Prevention Science, Vol. 5, No. 3, September 2004 (© 2004)

A Conceptual Framework for Adaptive Preventive Interventions

Linda M. Collins,[1,4] Susan A. Murphy,[2] and Karen L. Bierman[3]

Recently, *adaptive* interventions have emerged as a new perspective on prevention and treatment. Adaptive interventions resemble clinical practice in that different dosages of certain prevention or treatment components are assigned to different individuals, and/or within individuals across time, with dosage varying in response to the intervention needs of individuals. To determine intervention need and thus assign dosage, adaptive interventions use prespecified decision rules based on each participant's values on key characteristics, called tailoring variables. In this paper, we offer a conceptual framework for adaptive interventions, discuss principles underlying the design and evaluation of such interventions, and review some areas where additional research is needed.

KEY WORDS: adaptive interventions; prevention; research design.

For most of the history of research-based interventions aimed at prevention and treatment, the composition and dosage of these interventions have been *fixed*, in other words, a single composition and dosage has been offered to all program participants. For example, a school-based drug abuse prevention curriculum might be delivered to all sixth graders. Every component of the intervention that may be necessary for any particular participant is included in the curriculum, and each child is given the same intervention. Although it is recognized that individuals may have different intervention needs, it is expected that the intervention is in no way diluted or made counterproductive if components that are particularly relevant for an individual are combined with components that may have less, or even no, relevance for that individual.

[1]The Methodology Center and Department of Human Development and Family Studies, The Pennsylvania State University, University Park, Pennsylvania.
[2]Institute for Social Research and Department of Statistics, University of Michigan, Ann Arbor, Michigan.
[3]Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania.
[4]Correspondence should be directed to Linda M. Collins, 159 S. Henderson, University Park, Pennsylvania 16802; e-mail: lmcollins@psu.edu.

Recently, *adaptive* interventions have emerged as a new perspective on research-based prevention and treatment. According to this perspective, the varying intervention needs of individuals may not be met optimally by using a single uniform composition and dosage. For this reason, an adaptive intervention assigns different dosages of certain program components across individuals, and/or within individuals across time. Dosage varies in response to the intervention needs of individuals, and dosages are assigned based on decision rules linking characteristics of the individual with specific levels and types of program components. In some adaptive interventions a dosage of zero is possible on a given component. This implies that there may be individuals who do not receive certain components at all, and that different types or versions of program components may be assigned to different individuals. Part of the conceptual appeal of the adaptive approach is its clear resemblance to clinical practice. However, in order to maintain replicability (see below), adaptive interventions entail the use of explicit decision rules, thus differing from most clinical practice.

Adaptive interventions are becoming more common, as prevention programs move in the direction

1389-4986/04/0900-0185/1

of more comprehensive, multilayered systems of prevention services (Weissberg & Greenberg, 1998). Many adaptive interventions have appeared in the prevention and treatment literature in the last 15 years. Throughout this paper we will use one such adaptive intervention[5] as an example, namely Fast Track (Conduct Problems Prevention Research Group, 1992, 1999a, 1999b). Fast Track is a multiyear, multicomponent program designed to prevent conduct disorders in high-risk children. In addition to core intervention components delivered to all study participants (e.g., parent training and child social skill training groups), some components were delivered adaptively. For example, the number of home-based counseling visits assigned to each family varied depending upon level of parental functioning, and reading tutoring was assigned only to children who were demonstrating academic difficulties. The dosage assignment of the adaptive components in Fast Track was time-varying, that is, the dosage was adjusted up to three times per year. Other examples of adaptive interventions and treatments include Borhani *et al.* (1991), Breslin *et al.* (1999), Brooner and Kidorf (2002), Dishion and Kavanagh (2000), Kreuter and Strecher (1996), Prochaska *et al.* (2001), and Sobell and Sobell (1999, 2000). Lavori and Dawson (1998) and Lavori *et al.* (2000) discuss advantages of adaptive treatments in medical research.

Despite the increased interest in, and implementation of, adaptive interventions, there has been relatively little attention paid to some important conceptual and methodological issues. In this paper, we offer a conceptual framework for adaptive interventions. We discuss principles underlying the design and evaluation of adaptive interventions; use of these principles to help create an effective adaptive intervention, and to help understand where an intervention may have gone wrong; statistical analysis of adaptive interventions; and the potential strengths and limitations of adaptive interventions, as compared to fixed interventions. We conclude by reviewing several open topics related to adaptive interventions.

[5]Many adaptive interventions incorporate a universal intervention that is delivered to all participants, making then hybrids of adaptive and fixed interventions. We will adopt the convention here of calling any intervention that contains adaptive components an adaptive intervention, because all of the design considerations we raise in this paper apply to these interventions.

DESIGN PRINCIPLES FOR ADAPTIVE INTERVENTIONS

In an adaptive intervention, the assignment of a particular level of dosage and/or type of treatment is based on the individual's values on variables that are expected to moderate the effect of the treatment component. We will refer to these as *tailoring variables*. The list of candidate tailoring variables is almost endless, and naturally will depend on the study. Common types of tailoring variables include individual, family, or context characteristics representing risk or protective factors that influence responsivity to (or need for) various types or intensity of preventive intervention. The logic is that the level or type of intervention required to address the needs of individuals varies according to these tailoring variables. For example, individuals who are characterized by a particular risk factor may require an intensive intervention, whereas less intervention will be sufficient for individuals who do not have this characteristic. In time-varying adaptive interventions, the intervention is sustained over a period of time, and tailoring variables are assessed periodically, so that the intervention can be adjusted on an ongoing basis according to individual changes on the tailoring variable. In this case, the tailoring variable might reflect treatment responsivity or a proximal outcome, with dose adjusted based upon each participant's progress toward a prespecified threshold representing a "successful" outcome. For example, in the Fast Track Project level of parental functioning was the tailoring variable determining recommended dose of home visiting, and academic performance was the tailoring variable determining whether a child received tutoring.

