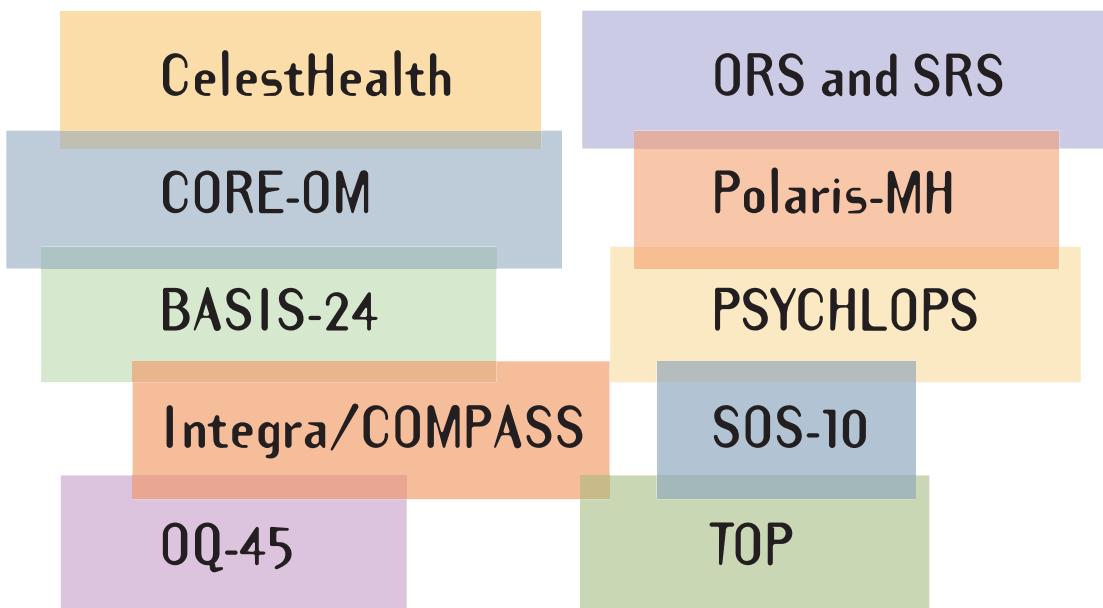


INTEGRATING Science and Practice

VOLUME 2 NUMBER 2 NOVEMBER 2012

10 Tools for Progress Monitoring in Psychotherapy



MISSION OF INTEGRATING SCIENCE AND PRACTICE

Integrating Science and Practice is published twice yearly by the *Ordre des psychologues du Québec*. The goal of the journal is to provide syntheses of scientific knowledge in the area of psychology and to facilitate the transfer of scientific knowledge to the field of practice. The journal aims to give practitioners in psychology, from all areas and fields of practice, the tools they need by providing them with critical reviews of the literature and brief syntheses of knowledge on specific themes. The journal is further intended to inform the public and professionals who work in collaboration with psychologists about recent scientific and clinical developments in psychology and about the contribution of psychologists towards improving people's quality of life.

The journal publishes articles by invitation only, following a call for proposals. Independent submissions are neither considered nor accepted. However, the editorial board may receive suggestions for themes. The choice of themes is made on the basis of their clinical relevance and their scientific, social and political relevance. Preference is given to articles that propose best practices in a specific field or context, or that question existing practices or policies based on available research findings. In every instance, the value of an article is assessed on the basis of its scientific merit and its potential for improving practices. All articles undergo anonymous peer review before being accepted and published.

Integrating Science and Practice is published by the Communications Department of the *Ordre des psychologues du Québec*. The reproduction and distribution of texts are authorized for non-profit purposes with reference to the source.

Editor: Martin Drapeau, Psychologist, Professor of Counselling Psychology and Psychiatry, McGill University

Publication Coordinator: Diane Côté, Director of Communications,
Ordre des psychologues du Québec

Editorial Board:
Rose-Marie Charest, President, *Ordre des psychologues du Québec*
Diane Côté, Director of Communications, *Ordre des psychologues du Québec*
Pierre Desjardins, Psychologist, Director of Quality and Development of Professional Practice, *Ordre des psychologues du Québec*

Martin Drapeau, Psychologist, member of the *Ordre des psychologues du Québec* board of directors and Professor at McGill University

Nathalie Girouard, Psychologist, Advisor on Quality and Development of Professional Practice, *Ordre des psychologues du Québec*

Special thanks to: Dr. Marilyn Fitzpatrick, Louise Overington, and Gabriela Ionita from McGill University, for their contribution to this issue.

