FROM: Longabaugh R, Wirtz PW (eds.). Project MATCH Hypotheses: Results and Causal Chain Analyses. Project MATCH Monograph Series v. 8 (NIH Pub. No. 01-4238). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism (NIAAA), 2001.

Part VII: Conclusions and Implications

Methodological Assessment and Critique

Substantive Review and Critique

Methodological Assessment and Critique

Philip W. Wirtz, Ph.D., and Richard Longabaugh, Ed.D.

ABSTRACT

The preceding chapters have as a common goal the identification of what went right and what went wrong in the theories leading to the matching hypotheses. The authors of these chapters took a number of distinctly different approaches in meeting this common goal, with varying levels of success. In the present chapter, we review some of these alternative approaches with an eye toward identifying the strengths and weaknesses of each. In the process, we address some of the difficulties inherent in the shift from testing the mediation of main effects to testing the mediation of moderator (or interaction) effects (as is a primary objective in matching studies). We consider in this chapter the manner in which some of these difficulties can be overcome and some of the challenges introduced in the causal testing of latent growth models. We conclude with prescriptive guidance which we hope will inform the field as it continues the fruitful pursuit of causal chain analysis.

Testing Causal Chains

The procedure for testing mediation has been formalized by Baron and Kenny (1986) and further explicated by Holmbeck (1997). The procedure is based upon empirical tests of four simple conditions. Following these guidelines, in order to conclude that variable B mediates the relationship between variables A (independent) and C (dependent; see figure 1), four conditions must prevail: (1) A and C must be related in the hypothesized direction, (2) A and B must be related in the hypothesized direction, (3) B must be related to C (in the hypothesized direction) after controlling for A, and (4) the relationship between A and C must be smaller after controlling for B than it is before controlling for B. In practice, the first three conditions require the relationship between the two variables to be directionally statistically significant at some preordained level of α (conventionally, 0.05). Condition 4 is satisfied if the parameter estimate obtained by regressing C on A (controlling for B) is smaller than the parameter estimates obtained by regressing C on A without controlling for B.

Note that the Baron and Kenny formulation is just as important for models in which mediation is hypothesized and is *not* found as it is for models in which mediation is hypothesized and *is* found, because strict adherence to the four steps provides an indication of where the purported causal chain broke down if the hypothesized mediation cannot be empirically verified. Relying on figure 1 again for example, if A is found not to be significantly related to B (and/or if C is found not to be significantly related to B after controlling for A), it provides an immediate indication of the locus of a logical flaw in the hypothesized causal chain. Thus, Baron and

Philip W. Wirtz, Ph.D.
Department of Management Science
George Washington University
2115 G Street NW, #403
Washington, DC 20052
Email: pww@gwu.edu

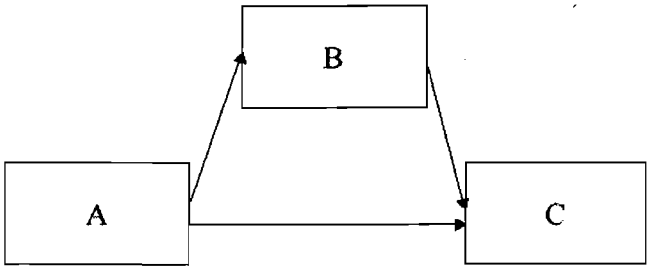

Figure 1. Model of a mediated effect, where B is a purported mediator of the relationship between A and C.

Kenny (1987) and Holmbeck (1997) understate the value of the four-step procedure to the process of causal chain analysis: in addition to providing a rigorous test of whether mediation occurs, these four steps also provide an indication of where the purported model failed when hypothesized mediation is found *not* to occur.

One implication of this four-step procedure is that relationships which might appear to be mediational are not. Consider the case where we know a priori that some factor A is a strong causal factor for C, and it does so through only one mechanism, which is to induce variation in B, with A being the sole cause of B. For example, suppose that there was an exact dose response between amount of smoking in a 6-month period (A) and buildup of plaque on arteries (B). Second, let's assume that there is a strong dose response between the thickness of the plaque in the arteries and resting blood pressure (C), although this is not the only cause of C. Suppose this leads to a correlation between A and B of 1, between B and C of 0.5, and between A and C of 0.5. If we calculate the semipartial correlation of B and C, controlling for A, it will be zero. Initially, this might seem to be an example where full causal mediation is in place, but condition 3 does not hold, because B is not related to C after controlling for A.

The problem here is that the standard that was set in answer to the question of "what constitutes mediation?" is lower than that required by Baron and Kenny. Under this example, Baron and Kenny would insist that before B (plaque buildup in the arteries) is called a mediator, it is necessary to establish that it is a causal factor of C (high blood pressure). When A is the sole causal antecedent of B (as in this example), then while it *might* be the case that B (plaque) is a causal factor of C (high blood pressure), it might alternatively be the case that *smoking* is the true causal factor and that plaque is just "along for the ride". Essentially, Baron and Kenny require that B be established as a true causal factor of C before it is called a mediator, and this is not possible if A and B are perfectly (or very highly) correlated.

To drive the point to an absurdity, suppose that there is a fourth variable (D) which is highly correlated with both A and B in this example: for instance, suppose that D is "percentage of friends who are smokers". Now, suppose we have measured only A, D, and C, and we wish to know if D mediates the relationship between A and C. If condition 3 was merely bivariate and did not control for A, it would allow us to reach the dubious conclusion that percentage of friends who are smokers mediates the relationship between smoking and blood pressure. (This is a dubious conclusion because it is doubtful that having a greater percentage of friends who smoke *causes* high blood pressure.) By controlling for A, condition 3 would eliminate this variable as a potential mediator, because it has not been proven that percentage of friends who are smokers (D) is *causally* related to high blood pressure. In the same way that D is eliminated as a potential mediator by Holmbeck's condition 3 under this model, B would be eliminated as a potential mediator under the model of the original example—and for the same reason.

The difference here is one of *could be* versus *is* a mediator. In the example, plaque buildup *could be* a mediator but we are not sure it is causally linked to high blood pressure, so we do not call it a mediator under the Baron and Kenny standard. Under Baron and Kenny, a higher standard needs to be met in order to establish mediation: there needs to be evidence of a causal link between the putative mediator and the outcome before mediation can be established.

A potential weakness in this four-step approach lies with its rather cavalier treatment of variables outside the model which could influence the outcome. For example, the under-

pinning of the condition-2 requirement that A and B be related lies with the logical assertion that, in order for B to mediate the relationship between A and C, B has to be *caused* by A. However, merely establishing that A and B are related provides a necessary but totally insufficient empirical test for the causal assertion because it does not account for the possibility that A and B are not causally linked but are rather both consequences of some antecedent variable. Thus, a much stronger test of condition 2 than is found in the typical application of the Baron and Kenny formulation would require A and B to be related *after controlling for other variables which might represent alternative explanations for the existence of the relationship.* Similarly, although it is not frequently cited in applications of the Baron and Kenny formulation, condition 3 would be much stronger if, in addition to controlling for A, additional control variables were added to the model which would refute potential threats to internal validity.

A fully comparable alternative approach for testing mediation employs structural equation modeling. Under this approach, the direct A–C effect is initially estimated by omitting B from the model. Following this, a full model containing both the direct (A–C) and indirect (A–B–C) linkages is tested. Mediation occurs when (1) the A–C effect in the initial model is directionally significant, (2) the A–B and (3) B–C effects in the second model are directionally significant, and (4) the A–C effect in the second model is less than the A–C effect in the first model. These conditions are exactly analogous to the Holmbeck (1997) explication of Baron and Kenny's (1986) formulation.

The choice between these two alternatives is often driven by whether the investigator has multiple indicators for A, B, and/or C. The two approaches are fully equivalent when each of the variables is measured by a single indicator. Multiple indicator models dictate the selection of a structural equation modeling procedure as the analytical procedure of choice.

While these two approaches have been widely adopted for testing simple mediation hypotheses, the formal test of a matching causal chain is one level more sophisticated, in that it requires testing for mediation of a *moderator* (i.e., interaction) effect rather than of the *main* effect that is addressed under the Baron and Kenny formulation and by a structural equation modeling approach. The additional challenge imposed by a matching hypothesis is further heightened under a structural equation modeling approach, where detection of interactions involving latent constructs has proven particularly daunting.

Even under the single-indicator Baron and Kenny formulation, the purported causal path can take any of several different forms. The potential for these different forms imposes a considerably greater burden on the investigator, who must explicate the model to be tested in advance of formally testing any causal chain hypotheses. As described subsequently, the application of the Baron and Kenny formulation for single-indicator models to mediated moderation models is relatively straightforward, given that the investigator has adequately explicated the postulated model in advance.

Some authors in this monograph approached the mediated moderation question by testing each treatment condition separately, similar to the "within-groups" approach critiqued by Finney and associates (1984). Such an approach carries the advantage of being arguably simpler to understand and to execute than following the Baron and Kenny four-step formulation. Unfortunately, however, the use of this approach does not permit either a formal test for the existence of mediation or a reflection of the locus of failure in a failed causal chain. The anger matching hypothesis, for example, is predicated on the presumption that pretreatment anger leads to resistance, and that MET more effectively deals with resistance than does (for example) CBT. One approach for testing such a hypothesis would require the assessment of pretreatment anger and posttreatment resistance. A causal diagram of this hypothesized model is presented in figure 2. While separate testing of an anger →posttreatment resistance link for MET versus CBT clients might reveal a significant relationship among CBT clients and a nonsignificant relationship among MET clients (as hypothesized), the magnitude of the between-treatment *difference* in the strength of the link might be small and nonsignificant—especially in the case

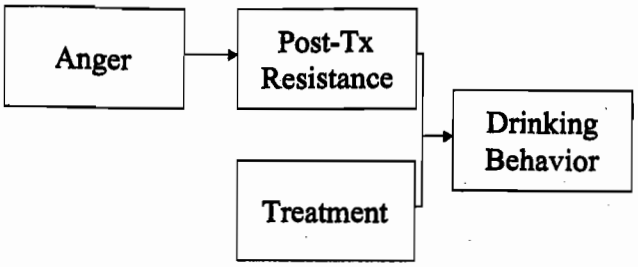

Figure 2. Example of a purported mediated moderator relationship

where CBT reflects marginal significance (i.e., p slightly below α) and MET reflects marginal nonsignificance (i.e., p slightly above α). Furthermore, if separate testing of MET versus CBT clients revealed that *both* groups reflected a significant relationship between anger and posttreatment resistance (as would also be consistent with the theory), there would be no mechanism using this approach for determining whether the anger-resistance link was *significantly* weaker for MET clients than for CBT clients. Thus, separate analysis of the treatments can actually be counterproductive, incorrectly suggesting (in some situations) a mediating effect that does not exist and incorrectly failing to validate the existence of a mediating effect when it does exist.

There is a deceptive attraction of analyzing treatment groups separately for the purposes of determining the locus of causal chain breakdowns. If the results of separate-group analysis fail to reveal a relationship (e.g., between anger and resistance) in a group where it was hypothesized, or if the relationship is found in a group where it was not hypothesized, this would at first glance seem to provide an indication of where the causal chain failed. However, the former case requires the methodologically incorrect practice of accepting the null hypothesis. The "failure to reveal" a relationship (at some preordained level of certainty, as manifested in the investigator-specified type 1 error level) is not the logical equivalent of "determining with certainty that the relationship does not exist." The latter case is also logically flawed: *any* non-zero relationship can be found to be statistically significant if the group size is large enough, and therefore a finding of an unanticipated "significant" relationship may be nothing more than an artifact of sample size. Thus, the practice of separately analyzing treatment groups in order to determine the locus of failures in the hypothesized causal chain is fraught with methodological shortcomings.