The first principle we wish to point out is an essential difference between fixed interventions and adaptive interventions with respect to exactly what constitutes the intervention. In the adaptive case, the intervention consists of not only the treatment components, but the treatment components inextricably coupled with the entire system for assigning dosage. In other words, the choice of tailoring variables, the measures of the tailoring variables, the decision rules linking tailoring variables to dosage assignment, and the implementation of these rules *are a part of the intervention itself*. (Note that according to this framework, aspects of the intervention, such as individual staff, schools, treatment sites, etc., are not part of the intervention. Rather, they are sources of extraneous variance.) Treatment, tailoring variables, measurement of tailoring variables, decision rules, and implementation of decision rules are interdependent; for example, decision rules will not be effective unless they are based on well-measured tailoring variables. Furthermore, these five aspects of adaptive interventions constitute a chain that is only as strong as its weakest link. In this paper we do not discuss the treatment itself; this is an ongoing topic covered at length elsewhere, and in any case is content-specific. We describe design considerations aimed at maintaining the strength of the remaining links in the chain, to wit (1) identification of adaptive components and related tailoring variables, (2) measurement of tailoring variables, (3) derivation of decision rules, and (4) implementation of decision rules.

The design principles discussed in this paper are aimed at achieving two general objectives. The first objective is maximizing the strength of the adaptive intervention. Even an adaptive intervention that enjoys a potentially powerful preventive treatment can be weakened by *any one* of poorly chosen tailoring variables, poorly measured tailoring variables, poorly conceived decision rules, or poorly implemented decision rules. A well-designed intervention avoids these weaknesses, maximizing the potential of the treatment.

The second objective is to maximize *replicability*. Replicability means that when a study is repeated on different samples, the same population-level treatment effect is being estimated in each sample.[6] The idea of replicability is an important one in prevention. We have the most confidence in a preventive intervention when its effects are replicable with different experimenters, different clinical staff, different locations, etc. In fact, one aspect of replicability is what Flay (1986) has termed "effectiveness," the ability of an intervention to maintain the desired effect under real-world implementation conditions. This is the ultimate goal of most preventive interventions.

Replicability in an adaptive intervention is closely linked to fidelity of implementation of decision rules. When the decision rules in an adaptive intervention are not well implemented, there is a resulting reduction in replicability. This is because it is possible to attribute the obtained results to factors other than the intervention. These factors are called alternative explanations or confounders. Alternative explanations stem from unknown or known reasons for implementation infidelity. For example, suppose program staff of an adaptive intervention sometimes use considerations other than the established decision rules to make dosage assignments. Then treatment–control differences (or lack thereof) may be in part attributable to any undocumented and unplanned procedures followed by program staff, rather than to the intervention. Unless all program staff in all other implementations of the intervention will make use of these same considerations, the results obtained in this study are not replicable. The principles outlined in the present paper can be used to establish clear definitions of fidelity, thereby helping researchers to encourage and maintain implementation fidelity and, by extension, replicability.

If the intervention is not time-varying, standard methods can be used to adjust for known confounders, although even then post hoc statistical methods are rarely as effective as using the appropriate research design at the outset (Winship & Morgan, 1999). However, most adaptive interventions are time-varying. At this writing, statistical methods for dealing with confounders in time-varying interventions are just being developed and are not yet available for every situation that may arise. Indeed, there is presently a great deal of controversy in the statistical field and elsewhere concerning the appropriate adjustment for confounding when an intervention is time-varying. This lends added importance to design-related decisions that can minimize the impact of confounders. We discuss the statistical analysis of well-implemented adaptive interventions later in this paper.

IDENTIFICATION AND MEASUREMENT OF TAILORING VARIABLES

Identification of Adaptive Components and Tailoring Variables

Most preventive interventions are based on a model that identifies key risk and protective factors and developmental processes associated with the maladaptive outcome they target. The preventive intervention includes individual components designed to impact different critical risk and protective factors based on this model. When investigators are considering an adaptive intervention, they identify which treatment components are to be delivered at the same dosage to all participants, and which ones are to be

[6]Note that because of sampling variability, this does not mean that the same estimate of the population-level treatment effect will be obtained in each sample. Rather, replicability implies that the same population parameter is being estimated in each sample.

delivered adaptively. As discussed above, adaptive components should be considered when the effects of a fixed intervention are expected to vary significantly for individuals who differ on certain characteristics. The identification of key individual (or group) characteristics that would be associated with different responses to treatment outcome in a fixed intervention, and that can serve as tailoring variables, is an important factor leading to a strong adaptive intervention.

In Fast Track, parental functioning was expected to moderate the effect of home visiting, and child reading skill was expected to moderate the impact of tutoring. Each family could be assigned low (monthly), medium (biweekly), or high (weekly) levels of home visits per semester. Dose recommendations were based upon ratings of parental functioning, which included empirically validated risk factors for child aggression (e.g., parent–child conflict, harsh and punitive discipline, maternal depression, family instability, and problematic home–school involvement). It was expected that for families with many of these problems, a high level of home visiting was needed to promote positive intervention effects. In contrast, for families with few of these problems, a low level of home visits would be sufficient to promote positive child outcomes, and higher levels might have a negative impact (e.g., stigmatizing families, reducing parent self-efficacy, fostering dependence on home visits for solving everyday problems). An additional risk was that families might feel burdened by home visits they felt were excessive and intrusive, fueling resentment of the program and reducing participation in other intervention components, thereby reducing overall intervention effects. Hence, the optimal impact of intervention was expected when the level of home visits was tailored to family need, avoiding the potential loss of intervention effects associated with either insufficient or excessive home visiting.

