Selected Papers of William L. White

www.williamwhitepapers.com

Collected papers, interviews, video presentations, photos, and archival documents on the history of addiction treatment and recovery in America.

Citation: White, W. & Godley, M. (2005). Addiction treatment outcomes: Who and what can you believe? *Counselor*, 6(3), 52-55. Posted at **www.williamwhitepapers.com**

Addiction Treatment Outcomes: Who and what can you believe?

William L. White, MA and Mark Godley, PhD
Chestnut Health Systems

Addiction counselors are sometimes asked, "How successful is your program?" or more globally, "Does addiction treatment *really* work?" Responses to such questions can be drawn from one's clinical experience, a study of one's own program, or from studies reported at professional conferences, in professional trade journals or in scientific publications. The problem is that the conclusions drawn from these multiple sources may be inconsistent or even contradictory.

Reporting addiction treatment outcomes has a long, problem-filled history. The first addiction treatment outcome study was conducted in 1874 by Dr. Joseph Turner, founder of the New York State Inebriate Asylum. Subsequent studies, some involving thousands of treated patients (Chamberlain, 1891; Crothers, 1893), became commonplace in the nineteenth century. The percentages of claimed "cures" declined as studies improved methodologically, but honest reporting of these outcomes conflicted with business interests as new competitors (private addiction cure institutes, private sanatoria, and bottled home cures) entered the field claiming 90-100% cure rates. The addictions treatment field in the late nineteenth century was plagued by a tension between the need for objective data to advance scientific knowledge and improve treatment protocol versus the need to claim high success rates to market services and raise funds.

More than a century later, discrepancies remain between addiction treatment outcomes reported in peer-reviewed scientific journals and the outcomes claimed by treatment industry representatives and the representatives of particular treatment programs. The goals of this article are to explore the roots of such discrepancy, to help addiction counselors critically evaluate treatment outcome claims, and to raise awareness of scientifically defensible methods for evaluating treatment outcomes.

Variability of Reported Outcomes

The authors' evolving careers over the past three decades as frontline clinicians, clinical and program directors, and clinical researchers have provided a unique vantage point to examine the discrepancies in outcome reports across the worlds of clinical practice and scientific research. Claims of 50-70%+ success rates for particular programs are common in our

queries of treatment program representatives, and the Internet is filled with treatment claims of 70-100% success rates—rates that far exceed those reported at scientific conferences and in the scientific literature. (A 2001 review of the largest and methodologically rigorous alcoholism treatment outcome studies by Miller, Walters and Bennett reported an average one-year continuous abstinence rate of 24%.) At the marketing and public relations level, we have witnessed promotional material that oversimplifies ("Treatment works") and overstates the complex and highly variable outcomes of addiction treatment (White, 2004). These discrepancies raise the question of who and what one can believe about the effectiveness of addiction treatment.

There are several possible sources of this discrepancy. Clinical outcomes do differ across client populations, addiction treatment programs and even across addiction counselors (Wilbourne & Miller, 2003; McLellan, Woody, Luborsky, & Goehl, 1988). It is certainly possible that some programs simply have far superior treatment services, have more competent individuals delivering these services, treat clients with better prognoses for recovery or some combination of these characteristics. We suspect, however, that the primary cause of these discrepancies lies in the quality of the study methods and procedures upon which statements about effectiveness are made.

Methodological Weaknesses in Local Follow-up Studies

Some claims of treatment success rest on no scientific foundation and instead represent everything from honest estimates to self-serving fabrications. However, the more common problem is that well-intentioned individuals within local programs conduct follow-up studies of discharged clients using methods that are so different from those used in scientific studies as to render the results across these worlds incomparable. Below we identify 10 of the most important criteria that differentiate scientifically rigorous studies of treatment effectiveness from the "home grown," or less rigorous studies.

Motivation for the Study: Addiction treatment programs are under ever-growing pressure to answer questions about their effectiveness. Their need to defend or extol the benefits of treatment can influence the design and conduct of a study in ways that bias study findings. In contrast, the best scientific studies are placed in the hands of trained clinical researchers whose primary goal is to implement the best possible scientific methods to evaluate the effects of particular interventions, not justify the legitimacy of treatment as a cultural institution.

Study Design and Treatment Specification: The question of "Is this treatment intervention effective?" is best answered in the context of, "Compared to what?" Follow-up studies of a single treatment program fail to answer the latter question. In contrast, the strongest studies answer these questions via random assignment of clients to different interventions whose active ingredients are specified in a treatment manual, delivered by staff who have completed competency-based training on the manual, and who are then monitored over time by a clinical supervisor to assure implementation fidelity. This approach allows comparison of effects for two or more treatments.