How, then, does a researcher who wishes to investigate the loci of a causal chain failure (or to statistically validate a hypothesized causal chain) proceed in the context of a matching (or, more generally, moderator) hypothesis? The answer lies in a *joint* analysis of *all* treatment groups using an extension of the Baron and Kenny formulation. In the anger-resistance example, because anger is hypothesized to lead to resistance, the interaction between *anger* and treatment is hypothesized as a distal reflection of the more proximal interaction between *resistance* and treatment. Phrased more formally, the resistance-treatment interaction is hypothesized to mediate the anger-treatment interaction. This provides a slightly more sophisticated application of the Baron and Kenny formulation, where A in figure 1 represents the anger-treatment interaction, B represents the interaction between resistance and treatment, and C represents drinking behavior (figure 3).

With these simple representations in place, the four-step Baron and Kenny formulation (as elucidated by Holmbeck) can now be applied in

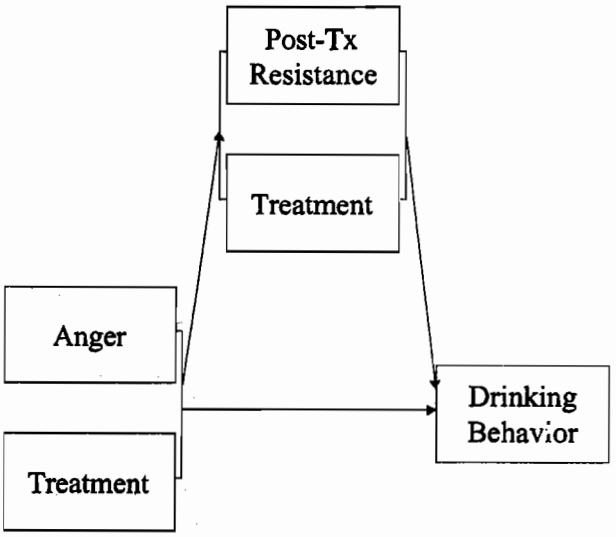

Figure 3. Testing a purported mediated moderator relationship

order to obtain a formal *statistical* test of the hypothesis of mediation. The only challenging part of this formulation lies in step 2 (establishing a relationship between A and B), because B is now an interaction. If the hypothesized interaction involves only two groups, B is a vector composed of the product of the two constituent variables. In this case, step 2 of the Baron and Kenny formulation can be implemented under a standard general linear model. If the hypothesized interaction involves more than two groups, B is an array composed of k - 1 product variables (where k is the number of groups specified in the hypothesis). In this case, step 2 of the Baron and Kenny formulation can be implemented under a standard multivariate general linear model.

We noted earlier that the Baron and Kenny formulation is just as important in determining where a purported mediational model failed as it is in determining the empirical plausibility of a hypothesized mediational model. The quintessential importance of this dual applicability is no less salient in the extension of the Baron and Kenny formulation to the mediation of moderator models than it is to the base applications referenced by Baron and Kenny and by Holmbeck. The simple substitution of interaction terms for A and for B in the Baron and Kenny formulation underscores two critical issues for investigators wishing to employ causal chain analysis. First, it is crucial that the investigator rigorously specify, in advance, the *specific* causal chain(s) that is (are) hypothesized, and then proceed to test *exactly* that chain. Second, in matching (or attribute-by-treatment) models, where the focus is on hypothesizing and testing group differences in the relationship between an attribute and an outcome variable, analyzing the groups separately will lead to a suboptimal (and, in all likelihood, inconclusive) result.

Canonical Models and the Testing of Causal Chains

As described earlier (pp. 21–26), the expansion of the Baron and Kenny (1986) formulation to include the mediation of moderator (interaction) effects carries with it the concomitant expansion of the types of models that can be investigated. Figure 2, for example, represents a model in which the purported matching effect is hypothesized to occur as a result of the surrogation of a matching variable (anger) for a more proximal variable (resistance) which interacts with treatment. Alternatively, the investigator might hypothesize that the treatment leads to a proximal outcome which interacts with the matching variable vis-a-vis drinking behavior. The Typology Hypothesis Team, for example, hypothesized that CBT and TSF would differ from MET in the amount of structure inherent in the treatment modality, and that degree of structure would interact with Typology (Type A vs. Type B) in affecting drinking behavior.

It should be emphasized that the value of distinguishing between canonical models is primarily descriptive. Whereas Canonical Type 1 focuses on a proximal effect of treatment to explain the hypothesized interaction, Canonical Type 2 focuses on a proximal effect of the matching variable. The empirical test of a Canonical Type 1 model, however, continues to follow the four-step Baron and Kenny (1986) formulation. In the case of Typology, for example, A would be replaced by the Typology-Treatment interaction, and B would be replaced by the Typology-Structure interaction (see figure 1).

A third canonical form, which we identify as Canonical Type 3A, is exemplified in the motivational readiness causal chain. Here, the interaction between treatment and readiness to change was hypothesized to affect alcohol abstinence self-efficacy (the putative mediator) which, in turn, was hypothesized to affect drinking outcome. In terms of the Baron and Kenny (1986) formulation, the formal testing of this type of model is slightly simpler than the other two canonical forms we have introduced: A is replaced by the original matching interaction (e.g., treatment-readiness to change), and B is replaced by the putative mediator (e.g., alcohol abstinence self-efficacy).

The fourth canonical form, which we identify as Canonical Type 3B, is tested in the same way as model 3A. What is different is that the content of the putative mediator in 3B is a change in the therapeutic implementation hypothesized to occur as a result of the original matching interaction.

Inherent Difficulties

We have argued above that the formal procedure for testing a putative causal chain involving a matching effect is not, if applied with sufficient rigor, arduously more difficult than the formal procedure for testing the type of mediational model described by Baron and Kenny (1986). However, there are analytical considerations which make the detection of interaction effects considerably more difficult than the detection of main effects.

Although interaction effects are frequently found in experimental studies, they are notoriously more difficult to detect in field settings. A number of reasons for this difficulty have been cited, including the covariance of the interaction term with its component variables, differences in measurement error, the use of nonlinear scales, and differential residual variances of interactions once the component main effects have been partialed out. McClelland and Judd (1993) have shown that tests of interactions in field studies will often have less than 20 percent of the efficiency of optimal experimental tests. Centering the component variables prior to creating a product-term interaction may reduce the problem (see Aiken and West 1991), although the usefulness of this procedure in significantly reducing the problem remains unclear (see Finney et al. 1984). In general, the powerful detection of interaction effects in field settings remains a highly elusive goal, the careful design and large sample size of Project MATCH notwithstanding.

Statistical Considerations

It has been noted with some alarm that the number of Project MATCH a priori hypotheses which were confirmed empirically is substantially less than what would be expected by chance (using a type 1 error rate of 0.05). Does this reflect negatively on our ability to understand the complex relationships involved in treatment matching for alcoholism? The preceding chapters of this monograph would certainly underscore just how complex some of these relationships can be. But it is also possible that statistical issues reduced the power of the study to identify hypothesized matching relationships. Three issues are of particular interest: the nature of the outcome measures, the inability to capitalize on the virtues of a latent growth model, and the possibility of undiscovered site differences. We consider each of these in turn.

The Nature of Outcome Measures.

Alcohol data are notorious in their departure from the assumptions of the general linear model. There are often clusters of "extreme" values (e.g., abstinence), the observations are rarely normally distributed around the regression line, and the relationships may be curvilinear. In contrast, most of the analytical procedures employed in alcohol research require nicely distributed mound-shaped data for which p values are reasonably accurate.

In order to bridge the gap between data as we would like them to be and data as they present themselves, it is not uncommon to attempt to induce normality and homoskedasticity through some sort of transformation. For example, a log transformation can often be helpful when the data are severely skewed to the right; a negative log transformation can be employed when the data are severely skewed to the left. Square root transformations are frequently of use when the skew is not quite so severe.

Transforming data presents two principal problems. First, interpreting the transformed data can often be challenging in the absence of obvious theoretical relevance, and therefore the results frequently have to be "back translated" by employing a reverse (or reciprocal) transformation in order to bring the units back into an interpretable metric. This is particularly challenging with alcohol treatment outcome data, where the nature of the skew is likely to change across time. For example, prior to treatment, drinking frequency measures (such as percentage of days abstinent) are likely to be positively skewed, with most individuals reflecting comparatively low values and some outliers reflecting comparatively high values of abstinent days. However, posttreatment (and particularly *immediately* posttreatment), these same measures are likely to be *negatively* skewed, with most individuals reflecting comparatively high values and some outliers reflecting compara-

tively low values. This nonnormality problem then poses a challenge to the researcher: should the same transformation be applied both at baseline and posttreatment (in order to preserve the integrity of the interpretation of the transformed variable), or should different transformations be applied to the same measure taken at different times in order to respond to the requisite assumptions under the general linear model?

In Project MATCH, the former approach was ultimately selected, although the question was thoroughly debated. In the end, employing different transformations on the same variable at different points in time was viewed as indefensible. However, this decision necessarily moved the analysis plan away from the assumptions necessary for the correct interpretation of the *p* values that were critical to the interpretation of the results. Alternative operationalizations of the primary outcome measures were considered (e.g., a dichotomization such as abstinent/nonabstinent or heavy drinking/not-heavy-drinking), but the loss of information in such alternative operationalizations was considered too great to base the entire trial on it.

A second problem associated with transforming the data to conform to the requisite assumptions of the general linear model was that, even after transformation, the data did not conform very well to these assumptions. When you have a large group of heavy drinkers (at the beginning of the trial) or of light drinkers (immediately following the end of treatment), normality cannot be even reasonably approximated in measures such as percentage of days abstinent. Thus, the *p* values on which the trial results were based were predicated on untenable assumptions. The exact effect of these departures from assumptions is unknown, but is likely to have contributed at least somewhat to the inability to detect a large number of "statistically significant" matching effects.

It should also be recognized that the primary intensity outcome measure (drinks per drinking day) possesses a property that limits its interpretability. In order to assign a value to this variable for all subjects in the trial, it was necessary to define a value for this indicator to take on when the individual did not drink at all during the period under study. After considerable debate, the Project MATCH Steering Committee decided to retain this variable as one of the two primary outcome variables and to assign the value zero to anyone whose number of drinking days during the period was zero. Thus, low scores on this variable could indicate either of two things: a truly low intensity or no drinking days at all (and therefore no basis for computing the intensity). The potential multidimensionality associated with this variable may also have been a contributor to the dearth of significant findings. (In fact, fewer matching effects were observed for the drinks per drinking day outcome variable than for the percentage of days abstinent matching variable.)

Latent Growth Modeling

This study presented one of the largest opportunities to introduce the features of latent growth modeling to the domain of randomized clinical trials. Latent growth modeling provides a feature not available in more traditional forms of the general linear model: the ability to remove random between-subject variation from the error term when testing hypotheses about fixed effects (such as the matching hypotheses). The approach is particularly useful when the individuals in a well-defined subgroup (such as those who receive a certain form of treatment modality) follow a similar pattern of increase or decline in drinking behavior across time. When this is the case, removing the between-subject variation from the error term of the *F* test of an a priori contrast can greatly increase the power of the analysis to detect hypothesized relationships.

Unfortunately, the subjects in Project MATCH revealed remarkable heterogeneity in drinking behavior across time, even among subjects who were in identical treatment modalities, identical sites, similar baseline drinking levels, and similar demographic profiles. While a clear-cut overall decrease in drinking was evident across the 12-week treatment period, and a slow regression toward more drinking was evident across the posttreatment followup period, there was little evidence to suggest that well-defined subgroups of individuals followed a similar (and distinguishable) growth trajectory.

