Addiction (1999) 94(1), 31–69

COMMENTARIES

Comments on Project MATCH: matching alcohol treatments to client heterogeneity

Editor's introduction

Project MATCH is an 8-year multi-site study of how patients respond to different treatment approaches designed to help them recover from alcohol problems. This nationwide US clinical trial has involved 1726 patients, 25 senior investigators, 80 therapists and many more research assistants and support staff, at over 30 participating institutions and treatment agencies. As the largest trial of psychotherapies ever undertaken, both the design and the findings of this study are worthy of commentary by experts in the field. The following Summary is intended to provide a synopsis for the major findings of Project MATCH. Details of these findings have been reported more fully in the references listed at the end of the summary. Four of these articles (Project MATCH Research Group, 1997a, 1997b, 1998a, 1998b) were used as the basis for commentaries which follow and were solicited from leading authorities in the field of treatment research.

Summary of Project MATCH

Research reports from the past 25 years have suggested that treatment outcomes can be improved by carefully matching individuals, based upon their personal characteristics, to specific therapeutic approaches. In a 1989 report, the Institute of Medicine of the US National Academy of Sciences strongly advocated research on patient-treatment matching. Project MATCH was designed to test the general assumption that matching would improve treatment outcomes, and in particular to test specific matching effects hypothesized on the basis of prior matching findings (Project MATCH Research Group, 1993).

Methods

The trial employed three individually delivered treatments that differed widely in philosophy and practice: (1) a 12-session Twelve-Step Facilitation Therapy (TSF) designed to help patients become engaged in the fellowship of Alcoholics Anonymous; (2) a 12-session Cognitive–Behavioral Therapy (CBT) designed to teach patients coping skills to prevent relapse to drinking; and (3) a Motivational Enhancement Therapy (MET) designed to increase motivation for and commitment to change, consisting of four sessions scheduled over 12 weeks. A total of 1726 individuals, varying widely in personal characteristics and alcohol problem severity, were assigned randomly to the three treatments at sites located in nine communities across the United States. The three treatments were tested in parallel studies in two types of settings: outpatient and aftercare. There were 952 outpatients (72% males), and 774 aftercare patients (80% males) recruited immediately following inpatient or intensive day hospital treatment. Specific a priori hypotheses were derived from previous research to predict which individuals would respond best to the three treatments. The following patient characteristics were investigated: severity of alcohol involvement, cognitive impairment, conceptual level, gender, meaning-seeking, readiness for change, psychiatric severity, social support for drinking, sociopathy, typology classification (Type A-Type B), alcohol dependence, anger, antisocial personality, assertion of autonomy, psychiatric diagnosis, prior engagement in AA, religiosity, self-efficacy and social functioning.

Outcome evaluations were conducted at 3-month intervals during the first 15 months of follow-up at all sites. In addition, 39-month fol-

0965–2140/99/010031–39 $9.50

Carfax Publishing Limited

low-ups were completed at the five outpatient sites.

The Project MATCH study was carefully designed and successfully implemented. Patients' participation in treatment was excellent. Patients who entered the study attended, on average, over two-thirds of their scheduled treatment sessions. Over 90% of the patients provided data for each follow-up point. The content of the treatments was carefully controlled, and analyses showed that the three treatments as delivered were very different from each other in expected ways (Carroll *et al.*, 1998). Blood tests as well as interviews with patients' families and friends confirmed patient self-reports of drinking (Babor *et al.*, in press).

Results

Patients in all three treatment conditions showed major improvement not only on drinking measures, but in many other areas of life functioning as well (Project MATCH Research Group, 1997a). Before treatment, Project MATCH patients averaged about 25 drinking days per month. The frequency of drinking decreased four-fold to fewer than 6 days per month after treatment. The volume of drinking also decreased dramatically. Before treatment, Project MATCH patients averaged about 15 drinks per day when drinking. This decreased five-fold to about three drinks on an average drinking day. Project MATCH patients showed significant decreases in depression, alcohol-related problems and in the use of other drugs, as well as improvement in liver function. Improvements that occurred during treatment were well maintained throughout the 12 months following the end of treatment, the period during which most relapses typically occur. At 1 year after treatment, for example, Project MATCH patients were still averaging more than 25 alcohol-free days (85%) in a 30-day month. A 39-month follow-up of the outpatient sample indicated continued maintenance of these high abstinence rates (Project MATCH Research Group, 1998b).

Although Project MATCH was intended to study patient–treatment interactions, the design did permit comparisons across treatments. There were few clinically significant outcome differences among treatments in either the outpatient or aftercare arm of the study. One exception is that outpatients who received TSF were more likely to remain completely abstinent (24%) during the year after treatment than those in the other two groups (14% and 15%). Also, during the treatment phase, small but statistically significant differences among treatments were found only in the outpatient arm on measures of alcohol consumption and alcohol-related negative consequences. Forty-one per cent of CBT and TSF clients were abstinent or drank moderately without alcohol-related consequences compared with 28% of MET clients (Project MATCH Research Group, 1998a).

The central purpose of Project MATCH was to determine whether patient–treatment matching improves outcome. Of the first 10 matching variables, only one a priori prediction was supported (Project MATCH Research Group, 1997a). Outpatients with few or no psychological problems had more abstinent days during most of the year following treatment when given Twelve-Step Facilitation treatment than when given Cognitive–Behavioral Therapy. For example, during the sixth month after treatment, patients assigned to TSF were abstinent on 87% of days, compared to 73% for those assigned to CBT. This effect persisted through most of the year following treatment, but by the end of the follow-up period there were no significant differences (83% for TSF compared to 80% for CBT). Contrary to prediction, no significant differences were seen among the three treatments for patients with moderate to severe psychological problems.

Several additional matching predictions were supported from the second set of analyses (Project MATCH Research Group, 1997b): (1) outpatients high in anger and treated in MET had better post-treatment drinking outcomes than outpatients treated in CBT; (2) aftercare clients high in alcohol dependence had better post-treatment outcomes in Twelve-Step Facilitation; low dependence clients did better in CBT. Finally, one matching effect, seen in outpatients, did not occur until the 3-year follow-up (Longabaugh *et al.*, 1998). Clients with a social network supportive of drinking did better in TSF than in MET. Among clients in the upper half of the distribution of network support for drinking, TSF patients reported abstinence on 83% of the days versus 66% for the MET clients. This differential of 17 percentage points was the largest matching difference observed in the trial.

Beyond the testing of specific matching hypotheses, the following findings have also been reported.

- Patients given MATCH treatments as aftercare showed about 15% higher rates of abstinence than did patients treated only as outpatients, despite the fact that aftercare patients had substantially more severe problems on nearly every pre-treatment measure. The reasons for this difference are unclear. It might be due to pre-existing differences between aftercare and outpatients, to motivational self-selection of aftercare patients (who had to complete intensive treatment before entering Project MATCH), to a period of abstinence for aftercare patients prior to Project MATCH treatment, or to the prior treatment received by aftercare patients.
- The percentage of continuous abstainers, while informative, underestimates the rate of favorable outcomes. Among aftercare patients, 35% sustained complete abstinence throughout 1 year after the end of treatment, but 60% never had three consecutive days of heavy drinking. Similarly for outpatients, 20% maintained complete abstention through 12 months of follow-up, but 50% never had three consecutive heavy drinking days.
- Some personal characteristics of patients seem to impact outcomes regardless of the type of treatment given. Better overall outcomes occurred for patients who reported less social support for continued drinking. Among outpatients, higher initial motivation for change was a strong predictor of better treatment outcomes.
- As found in many previous studies, patients who continued longer in treatment showed better outcomes. The more sessions a patient attended in Cognitive–Behavioral or Twelve-Step treatment, the better the outcome. For the four-session Motivational Enhancement Therapy condition, however, there was no relationship between length of attendance and positive outcomes.
- There were relatively few outcome differences among three treatments designed to differ dramatically in philosophy and procedures.
- Wherever differences were observed between treatments, they favored the Twelve-Step Facilitation therapy. These differences were generally modest in size, and only one (continuous abstinence for outpatients) endured throughout all 12 months of follow-up. Project MATCH demonstrated that outcomes from TSF are as favorable as those for other well-tested approaches in the treatment of alcohol problems.
- After completion of treatment, there were only a few outcome differences between the four-session Motivational Enhancement Therapy and the two 12-session treatments. Three-quarters of the predicted matching effects were based on the expectation of different (usually poorer) outcomes for selected patients in the less intensive Motivational Enhancement Therapy condition. None of these predictions was confirmed in the year after completion of treatment, and no relationship was found between severity of alcohol problems and response to Motivational Enhancement Therapy versus other treatment.

Conclusions

Viewpoints differ on how clinically significant these single characteristic matches are, given their overall variability over time and the rather modest size of most of the effects. Matches that may have clinical significance include psychiatric severity, anger, social support for drinking and alcohol dependence. The results suggest that triaging clients to individual therapy, at least based on the attributes and treatments studied in Project MATCH, is not a compelling requirement for treatment success as previously believed. The matches found, however, are reasonable considerations for clinicians to use as starting points in the treatment planning process.

Acknowledgement

Project MATCH was supported by grants under a cooperative agreement funded by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and implemented by nine clinical research units and a data coordinating center. The project was initiated and is administered by the Treatment Research Branch of NIAAA. The Project MATCH Research Group consists of John P. Allen, Margaret E. Mattson, William R. Miller, J. Scott Tonigan, Gerard J. Connors, Robert G. Rychtarik, Carrie L. Randall, Raymond F. Anton, Ronald M. Kadden, Mark Litt, Ned L. Cooney, Carlo C. DiClemente, Joseph Carbonari, Allen Zweben, Richard H. Longabaugh,

Robert Stout, Dennis Donovan, Thomas F. Babor, Frances K. DelBoca, Bruce Rounsaville, Kathleen Carroll and Philip W. Wirtz.

*Selected Project MATCH Publications**

BABOR, T. F., STEINBERG, K., ANTON, R. & DEL BOCA, F. (1999) Talk is cheap: measuring drinking outcomes in clinical trials, *Journal of Studies on Alcohol*, in press.

CARROLL, K., CONNORS, G. & COONEY, N. *et al.* (1998) Discriminability and integrity of treatments for alcoholism: results from Project MATCH, *Journal of Consulting and Clinical Psychology*, 66, 290–303.

DEL BOCA, F. K. & BROWN, J. M. (1996) Issues in the development of reliable measures in addictions research: introduction to Project MATCH assessment strategies, *Psychology of Addictive Behaviors*, 10, 67–74.

DONOVAN, D. & MATTSON, M. E. (Eds) (1994) Alcoholism treatment matching research: methodologic and clinical approaches, *Journal of Studies on Alcohol*, suppl. 12, pp. 171

KADDEN, R. P., CARROLL, K. & DONOVAN, D. *et al.* (1992) *Cognitive–Behavioral Coping Skills Therapy Manual: a clinical research guide for therapists treating individuals with alcohol abuse and dependence*, Project MATCH monograph series, vol. 3, DHHS Pub. No. (ADM) 92–1895.

LONGABAUGH, R., WIRTZ, P. W., ZWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, Alcoholics Anonymous and long-term matching effects, *Addiction*, 93, 1313–1333.

MILLER, W. R., ZWEBEN, A., DICLEMENTE, C. & RYCHTARIK, R. (1992) *Motivational Enhancement Therapy Manual: a clinical research guide for therapists treating individuals with alcohol abuse and dependence*, Project MATCH monograph series, vol. 2, DHHS Pub. No. (ADM) 92–1894.

NOWINSKI, J., BAKER, S. & CARROLL, K. (1992) *Twelve Step Facilitation Therapy Manual: a clinical research guide for therapists treating individuals with alcohol abuse and dependence*, Project MATCH monograph series, vol. 1, DHHS Pub. No (ADM) 92–1893.

PROJECT MATCH RESEARCH GROUP, PROJECT MATCH (1993) Rationale and methods for a multisite clinical trial matching patients to alcoholism treatment, *Alcoholism: Clinical and Experimental Research*, 17, 1130–1145.

PROJECT MATCH RESEARCH GROUP (1997a) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes, *Journal of Studies on Alcohol*, 58, 7–29.

PROJECT MATCH RESEARCH GROUP (1997b) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.

PROJECT MATCH RESEARCH GROUP (1998a) Matching alcoholism treatments to client heterogeneity: treatment main effects and matching effects on drinking during treatment, *Journal of Studies on Alcohol*, 59, 631–639.

PROJECT MATCH RESEARCH GROUP (1998b) Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes, *Alcoholism: Clinical and Experimental Research*, 22, 1300–1311.

*To obtain copies of these publications, contact the Scientific Communications Branch, National Institute on Alcohol and Alcoholism, Willco Building, Suite 409, 6000 Executive Blvd, Bethseda, MD 20892–7003, USA.

The unsinkable Project MATCH

Frederick B. Glaser

Project MATCH is a noble enterprise. Its dimensions are awesome. By its own admission it is "the largest, statistically most powerful, psychotherapy trial ever conducted" ([1], p. 25). Carried out at nine separate sites by 22 principal investigators, it involved almost 2000 patients, many of whom were followed for more than 3 years. The echelons of its administrative structure and the complexity and rigor of its statistical analyses are unparalleled. Its preparatory papers alone occupy an entire journal supplement, and exude enthusiasm that the mysteries of patient–treatment matching were soon to yield to science.

Alas, it was not to be. Announced initially at a press conference, the early results suggested that, although a high proportion of patients in all three interventions achieved positive outcomes, the findings were essentially negative with regard to matching. Nine of 10 carefully selected primary matching variables failed to exert any significant effect on outcome at 1 year. The general reaction was one of shock and dismay. In a communiqué electronically disseminated by a conference attendee, those who worked in the field were urgently advised to seek employment elsewhere. Insurance companies would interpret the findings as demonstrating the superior cost effectiveness of Twelve-Step programs and would withdraw reimbursement for other forms of treatment.

Such a scenario of pride and high hopes (*hubris*) cruelly and suddenly dashed (*ate*) has been familiar since classical times. A possibly whimsical simile suggests itself: Project MATCH as the *Titanic* of treatment outcome studies. Like Project MATCH, the great ship was large, the largest man-made object to that point in history. It was complex; it generated enormous enthusiasm; and it sank like a stone on its maiden voyage with great loss of life.

Yet the *Titanic*, like the fabled phoenix, rose

from its ashes. Although many lives were lost, some passengers survived the tragedy. Lessons were learned in such areas as naval architecture and communications that contributed significantly to seaworthiness. The legend of the *Titanic* has recently been embodied in the most successful motion picture to date, attesting to its continuing currency. There are rumors that an attempt may be made to raise it from the bottom of the ocean.

Project MATCH has undergone resuscitation within a much shorter time frame. An analysis of the *secondary* matching variables[2] has yielded substantially more positive results than that of the primary variables; many hypotheses were lost, but some survived. A 3-year follow-up study of those assigned to the outpatient arm of the study[3] has yielded a robust matching effect that was not apparent at 1-year follow-up.[4]

Based partly upon this field reversal, an educated guess can be essayed regarding the impact of Project MATCH. It seems likely to vary over time. Initially, the impact may be negligible, perhaps even negative. As time passes, however, the impact may become increasingly positive. Ultimately, the study may be recognized as the crucial investigation it has truly been.

Most research studies have little immediate effect upon the treatment enterprise. The RAND report, an earlier large-scale research effort, received wide publicity because of its findings in the area of moderate drinking. Far from having the horrendous impact that was feared by some, it turned out that the vast majority of treatment personnel, as well as patients, had never heard of the study.[5] The same is likely to be true of Project MATCH. It appears at this point to have a small audience except in the research community. Not all researchers have reacted positively to the study. In a recent (July, 1997) review of treatment-matching,[6] the authors pass over Project MATCH with a single sentence (p. 946) and go on to assert that matching research is "expanding in size and complexity" and is "an essential public health need" (p. 961) crucial to the future of the treatment enterprise.

The bleak negativity of the initial report of results may particularly delay the impact of Project MATCH. Having negotiated the tortuous prose of this lengthy report, and come away with very little of immediate utility, many will not read the subsequent papers. Who needs more bad news? They will thus be left with an incomplete and incorrect understanding of Project MATCH. It is a situation reminiscent of one in which allegations of scientific fraud were made against two eminent researchers in our field. Many are aware of the initial allegations; few are aware that the researchers have subsequently been totally (cf.[7]) and repeatedly exonerated.

Other factors will contribute to the initial lack of impact. Except in a few academic settings, the interventions chosen for study are not in general use. In the real world of treatment it is group approaches, rather than the individual approaches utilized exclusively in the study, that are the stock in trade. Current interventions also lack the detailed specificity of the Project MATCH treatment manuals, making it difficult to be certain to what species of intervention one is matching patients.

Whatever the validity of the Project MATCH experimental design may be, a study of matching in which none of the multitude of subjects was actually matched to a particular treatment lacks persuasiveness. The thicket of highly technical statistical procedures required to extract meaning from the study data impart an aura of unreality. In the face of widespread appreciation of such concepts as discontented sobriety and the multi-faceted nature of treatment outcome, the study's considerable reliance upon two measures of alcohol consumption as its principal outcome criteria is unconvincing. Perhaps at some future point a more direct and understandable report of the study as a whole, with its technical aspects relegated to appropriate appendices, will remedy these difficulties.