Legal deposit
Bibliothèque nationale du Québec
ISBN 978-2-923164-49-6
(Original version: 978-2-923164-48-9)
ISSN 1923-189X

Ordre des psychologues du Québec
1100 Beaumont avenue, suite 510
Mont-Royal, QC H3P 3H5
www.ordrepsy.qc.ca/en



TABLE OF CONTENTS

INTEGRATING SCIENCE AND PRACTICE NOVEMBER 2012

The Value of Progress Tracking in Psychotherapy Martin Drapeau	5
The CelestHealth System Craig J. Bryan, Stephen Mark Kopta, Bryan D. Lowes	7
The CORE-OM (Clinical Outcomes in Routine Evaluation) and Its Derivatives Chris Evans	12
The BASIS-24 Behavior and Symptom Identification Scale Thomaskutty B. Idiculla, Susan V. Eisen	16
The Integra /COMPASS Tracking Assessment System Robert J. Lueger	20
The Outcome Questionnaire-45 Michael J. Lambert	24
The Outcome and Session Rating Scales Scott D. Miller, Susanne Bargmann	28
The Polaris-MH Grant Grissom, Tina Harralson, Jesse Nankin	32
The PSYCHLOPS (Psychological Outcome Profiles) Mark Ashworth, Maria Kordowicz, Peter Schofield	36
The Schwartz Outcome Scale - 10 (SOS-10) Mark A. Blais	40
The Treatment Outcome Package (TOP) David R. Kraus	43

Every English and French issue of *Integrating Science and Practice* is available on the *Ordre des psychologues du Québec* website:
www.ordrepsy.qc.ca/scienceandpractice

Integrating Science and Practice

- _ Documenting the effectiveness of Psychotherapeutic Interventions
- _ Depression: Considerations Surrounding Treatment Choices
- _ Addressing Childhood Suffering

Cahier recherche et pratique

- _ Documenter l'efficacité des interventions en psychothérapie
- _ La dépression : considérations autour des choix de traitements
- _ Reconnaître et traiter la souffrance des enfants

The Value of Progress Tracking in Psychotherapy

A previous issue of *Science and Practice* (March 2010) was dedicated to the assessment of one's clinical practice, more specifically to the value of systematically documenting the effects of the treatment one offers. The value of progress tracking is unquestionable, as it addresses the question that ultimately matters most: is this treatment, as I am delivering it now, helping this patient sitting in front of me?

Asking this question at an individual level, for each patient, and not only at a populational level, is certainly something that distinguishes psychologists from most policy makers. Indeed, psychologists often move well beyond a patient's diagnosis and treatment recommendations that are based exclusively on this diagnosis to also take into consideration a number of other important factors that can affect treatment outcome. There are of course good reasons for this. First, the DSM does not aim to explicitly and systematically account for extraneous and personal factors, such as social support available, marital status, or other factors that may affect a patient's prognosis, which in and of itself does raise some questions about psychotherapy recommendations that are based solely on diagnosis. The need to move beyond diagnosis and diagnosis-based treatment recommendations is also supported by research. A number of studies have shown that numerous patient characteristics account for variance in treatment outcome well beyond the effects related to fit between treatment type and diagnosis (e.g., Beutler et al., 2011; Joyce et al., 2007).

It is no surprise, then, that the conclusions of the American Psychological Association (APA) Presidential Task Force on Evidence Based Practice in Psychology (2006¹) emphasized the integration of patient, relational and treatment variables. Likewise, the APA and the Order of psychologists of Quebec statements on evidence based practice define evidence-based practice in psychology as the integration of the best available research with clinical expertise in the context of *patient characteristics, culture, and preferences*. This statement was reiterated and in many ways

expanded upon by the APA Council of Representatives in 2012². Likewise, in 2012, the Canadian Psychological Association (CPA) struck a Presidential Taskforce on Evidence Based Practice³ that defined evidence-based practice as "the conscientious, explicit and judicious use of the best available research evidence to inform each stage of clinical decision making and service delivery (... and the application of this knowledge...) in the context of specific *client characteristics, cultural backgrounds, and treatment preferences*". However, the Taskforce also innovated by adding to this that "following the initiation of treatment, data obtained from the ongoing monitoring of clients' reactions, symptoms, and functioning should be used to modify or discontinue the selected treatment". Hence, the very definition of evidence-based practice includes ongoing treatment monitoring and progress tracking. Good practice is not only evidence based; it is also practice based.

This is an important step forward, which is perfectly congruent with what science has taught us to date. A decade of research has shown the value of progress tracking, and its added benefit to the practice of psychotherapy. It can help clinicians who do not notice when a patient is deteriorating and even predict poor outcome before it happens (e.g., see Hannan et al., 2005). Second, progress tracking can improve retention and adherence to treatment and even improve treatment outcome (e.g., Anker et al., 2009; Bickman et al., 2011; Reese et al., 2010; Shimokawa et al., 2010; see also the special issue of *Canadian Psychology* on progress tracking⁴). It can also help to direct the clinicians' attention to areas and domains where they may require additional training or supervision, or help them identify those patients with whom they are most effective. Unlike symptom measures, progress tracking methods indicate areas that are problematic in the patient's life (e.g. family, etc.), by such providing the clinician with clinically useful information that can translate into specific techniques or lead to the discussion of specific topics within therapy. Furthermore, patients respond favourably to progress tracking (e.g., Anker et al., 2011),

¹ See <http://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf>

² See <http://www.apa.org/about/governance/council/index.aspx>

³ See <http://www.cpa.ca/researchers/sciencedirectorate/defebp/>

⁴ See *Canadian Psychology* (2012), vol. 53, issue 2.