After extensive analysis, it became clear that the best theoretically grounded model to fit these data was quadratic (allowing for a curvilinear growth pattern across time), but even under such a model the between-subject heterogeneity was quite high. As a result of this heterogeneity, there was little to commend this approach over the more traditional multivariate analysis of covariance procedure—a fact which was born out in the highly similar results across the two procedures.

Do these results suggest that latent growth analysis (also known as hierarchical linear modeling) has a limited (if any) role in longitudinal studies of drinking behavior? Probably not: the promise of latent growth is still very real. The failure here was likely attributable in large part to the previously noted aberrational distributional characteristics of the primary outcome measures. With measures which provide clearer growth patterns of individual change across time, it is likely that latent growth analysis will prove to be considerably more powerful than the classical multivariate analysis of covariance approach. But with outcome measures as heterogeneous as the Project MATCH primary outcome variables, there is little apparent justification for more sophisticated analytical techniques such as latent growth analysis.

Undiscovered Site Differences

The Project MATCH analysis plan was centered around the assumption of a common model of drinking behavior influences for *all* outpatient subjects (and a distinct common model for *all* aftercare subjects), with site adjustments for baseline and demographic differences. This assumption provided a large pool of subjects (with commensurately large power) for testing the primary matching hypotheses.

The common-model assumption is, however, a double-edged sword: while it provides the basis for powerful detection of hypothesized matching effects, it also opens up the possibility of considerable within-sample heterogeneity. In essence, after minor adjustments for cross-site differences in baseline drinking and demographic characteristics, the assumption views the data as if the subjects were all selected from a *single* source, which they were not.

An extensive analysis provided no indication of systematic cross-site differences which, if not accounted for, would influence the results of the study. Nevertheless, the question remains whether focusing the analysis on individual sites would have reduced the heterogeneity to a sufficient level to overcome the resulting loss in power.

Conclusion

Why the matches believed to be so promising investigated in Project MATCH were not supported remains an uncertainty. Although those hoping for recommendations on robust clinical matching algorithms were disappointed, a great deal of guidance and wisdom regarding the design and analysis of subsequent matching studies has been achieved. A decade ago, when MATCH was conceptualized and implemented, the methodology developed by Baron and Kenny (1986) for testing for mediators was not well understood and had not been disseminated to the alcohol treatment research community. By the time Holmbeck (1997) published his paper clarifying this methodology, MATCH had been completed. Now that mediator analysis has become well known to the field, future studies will benefit from these clarified procedures. The MATCH causal chain analysis, despite its early inception, provides direction to those seeking to undertake mediation analyses of moderator variables.

Five lessons learned from Project MATCH pertain to the conduct of future treatment outcome studies.

First, it is absolutely critical that the investigator be clear not only about *what* is being hypothesized but also about *why* that matching hypothesis is being proposed. Advancement of theory is not well served by being right for the wrong reason.

Second, it is equally important that the research test both the hypothesis itself and the purported causal chain in a statistically defensible manner. As we examined each of the causal chains presented in this monograph, these were the two tenets of causal chain testing protocol that were commonly violated. Even when the hypothesized chain was structurally sound and

well defended, the failure to follow the four steps of the Baron and Kenny formulation (or their equivalents under a structural equation model) often led to a disappointing inability to identify where the chains failed. Although requiring minor modification for nonadditive models, the seminal foundation laid by Baron and Kenny for testing mediational models provides a solid framework for not only determining whether the purported causal model of a matching hypothesis is tenable, but also for identifying the locus of failure in those situations where the hypothesized causal model is not supported.

In our review of the preceding chapters, we found two tenets of the causal chain testing protocol that were frequently violated. The first was the failure to rigorously specify, in advance, the specific causal chain(s) that was (were) hypothesized and then to proceed to test exactly that chain. The second was the failure to follow the four steps of the Baron and Kenny formulation (or their equivalents under a structural equation model), often leading to an inability to identify where the chains failed.

Third, particularly in alcohol research (where the outcome measures often defy requisite distributional assumptions of the analytical technique), the distributional characteristics of the data are important factors in considering the choice of analytical technique. Sophisticated analytical approaches such as latent growth analysis will be of little value in the context of data which do not support the requisite distributional assumptions.

Fourth, the above-mentioned point notwithstanding, the promise of latent growth analysis looms quite large in longitudinal alcohol research where (1) the outcome measures are not characterized by gross skewness and (2)individuals are likely to be characterizable by their unique growth pattern. Project MATCH demonstrated the feasibility of latent growth analysis in a multisite randomized clinical trial.

Fifth, while a larger number of subjects means greater power to detect hypothesized effects, larger is not always synonymous with better—especially when the subjects come from multiple sites. The increased heterogeneity may offset the gains in power. To the extent that a smaller sample accurately reflects the systemic behavior of a smaller definable population of theoretical or practitioner interest, the investigator might be well advised to move in this direction rather than focusing on a larger, multisite, sample.

Some have criticized Project MATCH because it did not support large numbers of matching hypotheses. While this was disappointing to investigators and interested parties alike, we believe that the number of lessons learned from this study have moved the field forward. We do not agree with those who view the negative results from this trial as a widespread reflection of the futility of matching studies. With the lessons learned from this trial in hand, we believe that the future of alcohol matching studies has been considerably strengthened.

Acknowledgments

This project was supported by a series of grants from the National Institute on Alcohol Abuse and Alcoholism as part of the Cooperative Agreement on Matching Patients to Alcoholism Treatments. The authors wish to extend their thanks to George Howe and Paul Poppen for their comments on the mediation section of this document.

References

Aiken, L., and West, S. *Multiple Regression: Testing and Interpreting Interactions.* Newbury Park: Sage Publications, 1991.

Baron, R.M., and Kenny, D.A. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51:1173–1182, 1986.

Finney, J.W.; Mitchell, R.E.; Cronkite, R.C.; and Moos, R.H. Methodological issues in estimating main and interactive effects: Examples from coping/social support and stress field. *Journal of Health and Social Behavior* 25:85–98, 1984.

Holmbeck, G.N. Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology* 65: 599–610, 1997.

McClelland, G.H., and Judd, C.M. Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin* 114(2):376–390, 1993.

Substantive Review and Critique

Richard Longabaugh, Ed.D., and Philip W. Wirtz, Ph.D.

ABSTRACT

The first part of this chapter summarizes all of the client-treatment interactions observed in Project MATCH, irrespective of whether they were hypothesized. For those hypothesized, we examine the extent to which their underlying causal chains were supportive or still remain unknown. For emergent interactions, we evaluate their plausibility as matching hypotheses to be tested in future investigations. We conclude that MATCH detected evidence for several small interaction effects. We then address the lessons learned from the causal chain analyses. In the second part, we examine implications of our findings, first for future matching research and then for alcohol treatment research more generally. We conclude that Project MATCH's investigation of the mediation of matching effects has provided guidance for future alcohol treatment research.

We are now reaching the completion of our journey. The aims of this volume were first to present the rationale for each of the a priori matching hypotheses, the hypotheses themselves, and results. The second aim was to present the theory underlying each of the matching hypotheses and the results of testing these theories by causal chain analysis. For unsuccessful predictions, the causal chain analysis identified how the theory failed. For supported matching predictions, the causal chain analyses tested whether the underlying theoretical assumptions were the mechanisms through which the hypothesized interaction occurred. These analyses and results have now all been reported. The preceding chapter provides a methodological critique of the various approaches taken to testing the causal chains and matching predictions.

In the first part of this chapter, we attempt an integration of the overall effort, incorporating all data from the observed client-treatment interactions and causal chain analyses. Next, we examine the unique contribution of the causal chain analysis to understanding Project MATCH results. In the second part, we offer our recommendations and perspectives on treatment matching specifically and future treatment research more generally.

Summary of Observed Client Attribute-Treatment Modality Interactions

Before undertaking this review, introductory comments are in order. The presentation is organized by arm of study, and within arm of study, by phase of treatment—within treatment followed by posttreatment effects. Within each section, first the a priori hypothesized interactions that were supported are addressed in detail. Here our assumption is that these hypotheses have been supported. Therefore, our focus is in critiquing how well we understand the theory underlying the matching prediction. Next we describe the observed interactions that were not hypothesized. Here our focus is on exploring the plausibility of the interaction. Is it simply a chance occurrence or is it possible that it reflects a true effect? Causal chain analysis is used to

Richard Longabaugh, Ed.D.
Center for Alcohol and Addiction Studies
Brown University, School of Medicine
800 Butler Drive, Potter Building, Room 204
Providence, RI 02906
E-mail: Richard_Longabaugh@Brown.edu

evaluate the plausibility of the observed interaction. As the probability of finding both an observed interaction and a supporting causal chain by chance alone would be extremely low, this observation would suggest further research may be warranted.

It should be noted that in many instances observed interactions were not involved in an a priori contrast. Consequently, pertinent causal chain analyses were not conducted. In other instances, however, the treatment contrasts were involved in an a priori hypothesis, but the observed outcome was *opposite* that predicted. Here causal chain information is more likely to be available and pertinent.

Finally, given our evaluation in the last chapter that few of the underlying theories for the a priori matching hypotheses were both conceptually explicit and clearly and correctly operationalized, inferences drawn from the causal chain analyses are necessarily speculative. No supporting causal chain analysis was carried through to the last step of ruling out competing interpretations for the observed matching effect. Further, causal chain analyses of failed matching hypotheses were not required to identify all of the linkages in which the causal chain was unsupportive—only one failed link was needed. Thus, complete information necessary to critique each matching hypothesis is not available. Despite these limitations, we believe there is sufficient information from the causal chain analyses to enrich our understanding of the client attribute-treatment interactions tested.

The Observed "Hit Rate"

Considering the large number of interactions tested, it is entirely possible that all of the interactions observed may be attributed to chance. Leaving aside interactions involving time, only 3 percent of the a priori hypothesized contrasts were supported at the Bonferroni level of adjustment. Twelve contrasts met our criteria for significance out of 400 tested (2 dependent variables × 40 hypothesized contrasts, each tested 5 times—aftercare within- and posttreatment and outpatient within- and posttreatment and months 37–39).

This suggests two possibilities. Perhaps all a priori hypotheses are untrue, with those appearing to be supported being attributable to chance alone. Another possibility is that we made our criteria for rejecting the null hypothesis too stringent. The familywide Bonferroni correction may have led us to underdetect valid client-treatment interactions.

In order to provide a context for assessing these alternative interpretations, it is useful to provide the studywide hit rate for tests of interactions. Had we specified a studywide uncorrected alpha level of 0.05, we would expect on average 5 percent of all nonexistent interactions tested to meet our criteria by chance. Overall, across the 2 study arms, 2 dependent variables, 21 matching variables, and 3 treatment contrasts per matching variable (CBT vs. MET, CBT vs. TSF, MET vs. TSF), 504 tests of attribute by treatment modality interactions were conducted (excluding those involved in either a linear or quadratic interaction with time). Of these, 38 were observed to be significant at $p \leq .05$, 7.5 percent of those tested. This hit rate is somewhat greater than that expected by chance alone.

A possible implication of this comparison is that our attempt to capitalize on what we thought we already knew to develop a priori predictions actually impeded our ability to uncover evidence for matching. Had we been less confident of our predictions, we might have conducted an exploratory hypothesis-seeking analysis with one portion of the data and used the other as a holdout sample upon which to test the emergent hypotheses. The Steering Committee considered this alternative but ultimately rejected it because of our (in retrospect unfounded) confidence in the predictions that were developed.

