ELSEVIER

General Hospital Psychiatry 25 (2003) 115–123

General Hospital Psychiatry

Commentary and Perspective

From time to time, the Journal receives manuscripts that can be thought of as opinion pieces, essays, or editorial comment on matters of topical interest. Such submissions will be refereed in the usual fashion and, if suitable, published in this section. The Editorial Board invites Letters to the Editor or rebutting commentary with the understanding that all submissions are subject to editing.

Research to improve the quality of care for depression: alternatives to the simple randomized clinical trial

Thomas R. TenHave, Ph.D.[a,*], James Coyne, Ph.D.[a], Mark Salzer, Ph.D.[a], Ira Katz, M.D., Ph.D.[a,b]

[a]*University of Pennsylvania Medical School, Philadelphia, PA*
[b]*Veterans Administration Medical Center, Philadelphia, PA*

Abstract

Recognition of gaps between evidence gained from mental health research and clinical practice in the community together with changes in treatment patterns and patient/provider preferences for care have led to interest in enhancements in the designs and analyses of clinical and community trials of mental health interventions. Gaps between clinical trials and community care include differences in populations and treatment strategies. To bridge these gaps, we propose enhancing the simple randomized trial with several different designs with the immediate aims of improving patient recruitment and adherence in psychiatric intervention studies thus bringing study designs more in line with clinical practice. The goals are to estimate treatment efficacy and effectiveness so that both internal and external validity are optimized. In this discussion, we address design and analytic issues with respect to a number of enhancements of the randomized trial design, including partial patient-provider preference designs, randomized encouragement and consent designs, fixed adaptive design, and random between- and within-patient adaptive designs. Each has advantages and disadvantages depending on the effect under investigation. Some of these enhancements, such as the fixed adaptive design, have begun to be implemented in effectiveness trials in mental health services research, but all are worthy of more attention.

Keywords: Adaptive randomized; Partial patient preference; Randomized encouragement; Randomized consent

1. Introduction

Recent pressures to improve patient care in psychiatry in a cost-effective way have brought to light the gulf between the relatively simple and fixed treatments typically assessed in randomized clinical trials and the complex exigencies of routine care. There are two important aspects to the gap between the results of randomized clinical trials and what occurs in clinical practice, especially in treating psychiatric disorders: 1) the gap between populations receiving psychiatric care and the highly select study samples recruited to clinical trial research; and 2) the gap between the complex individualized and sequential treatment required for psychiatric disorders and the relatively fixed and simple interventions examined in clinical trials research. Other factors include attrition and nonadherence to treatment, which is often not addressed in the literature reports of clinical trials.

Evidence for the gap in outcomes is well-documented. It has been reported that 80–90% of persons with major depressive disorder can be treated successfully based on efficacy studies, [1] whereas, among patients receiving usual care in primary care settings, as many as 80% of those with depression fail to improve, and, among those who do, relapse is common. [2–7] Many clinical trials demonstrating higher success rates have typically enrolled highly selected and motivated patients who become the target of resource-intensive interventions.

More recent effectiveness health services trials have been reporting better response rates in usual care as well as recruitment rates, but these have been paralleled by substantial increases in the improvement observed with placebos in clinical trials [8]. A substantial gap remains between what is

* Corresponding author. Tel.:+1-215-573-4885; fax:+1-215-573-4865.

E-mail address: ttenhave@cceb.upenn.edu (T.R. TenHave).

0163-8343/03/$ – see front matter
doi:10.1016/S0163-8343(02)00275-X

obtained for active treatment in clinical trials and what is observed in routine care outside of specialty settings. Furthermore, although data remain limited, the gap with what is obtained in clinical trials seems to extend to the routine care provided by psychiatrists as well as primary care physicians [9].

One aspect of the gap between clinical trials and practice pertains to differences in treatment patterns. It has been reported that approximately 40–50% of patients fail to remit on the initial antidepressant prescribed to them [10]. Although care often requires combinations or sequences of treatments, seldom are the full algorithms that guide decisions subject to empirical tests in clinical trials. When such complexity is accommodated, seldom is it done in a way that allows examination of the contributions of specific components of the algorithm. Another aspect of the outcomes gap entails selection bias or sample versus population differences [11–15]. Zimmerman, Mattia and Posternak [16] reported that 85% of depressed outpatients would be excluded from the typically antidepressant randomized trial. In a study of screening for depression, 60% of primary care patients screening positive failed to complete a follow-up interview [17]. Using waiting room screening as a basis for identifying eligible patients, Wells, Sherbourne, Schoenbaum, Duan, Meredith, Unutzer et al. [18] found that less than half were willing to enroll in a clinical trial aimed at improving the outcome of depression. Some of these shortcomings of randomized trials are not apparent because of the dearth of trials reporting selection bias indicators.

Further complicating matters, the resulting population of potential depression study participants is constantly changing. There is a shrinking pool of persons who have never been exposed to any treatment and who are not currently receiving an antidepressant at the point of efforts to recruit them into clinical trials. As antidepressant prescribing has increased, rates of treatment may now equal or exceed the prevalence of depression [19]. Decreased stigma, increased treatment availability, and direct to consumer marketing of pharmaceuticals are all associated with increased identification and treatment of depression. They also affect the characteristics and preferences of potential study populations.

