from combined treatment trials. Am J Psychiatry. 1997; 154(6):773–781. [PubMed: 9167504]
46. Watanabe N, Churchill R, Furukawa TA. Combined psychotherapy plus benzodiazepines for panic disorder. Cochrane Database Syst Rev. 2009; (1):CD005335. [PubMed: 19160253]
47. Eagle KA, Montoye CK, Riba AL, DeFranco AC, Parrish R, Skorcz S, Baker PL, Faul J, Jani SM, Chen B, Roychoudhury C, Elma MA, Mitchell KR, Mehta RH. Guideline-based standardized care is associated with substantially lower mortality in medicare patients with acute myocardial infarction: the American College of Cardiology's Guidelines Applied in Practice (GAP) Projects in Michigan. J Am Coll Cardiol. 2005; 46(7):1242–1248. [PubMed: 16198838]
48. Brignole M, Ungar A, Bartoletti A, Ponassi I, Lagi A, Mussi C, Ribani MA, Tava G, Disertori M, Quartieri F, Alboni P, Raviele A, Ammirati F, Scivales A, De Santo T. Standardized-care pathway vs. usual management of syncope patients presenting as emergencies at general hospitals. Europace. 2006; 8(8):644–650. [PubMed: 16864618]
49. Sofaer N, Eyal N. The diverse ethics of translational research. Am J Bioeth. 2010; 10(8):19–30. [PubMed: 20694901]
50. Pace C, Miller FG, Danis M. Enrolling the uninsured in clinical trials: an ethical perspective. Crit Care Med. 2003; 31(3 Suppl):S121–S125. [PubMed: 12626956]
51. Davis RM, Hitch AD, Salaam MM, Herman WH, Zimmer-Galler IE, Mayer-Davis EJ. TeleHealth improves diabetes self-management in an underserved community: diabetes TeleCare. Diabetes Care. 2010; 33(8):1712–1717. [PubMed: 20484125]
52. Haskell WL, Berra K, Arias E, Christopherson D, Clark A, George J, Hyde S, Klieman L, Myll J. Multifactor cardiovascular disease risk reduction in medically underserved, high-risk patients. Am J Cardiol. 2006; 98(11):1472–1479. [PubMed: 17126653]
53. Goodman SG, Menon V, Cannon CP, Steg G, Ohman EM, Harrington RA. Acute ST-segment elevation myocardial infarction: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest. 2008; 133(6 Suppl) 708S-75S.
54. Verheugt FW. Acute myocardial infarction associated with ST segment elevation and the new European Society of Cardiology guidelines. Heart. 2009; 95(13):1112–1117. [PubMed: 19525260]
55. Kushner FG, Hand M, Smith SC, King SB III, Anderson JL, Antman EM, Bailey SR, Bates ER, Blankenship JC, Casey DE Jr, Green LA, Hochman JS, Jacobs AK, Krumholz HM, Morrison DA, Ornato JP, Pearle DL, Peterson ED, Sloan MA, Whitlow PL, Williams DO. 2009 Focused Updates: ACC/AHA Guidelines for the Management of Patients With ST-Elevation Myocardial Infarction (updating the 2004 Guideline and 2007 Focused Update) and ACC/AHA/SCAI Guidelines on Percutaneous Coronary Intervention (updating the 2005 Guideline and 2007 Focused Update): a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation. 2009; 120(22):2271–2306. [PubMed: 19923169]
56. Crowe E, Lovibond K, Gray H, Henderson R, Krause T, Camm J. Early management of unstable angina and non-ST segment elevation myocardial infarction: summary of NICE guidance. BMJ. 2010; 340:c1134. [PubMed: 20335335]
57. Kerr EA, Fleming B. Making performance indicators work: experiences of US Veterans Health Administration. BMJ. 2007; 335(7627):971–973. [PubMed: 17991979]
58. Tu JV, Khalid L, Donovan LR, Ko DT. Indicators of quality of care for patients with acute myocardial infarction. CMAJ. 2008; 179(9):909–915. [PubMed: 18936456]
59. Lewis MH, Gohagan JK, Merenstein DJ. The locality rule and the physician's dilemma: local medical practices vs the national standard of care. JAMA. 2007; 297(23):2633–2637. [PubMed: 17579232]
60. Merenstein D. A piece of my mind. Winners and losers. JAMA. 2004; 291(1):15–16. [PubMed: 14709561]

