



Working Paper 03-56
Economics Series 21
October 2003

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624 98 75

SUBSTANCE ABUSE TREATMENT: WHAT DO WE KNOW? AN ECONOMIST'S
PERSPECTIVE*

Matilde P. Machado¹

Abstract

The substance abuse treatment literature has basically dealt with four important questions: 1) Is treatment effective? 2) Are all programs equally effective? 3) why do programs differ in their effectiveness? and 4) which treatments are most cost-effective?. This paper reviews the substance abuse literature around these four questions.

Keywords: Substance and Alcohol Abuse Treatment, Effectiveness, Cost-Effectiveness, Survey.

JEL Classification: I12, H42, H51, H53..

¹Departamento de Economía, Universidad Carlos III de Madrid. E.mail: mmachado@eco.uc3m.es

* Part of this research was done while I was at the IAE (Barcelona) as a Marie Curie postdoctoral fellow (European Community grant #ERBFMBICT972472). I also thank the research support from the Merck Foundation. Useful comments came from Daniel A. Ackerberg and Michael H. Riordan. Mercedes Cabañas provided invaluable assistance as a documentalist. The responsibility for all errors and omissions remains with the author.

Substance Abuse Treatment, what do we know? an Economist's perspective.

Matilde P. Machado*
Universidad Carlos III de Madrid

October 28, 2003

1 Introduction

The take off of substance abuse treatment programs, as we know them today, occurred during the opioid epidemic of the 60's when the US Federal Government released substantial funds to these programs.¹ Since then state governments and private sources have also contributed to the funding of these programs. The availability of research money and data drew the attention of academics and the literature on the subject mounted.

Forty years later, it is hard to survey the literature on substance abuse treatment without getting lost. I did! I will not aim, therefore, to survey the literature extensively. Instead, this paper, is my attempt to categorize and summarize, from the point of view of an economist, its main findings.

The literature on substance abuse treatment has focused on four questions: 1) is treatment Effective? 2) are all programs equally effective? 3) why do programs differ in their effectiveness? 4) which treatments are more cost effective? It may come as a surprise that it has taken almost forty years, and counting, to provide answers to these four questions. The reason lies in a range of methodological issues that have constrained the interpretation and generalization of results from previous research and, therefore, their capability to provide definitive answers. Section 2 of the paper addresses the four questions raised above while Section 3 offers a

*Part of this research was done while I was at the IAE (Barcelona) as a Marie Curie postdoctoral fellow (European Community grant #ERBFMBICT972472). I also thank the research support from the Merck Foundation. Useful comments came from Daniel A. Ackerberg and Michael H. Riordan. Mercedes Cabañas provided invaluable assistance as a documentalist. The responsibility for all errors and omissions remains with the author.

¹By 1996, the Federal Government was spending \$2.8 billion in illicit drugs treatment programs alone (French et. al, 1997).

brief discussion of the main methodological issues involved. Section 4 concludes.

2 Are there answers?

2.1 Is treatment Effective?

To address the question “is treatment effective?” one must define effectiveness. This is not straightforward. Fortunately, McLellan et al. (1997) provide us with an answer. As they point out “effectiveness” does not necessarily have a common definition for everyone but instead depends on the expectations of each agent involved in the treatment process or treatment financing (e.g. patients, society, judicial system, insurance companies, etc.). They define three objectives that represent the spectrum of agents’s expectations about treatment outcomes and these are: 1) reduction of alcohol and/or drug use; 2) improvement in personal and social functioning; 3) improvement in public health. McLellan et al. (1997) revise the literature and conclude that in spite of frequent methodological inaccuracies the weight of the evidence supports the conclusion that treatment is effective according to each of the three dimensions.

2.2 Are all programs equally effective?

The next obvious question, once we know that treatment is indeed effective, is whether all programs are equal in their effectiveness. The attempts to answer this important question probably started with Emrick (1975) who argued that the literature did not sustain evidence that programs were different in their performance. In contrast, the more recent literature, has innumerable references to big differences in effectiveness of substance abuse treatment programs even within the same treatment modality (e.g. Ackerberg et al., 2001, Anglin and Hser, 1990, Ball and Ross, 1991, McLellan et al., 1997, McLellan et al. 1993a). The methodological sophistication of the more recent studies allows us to give a definite answer to the question above and say: NO, not all programs are equally effective.

As an illustration let me describe McLellan et al. (1993a) and Ackerberg et al. (2001). McLellan et al. (1993a) restrict their study to four top private programs (2 outpatient, 2 inpatient), with no observable significant differences in their patient populations. The four programs shared the same philosophy and treatment goals. They find significant differences both in the number and type of services provided to

patients and the programs' effectiveness. The authors stress that amongst a more general pool of programs, in particular public programs, the differences are likely to be even larger as funding for public programs can be very unstable. Ackerberg et al. (2001) restrict their analysis to alcoholic patients in publicly funded programs of the same treatment modality, i.e. outpatient treatment. They control for observable patient characteristics as well as for unobservable heterogeneity that may lead to biases. One of their results is that programs vary quite substantially in their effectiveness to reduce patients' alcohol consumption.

2.3 Why do programs differ?

The question "why do programs differ?" is fairly recent in the literature and will most likely characterize the next decade on substance abuse treatment research. It deals with opening what the literature has labeled the "black-box" and inquires about the "active ingredients" of treatment (e.g. Ball and Ross, 1991, Finney et al., 1996, McLellan et al. 1993a), McLellan et al. 1993b), Moos, Finney and Cronkite, 1990, Morgenstern and Longabaugh, 2000).