Beyond this, there are multiple aspects of the current service delivery system that will inhibit a high level of impact for the study. The use of client variables in matching requires a careful assessment prior to treatment (cf.[6]); such an assessment was a major part of the overall effort of Project MATCH. Most existing treatment programs would not be able to carry out such an assessment. Among other difficulties, comprehensive pre-treatment assessment is not covered by current insurance reimbursement policies. To alter this will be a particular problem for the United States; it is much more difficult to secure the application of results from treatment research studies here than in countries that have a single, universal health insurance scheme.[8]

Most existing programs do not offer more than a single treatment option. In such circumstances assessment for the purpose of matching is

superfluous. Nor do most treatment programs systematically and objectively examine outcome. If only as a continuing check on the cogency of matching paradigms, regular outcome monitoring, at the very least, is a necessity; at present, it is the exception rather than the rule. Longer periods of more detailed follow-up are still largely limited to research studies.

However, the treatment enterprise continues to evolve, under a variety of pressures, and its evolution appears to be in directions that may ultimately render the results of Project MATCH increasingly relevant. Such essential processes as comprehensive assessment, matching to a spectrum of highly specific interventions, and regular outcome determination may at present defy the capabilities of individual treatment programs; but they could logically and efficiently accrue to aggregations of treatment programs—that is, to treatment systems. The future of treatment may lie in the development of such systems (cf.[9]).

When and if such a future evolves, Project MATCH will come into its own. As the *Titanic* has become emblematic of both the glories and the hazards of transoceanic travel, both Project MATCH's successes and its shortcomings may in time prove to be seminal. The study may ultimately be seen as the critical juncture at which treatment research turned the corner and gathered momentum for the quantum leap required for the evolution of truly effective and efficient treatment.

FREDERICK B. GLASER
c/o Walter B. Jones Alcohol and Drug Abuse, East Carolina University School of Medicine, Treatment Centre, 2577 West Fifth Street, Greenville, North Carolina 27834, USA

1. PROJECT MATCH RESEARCH GROUP (1997a) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes, *Journal of Studies on Alcohol*, 38, 7–29.
2. PROJECT MATCH RESEARCH GROUP (1997b) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.
3. PROJECT MATCH RESEARCH GROUP (1998). Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes, *Alcoholism: Clinical and Experimental Research*, 22, 1300–1311.
4. LONGABAUGH, R., WIRTZ, P. W., ZWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, Alcoholics Anonymous, and long-term matching effects, *Addiction*, 93, 1313–1333.
5. HINGSON, R., SCOTCH, N. & GOLDMAN, E. (1977) Impact of the "RAND Report" on alcoholics, treatment personnel, and Boston residents, *Journal of Studies on Alcohol*, 38, 2065–2976.
6. GASTFRIEND, D. R. & MCLELLAN, A. T. (1997) Treatment matching: theoretical basis and practical implications, *Medical Clinics of North America*, 81, 945–966.
7. DICKENS, B. M., DOOB, A. N., WARWICK, O. H. & WINEGARD, W. C. (1982) *Report of the Committee of Enquiry into Allegations Concerning Drs. Linda and Mark Sobell* (Toronto, Addiction Research Foundation).
8. SPITZER, W. O. & STARFIELD, B. (1977) Health services research can make a difference! *New England Journal of Medicine*, 297, 1406.
9. INSTITUTE OF MEDICINE (1990) *Broadening the Base of Treatment for Alcohol Problems* (Washington, DC, National Academy Press).

Some common methodological criticisms of Project MATCH: are they justified?

Nick Heather

As befits a project that has rightly captured the attention of almost everyone in the alcohol problems treatment field, Project MATCH has already been subject to a number of methodological criticisms. I have not yet seen these criticisms in print but they are to be heard whenever the project is discussed at conferences and other meetings of interested parties. My intention here is to describe these criticisms and give a personal view on them. I shall assume for the purposes of this exercise that, although the matching findings that did emerge at 1- and 3-year post-treatment follow-ups (Project MATCH Research Group, 1997a; 1998b) may well be clinically useful, the *general* hypothesis that matching would improve overall success rates of treatment was not confirmed.

(1) *Too much assessment and too many follow-ups.* A common criticism refers to the very large amount of assessment time associated with the project (e.g. 8 hours of initial assessment) and the frequent research follow-ups (e.g. five follow-up interviews in the first year post-treatment). The impact of the follow-ups is illustrated by the anecdote that a client who had received MET could remember the name of her follow-up researcher but

not of her therapist. The significance of the criticism is that the intensive research contact, which could be presumed to have had some kind of therapeutic benefit, reduced possible differences between the effects of the three treatments and may have prevented matching findings from emerging.

It is clear that, in designing the trial, the Project MATCH investigators were concerned primarily with maximizing internal validity and only secondarily with external validity. This is a justifiable position, especially in view of the considerable resources used by the project. Certainly, to provide a proper test of all the specific matching hypotheses that were identified, a large amount of assessment time was inevitable. So too, frequent follow-ups enabled the investigators to conduct the latent growth analyses that provided stringent tests of hypothesized matching effects over time. Nevertheless, there is probably some force in this criticism; it may be that possible matching effects were swamped in this way. If so, however, they cannot surely have been very substantial or robust effects.

(2) *The treatment received was "too good".* A related criticism is that, owing to the presence of highly trained and skillful therapists, strict quality control over manual-driven treatments and possibly some kind of Hawthorne effect, the treatment clients received in the trial was much better than that provided in the "real world" of alcohol problems treatment. This produced a ceiling effect in treatment outcome which again may have obscured possible matching effects.

This criticism seems to rest on two misunderstandings. First, the therapists, although given a thorough preparation for the trial, were not trained beyond a level that should be expected in a well-run treatment service. In any event, it would have been a strange idea to have deliberately delivered treatment of inferior quality just because it was thought that this is what happens in the real world; if quality-controlled, manual-driven treatments are not delivered by competent therapists in the real world, the simple answer is that they should be.

Secondly, although treatment outcome was good, it was not all that good! In other words, there was sufficient variation in treatment outcome to allow matching effects to have appeared if they existed. Part of the problem here is the way in which outcome is presented in Project MATCH publications—as percentage of days abstinent and drinks per drinking day averaged over clients. It would have been more meaningful to clinicians to have also presented conventional outcome categorizations (e.g. successful/improved/unimproved) purely for descriptive purposes.

(3) *Attendance at Alcohol Anonymous may have confounded treatment effects.* This criticism works in two ways. First, clients from all three treatment groups made use of AA during and after treatment, with a mean attendance of three meetings per month (Longabaugh *et al.*, 1998). Once more, this may have blunted differences in treatment effects and obscured matching effects. However, it is difficult to know what could have been done about this problem; in the cultural context of the United States, it would presumably have been impossible to try to limit AA attendance among CBT or MET clients.

The opposite kind of criticism is to say that TSF clients had more AA attendance than those in the other two treatment groups (over six meetings per month for TSF clients compared with just over two per month for MET and one per month for CBT), not surprisingly in view of the fact that part of the treatment method in TSF was to encourage AA attendance. Thus it is claimed that the TSF group received more "treatment" than the other two groups, a confound relevant to the discovery of possible matching effects as well as main effects of treatment. Again, it is difficult to know how this problem could have been resolved. Perhaps clients in CBT could have been encouraged to attend Rational Recovery, Secular Organizations for Sobriety or some other, secular form of mutual aid influenced by cognitive–behavioural principles, assuming that these groups had a sufficient presence in the locations of the treatment sites.

(4) *Polydrug users were excluded from the trial.* This is another commonly voiced criticism, with the implication that Project MATCH results are irrelevant to treatment services in

the United States and elsewhere where polydrug users are the typical clientele. The first thing to note is that the criticism is not true: clients with DSM III-R diagnoses of substance *abuse* were not excluded (see Project MATCH Research Group, 1997a). Further, clients with *dependence* on cannabis were also not excluded; but should those with other substance dependence or any intravenous drug use in the previous 6 months have been excluded?

This is again a matter of the optimal trade-off between internal and external validity. The inclusion of drug dependent individuals may well have increased the latter but would undoubtedly have decreased the former, to the point perhaps where clearly interpretable findings would not have been possible. For example, to have included clients currently receiving some form of methadone treatment for opioid dependence would likely have produced a methodological and analytic nightmare. There are, however, serious methodological problems here which should receive close attention in future research.

(5) *Total abstinence was the exclusive goal of treatment*. This criticism is quickly dealt with. In the cultural context of alcohol problems treatment in the United States, and in a trial which aimed to replicate commonly used treatment modalities, it would not have been possible to use a moderation goal, even for clients with low levels of dependence. In the different cultural context of Britain, the recently-funded UK Alcohol Treatment Trial will include both abstinence and moderation goals in the design.

An incidental observation is that, in the Project MATCH treatment procedures, abstinence was more strongly encouraged in TSF and CBT than in MET, where treatment goal was to some extent seen as the client's responsibility. This may partly have accounted for the higher level of drinking and associated problems found in MET clients than in the other two groups during the treatment period (Project MATCH Research Group, 1998a). However, it is interesting and relevant to ancient disputes about the relative merits of the two goals that these differences had disappeared by the 1-year post-treatment follow-up.

(6) *The matching contingencies studied should either have been "thinner" or "fatter"*. This basic criticism has to do with the kind of matching it was decided to study. On the one hand, some clinicians contend that the failure of Project MATCH to find much evidence in favour of treatment matching flies in the face of their clinical experience, but this probably involves a misunderstanding of Project MATCH findings. As I interpret them, these findings apply only to *systematic* matching, in the sense of a formal treatment system with rules to channel clients into specific types of therapeutic approach. They have little or no bearing on the traditional clinical skill of tailoring treatment to the unique needs, goals and characteristics of a particular client in the individual case (i.e. "thinner" matching). Whether or not this kind of clinical skill adds to the effectiveness of treatment is unknown but confidently assumed. In any event, the place in treatment of this clinical skill is untouched by the results of Project MATCH. Neither do the results affect the kind of client–treatment matching that occurs informally when services dealing with medical, financial, psychiatric, family or legal problems are *added on* to a basic treatment programme (see McKay & McLellan, 1998).

The "fatter" matching criticism asserts that it would have more instructive to have looked for client–treatment matches in comparisons involving broader clinical or organizational categories of treatment delivery—inpatient vs. outpatient, group vs. individual therapy, pharmacotherapy vs. psychosocial therapy in general. Such studies would no doubt yield interesting information and should be carried out, but one cannot study everything at once. Moreover, the type of matching that was studied in Project MATCH was of much greater theoretical interest than the broader sorts of matching contrast just listed and therefore had the potential to advance understanding of what happens in treatment and why clients improve or fail to improve. This is exemplified in the causal path analyses of Project MATCH data being carried out by Longabaugh and colleagues (e.g. Longabaugh *et al.*, 1998).

(7) *The design should have included a control group*. This last criticism stems from the fact

that few matching effects and no clinically significant main treatment effects were found. In the absence of a control group, it is claimed, it is impossible to form any sound conclusions as to the effectiveness of treatment for alcohol problems. Thus a huge amount of time and money was expended with very little return.

Quite apart from any ethical difficulties involved in forming control groups among people seeking treatment, this criticism is very much a matter of being wise after the event. The design of a study can only properly be criticized up to the point at which data collection begins; it cannot be criticized retrospectively in the light of findings or non-findings that eventually emerge. Project MATCH was designed, obviously, to study treatment matching and was carefully and rigorously developed to meet that objective. Given the mood of high optimism in the field during the 1980s about the possible benefits of matching (see Institute of Medicine, 1990), it was important to obtain a clear answer, in an adequately rigorous and large trial, to this question: Can treatment matching (of the kind studied) substantially improve the overall effectiveness of treatment for alcohol problems? We do have a clear answer to this question: No. Or in the more measured words of the Project MATCH investigators themselves: "Despite the promise of earlier matching studies ..., the intuitively appealing notion that matching can appreciably enhance treatment effectiveness has been severely challenged" (Project MATCH Research Group, 1997b, p. 1690). We needed to know that.

NICK HEATHER
Centre for Alcohol and Drug Studies,
Plummer Court,
Carliol Place,
Newcastle upon Tyne NE1 6UR,
UK

INSTITUTE OF MEDICINE (1990) *Broadening the Base of Treatment for Alcohol Problems* (Washington, DC, National Academy Press).

LONGABAUGH, R., WIRTZ, P. W., ZWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, Alcoholics Anonymous and long-term matching effects, *Addiction*, 93, 1313–1333.

MCKAY, J. R. & MCLELLAN, A. T. (1998) Deciding where to start: working with polydrug individuals. In W.R. MILLER & N. HEATHER (Eds.) *Treating Addictive Behaviors* (2nd edition). New York: Plenum.

PROJECT MATCH RESEARCH GROUP (1997a). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol*, 58, 7–29.

PROJECT MATCH RESEARCH GROUP (1997b). Project MATCH secondary a priori hypotheses. *Addiction*, 92, 1655–1682.

PROJECT MATCH RESEARCH GROUP (1998a). Matching alcoholism treatments to client heterogeneity: treatment main effects and matching effects on drinking during treatment. *Journal of Studies on Alcohol*, 59, 631–639.

PROJECT MATCH RESEARCH GROUP (1998b). Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. *Alcoholism: Clinical & Experimental Research*, 22, 1300–1311.

Treatment research in the wake of Project MATCH

D. Colin Drummond

As the authors say, "Project MATCH is the largest, statistically most powerful, psychotherapy trial ever conducted".[1] Indeed, it is difficult to imagine how a study of such Titanic proportions will ever be conducted in the future. Never has an alcoholism controlled trial involved such a large number of subjects, and perhaps not since the Rand Report[2] has a study aroused so much interest from the alcoholism treatment and research community. There is no doubt that a study of this impressive magnitude, quality and rigour needs to be taken very seriously by those in the field, including policy makers, purchasers of health care, clinicians, researchers and those who fund research. There is likely to be considerable debate about the clinical implications of the Project MATCH findings. In research terms the important question now to be addressed is whether all future alcoholism treatment research will be swamped in the wake of Project MATCH, unless it is conducted to the same scale or rigour, or is there a future for equally important research questions to be addressed in more modest projects?

There are many reasons for the research field to welcome Project MATCH. Against a background trend of gradual improvement in methodological research quality[3,4] Project MATCH represents a quantum leap. Recruiting

a large enough sample size to provide sufficient statistical power to assess treatment matching effects on a wide range of variables allows significant doubt to be cast upon promising matches found in earlier studies, many of which were based on *post hoc* analyses. Future matching research will need to pay particular attention to the issue of statistical power.

The Project MATCH Research Group is a paragon of multi-centre research collaboration and much can be learned from the organization of the project. This aspect also highlights the unique value of a national government agency responsible for alcoholism treatment and research in the form of the National Institute of Alcohol Abuse and Alcoholism (NIAAA) which commissioned the study. The UK government would do well to consider this approach, given the enormous economic impact of alcohol problems on society and the missed opportunities for co-ordination of alcoholism treatment and research efforts. Similarly, there have been calls for a European agency with an equivalent remit to NIAAA.[5]

In a host of more specific areas future treatment research in the alcohol field should be informed by the methods of Project MATCH. Tightly specified and monitored manual-based treatment methods allow treatment fidelity to be protected even in a trial of this scale. The use of standardized, validated research instruments to measure outcome, minimizing missing data by assiduous follow-up and assessment of the validity of self-reports by collecting collateral reports and blood specimens for analysis of markers of heavy drinking all reduce potential sources of bias and increase the internal validity of the study. All these factors have previously been identified as important issues in research quality, but have seldom been incorporated into previous alcohol clinical trials.[4]

However, research quality comes at a price. At around $26m[6] this is an expensive project even by US research standards, and it appears very unlikely that another project of this scale would be conducted, at least not for a very long time. Although it is worth remembering that the total cost of alcohol misuse in the United States is estimated to be in the region of some $98.6bn in 1990, including some $10.5bn on funding treatment approaches.[7] From this perspective Project MATCH should be viewed as an appropriate research and development investment by NIAAA. Further, those who aim to conduct treatment research in the future on a more modest budget should take heart in the knowledge that the sample size in Project MATCH was chosen not simply to compare the relative efficacy of three treatment modalities, but rather to test for matching effects between 10 client characteristics and three treatments, involving 16 separate hypotheses. Thus considerably smaller sample sizes would still allow sufficient statistical power to test fewer hypotheses in the one study.

It is also important to note that the level of scientific rigour for which Project MATCH is an exemplar can bring its own problems in terms of external validity or generalizability. For a variety of reasons only 39% (1726) of the total initially screened sample (4481) were randomized. Reasons included failure to complete the 8-hour initial assessment battery, residential instability, legal problems, co-morbid diagnosis, anticipation of concurrent treatment and inability to nominate a "locator". None of these exclusions are unreasonable in order to protect the internal validity of the study or particularly stricter than most alcoholism treatment outcome studies.[4] However, this high level of selection both limits the generalizability of the Project MATCH findings to the broader treatment-seeking population, serves to reduce the heterogeneity of the sample and, hence (as the investigators acknowledge), to work against finding matching effects. Compared with the typical treatment-seeking population, Project MATCH subjects are likely to be more compliant, more highly motivated, have lower levels of comorbidity, and have greater social stability and support. Further, Project MATCH subjects received more extensive assessment, more attractive and possibly better quality controlled treatments and more aggressive follow-up than would be typical for alcoholism treatment in the United States. Future treatment-matching research (and treatment outcome studies in general) would benefit from finding methods of incorporating a wider range of subjects (including those with higher levels of co-morbidity, lower social stability, less motivation for treatment), in more typical treatment environments, without unduly compromising the internal validity of the research. Finding such a balance will always be a compromise, as is the case with sensitivity and specificity of screening instruments, for example, but clinicians who have to make decisions about matching treat-

ments to client characteristics need to be sure that the research evidence is applicable to their treatment population. The methodology to study complex interactions between client and treatment programme variables developed by Moos, Finney & Cronkite[8] has much to commend it to the natural clinical setting.