REFERENCES

- Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, 77, 693–704.
- Anker, M., Sparks, J., Duncan, B., & Stapnes, A. (2011). Footprints of couple therapy: Client reflections at follow up using a mixed method design in routine care. *Journal of Family Psychotherapy*, 22, 22-45.
- Beutler, L. E., Harwood, T. M., Kimpara, S., Verdirame, D., Blau, K. (2011). Coping style. *Journal of Clinical Psychology*, 67(2),176-183.
- Beutler, L. E., Harwood, T. M., Bertoni, M., & Thomann, J. (2006). Systematic treatment selection and prescriptive therapy. In G. Stricker et coll. (éd.), *A casebook of psychotherapy integration*, p. 29-41, Washington, DC: American Psychological Association.
- Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R., & Riemer, M. (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. *Psychiatric Services*, 62(12),1423-1429.
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, 61, 155–163.
- Joyce, A. S., Piper, W. E., Ogrodniczuk, J. S., & Klein, R. H. (2007). Patient characteristics and variations in termination processes and outcomes. In A. S. Joyce et coll. (éd.), *Termination in psychotherapy: A psychodynamic model of processes and outcomes*, pp. 109-131, Washington, DC: American Psychological Association.
- Reese, R. J., Toland, M. D., Slone, N. C., & Norsworthy, L. A. (2010). Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy*, 47(4), 616-630.
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78, 298–311.

and, finally, a number of progress tracking methods provide tools and advice to help the clinician adjust the treatment on an ongoing basis.

With this issue, our goal is to provide readers with an easy to use introduction to the most popular progress tracking measures. Each paper presents one measure and is written by the author of, or important contributors to that measure. All authors kindly agreed to structure their paper in the same manner. Hence, each paper succinctly presents the measure, the populations it can be used with, the languages it is available in, the domains it assesses, how it is used, and how it can be useful in treatment planning. It also briefly presents the psychometric properties of the

measure and the technical considerations tied to using the measure, and provides a brief overview of the settings and institutions that have chosen that measure for their clinicians to use. Each of these measures is short, both clinician and patient friendly, pan-theoretical, and can be used in private practice as well as in other clinical settings.

We had received very positive feedback about the previous issue of *Science and Practice* on progress tracking. Along with this feedback, readers also asked where they can find progress tracking measures and what they actually look like. This issue, which presents 10 of the most widely used methods, is our response to those questions.



Martin Drapeau, Ph.D., Psychologist
Editor,
Integrating Science and Practice
Professor of Counselling Psychology
and Psychiatry,
McGill University

The CelestHealth System



Prof. Craig J. Bryan, PsyD, ABPP
The University of Utah &
The National Center for Veterans Studies

craig.bryan@utah.edu

Prof. Stephen Mark Kopta, Ph.D., H.S.P.P.,
University of Evansville

Bryan D. Lowes, NSCI Group, Washington Crossing,
Pennsylvania

The authors present the CelestHealth System, which can be used to monitor patient progress during psychotherapy. The System includes four instruments: the 43-item Behavioral Health Measure-43 (BHM-43), which assesses overall mental health functioning; the 20-item Behavioral Health Measure-20 (BHM-20), a shorter version of the BHM-43; the 5-item Psychotherapy Readiness Scale, which can be used to predict risks of poor treatment outcomes; and the 6-item Therapeutic Bond Scale, which evaluates the relationship between the psychotherapist and patient. Clinicians can select within the system which instruments to use and how frequently they are administered. The system is appropriate for adults 18 years and older of normal or greater intelligence, and can be used in outpatient mental health settings, primary care medicine, and college counseling center settings. This paper presents the system, reports on its psychometric properties, and describes how the system can assist in treatment planning and delivery.