While the MATCH requirement of a priori hypotheses had the great virtue of sharpening the focus of study, it also precluded examination of emergent interactions that had not been predicted. Now that all of the hypotheses have been tested, exploratory investigations of other observed interactions are appropriate. With this perspective in mind, we examined the tables in the separate chapters reporting the results for interactions tested though not hypothesized. We extracted any interaction that met the $p \leq .05$ alpha level and incorporated these results

into appendices 1 and 2. These show all interactions that occurred, whether hypothesized or not, including those involving linear and quadratic time. Appendix 1 summarizes the outpatient arm, appendix 2 the aftercare arm.

As this information, though of potential significance, is too vast to address, table 1 presents for both arms of study only those interactions that did not involve a time dynamic (i.e., an interaction of the matching effect with time). Treatment contrasts involving 14 of the 21 matching variables met the criteria specified. While several of the interactions involving change over time are quite interesting and even comprehensible (i.e., the matching effect increases or decreases over time), we have not yet identified the tools necessary to undertake a causal chain analysis that could explain these dynamics.

Outpatient Arm

Within-Treatment Interactions

No matching prediction was observed during the within-treatment period that met the Bonferroni level of adjustment. Four interactions were observed with p's $\leq .05$.

Psychopathology and Psychiatric Severity. During treatment, there was evidence supporting an a priori matching hypothesis that CBT would be more effective than MET for more psychologically impaired clients. Not hypothesized but observed, TSF was also more effective than MET for more impaired clients. This suggests either that CBT and TSF each may have a distinct component that is helpful to more impaired participants which MET lacks, or that CBT and TSF may share an active ingredient lacking in MET that may help high psychopathology clients. One obvious active ingredient shared by CBT and TSF was the greater number of treatment sessions available (12 vs. 4). This may have accounted for this beneficial effect during the treatment period (rather than CBT's hypothesized greater focus on psychopathology).

Gender. It had been hypothesized that women would do better in CBT than in TSF, relative to men. In fact, the opposite was observed. Women treated in TSF had more abstinent days during treatment than when treated in CBT. Causal chain analysis revealed that most of the differences related to instrumentality and expressiveness upon which the gender matching hypothesis was predicated were in fact present. Where the causal chain appeared to fail was in CBT's inability to influence these characteristics as expected. Instead, one of the causal chains tested (involving both canonical models 1 and 2 (see pp. 21–26)) showed that females in CBT reduced their psychiatric severity less than did males in CBT, and less than either males or females in TSF. This would suggest that CBT's relative failure to diminish women's psychiatric severity may have accounted for their lesser improvement in drinking while in treatment. This causal chain is not completely supportive of the theory, however, as posttreatment psychiatric severity did not predict days abstinent during treatment for women. Thus, the puzzle remains. It should be noted that several of the hypothesized causal chains were not directly tested, so the mediating mechanism for TSF's superiority for women may yet be detected.

Sociopathy. It was hypothesized that CBT would be more effective than MET for clients high in sociopathy. Instead, during treatment, clients with high sociopathy fared equally well in CBT and MET. However, clients low in sociopathy had a higher percentage of days abstinent (PDA) when treated in CBT than their MET counterparts. This would suggest that CBT is either mismatched to high sociopathy, or MET is mismatched to clients low in sociopathy.

While the well-articulated causal chains did not attempt to account for drinking during treatment, indirect support for the interaction was observed in one of the causal chains conducted to account for posttreatment drinking. This type 3A canonical chain involved the working alliance. CBT clients low on sociopathy were more likely to have a better working alliance, and a better working alliance was related to better PDA. For MET clients, working alliance was unrelated to drinking outcomes. The linkages in the causal chain are thus consistent

Table 1. Observed interactions

Matching variable	Study arm[1]	Observation period[2]	Hypothesized contrast[3]	Observed contrast[4]	Predicted[5]	Dependent variable[6]	Causal chain supportive?[7]
Psychiatric Severity	OPT	Treatment	CBT>MET	CBT>MET	yes*	PDA	no
Gender	OPT	Treatment	CBT>TSF	TSF>CBT	opposite	PDA	no, partially tested
Sociopathy	OPT	Treatment	CBT>MET	MET>CBT	opposite	PDA	yes
Prior AA	OPT	Treatment		MET>CBT	no	PDA, DDD	not tested
Psychiatric Severity	OPT	year 1+	CBT>TSF	CBT>TSF	yes	PDA	no
Anger	OPT	years 1, 3	MET>TSF, CBT	MET>TSF	yes	PDA, DDD	no
				MET>CBT	yes	PDA, DDD	no
Self-Efficacy	OPT	year 1	MET>CBT	CBT>MET	opposite	DDD	yes
Social Functioning	OPT	year 1	CBT>MET	MET>CBT	opposite	PDA, DDD	no
			CBT>TSF	TSF>CBT	opposite	PDA, DDD	no
Prior AA	OPT	year 1		MET>CBT	no	PDA	not tested
Interpersonal Dependency	OPT	year 1		CBT>MET	no	PDA	no
		year 3		CBT>MET	no	DDD	no
		year 3	MET>TSF	TSF>MET	opposite	DDD	no
Network Support	OPT	year 3	TSF>MET	TSF>MET	yes	PDA, DDD	yes
A vs. B Typology	OPT	year 3	MET>TSF	TSF>MET	opposite	DDD	no
			MET>CBT	CBT>MET	opposite	DDD	no
Self-Efficacy	AFT	Treatment	MET>CBT	MET>CBT	yes*	PDA, DDD	no
		Treatment	MET>TSF	MET>TSF	yes	PDA, DDD	no
ASPD	AFT	Treatment	CBT>TSF	CBT>TSF	yes*	DDD	not tested
Temptation	AFT	Treatment		CBT>MET	no	PDA	not tested
Readiness	AFT	Treatment	CBT>MET	MET>CBT	opposite	PDA	no
Readiness	AFT	year 1		TSF>CBT	no	PDA	not tested
Alcohol Dependence	AFT	year 1	TSF>CBT	TSF>CBT	yes	PDA, DDD	yes

1. OPT = outpatient, AFT = aftercare
2. Treatment = months 1–3, year 1 = first year after treatment completion (months 4–15), year 3 = 37–39 months after treatment initiation
3. A priori
4. $p \leq .05$, 2-tailed test
5. yes = difference in slopes same as hypothesized, opposite = observed contrast in slopes opposite what was predicted, no = no a priori contrast hypothesized
6. PDA = percentage of days abstinent, DDD = drinks per drinking day
7. yes = causal chain supports the observed interaction (irrespective of what was hypothesized), no = there is not a complete linkage supporting the observed interaction even though one or more was tested, not tested = a complete linkage of the causal chain to the observed interaction was not tested.

* = hypothesized contrast but significant only when unprotected ($p \leq .05$)

\+ = There was both an attribute by treatment interaction and a time by attribute by treatment interaction. The contrast was significant at months 5–11. The finding is included because of the fact of the attribute by treatment interaction.

with the direction of the observed sociopathy-treatment modality interaction. This supports the credibility of the interaction observed despite its variance from the original hypothesis.

Prior AA. No matching effect had been hypothesized for the interaction of prior involvement in Alcoholics Anonymous (AA) and MET versus CBT treatment assignment. However, it was observed that for those with prior AA involvement, MET is predictive of higher PDA and fewer drinks per drinking day (DDD), while for CBT clients, prior AA is predictive of poorer drinking outcomes. As this contrast was not a candidate for causal chain analysis, only speculation can be offered. MET encourages clients to develop their own change plans, which may lead to utilization of AA during and after treatment. In contrast, CBT, with its own prescriptions for client change, may inadvertently work against whatever predilections some clients would have for invoking AA prescriptions for recovery.

Posttreatment Interactions

Psychiatric Severity and Psychopathology. An ordinal interaction effect was hypothesized for psychopathology. It was expected that for low levels of psychopathology, treatment assignment would not make a difference, but as psychopathology increased, clients assigned to CBT would have better drinking outcomes than those assigned to either TSF or MET. An ordinal interaction was observed; however, contrary to expectations, clients with low psychopathology had better drinking outcomes in TSF than in CBT. For those with high psychopathology, CBT was neither better nor worse than TSF.

Despite the unanticipated implication for clinical triaging—assign low psychopathology clients to TSF—the theoretical premise underlying the hypothesized interaction was supported. As CBT client psychopathology increased, drinking outcomes improved. As this was not so for TSF, where clients outcomes were unaffected by their level of psychopathology, it could be concluded that CBT was helping clients with greater psychopathology more than it was helping those with less. However, CBT's increased efficacy for clients with higher psychopathology raised their level of outcome sufficiently only to equal that of comparable TSF clients. This suggests that TSF has one or more active ingredients more helpful to all alcohol clients than does CBT. CBT's increased efficacy for the subset of clients with high psychopathology appears to compensate for the absence of these unidentified active ingredients.

The matching effect observed was no longer significant after the tenth month of posttreatment, suggesting that the differential effect of CBT versus TSF on clients varying in their psychopathology fades over time.

Multiple attempts to identify one or more underlying supportive causal chains for this interaction, both a priori and post hoc, were unsuccessful. These attempts, when decomposed, involved canonical models 1 and 3A. The causal chains indicated that, contrary to what was hypothesized, there was neither greater attention to client psychopathology in CBT nor did CBT clients experience a greater reduction in psychopathology following treatment. The mechanisms by which this interaction is produced are yet to be identified.

Anger. Client anger was observed to interact with MET versus TSF/CBT treatment modalities to produce a disordinal interaction during followup. This matching effect was the most consistent one observed in the entire study. Evidence for the matching effect was present at all followup points—at 3-year followup as well as during the first year of followup. However, the theoretical underpinnings for this effect could not be identified in the causal chain analyses conducted, which involved canonical models type 2 (for taking steps and problem recognition) and type 3 (for working alliance).

Also problematic, the a priori matching hypothesis proposed was ordinal: that angry clients would have better drinking outcomes in MET than in either CBT or TSF. The theory underlying this matching hypothesis was that high anger clients would be more resistant to treatment than would low anger clients. It was expected that MET, with its nonconfrontational stance, would reduce the client's resistance to treatment, and by doing so, produce better drinking outcomes. Unanticipated was that low

anger clients would have poorer drinking outcomes in MET than in CBT/TSF.

Nevertheless, a review of treatment research findings for other psychological dysfunctions indicates that a disordinal matching effect similar to this one has often been observed. If the state-trait anger variable is considered to be an index of reactance (Brehm 1976; Brehm and Brehm 1981), it has been found that psychotherapy clients with high resistance (also an index of reactance) are most effectively treated with a supportive, nonconfrontational, low directive therapy, while those low in reactance (or resistance) are most effectively treated with a more directive therapy (Beutler et al. 2000).

The hypothesis team was at a disadvantage regarding the selection of the putative mediators of this matching prediction. The assessment battery had already been decided upon by the time Secondary Matching Hypothesis Teams had completed development of their causal chains. The Anger Matching Hypothesis Team selected five indices hypothesized to characterize an *absence* of client resistance—two measures from the SOCRATES readiness scale: Problem Recognition and Taking Steps for Change, and three measures from the Working Alliance: Agreement on Treatment Goals, Tasks, and Bonding.

A first problem with these indices was that only one, Working Alliance Goals, was negatively related to client anger as would be expected and necessary for the underlying causal chain to be supportive. The two other indices of working alliance were unrelated to client anger. Problem Recognition was actually significantly related to client anger in the opposite direction predicted. Thus, there was a breakdown in the causal chain in the first linkage.