Another method of studying matching that has so far received little attention in the alcoholism treatment field is that of stepped care. Typically, matching research involves randomization of subjects to two or more treatments and the interaction between subject characteristics and treatment modality are studied on the basis that a proportion of subjects will be "correctly matched" and a proportion will be "mismatched". Such an approach is attractive in research methodology terms in that correct matches ought to have a better outcome than mismatches. However, returning to the issue of validity, in clinical practice "deliberate" mismatching is never on the agenda. In fact, it is difficult to predict exactly what the outcome of any given mismatch ought to be, or why. Clinicians tend to deliver interventions on the basis of presenting need and subsequent response to initial intervention. Stepped care is an alternative approach that has gained currency in the smoking field but is relatively new to the alcohol field.[9–11] Stepped care provides a practical clinical algorithm for the delivery of interventions based upon individual treatment response. Simply, more "difficult cases" who fail to respond to low intensity interventions are offered more complex or intensive (and expensive) interventions. Not only is this a potentially resource efficient way of delivering interventions, it also takes account of the dynamic interaction between client and treatment. Whether stepped care is studied as part of a randomised controlled trial or in its own right, it provides an opportunity to study the interaction between a range of client characteristics and care step utilisation in a way that is more faithful to the natural clinical environment (i.e. a form of naturalistic matching).

It is also important to note that Project MATCH studied only three of a wide potential range of available treatment approaches: clearly even a study of this magnitude could not evaluate every treatment option. The three approaches studied were relatively similar in intensity, duration, setting and method of delivery. There remains an opportunity to study matching effects in a range of different and widely available treatment approaches (e.g. brief versus intensive treatment, individual versus group, residential versus day care).

An important gap in the alcoholism treatment research literature is evidence on cost effectiveness of treatments.[12] The alcohol field is at a relatively early stage in moving from efficacy research to developing cost–effectiveness evaluation methodologies. It is encouraging that Project MATCH included health economic measures for subsidiary analyses. The field needs to progress from studying drinking-related outcomes to incorporate quality of life and disability measures comparable to other areas of healthcare research. Only when the economic impact of alcoholism treatment has been demonstrated will much-needed investment in the treatment system be possible.

Finally, there is a continuing need for the development of new treatment technologies. The fact that the three psychotherapies in Project MATCH fared equally well (or were equally ineffective) is not a reason to abandon the search for effective treatment approaches. Project MATCH did not, for example, include evaluation of emerging new drug treatments for alcoholism which also have potential for matching to client characteristics.[13] Nor did Project MATCH evaluate combinations of treatments that are more commonly delivered in the natural clinical setting. However, the development of effective new treatments needs to be grounded in sound theoretical models and basic research. Theoretical advances need to precede improvements in treatment technologies.

In summary, the alcoholism treatment research field has much to be grateful to the Project MATCH team (and NIAAA) for. The methodological advances developed in Project MATCH should have a positive effect on the overall quality of future treatment research. In terms of the funding of future research, Project MATCH highlights the need to make an adequate investment in research in order adequately to test hypotheses of major clinical importance. Instead of blindly spending considerable resources on approaches of unknown effectiveness, a greater proportion of the overall alcoholism treatment spend should be devoted to research and development.

As a treatment researcher it would be understandable to feel in awe of such a large, expen-

sive and impressively conducted study. In the wake of such a study it would be all too easy to pack up one's research tools and pursue a different career. However, just as the Titanic did not dampen man's desire to conquer the sea or to build more (albeit smaller) ships, so too should the alcoholism research field be spurred on by the advances of Project MATCH in the search for cost–effective treatments that meet individual social and health needs. This piece is intended to provide a pointer to some of the directions in which that search could lead.

D. COLIN DRUMMOND
Department of Psychiatry of Addictive Behaviour, St George's Hospital Medical School, University of London, Cranmer Terrace, London SW17 ORE, UK

1. PROJECT MATCH RESEARCH GROUP (1997) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes, *Journal of Studies on Alcohol*, 58, 7–29
2. ARMOR, D. J., POLICH, J. M. & STAMBUL, H. B. (1978) *Alcoholism and Treatment* (New York, Wiley).
3. SOBELL, M. B., BROCHU, S., SOBELL, L. C., ROY, J. & STEVENS, J. A. (1987) Alcohol treatment outcome evaluation methodology: state of the art 1980–1984, *Addictive Behaviors*, 12, 113–128.
4. MONCRIEFF, J. & DRUMMOND, D. C. (1998) The quality of alcohol treatment research: an examination of influential controlled trials and development of a quality rating system, *Addiction*, 93, 811–823.
5. EDWARDS, G. (1996) Toward a European response to alcohol: getting research onto the map, in: PETERS, T. J. (Ed.) *Alcohol Misuse: a European perspective*, pp. ix–xii (Amsterdam, Harwood).
6. GORDIS, E. (1996) Discussant. Project MATCH: treatment main effects and matching results, *Research Society on Alcoholism and the International Society for Biomedical Research on Alcoholism Joint Scientific Meeting*, 25 June 1996, Washington DC, USA.
7. RICE, D. P. (1993) The economic cost of alcohol abuse and alcohol dependence: 1990, *Alcohol and Health Research World*, 17, 10–11.
8. MOOS, R. H., FINNEY, J. W. & CRONKITE, R. C. (1990) *Alcoholism Treatment: context, process, and outcome*, (New York, Oxford University Press).
9. ORLEANS, T. C. (1993) Treating nicotine dependence in medical settings: a stepped care model, in: ORLEANS, T. C. & SLADE, J. (Eds) *Nicotine Addiction: principles and management*, pp. 145–161 (New York, Oxford University Press).
10. DRUMMOND, D. C. (1997) Alcohol interventions: do the best things come in small packages? *Addiction*, 92, 375–379.
11. BRESLIN, F. C., SOBELL, M. B., SOBELL, L. C., BUCHAN, G. & CUNNINGHAM, J. A. (1997) Toward a stepped care approach to treating problem drinkers: the predictive utility of within treatment variables and therapist prognostic ratings, *Addiction*, 92, 1479–1489.
12. HOLDER, H. D., LONGABAUGH, R., MILLER, W. R. & RUBONIS, A. V. (1991) The cost effectiveness of treatment for alcohol problems: a first approximation, *Journal of Studies on Alcohol*, 52, 517–540.
13. MONCRIEFF, J. & DRUMMOND, D. C. (1997) New drug treatments for alcohol problems: a critical appraisal, *Addiction*, 92, 939–947.

Some treatment implications of Project MATCH

John W. Finney

Project MATCH (Project MATCH Research Group (PMRG), 1997a), the largest, most rigorous alcohol treatment efficacy trial ever conducted, examined the main and interactive effects of Twelve-Step facilitation (TSF), cognitive–behavioral (CBT) and motivational enhancement (MET) therapy. An efficacy trial maximizes methodological rigor. Closely monitored therapists present a standard dose of specified treatment to carefully selected patients (with minimal co-morbidities) who have the target disorder, and who agree to random assignment to one of the available treatment conditions. As such, one can have considerable confidence that findings on the main and interactive effects of treatment actually reflect the (differential) impact of treatment conditions. The drawback is that it is often difficult to estimate how well the findings from an efficacy trial will generalize to "real-world" clinical situations (PMRG, 1997b).

Fortunately, findings have been published recently from a large-scale "effectiveness" evaluation of Department of Veterans Affairs (VA) Twelve-Step and cognitive–behavioral substance abuse treatment programs in the United States (Ouimette *et al.*, 1997). The evaluation focused on the outcomes of treatment delivered to patients under normal conditions of care. This commentary considers the implications of Project MATCH findings regarding the main effects for type of treatment and patient–treatment interaction (matching) effects, in the context of findings from the VA effectiveness evaluation.

An important finding in Project MATCH was that the patients exposed to TSF, CBT and MET generally had similar outcomes, although TSF patients had more abstinent days than the other two aftercare groups toward the end of the first follow-up year, more abstinent days and fewer drinks per drinking days than CBT outpatients, and were more likely to be abstinent in the 3 months prior to the 1-year follow-up than the other two outpatient treatment groups. Prior reviews of the alcohol treatment research literature (e.g. Miller *et al.*, 1995; Finney & Monahan, 1996) have pointed to considerable empirical evidence supporting the efficacy of cognitive–behavioral treatment approaches versus little controlled research for a Twelve-Step-based self-help group—Alcoholics Anonymous (AA). The findings from Project MATCH are important in that they show that a treatment based on Twelve-Step principles can perform as well as, or better than, a cognitive–behavioral intervention.

At the same time, it is important to note (as has been done by the PMRG (1997a)) that TSF in Project MATCH was not AA or simply referral to AA, and not the same as "real-world" Twelve-Step treatment programs. TSF was offered by professional therapists in individual (one-on-one) treatment sessions. Nevertheless, the results of the multi-site effectiveness evaluation of VA substance abuse treatment programs were similar to those of Project MATCH—few outcome differences at a 13-month follow-up, but a greater likelihood for patients exposed to Twelve-Step treatment to be abstinent compared with those provided with cognitive–behavioral treatment (Ouimette *et al.*, 1997).

Another provocative finding from Project MATCH was that MET, with four planned sessions, was as efficacious as CBT and TSF, with 12 planned sessions, even though most Project MATCH patients were diagnosed with alcohol dependence (as opposed to milder "drinking problems"). This result might lead some to conclude that four sessions of MET would yield the same results in standard treatment situations as were found in Project MATCH, and the same results as would 12 sessions of either TSF or CBT. If so, the treatment of choice would be MET, due to its lower cost. A more general conclusion could be that brief interventions are as effective as more intensive interventions.

For several reasons, I believe the PMRG (1997a) is correct in cautioning against drawing these conclusions. First, the relative treatment intensities of MET versus CBT and TSF in Project MATCH were not as great as the 1:3 difference in the planned number of sessions would suggest. In the outpatient and aftercare arms, the average number of MET sessions actually attended was 3.3 and 3.1; the corresponding figures were 8.3 and 8.0 sessions in CBT, and 7.5 and 7.3 sessions in TSF (Carroll *et al.*, 1998).

Moreover, if one combines the 8 hours of assessment that preceded the treatment phase, and the multiple follow-up contacts (that presumably had some therapeutic impact—Breslin *et al.*, 1996) with patients every 3 months in the year after treatment, one has to conclude that the MET condition constituted more therapy/contact hours than is normally conveyed by the term "brief". In addition, the heavy assessment across the three treatment conditions further reduces their overall variation in "treatment" intensity.

A provocative idea raised by Project MATCH's findings on the main effects of treatment is that what Brekke *et al.* (1997) referred to as the "longitudinality" of treatment (its dispersion over time), rather than its intensity, may be a critical treatment dimension. In this regard, Project MATCH clients attended MET over an average period of 8.4 weeks, CBT over 9.3 weeks, and TSF over 8.3 weeks (PMRG, 1997a). Also, the somewhat stronger effects of TSF in Project MATCH may have been due to the "treatment extension" afforded by patients attending more Twelve-Step meetings in the year after treatment ended relative to the number attended by patients in the other two treatment conditions (Longabaugh *et al.*, 1998; PMRG, 1998). In the VA substance abuse treatment evaluation, findings supported the idea that post-treatment Twelve-Step group attendance was a mediator of the effect of Twelve-Step (versus cognitive–behavioral) treatment on abstinence (Humphreys *et al.*, 1999).

Project MATCH may disillusion some researchers and clinicians with regard to the general idea of patient–treatment-matching, given the range of matching hypotheses examined in the primary and secondary analyses (PMRG, 1997a; 1997b) relative to the small number of (expected) significant results. To some unknown extent, however, the ratio of tested to supported hypotheses is misleading. The number of matching hypotheses that had the potential to be sup-

ported appears to have been somewhat less than the number tested.

Although the mathematical formulations of the hypotheses tested in Project MATCH specified nothing more than slope differences, published verbal descriptions of some of the hypotheses provide additional information. For example, the PMRG (1997a) wrote that "clients who had greater psychiatric severity were expected to have better outcomes in CBT compared to those in either TSF or MET, since CBT taught skills for coping with social and emotional cues to drink" (p. 8). Similarly, "lower levels of readiness to change were predicted to be associated with better outcomes for clients in MET, a motivation enhancement intervention, when contrasted with clients in CBT, a skills-based intervention" (PMRG, 1997a, p. 8).

Because expected "cross-over" effects or "disordinal" interactions were specified verbally for some other hypotheses (PMRG, 1997b), I assume that for the two hypotheses above the expectation was that treatment effects would not differ for patients with less psychiatric severity in the first instance and for patients with greater readiness to change in the second instance (i.e. "ordinal" interactions were anticipated—for an explicit example, see Longabaugh *et al.*, 1998). Thus, the first matching hypothesis implies a main effect for type of treatment favoring CBT over MET, whereas the second matching hypothesis implies a main effect for treatment type favoring MET over CBT. It is not possible for both main effects to emerge; thus, both matching hypotheses could not have been supported. We realized that there was a similar inconsistency between our hypotheses of no main effect of type of treatment in our VA evaluation and our hypotheses about specific interaction effects (Ouimette *et al.*, 1999).

Even with this perspective of a reduced number of matches that could have emerged (the number is unknown because verbal descriptions of all of the matching hypotheses have been not presented in the Project MATCH reports published to date), Project MATCH provides limited support for matching patients to the three treatment conditions based on the set of matching variables examined. In this regard, its findings are quite consistent with those from the multi-site evaluation of VA Twelve-Step and cognitive–behavioral treatments, where none of six similar matching hypotheses was supported (Ouimette *et al.*, 1999), as well as with extensive bodies of research in education (Cronbach & Snow, 1977) and psychotherapy (Dance & Neufeld, 1988). Certainly, other treatment variables (e.g. therapist characteristics, setting of treatment, etc.) should be examined for interactions with patient characteristics (PMRG, 1997a), but interactions are difficult to detect and replicate.

Nevertheless, I think most researchers and alcohol treatment providers believe that patient and treatment factors interact in producing outcomes. The belief in the existence of such interactions is acknowledged in assertions that the treatment of, and recovery from, alcohol use disorders are complex processes. Various patient, treatment and life context factors come into play (and probably interact). The complexity of the matching process is discussed by the PMRG (1997b), which suggests that patient–treatment interactions are likely to be of a higher order than the simple, single patient variable × treatment modality interactions examined thus far in Project MATCH and most other studies. Whether those anticipated interactions are of such a complex form that they cannot be addressed by current research technology remains to be determined. To the extent that they are, matching individual patients to treatment will remain more of a clinical art than a science.

JOHN W. FINNEY
Program Evaluation and Resource Center,
HSR&D Center for Health Care Evaluation,
VA Palo Alto Health Care System and Stanford University Medical Center,
Menlo Park Division (152), 795 Willow Road,
Menlo Park,
Palo Alto,
California 94025,
USA

BREKKE, J. S., LONG, J. D., NESBITT, N. & SOBEL, E. (1997) The impact of service characteristics on functional outcomes from community support programs for persons with schizophrenia: a growth curve analysis, *Journal of Consulting and Clinical Psychology*, 65, 464–475.

BRESLIN, C., SOBELL, L. C., SOBELL, M. B., BUCHAN, G. & KWAN, E. (1996) Aftercare telephone contacts with problem drinkers can serve a clinical and research function, *Addiction*, 91, 1359–13364.

CARROLL, K. M., CONNORS, G. J. & COONEY, N. L. *et al.* (1998) Internal validity of Project MATCH treatments: discriminability and integrity, *Journal of Consulting and Clinical Psychology*, 66, 290–303.

CRONBACH, L. & SNOW, R. L. (1977) *Aptitudes and Instructional Methods: a handbook for research on interactions* (New York, Irvington).

DANCE, K. A. & NEUFELD, R. W. J. (1988) Aptitude–treatment interaction research in the clinical setting: a review of attempts to dispel the "patient uniformity" myth, *Psychological Bulletin*, 194, 192–213.

EDWARDS, G. (1989) As the years go rolling by: drinking problems in the time dimension, *British Journal of Psychiatry*, 154, 18–26.

FINNEY, J. W. & MONAHAN, S. C. (1996) The cost effectiveness of treatment for alcoholism: a second approximation, *Journal of Studies on Alcohol*, 57, 229–243.

HUMPHREYS, K., DEARMIN-HUEBSCH, P. L., FINNEY, J. W. & MOOS, R. H. (1999) A comparative evaluation of substance abuse treatment. IV. Substance abuse treatment can enhance the effectiveness of self-help groups, submitted to *Alcoholism: Clinical and Experimental Research*.

LONGABAUGH, R., WIRTZ, P. W., SWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, Alcoholics Anonymous, and long-term matching effects, *Addiction*, 93, 1313–1333.

MILLER, W. R., BROWN, J. M. & SIMPSON, T. L. *et al.* (1995) What works? A methodological analysis of the alcohol treatment outcome literature, in: HESTER, R. K. & MILLER, W. R. (Eds) *Handbook of Alcoholism Treatment Approaches: effective alternatives*, 2nd edn, pp. 12–44 (Boston, MA, Allyn and Bacon).

OUIMETTE, P. C., FINNEY, J. W. & MOOS, R. H. (1997) Twelve step and cognitive–behavioral treatment for substance abuse: a comparison of treatment effectiveness, *Journal of Consulting and Clinical Psychology*, 65, 230–240.