61. Lurie P, Wolfe SM. Unethical trials of interventions to reduce perinatal transmission of the human immunodeficiency virus in developing countries. N Engl J Med. 1997; 337(12):853–856. [PubMed: 9295246]
62. Varmus H, Satcher D. Ethical complexities of conducting research in developing countries. N Engl J Med. 1997; 337(14):1003–1005. [PubMed: 9309109]
63. Kent DM, Mwamburi DM, Bennish ML, Kupelnick B, Ioannidis JP. Clinical trials in sub-Saharan Africa and established standards of care: a systematic review of HIV, tuberculosis, and malaria trials. JAMA. 2004; 292(2):237–242. [PubMed: 15249573]
64. Cohen MG, Fonarow GC, Peterson ED, Moscucci M, Dai D, Hernandez AF, Bonow RO, Smith SC Jr. Racial and ethnic differences in the treatment of acute myocardial infarction: findings from the Get With the Guidelines-Coronary Artery Disease program. Circulation. 2010; 121(21):2294–2301. [PubMed: 20479153]
65. Werner RM, Goldman LE, Dudley RA. Comparison of change in quality of care between safety-net and non-safety-net hospitals. JAMA. 2008; 299(18):2180–2187. [PubMed: 18477785]
66. Colon-Otero G, Smallridge RC, Solberg LA Jr, Keith TD, Woodward TA, Willis FB, Dunn AN. Disparities in participation in cancer clinical trials in the United States : a symptom of a healthcare system in crisis. Cancer. 2008; 112(3):447–454. [PubMed: 18085590]
67. Gross CP. Racial disparities in clinical trial enrolment. Lancet. 2008; 372(9651):1713–1714. [PubMed: 19013309]
68. Timmermans S, McKay T. Clinical trials as treatment option: bioethics and health care disparities in substance dependency. Soc Sci Med. 2009; 69(12):1784–1790. [PubMed: 19819059]
69. Byrne MM, Thompson P. Collective equipoise, disappointment, and the therapeutic misconception: on the consequences of selection for clinical research. Med Decis Making. 2006; 26(5):467–479. [PubMed: 16997925]
70. Holroyd KA, Powers SW, Andrasik F. Methodological issues in clinical trials of drug and behavior therapies. Headache. 2005; 45(5):487–492. [PubMed: 15953265]
71. Sheps SB. Research methods for surgeons: an overview. J Invest Surg. 1993; 6(4):321–328. [PubMed: 8399003]
72. Bass NM, Mullen KD, Sanyal A, Poordad F, Neff G, Leevy CB, Sigal S, Sheikh MY, Beavers K, Frederick T, Teperman L, Hillebrand D, Huang S, Merchant K, Shaw A, Bortey E, Forbes WP. Rifaximin treatment in hepatic encephalopathy. N Engl J Med. 2010; 362(12):1071–1081. [PubMed: 20335583]
73. Donta ST, Clauw DJ, Engel CC Jr, Guarino P, Peduzzi P, Williams DA, Skinner JS, Barkhuizen A, Taylor T, Kazis LE, Sogg S, Hunt SC, Dougherty CM, Richardson RD, Kunkel C, Rodriguez W, Alicea E, Chiliade P, Ryan M, Gray GC, Lutwick L, Norwood D, Smith S, Everson M, Blackburn W, Martin W, Griffiss JM, Cooper R, Renner E, Schmitt J, McMurtry C, Thakore M, Mori D, Kerns R, Park M, Pullman-Mooar S, Bernstein J, Hershberger P, Salisbury DC, Feussner JR. Cognitive behavioral therapy and aerobic exercise for Gulf War veterans' illnesses: a randomized controlled trial. JAMA. 2003; 289(11):1396–1404. [PubMed: 12636462]
74. 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(4):295–309. [PubMed: 18283207]
75. Powell LH, Calvin JE Jr, Richardson D, Janssen I, Mendes de Leon CF, Flynn KJ, Grady KL, Rucker-Whitaker CS, Eaton C, Avery E. Self-management counseling in patients with heart failure: the heart failure adherence and retention randomized behavioral trial. JAMA. 2010; 304(12):1331–1338. [PubMed: 20858878]
76. Woody GE, Poole SA, Subramaniam G, Dugosh K, Bogenschutz M, Abbott P, Patkar A, Publicker M, McCain K, Potter JS, Forman R, Vetter V, McNicholas L, Blaine J, Lynch KG, Fudala P. Extended vs short-term buprenorphine-naloxone for treatment of opioid-addicted youth: a randomized trial. JAMA. 2008; 300(17):2003–2011. [PubMed: 18984887]
77. Thiels C, Schmidt U, Treasure J, Garthe R, Troop N. Guided self-change for bulimia nervosa incorporating use of a self-care manual. Am J Psychiatry. 1998; 155(7):947–953. [PubMed: 9659862]