Patient/provider preference arguably is one factor underlying these gaps. Here, the term “preference” refers to a valenced (positive or negative) attitude toward a particular treatment. Such preferences diminish the ability of clinical research to detect treatment effectiveness and accurately estimate treatment effects. Preferences are likely to affect study recruitment, because patients and their providers self select as to whether they want to participate in a study that may involve being assigned to treatment arms that vary in their desirability. Preferences may also affect the degree to which patients engage in or adhere to the treatment, and their willingness to remain in the study [20].

Patient/provider preference can clearly lead to both nonadherence to treatment and study attrition. Schulberg, Block, Madonia, Rodriguez, Scott and Lave, in one of the largest efficacy studies of guideline-level treatment in primary care, screened over 10,000 primary care patients aged 18–64 years and randomized 283 patients with major depressive disorder to treatment with nortriptyline, interpersonal psychotherapy, or usual care. [21–22] Whether randomized to antidepressant or interpersonal therapy, 15% of patients dropped out immediately after randomization. Despite the team resources afforded the intervention patients in this study, including pharmacotherapists and psychotherapists dedicated to the intervention, only one in three of the intervention patients completed a full course of therapy. However, among those completing treatment, approximately 70% of intervention patients compared with 20% of usual care patients were recovered at 8 months. Moreover, in another study, Oxman, Barrett, Sengupta, Katon, Williams, Frank and Hegel [23] found that the majority of primary care patients who completed either pharmacotherapy or brief problem-solving therapy opted not to continue treatment at the end of randomized clinical trial targeting minor depression and dysthymia. This is consistent with the high rates of discontinuation of treatment with antidepressants in routine care [24] or open label studies [25].

Such resistance to treatment occurs in spite of the popularization of antidepressants through the media and word of mouth, and to increased confidence among patients and practitioners about their safety, acceptability, and effectiveness. The net result is that depressed persons are more likely to have a history of at least partial treatment that affects preferences and expectations. Those who might have once been interested in participating in research to obtain a desired treatment, such as medication, can now more readily obtain these treatments without the burden of participating in trials. The implications for conducting research in such a changing environment are profound. For example, in a randomized trial performed in the United Kingdom that compared antidepressant medication with marital therapy, Leff, Vearnals, Brewin, Wolff, Alexander, Asen et al. reported that marital therapy resulted in better outcomes [26]. However, this positive result was qualified by a 55% drop-out rate in the antidepressant group compared to 15% in the marital therapy group. One explanation for this discrepancy may have been that study enrollees had greater preferences for marital therapy, a treatment that is not widely available in the UK. They may have been interested in study participation because it provided them with an opportunity to receive marital therapy. Antidepressant treatment is widely available and is part of the national health care coverage and there would be less incentive to participate in the study if one were more inclined to obtain pharmacological treatment. Once enrolled and assigned to a treatment condition those assigned to receive marital therapy may have been more inclined to adhere to treatment and to remain in the

study than those assigned to the less preferred medication treatment arm.

Closing these large gaps between clinical trials research and outcomes in practice has proven to be very difficult because there may be significant interactions between a number of equally salient and challenging factors: patient/provider self-selection, nonadherence, and attrition. Self-selection, nonadherence and attrition can work in a synergistic way to produce misleading conclusions from these studies. When this occurs, findings from clinical trials can actually widen the gap between research and clinical practice. Intent-to-treat analyses are typically performed with the intention of producing tests and estimates of treatment efficacy or the "planned treatment effect" [27] (i.e., the impact of an intervention under ideal conditions). However, when there is treatment nonadherence, intent-to-treat analyses do not produce tests or estimates of treatment efficacy; rather they produce estimates and tests of the "implemented treatment effect" [27] or "public health benefit" [28] (i.e., the impact of an intervention in a population with the same pattern of patient behavior as in the study sample). In combination with the biased selection of study samples for clinical trials, treatment nonadherence can result in inferences about treatment effects limited by the adherence behavior that occurs in very exclusive populations [20,28]. The external validity of clinical trials is affected by the triple blow of biased patient selection, treatment nonadherence and attrition, which are all functions of patient preference and treatment patterns in the population.

We note that there are now effectiveness trials being designed and implemented that begin to incorporate attention to patient preference and clinician flexibility in the delivery of treatment [29,30], which fall under the class of "fixed adaptive designs" in the subsequent section on "Alternative Study Designs". However, the full range of options available for addressing such issues have not been articulated, and the data analytic strategies available in published studies do not yet exploit their full potential. Namely, results from designs allowing clinician discretion in the delivery of treatment are analyzed in a molar, blackbox fashion– in a way that does not allow evaluation of the effectiveness of particular treatment choices. Such evaluations are now facilitated in more valid ways than previously by new statistical methods discussed in the "Emerging Methods of Analysis Section".