OUIMETTE, P. C., FINNEY, J. W., GIMA, K. & MOOS, R. H. (1999) A comparative evaluation of substance abuse treatment: III. Examining mechanisms underlying patient–treatment matching hypotheses, submitted to *Alcoholism: Clinical and Experimental Research*.

PROJECT MATCH RESEARCH GROUP (1997a) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes, *Journal of Studies on Alcohol*, 58, 7–29.

PROJECT MATCH RESEARCH GROUP (1997b) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.

PROJECT MATCH RESEARCH GROUP (1998) Matching alcoholism treatment to client heterogeneity: Project MATCH three-year drinking outcomes, *Alcoholism: Clinical and Experimental Research*, 22, 1300–1311.

Life is short, the Art long

Lars Lindström

"Life is short, the Art long, opportunity fleeting, experience treacherous, judgement difficult. The physician must be ready, not only to do his duty himself, but also to secure the co-operation of the patient, of the attendants and of externals." These notes, jotted down 2400 years ago by Hippocrates (1923), occur to me as an eloquent summary of the findings by Project MATCH and their treatment implications.

First, the Research Group repudiates simplistic approaches. Their a priori matching hypotheses, trying to identify first-order matching effects involving three individual treatment modalities and 21 single client attributes, did not explain the complexity of the findings. The number of successful matches are small, the magnitudes of the matching effects are modest, and those effects that were supported lack consistency across study arms (outpatient treatment vs. aftercare), dependent variables and periods of observation.

Should these findings strike us as a surprise or even as a disaster, as implied by those who compare the fate of Project MATCH with that of *Titanic*? Certainly not! Considering the complexity of the task, previous researchers have emphasized that no dramatic breakthrough should be expected as a consequence of the matching of clients to treatments (Finney & Moos, 1986; Lindström, 1992). The possibility of generalizing findings is always limited by interactions which—like reflections in a "hall of mirrors" (Cronbach, 1975)—enter into systems of higher order.

The contributions of the social sciences may largely be of an indirect nature. Instead of hoping for enduring systematic theories about man in society (so-called "grand narratives"), Cronbach (1975) identified two reasonable aspirations. One is to "assess local events accurately, to improve short-run control"; the other one is to "develop explanatory concepts, concepts that will help people to use their heads". Even if matching effects may not be as generalizable across treatment sites or settings (e.g. outpatient treatment vs. aftercare) that we would like them to be, Project MATCH and similar studies may inspire and guide site-specific efforts to achieve an efficient short-term utilization of treatment resources. Moreover, the fertility of the matching hypothesis cannot be measured solely in terms of the number of significant matching effects observed. Its value will rather be judged by whether it is able to challenge and refine the ways in which we are thinking about alcoholism treat-

ment. In these respects at least, Project MATCH should be regarded as a success.

Secondly, Project MATCH will probably encourage a fresh look at treatment commonalities. Shared active ingredients are discussed as one plausible explanation of the results. The Hippocratic writers realized the need to "secure the co-operation of the patient". The Project MATCH Research Group notes that the therapist perception of the "working alliance" is a predictor of change across treatment conditions. Previous research (reviewed by Lindström, 1992) suggests that three conditions have to be fulfilled in order for a helpful alliance to be established:

- *The client must be given the opportunity to experience success.* This experience may activate self-healing processes, eventually enabling the client to master his entire life situation more effectively than before; but the process will be hampered if the therapeutic task is construed as being irrelevant or too threatening. The Project MATCH finding that "angrier" outpatients fared better when treated in the non-confrontational motivational enhancement therapy (MET) is consistent with this hypothetical mechanism of change.
- *Barriers to a favourable interpersonal relationship must be eliminated.* Previous matching studies have found effects of an interaction involving clients' cognitive styles and the degree of structure and directiveness provided by their therapists. These results indicate that effective communication is a prerequisite for an emotionally satisfying relationship, and that it is too easy to talk over people's heads, or else to talk down to them. By contrast, Project MATCH did not find any matching effects related to either "cognitive impairment" or "conceptual level". This result is unexpected, given that one programme (MET) seems to have been less structured and directive than the other two. On the other hand, the Research Group speculates that the participant team of highly trained therapists may have tailored their approach to fit those clients that might be poorly suited to the therapy. Adaption to cognitive style may have been facilitated by the fact that varying the degree of structure and directiveness does not seem to have been a major issue in Project MATCH.
- *A secure base must be provided in order for a therapeutic process to take place.* It seems that a small group of persons need to spend some time in inpatient care in order to be able to concentrate their energies on treatment. The significance of the treatment setting was not investigated by Project MATCH, but has been discussed in a research review by Finney *et al.* (1996).

Thirdly, the Hippocratic Corpus emphasizes the importance of securing the co-operation of "the attendants and of externals". In fact, the most intriguing finding by the Project MATCH Research Group is the late appearance of a matching effect indicating that clients with a social network that supports drinking fared better in the twelve-step facilitation therapy (TSF) than in other treatments. The effect is partially attributable to an increased involvement in AA. Longabaugh *et al.* (1998) envisages the "exciting possibility that therapeutic ingredients (e.g. AA involvement) that interact with contextual variables (e.g. social support) can have effects that increase in magnitude over time, rather than diminish as is so often the case when treatment focuses mainly on the individual". This implied shift towards a long-term perspective with a focus on the social contexts of drinking and recovery is long overdue. If Project MATCH contributes to such a re-orientation of treatment research and practice, it has certainly been a good investment.

LARS LINDSTRÖM
Visual Arts/Drama/Athletics/Music,
Stockholm Institute of Education,
Box 34103,
Stockholm S-100 26,
Sweden

CRONBACH, L. (1975) Beyond the two disciplines of scientific psychology, *American Psychologist*, 30, 116–127.

FINNEY, J. W., HAHN, A. C. & MOOS, R. H. (1996) The effectiveness of inpatient and outpatient treatment for alcohol abuse: the need to focus on mediators and moderators of setting effects, *Addiction*, 91, 1773–1796.

FINNEY, J. W. & MOOS, R. H. (1986) Matching patients with treatments: conceptual and methodological issues, *Journal of Studies on Alcohol*, 47, 122–134.

HIPPOCRATES (1923) *Hippocrates*, vol. 2 (Cambridge, MA, Harvard University Press).

LINDSTRÖM, L. (1992) *Managing Alcoholism: matching clients to treatments* (Oxford, Oxford University Press).

LONGBAUGH, R., WIRTZ, P. W., ZWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, Alcoholics Anonymous and long-term matching effects, *Addiction*, 93, 1313–1333.

Project MATCH and the stages of change

Stephen Sutton

In this commentary I briefly consider the implications of selected Project MATCH findings for the transtheoretical model, and then the implications of the transtheoretical model for future treatment matching studies.

The transtheoretical model (TTM), popularly known as the "stages of change" model, assumes that behaviour change involves movement through a sequence of discrete, qualitatively distinct, stages: precontemplation, contemplation, preparation, action, maintenance and termination (Prochaska, DiClemente & Norcross, 1992; Prochaska & Velicer, 1997). Although Project MATCH was not based on the TTM, it did include as matching variables two measures of motivational readiness to change derived from the URICA and the SOCRATES, respectively. Although these are both multi-dimensional scales, in the analysis of the Project MATCH data a single score was derived for each individual from each of the two questionnaires. The score based on the URICA was treated as a primary a priori matching variable. It was predicted that clients with relatively low readiness to change scores would do better if they received motivational enhancement therapy (MET) than if they received cognitive—behavioural coping skills therapy (CBT). (More formally, it was hypothesized that the slope of the regression line for drinking outcome on readiness score would be greater in the CBT condition than in the MET condition; Project MATCH Research Group, 1997a, Table 1, p. 9.) A similar prediction was made for the alcohol-specific readiness to change score derived from the SOCRATES, which was treated as a secondary a priori matching variable (Project MATCH Research Group, 1997b). Neither hypothesis received consistent support, although both motivation scores were found to be predictive of drinking outcomes.

These findings have little bearing on the validity or otherwise of the TTM. Treating readiness to change as a continuous dimension is not consistent with the model's assumption that the stages of change are discrete and qualitatively distinct (Sutton, 1996). Methods of measuring stage of change that classify clients into a set of mutually exclusive and exhaustive categories—"staging algorithms"—are more in keeping with stage model assumptions, but such a categorical measure was apparently not included in the Project MATCH pre-treatment assessment battery.

The TTM, in fact, implies a very different kind of matching study. In particular, it implies that a different treatment should be developed for each stage in the model: one treatment for pre-contemplators, to move them to the contemplation stage; another for contemplators, to move them to the preparation stage; and so on. The strongest research design for testing the effects of such treatments would involve assigning clients randomly in a given stage to receiving either the treatment intended for people in that stage (the stage-matched treatment) or a treatment designed for people in a different stage (stage-mismatched treatment). The TTM would predict stage by treatment interaction effects. Contemplators, for example, should do better—in terms of the proportion who move to the next stage—if they receive the treatment designed for people in their stage than if they receive a treatment tailored to a different stage. Demonstrating in randomized studies that stage-matched treatments are consistently more effective than stage-mismatched treatments would constitute strong evidence for the TTM—much stronger than the evidence collected to date. An extension of the strategy of comparing stage-matched and stage-mismatched treatments is to compare different sequences of treatments. The TTM would predict that treatments delivered in the "correct" sequence (i.e. corresponding to the stage sequence postulated by the model) would be more effective than treatments delivered in any other order (Weinstein, Rothman & Sutton, 1998).

There are two reasons why it would be premature to embark on such studies at the present time. First it should be remembered that, in spite of the wide currency of stage ideas in the addictions field, the vast majority of empirical applications of the TTM have been to smoking; to date, few studies have applied the model to drinking. The second, and arguably more serious, reason is a theoretical one. The TTM is underspecified. Expositions of the model do not clearly specify,

or hypothesize, what factors are important in influencing each stage transition. The model includes a number of constructs that are potential candidates for such causal factors: the 10 processes of change, the pros and cons, and confidence and temptation; but the model does not specify the causal relationships between these constructs and the stages of change. A fully specified stage model would make predictions of the form "Factor A is important in influencing the transition from stage I to stage II whereas factor B is important in influencing the transition from stage II to stage III." Such a prediction could be tested in prospective studies by examining the extent to which factor A predicts movement to stage II among people in stage I and factor B predicts movement to stage III among people in stage II. Stronger evidence would come from experimental demonstrations (i) that modifying A increased the likelihood of the first transition while modifying B increased the likelihood of the second transition and (ii) that receiving the treatments in the sequence specified by model was more effective than receiving the treatments in a different sequence.

Such studies of stage-matched and stage-mismatched treatments delivered in different orders would be difficult and costly to conduct. Nevertheless, if based on a properly specified stage model, this approach offers one possible way of keeping alive the intuitively plausible notion that treatments should be matched to client characteristics—a notion that has, in the MATCH researchers' own words, been severely challenged by the Project MATCH findings.

STEPHEN SUTTON
Health Behaviour Unit,
University College London,
Brook House,
2–16 Torrington Place,
London WC1E 6BT,
UK

PROCHASKA, J. O., DICLEMENTE, C. C. & NORCROSS, J. C. (1992) In search of how people change: applications to addictive behaviors, *American Psychologist*, 47, 1102–1114.

PROCHASKA, J. O. & VELICER, W. F. (1997) The transtheoretical model of health behavior change, *American Journal of Health Promotion*, 12, 38–48.

PROJECT MATCH RESEARCH GROUP (1997a) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes, *Journal of Studies on Alcohol*, 58, 7–29.

PROJECT MATCH RESEARCH GROUP (1997b) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.

SUTTON, S. R. (1996) Can 'stages of change' provide guidance in the treatment of addictions? A critical examination of Prochaska and DiClemente's model, in: EDWARDS, G. & DARE, C. (Eds) *Psychotherapy, Psychological Treatments and the Addictions*, pp. 189–205 (Cambridge, Cambridge University Press).

WEINSTEIN, N. D., ROTHMAN, A. J. & SUTTON, S. R. (1998) Stage theories of health behavior: conceptual and methodological issues, *Health Psychology*, 17, 290–299.

Efficacy of outpatient alcoholism treatment

M. Soyka

Project MATCH is an impressive and wonderfully designed treatment study to test specific matching effects and also the efficacy of different treatment approaches in a total of 1726 individuals who were randomly assigned to three treatments which were tested in parallel studies in two types of settings: 952 patients were treated in an outpatient setting, 774 patients in an aftercare setting immediately following inpatient or intensive day hospital treatment.

Many clinical questions can be adressed after reviewing the published and unpublished data of this study. My comment will focus basically on the abstinence rates achieved in both treatment arms. The question which outcome variables should be chosen (total abstinence vs. reduction of heavy drinking/drinks per day) and how to measure treatment success is discussed controversy in the literature. In the survival analysis of Project MATCH both alternatives are given: time to first drink and time to first heavy-drinking period defined as 3 consecutive days of heavy drinking (>6 resp. 4 drinks/day). Since participants in all of the therapies had been informed that abstinence was the treatment goal the first evaluation strategy seems to match the treatment goals the best. Relapse estimates of alcohol consumption were obtained by Form 90 (baseline and follow-up) and the time-line follow-back methodology, among others. Liver enzymes and CDT (assessed only in the 15-month examination) were also measured and the authors reported that collateral and biochemical measures indicate a high degree of confidence in the verbal report data obtained. Taken together the data on abstinence rates can be considered as reliable.

With regard to treatment procedure, what I had learned from oral presentations about Project Match and what is not consistently reported in the publications available so far—to say the least—is that patients in Project Match have been paid for research interviews. This may influence compliance to treatment and abstinence rates.

In the 1-year follow-up in the aftercare arm 35% of subjects reported continued complete abstinence throughout the 12 follow-up months compared to 19% in the outpatient treatment group. In the aftercare arm 40% of patients returned to heavy drinking compared to 46% in the outpatient treatment group. In the 3-year follow-up a number of secondary outcome measures were also evaluated but will not be discussed here. In the 3-year follow-up for the outpatient treatment group ($N = 806$) 29% of patients reported complete abstinence during months 37–39.

In the outpatient group only patients low in psychiatric severity had more abstinent days after Twelve-Step facilitation than after cognitive behavioral therapy. For patients with higher levels of psychiatric severity none of the treatment approaches were found to be superior. With regard to the overall outcomes there were few differences among the three treatments although Twelve-Step-Facilitation showed a slight advantage.

My point is: are these favourable clinical results? While the majority of alcohol treatment programmes, at least in Germany, are still inpatient programmes in recent years a certain shift from inpatient to outpatient treatment can be noticed (Knowles, 1983). A number of treatment programmes have been advocated but, different from inpatient treatment programmes of alcoholism, few catamnestic studies have been conducted concerning the efficacy of outpatient rehabilitation in alcoholism. Therefore Project MATCH makes a significant contribution to our understanding of the efficacy of outpatient alcohol treatment. Current opinion on the efficacy of outpatient treatment compared to inpatient treatment is mixed (Collins, 1997a,b; Cole *et al.*, 1981; Öjehagen *et al.*, 1987; Pettinati *et al.*, 1993). While many clinicians in Europe feel that the abstinence rates in alcoholics who participated in inpatient treatment of alcoholism are superior to outpatient treatment other, predominantly US, researchers do not unanimously share this opinion (Schuckit, 1992). Finney *et al.*, (1996) reviewed 14 studies and concluded that in seven studies no differences between inpatient and outpatient treatment could be found, in five studies inpatient treatment resulted in better outcome, in two studies outpatient treatment did.

Other studies have shown more favourable results for outpatient rehabilitation programmes compared to Project MATCH. In a recent 18–24-month follow-up study 65 patients who took part in an intensive 8-month outpatient treatment programme were examined (Soyka *et al.*, 1997). During the treatment phase patients were seen on a regular basis and participated in various individual and group therapies (7 hours/week on average). Fifty-one of the 65 patients who had participated in the programme could subsequently be personally interviewed (seven patients refused to take part, six could not be reached, one had died). Forty of the 51 patients had completed the outpatient treatment. Assuming that all patients who could not be interviewed, or refused, were relapsers, the abstinence rate was found to be 48%. Although the patients included in the programme might be considered to be socially more stable compared to other alcoholics these data indicate that better abstinence rates in outpatient treatment can be achieved by a more intensive treatment setting. We will examine this treatment programme further in future years.

Compared to the treatment setting described above the intensity of theareputic interventions was much lower in Project MATCH (four–12 treatment sessions). The abstinence rates found in the 1-year follow-up for the outpatient treatment are similar to those found in pharmaceutical trials for the placebo group (Sass *et al.*, 1996). "Placebo" in this respect does not mean that patients did not receive any kind of treatment, but were usually seen on a regular outpatient basis. The placebo problem in pharmaceutical trials in alcoholism has been adressed by Moncrieff & Drummond (1997), among others.

The good news is: outpatient treatment of alcoholism is effective. The bad news is: treatment results in the aftercare group were shown to be much better, at least with respect to continous abstinence. Twelve sessions of any kind of psychotherapy may not be sufficient for patients with a more severe alcohol problem.

The important question is not only which treatment approach might be the most effective and which client should be transferred to which programme, but also which client should be treated as an inpatient first and which patient should be predominantly treated as an outpatient. This question cannot be answered yet, but Project MATCH gives us some baseline data on what we can expect from twelve sessions of psychotherapy in alcoholics: apparently not too much.

M. SOYKA
Psychiatric Hospital University of Munich,
Nussbaumstr. 7,
80336 Munich,
Germany

COLE, S. G., LEHMAN, W. E., COLE, W. E. & JONES, A. (1981) Inpatient vs. outpatient treatment of alcohol and drug abusers, *American Journal of Drug and Alcohol Abuse*, 8, 329–345.