Students whose academic performance was above the 33rd percentile were expected to benefit little from the reading tutoring intervention in Fast Track. Thus tutoring was not viewed as necessary for them to achieve the protective factor of grade-level reading ability. Moreover, if tutoring removed them from class, the missed class time could put them at risk for developing problems in another subject. In contrast, it was anticipated that students with academic performance below the 33rd percentile would require this tutoring to improve their reading abilities and thereby benefit from classroom instruction.

Tailoring Variables as Moderators, Mediators, and Outcomes

Depending on the context or the occasion, a tailoring variable may serve as a moderator, a mediator, a short-term outcome, or even the ultimate outcome of interest. As discussed above, a variable is selected as a tailoring variable in an adaptive intervention precisely because it would moderate treatment in a hypothetical fixed intervention. A tailoring variable that represents malleable characteristics of the individual, as opposed to more or less immutable characteristics of the individual such as gender, ethnicity, status as a dyslexic, and so on, may play several roles in an adaptive intervention. For example, in Fast Track parental functioning played the role of a tailoring variable, because it was expected to moderate the impact of home visiting. However, because the theoretical model underlying the Fast Track study also specified that the intervention would improve parental functioning, and thereby prevent antisocial behavior, parental functioning also played the role of a mediator. In addition, because parental functioning was a proximal target of the intervention, it served as a short-term outcome. In other studies, a pretreatment measure of the primary outcome variable may serve as a tailoring variable. For example, a program to prevent adolescent drug use may vary certain components depending upon each adolescent's pretreatment experience with substance use. In this case, drug use behavior measured before the intervention is a tailoring variable; measured after the intervention, it is an outcome. In time-varying interventions, the same variable may play tailoring variable, mediator, and outcome roles at different times.

Measurement of Tailoring Variables

Every dosage assignment decision made about an individual in an adaptive preventive intervention begins with the individual's value on the relevant tailoring variable. To the extent that the tailoring variable is measured well, the appropriate dose of the intervention will be assigned; to the extent that the tailoring variable is measured poorly, it is possible that inappropriate doses will be assigned, resulting in an ineffective intervention. Thus the quality of the measurement of tailoring variables in an adaptive intervention is critical.

In the behavioral sciences, measurement instruments are usually evaluated by two criteria, reliability and validity. Reliability is the amount of variance in the instrument that is not due to random error, in other words, the amount of "signal" as opposed to "noise." Unreliability in the measurement of tailoring variables introduces random error into the dosage assignment decision, and thus into the intervention itself. If an instrument measuring a tailoring variable is highly unreliable, that is, a large proportion of its variance is attributable to random error, then individuals will be assigned dosages unsystematically. To take this to its logical extreme, if a tailoring variable is measured completely unreliably, any resulting dosage decision is effectively random. Thus the presence of random error in measurement of tailoring variables greatly reduces the strength of the adaptive intervention, because of the imprecision that is introduced into the dosage decision. Validity is the extent to which an instrument measures the attribute it is employed to measure (McDonald, 1999), in other words, the extent to which the instrument is unbiased. Invalid measurement of tailoring variables in an adaptive intervention can systematically point to an inappropriate dosage under some circumstances, depending on the type of bias. Although biased measurement of tailoring variables does not introduce random error into the dosage decision in the same way that unreliable measurement does, if biased measurement results in inappropriate dosage the treatment effect will be weakened, or in extreme cases can even be negative.

DERIVATION OF DECISION RULES

Characteristics of Good Decision Rules

Decision rules form the basis for assigning the appropriate dose or type of each intervention component to each participant, based on that participant's values on relevant tailoring variables. With effective decision rules, components of the intervention are delivered in the intended intensity to the intended individuals. With ineffective decision rules, some individuals will receive an inappropriate dosage of some components, or possibly even an inappropriate treatment. Thus ineffective decision rules reduce the effectiveness of adaptive interventions.

Good decision rules have three important characteristics. First, they are based on an accurate model of the relations among tailoring variables, treatment dosage, and outcome. Thus a clear and thoughtful articulation of this model is very important. Second, good decision rules are objective. They clearly operationalize the dosage to be given and the value (or range of values) on the measure of the tailoring variables. For example, a decision rule that states "poor readers will receive reading tutoring" is not sufficient; a better decision rule states "readers who are in the thirty-third percentile or below in reading will receive three hours per week of reading tutoring." Third, good decision rules are comprehensive, covering anticipated situations that can occur in practice, including situations where the measure of the tailoring variable is missing or ambiguous.

Articulating the Model Relating Tailoring Variables, Dosage, and Outcome

Articulating a theoretical model of exactly how the effect of a particular treatment is expected to differ across values of a tailoring variable is an important first step in deriving decision rules. As discussed above, the philosophy underlying adaptive interventions is that a given treatment will not have the same effect for all individuals. Instead, individuals with certain characteristics on a tailoring variable will enjoy a more beneficial treatment effect, or suffer a less beneficial effect, than individuals with other characteristics. Another way to think of this is that in order to achieve a particular desired treatment effect, different dosages or types of treatment may be needed for different individuals. Here treatment effect is defined as the difference between two hypothetical outcomes for the same individual: the outcome expected if the individual were assigned to the intervention condition, and the outcome expected if the individual were assigned to the control condition. (For purposes of this exercise, we assume that the effects of any other treatment components are constant, so we assign their effects a value of zero without loss of generality.)

In a complete model the expected relation among tailoring variable, treatment, and outcome is expressed for all values of the tailoring variable, all dosages and types of treatments under consideration, and for treatment effects on all important outcome variables. The purpose of articulating this model is to identify for which dosages and types of treatment the effect is optimized, for a given value of the tailoring variable, in order to provide a scientific rationale for the decision rules. The idea of articulating the full range of the relation among tailoring variables, treatment, and outcome will be daunting in many situations. However, the more accurately this relation is represented, the greater is the potential for efficacious intervention. A prevention scientist seriously considering an adaptive intervention will usually have sufficient background to produce an intelligent and informed, if imperfect, model. Prior research may be especially valuable in articulating the model, but it is not the only appropriate source of information, or even necessarily the best in all cases; other sources likely to be helpful in articulating the model are scientific theory in the area and prior clinical or prevention experience. In most instances, the task will require gathering any and all available information, assembling the research team and clinical staff, and carefully thinking through and discussing "If we were to give this dosage to people with this characteristic on the tailoring variable, what treatment effect would be expected?"