78. Carroll KM, Connors GJ, Cooney NL, DiClemente CC, Donovan DM, Kadden RR, Longabaugh RL, Rounsaville BJ, Wirtz PW, Zweben A. Internal validity of Project MATCH treatments: discriminability and integrity. J Consult Clin Psychol. 1998; 66(2):290–303. [PubMed: 9583332]

79. Allen LA, Woolfolk RL, Escobar JI, Gara MA, Hamer RM. Cognitive-behavioral therapy for somatization disorder: a randomized controlled trial. Arch Intern Med. 2006; 166(14):1512–1518. [PubMed: 16864762]

80. Phillips LS, Branch WT, Cook CB, Doyle JP, El-Kebbi IM, Gallina DL, Miller CD, Ziemer DC, Barnes CS. Clinical inertia. Ann Intern Med. 2001; 135(9):825–834. [PubMed: 11694107]

81. Safford MM, Shewchuk R, Qu H, Williams JH, Estrada CA, Ovalle F, Allison JJ. Reasons for not intensifying medications: differentiating "clinical inertia" from appropriate care. J Gen Intern Med. 2007; 22(12):1648–1655. [PubMed: 17957346]

82. Henke RM, Zaslavsky AM, McGuire TG, Ayanian JZ, Rubenstein LV. Clinical inertia in depression treatment. Med Care. 2009; 47(9):959–967. [PubMed: 19704353]

83. Katon W, Russo J, Lin EH, Heckbert SR, Karter AJ, Williams LH, Ciechanowski P, Ludman E, Von KM. Diabetes and poor disease control: is comorbid depression associated with poor medication adherence or lack of treatment intensification? Psychosom Med. 2009; 71(9):965–972. [PubMed: 19834047]

84. Rose AJ, Berlowitz DR, Manze M, Orner MB, Kressin NR. Comparing methods of measuring treatment intensification in hypertension care. Circ Cardiovasc Qual Outcomes. 2009; 2(4):385–391. [PubMed: 20031865]

85. Katon WJ, Von KM, Lin EH, Simon G, Ludman E, Russo J, Ciechanowski P, Walker E, Bush T. The Pathways Study: a randomized trial of collaborative care in patients with diabetes and depression. Arch Gen Psychiatry. 2004; 61(10):1042–1049. [PubMed: 15466678]

86. Fraker TD Jr, Fihn SD, Gibbons RJ, Abrams J, Chatterjee K, Daley J, Deedwania PC, Douglas JS, Ferguson TB Jr, Fihn SD, Fraker TD Jr, Gardin JM, O'Rourke RA, Williams SV, Smith SC Jr, Jacobs AK, Adams CD, Anderson JL, Buller CE, Creager MA, Ettinger SM, Halperin JL, Hunt SA, Krumholz HM, Kushner FG, Lytle BW, Nishimura R, Page RL, Riegel B, Tarkington LG, Yancy CW. 2007 chronic angina focused update of the ACC/AHA 2002 Guidelines for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines Writing Group to develop the focused update of the 2002 Guidelines for the management of patients with chronic stable angina. Circulation. 2007; 116(23):2762–2772. [PubMed: 17998462]