COLLINS, G. B. (1997a) Inpatient and outpatient treatment of alcoholism—essential features, in: MILLER, N. S. (Ed.) *The Principles and Practice of Addictions in Psychiatry*, pp. 319–330 (Philadelphia, WB Saunders).

COLLINS, G. B. (1997b) Emerging concepts of alcoholism treatment: challenges and controversies, in: MILLER, N. S. (Ed.) *The Principles and Practice of Addictions in Psychiatry*, pp. 331–342 (Philadelphia, WB Saunders).

FINNEY, J. W., HAHN, A. C. & MOOS, R. H. (1996) The effectiveness of inpatient and outpatient treatment for alcohol abuse: the need to focus on mediators and moderators of setting effects, *Addiction*, 91, 1773–1796.

KNOWLES, P. L. (1983) Inpatient versus outpatient treatment of substance misuse in hospitals, 1975–1980, *Journal of Studies on Alcohol*, 44, 384–387.

MAMMO, A. & WEINBAUM, D. F. (1993) Some factors that influence dropping out from outpatient alcoholism treatment facilities, *Journal of Studies on Alcohol*, 54, 92–101.

MONCRIEFF, I. & DRUMMOND, D. C. (1997) New drug treatments for alcohol problems. a critical appraisal, *Addiction*, 91, 787–805.

ÖJEHAGEN, A., SKJAERRIS, A. & BERGLUND, M. (1987) Prediction of posttreatment drinking outcome in a 2-year out-patient alcoholic treatment program. A follow-up study, *Alcoholism: Clinical and Experimental Research*, 14, 1–9.

PETTINATI, H. M., MEYERS, K., JENSEN, J. M., KAPLAN, F. & EVANS, B. D. (1993) Inpatient vs outpatient treatment for substance dependence revisited, *Psychiatry Quarterly*, 64, 173–182.

SCHUCKIT, M. A. (1992) Treatment of alcoholism in office and outpatient settings, in: MENDELSON, J. H. & MELLO, N. K. (Eds) *Medical Diagnosis and Treatment of Alcoholism*, pp. 363–392 (New York, McGraw Hill).

SASS, H., SOYKA, M., MANN, K. & ZIEGLGAENSBERGER, W. (1996) Relapse prevention by acamprosate: results from a placebo-controlled study on alcohol dependence, *Archives of General Psychiatry*, 53, 673–680.

SOYKA, M., KIRCHMAYER, C., KOTTER, G., JOHN, C., LOEHNERT, E. & MOELLER, H.-J. (1997) New possibilities of treatment and rehabilitation of alcohol-dependent patients—catamnestic study on the efficacy of outpatient treatment programmes demonstrated by a model procedure [in German], *Fortschritte der Neurologie und Psychiatrie*, 65, 407–412.

Should a few hours of therapy change behaviour 3 years later?

Tim Stockwell

If someone had told me 10 years ago that an alcoholism treatment trial would one day be conducted with a sample size of 1726, with random allocation to the three treatment groups and with follow-up rates of around 90% for up to 3 years I would probably, when I stopped laughing, have mused with fascination about what the field might learn from such a priceless opportunity. Astonishingly, it really has happened and, when the funding body and the many co-investigators have been deservedly congratulated, what indeed have we learned so far from this monumental study?

To some the most surprising result is that there were so few results. Put in the most depressingly negative way, not one of 21 empirically derived predictions about the types of client who would do best with which treatment was consistently supported across both arms of the study—outpatient and aftercare. Only a handful of predictions were supported in one or other arm and usually at only one of the three main assessment periods for outcome data. Most of these small but significant effects suggest an advantage for Twelve-Step facilitation therapy, especially for clients who are more severely alcohol dependent, are more likely to show signs of Antisocial Personality Disorder as well as poor social functioning and have social networks which support continued drinking. Attendance at AA meetings during the 3 years of the study was shown to partially mediate this effect.

What does this failure to find many significant matches between client characteristics and types of treatment mean? This is discussed extensively

in the various excellent reports of the study and I will only offer here a few different perspectives and consider a few heretical explanations.

(i) Treatment for severe alcohol dependence does not work

There are echoes in this null result from other, earlier major studies of alcohol treatment. The systematic reviews of Emrick (1975) and Costello (1975), for example, and the randomized controlled trial of a single session of 'advice' versus intensive inpatient and outpatient treatment by Edwards and colleagues (1977) are all remarkable for failing to identify significant main treatment effects. If treatment for alcohol dependence has very little chance of having a significant impact it is unlikely that there will be major differences between different varieties of treatment or interactions of these with different client characteristics.

These earlier studies were influential in creating a mind-set which explained null results from treatment studies in terms of subtle matching effects otherwise lost in the search for an overall effect of treatment across all types of clients. Indeed, the Costello (1975) and Edwards *et al.* (1977) studies both had things to say about intensity of treatment and problem severity as critical variables to measure. In many ways, these early studies created the momentum for Project MATCH and now, it would seem, one possible heretical interpretation of the results is that, apart from a small advantage to people encouraged to join a self-help organization, there is little of enduring value from intensive treatment for alcohol dependence.

The confident assertion of the Project MATCH group regarding the 3-year outcomes that the overall outcomes were so good that something must have benefited some clients in achieving some abstinence from alcohol is an interesting statement of faith in the power of 'therapy' to influence behaviour several years down the track. We now know too much about selection processes and factors totally outside the realm of treatment as influences on long-term drinking outcomes to accept such a view uncritically. However, I suggest there are still other explanations to explore before we submit to the sad conclusion that alcohol treatment has little or no effectiveness.

(ii) The most powerful intervention was the research

In many ways it is more plausible to suggest that with a chronically relapsing condition like alcohol dependence a few therapeutic sessions over 12 weeks will have less long-term impact on drinking than a series of follow-up interviews strategically placed over a 3-year period. Many of the ingredients of what is believed to be an effective motivational intervention are contained in such a series of research interviews: e.g. a non-judgemental focus on recent drinking behaviour and related harms and the expectation of this being repeated over an extended time period. Some time ago I published a study which found that attendance at a prior research interview was more likely to result in outpatient appointments at an alcoholism treatment unit being kept (Sutherland *et al.*, 1985). I also recall an evaluation of the Accept Day Centre in London which asked clients what aspects of the treatment programme they most benefited from and found that the most appreciated single intervention was the follow-up visits—which were actually research interviews (Potamianos & Papadatos, 1987). The Project MATCH team notes that there was slightly more contact time (5 hours) spent on follow-up assessments over the 3 years of the study than there was in one of the treatments, 'Motivational Enhancement Therapy' (Project MATCH Research Group, 1997). They also acknowledge that if this has a therapeutic benefit then it greatly reduces the possibility of finding matching effects—simply because all treatment groups received identical amounts of follow-up assessment.

(iii) The interpersonal dynamics of the treatment sessions were more important than the type of manual being followed

There have been some reports in the literature of good outcomes being best predicted by the quality of the client–therapist interaction. Perhaps we need to be better able to specify the styles of communication of clients and therapists so as to both 'pick' good therapists and also match them to clients with a complementary communication style. Indeed, there is a suggestion that this may have occurred in relation to Project MATCH with 'angry' clients faring better with non-confrontational MET than either of the other two treatments.

(iv) Conclusions and suggestions for future research
In light of the above considerations, I am inclined to believe that it was remarkable that any interaction effects were found at all in this study. Special attention should be given to the few findings that were significant because of the factors mitigating against any differences between these treatments being detected. Indeed, the finding that Twelve-Step facilitation has some extra benefits fits within an overall view that a severe and long-term relapsing condition you need to be able to provide an intense and long-term support system. Consistent with this was the temporary result that during the 12-week course of treatment the relatively brief MET modality resulted in less abstinence than the other more intensive treatments, i.e. the longer treatment course may have maintained more abstinence during this more limited period.

In passing I would note that these findings mainly apply to a traditional treatment goal of total abstinence and that the outcome measures have not been used to examine the variables which predict the possibility of moderate drinking with problem-free outcomes. Doubtless, future reports from MATCH will re-examine this old chestnut regarding the types of clients who are able to achieve such 'controlled drinking' outcomes.

Will there ever be another Project MATCH? I hope so, and I hope that both the treatments and evaluation designs include a consideration of the enduring nature of severe alcohol problems and the myriad of factors outside of the experience of treatment which continue, even after the end of a 12-week course of treatment. There will need to be a control for the potentially therapeutic effects of the research interviews themselves, perhaps by varying the length, frequency and format of these. I hope also that there would be an attempt to specify and measure personal styles of clients and therapists and examine these as possible interactions with each other and with treatment modality.

TIM STOCKWELL
National Centre for Research into the Prevention of Drug Abuse, Curtin University of Technology, GPO Box U 1987, Perth 6001, Western Australia

COSTELLO, R. M. (1975) Alcoholism treatment and evaluation: in search of methods, I. Collation of two-year follow-up studies, *International Journal of Addictions*, 10, 857–867.

EDWARDS, G., ORFORD, J. & EGERT, S. *et al.*, (1977) Alcoholism: a controlled trial of 'treatment' and 'advice', *Journal of Studies on Alcohol*, 38, 1004–1031.

EMRICK, C. D. (1975) A review of psychologically oriented treatment of alcoholism, II. The relative effectiveness of different treatment approaches and the effectiveness of treatment versus no treatment. *Journal of Studies on Alcohol*, 36, 88–108.

POTAMIANOS, G. & PAPADATOS, Y. (1987) The 'Accept' community programme for problem drinkers: a comparison with general hospital care, in: STOCKWELL, T. & CLEMENT, S. (Eds) *Helping the Problem Drinker: new initiatives in community care* (London, Croom Helas).

PROJECT MATCH RESEARCH GROUP (1997) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.

SUTHERLAND, G., STOCKWELL, T. & EDWARDS, G. (1985) Impact of a research interview on clinic re-attendance: a happy finding, *British Journal of Addiction*, 80, 211–213.

Patient matching in treatment for alcohol dependence: is the null hypothesis still alive and well?

Wayne Hall

In the light of Project MATCH, is it reasonable to accept the null hypothesis that there are no clinically significant matching effects between patient characteristics and cognitive–behaviour therapy (CBT), motivational enhancement therapy (MET) and Twelve-Step facilitation therapy (TSF)? The Project MATCH investigators considered the null hypothesis but preferred the alternative hypothesis that further analysis may reveal combinations of patient and therapist characteristics that show more substantial matching effects than any of the variables that they have examined to date.[1]

Their reluctance to accept the null hypothesis is understandable. Rarely has so much effort and intellectual horsepower produced so few results in favour of the matching hypothesis. The few matching effects identified at the 1-year post-treatment were small and conditional upon treatment arm, the outcome that was measured, and when it was measured.[1,2] One of the two matching effects that was discernible 3 years after treatment, that between social network support for drinking and TSF, was partly mediated by

post-treatment involvement with AA.[3] On the whole, this is not the stuff of which clinically useful treatment matches are made, as the Project MATCH investigators acknowledged.[1,2]

In declining to accept the null hypothesis, the Project MATCH investigators invoked the old shibboleth that one cannot prove the null hypothesis, a special case of the more general claim that it is "impossible to prove a negative".[1] This is one of these widely believed statements that happens to be false.[4,5] We do, and we must, decide what *not* to believe, including that the null hypothesis is true.[6,7]

It is reasonable to conclude that there are no matching effects if we can show: (1) that we have carefully looked for and failed to find matching effects in studies that were well designed to detect them; and (2) that we have ruled out plausible rival explanations of our failure to observe matching effects, such as inadequate statistical power, measurement error, inappropriate statistical analyses and the presence of confounding factors that may have obscured or attenuated any matching effects.[4,6]

The extraordinary effort that went into the formulation and testing of the matching hypotheses makes Project MATCH's failure to find matching effects a convincing null result. Exemplary care was taken in identifying the most plausible and empirically promising matching hypotheses. The outcome measures were validated and alternative measures of key matching variables were used. The three forms of treatment were faithfully implemented to ensure that they differed in the ways that they should. Subject attrition at the 15- and 39-month follow-ups was minimal. The statistical tests of the matching hypotheses used state of the art methods to test hypotheses that were specified a priori. A large sample was used that enabled the study to detect very small matching effects. Sensitivity analyses tested plausible explanations of failures to find matching effects without any change in overall result (e.g. use of MANCOVA in addition to latent growth curve analyses and analyses of other outcome measures).

The Project MATCH investigators remain hopeful that further analyses will identify complex patient profiles, or combinations of patient and therapist characteristics that show clinically or theoretically useful matching effects. The former seems unlikely, given the nugatory results of efforts to identify prognostically useful typologies of alcohol dependence, and the fact that one of the most promising typologies (Type A–B) included in Project MATCH failed to show matching effects.[2] The pursuit of therapist–patient-matching effects seems even less hopeful; it presupposes that the average effects of treatment are so modest that they depend upon who delivers them. Matching that requires assessments of patients and therapists is also impractical unless, as some have suggested, the assessment in Project MATCH was so therapeutic that it overwhelmed treatment effects (in which case we should assess rather than treat our patients).

The most reasonable conclusion from Project MATCH is that there are few, if any, practically or theoretically interesting matches between most patient characteristics and the psychological interventions that have been often advocated for the treatment of alcohol dependence. At the very least, the burden of proof now rests with those who believe that the modest average benefits of psychological interventions for alcohol dependence can be improved substantially by matching patients to treatment.

Project MATCH does not, of course, rule out the possibility that there will be patient matches with psychological and pharmacological treatments for alcohol dependence. Even so, we may be better to invest in research on treatments that have substantial effects on *most* of those who receive them, rather than assuming that average treatment benefits will be so modest that matching is required to maximize outcome.

Lastly, we should not allow the pervasive prejudice against the null hypothesis[8] to underestimate the value of Project MATCH. Convincing null results may be less exciting than novel positive findings but they still inform the field about the best way to invest limited research resources. Project MATCH has certainly fulfilled that role. In doing so, it has set new standards of excellence in treatment outcome studies of addictive behaviour.

WAYNE HALL
National Drug and Alcohol Research Centre
University of New South Wales
PO Box 1, Kensington, NSW 2033,
Australia

1. PROJECT MATCH RESEARCH GROUP (1997) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.

2. PROJECT MATCH RESEARCH GROUP (1997) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcome, *Journal of Studies on Alcohol*, 58, 7–29.
3. LONGABAUGH, R., WIRTZ, P. W., ZWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, Alcoholics Anonymous and long-term matching effects, *Addiction*, 93, 1313–1333.
4. HANSON, N. R. (1971) What I don't believe, in: TOULMIN, S. & WOOLF, H. (Eds) *What I Do Not Believe and Other Essays*, pp. 309–328 (Dordrecht, Holland, D. Reidel Publishing Company).
5. PASQUARELLO, T. (1984) Proving negatives and the paranormal, *Skeptical Inquirer*, 8, 259–270.
6. HALL, W. & EINFELD, S. (1990) On doing the "impossible": proving the non-existence of a putative causal relationship, *Australian and New Zealand Journal of Psychiatry*, 24, 217–226.
7. ROGERS, J. L., HOWARD, K. I. & VESSEY, J. T. (1993) Using statistical significance tests to evaluate equivalence between two experimental groups, *Psychological Bulletin*, 113, 553–565.
8. GREENWALD, A. G. (1975) Consequences of prejudice against the null hypothesis, *Psychological Bulletin*, 82, 1–20.

The "value" of Project MATCH for service provision

Christine Godfrey

Project MATCH was designed as a study to address an important question arising from a review of service provision in the United States (Institute of Medicine, 1990). This question was whether patient characteristics could be used to identify the most appropriate treatment. The funder of the project, NIAAA, expected the research findings to relate directly to practice in the United States. Matching client characteristics to different treatments seemed an important area to investigate not only for improving overall effectiveness rates of treatment but also in terms of reducing costs, or at least increasing the cost-effectiveness of treatment. Designing systems of services which assess individuals and assign them to specific packages of care, according to their characteristics, would seem to have much to offer to planners and funders of services. However, in the absence of most of the expected matching effects from the findings of the study, how can the results inform some of the practical concerns of planners and providers?

Providing information on the value for money of alcohol services has emerged as a major concern across the world in recent years, whether the underlying health and social care system is funded through taxation or private insurance, and whether agencies providing services are in the public sector, profit-making or in the voluntary or not-for-profit sector. For an economist, it was very disappointing that such a large research study was not designed with a primary economic evaluation component. A greater emphasis on an economic approach may well have injected a more pragmatic focus to the whole study, as well as providing useful results.

The study will, however, yield some results based on economic analyses. One paper on the direct provision costs of the three therapies in the trial has been published (Cisler *et al.*, 1998) and further analysis of the impact of treatment on future health care costs is under way. This element of the research, as with the many other components, merits careful consideration, especially by those looking for practical lessons to help their own service provision.

The first interesting finding is that the shorter MET treatment, with four planned sessions as opposed to 12 for TSF or CBT, is not as cheap as might be thought. An important factor is the actual amount of treatment received by the different groups. While all groups received substantial amounts of treatment, the MET group had a higher percentage of sessions attended in both the outpatient and aftercare arms of the trial (Project MATCH Research Group, 1997). Taking this into account, the contrast is not between four and 12 sessions, but more like three MET to seven or eight sessions of TSF or CBT.