While the hypothesis team observed that the indices of working alliance partially mediated the relationship of the anger-treatment matching variable to drinking, the analysis conducted included working alliance variables as both main effects and product terms. As the separate effects were not reported, their status as putative mediators of the moderating effect is confounded with their main effects.

As discussed in the preceding chapter, the fact that the relationships were reduced from significant to nonsignificant when these mediators were partialed out might also be attributable to the greater number of degrees of freedom utilized in the mediator analyses.

The one putative mediator of the matching effect that appears to be operative is Problem Recognition, which reduces the relationship of the treatment-anger variable to PDA to nonsignificance and is also related to client anger. However, this relationship is positive rather than negative. This suggests that angry clients are more likely to report problem recognition, which when paired with MET as opposed to CBT and TSF, leads to better PDA.

This causal chain involving problem recognition has migrated considerably from the causal chain from which the hypothesis team started. Waldron et al. (this volume) conclude that the data at hand do not provide an adequate operationalization of Resistance. They expect that tape ratings of therapy sessions may provide such an index in the future. Beutler's research is supportive of this interpretation. He and his colleagues have found that observer ratings of therapist directiveness and client resistance produce the anticipated disordinal matching effect (Beutler et al. 2000; Karno et al. in press). Thus, despite the lack of a supporting causal chain, the disordinal matching effect is highly credible.

Network Support for Drinking. As hypothesized, clients with network support for drinking prior to treatment had better drinking outcomes at 3-year followup than those assigned to MET. Although hypothesized to be ordinal, the interaction actually gave indication of being disordinal, as clients with little support for drinking prior to treatment had significantly (but only slightly) better outcomes in MET than in TSF.

The causal chain model tested was type 1. The analyses indicated that participation in AA was in part responsible for the observed attribute-treatment interaction. Clients treated in TSF were more likely to participate in AA, which in turn was differentially associated with better drinking outcomes, with those high in network support benefiting the most. However, the significance of the matching effect was only

substantially reduced, not eliminated, when the effect of AA participation was partialed out. So other mediators, as yet unidentified, also contribute to this attribute-treatment interaction.

Another point of interest here is that the mediator variable itself turned out to be a moderator variable, the product term of network support for drinking and AA participation and treatment modality. The enhanced benefit of AA participation for clients with networks supportive of drinking was most apparent in TSF. This suggests that the compatibility of treatment modality and self-help group belief system is likely to be an important factor to consider in matching. Research by McCrady and colleagues (1999) bolsters this speculation. These investigators found that a cognitive-behavioral Alcohol Behavior Couples Therapy followed by AA involvement was less helpful than a cognitive-behavioral relapse prevention aftercare model.

Finally, two other aspects of this attribute-treatment modality interaction have yet to be satisfactorily accounted for. First, why did clients with networks unsupportive of drinking actually do somewhat better in MET than in TSF? No explanation or analysis was offered to account for this tail of the disordinal interaction. Second, it is not clear why the interaction that eventually emerged for TSF's superiority with clients having networks supportive of drinking did not do so earlier in the followup period. Presumptively, participation in AA after treatment completion explains the delayed matching effect, but this has not been tested.

Nonpredicted Outpatient Interactions

Nonhypothesized interactions involving five matching variables were observed during the posttreatment period: Prior AA, Self-Efficacy, Social Functioning, Interpersonal Dependence, and Typology.

Prior AA. The effects of prior AA involvement in the MET and CBT treatment contrast observed during treatment continued into the first year of posttreatment. MET clients who had been involved in AA prior to MATCH treatment continued to have more abstinent days than those not so involved, while for CBT clients, the reverse was observed. The pervasiveness of this interaction into the posttreatment period strongly suggests that further study is warranted.

Self-Efficacy. It was hypothesized that self-efficacy would interact with treatment modality such that clients low in self-efficacy would do better in CBT than in MET. In the outpatient arm (in contrast to the aftercare arm), an interaction occurred that suggested the opposite. Low self-efficacy clients treated in MET had fewer drinks per drinking day than those treated in CBT.

The causal chain analysis supported this finding. Employing a canonical model 3A, an interaction effect involving change in self-efficacy in CBT versus MET was tested. It was found that low self-efficacy MET clients had a greater increase in self-efficacy before to after treatment than did CBT clients. This increase in self-efficacy predicted end of treatment drinking which in turn predicted self-efficacy at 9 months and drinks per drinking day in the year following treatment. Thus, the causal chain analysis, by supporting the observed interaction, strengthens the credibility of this finding, despite its opposition to what had been hypothesized.

Social Functioning. It was hypothesized that CBT would be more effective than either TSF or MET for clients who had poor social functioning. In fact, the opposite effect was observed. Clients with poor social functioning who were assigned to CBT had fewer abstinent days and more drinks per drinking day throughout the first year of followup. Clients with high social functioning initially did better in CBT, but this superiority dissipated as the year went on.

The causal chain analysis was limited to determining where the a priori theory had broken down. While it was found that, as expected, posttreatment social functioning predicted posttreatment drinking, CBT did not improve social functioning more than did MET or TSF. This explained why CBT was not superior to MET and TSF for poorly functioning clients. However, it does not explain why it should be significantly worse. Post hoc causal chain analyses are necessary to explore this finding.

Interpersonal Dependency. No a priori predictions were made regarding contrasting effects of CBT and MET on clients varying in their interpersonal dependency. Nevertheless, interactions were observed for PDA in the year following treatment and for DDD at the 3-year mark. During the year following treatment, MET clients with higher dependency had fewer days abstinent than those with low dependency. At 3 years out, higher dependency in CBT was associated with fewer drinks per drinking day. Thus, in one case, high-intensity treatment was beneficial to more dependent clients, and in the other, low intensity treatment was beneficial to those with low dependency.

At 3 years, as had been hypothesized, TSF was associated with fewer drinks per drinking day for more highly dependent clients. Combined with the results for CBT, this suggests that interpersonal dependency is predictive of fewer drinks per drinking day in higher intensity treatments, but unrelated to DDD in low intensity treatments. The causal chain analyses, involving treatment completion and treatment satisfaction in a canonical model 3A do not provide any clues for these observed differences. Interpersonal dependency is unrelated to both treatment completion and satisfaction in all three treatments.

A Versus B Typology. It was predicted that clients having a type A typology would have better drinking outcomes in CBT and TSF than in MET, while the reverse would be true for type B clients. At 3-year followup, the opposite effect was reported. Three causal chains were tested. The first tested amount of therapeutic structure and cognitive change in a canonical model 1. The second and third were type 3A canonical models and tested working alliance and change in psychopathology as potential mediators. The models indicated various links in which the causal chains broke down but did not provide any clues as to why the unpredicted interactions occurred. Given the gap in time between treatment and 3-year followup and the absence of a connecting causal chain, it is likely that this finding is spurious.

Summary of Outpatient Arm

During the within-treatment period, none of the a priori predictions was supported at a protected level of confidence. Therefore, nothing definitive can be said about matching clients to treatment modalities during outpatient treatment. If all observed interactions are taken into account, a multidimensional typology would be necessary to develop useful hypotheses about which clients would do better and worse in these treatments. For example, it might be hypothesized that for women, those who are high in psychopathology would do best in TSF and less well in CBT. For men high in psychopathology and low in sociopathy, CBT might be optimal, while MET could be contraindicated.

In the posttreatment period of observation, 13 attribute-treatment interactions were observed. Four of these were predicted, involving psychiatric severity, anger (in two treatment comparisons), and network support for drinking. Six, involving four matching variables—self-efficacy, social functioning, interpersonal dependency, and typology—were opposite the a priori predictions. Three others emerged in the absence of any predictions. One interaction involving prior AA that emerged during treatment persisted through the first year of followup. Two other interactions, involving anger and interpersonal dependency, persisted from the first year of posttreatment followup through to the third.

Whether the criterion be a priori predictions or observed interactions, evidence is sufficient for posttreatment matching effects in the outpatient arm of study.

Aftercare

Within-Treatment

Self-Efficacy. The self-efficacy matching hypothesis was strongly supported in the aftercare arm during treatment. As the effect disappeared after treatment had been completed, the practical advantage to be gained for clinical purposes is diminished. Nevertheless, it has theoretical significance.

It was hypothesized that MET would be more effective than either CBT or TSF for clients with higher self-efficacy. However, with lower self-efficacy, it was predicted that CBT and TSF would be more effective. Results indicated that client

self-efficacy made a difference in MET, such that the lower the client's abstinence self-efficacy, the poorer the drinking outcomes. In contrast, client self-efficacy did not appear to affect treatment success in either CBT or TSF. The net result of these effects was that low self-efficacy MET clients did more poorly during treatment than their counterparts in CBT and TSF. At higher levels of self-efficacy, treatment modality did not differentiate drinking patterns during treatment.

The causal chain underlying the hypothesis anticipated that treatment modality and pretreatment self-efficacy would interact to affect self-efficacy during treatment (canonical model 3A). This differential change in self-efficacy would in turn interact with pretreatment self-efficacy to influence drinking. The causal chain failed in that CBT increased the self-efficacy of all clients more than did MET, thus failing to account for the matching effect.

Nonpredicted Within-Treatment Interactions

In addition to the predicted interaction involving self-efficacy and treatment modality, three other client variables were observed to interact with treatment modality: motivational readiness, antisocial personality disorder (ASPD), and temptation minus confidence.

Antisocial Personality Disorder. It was hypothesized that CBT would be more effective than TSF in treating clients with ASPD. For drinks per drinking day, this contrast was significant at $p<.024$ (a value larger than what was required to meet the Bonferroni correction). As causal chain analyses were conducted only with the related variable, sociopathy, no pertinent information is available to evaluate this attribute-treatment modality interaction.

Temptation Minus Confidence. The construct of temptation minus confidence (closely related to self-efficacy, which did show the hypothesized interaction during treatment) was involved in an interaction contrasting CBT and MET during treatment. Highly tempted CBT clients had more abstinent days than did their MET counterparts. As this contrast was not involved in an a priori hypothesis, no causal chain was developed or tested for this effect.

Motivational Readiness. Motivational readiness was hypothesized to interact with CBT and MET because clients with low readiness were expected to respond to MET more than to CBT. It was expected that highly motivated clients would respond equally well to the two treatments. Thus, an ordinal interaction was hypothesized. The interaction observed was opposite that predicted. Clients with low motivational readiness apparently achieved higher PDA when treated in CBT versus MET. For those with high motivation, treatment assignment made less of a difference. Causal chain analyses failed to support the observed interaction, reducing its credibility.

Posttreatment Interactions

During the 1-year posttreatment period, one hypothesized interaction was observed, involving alcohol dependence. A second, involving motivational readiness, emerged in the absence of a prediction.

Alcohol Dependence. A hypothesized disordinal matching effect was observed. The higher the clients' alcohol dependence, the more likely they would achieve a higher percentage of abstinent days and fewer drinks per drinking day when treated in TSF versus CBT. Conversely, those with lower alcohol dependence would achieve more PDA and fewer DDD when treated in CBT versus TSF. This disordinal interaction met all criteria for matching— clients high on dependence had significantly better drinking outcomes when treated with TSF, while those low on dependence had significantly better drinking outcomes when treated with CBT.

Two causal mechanisms were hypothesized to be responsible for this interaction—therapist emphasis on abstinence and client participation in AA. In both cases, the canonical causal chain model tested was type 1. Only therapist emphasis on AA was supported. Lack of therapist emphasis on abstinence, associated with CBT, was

found to explain the superiority of CBT for clients with low alcohol dependence. As alcohol dependence increased, the superiority of CBT diminished, so that at high levels of dependence, the treatments were not distinguishable in their effectiveness.