87. Kowalski CJ. Pragmatic problems with clinical equipoise. Perspect Biol Med. 2010; 53(2):161–173. [PubMed: 20495255]

88. Miller FG, Brody H. What makes placebo-controlled trials unethical? Am J Bioeth. 2002; 2(2):3–9. [PubMed: 12189059]

89. Appel LJ, Champagne CM, Harsha DW, Cooper LS, Obarzanek E, Elmer PJ, Stevens VJ, Vollmer WM, Lin PH, Svetkey LP, Stedman SW, Young DR. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. JAMA. 2003; 289(16):2083–2093. [PubMed: 12709466]

90. Svetkey LP, Harsha DW, Vollmer WM, Stevens VJ, Obarzanek E, Elmer PJ, Lin PH, Champagne C, Simons-Morton DG, Aickin M, Proschan MA, Appel LJ. Premier: a clinical trial of comprehensive lifestyle modification for blood pressure control: rationale, design and baseline characteristics. Ann Epidemiol. 2003; 13(6):462–471. [PubMed: 12875806]

91. Snapinn SM. Noninferiority trials. Curr Control Trials Cardiovasc Med. 2000; 1(1):19–21. [PubMed: 11714400]

92. Califf RM. A perspective on the regulation of the evaluation of new antithrombotic drugs. Am J Cardiol. 1998; 82(8B) 25P–35P.

93. Mohr DC, Vella L, Hart S, Heckman T, Simon G. The effect of telephone-administered psychotherapy on symptoms of depression and attrition: A meta-analysis. Clinical Psychology: Science and Practice. 2008; 15(3):243–253. [PubMed: 21369344]

94. Mohr DC, Siddique J, Ho J, Duffecy J, Jin L, Fokuo JK. Interest in Behavioral and Psychological Treatments Delivered Face-to-Face, by Telephone, and by Internet. Ann Behav Med. 2010

95. Freedman B. Equipoise and the ethics of clinical research. N Engl J Med. 1987; 317(3):141–145. [PubMed: 3600702]
96. Miller FG, Brody H. A critique of clinical equipoise. Therapeutic misconception in the ethics of clinical trials. Hastings Cent Rep. 2003; 33(3):19–28. [PubMed: 12854452]
97. Miller FG, Brody H. Clinical equipoise and the incoherence of research ethics. J Med Philos. 2007; 32(2):151–165. [PubMed: 17454420]
98. Killaspy H, Kingett S, Bebbington P, Blizard R, Johnson S, Nolan F, Pilling S, King M. Randomised evaluation of assertive community treatment: 3-year outcomes. Br J Psychiatry. 2009; 195(1):81–82. [PubMed: 19567902]
99. Jensen PS, Weersing R, Hoagwood KE, Goldman E. What is the evidence for evidence-based treatments? A hard look at our soft underbelly. Ment Health Serv Res. 2005; 7(1):53–74. [PubMed: 15832693]
100. Eisenberg DM, Delbanco TL, Berkey CS, Kaptchuk TJ, Kupelnick B, Kuhl J, Chalmers TC. Cognitive behavioral techniques for hypertension: are they effective? Ann Intern Med. 1993; 118(12):964–972. [PubMed: 8489111]
101. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: variable, substantial, and growing. JAMA. 2002; 287(14):1840–1847. [PubMed: 11939870]
102. Fawcett J, Epstein P, Fiester SJ, Elkin I, Autry JH. Clinical management--imipramine/placebo administration manual. NIMH Treatment of Depression Collaborative Research Program. Psychopharmacol Bull. 1987; 23(2):309–324. [PubMed: 3303100]
103. Glassman AH, O'Connor CM, Califf RM, Swedberg K, Schwartz P, Bigger JT Jr, Krishnan KR, van Zyl LT, Swenson JR, Finkel MS, Landau C, Shapiro PA, Pepine CJ, Mardekian J, Harrison WM, Barton D, Mclvor M. Sertraline treatment of major depression in patients with acute MI or unstable angina. JAMA. 2002; 288(6):701–709. [PubMed: 12169073]
104. Brent DA, Holder D, Kolko D, Birmaher B, Baugher M, Roth C, Iyengar S, Johnson BA. A clinical psychotherapy trial for adolescent depression comparing cognitive, family, and supportive therapy. Arch Gen Psychiatry. 1997; 54(9):877–885. [PubMed: 9294380]
105. Edinger JD, Wohlgemuth WK, Radtke RA, Coffman CJ, Carney CE. Dose-response effects of cognitive-behavioral insomnia therapy: a randomized clinical trial. Sleep. 2007; 30(2):203–212. [PubMed: 17326546]
106. Safren SA, Sprich S, Mimiaga MJ, Surman C, Knouse L, Groves M, Otto MW. Cognitive behavioral therapy vs relaxation with educational support for medication-treated adults with ADHD and persistent symptoms: a randomized controlled trial. JAMA. 2010; 304(8):875–880. [PubMed: 20736471]
107. Wampold, BE. The great psychotherapy debate: models, methods, and findings. Mahwah, N.J.: L. Erlbaum Associates; 2001.
108. DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, O'Reardon JP, Lovett ML, Gladis MM, Brown LL, Gallop R. Cognitive therapy vs medications in the treatment of moderate to severe depression. Arch Gen Psychiatry. 2005; 62(4):409–416. [PubMed: 15809408]
109. Hollon SD, DeRubeis RJ, Shelton RC, Amsterdam JD, Salomon RM, O'Reardon JP, Lovett ML, Young PR, Haman KL, Freeman BB, Gallop R. Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. Arch Gen Psychiatry. 2005; 62(4):417–422. [PubMed: 15809409]
110. Klein DF. Preventing hung juries about therapy studies. J Consult Clin Psychol. 1996; 64(1):81–87. [PubMed: 8907087]
111. Cochran BN, Pruitt L, Fukuda S, Zoellner LA, Feeny NC. Reasons underlying treatment preference: an exploratory study. J Interpers Violence. 2008; 23(2):276–291. [PubMed: 18162641]
112. King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, Sibbald B, Lai R. Conceptual framework and systematic review of the effects of participants' and professionals' preferences in randomised controlled trials. Health Technol Assess. 2005; 9(35):1-iv. [PubMed: 16153352]