Another factor is the cost of different therapists. In the study, the MET and CBT therapists had higher levels of qualifications than the TSF therapists. Such qualification differences inevitably lead to differences in incomes and hence the relative costs of therapists. Another factor is that an integral part of the MET therapy is the feedback of a number of tests, including blood tests. The instrumentation used in the trial was certainly extensive and consequently costly. Some savings may be made in practice and, given the demands of those funding service provision, more treatments may require at least some assessment and outcome measures. The important overall finding is that delivering a short but effective treatment may not bring a pro

rata reduction in cost in comparison to a longer therapy.

A more difficult question to address within this particular study is the cost of switching to manual driven therapies of the sort used in the trial, compared to current practice. The advantage of the Project MATCH approach is that treatments are "contained" in time and intensity. The costs of treatment in practice can be very variable: see Coyle *et al.* (1997). Some of this variation may come from the type of client excluded from Project MATCH but much would remain. There may be other cost advantages in being able to plan more accurately the throughput of clients if a manual-based approach is adopted. There is, however, a considerable amount of investment and continuing costs required if treatments are to be delivered at the same level of quality and consistency as achieved in the research project. It is to be hoped the NIAAA will fund some implementation studies to investigate both the costs and effects in practice.

There remain some important questions. One of the most important findings from Project MATCH, along with other research studies, is that alcohol treatment can achieve good outcomes. This cannot be overemphasized in the environment where most health care funders are attempting to restrict the types of intervention for which they are willing to pay. Well-trained therapists and a high quality of delivery of treatments, however, may not be the cheapest alternative. Can those who provide the finance for alcohol treatments be persuaded that extra outcomes are worth the cost? How does alcohol treatment compare to other areas of health and social care? In order to help answer these types of question, alcohol researchers may have to be willing to consider broader outcome measures than simply effects on drinking and alcohol-related problems. Some of the further analysis of the Project MATCH data may help clarify these issues—for example, if treatments result in lower future demands on the health care system.

Clearly, there is a wealth of findings already published from this study and more to come, which will continue to impact on those delivering alcohol treatments in many countries. These effects are in addition to the formidable advances in treatment research. Those funding treatment will not have the time, or possibly the expertise, to digest all these findings. It remains a challenge to ensure that the findings are not interpreted as a means to cut funding but rather that they are used to establish that treatment can be good value for money.

CHRISTINE GODFREY
Professor of Health Economics,
Centre for Health Economics,
Department of Health Sciences and Clinical Evaluation,
University of York,
Heslington,
York YO10 5DD, UK

CISLER, R., HOLDER, H. D., LONGABAUGH, R., STOUT, R. L. & ZWEBEN, A. (1998) Actual and estimated replication costs for alcohol treatment modalities: case study from Project MATCH, *Journal of Studies on Alcohol*, 50, 503–512.

COYLE, D., GODFREY, C., HARDMAN, G. & RAISTRICK, D. (1997) Costing substance misuse services, *Addiction*, 92, 1007–1015.

INSTITUTE OF MEDICINE (1990) *Broadening the Base of Treatment for Alcohol Problems* (Washington, DC, National Academy Press).

PROJECT MATCH RESEARCH GROUP (1997) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.

Future research directions and the impact of the MATCH project on research technology in the addictions

Luis San

The hypothesis of improving treatment outcome based on the careful assignment of patients to specific therapeutic approaches has always been present in the mind of many clinicians. So far, more than 30 studies on the matching hypothesis have been published. In statistical terms, matching research focuses on patient–treatment interaction effects rather than on treatment or patient main effects. In general, these studies show interesting results although with some methodological limitations that compromise interpretations of the findings or limit their generalizability. Project MATCH was designed to address many of the limitations of prior matching studies including a larger sample size, a wider range of clients and a more rigorous research design (Project MATCH, 1997a, 1997b, 1998a, 1998b, 1999; Longabaugh *et al.*, 1998).

Despite this attractive hypothesis Project MATCH has found similar efficacy rates for the various alcohol treatment modalities, independent of patient characteristics. Matching effects are not robust, with the authors giving different explanations for these unconvincing results (Project MATCH, 1997b). The three psychotherapy treatments used in Project MATCH are highly distinct from one another, and comparable regarding non-specific dimensions such as therapist skill (Longabaugh *et al.*, 1998). However, it is surprising that only individual therapy techniques were used, with patients included in either group allowed to become engaged in Alcoholics Anonymous (AA) meetings. It is also important to note the short duration of the psychotherapy, as a 3-month treatment in a chronic relapsing disease such as alcohol dependence seems a limited period to allow the observation of outcome modifications. Brief therapy has shown efficacy in patients with high motivation levels.

On the other hand, it is necessary to test these conjectures with long-term studies, as some treatment effects only emerge after some years have elapsed (Longabaugh *et al.*, 1998; Project MATCH, 1998a, 1998b). Thus, some type of post-treatment evaluation is in order to support a causal chain. Conversely, some significant matching effects appearing in the first year post-treatment—such as psychiatric severity—were not observed after 3 years (Project MATCH, 1998e) and other outcomes—e.g. reduction in drinking—were observed in the first year after treatment and sustained over the follow-up period.

A recently published paper (Geddes & Harrison, 1997) stated "Clinical psychiatry involves making difficult decisions about diagnosis, therapy and prognosis. Sometimes we may be entirely confident about our decisions, but often we are uncomfortably aware that we are making a choice without being sure there is convincing evidence to justify it. May be we don't know or have forgotten what the evidence is, or perhaps there isn't any." Alcohol use disorders have a variable course that is frequently characterized by periods of remission and relapse. Although some patients with alcohol dependence achieve long-term sobriety without active treatment, many others need treatment to stop the cycles of remission and relapse. The long-term goals of treatment include abstinence or reduction in use and effects, relapse prevention and rehabilitation.

So far, there is no agreement about which types of treatment are more effective, how long treatment should be, whether individual or group therapy is preferable or what kind of medication should be used for a particular type of patient. The search for effective treatments for substance abuse/dependence will continue to focus on the development of new agents, the refinement of the use of existing agents and the clarification of the appropriate role of psychological treatments that accompany pharmacological therapies (O'Connor & Schottenfeld, 1998). Although no single approach has been demonstrated clearly to be universally efficacious, several strategies appear promising.

In this respect, the Food and Drug Administration (FDA) has approved two medications as adjunctive treatments to decrease the likelihood of relapse in alcohol-dependent patients: disulfiram and naltrexone. Other drugs, including acamprosate, selective serotonin-re-uptake inhibitors, serotonin antagonists, GABAergic agents and dopaminergic agents, have also been studied for the prevention of relapse in disorders of alcohol use. However, their role in patient treatment remains to be determined. The choice of a therapeutic setting depends on many variables such as the clinical characteristics of the patient, the patient's preference, treatment needs and available alternatives. Different trials have to be performed in order to identify, with a high degree of accuracy, the patient most responsive to the different drugs, its optimal dosage and duration ranges and its effects when used in conjunction with alternative psychological interventions.

Addiction is a chronic disorder, but it also is a way of life with a clear influence on family, job and community. Therefore treatment must also be addressed to all these functional outcomes (Editorial, 1997). Another consideration to keep in mind is the reality of polydrug use. While most addicts use more than one drug most of our studies focus on only one substance, possibly explaining the lack of coincidence between research findings and treatment practice in the addiction field. It is clear that changing treatment systems needs more than good research evidence (Hodgson, 1994).

There are many difficulties inherent in the use of complex composite measures and we do not

know their exact significance. In the future we will probably need to increase our knowledge of patients and treatment, while applying new instruments of assessment and outcome (Ball, 1994).

Project MATCH will have an impact on clinical research, with the understanding that a new generation of drug treatment studies will have to recognize the need for and focus on a greater set of outcomes and functional measures. Despite considerable efforts made in treatment-matching, there is relatively little research to support the concept of matching, the studies published so far have not been conclusive, and to date we cannot know for sure the profile of the patient likely to respond to a particular treatment modality.

LUIS SAN
Benito Menni Mental Health Institute,
Barcelona,
Spain

EDITORIAL (1997) Drug abuse and addiction treatment research, *Archives of General Psychiatry*, 54, 691–694.

BALL, J. C. (1994) Why has it proved so difficult to match drug abuse patients to appropriate treatment? *Addiction*, 89, 263–265.

GEDDES, J. R. & HARRISON, P. J. (1997) Closing the gap between research and practice, *British Journal of Psychiatry*, 171, 220–225.

HODGSON, R. (1994) Treatment of alcohol problems, *Addiction*, 89, 1529–1534.

LONGABAUGH, R., WIRTZ, P. W., ZWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, Alcoholics Anonymous and long-term matching effects, *Addiction*, 93, 1313–1333.

O'CONNOR, P. G. & SCHOTTENFIELD, R. S. (1998) Patients with alcohol problems, *New England Journal of Medicine*, 338, 592–602.

PROJECT MATCH (1999) Editor's Introduction, *Addiction*, 94, 31.

PROJECT MATCH RESEARCH GROUP (1997a) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes, *Journal of Studies on Alcohol*, 58, 7–29.

PROJECT MATCH RESEARCH GROUP (1997b) Project MATCH secondary a priori hypothesis, *Addiction*, 92, 1671–1698.

PROJECT MATCH RESEARCH GROUP (1998a) Matching alcoholism treatments to client heterogeneity: treatment main effects and matching effects on drinking during treatment, *Journal of Studies on Alcohol*, 59, 631–639.

PROJECT MATCH RESEARCH GROUP (1998b) Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes, *Alcoholism: Clinical and Experimental Research*, 22, 1300–1311.

Project MATCH

Enoch Gordis & Richard Fuller

In any discussion of Project MATCH, the most important fact to keep in mind is the methodological rigor of the study. The most remarkable features of this rigor are the sample size, the procedures to assure fidelity of treatment and the follow-up rate.

With 1726 patients, Project MATCH is the largest controlled trial of psychotherapies conducted to date. This sample size makes it unlikely that Type II errors occurred, i.e. significant patient–treatment matches were missed. The sample represented a broad range of patients who were generally representative of patients treated for alcoholism in the United States.

Detailed procedural manuals were developed for the three treatments. (While developed for research purposes, these manuals met a provider need for well-specified therapies as evidenced by the over 35 000 requests for them received by the National Institute on Alcohol Abuse and Alcoholism.) The therapists were trained in the use of the manuals and received weekly on-site supervision, therapy sessions were videotaped and random tapes were reviewed by a central supervision unit. These procedures ensured that the therapies were delivered in a consistent and replicable manner.

The most remarkable achievement of Project MATCH was that 90% of the patients participated in *all* five scheduled assessment interviews during the year after treatment ended. Patients' reports of drinking status were verified by collateral reports, breath alcohol tests and liver function tests.

Other features of Project MATCH's methodological rigor included the use of innovative statistical procedures such as urn randomization to assign patients to treatments and hierarchical linear modeling techniques to analyze the data. A substudy ensured the reliability of interviews and questionnaires used to measure drinking, drug use, consequences of drinking, dependence, motivation, religiosity and AA involvement. Because of this extraordinary rigor, the results of Project MATCH are highly credible and merit serious consideration.

With regard to the results, only four matches were found of the 21 tested. Work that preceded Project MATCH had suggested that

patient–treatment-matching would improve treatment outcome. From these preliminary findings, Project MATCH investigators spent a great deal of effort and time selecting 21 a priori matching hypotheses to test. Finding only four matches was a disappointment to study investigators. However, these four matches, although modest, are worthy of clinical consideration.

Three of the four matches identified were found among the outpatients. The most consistent match was that motivational enhancement therapy (MET) was more effective for patients high in anger than either cognitive–behavioral therapy (CBT) or Twelve-Step facilitation (TSF), but less effective than CBT or TSF for those low in anger. This match was found at both the 1-year[1] and the 3-year follow-ups.[2] Patients in the highest third of the anger variable who were treated in MET had on average 76% abstinent days compared to 66% for those treated with CBT or TSF.[2]

A second match emerged by the 3-year follow-up. TSF was found to be more effective than MET for patients whose social networks were supportive of drinking.[2] Patients who where in the highest third of the support for drinking variable and treated with TSF had 16% more abstinent days than those treated in MET (77% vs. 61%). Conversely, MET was more effective than TSF for individuals with social networks not supportive of drinking. The Project MATCH investigators performed a causal chain analysis to identify what factor(s) explained why this match worked. This analysis found that AA involvement partially mediated this effect and was associated with better 3-year outcomes.[3] Lastly, TSF was more effective than CBT for those low in psychiatric severity, but this match began to wane by 9 months post-treatment and was no longer significant at 1 year.[4]

The fourth match was found in the aftercare patients. In this match, TSF was more effective than CBT for patients high in alcohol dependence. Patients in the highest decile of dependence and treated with TSF had 10% more abstinent days (94% vs. 84%) than those treated with CBT.[1] Conversely, CBT was better than TSF for those lower in alcohol dependence.

These results indicate that treatment staff should assess patients for these attributes upon entry into treatment programs. The MATCH investigators used the Spiegelberger State–Trait Anxiety Scale; the Important Persons and Activities interview; the psychiatric severity subscale of the Addiction Severity Index; and the Edinburgh Dependence Scale. It takes approximately 60 minutes to administer this battery of tests. A small investment in time could enable more appropriate treatment for patients. For example, assessing the social networks of patients and facilitating involvement in AA for those whose networks are supportive of drinking would be clinically relevant.

While evaluating patient–treatment-matching was the primary aim of Project MATCH, it was possible to compare outcomes of the three treatments because patients were randomly assigned the therapies. During treatment MET did somewhat worse than CBT or TSF. In CBT and TSF 41% of the outpatients were either abstinent or drank moderately without problems compared to 28% of the MET patients. This is not surprising because MET consisted of one-third the number of sessions of the other two treatments. However, after treatment MET achieved abstinent days and drinking intensity when an individual drank similar to the other two therapies. In terms of sustained abstinence, TSF produced better results in the outpatients (10% more patients achieving year-long continuous abstinence) than CBT or MET.

Interestingly, those patients who received a period of residential or day-hospital treatment immediately before enrolling in Project MATCH had more sustained abstinence and more abstinent days than those treated on an outpatient basis only.[4] This occurred despite the fact that aftercare patients scored higher on dependence than outpatients. It cannot be stated conclusively that residential/day-hospital treatment is superior to outpatient-only treatment because patients were not randomly assigned to aftercare or outpatient treatment. However, these results are similar to those of Walsh *et al.*[5] who found inpatient treatment superior to AA alone in a randomized study. The aftercare patients likely did better because they had more previous treatment or because selection bias occurred, but the results raise the possibility that a period of assured abstinence is important to recovery for many alcoholics. The haste of managed-care organizations in the United States to move treatment from a residential setting to an

outpatient setting may not be in the best interest of many alcoholic patients.

Because there was a rapid and sustained striking increase in abstinent days and marked reductions in drinks per drinking day overall, the Project MATCH investigators have concluded that patients did well. However, the MATCH team has been criticized for assuming this despite the absence of a no-treatment control. A no-treatment control was considered during the Project MATCH planning phase. However, it was not initiated because (1) the goal of the study was to study matching patients to treatments and (2) the investigators believed that their Institutional Review Boards would not permit denial of treatment to individuals seeking treatment. The criticism of lack of a no-treatment control has methodological merit, but the Project MATCH main effect results are important from a public health perspective. The results suggest that one-third to one-half of those seeking treatment in quality treatment programs will either be continuously abstinent or be drinking moderately without problems for a year after treatment. Similar results were seen at the 3-year follow-up. This is in stark contrast to the impression that many policy makers in the United States have that alcoholism treatment is not effective.

While the four matches discussed have some clinical relevance, it must be concluded that the overall findings from Project MATCH refute the appealing hypothesis that patient–treatment-matching will substantially improve treatment outcomes. While four matches were found, there were only four of 21. Three were in one setting and one in another, and one match was no longer present 1 year after treatment.

It is possible that matches might have been found by studying matching to settings or other forms of treatment (e.g. pharmacotherapies, family therapy). However, the Project MATCH results suggest that our knowledge of the mechanisms responsible for some individuals becoming addicted to alcohol is not as complete as we would like to think. The numerous studies under way in genetics and neuroscience should improve our knowledge. Building on this new information has great promise for developing new effective therapies, both pharmacological and psychological, to treat alcohol dependence. Once this happens, new patient–treatment matches may surface making re-visiting this issue worthwhile.

E. Gordis & R. Fuller
National Institute on Alcohol Abuse and Alcoholism, NIH, 6000 Executive Blvd, Willco Bldg, Ste 400, Rockville, MD 20892–7003, USA

1. Project MATCH Research Group (1997) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.
2. Project MATCH Research Group (1998) Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes, *Alcoholism: Clinical and Experimental Research*, 22, 1300–1311.
3. Longabaugh, R., Wirtz, P. W., Zweben, A. & Stout, R. L. (1998) Network support for drinking, Alcoholics Anonymous, and long-term effects, *Addiction*, 93, 1313–1333.
4. Project MATCH Research Group (1997) Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes, *Journal of Studies on Alcohol*, 58, 7–29.
5. Walsh, D. C., Hingson, R. W. & Merrigan, D. M. *et al.* (1991) A randomized trial of treatment options for alcohol-abusing workers, *New England Journal of Medicine*, 325, 775–782.

A contrast in treatment philosophies

Juan C. Negrete

One can only admire the magnitude of the resources invested in the Project MATCH study; the depth and breath of scientific expertise it commanded; the painstakingly fastidious design; the rigorous execution and, of course, its successful completion. The ability to mount a study of this sort is what makes the US scientific community the envy of researchers in other parts of the world. Alas, research on the effects of psychological treatments is a most difficult enterprise, and it would seem that even well-resourced projects such as Project MATCH cannot overcome the frustrating limitations which are inherent to the task. For all the mammoth effort it represents, this study has yielded a disappointingly meagre amount of conclusive evidence.