Keywords: CelestHealth System; CHS-MH; Behavioral Health Measure; BHQ; Psychotherapy Readiness Scale; Therapeutic Bond Scale; treatment outcome; progress monitoring; psychotherapy

The CelestHealth System (CHS-MH) is the evolved tool of over 30 years of mental health treatment outcomes assessment. In the early days, the belief was more items meant better assessment. In the late 1970's, paper and pencil questionnaires became popular with research studies investigating self-reported outcomes. In the early 1980's, clinical tracking began with paper and pencil questionnaires given at every session. This tracking approach was incorporated into studies investigating dose-effect relationships across sessions, where dose is the number of sessions and effect was the probability of improvement (e.g., Howard, Kopta, Krause & Orlinsky, 1986; Kopta, Howard, Lowry, and Beutler, 1994). Offered by Integra Incorporated, COMPASS (Howard, Brill, Lueger, O'Mahoney & Grissom, 1995; Sperry, Brill, Howard & Grissom, 1996) arose in the 1990's as the most comprehensive outcomes assessment system featuring measures for mental health, psychotherapeutic bond, patient satisfaction, treatment need/expectations, and presenting problems. It totaled 123 items for the patient to complete and 7 life functioning items that the clinician completed. For a fee, COMPASS responses were faxed to an Integra office and a report was returned by fax to the clinician. At each session, the clinician reviewed the new clinical report with the patient. However, in the late 1990's, the demand for briefer questionnaires and more immediately available outcomes increased.

To answer this demand, the CHS-MH was created in part from COMPASS, with updated technology using computerized systems to provide a quicker clinical report to the clinician. The System features four instruments: (a) the 20-item *Behavioral Health Measure-20* (BHM-20) assesses mental health and takes 90 seconds to complete; (b) the 43-item *Behavioral Health Measure-43* also assesses mental health with additional subscales, and takes 3 minutes to complete; (c) the 5-item *Psychotherapy Readiness Scale* predicts risk to do poorly in psychotherapy and takes 30 seconds to complete; and (d) the 6-item *Therapeutic Bond Scale*, evaluates the relationship between the psychotherapist and patient in 30 seconds. All four instruments are optional, with the clinician selecting within the system which instruments to use and how frequently they are administered (e.g. every session, only pre-post treatment). The CHS-MH is appropriate for adults 18 years and older of normal or greater intelligence. It has been used in outpatient mental health, primary care medicine, and college counseling center settings. It is available in English, Spanish, and Vietnamese.

Domains Assessed

The BHM-20 measures the most frequently endorsed symptoms in outpatient psychotherapy consistent with the three phases of mental health change. This phase model proposes that improvement in behavioral health occurs in three progressive, sequential stages across therapy sessions, with improvement at each phase being contingent upon improvement in the previous stage. First, the client obtains a greater sense of *well-being* with increased optimism and hope. Next, specific *symptoms* such as panic attacks, depressive thinking, episodes of binge eating, and sleep disturbance diminish. Finally, *life functioning* improves across areas such as work, as a parent and partner, and in life enjoyment. The phase model of psychotherapy has been validated in several research studies (e.g., Howard, Lueger, Maling & Martinovich, 1993; Leon, Kopta, Howard & Lutz, 1999; Lutz, Lowry, Kopta, Einstein & Howard, 2001; Stulz & Lutz, 2007).

Within the phase domains, the BHM-20 assesses several mental health problems: (a) *well-being* (distress, life satisfaction, motivation), (b) *psychological symptoms* (depression, anxiety, panic disorder, mood swings associated with bipolar disorder, eating disorder, alcohol/drug abuse, suicidality, risk of violence), and (c) *life functioning* (work/school performance, intimate relationships, social relationships, life enjoyment). Within each phase domain, the longer BHM-43 has more subscales comprised of less frequently endorsed problems in outpatient psychotherapy. For example, the *symptoms scale* measures hostility, sleep disorder, obsessive-compulsive symptoms, and psychotic symptoms, and the *life functioning scale* additionally measures physical

health, self-management, and sexual functioning concerns. The CHS-MH is compatible with most clinical theories and practices.

Use and Procedures

The CHS-MH input and output information is securely communicated across computer networks to a centralized processing system. Before the session, the patient enters his/her responses using a computer (e.g., Netbook, iPad, or desktop computer) that is available in the waiting room. Using simple and familiar browser-based interfaces, patients respond to a maximum of 31 multiple choice items; the typical time for completion of the entire CHS-MH is 2.5 minutes. The patient's responses are analyzed and scored in secure CHS-MH servers; a complete, formatted report is immediately available to the clinician as soon as the patient finishes the assessment. The report produces several color-coded graphs and tables. Dose-Outcome graphs for the subscales show the patient's progress across sessions, the Behavioral Health Profile displays color-coded subscale scores based on normative data, and the Suicide Monitoring Scale helps clinicians to consider current suicide risk level. The clinician discusses with the patient his/her distress levels (i.e., severe, moderate, mild, normal) and level of suicide risk (high, moderate, low, no risk) as indicated by the Behavioral Health Profile. If chosen as an option, the clinician also discusses results from the Bond Scale and the Psychotherapy Readiness Scale. With this method, both patient and clinician can see which problems need to be targeted and followed as treatment moves forward in time. Based on this feedback, technical adjustments within therapy can be made to further reduce the patient's psychopathology and increase motivation.