Left unexplained by this mediator analysis is why TSF clients with high dependence had better drinking outcomes than comparable CBT clients. It would be expected that as dependence increases, therapist emphasis on abstinence would enhance drinking outcomes. From figure 5 in the Cooney and Babor chapter, it can be seen that this was not the case. While lack of emphasis on abstinence enhanced PDA for those less dependent, emphasis on abstinence did not enhance outcomes for those more dependent. The implication is that some other active ingredient, associated with TSF, was responsible for increasing the PDA of highly dependent clients. This ingredient is yet to be identified.

Nonpredicted Posttreatment Interactions

Motivational Readiness. During the year following treatment, MET and TSF were observed to interact with motivational readiness. No hypothesis had been offered for this contrast. Clients with low motivation treated in TSF had fewer drinking days than those treated in MET. As motivation increased, the treatment differences in PDA decreased. As this did not pertain to a hypothesized contrast, no pertinent causal chain analyses were conducted. It is of interest to note, however, that in both the within-treatment contrast with CBT and the posttreatment contrast with TSF, clients with *low* motivation who were treated in MET had more drinking days. This is inconsistent with the notion that MET is helpful because it increases the motivation of less motivated clients.

Summary of Aftercare Arm

Five interactions were observed during the aftercare treatment. Three were predicted a priori (although only one achieved a protected level of significance), one was opposite that predicted, and one was observed in the absence of a prediction. Within-treatment matching effects occur as often in aftercare treatment as in standalone outpatient treatment. Of note, however, of the aftercare interactions observed, four of the five involved variables as much reflective of state as trait (i.e., readiness, self-efficacy, and temptation minus confidence). It is plausible that client states would be more responsive to ongoing aftercare treatment than would traits.

In contrast, only two matching effects following aftercare treatment were observed. It may be that the confounding of more intensive treatment with the MATCH aftercare treatment diluted posttreatment matching effects that otherwise might have occurred had MATCH been a standalone treatment.

Commentary

In summary, this review of observed client attribute-treatment modality interactions calls attention to several that had not been previously addressed because either:

- Although they were hypothesized a priori, they did not achieve a familywide protected level of significance using a Bonferroni adjustment.
- They were not hypothesized a priori.
- The results observed were in the direction opposite that hypothesized.

Review of all 21 matching variable candidates points to 7 that were not involved in an interaction affecting drinking outcomes (excluding interactions involving a time dynamic): alcohol involvement, psychiatric comorbidity, cognitive impairment, conceptual level, meaning seeking, religiosity, and problem recognition. However, some of these variables had prognostic value, unaffected by treatment assignment: alcohol involvement, meaning seeking, religiosity, and problem recognition (see appendix 3). In such instances, it can be concluded that their prognostic effect was not moderated by this set of treatment modalities, posing a challenge to those trying to develop more effective treatments.

Only two client variables, conceptual level and cognitive impairment, appear to be irrelevant for prognostic or treatment assignment

purposes. Thus, it would appear that all of the variables included in Project MATCH's selection as candidates for matching are, in at least some very small degree, likely to be pertinent to assessing the effectiveness of treatment.

Implications of Matching Results

Having completed our review of all attribute-treatment interactions not involving a time dynamic (i.e., an interaction with time), what implications can be drawn for the idea of matching clients to treatments? Does client-treatment matching affect drinking outcomes? If so, is it possible to conduct treatment outcome research that will demonstrate these matching benefits?

Does Client-Treatment Matching Occur?

Does client-treatment matching occur? Intuitively, the answer is yes. At the most general level, clients treated for one disorder who in fact have a different disorder are less likely to respond to the treatment. Confining our scope to clients treated for alcohol problems, there is evidence that treatment service matching to client profile affects outcomes (e.g., McLellan et al. 1983). The scope of MATCH, however, was more narrowly focused. Our aim was to compare three individually delivered, manual guided and structured psychosocial therapies in their capacity to be matched or mismatched to 21 client attributes, mostly traits and a few states.

Within this limited scope, what do we conclude about matching in Project MATCH? Having reviewed all of our matching results, either there is little evidence for matching at this level of specificity or there is ample evidence for multiple small matching effects.

Argument for No Evidence for Matching

The results observed in Project MATCH may indicate that client attribute-treatment interactions do not occur in a way that would affect client drinking behaviors in a predictable way. We failed to find more than 3 percent of the hypothesized interactions significant at the protected level. Moreover, those that were significant accounted for small portions of the variance. No interaction was significant at all periods of observation nor was any hypothesized interaction significant in both arms of the trial. There is no indication in these data that strong matching effects exist, indicating that the matching paradigm may have outlived its heuristic value.

Argument for Multiple Small Matching Effects

An alternative conclusion is that MATCH results suggest the presence of several small matching effects between single client attributes and these three treatment modalities. As the number of observed interactions may exceed what could be attributed to chance alone, it is likely that something is going on, but we have not been able to comprehend what that something is.

Our inclination is to accept this interpretation of multiple small matching effects. But then, why was this pattern of multiple small interactions observed rather than several strong matching effects? One explanation may lie in the design. Three treatments developed for all patients with alcohol dependence or abuse were compared with one another for their matching effects on clients contrasted in 21 different ways. As each of these client characteristics was believed to be involved in one or more credible matching hypothesis, in effect this meant that they were competing with one another for the same outcome variance.

As only three treatment contrasts were possible, TSF versus CBT, TSF versus MET, and MET versus CBT, on average seven matching variables were involved in each of these contrasts. Given their a priori plausibility, how likely is it that any one matching variable would show a strong matching effect if the others were also influential? If the matching variables were moderately correlated with one another (as most were), it is conceivable that they may have been proxies for one another or for more general underlying constructs.

Under these conditions we might expect to see the pattern observed: small matching effects involving alternative matching variables in different interactions observed during different

periods of observation and in different arms. When it is noted (as we discuss later) that many of the matching predictions relied upon the same putative mediator variables to explain the matching effect, this explanation becomes more likely.

Finally, some matching hypotheses assumed main effects of treatment modality or client matching variable on mediating mechanisms (canonical models 1 and 2, respectively), while others assumed and found these same mediating mechanisms being affected by modality-matching variable interactions (models 3A and 3B). The latter effects would undermine the assumptions of the former. If one matching hypothesis assumes treatment modality has a consistent effect on treatment process, whereas results from testing the causal chain of another matching hypothesis shows the treatment process to be affected by the second matching variable, the conditions necessary for mediation of the first matching prediction may be precluded.

This said, it is nevertheless clear that if matching research is to be productive at this level of specificity, a major shift in approach is necessary. MATCH represented the best effort that could be put forth using this type of approach where single client characteristics were expected to interact with two or more treatments to affect drinking outcomes.

Causal Chain Analyses

A comprehensive review and critique of the causal chain analyses is clearly beyond the scope of this chapter. Instead, we limit our comments to a few of the more salient points to emerge.

First, Project MATCH's use of causal chain analysis increased our understanding of why hypothesized interactions emerged, as in the case of alcohol dependence and network support for drinking. As importantly, the causal chain analysis threw into question assumptions we had as to why other observed matching effects that were hypothesized did in fact occur: anger, psychiatric severity, self-efficacy.

The causal chain analyses also helped to clarify our understanding of why so many of the hypothesized interactions failed to emerge. Causal chain analyses also provided plausibility for a few observed interactions that were not hypothesized.

Finally, prior to undertaking this review, we suggested that it was possible for a causal chain to emerge in the absence of an interaction. This too, in fact, appears to have occurred. For example, in both arms of the study, the amount of structure in therapy interacted with sociopathy to affect drinks per drinking day, and in the outpatient arm, TSF was found to have more structure than MET. Thus the causal linkage appears to be complete, despite the fact that TSF versus MET did not affect drinks per drinking day when clients high and low in sociopathy were contrasted. (The analytic approach used to test these causal chains precluded a test of the significance of the interaction between the two treatment groups. Thus, this inference is tentative).

The most serious limitation apparent from our critique of the causal chain analysis was that all too often the underlying theory was either not made sufficiently explicit so that it could be operationally tested, or if conceptually clear, its operationalization was either not totally accurate or was incomplete. It is now apparent that mediator analysis of moderating effects is a significant challenge. Nevertheless, once beyond these front-end limitations, the causal chain analyses conducted yielded considerable information.

Mediator Variables

The variables hypothesized to act as mediators turned out to be a surprisingly small set of indicators of the treatment process and a somewhat larger number of client responses to treatment.

Treatment Processes

Only six constructs characterizing treatment were tested as hypothesized mediators. Working alliance, amount of structure in treatment, the amount of treatment offered or received, and AA involvement and/or attendance were each tested in the causal chains for several of the hypotheses. Two indices of treatment content were each used once, treatment emphasis on psychopathology and emphasis on abstinence.

Working Alliance. Working alliance was the most frequently tested measure of treatment process. The factors thought to affect it varied across matching hypotheses. And, in fact, analyses indicated that it was affected by treatment modality *and* client matching variable *and* the interaction of treatment modality and client matching variable. For example, in the outpatient arm, MET enhanced clients' experience of the working alliance more than did CBT or TSF in one causal chain analysis, while in another, TSF clients reported better working alliance indicators than did CBT or MET clients.

Such inconsistencies become understandable when the evidence indicates that working alliance is also influenced by the interaction of treatment modalities and some matching variables (cognitive impairment, religious beliefs and background, meaning seeking, and problem recognition). Finally, working alliance is also affected directly by client matching variables (sociopathy, anger, motivational readiness). Given these complexities, it is clear that a matching effect predicated upon working alliance being solely an effect of but one of these three sources of variance is inherently weakened.

While it was expected that motivational readiness, prior AA, typology, anger, and problem recognition each would interact with treatment modality to affect working alliance, this turned out to be so only for problem recognition, which interacted with MET versus CBT/TSF in both the outpatient and aftercare arms.

There was evidence that working alliance was directly affected by client matching variables more often in the outpatient treatments (sociopathy, anger, and readiness) than in aftercare (readiness). Finally, working alliance was observed to be directly affected by treatment modality in both outpatient and aftercare treatments. For example, MET led to a better working alliance than TSF or CBT, and TSF appeared to lead to better agreement on the task of therapy than did either CBT or MET.

Working alliance was also conceptualized as having an effect on drinking outcome. In the outpatient arm, working alliance did indeed have a main effect on drinking outcome. However, its effect was also observed to be moderated by treatment modality (CBT vs. MET) and client matching variable (motivational readiness and typology). In the aftercare arm, working alliance did not have a main effect on drinking outcome but did interact with treatment modality (CBT vs. MET) and matching variable (motivational readiness) to affect drinking outcomes.

Readily apparent from these results, working alliance is influenced by treatment modality and matching variables in multiple ways and in turn affects drinking outcomes either directly or in combination with treatment modality or matching variables.

Treatment Structure. Structure, also invoked as a mediator to explain matching predictions, proved to be less of a discriminator between treatment modalities than expected. For the most part, MET proved not to be less structured than either CBT or TSF.

In the outpatient arm, structure was not affected by treatment modality. In the aftercare arm, structure was reported to be affected by the interaction of treatment modality (MET vs. TSF and CBT) and client typology (A vs. B). This finding indicates that, contrary to our best intentions, the delivery of treatment modality was influenced by client characteristics in the aftercare arm (canonical model type 3B). It appears that type B clients influenced MET to become more structured than when delivered to type A clients.

When structure is viewed as a factor affecting drinking outcome, it appears that in some instances it directly affects drinking (PDA in outpatient), while it may also interact with sociopathy to influence drinking outcome (DDD in outpatient, PDA in aftercare).