The rationale

Whether alcoholics with different individual "attributes" would be better served by treatment interventions which are distinct in content and method is, of course, a legitimate research question. The scientific community had grown weary of treatment systems which tended to be too stereotyped and undifferentiated. Such concerns, which were particularly topical in the 1970s and 1980s, have lost some of their justification in recent years as a result of the vigorous development of a variety of treatment practices which offer a much wider choice for patient placement. More alcoholics can now be treated in ambulatory programs of varying degrees of intensity; and there are at present specific services for the dually diagnosed which did not exist as such until recently.

The availability of more treatment choices, of course, renders the need to define patient placement criteria all the more pressing. Major efforts to that effect are being made in the United States; both the American Society of Addiction Medicine and the American Association of Community Psychiatrists have issued "level of care" guidelines.[1] These instruments are designed to help clinicians perform an adequate assessment of relevant clinical dimensions and select the type of services that would be appropriate in each case, according to the patient's condition at the time of referral.

Guidelines of this sort do address some basic treatment matching questions: who should be treated in which type of service; the selection between out-patient and residential care; and the decision on whether psychiatric intervention is warranted. Their helpful operational features notwithstanding, these American guidelines cannot claim scientific validity. They represent the consensus opinion of expert practitioners, but their accuracy and predictive value as clinical instruments have yet to be properly tested. Moreover, while providing criteria for patient placement in terms of the general structure of services and the professional skills required of the intervening agents, they do not deal with the choice of specific treatment philosophy and delivery method. It is precisely because of its potential contribution to clinical decision-making about treatment content that the results of the Project MATCH study were awaited with great expectation.

The scope

The Project MATCH trial did indeed offer a choice of three fairly distinct therapy interventions, and patients with a variety of individual attributes were randomly exposed to each of them. However, it was not designed to test for some of the matching issues clinicians most often have in mind when considering treatment options. To note some of the study's limitations: all Project MATCH treatments were given on individual, outpatient and short-term bases only. Basic treatment selection questions such as who is unsuitable for group therapy (the staple in most treatment services), who would benefit more from receiving therapy in the enforced abstinence conditions of an inpatient setting or who would do better if kept in therapy longer than 12 weeks were not tested in this study. Similarly, the criteria followed in the selection of subjects excluded some alcoholics whose characteristics often require treatment-matching decisions. Individuals with concurrent misuse of other drugs (except for cannabis) and those with significant psychiatric co-morbidity were not represented in the Project MATCH samples. Two categories which constitute a sizeable percentage of the cases seen in an average alcoholism treatment service.

The issue of psychiatric co-morbidity is particularly topical these days, but the approach adopted in the study would not satisfy current expectations in this area. Alcoholics with a history of psychosis and those feeling suicidal at the time of the baseline assessment were excluded outright. The rest were psychiatrically evaluated (only for mood and anxiety disorders) through the self-administered, computerized version of the DIS (DSM-III-R criteria). Some psychiatric conditions of much relevance to the selection of specific treatment modalities, such as personality or post-traumatic stress disorders, were not diagnosed and therefore not included among the "attributes" that could be matched to any particular form of therapy. Even in the case of the disorders which were explored, the study does not seem to have gathered the most useful information, for it reports life-time rather than current occurrence.

The Project MATCH researchers appear to have adopted a "dimensional" rather than diagnostic approach to the assessment of psychopathology, and some may find that choice somewhat less than clinically meaningful. For

instance, subjects were classified on the bases of a non-specific parameter such as "anger". Of course, the nature of this affect does vary widely, for it sometimes reflects a permanent character trait and for others a temporary psychopathological state. Unless properly understood within a more complete clinical formulation, the dimension "anger" is unlikely to serve the purpose of treatment-matching.

The scores obtained at the Beck Depression Inventory would suggest that the subjects randomized into all three treatment modalities were not clinically depressed. This observation, and the exclusion criteria mentioned before, lead to the conclusion that Project MATCH dealt mainly with alcoholics with rather low levels of psychiatric disturbance so that, in spite of its intention to match for psychiatric "severity", the study does not seem to have tested a range of pathology wide enough to respond to the treatment selection needs of the average alcoholism clinician.

The findings

The main goal was not, the Project MATCH group asserts, to compare the effectiveness of the three therapies employed in the study, but since the individual patient attributes they defined a priori do not seem to account for outcome variance to any significant extent, it is justifiable to focuse on group findings which depict the clinical response to each treatment approach. The most striking evidence is the success rate of the TSF (Twelve-Step facilitation) therapy which did as well as the the other two with all categories of patients, and significantly better with cases of more severe alcohol dependence and with those exposed to environments which did not support abstinence. Consistent with that observation is the finding that a better long-term outcome is linked to higher levels of involvement in the AA program, regardless of the type of therapy received during the active treatment phase of the project. Thus, the Project MATCH study confirms the recently published findings[2] of another large-scale comparison between treatments based on the AA or a CBT type of philosophy.

Having been delivered by a professional therapist and on a one-to-one basis, the TSF treatment offered in Project MATCH was not, nor did it intend to be, an adequate test of the AA approach. It lacked, among other things, the essential elements of group participation and peer interaction; but in helping the subject work through the first three steps of the AA program, it did mean to convey its fundamental principles of acceptance of alcoholism as a pathological condition over which the patient cannot hope to gain control, and of an alcohol-free life-style as the only viable solution.

These principles are in clear contrast with the message of the other two Project MATCH therapies, which assumed that alcoholics have the ability to make wise decisions with respect to drinking and the capacity to carry them out. Project MATCH was indeed a comparison between two very different conceptions of the problem, and it seems that the more categorical one has an advantage. The Twelve-Step recovery program is a long-term exercise in self-improvement; those who embrace it continue their therapeutic work well after succeeding in their effort to abstain from drinking. As such, the program involves the making of existential choices and the adoption of a well-defined philosophy of life. Changes that go beyond the mere modification of behaviour *vis-à-vis* alcohol.

However, the AA method also contains cognitive–behavioural features which are not widely recognized;[3] singleness of purpose; simplicity in the definition of the problem; clear and attainable objectives (e.g. 1 day at a time), reward for participation at any level (e.g. mere presence at a meeting greeted with expressions of support; the 1-month medal; the 1-year celebration cake, etc.).

In the specific instance of the Project MATCH study, it has been shown that the AA cognitive approach in the early phases of treatment is valid and capable of benefiting a wide variety of alcoholics without severe psychiatric impairment.

JUAN C. NEGRETE
Professor and Head, Addiction Psychiatry Program,
University of Toronto and
Addiction Research Foundation,
Division of the Centre for Addictions
and Mental Health,
33 Russell Street,
Toronto, Canada M5S 2S1

1. American Association of Community Psychiatrists (1998) LOCUS (Level of Care Utilization System for Psychiatric and Addiction Services), unpublished monograph (Pittsburgh, PA, AACP).
2. Ouimette, P. C., Finney, J. W. & Moos, R. H. (1997) Twelve-step and cognitive–behavioural treatment for substance abuse: a comparison of treatment effectiveness, *Journal of Consulting and Clinical Psychology*, 65, 230–240.
3. Edwards, G. (1996) Alcoholics Anonymous as mirror held to nature, in: Edwards, G. & Dare, C. (Eds) *Psychotherapy, Psychological Treatments and the Addictions*, pp. 220–239 (Cambridge, UK, Cambridge University Press).

Future research directions: a commentary on Project MATCH

Jim Orford

The Project MATCH investigators have set us a high standard for the conduct of randomized controlled trials (RCTs) in the addiction field. For the foreseeable future no one using an RCT design to study the outcome of treatment for people with addiction problems should fail to study their work with the utmost care and to use it as a benchmark against which to design their own research. There are numerous strengths to the Project MATCH design of which future researchers should take due not, and I can only mention a few here. Among the highlights for me are the following. First, there is the sheer *N* of it. A strong design in the psycho-social treatments research field requires a cast of several hundreds, and that requires a multi-site design. This has allowed a very large dataset to be explored thoroughly using elegant statistical methods such as the latent growth approach, which identifies the line of best fit to a single participant's data from different follow-up points (3, 6, 9, 12 and 15 months in this case) and uses that line in subsequent analysis.

The second highlight for me, for future note, is the immense care and thoroughness with which the whole concept of treatment type × client attribute-matching was examined, the literature on matching picked over and matching hypotheses chosen and, most importantly, the evidence examined rigorously to support or refute not only the basic matching hypotheses themselves but also hypotheses about the mechanisms or causal chains thought to underlie the existence of a match. Choosing an hypothesis on the basis of previous research, collecting as large and as complete a dataset as Project MATCH did in order to test it, finding statistical support when other variables were controlled and due caution was exercised to allow for multiple statistical tests carried out, and then to find support for a mediational, causal model of what is going on, is impressive indeed. This appears to have been the case in two instances. One involved support for a match between Twelve-Step facilitation (TSF) and higher alcohol dependence, predictive of a good outcome in the after-care arm of the study (a modest matching effect), which was traced to the greater focus of TSF on an abstinence goal. The second involved the match between TSF and higher social support for drinking which was related to a good 3-year outcome in the outpatient arm (after-care clients were not followed-up to 3 years). This, the strongest matching effect found, was partly accounted for by the higher AA participation of TSF clients (Longabaugh *et al.*, 1998). This is good science indeed.

Or, rather, it is good science of a certain type. Indeed, it is exemplary of a way of carrying out science in our field which is currently dominant and privileged and particularly favoured by grant-giving bodies such as NIAAA in the United States. Like all paradigms it has strengths and weaknesses, and is not the only way of doing things. Future researchers should not, in my view, be blinded by the very evident strengths of Project MATCH: the field as a whole should engage in open, wide-ranging debate about the best ways forward. The remainder of this commentary will be devoted to elaborating this point. Following Campbell & Stanley (1963) the distinction is often drawn between the *internal* and *external* validity of an experiment. Broadly speaking, internal validity refers to the degree to which the study was sufficiently rigorously designed so that conclusions can confidently be drawn about hypotheses (main effects of treatments, matching effects, etc.) within this *specific experiment*. External validity, on the other hand, refers to the degree to which the results can be generalized beyond the confines of the specific study, for example to other populations or settings. No single study can be perfect on all counts, and there is always a trade-off. A field as complicated as psycho-social treatment of an addiction problem is full of threats to both internal and external

validity. Too rigorous, and one can be accused of not being relevant to the real world. Too much in the real world, and one's research is likely to be insufficiently tight to be conclusive. The Project MATCH Research Group (1998) have already admitted, as is their frank and open style, that they sought to maximize internal validity perhaps at the expense of external validity. In my judgement this is certainly the case.

It cannot really be said that the Project MATCH investigators themselves were in error, but rather that the dominant research paradigm in which they, and most of the rest of us, have been schooled, strongly reinforced by their funders, dictates a strong bias towards internal validity. Let me illustrate what I mean by this. Over-concern with internal validity, at the expense of external, can result in close attention being paid to the internal minutiae of the experiment while the wider aspects of the context in which the study is set are ignored. The most obvious area of neglect is a consideration of the whole experience, for a client, of receiving treatment. Although clients were asked, via the Working Alliance Inventory, to give their views on their relationships with therapists, the dominant paradigm allows no room for exploring the meaning to clients of the treatment experience within the wider context of their own lives and the treatment system of which the experimental treatment is just a part.

Some of that context, such as family or work pressures upon a client to enter treatment, life events that may have influenced treatment seeking, the process of referral to the treatment site, and expectations of treatment, are very closely related to the treatment experience, but the dominant research paradigm puts little weight upon them and Project MATCH reports tell us little about them. These wider contextual issues have been studied by others, sometimes in the context of successful change without treatment, and sometimes asking treatment clients about their experiences (e.g. Edwards *et al.*, 1992; Tucker *et al.*, 1994).

Nor does this type of research tell us very much about the wider context of the service setting in which the specific treatments offered are embedded. References are frequently made in Project MATCH reports to site effects or site × treatment interactions, but these effects are peripheral to the main theme, matters to be controlled away statistically. No description is provided of the sites and no hypotheses advanced in order to interpret site effects. In particular no clear information is provided about the total population served by the treatment centres, the nature of their total case loads and the representativeness of the study sample. Such information is essential for judging external validity.

Not only does this leave us knowing very little about the meaning of the treatment experience for clients, but it may also result in a more fundamental design weakness. In order to gather all the information necessary for testing hypotheses, all clients in Project MATCH went through a very long, several-session, initial assessment, and a number of follow-up assessments in the 12 months after treatment. Informally, Project MATCH investigators admit that many of the clients may have known the researchers rather better than their therapists! Furthermore, in order to carry out a really controlled trial, it was decided to compare three forms of individual psychotherapy. Although the conceptual underpinnings of the three treatments were very different, and great pains were taken to ensure and to test that therapists really did carry out the treatment they were meant to be giving, general therapy skills were high for therapists in all three groups. A high level of working alliance was achieved in all three groups. Hence, despite theoretical differences, it is very likely that the three treatments were in most important respects fairly similar not just in outcome, but in design. Despite all the other strengths of the research design, similarity of the treatments and the large amount of pre- and post-assessment make the design, strangely, a rather weak one.

This concentration on that aspect of people's experience that is construed as "therapy" and which can be manualized is odd. It is as if a powerful magnifying glass is being held over one small piece of an obscure picture while the rest is ignored. The reason for this strange state of affairs is not hard to discern: it lies in the immense professional and cultural support for counselling and psychotherapy in the West, some would say the arrogance of the therapy industry, based upon what is essentially a medical model of psychological change (Pilgrim, 1997). As one would expect of research based upon such a tradition, Project MATCH and other research like it is virtually silent on questions of social class, sex, race and culture, at-

tributes of clients that are related to their social positions.

This debate is well enjoined in the wider psychotherapy literature, where both the extension of a simple "drug metaphor" for conceiving of therapy outcome and process (Shapiro *et al.*, 1994) and the dominance of RCT methodology (Goldfried & Wolfe, 1998) have come in for serious and considered criticism. Goldfried & Wolfe are of the opinion that there has been an, "... increasing medicalization of psychotherapy outcome research" (p. 145) under the influence of the US National Institute of Mental Health which has been increasingly oriented towards a biological model of treatment since the 1970s. Along with this trend has come a number of other tendencies, noticeable in Project MATCH, such as a focus on DSM diagnostic criteria and an over-emphasis on symptom reduction as the primary outcome measure. One recommendation I would make to future researchers in the addiction treatment field is to make it a rule to read one item of literature from the more general literature on psychological change and psychotherapy for every item of literature they read about addiction. There has long been a tendency for the addiction field to be parochial (Orford, 1975) and the signs are that this tendency is, if anything, intensifying. Most of the issues that concern us have been addressed in the broader literature. For example, the absence of differential outcomes from different types of psychotherapy (the "equivalence paradox") has been well recognized in psychotherapy research for some time.

However, Project MATCH was principally a thorough and powerful attempt to test the idea of treatment–client attribute matching which has been one of the great hopes for finding positive results in the alcohol treatment research field for a number of years past. The group's conclusion now is that there is probably not much in the matching idea. This could be because the beneficial effects of giving particular people a particular treatment, or the harmful effects of giving others that treatment, are only apparent towards the extreme ends of a client attribute continuum, and studies such as Project MATCH that are designed to be strong on internal validity may exclude just those participants whose characteristics are comparatively extreme, thus reducing the prospects of finding matching effects (Project MATCH Research Group, 1998). Alternatively there may exist important matching effects, but these may be more complicated still, involving different clients, different therapists, different treatments and different outcome criteria (Project MATCH Research Group, 1997). A more likely explanation—and this is the direction in which my argument runs, and I believe that of the general psychotherapy literature—is that different forms of individual psychotherapy, provided that they are competently delivered by skilled therapists operating in the context of an efficient, evaluation research-orientated, service delivery organization with sound assessment and follow-up procedures, are so functionally equivalent that differences between treatments—main or matching—are unlikely.

Whichever of these explanations for the failure of matching is correct, the overall conclusion that I come to is that the change process is a highly complicated one and its understanding is beyond the reach of research designed solely according to the traditional, dominant model. If there are complicated matches to be found then even the statistical power, and sampling of sites and therapists, of a study as comprehensive as Project MATCH are insufficient to find them with any confidence. What is more, complicated matches are very unlikely to be generalizable beyond a particular culture, epoch, or even type of treatment delivery system (the same match was never found in both arms of the Project MATCH study, for example). If, on the other hand, as I believe, it is largely factors outside of the relationship between an individual therapist and an individual client that have differential impact on the change process, then studies of change will always be inadequate if they focus solely on that relationship.

Either way, what we have here is, of course, a familiar epistemological dilemma. Human life, including changing an addictive behaviour, is a rich, varied and uncertain thing, in some ways not unlike the weather and the economy. Our research methods have to simplify the picture. Sometimes this works, but sometimes the simplification is too great to be useful (Elliott & Anderson, 1994). This danger is clearly recognized by the Project MATCH team. Babor at a recent meeting has described the model on which the project was based as a "technological, medical model", emphasizing the correction of individual client "liabilities" such as behavioural

deficits and negative personality characteristics. Elsewhere, the Project MATCH Research Group (1997) have acknowledged that, "... the simplicity of the a priori matching hypotheses was overwhelmed by the complexity of the findings".

So what is to be done? I see two ways forward, and I recommend we should take both.