2. Alternative study designs

In response to the above concerns about patient recruitment, nonadherence and attrition and their impact on validity, we consider enhancements in the designs and analyses of randomized clinical or community trials for evaluating interventions in the study of depression that can be extended to other mental health conditions. The aims of these newer designs are to improve patient recruitment and treatment adherence and to reduce attrition in psychiatric intervention studies, and to bring study designs more in line with clinical practice in estimating treatment efficacy and effectiveness. Such modifications may make trials more attractive to participants and make the findings more useful to all stakeholders.

Four different types of designs will be reviewed as enhancements of the traditional randomized trial to formally accommodate patient/provider preferences and/or nonadherence: 1) "fixed adaptive designs" in which patients are randomized to treatment arms that entail algorithmic sequences of treatment changes to adapt to intermediate patients outcomes (e.g., augmentation with a second line antidepressant if the patient does not respond to the initial antidepressant); 2) "randomized adaptive designs," which represent a sequentially randomized version of fixed adaptive designs, and in which all subjects are sequentially randomized, so that at each visit they are randomized to a specific component of a regimen for the subsequent period, but in such a way ("biased coin toss") that previous outcomes and patient and provider preference influence the probabilities of assignment to specific treatment regimens; 3) "randomized consent" or "encouragement" designs in which all subjects are randomized, but then are given the option to adhere to their assigned treatment or switch treatments if a different treatment is preferred; and 4)"partially randomized patient preference" (or just "patient preference") designs where subjects decide if they are to be randomized to the interventions or allowed to select the intervention assignment.

The above designs extend traditional randomized clinical trial designs by incorporating understandings of treatment preferences and their impact on study behaviors into study designs in order to allow more ecologically informed inferences. While the study designs of interest in this paper all entail randomization of study participants to interventions, they differ with respect to the timing of patient/provider choice, outcomes, and randomization (See Table 1). Several designs allow patient/provider choice prior to randomization, whereas others designs provide for patient or physician-based decisions after randomization.

Methodologically, the fixed adaptive treatment and randomized consent designs are the most closely related to the traditional randomized trial designs. They involve randomizing all study subjects once, and then allow different intervention protocols depending on preferences. The randomized adaptive and patient preference designs differ the most from traditional randomized designs, because they entail either randomization to interventions of only a sub-set of patients (patient preference) or a sequence of randomization for each subject (randomized adaptive). Among these designs, the patient preference designs are susceptible to the most bias in estimating differences between interventions. Such designs have evolved in a number of different forms in other areas of biomedical research to limit this susceptibil-

Table 1
Comparison of different designs with respect to timing of assessment of patient preference and randomization

Design	Time 1	Time 2
Partial patient preference	Patient preference	Randomization
Randomized encouragement/consent	Randomization	Patient preference
Randomized adaptive	Randomization/patient preference, outcome	Randomization/patient preference, outcome
Fixed adaptive	Randomization	Patient preference, outcome
Traditional randomized	Randomization	Patient outcome

ity. However, as long as they entail a patient choice prior to randomization, they remain plagued by bias.

Essentially, all of the designs considered here including the patient preference designs rely on perceptions of equipoise holding among study subjects. Otherwise, there will be very few subjects actually receiving the intervention that is not preferred by subjects, and regardless of design, estimates of treatment effect will consequently be difficult to obtain.

These designs differ with respect to the treatment effect of interest in studies. As indicated above, we distinguish between the "planned treatment effect" [27] and the "implemented treatment effect" [27] or "public health benefit" [28]. The planned treatment effect is usually estimated using interventions studies under conditions in which considerable control is exerted over patient (e.g., highly restricted, adherent populations with noncomplicated illnesses), provider (e.g., well trained and supervised in delivering the specific intervention, oftentimes with limited caseloads), and intervention (e.g., sufficient dose, high degree of fidelity in delivery of intervention). In contrast, the implemented treatment effect pertains to much more variable conditions such as diverse, less adherent patients, more complicated illnesses in terms of severity or co-occurring disorders, providers (e.g., less supervision and oversight in delivering the intervention, caseloads that are similar to what is found in usual care settings), and intervention (e.g., great variability in dose, less fidelity to the model form of the intervention based on patient, provider, and setting characteristics).

2.1. Fixed adaptive designs

Such designs are becoming more prevalent in the mental health services research field in attempting to ensure that evidence-based treatments are administered to the appropriate patient population in community practice [29–31]. The common feature of these studies is that all patients in a given study are randomized to treatment arms entailing algorithmic sequence of treatment changes to adapt to intermediate patients outcomes (e.g., augmentation with a second line antidepressant if the patient does not respond to the initial antidepressant). Such a design is referred to as a "fixed adaptive" design because the range of optional algorithmic sequences to which subjects can be randomized is predetermined prior to randomization. It is conceptually distinct from the randomized adaptive design, under which the algorithmic sequence is specified as the trial proceeds with a sequence of randomizations or coin-tosses.