As mentioned in the previous subsection, restricting to a single modality is no guarantee of homogeneity. The Institute of Medicine report (Gerstein and Hardwood, 1990) when defining outpatient non-methadone treatment programs states that they "display a great deal of heterogeneity in their treatment processes, philosophies, and staffing" as well as on their patient population. Another illustration of program heterogeneity and its sources is Ball and Ross (1991) evaluation of six methadone treatment programs. Methadone treatment is, probably, the most homogenous treatment modality, nonetheless, the authors point out "[...] it has not been sufficiently recognized, however, that individual programs rather than an entire modality may well be the crucial unit of analyses when determining treatment effectiveness."

In this subsection I will start by describing some of the factors traditionally used by the "black-box" literature to explain differences across programs. These are patient characteristics and a number of treatment characteristics such as treatment duration, intensity or special treatment services. In general, these characteristics are insufficient to explain the amount of heterogeneity across modalities and across programs within the same modality. Later in this subsection, I will mention some of the studies that concentrated on finding the "active ingredients" of the treatment process.

Most of the literature based on the “black box” type of approach has relied on patient characteristics as the main predictors of differences across programs. Usually, characteristics such as less severity of dependence, intact marriage, lower psychiatric symptoms, job, less family problems, minimal criminal activity, have been associated with better outcomes (Ackerberg et al., 2001, Anglin and Hser, 1990, Apsler and Harding, 1991, Lu and McGuire, 2002, McLellan et al., 1997). Moos, Finney and Cronkite (1990) reached the conclusion that in their sample of 5 residential programs, social background accounted only for less than 2 percent of the variance in drinking outcomes.

Treatment duration is another factor that is usually mentioned in the traditional “black-box” literature as accounting for differences across programs. The literature repeatedly refers to the “time effect” as the phenomenon whereupon patients that stay longer in treatment do better on average (Ball and Ross, 1991, Gerstein and Hardwood, 1990, Longabaugh et al. 1993, Lu and McGuire, 2002, McLellan et al. 1997, Moos, Finney and Cronkite, 1990, Tims et al., 1991 in reference to the DARP and Phoenix house studies²). Some of the studies that use time in treatment have not taken into account its potential endogeneity. Endogeneity is present if patients who dropout of treatment prematurely, for example, are harder to treat. This self-selection of patients into out-of-treatment would lead to overestimation of the effect of treatment duration on treatment outcomes. There are some exceptions where treatment duration is treated as an endogenous variable: Ackerberg et al. (2001) is a nice illustration where time in treatment is endogenized within a structural model of the treatment process; French et al. (1991) make use of first differences to eliminate error components that are correlated with treatment duration. Their technique only works if strong parameter assumptions hold; Finally, Lu and McGuire (2002) use instrumental variables to eliminate the bias introduced by time in treatment.

Time in treatment is also often used as a proxy for services received, under the premise that longer stays should yield a higher exposure to treatment. Again, it is important to control for endogeneity if one wishes to draw conclusions regarding causal effects. There are several references where one can find associations between better outcomes and the amount of services used. McLellan et al. (1993a) associated differences in outcomes across programs with type and quantity of treatment services. The authors reached the conclusion

²There is a new trend analyzing the effects of “brief interventions” for alcoholics and a number of studies claim their effectiveness relative to longer treatment. Drummond (1997), however, refutes this new literature mainly on the basis of sample selection. Trent (1998) is another example where shorter programs seem as effective as longer ones. Yet, in Trent’s study there is no randomization of patients into treatment alternatives and the sample is composed of rather stable individuals.

that programs that directed more services to a particular area had better patient outcomes in that area and that differences across programs were less accounted for by patient characteristics than by differences in the treatment programs. On another study, this time using randomized clinical trials, McLellan et al (1993b) found evidence that the addition of extra services (e.g. medical, psychiatric, employment, family therapy, etc.) to the baseline methadone programs proved to have a positive effect on outcomes. This piece of evidence is particularly important since the use of random clinical trials avoids, at least in theory, endogeneity and selection problems that are likely to produce biases in estimations using data from field studies. In 1998, McLellan and coauthors, showed that the introduction of Case Management doubled the patients' likelihood of abstinence. An association between participation in some treatment components such as therapy sessions, AA meetings, films, attendance to church services, etc. and outcomes is shown in Moos, Finney and Cronkite (1990) but no causation can be established as these are very endogenous events. Ball and Ross (1991) arrive to similar conclusions where again causality cannot be proved due to the endogeneity of their variables. Finally, Lu and McGuire (2002) use instrumental variables to control for the potential endogeneity of units of treatment received (as well as for treatment duration). They find a positive and significant effect of more units of treatment for moderate and heavy substance abusers although these effects vanish once they control for an interaction term of time and units of treatment.

Some studies have concentrated more on treatment intensity rather than time in treatment as a better predictor of differences in outcomes. Finney et al. (1996) review the fourteen best studies on the comparison of outpatient and inpatient programs. They conclude that only seven, i.e. half of the studies reviewed, found significant differences on drinking related outcomes between inpatient and outpatient programs. Out of these seven studies, five favored the inpatient programs, which proved to be the most intensive. Walsh et al. (1991) is another interesting example where patients are randomly assigned to 3 treatment modalities: regular AA meetings, AA meetings with 2 weeks of prior hospitalization and client's choice of treatment. They find that the 2 week prior hospitalization group addressed drinking problems significantly better. Yet, the large scale study MATCH (Matching Alcohol Treatment to Client Heterogeneity) addressed to alcoholics found no significant differences in the drinking behavior between the less intensive treatment modality and the more intensive ones (Brown and Wood, 2001). In a sample of residential treatment also for alcoholics, Moos, Finney and Cronkite (1990) found that treatment intensity alone explained around 1.4 to 2.9 percent of the

outcome variance. Interestingly, they also find that treatment intensity was mainly explained by a program dummy and that neither patients' consumption level at admission nor social demographic variables were highly associated with it. Although they do not conduct a formal test, this program idiosyncrasy constitutes evidence (but not proof) in favor of the exogeneity of treatment intensity in their study. It is possible, that the absence of a significant impact of intensity in the MATCH study was driven, as their critics point out, by the rather strict selection of patients into the study.