We can try to have our cake and eat it

We can continue to use RCT methods, but if we do we must try to make the treatments as truly distinct as possible, we should reduce assessment and follow-up time as much as possible, we must keep exclusions of clients to a minimum, monitor exclusions very carefully, and study the treatment service-providing units and systems very carefully, and do far more than has been done in the past to study simultaneously what else is going on in clients' lives and how they perceive treatment and its impact in the context of their lives as a whole.

This is our strategy, for example, in the United Kingdom Alcohol Treatment Trial (UKATT) which is about to begin. Funded by the Medical Research Council (UK), the study is an RCT, comparing two treatments that we believe to be very distinct—motivational enhancement therapy (MET) and social behaviour and network therapy (SBNT). Whether we can, within one study, meet "The challenge for effectiveness research ... to add the component of external validity whilst still preserving internal validity" (Goldfried & Wolfe, 1998) remains to be seen.

We can try to topple the dominant paradigm

It is, as ever, comparatively easier to criticize than to be constructive, and critics of the dominant paradigm have admitted that an alternative way forward is not yet clear (e.g. Goldfried & Wolfe, 1998). Among existing methodologies there are some, however, that have been championed by groups of researchers, mostly from outside the addiction field, and these could be far more thoroughly exploited than at present. They include: the *events paradigm* that focuses analysis on specific incidents within therapy sessions chosen by clients and/or therapists as critical to change; *replicated single case design* studies, which start by exploring and testing hypotheses about process and outcome in individual cases, building more general knowledge by replication and comparison; and use of a *naturalistic research* design, whereby treatments are studied as they naturally occur, including detailed analysis of the wider treatment and social contexts and relying on correlational analysis to test hypotheses (used to good effect in the alcohol field by Moos, Finney & Cronkite, 1990).

Not inconsistent with the above suggestions would be a more radical shift towards a greater reliance upon qualitative research and participant/service user involvement in the development, execution and/or interpretation of the findings of research. There is already a growing tradition of qualitative exploration of the change process in the addictions, focusing either upon treated samples (e.g. Edwards *et al.*, 1992), or untreated populations (e.g. Biernacki, 1986). Inviting the participation of the recipients of treatments themselves, in a more active way than simply as providers of pre-determined types of data, would be more of a break with tradition in our field, but might be more in keeping with a human agency model of change which supposes that people are active shapers of change processes and outcomes rather than simply passive recipients of treatment ingredients designed and delivered by others.

All these suggestions for challenging the dominant paradigm are about coping with complexity. Elliott & Anderson (1994) put it elegantly when they stated:

> Indeed, the dialectic seems to be between those who like to keep their research problems neat and those who seek to go "back to the things themselves" in all their muddledness. For the most part, the simplifiers have been in ascendence in psychotherapy research, but there are signs that this is changing now (p. 63).

My point is that we need to have the debate, and it needs to be as open as possible. One extreme view would be that the Project MATCH methodology is the one and only gold standard, the one model that sets the way for the field to follow in the years to come. At the other extreme

would be a view that Project MATCH marks the end of the line for that kind of research which has gone as far as it can go: a new way of performing research should be privileged in the future.

JIM ORFORD
School of Psychology,
The University of Birmingham,
Edgbaston,
Birmingham, B15 2TT,
UK

BIERNACKI, P. (1986) *Pathways from Heroin Addiction: recovery without treatment* (Philadelphia, Temple University Press).

CAMPBELL, D. & STANLEY, J. (1963) Experimental and quasi-experimental designs for research on teaching, in: GAGE, N. (Ed.) *Handbook of Research on Teaching* (Chicago, Rand McNally).

EDWARDS, G., OPPENHEIMER, E. & TAYLOR, C. (1992) Hearing the noise in the system. Exploration of textual analysis as a method for studying change in drinking behaviour, *British Journal of Addiction*, 87, 73–81.

ELLIOTT, R. & ANDERSON, C. (1994) Simplicity and Complexity in Psychotherapy Research. In: R.L. RUSSELL (Ed.) *Reassessing Psychotherapy Research*, pp. 65–113 (New York, The Guilford Press).

GOLDFRIED, M. & WOLFE, B. (1998) Toward a more clinically valid approach to therapy research, *Journal of Consulting and Clinical Psychology*, 66, 143–150.

LONGABAUGH, R., WIRTZ, P. W., ZWEBEN, A. & STOUT, R. L. (1998) Network support for drinking, alcoholics anonymous and long-term matching effects, *Addiction*, 93, 1313–1333.

MOOS, R. H., FINNEY, J. W. & CRONKITE, R. C. (1990) *Alcoholism Treatment: context, process and outcome* (New York, Oxford University Press).

ORFORD, J. (1975) Alcoholism and marriage: the argument against specialism *Journal of Studies on Alcohol* 36, 1537–1563.

PILGRIM, D. (1997) *Psychotherapy and Society* (London, Sage).

PROJECT MATCH RESEARCH GROUP (1997) Matching alcoholism treatment to client heterogeneity: Project MATCH post-treatment drinking outcomes, *Journal of Studies on Alcohol*, 58, 7–29.

PROJECT MATCH RESEARCH GROUP (1998) Project MATCH secondary a priori hypotheses, *Addiction*, 92, 1671–1698.

SHAPIRO, D., HARPER, H., STARTUP, M., REYNOLDS, S., BIRD, D. & SUOKAS, A. (1994) The high-water mark of the drug metaphor: a meta-analytic critique of process–outcome research, in: RUSSELL, R. L. (Ed.) *Reassessing Psychotherapy Research*, pp. 1–35 (New York, Guilford Press).

TUCKER, J. A., VUCHINICH, R. E. & GLADSJO, J. A. (1994) Environmental events surrounding natural recovery from alcohol-related problems, *Journal of Studies on Alcohol*, 55, 401–411.

A study to remember: response of the Project MATCH Research Group

In a study as complicated, as long and as challenging as Project MATCH, there is always the dreaded possibility that hindsight will appear to be more wise than foresight. This is why it is refreshing to read so many well conceived comments that not only recognize the limitations of the scientific method and the fallibility of clinical research scientists, but also offer insights rather than insults, and congratulations as well as criticism. Instead of responding to each comment with the detail it deserves, we have chosen to focus on some general themes that will orientate people interested in the trial toward the considerable body of literature that has already been produced on these critically important topics.

Many of the concerns about Project MATCH are based on design features that after careful consideration were not included in the study: group treatment, an untreated control group, a longer duration of treatment, a broader spectrum of alcoholics, and the wider context of treatment seeking. These incisive methodological observations are as much a testimony to the challenges of treatment matching research as they are a critique of Project MATCH. In almost every respect, the comments reflect the numerous debates and compromises that the PMRG experienced on its way to designing and implementing the trial. Heather's comments, in particular, and the thoughtful answers he has provided to his own questions, are consistent with our own experience: there is no simple way to conduct a treatment-matching study.

Regarding the potential leveling effect of AA attendance on clients assigned to all three treatment conditions, it is clear that there is no easy way to control for this variable in treatment research in the USA. The approach we used—to *allow* MET and CBT participants to attend AA, and to encourage and facilitate TSF clients—proved to be one of the most intriguing aspects of the trial. Not only did we successfully manipulate AA attendance in the TSF condition (clients exposed to TSF did indeed attend more meetings), we also were able to study its repercussions on both short-term and long-term outcomes. It is important to note that most of the "matches" we observed (psychiatric severity, social support in the outpatient arm; alcohol dependence in the aftercare arm) were connec-

ted with TSF, and the treatment process and causal chain analyses we conducted suggest that certain of the hypothesized AA mechanisms (e.g. social network support for abstinence) but not others (e.g. spirituality) were instrumental in mediating the matching effects. This is consistent with a more dynamic concept of treatment that views the effects of a single episode of therapy as part of a sequence of interpersonal (e.g. motivation, self-efficacy) and environmental (a new and sober reference group) changes that together translate into reduced drinking and increased abstinence for a significant period of time.

Heather's remarks regarding "thinner vs. fatter" matching strategies are worth underscoring because they speak to many of the theoretical and practical issues that went into the design of the trial. Our choice of one particular matching strategy (i.e. matching to different types of short-term, outpatient psychotherapy) was based on feasibility, scientific evidence, clinical relevance and appropriateness for clinical trials methodology. Other types of matching, equally worthy but perhaps less amenable to scientific investigation, were left to future research. Whether our research hypotheses and methods were the right ones is a question perhaps best decided by future historians of science, rather than by us or our gentle critics.

Several commentators wonder whether we can infer from Project MATCH that there was any treatment effect at all, given the absence of an untreated control group. Moreover, if a treatment effect can be assumed, it is suggested that perhaps it had more to do with the client's research participation than the therapy received. To those who believe that it was the assessment rather than the treatment, we submit that the two are perhaps considered part of a more common culturally conditioned experience that cannot be separated in the eyes of a treatment-seeking client. The clients we studied were seeking treatment, not assessment, and not a research experience. None of the 1726 participants to our knowledge complained that they were not receiving treatment, and none suggested that the assessment and the follow-up evaluations were irrelevant to their recovery. As Lindström suggests, the rapid reduction of drinking (or the continuation of abstinence in the aftercare sample) following the initiation of Project MATCH treatment may be attributable to the fact that each of the Project MATCH therapies contained the essential ingredients of behavior change: an opportunity to experience success, a culturally approved mechanism to stop drinking and a basis with which to marshall environmental supports for continued sobriety.

As noted by one commentator (Finney), the findings of a recently completed large scale comparative study of cognitive–behavioral and Twelve-Step treatments were very similar to the results reported in Project MATCH, i.e. a slightly better prognosis for clients treated in Twelve-Step programs (primarily because of their post-treatment involvement in a social support network) and little evidence for matching. Our matching focused on basic baseline characteristics of clients. This type of matching is at the heart of the attribute by treatment interaction (ATI) hypothesis. ATIs represent the most logical and intuitive of the matching strategies offered prior to the Project MATCH trial. However, the research in educational and other areas of interactive interventions has found that this type of matching is very difficult to find and may represent belief rather than reality. Perhaps future research should be more attentive to the non-specific effects of treatment, or even the expectation of treatment, which seems to be a sufficient incentive for many alcoholics to initiate and maintain sobriety.

One commentator (Sutton) mentioned that we may not have tested accurately the type of matching suggested by the transtheoretical model proposed by Prochaska & DiClemente. Although motivational readiness to change was assessed in Project MATCH using measures that are capable of yielding stage-related profiles as well as a single score, our primary analyses used a single, second-order factor score for reasons of power. However, secondary analyses examined stage-specific subgroups and found results similar to those already reported, so there was no magic in stage status. Nevertheless, this commentator rightfully notes that the ATI type of matching in Project MATCH did not involve the process of change type of matching indicated by the transtheoretical model. Process measures in Project MATCH, however, did support the importance of motivation, self-efficacy and processes of change in predicting drinking behavior throughout the follow-up period. Matching interventions to the dynamic nature of the change process would be an important arena for investigation of possible matching effects.

Whether the approach we took to the investigation of matching was a wise investment of research funds has often been questioned. Although the total costs of the trial were approximately $28m, these costs should be considered in relation to the duration of the project (10 years), the number of subjects studied (1726), the extent of the follow-up evaluation (up to 39 months) and the publications (books, monographs, articles) and products (manuals, methodological innovations) that were produced. Clinical research tends to be expensive to conduct, and Project MATCH was no more expensive on a per subject basis than other clinical trials. Moreover, large-scale trials play an important, but by no means exclusive, role in advancing science. Many of the Project MATCH hypotheses were grounded in smaller, less methodologically rigorous studies. These were essential for the trial. The interplay between "big science" and "small science" is complex and the timing for moving from one to another is tricky. We believe that NIAAA invested wisely in the big science approach to treatment matching and that Project MATCH will prove to have a positive impact on subsequent research and clinical practice.

Another critique is that our results with regard to matching are unpersuasive because clients were not matched prospectively to treatments, or we chose the wrong attributes. There is an often-missed design point worth noting here. As we struggled with how best to design the study, we discovered that a clinical trial in which clients are assigned randomly to treatments is functionally equivalent to one in which clients are prospectively matched versus mismatched to treatments on a random basis. In both cases, clients stand an equal chance of being assigned to the hypothetically "right" or "wrong" treatment. This means that once clients have been randomized, one can test any number of possible matches retrospectively in the same manner as if the matching had been done prospectively. This also means that additional matches can be tested as long as the hypothesized matching attribute was measured in the trial. We took advantage of this fact to conduct post-hoc searches for any client–treatment matches that we might have missed, with maximal capitalization on chance. Even under these "fishing expedition" circumstances, we failed to uncover any client attributes that predicted differential response to our three treatments.

One frequently suggested methodological issue is the notion that the relative lack of matching is attributable to the overwhelming effect of the research assessments and frequent follow-up. In this view, the internal validity of the trial was undone by a combination of the Heisenberg Principle and the Hawthorne Effect. Although we do not agree entirely with this interpretation, we do agree that even if possible matching effects were "swamped" in this way, matching may not be expected to have much of an incremental contribution to treatment efficacy.

Although a number of commentators suggest that the results of Project MATCH were minimal, insubstantial, disappointing or irrelevant to clinical practice, at least four of our commentators (Lindström, Orford and Gordis & Fuller) believe that the matching effects we observed may be worthy of further investigation, if not actual application to routine clinical practice. We too take issue with the suggestion that the results were so negative that the matching hypothesis should be considered disconfirmed. First, the matches that were observed (anger, social support, dependence, psychiatric severity) were neither inconsequential nor uninteresting. Although the findings were disappointing to many, the modest evidence for some matching effects, combined with the emerging understanding of the treatment process itself, should not be considered the hydrologist's equivalent of a dry well (or Glaser's equivalent of a *Titanic* shipwreck). Although Project MATCH began with a relatively optimistic view of treatment-matching, as suggested by our peers and predecessors in the treatment literature, we have grown to appreciate the dynamic nature of treatment and recovery, and our data suggest that the process of treatment-matching is more complicated than we imagined. This is most apparent in our analyses of the causal mechanisms that were explored in supplementary analyses after matching had been confirmed or disconfirmed.

In all cases where matching was not supported, it was clear from the causal chain analyses that the hypothesized causal mechanisms were not operative. When matching hypotheses were supported it was as likely (alcohol dependence, support for drinking) as not (psychiatric severity, anger) that underlying mediators of the expected matching effects could be identified. Despite the considerable effort devoted to developing theory to support the matching predictions, it

was humbling to discover how wrong our theories could be. Our findings, as well as those of others, suggest that one direction for future treatment research is explicating the treatment process and its relationship to outcome. Relatedly, as suggested by Orford and other commentators, the relationship between treatment and the context in which it occurs must be brought to the forefront in our conceptualizations.

The ultimate significance of Project MATCH may still elude us as the findings unfold and their significance is understood. Orford's comments provide just this kind of philosophical view of Project MATCH, one that raises fundamental questions about the purpose and nature of empirical research. It has taken more than 6 years to collect the data and 3 years to carry out the data analyses, but our methodical approach has at times confounded our critics and rewarded our patient supporters. Because the findings are difficult to present within the context of a single paper, we have chosen to present the major results of the trial in the form of two books, one devoted to the causal mechanisms underlying the effectiveness of the therapies, the other to a complete synthesis of the findings on matching, therapist effects, treatment process and causal mechanisms. Only then can the full story of Project MATCH be told, and then only from a still-limited time perspective. As our understanding of treatment efficacy moves from a more static, episodic model to one in which systems of care and their community context are conceived as more dynamic and sequential, we may learn that Project MATCH was the beginning of a new way to understand individual paths to recovery rather than the end of our misplaced faith in a mechanical model of treatment-matching.

With his matchless wit, Glaser proposes that the grand scale and jarring climax of Project MATCH make the voyage of the *Titanic* an apt metaphor, and his imagery is echoed by Drummond. We had steeled ourselves to expect stern criticism, but the bridge offered by this first-class metaphor propels us to new depths of contemplation.

No one claims that the *Titanic* sank because it was too large or too well-designed. Steaming ahead in the dark, it hit an obstacle that would not move. While external validity can always be questioned, we believe it likely that other well-designed vessels hitting the same iceberg might encounter similar difficulties. Orford suggests that it might be wise to study the iceberg and, indeed, that the findings of Project MATCH may only reflect the tip of a much larger problem. From this perspective, prior treatment researchers may have been rearranging deckchairs while missing the big picture.

Then there is the disagreement among commentators as to whether the effectiveness of alcoholism treatment is buoyed or swamped by the findings of Project MATCH. While our findings do indicate that it makes relatively little difference who gets into which lifeboat, we admit that we are unable from this study to provide conclusive proof of the value of lifeboats in general.

Still other critiques are of the if-only variety: if only we had examined the data in another way, the results might have been different. The good news is that it is not too late. The whereabouts of the hull are known, and it may yet contain buried treasure. The maiden voyage is over, and the Project MATCH dataset was opened in January of 1998 for exploration by other treasure hunters. Maps are readily available for those who wish to tour its many chambers.

The full meaning of Project MATCH must be judged by history. Perhaps its legacy for the field will be in encouraging stronger designs, charting new courses, or inspiring greater humility and respect for the sea. We hope it will serve as a kind of anchor for future studies, providing a degree of stability when needed, yet not constraining further progress. If the excellent observations stimulated in the fourteen commentaries offer any indication of the ultimate impact of this trial on theory, research and clinical practice, then we did not embark on our voyage in vain.

Acknowledgement

This response was prepared on behalf of the Project MATCH Research Group by Thomas F. Babor, William R. Miller, Carlo DiClemente and Richard Longabaugh.