2.2. Randomized adaptive designs

Lavori and Dawson [32] have proposed a design that is appropriate to evaluating treatment algorithms that entail switching the dosage or nature of the treatment that individual patient are receiving based on the observed effects of previous assignments. With appropriate analyses, this design provides a modeling of the flexibility seen in clinical practice while allowing for the evaluation of the consequences of particular choices in treatment. These applications take findings from monitoring the clinical outcomes within individual patients as the basis for making decisions about subsequent care. The obvious applications are in the evaluation of rules for adjusting dosages of antidepressants, augmenting treatment with additional medications or psychotherapy or switching treatments altogether. It is also possible to use instances of nonadherence such as patients' failures to keep psychotherapy appointments as intermediate outcomes defining clinical decision points. The intervention could then consist of introduction of some support for patient adherence, whether this consists of simple prompts or more elaborate efforts to overcome logistical barriers such as provision of transportation. These additions to clinical trials are seldom evaluated in terms of their contribution to the overall treatment effect observed and often added in ways that preclude evaluation and limit the generalizability of results. Even worse, they are often added nonrandomly or over the course of a trial as the significance of these factors for nonadherence or attrition become apparent. Because they are not part of the formal intervention, they may not be given their due attention in the description of methods or the interpretation of results.

The within-subject biased-coin design proposed by Lavori and Dawson [32] is an adaptation of the "play-the-winner" design with ball-in-the-urn modification of randomization probabilities. The original "play-the-winner" strategy entailed randomizing only the first recruited subject and then basing each subsequent subject's treatment assignment on the success of the previous subject's assigned treatment [33]. For instance, if the previous subject exhibited a positive treatment response with the assigned treatment then the current subject would receive that treatment, otherwise the current subject would receive the other treat-

ment The ball-in-the-urn modification of this play-the-winner strategy involves randomizing each subject, but allowing the randomization probability of each subject to deviate from the first randomization probability. This deviation is based on a mathematical function of the previous outcomes, such as a probability-yielding transformation of the sum of the previous outcome scores [34].

Lavori and Dawson [32] extended this ball-in-the-urn modification of the play-the-winner strategy by applying it to a repeated measures analysis in the context of determining the optimal time for switching or augmenting antidepressant medication. More specifically, under their design all subjects begin a trial on an initial treatment, and at each of the multiple preplanned follow-up visits each subject is randomized to remain on initial treatment or switch/augment. The randomization probability for a given subject at a given follow-up visit is a function of the subject's outcomes in previous visits, again such as a probability-producing transformation of the sum of the previous outcome scores. The resulting probability is then the randomization probability at a current visit, thus allowing previous outcomes to influence the chances of starting anew treatment for a given individual at that visit. In the context of assessing the efficacy of more complex algorithmic treatment interventions, the sequential randomization aspects of this design also allow more valid estimation of overall treatment efficacy and the efficacy of the individual components of the algorithm, in addition to estimation of the optimal timing of particular regimes proposed by Lavori and Dawson [32]. These benefits are discussed in more detail below.

2.3. Partially randomized patient-preference designs

Strong evidence and persuasive arguments from the psychological literature suggest that treatment preferences influence study behaviors (e.g., Corrigan & Salzer, 2001) and potentially compromise research aimed at understanding treatment effects [20]. Another class of designs, partially randomized patient-preference designs, is particularly useful where strong preferences among some patients threaten either the ability to recruit an adequate sample size of representative patients or where such preferences threaten patients' acceptance of treatment assignment, adherence, or retention in the clinical trial, thus leading to possibly substantial nonrandom differences in outcomes between those who receive the treatment and those who do not. The partially randomized patient-preference designs entail strategies that anticipate this preference problem using prospective participant analysis prior to randomization, avoiding attempted comparisons of resulting groups that are likely to be noncomparable.

The class of patient preference designs also allows modeling the clinical context in which patients exercise the opportunity to make choices among available treatments. As well as offering some accommodation of these choices, it provides an evaluation of the extent to which the patient expectations reflected in preferences for particular treatments might influence the observed performance of the assigned treatment. Yet another set of applications come in situations where preferences on the part of providers or referral sources may affect recruitment and retention, and, here also, this class of designs allows modeling and evaluation of this common clinical reality.

These benefits of the patient preference designs impact the selection bias and adherence problems underlying randomized trials in psychiatry. Such designs may be more attractive to potential subjects than the traditional randomized trials for which treatment availability depends on chance, thus reducing the selection bias inherent in randomized trials. Additionally, it is thought that adherence and effectiveness may be greater for behavioral interventions if a patient has some influence in choosing the treatment [35]. However, there is also evidence that being allowed to choose treatment does not necessarily improve short term outcome in depressed patients in primary care given either antidepressants or counseling [36] or nondirective counseling, cognitive-behavioral therapy, or usual general practitioner care [37]. The justification for patient preference designs is therefore one of increased subject enrollment and increased generalizability to the mixture of patients who are involved in deciding on interventions and those who experience equipoise with respect to choice of intervention. Increased adherence may also result from the patient preference designs, thus increasing the internal validity of studies. Because of these benefits, this class of designs has received much attention in a number of different fields such as dentistry and obstetrics research [38–41]. However, applications in psychiatry have been relatively recent [42].