Inclusion Criteria: The high reported success rates of some programs may reflect a practice euphemistically known as "creaming"—admitting to treatment or including in the study only individuals with the best prognoses for long-term recovery, while excluding those with high problem severity. In contrast, scientific studies of program effectiveness attempt to assure that those being studied are representative of the larger pool of persons in need of such treatment or restrict statements of effectiveness to those persons who share the characteristics of those studied.

Intention to Treat: Reporting success rates solely of "graduates" or those "successfully completing continuing care"—a common practice in program marketing materials and local program evaluations--

inflates treatment outcomes by excluding those clients who failed to complete primary treatment or continuing care. In contrast, scientific studies base their conclusions about the degree of treatment effectiveness on "intent to treat" samples, meaning that once enrolled in the study, even clients who drop out after one day or one session are still included in the main analyses (for retrospective studies this would be the equivalent of including every client admitted during a specified time period, not just those who successfully completed treatment). This is a key point because treatment was *intended* to work for individuals who prematurely dropped out, therefore they must be included all the way through longitudinal follow-up and data analysis. This practice is essential in defining the success and limitations of interventions.

Sample Size: Small samples provide poor statistical power to detect treatment effects upon which conclusions can be drawn, e.g., is the treatment more effective for women than men, when women may make up a very small number of the overall sample. Alternatively, studies with small sample sizes may find a large effect (capitalizing on a few patients with excellent outcomes) that would later dissipate once, for example, 50 patients per condition or more were enrolled, treated, and followed-up. We have seen significant findings occur with 30 clients per condition only to be reduced to clinically trivial differences upon following up 50 clients per condition. Studies with small samples may be helpful in calling for large sample controlled replications, but they should not be relied upon for marketing, policy or program decision making.

Follow-up Rate: The low follow-up rates (less than 50%) that often characterize local treatment program follow-up studies artificially inflate recovery outcomes by excluding those individuals who couldn't be found at follow-up. Research shows that poorer functioning clients are more likely to be lost at follow up (Scott, 2004). The Center for Substance Abuse Treatment requires 80% follow up rates at 6 and 12 months for its discretionary grant programs and recently published follow-up technology is demonstrating consistent rates of 90+% several years after intake (Scott, 2004; Scott & Dennis, 2000; Cottler, Compton, Ben-Adallah, Horne, & Claverie, 1996).

Data Sources and Interviewer Reactivity: Most local program evaluations are based on client self-report data collected through telephone interviews or mail in surveys. In contrast, scientific studies often supplement client self-report with collateral interviews and chemical testing. Local evaluations often use clinical staff with a pre-existing relationship with the client to collect the data. This may introduce a reactive bias where the desire of the client to meet the clinical staff's expectations of complete abstinence leads to false negative self-reports of no or low use. In contrast, scientific studies use trained interviewers who are not affiliated with the treatment program to collect follow-up information.

Outcome Measures: A fundamental flaw in many local outcome studies is the use of a single "(and sometimes undefined) effectiveness measure. All references to "success rates" or "recovery rates" need to be operationally defined in clear and specific terms. The trend in scientific studies is to do pre- and post-treatment, and follow-up comparisons across multiple dimensions, e.g., percent abstinent during a defined period; percent of days abstinent per month, ounces of ethanol consumed per using episode, the number of problems related to use, and other measures of life-health functioning.

Peer-review, the "Gold-standard:" We have been surprised how many programs failed to provide us requested copies of the study upon which a claimed success rate was based. Such secrecy is the antithesis of good science. No study of treatment outcomes is credible until its methods and findings have been subjected to peer review, published, or at least made available for review by others.

Truth-in-Advertising: Some local programs base reports of their effectiveness on a single study from their distant past, even though the clinical design of the program has changed since the original study was conducted. Providers often make

changes to treatment models, including evidenced-based treatment manuals that are designed to be conducted according to the manual protocol! If we "drift" from the "xyz" manual or the treatment program that was evaluated previously---we should not claim that we are doing the "xyz program," or that our program will have a level of effectiveness based on a study for which the results no longer apply.