AA Attendance. As a final example of the complex ways in which putative mediators are both influenced by client matching variables, treatment modalities, and their interactions and in turn influence drinking outcomes, we cite AA attendance. In the outpatient arm, AA attendance was influenced by treatment modality (TSF vs. CBT, MET), client matching variable (alcohol dependence and network support for drinking), and the interaction of treatment modalities and client matching variables (TSF vs.

MET and/or CBT with network support, religiosity, meaning seeking, and alcohol dependence).

AA attendance, in turn, had both a direct effect on drinking outcome and a moderating effect in combination with TSF versus MET and/or CBT. In the aftercare arm, AA attendance was affected by treatment modality (TSF vs. MET and CBT), client attribute (gender and alcohol dependence), and the interaction of treatment modality with client attribute (TSF vs. MET and CBT with meaning seeking). AA in turn sometimes affected drinking outcome but in other analyses did not do so.

In aggregate, these examples indicate that the putative active ingredients of treatment are themselves influenced by multiple and complex factors and in turn influence drinking outcomes in variable and complex ways. Given this level of complexity, it is not surprising that matching predictions based upon simplifications of these complex dynamics would not be supported.

Client Response to Treatment

In contrast to causal chains involving treatment processes, those involving client response to treatment present a simpler picture. It was generally expected that improvement in the matching variable from before to after treatment would be predictive of better drinking outcomes. In fact, this usually turned out to be the case. Increased self-efficacy, readiness to change, and taking steps to change generally were predictive of better drinking outcomes, as was decreased temptation to drink. Anger reduction, a decrease in network support for drinking, increased social functioning, and a reduction in psychiatric symptoms were also predictive of better drinking outcomes

Nevertheless, when the effect of the interaction of improvement in the matching variable and baseline level of the matching variable was tested, it was found that the effect of the posttreatment score on drinking outcomes could be moderated by the client's baseline score. For example, while anger reduction was predictive of decreased drinks per drinking day in the outpatient arm, the effect of anger reduction on PDA was moderated by the client's baseline level of anger. In the outpatient arm, anger reduction's effect on both PDA and DDD was moderated by the client's baseline level of sociopathy. Thus, causal chains that did not incorporate the interaction of the baseline level of the matching variable into the affect of the posttreatment level on drinking would be more likely to fail.

If, for example, it is theorized that clients with networks supportive of drinking will benefit from a treatment that decreases their network's support of drinking, then it would be expected that those who had networks most supportive of drinking prior to treatment would benefit the most from this intervention. If baseline level of network support is not included as part of the interaction term for predicting drinking outcome, then the causal chain is simply predicting that those who have networks less supportive of drinking will have better drinking outcomes. The logic is, however, that those with networks supportive of drinking would be more affected by a change in network support from pretreatment to posttreatment.

While posttreatment response was usually predictive of drinking outcomes, the predictors of posttreatment response remain a mystery. Neither the matching variable itself, treatment modality, treatment process, nor the interaction of any combination of these variables predicted posttreatment response with any regularity.

Presumptive Causal Chains

A review of all tested causal chains suggests that perhaps as many as 15 may have successfully linked the interaction of treatment modality and matching variable to drinking outcome through a presumptive mediator. (It is not possible to conclude this with certainty because of various limitations in the causal chain analyses.) Again, a few summary comments are in order.

First, all canonical models (1, 2, 3A, and 3B) were represented in completely linked causal chains. Canonical model 3, which conceptualized an intervening variable as the consequence of treatment modality and client treatment matching variable, was most often successfully linked. Canonical model 2, which conceptualized a matching variable predictive of a client characteristic that interacted with a treatment process variable to affect drinking outcome, was

least often successfully linked. Model 1, which conceptualized a treatment modality leading to a treatment process that interacted with the client matching variable, was intermediate in success between models 2 and 3.

A second observation is that, as often as not, the causal chain was opposite the direction predicted. Of significance, when the causal chain was opposite that predicted, usually so was the matching effect that it was developed to explain (e.g., for sociopathy, self-efficacy, and gender). This consistency suggests that the matching effect was real; what was at variance with reality was the a priori theorizing.

A third observation, of considerable importance, is that (irrespective of whether they were hypothesized) most of the causal chains that appeared to be successfully linked occurred in the outpatient arm of study (13) rather than in aftercare (2). This is important because outpatient was a standalone treatment, whereas aftercare followed a more intensive treatment experience. It would be expected that a standalone treatment would be more likely to be amenable to a successful examination of mediators than would a treatment that was only the latter part of the whole treatment experience of the client. The greater robustness of the outpatient causal chain analyses corresponds with the greater robustness of the outpatient matching effects.

This concordance suggests that research utilizing standalone treatments is more likely to be informative.

Research Recommendations

Our interpretation of MATCH matching results is that we have observed several small single attribute by treatment modality interactions. From this conclusion, it follows that treatment strategies that rely on consideration of client attributes will need to find a better empirical foundation for their justification. We suggest some (nonmutually exclusive) avenues for future matching research.

Dedicated Treatments for Identified Clients

Treatment modalities can be developed to treat clients with singular outstanding characteristics, such as gender. Rather than expecting that CBT would be superior to TSF because of assumptions regarding its active ingredients that might affect women more than men, a treatment would be designed specifically to treat female alcohol abusers. Such treatments have been developed but have not been subjected to rigorous study as to whether they enhance treatment outcomes.

Matching Therapy Process to Client Attributes

The results of the causal chain analyses indicated that while treatment modality often did not relate to putative active ingredients in the treatment process as anticipated, these treatment process variables were themselves often predictive of client changes, including client drinking. This finding suggests the need to go beyond the "brand name" of the modality to identify differences in therapy behaviors that interact with different client attributes.

For example, in studying tapes of therapy sessions, Karno and colleagues (in press) found that over and above treatment modality, therapist behaviors interact with client characteristics to affect drinking outcomes. For clients who are assessed as high in emotional arousal, therapist behaviors that seek to focus on that arousal, as opposed to dampening the arousal, lead to better drinking outcomes. Similarly, for clients who are characterized by low emotionality, therapist behaviors that do not seek to enhance emotional arousal appear to lead to better drinking outcomes than therapist behaviors that seek to induce emotional arousal.

Karno et al. also found that therapist directiveness interacted with client resistance to affect drinking outcomes, such that clients with high resistance had poorer outcomes with directive therapists, while those with low resistance had poorer outcomes with less directive therapists. These interactions accounted for more variance than did treatment modality and suggest that variability in therapist behavior unrelated to treatment modality interacts with client characteristics to affect drinking outcomes. Karno is currently conducting a replication study of these findings with MATCH

audiovideo tapes (Karno and Longabaugh 2000).

Matching Multimodal Therapies to Multiple Patient Characteristics

Another approach to matching treatment to client attributes is to build decision trees into the therapy that will modify the treatment modalities to be received by the client on the basis of a multidimensional characterization of the client. This approach is illustrated by a clinical trial under way (Davidson et al. 2000). Gulliver and Longabaugh (2000) have developed a broad spectrum treatment that selects modules to be delivered to the client on the basis of assessments of functioning and alcohol-specific support in five domains: cognitive impairment, family relationships, occupation, residential stability, and social network. Dual classification of the client's level of functioning and support for abstinence in each of these domains leads to triage to different configurations of treatment modules.

A related approach is to match modules to clients on the basis of client choice. This approach is being used in Project COMBINE (2000). Here, a combined behavioral intervention employing the principles of motivational interviewing (Miller and Rollnick 1991) leads clients to develop their own change plans and then select from a large number of modules available to facilitate achievement of this change plan (Miller 2001). In this approach, client choice is viewed as the result of the interaction of client self-assessment and therapy alternatives offered.

Hierarchical Algorithms for Selecting Treatment Options

If we look outside the field of alcohol treatment to therapy for psychiatric problems more generally, a more elaborate theory of systematic treatment selection nests matching principles within a hierarchy of decision trees to be applied to a given patient's treatment. The systematic treatment selection model developed by Beutler and associates (2000) describes such an approach. For purposes of illustration, we will superimpose this model on our study.

Beutler and colleagues have proposed empirically supported principles for systematic selection of treatment for depression. To date, they believe that six dimensions are important in guiding selection of treatment for the patient: functional impairment, subjective distress, experienced social support, problem complexity/chronicity, level of resistance, and coping style. They find empirical support for relationships between these variables and treatment.

Seven variables of treatment believed to be important are treatment intensity/duration, emotional focus, interpersonal focus, insight versus behavioral focus, breadth of treatment focus, directiveness of therapy, and the extent to which therapy is symptom focused. They have observed that present treatment modalities are only loosely associated with these dimensions. Consequently, they believe that in order to find lawful relations between these treatment dimensions and patient characteristics, it is necessary to tailor treatment behaviors (rather than treatment modalities) to patient characteristics.

Project MATCH matching variables can be grouped within the six characterizations of patients by Beutler et al. Functional impairment subsumes alcohol involvement, alcohol dependence, psychiatric severity, cognitive impairment, social dysfunction, and by inference, lack of confidence in maintaining abstinence and temptation to drink. Subjective distress may subsume motivational readiness, alcohol problem recognition, and meaning seeking. Experienced social support could subsume network support for drinking, religious background and beliefs, and prior AA involvement. Problem complexity is poorly represented in the MATCH domain, perhaps indexed by Axis I-comorbidity. Level of resistance encompasses anger, interpersonal dependence, and as used in MATCH, gender. Finally, coping style (inwardly directed vs. externally focused) maps readily on to sociopathy, ASPD, and A versus B typology. In this conceptualization, MATCH variables within each of these domains could be considered alternative or overlapping proxy measures for each construct.

As anticipated, however, mapping the three MATCH treatment modalities into Beutler et al.'s seven dimensions of treatment is not possible. Only a few correspondences can be

estimated with any confidence. Treatment intensity/length can be mapped. Inpatient plus aftercare treatment is more extensive and intensive than standalone MATCH outpatient treatments. Within each treatment arm, MET is less extensive and intensive than either CBT or TSF. Because TSF aims to involve the client in AA during and after treatment, while CBT does not, TSF could be considered to be more intensive than CBT.

Treatment directiveness can also be mapped, with CBT and TSF considered to be more directive than MET. Symptom focus might be captured in the contrast of CBT versus MET, with CBT being more symptom focused. The three treatments cannot be differentiated with any confidence on the remaining dimensions.

Mapping Project MATCH interactions on to Beutler et al.'s interface of treatment intensity with social support and problem impairment supports their hypothesis that greater impairment and lesser support require greater treatment intensity. The interface of treatment directiveness and client resistance supports the MATCH interactions involving anger and interpersonal dependency, and MET versus CBT and TSF supports the hypothesis that greater resistance requires less directiveness. The interface of symptom focus with patient coping style is less clear, but MATCH interactions between CBT and sociopathy, ASPD, and typology were observed.

In summary, hypotheses derived from the Beutler et al. systematic treatment selection model appear to explain many of the interactions observed in MATCH. Characterization of the treatments by actually observed therapist behaviors (Karno 2000) will put several of Beutler's matching predictions to a direct empirical test.

Commentary

Irrespective of the theoretical approach taken to matching, it is clear that any model that rests on single client attribute-treatment interactions will not suffice. Rather, if matching is going to enhance treatment outcomes, multidimensional matching algorithms will be needed. It is likely that such models must encompass nonlinear relationships as well. Matching effects may not be simply additive. For example, if a client is mismatched with treatment on one important characteristic this may well nullify matching synergies on a number of other dimensions. Whether treatment research will ever reach this level of sophistication remains to be seen.