However, these benefits need to be weighed against several disadvantages. The stringent sample size requirements and the possibility that all patients in the preference arm may choose one treatment if the study patients do not perceive equipoise, thus precluding any comparison in the preference group that is parallel to the randomized comparison. In particular, a separate analysis of treatment differences needs to be performed in each of the preference and randomized group, leading to a second order comparison of treatment effects between the two types of patients, i.e., a test of group-treatment interaction. Such a comparison requires an approximate increase of 50–100% of the original sample size necessary for a completely randomized trial. Moreover, the comparison within the preference group and between the preference and randomized groups to assess for equal treatment effects in the two groups is confounded. However, the problems related to such confounding may be diminished, and the utility of the designs increased, if the preference groups are included primarily for comparisons in the postacute phases of treatment (e.g., continuation or maintenance phases), when the balancing benefits of preacute phase randomization are diluted due to drop-out

and nonresponse during the acute phase. That is, the confounding differences between the preference and randomized groups is most likely less substantial after the acute phase when the randomized groups are "contaminated" by losing drop-outs and nonresponders for the continuation or maintenance phases.

In an attempt to resolve the increased susceptibility to confounding, variants in patient preference designs that include two levels of randomization have been proposed [43–45]. In these designs, each enrolled subject is randomized to either: 1) a second randomization into one of two treatment arms; or 2) a patient preference arm under which a decision by the patient is made in terms of treatment assignment. Such a design also accommodates patient preference, but benefits from having one level of randomization that is applied to all subjects. However, the additional randomization is only beneficial in assessing differences between patients who choose their treatment and those who are randomized to their treatment. That is, the second order (i.e., interaction) assessment of treatment effect differences between preference and randomized groups of patients will not be as susceptible to confounding of unobserved patient differences. Nonetheless, within the patient preference group, any treatment effect between those who prefer one intervention over another will be subject to unobserved confounding. Hence, any second order difference between preference and randomized patients will be difficult to explain in terms of patient preference or unobserved patient differences in the patient preference group. Bradley [38] provided some other criticisms of this two-level randomized patient-preference design, in advocating the single-randomization patient preference design. Nonetheless, both types of patient preference designs are susceptible to the same type of unobserved confounding of comparisons between interventions in the patient preference group.

Lavori, Rush, Wisniewski, Alpert, Fava, Kupfer et al. [44] proposed a related design "equipoise-stratified randomization under which patients or their providers are allowed to choose from among several different overall strategies within which the patient and provider perceive equipoise among all treatments. After a general class of strategies is chosen, patients are then randomized to the individual treatment regimes within a class selected by the patient and provider. Lavori, Rush, Wisniewski, Alpert, Fava, Kupfer et al. [44] proposed applying such a scheme to the STEP-BD program, which offers a multitude of alternative treatments and possibilities for augmenting or switching between them. Once a patient and/or his/her provider have chosen a set of alternative treatments that are acceptable, patients are randomized to one of the target drugs either at the onset of treatment or when either switching/augmenting is required. As with the patient preference designs, the issue of confounding arises when comparing treatment effects among the different sets of equipoise treatments.

2.4. Randomized encouragement designs

Randomized encouragement or randomized consent designs offer another set of strategies for addressing patient preference or nonadherence. They involve randomizing all subjects but then giving them the option to adhere to their assigned treatment or switch treatments if a different treatment is preferred. The intent is to pre-empt nonadherence by allowing patients to choose according to protocol. Zelen [45] defined two types of randomized consent designs: 1) "single consent randomized design", in which only study subjects randomized to the experimental arm are asked to consent to the experimental treatment or to the standard treatment, whereas subjects randomized to the standard treatment are not asked for consent and expected to adhere to the standard treatment; and 2) the "double consent randomized design", in which study subjects in each of the experimental and standard arms are asked to consent to their respective treatments or to receiving the alternative treatment [46]. Duan, Wells, Braslow and Weisz [28] have proposed and implemented [47] an extension of this type of design to include encouragement strategies for the patients randomized to experimental treatment, as a way to formalize adherence strategies. In studies examining ways of improving the outcome of depression in primary care, the encouragement adaptation of the single consent randomized trial design may be applicable in cases where a care manager-aided intervention is compared to a usual care arm in the context of primary care practices, where care managers are typically not available without study resources. In the studies where different well-established therapies are under comparison (e.g., SSRI versus non-SSRI antidepressants), a double consent randomized design with an encouragement component may be more appropriate, as either treatment is publicly available to all patients.