Improving Local Evaluations of Addiction Treatment

Most local programs do not have the financial and technical resources to design and conduct large randomized clinical trials, but there is much that can be done to improve the scientific integrity of their client follow-up studies. We offer the following prescriptions toward this end:

- vow to conduct the study in the most objective manner possible, using trained experts in clinical research methods to oversee your procedures from *recruitment to treatment* to *follow-up* to *analysis* and *report writing*,
- define the active ingredients of the treatment that is to be evaluated, assure that staff can and do competently deliver those critical ingredients and report dosage (i.e., how much of the treatment did each participant receive),
- report characteristics of clients upon whose experiences clinical outcomes are based,
- use the Consort guidelines (http://www.nature.com/bdj/about/Consort.htm) to report client eligibility, recruitment, enrollment, and attrition; remember to use the "intent to treat" standard,
- use as large a sample as possible, especially if you plan subgroup comparisons,
- achieve at least an 80% follow-up rate or be prepared to carefully limit your conclusions,
- except for pilot or preliminary investigations, report outcomes for the longest possible period you can afford. Ideally, one year post-treatment; but longer is clearly better,
- validate self-report data with collateral interviews and, whenever possible, chemical testing,
- use independent interviewers rather than clinical staff with pre-existing relationships with the study population,
- evaluate outcomes across multiple dimensions (e.g., changes in quantity and frequency of primary and secondary drug consumption and related problems) and multiple measures of health and functioning as might be found in the Global Appraisal of Individual Needs or the Addiction Severity Index,
- be your own worst critic before someone else is—look for alternative explanations for the results obtained and enlist the help of experienced clinicians to help with this, and
- make all studies upon which success claims are made available for professional and public review.

Recruiting local behavioral scientists for consultation on research design, study procedures, analysis, and report writing is highly recommended. Such qualified and experienced consultants may be found at local colleges and universities. There are also excellent written resources for conducting rigorous program evaluation (e.g., Cook and Campbell, 1979) and study guidelines for randomized clinical trials (http://www.nature.com/bdj/about/Consort.htm) that provide expert guidance for conducting methodologically sound studies.

Closing Thoughts

Studies emanating from the research community confirm that addiction treatment can improve the lives of a significant portion of persons experiencing severe substance-related problems, but inflating the potential power of addiction treatment through flawed methods of calculating clinical outcomes can produce unintended harm. Such miscalculations can harm clients by punishing them for failure to achieve unrealistic expectations. Inflated expected outcomes can harm families by creating unrealistic expectations that lead to the abandonment of a family member for whom treatment did not "work" as expected. Such overestimates may also harm the addiction treatment field by contributing to cultural pessimism about the prospects of immediate and sustained recovery.

If we are to avoid such injury and seriously pursue the science to practice agenda, we will need to follow a consistent set of standards to guide both the local conduct of treatment effectiveness studies and the reporting of such studies within professional and public venues. It is time we openly talked about what and who can be believed in the reporting of addiction treatment outcomes.

About the Authors: William L. White (bwhite@chestnut.org) is a Senior Research Consultant at Chestnut Health Systems and the author of *Slaying the Dragon: The History of Addiction Treatment and Recovery in America*. Mark D. Godley (mgodley@chestnut.org) is the Director of Research at Chestnut Health Systems.

References

Chamberlain, H. (1891). Modern methods of treating inebriety. *Chautaquan* 13:494-499. (July).

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design & Analysis Issues for Field Settings.* Boston, MA: Houghton Mifflin Co.

Cottler, L. L., Compton, W., Ben-Adallah, A., Horne, M., & Claverie, D. (1996). Achieving a 96.6 percent follow-up rate in a longitudinal study of drug abusers*. Drug and Alcohol Dependence*, 41, 209-217.

Crothers, T.D. (1893). *The Disease of Inebriety from Alcohol, Opium and Other Narcotic Drugs: Its Etiology, Pathology, Treatment and Medico-legal Relations.* NY: E.B. Treat, Publisher.

McLellan, A. T., Woody, G. E., Luborsky, L., & Goehl, L. (1988). Is the counselor an "active ingredient" in substance abuse rehabilitation? An examination of treatment success among four counselors. *Journal of Nervous and Mental Disease, 176*(7), 423-430.

Miller, W.R., Walters, S.T. & Bennet, M.E. (2001). How effective is alcoholism treatment in the United States? *Journal of Studies on Alcohol*, March Issue, 211-220.

Scott, C.K., & Dennis, M.L. (2000). A cost-effective approach to achieving over 90% follow-up in outcome monitoring with substance abuse treatment clients. *Journal of Drug and Alcohol Dependence, 60* (Suppl. 1), s200.

Scott, C.K (2004). A replicable model for achieving over 90% follow-up rates in longitudinal studies of substance abusers. *Drug and Alcohol Dependence*, *74*, 21-36.

White, W. (2004). Treatment works: Is it time for a new slogan? Posted at www.bhrm.org.

Wilbourne, P., & Miller, W. (2003). Treatment of alcoholism: Older and wiser? *Alcoholism Treatment Quarterly, 20*(3/4), 41-59.