In some cases, as for example Ackerberg et al. (2001), adhering to the "black-box" type of evaluation is not only convenient but sufficient. These authors analysis has in mind the problem of allocating a fixed budget among a set of programs believed to follow roughly the same treatment philosophies and goals over time and patients. The authors do not ask explicitly about the active ingredients of treatment because their data is restricted to patient characteristics and broad characteristics of the treatment program/process. On the other hand, their dynamic structural model of the treatment process, something that is quite uncommon in the literature, unveils large differences between programs regarding observable and unobservable characteristics of patients at admission, patients' progress through treatment, patients' completion criteria, and programs' capacity to hold patients until treatment completion.³

As the need to study the active ingredients became more acute, researchers devoted themselves to the acquisition of very detailed data sets with information on all aspects of the treatment scenario (e.g. location, facility, staff, director, funding, staff enthusiasm and opinion about the program, treatment philosophy, etc.). Overall, the studies reviewed here reach the conclusion that big differences may follow from small details. A good reference is Ball and Ross' (1991) exhaustive study of six methadone treatment programs. In their study they find that leadership, organization, staffing patterns, and amount of services to patients were among the variables that accounted for a significant proportion of the variance across programs. In the same line of research, Joe, Simpson and Sells (1994) found evidence that both the admission staff and the staff responsible for the treatment plan were associated with the patient relapse rates at follow-up. They find that while psychiatrists and psychologists were associated with lower relapse rates, physicians were associated with higher relapse rates. This suggests that treatment quality may depend on the staff's specialty. Another study worth mentioning is Moos, Finney and Cronkite (1990) who obtain subjective indicators of quality of

³Joe, Simpson and Sells (1994) also model treatment in a dynamic framework but for their particular approach they need a truncated sample of people in treatment for at least 3 months.

treatment from patients and staff. Their questionnaire (COPEs) collects data on variables such as support, relationship between staff and patients, organization of the program, etc. One interesting finding is the idiosyncratic nature of these indicators, which were neither related to quantity/intensity of treatment, nor to the staff's background, nor to the usual indicators such as staff-patients ratio. The study by Moos, Finney and Cronkite (1990) offers, therefore, a very useful additional source of information regarding the active ingredients of treatment in the substance abuse field.

It is also worth emphasizing the state of the art research on evaluating the benefits of matching treatment to patient characteristics (e.g. Allen et al., 2003, Brown and Wood, 2001). Volume 94 (1) of *Addiction* is largely dedicated to the MATCH project, the first project where alcohol abusers were randomly assigned to three different outpatient interventions with the aim of looking not only for differences across interventions but also for differences across interventions resulting from particular patient characteristics. MATCH sought differences in treatment response for 10 such interaction or matching patient characteristics. The results showed that out of the 10 matching variables only "psychiatric severity" and "network supportive of drinking" turn out to be significant. The latter, however, was only found significant at the 3-year follow-up period while the former lost strength after the 1 year follow-up (Brown and Wood, 2001, Holder et al., 2000). Although these results were largely unexpected and disappointing there are important lessons to take from the MATCH project and its criticisms.

Finally, some isolated treatment strategies have been tried some with more success than others (e.g. Higgins et al., 1994, Longabaugh et al., 1998, Longabaugh et al. 1993). One strategy is to use medication to treat substance abuse. Acamprosate, for example, has proven very effective in preventing relapse in alcoholics (see the reviews by Mason, 2001 and Overman et al., 2003). Another strategy is to use employment as a treatment intervention (see Silverman and Robles, 1998). Silverman and Robles (1998) offer an explanation, based on income and substitution effects resulting from having a job, for why employment can actually be counter-productive. They suggest that a salary contingent on drug-free urine samples may be the incentive drug users need to stay abstinent.

2.4 Which programs are more cost-effective?

Cost-Effectiveness analysis of drug/alcohol treatment are much less frequent than effectiveness studies (see Apsler and Harding, 1991 for a review of some studies and their faults). Most studies on cost-effectiveness had in mind a comparison between the more expensive residential setting with the least expensive outpatient setting (Long et al., 1998, Longabaugh et al., 1983, Walsh et al., 1991, among others). The results are mixed as the next paragraphs will show. However, it is probably safe to conclude that for the majority of patients (i.e. except perhaps the most severe cases) the outpatient setting is more cost-effective. Below, I first describe a few of the studies that compare inpatient and outpatient cost-effectiveness and on the following paragraphs I concentrate on some studies that analyze cost-effectiveness of each modality separately.

Walsh et al. (1991), described in the previous subsection, reach the conclusion that costs of hospital treatment of the outpatient AA group was only 10% lower than the inpatient group while the latter did significantly better at reducing drinking. Quite contrary, Long et al. (1998) comparison of a 5 week long inpatient program with an outpatient program with 2 weeks residential treatment conclude that while the differences in outcome at 6 and 12 months are not significant, the revised shorter program was able to save an average of 33% on costs.⁴ Similarly to Long et al. (1998), Longabaugh et al. (1983) show that for a sample of fairly stable patients the charges of a Partial Hospital setting are significantly smaller (\$1658 less on average) than the charges of an Extended Inpatient treatment setting while there are no differences on post-treatment drinking, social functioning, rehospitalization or rehospitalization charges, and attendance to aftercare outpatient services. Finally, Holder et al (1991) review up to 33 sub-modalities for alcohol treatment. Their paper conveys a very strong message, that cheaper is most of the times also better.