All client raw responses and the calculated scale scores are stored in highly secure CHS-MH servers and can be used for research (with approval from an appropriate institutional review board) and administrative purposes. An important new feature is the CelestHealth Treatment Outcomes Profile, which uses collected data to provide an overview of the clinical status and progress for all clients within the center. This profile shows a variety of outcome variables including percentage of patients recovered, improved, unchanged, and deteriorated for the most recent session, as well as distress levels at intake and at most recent session. These variables are provided for all scales and subscales of the BHM-20, and the system is updated frequently throughout the day to provide a "real-time" snapshot of the center's client population. Date ranges can also be applied to look at the center's status over historical timeframes.

The annual license fee for centers that do not charge for services (e.g., college counseling centers) is \$US110 per full time equivalent clinician who uses information from the system, with a minimum cost of \$US550. For a private practice or center who charges for services, then the license fee is based on number of intakes because use of the system can be billed to insurance companies and Medicare/Medicaid. There are no set-up costs; consulting with CelestHealth staff is free by phone or email.

Assessment and Treatment Planning

The CHS-MH streamlines the assessment of clients' symptoms and functioning by helping clinicians to more rapidly identify problem areas and symptom domains of interest. From a clinical perspective, clinicians and clients can quickly determine

current clinical severity and historical trends over the course of treatment because of the CHS-MH's ability to immediately calculate and display scores using a color-coded scheme. Clinicians can use the CHS-MH's output to ask more targeted assessment questions and to focus conversations on specific factors that are more proximally related to treatment outcomes (e.g., interventions, life events). The color-coded feedback system also can be used to facilitate interventions themselves. For example, a client's responses or scores can be used to disconfirm their beliefs that treatment is not working, or to reinforce adherence and motivation for treatment (e.g., demonstrating improvement from one session to the next despite the patient's assertion that "things aren't getting any better"). The CHS-MH can therefore enhance clinical accuracy and efficiency, and can facilitate successful treatment outcomes.

Technical Support

Support includes three electronic manuals: (a) the *CelestHealth Getting Started Manual* describes the procedures for setting up administrators, clinicians, and clients on the CHS-MH; (b) the *CelestHealth Clinical Report Manual* explains to the clinician how to understand the CHS-MH output; and (c) the *CelestHealth Psychotherapist Manual* instructs the psychotherapist how to use the system with the client. Free online and telephone support is also available.

Psychometric Properties

The BHM-20 has demonstrated good reliability and validity. An initial psychometric evaluation was conducted by Kopta and Lowry (2002) using four samples (i.e., community adults, college students, college counseling clients, and psychotherapy

outpatients). Internal consistency coefficients ranged from .89 to .90 for the Global Mental Health score. For the three phases, the ranges were as follows: Well-Being, .65 to .74; Symptoms, .85 to .86; and Life Functioning, .72 to .77. Construct validity analyses using the discriminant validity method showed significant differences ($p < .001$) between the samples for all four scales, with each scale distinguishing clinical from nonclinical groups. Sensitivity to change using college counseling and the psychotherapy outpatient samples showed improved outcomes when comparing intake scores to session 3 scores for all scales. Concurrent validity is supported based on very high correlations of BHM-20 scales with the other well-known measures of mental health functioning including the BASIS-32 (-.83), COMPASS (-.76), OQ-45 (-.81), and the SCL-R-90 (-.85). More recently, the BHM-20's psychometric properties were investigated across seven separate samples (Blount et al., 2010): four primary care samples, two clinical samples of deployed military personnel, and one nonclinical sample of deployed military personnel. Internal reliability estimates were consistent with Kopta and Lowry's (2002) findings. The four scales also demonstrated medium to large correlations in the expected directions with more specific mood and symptom scales (e.g., happiness, fatigue, anxiety, depression, PTSD). Further psychometric evaluation of the BHM-20 is ongoing.

Patient /Client/Clinician Feedback

Positive feedback regarding the simplicity and practicality of the CHS-MH has been received from both clients and clinicians, especially regarding the color-coded feedback feature. Acceptability of the CHS-MH system is especially enhanced within medical settings when presented or described

as a method for quickly and reliably assessing clients' "vital signs" of mental health. Clients additionally report positive feelings about the use of the CHS-MH when clinicians review and refer to the clients' responses during the appointment. Clinicians have reported that the system's brevity and reliability/validity, as well as its separate measurement of life functioning from more generalized symptom clusters, are particularly attractive features. Primary care providers have similarly reported positive feedback because of the CHS-MH's separate measurement of daily functioning. The separation of a functional domain is not common among outcomes measurement tools, and is especially useful for clinicians' case conceptualization, diagnostic impressions, and treatment planning.