As an example, consider the reading tutoring treatment and home visiting treatment components in the Fast Track intervention. The reading tutoring treatment has only two possibilities, no tutoring and tutoring. The tailoring variable under consideration in this case is reading ability, which for purposes of dosage assignment has only two levels, below average and at or above average. We will consider the possible treatment effect on one example outcome variable in the Fast Track intervention, risk of special education placement. Of course, for those assigned to the no tutoring condition risk of special education placement is unaffected, whether or not they are below-average readers. There is also no change in risk of special education placement when reading tutoring is given to students who are average or above-average readers, because they do not need additional reading skills to keep up in their classes. However, reading tutoring is expected to result in a pronounced decrease in risk of special education placement for students who are below-average readers. Based on this expected relation among reading ability, treatment, and treatment effect, for average or above-average readers the treatment effect is optimized if no reading tutoring is assigned, whereas for below-average readers the treatment effect is optimized if tutoring is assigned. Now consider the model underlying the home visiting treatment component in Fast Track. In this case the tailoring variable being considered is parental functioning, assessed at three levels: high, medium, and low. The outcome variable we will discuss is teacher ratings of oppositional-aggressive behavior. For high-functioning at-risk parents, the optimal dose is monthly home visits. Biweekly and weekly home visits are associated with smaller treatment effects than monthly home visits for this group, because of the reactivity associated with giving a larger dose than is needed. For medium-functioning parents the best effect is expected with biweekly home visits, whereas a dose of weekly home visits is associated with the worst treatment effect; the effect for monthly home visits falls between the effect sizes for weekly and biweekly visits, because this dose is insufficient. For low-functioning parents, who are most in need of the intervention, the best effect is associated with weekly home visits, and the worst effect is associated with monthly home visits. Researchers considering adaptive interventions may find it a useful heuristic to sketch hypothetical relations such as these in graph form.

In reality, in most prevention studies there are several tailoring variables to consider. In addition, the adaptive intervention may be time-varying, which means that the history of previous dosage may also affect the outcome. Although it is possible to incorporate interactions among tailoring variables and effects of prior treatments, this represents a significant increase in complexity. We suggest that unless these effects are expected to be strong, examining tailoring variables one at a time and ignoring the effects of prior treatments is an acceptable tradeoff. For those who desire more precision, we recommend looking into research in statistics, medicine, computer science, and other fields, where determining the optimal decision rule is currently an extremely active area. For example, see Shachter (1986), Owens *et al.* (1997), Murphy (2003), Bather (2000), Cowell *et al.* (1999), and Bertsekas and Tsitsiklis (1996). In particular, the last four consider the highly complex time-varying setting, where interactions among tailoring variables and effects of prior decisions are incorporated.

Clinical Judgment

One particularly important issue to consider is the extent to which and ways in which clinical judgment will be used to contribute to dosage assignment. Often adaptive interventions are carried out by staff with extensive clinical training. Their expertise can be used in several ways. One role clinical judgment can play is to provide input into the development of the assessment package, broadly defined. Given that selection and measurement of tailoring variables and the decision rules regarding dosage assignment are part of the preventive intervention and will affect intervention impact, including the input of participating clinicians with appropriate expertise when specifying these prevention program parameters should strengthen the utility and practical applicability of the design. A second potential role for clinical judgment involves the participation of intervention staff in the assessment of the tailoring variables. For example, in the Fast Track Project, the prevention staff who conducted home visits also completed rating scales assessing parental functioning at regular intervals, which served as a basis for prescribing levels of home visiting. The decision to use prevention staff to assess parental functioning hinged on the belief that the expertise of these staff, along with their intimate knowledge of family contexts and family interaction patterns, placed them in a better position to provide valid assessments of variations in parental functioning than other alternatives. In a situation such as this, the reliability, validity, and relative superiority of staff ratings compared to other forms of measurement of tailoring variables could be evaluated empirically, developing clear guidelines over time regarding the degree to which clinical judgment should be included in the assessment of tailoring variables.

A third potential role for clinical judgment involves contributions to the decision process regarding dosage assignment. In this case, prevention staff are provided with latitude in the application of decision rules regarding dosage, using their expertise to make case-by-case decisions regarding the assignment of dose levels (Bierman *et al.*, 2001). Although using clinical judgment to inform dosage decisions in this way may seem useful from a clinical standpoint, it is important to consider that this procedure renders clinical judgment a part of the decision rules, and therefore a part of the overall treatment. At the same time, previous research raises serious questions about the reliability and validity of clinical judgment when compared to actuarial procedures in diagnostic assessment, when clinicians are asked to weigh multiple pieces of assessment information and make categorical decisions (Breslin *et al.*, 1997; Daws *et al.*, l989). This suggests that allowing clinicians to invoke judgment in the application of decision rules may introduce idiosyncratic variability in the application of these rules, threatening the replicability and validity of the prevention program. In addition, poor judgments by clinical staff can reduce intervention effectiveness. For example, one study found that allowing clinicians to individualize treatment packages for phobic patients was not beneficial relative to a fixed intervention and, in fact, was detrimental if clinicians chose to omit a particularly powerful component of treatment (Schulte *et al.*, 1992).