Some studies have concentrated on a single modality. Machado (2001) uses a sample comprised of all publicly funded outpatient treatment programs in Maine for which the Maine Office of Substance Abuse collects standardized data on costs and outcomes. She estimates the impact of a marginal dollar on the abstinence rate of the representative program and finds it to be very small and not statistically significantly different from zero, which suggests that the Maine Office of Substance Abuse could have reduced its budget

⁴The authors acknowledge the fact that patients were not randomized into treatment. Furthermore, they alert about the possible unintended effects that the change in program within the same facility may have had, for instance, on staff and/or patient mood.

for substance abuse treatment without compromising its outcome. Ackerberg et al. (2001) using a patient-level data set similar to Machado's support her suggestion and identify the least cost-effective programs in Maine. The authors also simulate a number of "transfers" of best practices between programs that would reduce costs at least by 9.2% without worsening patients' outcomes. The finding that programs differ in their effectiveness and cost effectiveness is consistent with previous literature (Holder, 1991). Years later, Holder et al. (2000) compare the cost-effectiveness of the three different modalities offered in the MATCH project. They find that total medical costs for the least intensive 4-session Motivational Enhancement Therapy (MET) were lower than the total medical costs of the alternatives (12-session Cognitive Behavior Therapy and Twelve-Step Facilitation) over a three year follow-up period for which the MATCH team had found no significant differences on drinking outcomes.⁵

Barnett and Swindle (1997) study inpatient substance abuse treatment programs exclusively.⁶ They use readmission rates within 6 months as the outcome measure.⁷ They conclude that shorter (21-days) inpatient treatment programs are more cost-effective than longer (28-days) inpatient treatment programs⁸, also larger programs are more cost-effective than smaller programs, although, consolidation of small programs, they caution, may restrict access to treatment.

Cost-effectiveness of brief interventions (BI) was studied by Wutzke et al. (2001) and by Shakeshaft et al. (2002) under two very different approaches. Wutzke et al. (2001) do an extrapolation exercise on cost-effectiveness of BI using the results on effectiveness derived from the World Health Organization study on Identification and Management of Alcohol Related Problems in Primary Health Care, which involved several countries (see Wutzke et al., 2001 for more references on the WHO report). Wutzke et al. (2001) use "the number of life years saved" as their outcome measure and extrapolate the costs and benefits of nationwide implementation of the brief intervention program. They conclude that the marginal cost of a year saved

⁵The MET's superior result does not hold for the severe psychiatrically ill patients and for patients with highly supportive network of drinking where CBT was the most cost-effective (Holder et al., 2000). Of course, critics of the MATCH project would probably argue that the differences in total medical costs are not really differences in cost-effectiveness but represent either differences in alternative outcome measures that the MATCH team did not care to record or is a result of patient selection into the study. See Brown and Wood (2001) for a review of some criticisms directed at the project MATCH design and methodology.

⁶Although, inpatient treatment programs are less common, they still account for half the funds spent in substance abuse treatment in the United States (Barnett and Rodgers, 1997 as cited by Barnett and Swindle, 1997).

⁷The authors acknowledge the shortcoming of being able to control only for readmissions into Veterans Affairs hospitals.

⁸In the cost and effectiveness estimations, the authors do not use the real patient's length of stay because this variable is intrinsically endogenous. Instead they use the "intended" length of stay which is set at the program level by the program director. It is not clear to the reader why the authors opted for this approach instead of using the "intended" length of stay as an instrument for the real length of stay.

is below 1873 Australian dollars, a small number by any counts, which, most likely, will have an impact in stimulating the implementation of these type of programs.⁹ Shakeshaft et al. (2002) compare the cost-effectiveness of BI with the more traditional cognitive behavioral therapy (CBT). In spite of the limitations of their study, such as high attrition rates, they are confident in the conclusion that BI is statistically significantly more cost-effective than CBT for, at least, those with a low degree of alcohol dependence.

Finally, Palmer et al. (2000) use a disease model to simulate the long-term costs and benefits of alcohol standard therapy relative to the standard therapy with Acamprosate. They conclude that in spite of the short-run costs of the medication, Acamprosate enhanced therapy is more cost-effective in the long-run since patients enjoy a longer life expectancy at the expense of lower lifetime costs for the German Health Insurance System.

It is likely that this next decade sees a surge in cost-effectiveness studies. An important step towards this surge is, of course, the standardization of data on costs. French and McGeary (1997), and the references therein, describe a data collection methodology that allows the computation of the opportunity costs of substance abuse treatment programs.

3 Methodology Issues

The objective of this Section is to name the main issues that have constrained the homogenization and comparison of results across studies. Some of the methodological issues are a matter of beliefs or preferences (e.g. the choice of output), others, however, may be the cause of biased results (e.g. self-reported data or timing of measurement).