Recommendations for Alcohol Treatment Research

Up until now we confined our commentary to recommendations regarding treatment matching and treatment matching research. In this section, we close with a series of recommendations concerning alcohol treatment research more generally.

Need for Study of the Treatment Process

The causal chain analyses conducted to test the theory underlying the matching predictions indicated that we have little idea what the active ingredients of treatment are nor how they affect patients more generally as well as patients with specific attributes. The "black box of treatment" identified by Moos and Finney 20 years ago (Moos et al. 1980) still remains pretty much a mystery. While we have been able to identify some general elements of treatment (e.g., working alliance) as modest predictors of enhanced outcome, as of yet we know little about the process that brings about this better alliance.

The examination of putative mediators in the present volume revealed that ingredients of a treatment that we expected to impact differentially on clients with certain attributes frequently failed to do so in the ways anticipated. This might be regarded as simply a shortcoming of matching theory. However, an investigation of the putative active ingredients of Cognitive-Behavioral Therapy (Morgenstern and Longabaugh 2000) indicated that the active ingredients of this popular and well studied therapy were also yet to be identified. So our ignorance of how treatments work extends beyond matching theories to our theories of treatment.

As the present volume has amply demonstrated, our theories about how treatment

works need to be operationalized and tested. Not to do so permits us to forge ahead while operating under false assumptions. That is no way to build a knowledge base about treatment.

So, the strongest recommendation we have regarding treatment research is that the treatment process itself has to be studied, through mediator analysis or other means, and related to treatment outcomes. We need to know what the mechanisms are that give rise to better outcomes. To do so will markedly affect the sophistication of theories about treatment.

Need for Study of the Treatment Context

As has been said on numerous occasions, treatment is but one small event in the life of an alcohol troubled person. Treatment takes place in a context: What factors bring the client to treatment, what is going on in the client's life while the treatment is delivered, and what environmental context awaits the client completing treatment? Until we understand these treatment context variables and bring them into our treatment models, we will be unable to account for major variance in treatment outcomes.

In the present study, such effects were implied by the disappearance of within-treatment matching effects after treatment was completed as well as the treatment by time interaction effects observed while treatment was going on. Also pertinent, the one matching effect that appeared 3 years after treatment identified variables outside of the treatment itself as a moderator (network support for drinking) and a mediator (attendance at AA meetings). Environmental context and interface needs to be brought directly into our models and either controlled for or systematically varied.

Measure Outcomes at the End of Treatment

In research on psychosocial treatments we usually designate "outcome" as a sustained period following treatment completion. We do so in the belief that the changes that occur because of treatment are sustaining and will enable the client to successfully cope with drinking during this posttreatment period. How long treatment effects last is an important question, but it is not the only important question. Also important is: Does treatment affect change in the client while it is ongoing or by its completion? If this does not happen, clearly the theory of treatment is either incorrect or has not been implemented as conceptualized. Thus, unless the theory underlying the treatment clearly specifies that its effects are not to become apparent until a specified time after treatment has been completed, the posttreatment observation is not a direct test of the theory.

This expectation of sustaining benefits of treatment is not present in most studies of pharmacotherapies. Here, the question is usually whether the client is changed while taking the medication. Whether the client continues to drink less after the completion of a trial of naltrexone, for example, is viewed as an interesting question but not one that speaks to the efficacy of the drug itself. Why is it that we should expect so much more from psychosocial therapies than from pharmacotherapies?

Studies of mechanisms should be designed to test as directly as possible the effects of these mechanisms on the designated dependent variable. Once mechanisms have been identified, they can be continued or terminated as part of the treatment. Studies of behavioral therapies (like studies of pharmacotherapies) suggest that their effects are likely to last as long as the treatment is ongoing. If that is so, what is it that precludes our conceiving of psychosocial treatments as having maintenance phases which might go on indefinitely? Treatment so conceived might well prove to be more cost effective than repeated "acute phase" treatments.

Measuring Outcome

In alcohol treatment outcome studies, the primary dependent variable is typically a measure of drinking (frequency, intensity, total volume, or some composite of these). Secondary measures of outcome may include measures of functioning, subjective well-being, service utilization, and the like. Project MATCH was prototypic in this regard. Yet many of our treatment theories do not specify a direct impact of the treatment on drinking per se. In MATCH, for example, it was expected that TSF clients would stop drinking as a consequence of working the

12 steps that require major cognitive, attitudinal, and behavioral changes. CBT was expected to improve the client's cognitive-behavioral coping skills, while MET was focused on increasing client motivation to change. Assuming these foci of treatment are accomplished, a change in drinking is expected to follow.

Depending upon the underlying theory of treatment, drinking may not be conceived of as the primary dependent variable. For example, drinking may be seen as one mediator of enhanced quality of life. If people drink less, they are expected to experience fewer negative consequences from drinking and positive consequences from not drinking. However, if cessation of drinking does not lead to one, the other, or both of these effects, enhanced quality of life will not occur. Thus, change in drinking status is one link in the supporting causal chain.

In contrast is a theory that relies on enhanced quality of life as a mediator of change in drinking. In this scenario, the treatment leads to increased quality of life that in turn results in a reduction in drinking. If quality of life does not improve, drinking will not diminish. Thus, depending upon the underlying causal chain, drinking may be a mediator variable, a dependent variable, or both.

Our interest in alcohol treatment is in reducing or eliminating the negative consequences of drinking rather than ending drinking itself. Thus, the negative consequences from drinking should be the primary measure of outcome. Measures of drinking may be important or critical to testing various parts of the theory of treatment, but they are neither necessary nor sufficient as primary measures of alcohol treatment outcomes.

The Need for Inclusion of Multiple Treatment Sites

If Project MATCH has taught us only one thing, it is the risk involved in conducting single-site studies. Had we not included multiple sites and different study arms and had conducted the study in just one treatment site, we would have found an effect that we assumed could be generalized. Instead, subtle and as yet unidentified variables led to inconsistencies of effects across sites and settings. In order to begin to test for generalizability of effects beyond a single treatment site, it would appear that, despite the additional cost, inclusion of at least two sites in a treatment outcome study are necessary.

Conclusion

Our predictions as to how treatments would be distinctive in ways that would differentially impact clients with specific attributes were woefully inadequate. A major conclusion is that we do not know yet how our treatments work. The theories developed to support matching predictions were not adequate for the task.

Several caveats are in order. While the development and testing of the a priori matching hypotheses was exceedingly rigorous and analytically appropriate, in hindsight, the development and testing of causal chains was not. There are several reasons for this.

First, when Project MATCH was initiated (1989), mediational analyses were not widely known to the alcohol treatment research community. The seminal paper by Baron and Kenny had only been published in the Journal of Personality and Social Psychology in 1986. Holmbeck's paper clarifying some of the ambiguities in testing for mediational effects was not published until 1997, after the MATCH causal chain analyses had been completed. Thus, from an historical perspective, the causal chain analyses were undertaken in Project MATCH before the methodological and statistical techniques were fully developed. Even to the present, analyses have been confined largely to testing for mediation of main effects. In Project MATCH, we sought to push the envelope by initiation of mediational analyses of moderator effects, a novel topic even a decade later.

Compounding this difficulty, the senior scientists selected to conduct the MATCH trial were trained in an era before mediational analysis had become prominent. As a group, we were not especially well trained in these new methodologies. Further, there were differences among MATCH investigators regarding the importance of a priori theory development. Some investigators believed, in reflection quite accurately, that elaborate theory development was

premature and expected that the facts would emerge in due course.

Because we were breaking new ground, there was not a consensus on the best analytic tools to test the causal chains for moderating variables. Consequently, each matching hypothesis team was free to adopt whatever analytic tools made sense to them for testing their causal chain. The variety of approaches adopted is quite apparent as one reads across the chapters.

Because of the need to implement Project MATCH on a timeline close to that projected and funded, the study was implemented before the theories underlying the matching predictions had been fully elucidated. This meant that matching hypotheses were guided by a theoretical orientation rather than derived from a completed theory of which they were a part. Clearly, many of the assumptions underlying the matching predictions proved to be incorrect. Had theories been fully explicated earlier, it would have been apparent that many involved assumptions that would be untenable or at least previously untested.

There was also the necessity of relying on the assessment battery that had been adopted and was already being implemented prior to the completion of causal chain development. This resulted in proxy indices for pivotal mediators for many of the causal chains tested. Further, the number of measures of treatment process that transcended the three treatments was limited by assessment time and MATCH priorities. As so many matching hypotheses were to be tested, the allocation of resources to assessing any one was necessarily limited. Finally, while the three treatments developed for testing matching hypotheses were believed to be distinctive from one another, they were not developed to be distinctive in their relations to the matching variables, which were still to be decided upon.

It is our hope that the Project MATCH study of mediational analyses of moderator effects, now concluded, will provide guidance to treatment researchers as they develop theories and test interventions to enhance treatment outcomes.

Acknowledgment

This project was supported by a series of grants from the National Institute on Alcohol Abuse and Alcoholism as part of the Cooperative Agreement on Matching Patients to Alcoholism Treatments.

References

Baron, R.M., and Kenny, D.A. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51:1173–1182, 1986.

Beutler, L.E.; Clarkin, J.F.; and Bongar, B. *Guidelines for the Systematic Treatment of the Depressed Patient.* New York: Oxford University Press, 2000.

Brehm, S.S. *The Application of Social Psychology to Clinical Practise.* New York: Halstead Press, 1976.

Brehm, S.S,. and Brehm, J.W. *Psychological Reactance: A Theory of Freedom and Control.* New York: Academic Press, 1981.

COMBINE Research Group. "Symposium on Design and Rationale for COMBINE, a Multisite Study on Combining Medications and Behavioral Interventions for Alcohol Dependence." Research Society on Alcoholism Annual Meeting, Denver, CO, June 2000.

Davidson, D.; Gulliver, S.; Longabaugh, R.; and Swift, R. Post-Treatment effects of Naltrexone. R01AA12696, NIAAA, 2000.

Gulliver, S., and Longabaugh, R. "A Manual for Broad Spectrum Treatment." Unpublished manuscript, 2000.

Holmbeck, G.N. Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology* 65: 599–610, 1997.

Karno, M., and Longabaugh, R. Psychotherapy Process in Alcoholism Treatment Matching. R01AA12155, NIAAA, 2000.

Karno, M.P.; Beutler, L.E.; and Harwood, T.M. Interactions between psychotherapy procedures and patient attributes that predict alcohol treatment effectiveness: A preliminary report. *Addictive Behaviors*, in press.

McCrady, B.S.; Epstein, E.E.; and Hirsch, L.S. Maintaining change after conjoint behavioral alcohol treatment for men: Outcomes at 6 months. *Addiction* 94(8):1381–1396, 1999.

McLellan, A.T.; Woody, G.E.; Luborsky, L.; O'Brien, C.P.; and Druley, K.A. Increased effectiveness of substance abuse treatment: A prospective study of patient-treatment "matching." *Jour-*

nal of Nervous and Mental Disease 171:597–605, 1983.

Miller, W.R., ed. "Combined Behavioral Intervention: Therapist Manual." Project COMBINE, unpublished manuscript, 2001.

Miller, W.R., and Rollnick, S. *Motivational Interviewing: Preparing People for Change.* New York: Guilford Press, 1991.

Moos, R.; Finney, J.W.; and Cronkite, R. The need for a paradigm shift in evaluations of treatment outcomes: Extrapolations from the Rand research. *British Journal of Addiction,* 75:347–350, 1980.

Morgenstern, J., and Longabaugh, R. Cognitive-behavioral treatment for alcohol dependence: A review of evidence for its hypothesized mechanisms of action. *Addiction* 95(10):1475–1490, 2000.