IMPLEMENTATION OF DECISION RULES

The final link in the chain constituting an adaptive intervention is the optimal implementation of the decision rules. The optimal way to implement decision rules is universally, in other words, to apply them consistently across study participants, time, implementation site, staff member, and every other set of circumstances, so that every dosage decision is made using identical rules, and the decision rules are applied identically to any participant with the same values on the tailoring variables. In an optimal intervention design, decision rules are established before the intervention begins, so that there is no variability or "drift" in how they are carried out as a study progresses. When decision rules are optimally implemented there are no changes or exceptions made on an ad hoc basis. In suboptimal implementation of decision rules, some persons are treated differently from others, because the dosage assignment is based in part on factors that do not figure in the decision rules and may be unique to a certain individual, time, or situation. Suboptimal implementation of decision rules can introduce random error into the preventive intervention, thereby lessening its effectiveness. It also can harm replicability by introducing confounders into the experimental comparison of the preventive intervention with other conditions.

Suboptimal implementation may occur because the clinical staff perceive that the decision rules are inappropriate or less appropriate in a particular case due to some extenuating circumstances, such as when a caseworker argues that a particular family needs weekly home visits, despite ratings of parental functioning that identify monthly visits as the appropriate dose. This may occur for several reasons. First, perhaps in addition to parental functioning, there are important moderators that would affect the impact of different levels of home visiting that have not been identified, and should be included as tailoring variables in a revised set of decision rules. A second alternative is that the measure of the tailoring variable (in this case the measure of parental functioning) may lack reliability or validity, and the caseworker perceives that the measure is not doing a good job of identifying the appropriate dosage. A third possibility is that the decision rules are stated ambiguously, or due to insufficient training or supervision, the caseworker lacks a clear understanding or acceptance of the rationale for the decision rules, and is advocating for more home visiting for the family on the basis of a personal belief that it is best for the family. In this case, the solution may involve additional staff training and supervision, or clarification of the rules, rather than any fundamental changes to the decision rules themselves.

When deviations from the decision rules are allowed, the project (and field) will benefit most if this is done systematically in ways that assist with the interpretation of the prevention trial, as well as help to inform models for future trials. One way to do this is for the research project to hold regular meetings of the scientific and clinical staffs, on an ongoing basis, for the express purpose of reviewing every dosage decision where there is any question about following the decision rules to the letter, where the dosage decision is ambiguous, or where there is conflicting or incomplete information. In many cases, this will lead to clarification of the decision rules. In some cases, it may be necessary to make an exception to the rules. If a careful log of such cases is kept, including a detailed explanation of why an exception was made, this information can be used to describe the implemented treatment with the aim of maintaining replicability, by using it to make sure that the same procedure is followed in any future implementations of the intervention. Furthermore, the information in this log will be helpful in fine-tuning the decision rules for future studies. However, to the extent that individuals with the same values on the tailoring variable are assigned dosages by relying on ad hoc procedures rather than the established decision rules, there will still be problems with replicability. The log will help to assess the extent of the problem, and possibly to prevent it in the future, but will not help to ameliorate it in the current study.

THE STATISTICAL ANALYSIS OF AN ADAPTIVE INTERVENTION: GENERAL CONSIDERATIONS

Despite the differences between adaptive preventive interventions and fixed preventive interventions, the approach to conducting a scientific evaluation of adaptive interventions is essentially the same as the approach required for fixed preventive interventions. Design considerations that apply to fixed preventive interventions, such as sampling, control groups, assignment to treatment, the use of multiple cohorts, statistical power, the timing and spacing of observation in a longitudinal study, and so on apply to adaptive interventions as well. Furthermore, the types of questions that can be answered when evaluating an adaptive intervention are similar to the types of questions that can be answered using a fixed preventive intervention. Generally, researchers want an answer to the question, "Is the preventive intervention reducing risk?" The term "risk" must be defined in relation to some baseline comparison group. The comparison could be made against a classic no-treatment (or treatment-as-usual) control in order to address the question, "Is the adaptive intervention reducing risk, as compared with no treatment (or treatment as usual)?" In some cases, the question of interest may be instead, "Is the adaptive intervention reducing risk, as compared with a fixed intervention?" Such a question calls for a design that includes a comparison condition receiving a fixed intervention, instead of, or in addition to, a no-treatment control.

Just as in fixed intervention evaluations, when evaluating adaptive interventions it is important to maintain random assignment to treatment conditions. Thus with adaptive interventions, assignment to treatment conditions means assignment to a particular "bundle" of treatment components, tailoring variables, measures of tailoring variables, and decision rules. If there is random assignment, we can compare the distribution of responses between treatment conditions and between any one treatment condition and control. For example, we may be interested in how the average level of conduct disorder at a particular age differs between treatment conditions or between a treatment condition and control. A more sophisticated analysis might use growth models to compare rate of growth of school behavior problems between treatments or between any one treatment group and the control group. Or we may compare the timing of conduct disorder milestones between treatment conditions and between a treatment condition and control condition, via survival analysis methods. In all of these cases, randomization to the treatment/control conditions assures us that the types of subjects are balanced between the different treatment and control conditions. Additionally, implementation fidelity to the treatment within a condition implies that such comparisons represent valid, unbiased estimates of the "planned treatment" effect. When there is infidelity in implementation then such comparisons result in estimates of the "implemented treatment" effect. Such estimates usually do not address the question that the researchers originally posed, and thus are less interesting. Methodology for assessing the "planned treatment" effect of adaptive interventions when there is implementation infidelity is in its infancy. Murphy *et al.* (2001; see also Robins, 1986, 1989, 1993, 1997) illustrate a method that reweights the individuals' responses, assigning higher weight to the individuals whose treatment patterns more closely follow the decision rules.