113. Robinson EJ, Kerr CE, Stevens AJ, Lilford RJ, Braunholtz DA, Edwards SJ, Beck SR, Rowley MG. Lay public's understanding of equipoise and randomisation in randomised controlled trials. Health Technol Assess. 2005; 9(8):1-iv. [PubMed: 15763039]

114. Bedi N, Chilvers C, Churchill R, Dewey M, Duggan C, Fielding K, Gretton V, Miller P, Harrison G, Lee A, Williams I. Assessing effectiveness of treatment of depression in primary care. Partially randomised preference trial. Br J Psychiatry. 2000; 177:312–318. [PubMed: 11116771]

115. Adamson SJ, Sellman DJ, Dore GM. Therapy preference and treatment outcome in clients with mild to moderate alcohol dependence. Drug Alcohol Rev. 2005; 24(3):209–216. [PubMed: 16096124]

116. Hegerl U, Hautzinger M, Mergl R, Kohnen R, Schutze M, Scheunemann W, Allgaier AK, Coyne J, Henkel V. Effects of pharmacotherapy and psychotherapy in depressed primary-care patients: a randomized, controlled trial including a patients' choice arm. Int J Neuropsychopharmacol. 2010; 13(1):31–44. [PubMed: 19341510]

117. King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, Sibbald B, Lai R. Impact of participant and physician intervention preferences on randomized trials: a systematic review. JAMA. 2005; 293(9):1089–1099. [PubMed: 15741531]

118. Chilvers C, Dewey M, Fielding K, Gretton V, Miller P, Palmer B, Weller D, Churchill R, Williams I, Bedi N, Duggan C, Lee A, Harrison G. Antidepressant drugs and generic counselling for treatment of major depression in primary care: randomised trial with patient preference arms. BMJ. 2001; 322(7289):772–775. [PubMed: 11282864]

119. Preference Collaborative Review Group. Patients' preferences within randomised trials: systematic review and patient level meta-analysis. BMJ. 2008; 337:a1864. [PubMed: 18977792]

120. Kocsis JH, Leon AC, Markowitz JC, Manber R, Arnow B, Klein DN, Thase ME. Patient preference as a moderator of outcome for chronic forms of major depressive disorder treated with nefazodone, cognitive behavioral analysis system of psychotherapy, or their combination. J Clin Psychiatry. 2009; 70(3):354–361. [PubMed: 19192474]