3.1 Outcome measure(s)

One of the main methodological issues that has constrained the literature's capability of giving definite answers to the four questions raised in the last Section is the researchers' choice of outcome measures to be

⁹Unlike other studies on the effectiveness of brief interventions, the WHO study is probably free from sample selection, for two reasons: First, because it only compares types of brief interventions with the default of "doing nothing" and not with totally different programs. Second, patients are classified according to their drinking abuse severity and results were obtained for "hazardous" and "harmful" drinking patients separately, where the "hazardous" and "harmful" classifications were obtained from a standardized screening test.

used on evaluations of substance abuse treatment programs effectiveness and cost-effectiveness. Within the three objectives proposed by McLellan et al (1997), stated in Subsection 2.1., there are still a wide range of possibilities. The recent years have been characterized by the redefinition of substance abuse treatment as a multi-product activity. Accordingly, the ASI (Addiction Severity Index), proposed by McLellan and his coauthors (McLellan et al., 1980), has been widely used in recent research. The ASI is composed of 7 distinct areas: employment; medical status; alcohol use; drug use; legal status; family and social relationships, and psychiatric symptoms. Several alternative sets of outcome indicators have been used as well such as the Maine Addiction Treatment System (MATS) implemented by the Maine Office of Substance Abuse (OSA) (Commons and McGuire, 1997, Commons et al., 1997), and the one in Moos, Finney and Cronkite (1990). MATS was designed to recover multiple outcomes at the patient level (i.e. abstinence, reduction in use, employability, job improvement, problems at job/school, problems with significant other/family, problems with the law and the judicial system, etc.) to inform the State agency about each program's performance. The intent to condition funding on performance represented a novelty in the financing of substance abuse treatment programs at the state level. Moos, Finney and Cronkite (1990) have a similar set of indicators in the following areas: alcohol consumption; abstinence; physical symptoms, depression; social functioning; occupational functioning.

The interconnection between the different multiple areas used to evaluate the substance abuse treatment performance is still not well understood in the literature. While Jaffe (1984), as cited by Apsler and Harding (1991), find them to be independent in the short run, Moos, Finney and Cronkite (1990) find that improvement in one of their 6 outcomes is associated with improvement in all others.

Although abstinence and reduction of use are considered by most researchers the most important outcome variables (e.g. Lu and McGuire, 2002, Machado, 2001, the MATCH project),¹⁰ others have considered output measures such as "time in treatment" (e.g. Apsler, 1991), relapse rates (Barnett and Swindle, 1997, Lu, 1999), and even labor market outcomes at 1 year follow-up (French et al., 1991). But, precisely because different outcome variables contain different information regarding patients' overall health status, some authors consider it to be important to integrate several outcome measures in their analysis. Ackerberg et al. (2001) is an example where several outcome variables are used to measure different aspects of treatment

¹⁰The "Harm-Reduction" movement in some European countries is an exception to this belief. It aims mainly at decreasing the devastating consequences of substance abuse (see Marlatt, 1996).

performance. Using data from MATS, the authors explore reduction in alcohol use as well as “status at discharge” (i.e. completion of treatment, dropout) and “treatment duration”.

Choosing an outcome measure is a matter of taste or belief. There is no unique way of measuring effectiveness. The variety of output measures complicates the comparison across studies but it has certainly enriched our knowledge about the impact of treatment.

3.2 Dropouts

How to deal with treatment dropouts is also not standardized in the literature. While Ackerberg et al. (2001) use discharge status (completion of treatment/dropout) as one of the outcome measures, Apsler (1991) points out that it is not clear whether or not dropouts should be included in cost-effectiveness studies. He argues that since these patients have not been subject to “enough” treatment for treatment to be cost-effective including them penalizes those programs that devote reasonable resources to dropouts. On the other hand, Ball and Ross (1991) evaluation of methadone programs say “treatment cannot be considered effective if a patient leaves shortly after admission or is asked to leave because of no-compliance.”

The statistical treatment of dropouts in evaluations of substance abuse treatment programs is not a marginal issue. There is a wide spread perception that dropout rates are high in substance abuse treatment with the exception of methadone programs (Apsler and Harding, 1991, Anglin and Hser, 1990, Ball and Ross, 1991, IOM report 1990). A related issue, is what affects the decision to leave treatment prematurely. This has not been thoroughly studied. Ackerberg et al. (2001) model the dropout decision as a constant hazard rate that depends on patient characteristics and program dummies. Their results show that this parsimonious model fits the data relatively well. Age, female, white, low severity, higher income, drunk driving, probation/parole, are good predictors of treatment’s retention.

3.3 Clinical trials versus field studies

Another important methodological issue raised in the literature is the validity of random clinical trials versus field studies. Fortunately, the literature is rich on both type of approaches. The biggest advantage of random clinical trials is, of course, their internal validity, that is the possibility of establishing causality and isolate

a particular effect. On the other hand, several authors have referred to the limitations of studies based on this approach (e.g. Heckman and Smith, 1995, Strohmetz et al., 1990). Following Heckman and Smith (1995), the difficulty of prohibiting the control group from receiving outside treatment is one of the major problems of clinical trials. Others as Strohmetz et al. (1990) (as cited in Finney et al., 1996) offer evidence to the sample selection involved in the voluntary assignment to a study based on clinical trials, in particular, patients that accept random assignment tend to be the more impaired and to have fewer social resources. On the other hand, others have argue that the inclusion/exclusion criteria to be eligible to a clinical trial study results in sample selection since the most severe patients are rejected. The strict selection of patients into the randomized clinical trials was, according to critics of the project MATCH, one of the main reasons for the disappointing results obtained (Brown and Wood, 2001). Another example of evidence of sample selection in experimental studies is given in the review by Finney et al. (1996). They show that when inpatient treatment is found superior to outpatient, it is more likely that the study was based on a naturalistic (do not use pre-selection of patients) rather than an experimental design (use pre-selection). Furthermore, it was also observed that while dropout rates vary among treatment groups in field studies, they tend to be very similar in experimental studies.