If the statistical analysis fails to show a benefit attributable to an adaptive preventive intervention, there are numerous alternative possibilities for what went wrong. Of course, one or more components of the preventive treatment itself could be ineffective. However, it could be that the problem was with the selection of tailoring variables; with the measures of the tailoring variables, which may have been so unreliable that dosage was essentially random, or invalid in a way that rendered it counterproductive; or it could be that the decision rules were based on ill-conceived utility curves, or were not thought out completely; or it could be that the decision rules were good, but fidelity to the implementation was not maintained. If any of these problems exist, it is possible that the intervention would have been successful using the same treatment with more reliable and valid measurement of tailoring variables, or better decision rules and/or decision rule implementation. Post hoc analyses may shed light on some of these alternative hypotheses. For example, the reliability and validity of the measures of the tailoring variables can be examined after the fact. Furthermore, if careful records of implementation have been kept, it may be possible to sort out where implementation followed the prescribed decision rules and where it deviated from these rules, and examine treatment effects controlling for fidelity to the decision rules.

Mediation models (e.g., Collins *et al.*, 1998; Kenny *et al.*, 1998) are important in prevention science. Most preventive interventions operate by changing mediating variables that ultimately operate on the outcome variables. As discussed above, many of the variables that serve as tailoring variables in an adaptive intervention also play the role of mediators of the treatment. For example, reading ability was used as a tailoring variable in the Fast Track study, and also was expected to be a mediator of the treatment (i.e., assignment to treatment condition leads to improved reading ability, which in turn leads to reduced risk of special education placement). Mediation models can be fit in the usual way using data from adaptive interventions.

Although it is straightforward to address the overall question of "Is the intervention efficacious?" with an adaptive intervention, absent additional assumptions it is not possible to assess dosage response or to isolate the effects of one component within a single adaptive intervention treatment condition. (Note that this is also true within a single fixed intervention treatment condition!) It follows that dose cannot be used as a moderator in these designs. It is true that within an adaptive intervention treatment condition dosage is varied across individuals, in terms of how much of a particular treatment component an individual receives, and even whether the individual receives any of that component at all. But there are systematic preexisting differences between individuals in different dosage levels or dosage groups within one treatment condition. Thus if assessing the effects of different dosages, or the effects of individual components, is desired, it is best to build this into the design a priori. This can be done by creating conditions where dosage or exposure to select components is assigned randomly. For example, if the Fast Track study had wished to evaluate the incremental effect of the home visitation component, it could have continued with an adaptive intervention, but included a condition where no home visitation was offered. Then there would have been random assignment to one of three conditions: full intervention; intervention with no possibility of home visitation; and control. Similarly, if the relationship between dose and response on a particular intervention component is of interest, this can be assessed by creating conditions where dosage is systematically varied, and randomly assigning to these conditions.

FUTURE RESEARCH DIRECTIONS AND OPEN QUESTIONS

Methodological research relevant to adaptive interventions is currently at an early stage. We argue that a productive direction for future research would be to investigate the potential advantages of adaptive interventions as compared to fixed interventions. In general, an adaptive intervention has the potential to be an attractive alternative to a fixed intervention *if in a comparable fixed intervention, significant variation in treatment effects would be expected as a function of identifiable tailoring variables*, *across participants and/or within participants over time*; in other words, whenever identifiable characteristics of the individual would influence the effects of certain components of the intervention if the intervention were fixed. Conversely, if there is no reason to believe that treatment effects vary systematically across individuals to any appreciable degree, it is unlikely that an adaptive intervention presents advantages over a fixed intervention.

We see four potential specific advantages of adaptive interventions. First, a properly conducted adaptive intervention may *reduce negative effects* associated with doses of a treatment component that are inappropriate for certain individuals. An adaptive intervention is particularly attractive if individuals with certain values on a tailoring variable are expected to experience a negative treatment effect for a particular treatment dosage or type, even if the identical treatment is expected to produce a positive effect for individuals with other values on the tailoring variable. For example, for a high risk individual a large dose of a certain treatment may be beneficial. But an individual at low risk who is given the same dose may become bored or restless, or stigmatized by others, or even develop a negative self-image.

Second, adaptive interventions may *reduce waste.* At first glance, it may seem that the distinguishing characteristic of an adaptive intervention is the objective of ensuring that each individual receives the intervention he or she needs in order to optimize the treatment effect. However, fixed interventions also have this objective; they pursue it by delivering a uniform dose to all participants, chosen so as to be sufficiently large for most participants, even if for some individuals it exceeds the needed dose. Rather, the distinguishing characteristic of an adaptive intervention is the objective of ensuring that each individual receives an intervention dose sufficiently large to meet his or her needs, *but no larger*. Thus adaptive interventions can make better use of available resources by distributing them among intervention participants according to potential effectiveness. For example, if a finite number of hours per week of staff time is available to provide family counseling, the overall effectiveness of an intervention may be improved by devoting a larger proportion of those hours to the families in worst need of counseling.

Third, adaptive interventions can potentially *increase compliance* if, due to the individualization of treatment, recipients feel more comfortable with their level of participation. In fixed interventions, low-risk individuals may drop out if they feel they are wasting time participating at a level well beyond what they need, and high-risk individuals may drop out if they feel the dose they are receiving is insufficient. In a carefully formulated adaptive intervention, the probability of both of these events can be reduced.

Fourth, adaptive interventions may *enhance the potency* of the intervention as compared to a comparable fixed intervention, if they increase salience, eliminate negative effects, improve compliance, and/or the reduced waste makes it possible to devote additional resources to higher-risk individuals who can benefit from them. This is particularly the case when there are certain intervention components beneficial to some individuals but potentially harmful to others. For example, some individuals might benefit from additional individualized counseling, but if an attempt was made to administer this to all participants, those who do not need or want such counseling would soon lodge a complaint or drop out of the study. This might deprive these individuals of the benefits of other aspects of the intervention. In interventions that involve providing health information, Kreuter *et al.*, (1999) found that content tailored so as to provide only selected, personally relevant information tends to be attended to, thoughtfully processed and thus more efficacious. When an adaptive approach successfully enhances the potency of an intervention, the enhanced potency translates directly to a larger effect size and improved statistical power.