121. Kwan BM, Dimidjian S, Rizvi SL. Treatment preference, engagement, and clinical improvement in pharmacotherapy versus psychotherapy for depression. Behav Res Ther. 2010; 48(8):799–804. [PubMed: 20462569]

122. Lin P, Campbell DG, Chaney EF, Liu CF, Heagerty P, Felker BL, Hedrick SC. The influence of patient preference on depression treatment in primary care. Ann Behav Med. 2005; 30(2):164–173. [PubMed: 16173913]

123. Colagiuri B. Participant expectancies in double-blind randomized placebo-controlled trials: potential limitations to trial validity. Clin Trials. 2010; 7(3):246–255. [PubMed: 20421243]

124. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Engl J Med. 2001; 344(21):1594–1602. [PubMed: 11372012]

125. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. J Intern Med. 2004; 256(2):91–100. [PubMed: 15257721]

126. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. Cochrane Database Syst Rev. 2010; (1):CD003974. [PubMed: 20091554]

127. Olshansky B. Placebo and nocebo in cardiovascular health: implications for healthcare, research, and the doctor-patient relationship. J Am Coll Cardiol. 2007; 49(4):415–421. [PubMed: 17258086]

128. Kaptchuk TJ, Friedlander E, Kelley JM, Sanches MN, Kokkotou E, Singer JP, Kowalczykowski M, Miller FG, Kirsch I, Lembo AJ. Placebos without Deception: A Randomized Controlled Trial in Irritable Bowel Syndrome. PLoS ONE. 2010; 5(12):e15591. [PubMed: 21203519]

129. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. Eur Neuropsychopharmacol. 2009; 19(1):34–40. [PubMed: 18823760]

130. Rutherford B, Sneed J, Devanand D, Eisenstadt R, Roose S. Antidepressant study design affects patient expectancy: a pilot study. Psychol Med. 2010; 40(5):781–788. [PubMed: 19732481]

131. Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, Lanctot KL. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled

trials? Results from pooled and meta-analyses. J Clin Psychiatry. 2010; 71(3):270–279. [PubMed: 20122371]

132. Stone DA, Kerr CE, Jacobson E, Conboy LA, Kaptchuk TJ. Patient expectations in placebo-controlled randomized clinical trials. J Eval Clin Pract. 2005; 11(1):77–84. [PubMed: 15660541]
133. Colagiuri B, Morley K, Boakes R, Haber P. Expectancy in double-blind placebo-controlled trials: an example from alcohol dependence. Psychother Psychosom. 2009; 78(3):167–171. [PubMed: 19270472]
134. Davidson KW, Trudeau KJ, Ockene JK, Orleans CT, Kaplan RM. A primer on current evidence-based review systems and their implications for behavioral medicine. Ann Behav Med. 2004; 28(3):226–238. [PubMed: 15576261]

TABLE 1

Existing Practice Control Groups

Condition	Description	Stringency	Uniformity	Vulnerability to Unintentional Intensification
Naturalistic				
Inadequate Care (IC)	Inaccessible or poor quality nonstudy care	Low	Trial-Dependent	Low
Treatment As Usual (TAU), Usual Care (UC), Routine Care (RC), or Standard Care (SC)	Typical, normative, nonstudy care	Moderate	Trial-Dependent	High
Standard of Care (SOC)				
Individualized (iSOC)	Individualized, state-of-the-art, guideline-adherent nonstudy care	High	Low	Low
Uniform (uSOC)	Uniform, state-of-the-art, guideline-adherent nonstudy care	High	High	Low
Modified				
Standardized treatment regimen (STR)	Nonstudy care is standardized for the trial to ensure uniformity of care, and possibly higher quality care	Moderate to High	High	Low
Constrained Usual Care (CUC)	Specific aspects of usual nonstudy care are withheld or discouraged during the trial	Moderate	Trial-Dependent	Trial-Dependent
Enhanced Usual Care (EUC)				
Minimally Enhanced (mEUC)	Nonstudy care is enhanced in minor ways during the trial to address methodological or ethical issues	Moderate	Trial-Dependent	Trial-Dependent
Extensively Enhanced (eEUC)	Nonstudy care is substantially enhanced during the trial to bring it closer to the standard of care	High	Trial-Dependent	Trial-Dependent