3.4 Self-reported data

Many studies in the literature use self-reported data as opposed to urine samples or other objective measures of substance use. MATS data which fueled a lot of research on Maine's substance abuse financing system and on substance abuse treatment effectiveness (e.g. Ackerberg et al., 2001, Lu and McGuire, 2002, Lu, 1999, Machado, 2001, Shen, 2003) is based on interviews to patients at admission and discharge from treatment. MATS relies, therefore, on self-reported measures although filtered by the interviewer clinician, which presumably increases its reliability.¹¹ Safeguards to guarantee the accuracy of the data, such as independence of the interviewer (McLellan et al., 1997), were not taken in MATS, perhaps because these safeguards are not feasible to implement in real life scenarios.

Although there are many examples where self-reported data is considerable valid and accurate (Ball and Ross, 1991, Grant, 1997, Long et al., 1998, McLellan et al., 1997, Midanik, 1982, Moos, Finney and

¹¹Lu (1999) finds evidence that after the introduction of Performance Based Contracting between the Maine Office of Substance Abuse and the treatment agencies, agencies changed their reporting practices in order to affect their budgets.

Cronkite, 1990), there are studies that show evidence of underreporting close to the admission time in order to facilitate entry into treatment (Aitken, 1986 as referred by Commons and McGuire, 1997 and Maitso et al., 1990). Some argue that patients enter treatment when they “hit the bottom” (Aplser and Harding, 1991) and, therefore, even if accurately stated, their measures of consumption would most likely be overstated. For populations other than substance abusers, there are several cases of misreporting (e.g. Høyer et al. 1995, Butler et al., 1987). Høyer et al. (1995), for example, estimates that the general population (i.e. not exclusively alcohol abusers) reports alcohol consumption that only amounts to 38.7% of total consumption.

3.5 Timing of measurement

When to measure substance abuse status is also not well established by the scientific community. Measuring substance use at admission time may be biased because of the “hit the bottom” effect mentioned in the previous subsection (Aplser and Harding, 1991). Measuring substance abuse at discharge may not reflect the long lasting effects of treatment (Aplser and Harding, 1991). On the other hand, longer follow-up periods may have the problem of confounding the effects of several treatment episodes.

4 Conclusion

This paper reviews the main questions and issues raised by the substance abuse literature. It is clear that treatment is effective although not all programs are equally effective. Moreover, although treatment may work better for patients with certain characteristics, e.g. low psychiatric problems, the “active” ingredients of treatment seem to be very idiosyncratic to treatment programs. Some studies have revealed that, for example, staff qualifications and their rapport with patients are important ingredients. It is also clear, that stable alcoholic patients do well in short inexpensive programs (MATCH project) while patients with more serious conditions benefit from extra services and longer treatment.

A lot has been learned so far about substance abuse treatment but researchers do not have all the answers. There is a need for studies that assess the long-run benefits and costs of treatment. Researchers also hope to learn more about the benefits of matching patients to treatment and continue their search for the “active” ingredients of treatment. The search for the most effective treatment will also provide leads about the most

cost-effective treatment.

Proponents of incentive schemes, such as conditioning funding on performance, may realize from this paper's review that, perhaps even more than for other health services, the multi-product aspect of treatment of substance abuse, the potential for patient selection, and the potential for misreporting of data are challenges that need to be dealt with in order to guarantee the success of such schemes.

References

- [1] Ackerberg, Daniel A., Machado, Matilde P. and Michael H. Riordan (2001): "Measuring the relative performance of providers of a health care service" NBER WP #8385.
- [2] Aitken, L.S. (1986): "Retrospective Self-Reports by Clients Differ from Ongoing Reports: implications for the evaluation of drug treatment programs," *International Journal of Addictions*, 21(7), pp 767-788.
- [3] Anglin, M. D. and Y. Hser (1990): "Treatment of Drug Abuse," in: Tonry, M. and J.Q. Wilson (Eds), *Crime and Justice: an annual review of research*, vol 13. University of Chicago Press.
- [4] Apsler, Robert and Wayne Harding (1991): "Cost-Effectiveness Analysis of Drug Abuse Treatment: current status and recommendations for future research," in: *Drug Abuse Services Research Series*, No 1: Background Papers on Drug Abuse Financing and Services Research. National Institute on Drug Abuse. DHHS Pub. No. (ADM) 91-17777. Washington, D.C., pp 58-81.
- [5] Apsler, Robert (1991): "Evaluating the Cost-Effectiveness of Drug Abuse Treatment Services," in: *Economic Costs, Cost-Effectiveness, Financing and Community-Based Drug Treatment*. Rockville, MD: National Institute on Drug Abuse, Research Monograph Series 113, pp 17-26.
- [6] Allen, John P., Thomas F. Babor, Margaret E. Mattson, and Ronald M. Kadden (2003): "Matching alcoholism treatment to client heterogeneity, the genesis of project MATCH" in Babor, Thomas F. and del Boca, Frances K. (Eds), *Treatment Matching in Alcoholism*, Cambridge University Press, Cambridge, United Kingdom.
- [7] Ball, John C. and Alan Ross (1991): "The Effectiveness of Methadone Maintenance Treatment: patients, programs, services, and outcome" Springer-Verlag New York Inc.
- [8] Barnett, Paul G. and Ralph W. Swindle (1997): "Cost-Effectiveness of Inpatient Substance Abuse Treatment," *Health Services Research*, 32 (5), pp 615-629.
- [9] Brown, Thomas G. and Wendy-Jo Wood (2001): "Are all Substance Abuse Treatments Equally Effective?," Comité Permanent de Lutte à la Toxicomanie, Deuxième trimestre 2001, Gouvernement du Québec, Ministère de la Santé et des Services Sociaux.
- [10] Butler, J. S., and Richard V. Burkhauser, and Jean Mitchell and Theodore P. Pincus (1987): "Measurement Error in Self-Reported Health Variables," *The Review of Economics and Statistics*, 69, pp 644-650.
- [11] Commons, Margaret, Thomas G. McGuire, and Michael H. Riordan (1997): "Performance Contracting for Substance Abuse Treatment," *Health Sciences Research*, 32 (5), pp 631-650.