3. Emerging methods of analysis

Estimation of both efficacy and effectiveness is discussed here in two contexts:1) overall treatment regimens (e.g., a sequence of several medication and/or psycho-therapies through acute, continuation, and maintenance phases); and 2) individual components of treatment regimens (e.g., effect of a specific medication in the acute phase before any needed rescue medication or therapy is administered). With treatment nonadherence, intent-to-treat analyses of these efficacy studies provide estimates of treatment effectiveness, i.e., the average treatment effect in a population with the same mixture of adherers and nonadherers as in the study sample [48]. Estimates of treatment efficacy can be obtained from these clinical trials in the presence of treatment nonadherence when approaches such as the instrumental variable procedure are employed [48]. Here, the instrumental variable is randomization and is used to control for potential bias due to unobserved factors related to the out-

Table 2
Summarized comparison of alternative designs to randomized trial

Design	Benefits	Costs
Patrial patient preference	Generalizability to randomize nonconsenters	Much larger sample size required
Randomized encouragement	Formally accommodate patient/provider preference	Potential for absence of equipoise
Randomized adaptive	Testing efficacy of individual treatment components	Complicated logistics; sample size requirements
Fixed adaptive	Algorithms offering treatment choices	Black box comparison; less vulnerable to non-adherence and drop-out
Traditional randomized	Traditional design relatively easy to implement	Black box comparison; less vulnerable to non-adherence and drop-out

come and nonadherence to treatment. However, these approaches require assumptions, such as no direct effects of randomization apart from taking the assigned treatment that may not be verifiable from the collected data or difficult to satisfy in unblinded trials.

3.1. Overall effectiveness

For obtaining estimates of the overall treatment effectiveness or overall program effect, using the intent-to-treat analysis with the fixed adaptive or randomized consent designs is suitable, because the overall treatment regimens are randomized under these two designs. This is not the case for the randomized adaptive designs, in which individual components of a treatment regimen are randomized and hence intent-to-treat comparisons of overall regimens are difficult to perform if not precluded.

3.2. Overall "planned treatment effect"

Furthermore, the randomized or fixed adaptive designs may be used to estimate the "planned treatment effect" of the overall treatment (i.e., the effect of receiving a planned algorithm) with recently developed statistical methods for estimation of individual treatment effects at each stage, and then weighted averaging for combining these individual component effects to obtain estimates of overall efficacy due to a specific regimen of treatment [49]. Such estimates of efficacy are contingent on delineating the individual components of the overall treatment regimen and then defining and measuring adherence to these individual components. An example of a sequence of delineated treatment components may be: 1) the initiation of treatment in the acute phase with an SSRI; 2) continuing SSRI treatment for responders in the continuation phase; or initiating alternate medication for nonresponders still in the acute phase; and 3) for nonresponders who do not respond to the second medication, starting a third medication or interpersonal therapy.

The methods of analysis for estimating overall efficacy under randomized encouragement protocols represent extensions of instrumental variable and weighted regression methods. The weighted regression approaches, also known as marginal structural modeling [49], employ weights that are proportional to the probability of receiving the treatment. Such a probability is based on nonadherence levels, patient and provider preferences, and randomization probabilities (in the case of the randomized adaptive designs).

3.3. "Planned treatment effect" of individual components

For obtaining estimates of the "planned treatment effect" (i.e., efficacy) of individual components, the randomized adaptive design is most appropriate because the randomization at each stage of the treatment regimen reduces the chances of unmeasured confounding of the treatment effects. In contrast, other designs including the fixed adaptive design require certain assumptions such as the absence of unmeasured confounders to obtain estimates of the efficacy of individual components of a treatment regimen.

We note that possibilities for inference of any treatment effect under the patient preference design is limited, because of confounding due to unmeasured differences between the group that is not randomized and the group that is. The corresponding analysis approaches rely on regression adjustments such as propensity scores that adjust for potentially observed confounders. Alternatively, latent class models may be employed for identifying potential classes of subjects within which confounding is controlled [50].

Finally, not only is the randomized adaptive design beneficial with respect to estimating the efficacy of overall treatment regimens and their individual components, but the design also facilitates the estimation of the optimal timing of a particular regimen. Lavori and Dawson [32] proposed using a Bayesian procedure that relies on prior information provided by study investigators and previous and current treatment outcomes for individual patients in imputing unobserved outcomes under the alternative treatment that was not received by the current subject (e.g., watchful waiting if a subject was randomized to active treatment). However, one can also employ more parametric methods [49].

4. Conclusions and recommendations

The various designs discussed in this paper all have their unique benefits and costs that make them appropriate under different study contexts and represent methodological and analytic steps that can be taken to extend knowledge generated by more traditional randomized designs (See Table 2). For investigating individual components of complex interventions, it is

recommended that the randomized adaptive design be utilized. This randomized adaptive design can accommodate patient/provider inputs by altering randomization probabilities in light of previous outcomes and preferences in estimates of overall efficacy. It can also be useful for individual components. Use of the related Zelen's "Play the winner" design augmented with ball-in-the-urn modification of randomization probabilities may be appropriate for studies of simple interventions such as comparisons of face-to-face versus telephone-based approaches to disease management for depression. In this case, nonadherence on the part of those patients assigned to one intervention would increase the probability of randomization of future subjects to the other arm. The fixed adaptive randomization strategies would be useful for assessing black box interventions, in which the efficacy of individual components is not a high priority, and when patients have positive perceptions of equipoise. For more emphasis on patient preferences, either the patient preference or the randomized encouragement designs may be used with the available sample size the determining factor among these two patient preference-based approaches. More specifically, the patient preference design under which only a part of the sample is randomized allows the assessment of whether differences between groups randomized to receive alternative treatments also exist between groups that chose these interventions. However, such comparisons need large sample sizes on the order of tests for interactions. The bias problems with patient preference designs may be diminished if patient preference groups are introduced into the analysis in the maintenance or more subsequent phases of treatment when the benefits of baseline randomization are diminished. In contrast, the randomized encouragement design results in the randomization of all subjects. However, nonadherence can reduce effect sizes. Hence, if sample size is a limiting factor then randomized encouragement designs would be most useful; otherwise patient preference designs would be preferable. We also note that the randomized encouragement design is more susceptible to lack of perceived equipoise on the part of patients and their providers because it offers choices of treatments to randomized participants.