There are several important open questions related to adaptive interventions. A challenging and contentious issue is whether and where clinical judgement should be incorporated into the dosage assignment process. Should the role of clinical judgement be primarily in the formulation of the decision rules, or should the decision rules explicitly allow for clinical judgement in their execution? The dilemma for today's prevention scientist is that, given the current state of the art, most systems of decision rules are incomplete, and clinical judgement may be needed to make dosage decisions when a situation is confronted that is inadequately addressed by the rules. The choice of whether to incorporate clinical input into dosage assignment decisions in adaptive interventions involves a tradeoff between the value of the clinical information on the one hand and the potential threats to replicability and validity on the other. In our opinion, the reason for including clinical judgment at this level is to compensate for gaps in our science–gaps that will, it is hoped, become filled in future investigations. The ultimate goal is to use clinical judgment to refine and extend the reach of the decision rules, but to attain, in a finalized prevention program, decision rules that are comprehensive, clinically sensitive, and inviolable, avoiding the threats to validity introduced by allowing for variations in rule application determined by prevention staff. Little research exists examining the role of clinical judgement in adaptive interventions (but see Breslin *et al.*, 1997); more is needed.

Research is needed on how to improve the design and analysis of adaptive treatment trials. For example, how can we identify powerful tailoring variables? Results from similar fixed interventions conducted in the past may shed light on likely moderators of treatment. Information from prior studies where the intervention failed with a particular subgroup, or with individuals with particular characteristics, is likely to be highly valuable. Where a large body of literature is available, it may even be possible to conduct a meta-analysis in order to synthesize the evidence on potential tailoring variables. A related set of questions concerns measurement of tailoring variables. In many studies, tailoring variables will also play the role of mediator or intermediate outcome, or may be variables that are of considerable theoretical interest quite apart from their role in the dosage decision. In such studies it is likely that the researchers have already thought through how best to measure the tailoring variables. However, given the paucity of empirical work focused on the assessment of tailoring variables, making decisions regarding their measurement is frequently challenging. In some cases, well-established measures of important tailoring variables may not be available, and research is needed to develop reliable and valid measurement instruments. In other cases, measures validated in a research framework may need to be transformed or otherwise adapted for use in the context of clinical assessment for intervention planning. Furthermore, in time-varying adaptive interventions the tailoring variables must be assessed periodically. This raises issues such as practice effects, and the possibility of subject reactivity to the measurement.

Another area where research is needed is in the articulation of the relation among tailoring variable, treatment, and outcome. If time and resources are available, it may be possible to conduct an empirical study for the purpose of identifying optimal dosages. This could be done by randomly assigning individuals to different treatment dosages. As more scientists articulate these relations, perhaps information will accumulate that will be helpful to the prevention science community at large, including new procedures for establishing the relations empirically. Research is also needed on methods for estimating decisions rules that will optimize response (e.g., Murphy, 2003).

Little is currently known about the cost/benefit ratio of adaptive versus fixed interventions. As discussed above, adaptive interventions can potentially result in savings of resources as compared to fixed interventions. However, the extent of these savings in practice is unclear. It is also unclear the extent to which any savings are offset by resource demands in other areas. These demands can be divided conceptually into nonrecurring up-front costs and ongoing costs. One example of a nonrecurring up-front cost is the articulation of the expected relation among tailoring variable, treatment, and outcome. This is likely to require considerable staff time in the early stages of development of an intervention. Another possible up-front cost is that certain adaptive interventions may require more intensive staff training than comparable fixed interventions. Because fidelity to the decision rules is critical, monitoring and record keeping related to the dosage assignment procedure as part of the implementation process monitoring is important. Of course, careful advance planning, staff training, and process monitoring are equally important in high-quality fixed and adaptive interventions. We suspect that the cost/benefit ratio is advantageous in the long run for adaptive interventions, particularly after the nonrecurring up-front costs have been incurred. However, this is an empirical question.

Another open question is the extent to which fidelity of an adaptive intervention can be maintained after the intervention is handed over to an implementation in a community setting. It could be argued that the added complexity of adaptive interventions puts them at increased risk for implementation problems that can undermine effectiveness. The implementation of decision rules is probably the area in an adaptive intervention that is most at risk in an effectiveness trial. Adaptive interventions share this implementation issue with manualized behavioral therapies. If the manual, or in this case, the decision rules are not followed to the letter because of insufficient training, perceived inadequacy of the rules, resource limitations curtailing certain doses, or any other reason, an efficacious adaptive intervention may be undermined. However, it could also be argued that community personnel will resonate better to adaptive intervention delivery procedures than they will to fixed intervention procedures, because adaptive interventions resemble sensible clinical practice and also husband valuable intervention resources. If this is so, community personnel may follow adaptive intervention procedures more closely than they will fixed intervention procedures. Again, this is an empirical question, which, like many questions concerning implementation fidelity in general, deserves further study.

ACKNOWLEDGMENTS

The order of the first two authors is alphabetical. We wish to thank the following individuals for helpful contributions to this paper: Richard Catalano, John Coie, Kenneth Dodge, John W. Graham, Mark Greenberg, John Lochman, James McKay, Robert McMahon, Robert Nix, Ellen Pinderhughes, and two anonymous reviewers. Preparation of this paper was supported by National Institute on Drug Abuse grant P50 DA10075, Center for Prevention Methodology (Dr Collins and Dr Murphy), and by a Career Scientist Award K02-DA15674 (Dr Murphy).

REFERENCES

Bather, J. A. (2000). *Decision theory: An introduction to dynamic programming and sequential decisions*. New York: Wiley.

Bertsekas, D. P., & Tsitsiklis, J. N. (1996). *Neuro-dynamic programming.* Belmont, MA: Athena Scientific.