- [12] Commons, Margaret and Thomas G. McGuire (1997), "Some Economics of Performance Based Contracting for Substance Abuse Services," in: Egertson, Joel A, Daniel M. Fox, and Alan I. Leshner (Eds), *Treating Drug abusers Effectively*. Milbank Memorial Fund and Blackwell Publishers, pp 223-249.
- [13] Drummond, D. Colin (1997): What I Would Most Like to Know "Alcohol Interventions: do the best things come in small packages?," *Addiction* 92(4) pp 375-379.
- [14] Emrick, Chad D. (1975): "A Review of Psychological Oriented Treatment for Alcoholism II - the relative effectiveness of different treatment approaches and the effectiveness of treatment versus non-treatment," *Journal of Studies on Alcohol*, 36 (1), pp 88-108.
- [15] Finney, John W., Annette C. Hahn and Rudolf H. Moos (1996): "The Effectiveness of Inpatient and Outpatient Treatment for Alcohol Abuse: the need to focus on mediators and moderators of setting effects," *Addiction*, 91(12), pp 1773-1796.
- [16] French, Michael T. and Kerry Anne McGeary (1997): "Estimating the Economic Cost of Substance Abuse Treatment," *Health Economics Letters*, vol. 6, pp 539-544.
- [17] French, Michael T., Gary A. Zarkin, Robert L. Hubbard, and J. Valley Rachal (1991): "The Impact of Time in Treatment on the Employment and Earnings of Drug Abusers," *American Journal of Public Health*, vol 81, No.7, pp 904-907.
- [18] Gerstein, D. R. and H. J. Hardwood (Eds) (1990): *Treating Drug Problems*. Washington, D.C.: Institute of Medicine, National Academy Press.
- [19] Grant, Kathryn A., and Lisa T. Arciniega, J. Scott Tonigan, William R. Miller and Robert Meyers (1997): "Are Reconstructed self-reports of drinking reliable?," *Addiction*, 92(5), pp 601-606.
- [20] Heckman, J. J. and Jeffrey A. Smith (1995), "Assessing the Case for Social Experiments," *Journal of Economic Perspectives*, 9(2), pp 85-110.
- [21] Higgins, Stephen T., Alan J. Budney, Warren K. Bickel, and Gary J. Badger (1994): "Participation of Significant Others in Outpatient Behavioral Treatment Predicts greater Cocaine Abstinence," *American Journal of Drug and Alcohol Abuse*, 20(1), pp 47-56.
- [22] Holder, Harold D., Ron A. Cisler, Richard Longabaugh, Robert L. Stout, Andrew J. Treno, and Allen Zweben (2000): "Alcoholism Treatment and Medical Care Costs from project MATCH," *Addiction*, 95 (7), pp 999-1013.
- [23] Holder, Harold, Richard Longabaugh, William R. Miller and Anthony V. Rubonis (1991): "The Cost Effectiveness of Treatment for Alcoholism: a First Approach," *Journal of Studies on Alcohol* 52 (6), pp 517-540.
- [24] Høyer, Georg, Odd Nilssen, Tormod Brenn, and Helge Schirmer (1995): "The Svalbard Study 1988-89: a unique setting for validation of self-reported alcohol consumption," *Addiction*, 90, pp 539-544.
- [25] Jaffe, Jerome, H. (1984), "Evaluating Drug Abuse Treatment: a comment on the state of the art," in: Tims, F. M. and J.P. Ludford (Eds), *Drug Abuse Treatment Evaluation: Strategies, Progress and Prospects*. National Institute on Drug abuse Research Monograph No. 51. Rockville, MD, pp 13-28.
- [26] Joe, George W., D. Dwayne Simpson and S.B. Sells (1994): "Treatment Process and Relapse to Opioid Use during Methadone Maintenance," *American Journal of Drug and Alcohol Abuse*, 20(2), pp 173-197.