Unique Features of the Measure

Perhaps the CHS-MH's most distinguishing feature is its electronic, web-based format and immediate feedback system that can be implemented across computing platforms (e.g., PC, Apple, smart phone, tablets, etc.). This capability makes the CHS-MH system especially flexible and practical to use, and has contributed directly to its easy implementation across a wide range of clinical settings. For general clinical settings such as an outpatient psychotherapy practice, primary care, or emergency departments, the CHS-MH's design for measuring generalized mental health functioning, as opposed to diagnosis- or condition-specific symptoms, is especially beneficial. Within primary care, for example, the BHM-20 can be a better indicator of mental health functioning than more limited or restricted symptom measures of depression or anxiety given the wide spectrum of clinical issues that typically present in this setting (e.g.,

depression, anxiety, weight management, diabetes management, insomnia, chronic pain, etc).

The BHM-20 additionally includes screening items for suicidal ideation and impulses that have been shown to improve the detection of suicidal patients six-fold in primary care relative to standard interviewing and assessment approaches by primary care providers (Bryan et al., 2008). The CHS-MH's Suicide Monitoring System (SMS) was additionally developed in collaboration with suicide experts to aid clinicians in tracking and managing suicide risk over the course of treatment in a more reliable manner (Kopta, Mond, David, Potruzski & Doll, 2010), thereby helping clinicians to meet standards of care for suicide risk assessment and management.

The Psychotherapy Readiness Scale of the CHS-MH is unique as a reliable, valid alert for distinguishing patients who do poorly in psychotherapy.

Administered at intake, the client responds to five items in about 30 seconds that relate to the duration of presenting problems, previous psychotherapy experience, and motivation for treatment. Using Global Mental Health as the outcome variable, the scale has been shown to distinguish patients demonstrating poor psychotherapy outcomes from those showing good psychotherapy outcomes (Kopta, 2010).

Institutional Implementation

CHS-MH is used in college counseling settings including Harvard University, Johns Hopkins University, University of Minnesota, Indiana University, and the University of Florida. Other settings include primary care medical clinics at several U.S. Air Force Bases and university medical centers, as well as and private mental health clinics. The CHS-MH has been implemented by mental health professionals deployed to Iraq to track clinical outcomes.

Within primary care, the CHS-MH has been used to track outcomes for both general patient populations (Bryan, Morrow & Appolonio, 2009; Corso, Bryan, Corso, Morrow & Kanzler, 2010; Ray-Sannerud et al., 2011) and for specific subpopulations (e.g., PTSD; Cigrang et al., 2011; Corso et al., 2009). The CHS-MH's BHM-20 has additionally been used to improve the detection of suicidal patients in primary care clinics (Bryan, Corso, Rudd & Cordero, 2008), and is currently undergoing pilot testing as a suicide assessment aid in emergency departments. The CHS-MH's recent addition of an option for tracking psychotropic medication along with clinical outcomes provides an especially useful tool for health care providers of all disciplines, but most notably psychiatrists and nonpsychiatric prescribers (e.g., primary care physicians).