In working to close the gaps between current randomized clinical trials, and the need to deliver effective care to populations of patients, there is a need for research to develop hybrid designs that combine the best aspects of each of the approaches discussed above For example, one such design might entail randomizing each prospective randomized patient selection to either a preference or randomized group before treatment assignment is made. In each of the preference and randomization groups, the treatment under investigation could include encouragement strategies in which adherence to treatment is the focus of care management strategies. In groups receiving encouragement, the elements of care provided could be determined through an adaptive treatment format, so as to also allow assessment of timing of switching/augmenting and titration rates of antidepressant medications. Such a design would be complex, both conceptually and logistically. However, the difficulties mirror those encountered in delivering high quality care to real patients. If that goal is to be achieved, research must move beyond simple randomized clinical trials to address the real interventions that are required by representative patients and providers.

Acknowledgments

Grant support from R01-MH-61892 "Mixed Models for Discrete Data with Non-compliance," PI: Thomas R. Ten Have.

Grant support from P30-MH2129 "IRC: Depression in Late Life—Psychiatric/Medical Comorbidity," PI: Ira Katz.

References

[1] Ormel J, Kempen GI, Deeg DJ, et al. Functioning, well-being, and health perception in late middle-aged, and older people: comparing the effects of depressive symptoms, and chronic medical conditions. J Am Geriatr Soc 1998;46(1):39–48.

[2] Callahan CM, Hendrie HC, Dittus RS, et al. Improving treatment of late life depression in primary care: a randomized clinical trial. J Am Geriatr Soc 1994;42:839–46.

[3] Schulberg HC, Block MR, Madonia MJ, et al. The "usual care" of major depression in primary care practice. Arch Fam Med 1997;6:334–9.

[4] Schulberg HC, Mulsant B, Schulz R, et al. Characteristics and course of major depression in older primary care patients. Int J Psychiatry Med 1998;28(4):421–36.

[5] Thase ME. How should efficacy be evaluated in randomized clinical trials of treatments for depression? J Clin Psychiatry 1999;60(Suppl 4):23–31.

[6] Cole MG, Bellavance F, Mansour A. Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. Am J Psychiatry 1999;156(8):1182–9.

[7] Thase ME. Long-term nature of depression. J Clin Psychiatry 1999;60(Suppl 14):3–9, [discussion 31–5].

[8] Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression. Variable, substantial, and growing. JAMA 2002;287:1840–47.

[9] Simon GE, Von Korff M, Rutter CM, Peterson DA. Treatment process and outcomes for managed care patients receiving new antidepressant prescriptions from psychiatrists and primary care physicians. Arch Gen Psych 2001;58:395–401.

[10] AHCPR, Volume 2. AHCPR Guidelines (Depression Guideline Panel. Depression in Primary Care. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. AHCPR Publication No. 93–0551, 1993.

[11] Alvidrez J, Azocar F. Distressed women's clinic patients: preferences for mental health treatments and perceived obstacles. Gen Hosp Psychiatry 1999;21(5):340–47.

[12] Bakker A, Spinhaven P, van Balkom AJLM, Vleugel L, van Dyck R. Cognitive therapy by allocation versus cognitive therapy by preference in the treatment of panic disorder. Psychoth and Psychosom 2000;69:240–43.

[13] Bradley C. Psychological issues in clinical trial design. Irish J Psychol 1997;18:67–87.

[14] Morin C, Gaulier B, Barry T, Kowatch RA. Patients' acceptance of psychological, and pharmacological therapies for insomnia. Sleep 1992;15:302–5.