Bierman, K. L., Nix, R. L., Maples, J. J., Murphy, S. A., & Conduct Problems Prevention Research Group. (2004). *Evaluating the use of clinical judgment in the context of an adaptive intervention design: The Fast Track Prevention Program.* Manuscript submitted for publication.

Borhani, N. O., Applegate, W. B., Cutler, J. A., Davis, B. R., Furberg, C. D., Lakatos, E., Page, L., Perry, H. M., Smith, W. M., & Probstfield, J. L. (1991). Systolic hypertension in the elderly program (SHEP). Part 1: Rationale and design. *Hypertension*, *17*(Suppl II), 2–15.

Breslin, F., Sobell, M. B., Sobell, L. C., Buchan, G., & Cunningham, J. A. (1997). Toward a stepped care approach to treating problem drinkers: The predictive utility of within-treatment variables and therapist prognostic ratings. *Addiction*, *92,* 1479–1489.

Breslin, F., Sobell, M. B., Sobell, L. C., Cunningham, J. A., Sdao-Jarvie, K., & Borsoi, D. (1999). Problem drinkers: Evaluation of a stepped-care approach. *Journal of Substance Abuse*, *10,* 217–232.

Brooner, R. K., & Kidorf, M. (2002). Using behavioral reinforcement to improve methadone treatment participation. *Science and Practice Prespectives*, *1*(1), 38–46.

Collins, L. M., Graham, J. W., & Flaherty, B. P. (1998). An alternative framework for defining mediation. *Multivariate Behavioral Research*, *33,* 295–312.

Conduct Problems Prevention Research Group. (1992). A developmental and clinical model for the prevention of conduct disorders: The Fast Track Program. *Development and Psychopathology*, *4,* 509–528.

Conduct Problems Prevention Research Group. (1999a). Initial impact of the Fast Track prevention trial for conduct problems: I. The high-risk sample. *Journal of Consulting and Clinical Psychology*, *67,* 631–647.

Conduct Problems Prevention Research Group. (1999b). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology*, *67,* 648–657.

Cowell, R. G., Dawid, A. P., Lauritzen, S. L., & Spiegelhalter, D. J. (1999). *Probabilistic networks and expert systems*. New York: Springer.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243,* 1668–1674.

Dishion, T. J., & Kavanagh, K. (2000). A multilevel approach to family-centered prevention in schools. *Addictive Behaviors*, *25*, 899–911.

Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, *15,* 451–474.

Kenny, D. A, Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp. 233–265). Boston, MA McGraw–Hill.

Kreuter, M. W., & Strecher, V. J. (1996). Do tailored behavior change messages enhance the effectiveness of health risk appraisal? Results from a randomized trial. *Health Education Research*, *11,* 97–105.

Kreuter, M. W., Strecher, V. J., & Glassman, B. (1999). One size does not fit all: the case for tailoring print materials. *Annals of Behavioral Medicine*, *21,* 276–283.

Lavori, P. W., & Dawson, R. (1998). Developing and comparing treatment strategies: An annotated portfolio of designs. *Psychopharmacology Bulletin*, *34,* 13–18.

Lavori, P. W., Dawson, R., & Rush, A. J. (2000). Flexible treatment strategies in chronic disease: Clinical and research implications. *Biological Psychiatry*, *48,* 605–614.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Murphy, S. A. (2003). Optimal dynamic treatment regimes (with discussion). *Journal of the Royal Statistical Society, Series B*, *65,* 331–366.

Murphy, S. A., van der Laan, M. J., Robins, J. M., & Conduct Problems Prevention Research Group. (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association*, *96,* 1410–1423.

Owens, D. K., Shachter, R. D., & Nease, R. F. (1997). Representation and analysis of medical decision problems with influence diagrams. *Medical Decision Making*, *17,* 241–262.

Prochaska, J. O., Velicer, W. F., Fava, J. L., Rossi, J. S., & Tsoh, J.Y. (2001). Evaluating a population-based recruitment approach and a stage-based expert system intervention for smoking cessation. *Addictive Behaviors*, *26,* 583–602.

Robins, J. M. (1986). A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect. *Computers and Mathematics with Applications*, *14,* 1393–1512.

Robins, J. M. (1989). The analysis of randomized and nonrandomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In L. Sechrest, H. Freeman, & A. Mulley (Eds.), *Health service research methodology: A focus on AIDS* (pp. 113–159). NCHSR, U.S. Public Health Service, Washington, DC.

Robins, J. M. (1993). Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. *Proceedings of the Biopharmaceutical Section, American Statistical Association*, 24–33.

Robins, J. M. (1997). Causal inference from complex longitudinal data. In M. Berkane (Ed.), *Latent variable modeling and applications to causality: Lecture notes in statistics* (pp. 69–117). New York: Springer.

Schulte, D., Kunzel, R. l., Pepping, G., & Schulte-Bahrenberg, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behavior Research and Therapy*, *14,* 67–92.

Shachter, R. D. (1986). Evaluating influence diagrams. *Operations Research*, *34,* 871–882.

Sobel, M. B., & Sobell, L. C. (1999). Stepped care for alcohol problems: An efficient method for planning and delivering clinical services. In: J. A. Tucker, D. M. Donovan, & G. A. Marlatt (Eds.), *Changing addictive behavior: Bridging clinical and public health strategies* (pp. 331–343). New York: Guilford Press.

Sobel, M. B., & Sobell, L. C. (2000). Stepped care as a heuristic approach to the treatment of alcohol problems. *Journal of Consulting and Clinical Psychology*, *68,* 573–579.

Weissberg, R. P., & Greenberg, M. T. (1998). Prevention science and collaborative community action research: Combining the best from both perspectives. *Journal of Mental Health*, *7,* 479–492.

Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational longitudinal data. *Annual Review of Sociology*, *25,* 659–706.