REFERENCES

- Beck, A.T., Ward, C., & Mendelson, M. (1961). Beck Depression Inventory (BDI). *Archives of General Psychiatry*, 4, 561-571.
- Blount, T.H., Bryan, C.J., Kanzler, K.E., Morrow, C.E., Corso, K.A., Corso, M.L. (2010, November). *Psychometric properties of the BHM-20 in military samples*. Poster presented at the annual meeting of the Association for Behavioral and Cognitive Therapy, San Francisco, CA.
- Bryan, C.J., Corso, K.A., Rudd, M.D., & Cordero, L. (2008). Improving identification of suicidal patients in primary care through routine screening. *Primary Care and Community Psychiatry*, 13, 143-147.
- Bryan, C.J., Morrow, C.E., & Appolonio, K.A.K. (2009). Impact of behavioral health consultant interventions on patient symptoms and functioning in an integrated family medicine clinic. *Journal of Clinical Psychology*, 65, 281-293.
- Cigrang, J. A., Rauch, S. A. M., Avila, A. L., Bryan, C. J., Goodie, J. L., Hryshko-Mullen, A., Peterson, A. L., & the STRONG STAR Consortium. (2011). Treatment of active-duty military with PTSD in primary care: early findings. *Psychological Services*, 8, 104-113.
- Corso, K.A., Bryan, C.J., Morrow, C.E., Appolonio, K.K., Dodendorf, D.M., & Baker, M.T. (2009). Managing posttraumatic stress disorder symptoms in active duty military personnel in primary care settings. *Journal of Mental Health Counseling*, 31, 119-136.
- Corso, M., Bryan, C.J., Corso, K.A., Morrow, C.E., & Kanzler, K.E. (2010, April). *Mental health functioning recovery curves associated with behavioral health consultant services in integrated primary care*. Poster presented at the annual meeting of the Society of Behavioral Medicine, Seattle, WA.
- Derogatis, L.R. (1983). *Administration, scoring, and procedures manual—II*. Baltimore, MD: Clinical Psychometric Research.
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). The Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, 56, 885-895.
- Howard, K.I., Brill, P.L., Lueger, R.J., O'Mahoney, M.T., & Grissom, G.R. (1995). Integra outpatient tracking assessment. Philadelphia: Compass Information Services, Inc.
- Howard, K.I., Kopta, S.M., Krause, M.S., & Orlinsky, D.E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist*, 41, 159-164.
- Howard, K.I., Lueger, R.J., Maling, M.S., and Martinovich, Z. (1993). A phase model of psychotherapy: Causal mediation of outcome. *Journal of Consulting and Clinical Psychology*, 38, 139-149.
- Kopta, S.M. (2010, October). *The future is here for college counseling centers: Counseling works, preventing suicide, and improving outcomes*. Workshop presented at the Annual University and College Counseling Center Directors (AUCCCD) Conference, Portland, Oregon.
- Kopta, S.M., & Lowry, J.L. (2002). Psychometric evaluation of the Behavioral Health Questionnaire-20: A brief instrument for assessing global mental health and the three phases of psychotherapy outcome. *Psychotherapy Research*, 12, 413-426.
- Kopta, S.M., Howard, K.I., Lowry, J.L., & Beutler, L.E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 1009-1016.
- Kopta, S.M., Mond, M., David, L., Potruzski, N., & Doll, L. (2010, June). *Assessing suicidality: Validation of the Suicide Monitoring Scale of the Behavioral Health Measure-20 and implications for psychotherapeutic strategies*. Paper presented at the annual meeting of the International Society for Psychotherapy Research, Monterey.
- Leon, S.C., Kopta, S.M., Howard, K.I., & Lutz, W. (1999). Predicting patients' responses to psychotherapy: Are some more predictable than others? *Journal of Consulting and Clinical Psychology*, 67, 698-704.
- Lutz, W., Lowry, J.L., Kopta, S.M., Einstein, D., & Howard, K.I. (2001). Prediction of dose-response relations based patient characteristics. *Journal of Clinical Psychology*, 57, 889-900.
- Ray-Sannerud, B., Bryan, C.J., Dolan, D., Morrow, C.E., Corso, K.A., Kanzler, K.E., & Corso, M.L. (April 2011). *Preliminary evidence for long-term outcomes following brief behavioral health intervention in primary care clinics*. Poster presented at the annual meeting of the Society of Behavioral Medicine, Washington, DC.
- Sperry, L., Brill, P.I., Howard, K.I., and Grissom, G.R. (1996). Treatment outcomes in psychotherapy and psychiatric interventions. New York: Brunner/Mazel.
- Stultz, N. & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: The phase model revisited. *Journal of Clinical Psychology*, 63, 817-833.

The CORE-OM (Clinical Outcomes in Routine Evaluation) and its derivatives



Dr. Chris Evans¹, MRCPsych., M.Sc.,
Institute of Group Analysis,
Nottinghamshire Personality Disorder
and Development Network,
Mandala Centre

chris@psyctc.org

This paper presents the Clinical Outcomes in Routine Evaluation (CORE-OM), a 34 item self-report questionnaire designed to measure change in the mental health of adults in the context of psychotherapy service delivery. The questionnaire includes items with a primarily intra-personal focus and others with a primarily interpersonal focus. It assesses a number of different domains, including client well-being, problems and symptoms, functioning and risk. All CORE-OM derived measures can be scored easily by hand with scoring boxes on the instruments. Particular strengths of the CORE-OM include its broad domain coverage with risk "flag" items; the existence of well tested short forms; adaptations for families, young people and people with learning difficulties; the mapping of some items to health economic valuation; and the availability of the system in 22 languages. The psychometric properties of the questionnaire are discussed along with its use in treatment planning and delivery and the procedures, costs and training it requires.

Keywords: Clinical Outcomes in Routine Evaluation; CORE-OM; treatment outcome; progress monitoring; psychotherapy

The CORE-OM (Chris Evans et al., 2000) is a 34 item self-report questionnaire designed to measure change in mental health of adults, particularly change brought about by psychological therapies. It was launched in 1998 as a central part of the "CORE system" complemented by a practitioner completed instrument (CORE-A) comprising the CORE-TAF (Therapy Assessment Form) and CORE-EoT (End of Therapy) and by two 18 item shortened forms for repeated use (CORE-SF/A and SF/B; Michael Barkham, Margison, et al., 2001). Subsequent work produced another shortened version for general population surveys (GP-CORE; Sinclair, Barkham, Evans, Connell & Audin, 2005), and two more short forms for sessional and screening use: CORE-10 (in prep.) and CORE-5 (Wright, Bewick, Barkham, House & Hill, 2009). It has been translated into 22 languages to date, including a French version currently in progress, and adaptations exist for adolescents (YP-CORE; Twigg et al., 2010), for people with learning difficulties (CORE-LD Brooks & Davies, 2008; and in prep.) and for description of families (SCORE-15; Stratton, Bland, Janes & Lask, 2010 and SCORE-28; Cahill, O'Reilly, Carr, Dooley & Stratton, 2010). A health economic Quality of Life (QoL) scoring is emerging (Mavranezouli, Brazier, Young & Barkham, 2010) as are algorithms or lookup tables to map CORE-OM scores to and from scores on other questionnaires such as the BDI-II (Leach et al., 2006).