- [27] Long, Clive G., Martin Williams, and Clive R. Hollin (1998): "Treating Alcohol Problems: a study of Programme effectiveness and cost-effectiveness according to length and delivery of treatment" *Addiction* 93(4), pp 561-571
- [28] Longabaugh, Richard, and Philip W. Wirtz, Allen Zweben, and Robert L. Stout (1998): "Network Support for Drinking, Alcoholics Anonymous and Long-Term Matching Effects," *Addiction* 93(9), pp 1313-1333.
- [29] Longabaugh, Richard, Martha Beattie, Nora Noel, Robert Stout, and Paul Malloy (1993): "The Effect of Social Investment on Treatment Outcome," *Journal of Studies on Alcohol*, 54(4), pp 465-478.
- [30] Longabaugh, Richard, and Barbara McCrady, Ed Fink, Robert Stout, Thomas McAuley, Cathy Doyle and Dwight McNeill (1983): "Cost-Effectiveness of Alcoholism Treatment in Partial versus Inpatient Settings - six month outcomes," *Journal of Studies on Alcohol* 44, No 6.
- [31] Lu, Mingshan and Thomas G. McGuire (2002): "The Productivity of Ourpatient Treatment for Substance Abuse," *Journal of Human Resources*, 37 (2), pp 339-335.
- [32] Lu, Mingshan (1999), "Separating the "True Effect" from "Gaming" in Incentive Based Contracts in Health Care," *Journal of Economics and Management Strategy*, 8, issue 3, pp 383-431.
- [33] Machado, Matilde P. (2001): "Dollars and Performance: treating alcohol misuse in Maine," *Journal of Health Economics*, 20, pp 639-666.
- [34] Maitso, S. A., J. R. McKay, and G. J. Connors (1990), "Self-Report Issues in substance Abuse: state of the art and future directions," *Behavioral Assessment*, 12, pp 117-134.
- [35] Marlatt, G. Alan (1996): "Harm-Reduction: Come as You Are," *Addictive Behaviors*, 1, No. 6, pp 779-788.
- [36] Mason, B. J. (2001): "Treatment of Alcohol-Dependent Outpatients with Acamprosate: a clinical review," *Journal of Clinical Psychiatry*, 62, Suppl 20:42-8.
- [37] McLellan, Thomas A., Teresa A. Hagan, Marvin Levine, Frank Gould, Kathleen Meyers, Mark Ben-civengo, and Jack Durell (1998): "Supplemental Social Services Improve Outcomes in Public Addiction Treatment," *Addiction* 93 (10), pp 1489-1499.
- [38] McLellan, A. Thomas, George E. Woody, David S. Metzger, Jim McKay, Jack Durell, Arthur I. Alterman, Charles P. O'Brien (1997): "Evaluating the Effectiveness of Addiction Treatments: reasonable Expectations and Appropriate Comparisons," in *Treating Drug Abusers Effectively* ed. by Joel A. Ergertson, Daniel M. Fox, and Alan I. Leshner, Milbank Memorial Fund and Blackwell Publishers, pp 7-40.
- [39] McLellan, A. Thomas, Grant R. Grissom, Peter Brill, Jack Durell, David S. Metzger, Charles P. O'Brien (1993a): "Private Substance Abuse Treatments: Are Some Programs More Effective than Others?," *Journal of Substance abuse Treatment*, vol. 10, pp 243-254.
- [40] McLellan, A. Thomas, Isabelle O. Arndt, David S. Metzger, George E. Woody, Charles P. O'Brien (1993b): "The Effects of Psychosocial Services in Substance Abuse Treatment," *Journal of the American Medical Association*, April 21, 269 No. 15, pp 1953-1959.
- [41] McLellan, A. Thomas, L. Luborsky, C. P. O'Brien, and G. E. Woody (1980), "An Improved Evaluation Instrument for Substance Abuse Patients: the Addiction Severity Index," *Journal of Nervous and Mental Disorders*, 168, pp 26-33.

- [42] Midanik, Lorraine (1982): "The validity of Self-Reported Alcohol Consumption and Alcohol Problems: A Literature Review," *British Journal of Addiction*, 77, pp 357-382.
- [43] Moos, Rudolph, and John Finney and Ruth C. Cronkite (1990): *Alcoholism Treatment*, Oxford University Press.
- [44] Morgnstern, Jon and Richard Longabaugh (2000): "Cognitive-Behavioral Treatment for Alcohol Dependence: a review of evidence for its hypothesized mechanisms of action," *Addiction*, 95(10), pp 1475-1490.
- [45] Overman, Gerald P., Christian J. Teter, and Sally K. Guthrie (2003): "Acamprosate for the Adjunctive Treatment of Alcohol Dependence," *The Annals of Pharmacotherapy*, 37, pp 1090-1099.
- [46] Palmer, A. J., K. Neeser, C. Weiss, A. Brandt, S. Comte and M. Fox (2000): "The Long-Term Cost-Effectiveness of Improving Alcohol Abstinence with Adjuvant Acamprosate," *Alcohol and Alcoholism*, 35 (5), pp 478-492.
- [47] Shakeshaft, Anthony P., Jenny A. Bowman, Sally Burrows, Christopher M Doran and Rob W. Sanson-Fisher (2002): "Community-Based Alcohol Counseling : a randomized clinical trial," *Addiction*, 97, pp1449-1463.
- [48] Shen, Yujing (2003): "Selection Incentives in Contracting" *Health Services Research*, vol 38 (2), pp 535-553.
- [49] Silverman, Kenneth and Elias Robles (1998): "Employment as a Drug Abuse Treatment Intervention," NBER Working paper No. 6402.
- [50] Strohmetz, D. B., Alterman, A. I. and Walter, D. (1990), "Subject Selection Bias in Alcoholics Volunteering for a Treatment Study," *Alcoholism: Clinical and Experimental Research*, 14, pp 736-738.
- [51] Tims, Frank M., Bennett W. Fletcher, and Robert L. Hubbard (1991): "Treatment Outcomes for Drug Abuse Clients" *NIDA REsearch Monograph 106* ed. by C. G. Pickens, C. G. Leukefeld, and C. R. Schuster. Rockville, MD: US Department of Health and Human Services.
- [52] Trent, L.K. (1998): "Evaluation of a Four - versus Six - Week Length of Stay in the Navy's Alcohol Treatment Program," *Journal of Studies on Alcohol*, 59 (3), pp 270-279.
- [53] Walsh, Diana Chapman, Ralph Hingson, Daniel Merrigan, Suzette Morelock Levenson, Adrienne Cupples, Timothy Heeren, Gerald Coffman, Charles Becker, Thomas Barker, Susan Hamilton, Thomas G. McGuire, Cecil Kelly (1991): "A Randomized Trial of Treatment Options for Alcohol-Abusing Workers," *The New England Journal of Medicine*, September 12, vol. 325 No. 11, pp 775-782.
- [54] Wutzke, Sonia E., Alan Shiell, Michelle K. Gomel and Katherine M. Conigrave (2001): "Cost-Effectiveness of Brief Interventions for Reducing Alcohol Consumption," *Social Science and Medicine* 52, pp 863-870.