[15] Walker J, Vincent N, Furer P, Cox B, Kjernisted K. Treatment preference in hypochondriasis. J Behav Ther Exp Psychiat 1999;30: 251–58.
[16] Zimmerman M, Mattia JI, Posternak MA. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? Am J Psychiatry 2002;159:469–73.
[17] Leon AC, Portera L, Olfson M, et al. Diagnostic errors of primary care screens for depression and panic disorder. Int J Psychiat Med 1999;29:1–11.
[18] Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. JAMA 2000;283:212–20.
[19] Coyne JC, Thompson R, Klinkman MS, Nease DE. Emotional Disorders in primary care. Emotional disorders in primary care. J Consult Clin Psychol 2002;70:798–809.
[20] Corrigan PW, Salzer MS. The conflict between random assignment and treatment preference: implications for internal validity. Evaluation and Program Planning, (in press).
[21] Schulberg HC, Block MR, Madonia MJ, et al. Applicability of clinical pharmacotherapy guidelines for major depression in primary care settings. Arch Fam Med 1995;4(2):106–12.
[22] Schulberg HC, Block MR, Madonia MJ, et al. Treating major depression in primary care practice. Eight-month clinical outcomes. Arch Gen Psychiatry 1996;53(10):913–9.
[23] Oxman TE, Barrett JE, Sengupta A, et al. Status of minor depression or dysthymia in primary care following a randomized controlled treatment. Gen Hosp Psych 2001;23(6):301–10.
[24] Katon W, Vonkorff M, Lin E, Bush T, Ormel J. Adequacy and duration of antidepressant treatment in primary care. Medical Care 1992;30:67–76.
[25] Stevens T, Katona C, Manela M, Watkin V, Livingston G. Drug treatment of older people with affective disorders in the community. Int'l J Ger Psychiatry 1999;6:467–72.
[26] Leff J, Vearnals S, Brewin CR, et al. The London depression intervention trial: randomised controlled trial of antidepressants vs couple therapy in the treatment and maintenance of people with depression living with a partner: clinical outcome and costs. Br J Psychiatry 2000;177:95–100.
[27] Collins LM, Murphy SA, Bierman KA, Design, and evaluation of adaptive preventive interventions, (unpublished manuscript).
[28] Duan N, Wells KB, Braslow JT, Weisz JR. Randomized encouragement trial: a paradigm for public health intervention evaluation (unpublished manuscript).
[29] Katon W, Von Korff M, Lin E, et al. Stepped collaborative care for primary care patients with persistent symptoms of depression: a randomized trial. Arch Gen Psychiat 1999;56:1109–15.
[30] Unutzer J, Katon W, Williams JW, et al. Improving primary care for depression in late life: the design of a multicenter randomized trial. Med Care 2001;39:785–99.
[31] Wells KB. The design of Partners in Care: evaluating the cost-effectiveness of improving care for depression in primary care. Sociological Psychiatr Epidemiol 1999;34:20–9.
[32] Lavori PW, Dawson R. A design for testing clinical strategies: biased adaptive within-subject randomization. J Roy Stat Soc A Sta 2000; 163:29–38.
[33] Zelen M. Play the winner rule and the controlled clinical trial. JASA 1969;64:131–46.
[34] Wei LJ, Smythe RT, Lin DY, Park TS. Statistical inference with data-dependent treatment allocation rules. JASA 1990;85:156–62.
[35] Editorial, British Journal of Obstetrics, and Gynaecology, 1997;104: 1332–35.
[36] Bedi N, Chilvers C, Churchill R, et al. Assessing effectiveness of treatment of depression in primary care. Partially randomised preference trial. Br J Psychiatry 2000;177:312–18.
[37] Ward E, King M, Lloyd M, et al. Randomised controlled trial of non-directive counseling, cognitive-behavior therapy, and usual general practitioner care for patients with depression. I. Clinical effectiveness. BMJ 2000;321:1383–8.
[38] Bradley C. Patient preferences and clinical trial design and interpretation: appreciation and critique of a paper by Feine, Awad & Lund. Community Dentistry and Oral Epidemiology 1999;27:85–8.
[39] Feine JS, Awad MA, Lund JP. The impact of patient preference on the design and interpretation of clinical trials. Community Dentistry and Oral Epidemiology 1998;26:70–4.
[40] Cooper KG. The impact of using a partially randomized patient preference design when evaluating alternative managements for heavy menstrual bleeding. Brit J Obstet Gyn 1997;104:1367–73.
[41] Brewin CR, Bradley C. Patient preferences and randomized clinical trials. BMJ 1989;299:313–15.
[42] Hunter MS, Liao KLM. Evaluation of a four-session cognitive-behavioral intervention for menopausal hot flashes. Br J Hlth Psych 1996;1:113–25.
[43] Chilvers C, Dewey M, Fielding K, et al. Antidepressant drugs and generic counseling for treatment of major depression in primary care: randomised trial with patient preference arms. BMJ 2001;322:1–5.
[44] Lavori PW, Rush JA, Wisniewski SR, et al. Strengthening clinical effectiveness trials: equipoise-stratified randomization. Arch Biol Psychiat 2001;50:792–801.
[45] Zelen M. A new design for randomized clinical trials. NEJM 1979; 300:1242–46.
[46] Zelen M. Randomized consent designs for clinical trials: an update. Statistics in Medicine 1990;9:645–56.
[47] Wells K. The design of partners in care: evaluating the cost-effectiveness of improving care for depression in primary care. Soc Psychiatry Epidemiol 1999;34:20–9.
[48] Rubin DB. More powerful randomization-based p-values in double-blind trials with non-compliance. Statistics in Medicine 1998;17:371–85.
[49] Murphy SA, Van der Laan MJ, Robins JM, CPPRG. Marginal mean models for dynamic regimes. J Am Stat Assoc 2001;96:1410–1423..
[50] Muthen B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics 1999;55:463–69.