The CORE-OM assesses change in as wide a group of clients as possible, from those with no problems to those with very serious thoughts of suicide, self-harm or other severe distress. It was not designed for forensic services nor for people with current paranoid disorders who might mistrust its use, though it is as unprovocative for such clients as possible. The measure was also not intended for use by adolescents and a 10 item YP-CORE (Young Person's CORE) has been derived by selection and simplification from the CORE-OM for ages 11-16 (Twigg et al., 2010). A 14 item related measure, LD-CORE for use with adults with mild to moderate learning difficulties (LD) has also been developed though this includes items not in the CORE-OM as the problems faced by people with LD are not the same as those without LD.

¹ Dr. Evans is a Trustee of CORE System Trust (CST), the not-for-profit company that holds the copyright on the CORE instruments. Like the other CST Trustees, Dr. Evans receives no income from the CORE-OM nor from CORE-IMS. This paper is a personal view, not the view of CST or CORE-IMS.

Domains Assessed

The design ensured some items of primarily intrapersonal focus and others primarily interpersonal and to cover well-being, problems/symptoms, functioning and risk.

An early challenge was to obtain a measure not too dominated by psychiatric diagnosis hence few items map to DSM or ICD though anxiety and depression are well covered. No self-report measure can directly measure unconscious functioning but many items fit with Freud's aims that a successful therapy replace neurotic misery with normal human unhappiness and his equally famous aim that a good therapeutic outcome enhances the ability to love and to work. Objections from CBT therapists were mainly that they preferred measures specific to presenting problems. For all modalities it has been important to explain that no measure should replace normal clinical information channels and the CORE-OM no more replaces specific measures in CBT than it replaces counter-transference in analytic work. Counsellors and humanistic therapists surprisingly proved the most rapid adopters perhaps reflecting that item wording which was kept in lay language.

Use and Procedures

All CORE-OM derived measures can be scored easily by hand with scoring and scoring boxes on the instruments. All measures have been formatted to be scanned for optical character reading (OCR) and most modern OCR systems should read scores without any illegal modification of the forms. All the CORE instruments are free for use on paper (www.coreims.co.uk/copyright.pdf). We have always given permission for the item text to be used in software for research but

only one company, CORE Information Management Systems (CORE-IMS, www.coreims.co.uk) has permission to put CORE instruments into software for non-research use.

CORE-IMS provide two computer solutions originally for CORE measures alone but they now support over 30 other measures, though sometimes with additional licence costs, and could be extended to cover any typical measure. Their systems are CORE-PC and CORE-Net, the former a standalone PC system and the latter a networked client-server system. CORE-Net uses a routine web browser to access a server, usually a CORE-IMS internet server, though dedicated local servers can be used. Costs depend on usage with CORE-PC available by annual license with a minimum license for 125 clients costing £250 (\$CA396) and additional clients charged from £2 (\$CA3) downwards as volume increases. CORE-Net is also provided by annual license with license costs averaging £3 (\$CA4.75) per client. Neither CST nor CORE-IMS dictate what services do with their data. The original CORE system handbook from 1998 had a chapter on appropriate informed consent to use identifiable data, congruent with European data protection law. CORE system advice has always been that data will be useful, and sometimes only legal, if the service/therapist have thought clearly about planned uses of the data so informed consent for use can be obtained. All data held centrally has always been scrupulously anonymised preventing central analysis from identifying services, therapists or clients. CORE-IMS and all the server companies used by CORE-Net are accredited to ISO/IEC27001, an international standard for information security.

CORE-IMS have supported session-by-session presentation of individual client data since 1999 and two shortened, 18 items forms (CORE-SF-A/B) were provided from the CORE-OM launch. The SF-A and SF-B have four well-being items in common and 14 other items that are different, minimising memory effects. Over the last decade, expected length of routine measures has fallen and two other short forms, the CORE-5 and CORE-10 are now available. In therapies appropriate to this, sessional use of the CORE-OM or shorter forms, supports score guided adaptive therapy as championed by Lambert (Lambert, 2010) and others.

Assessment and Treatment Planning

As for sessional use, our "bottom up" philosophy means that how the CORE-OM relates to assessment must be determined locally. Some services position it as an appraisal of the service and not part of the therapy, thus minimising the clients' use of the measure to communicate to their therapists and removing it from assessment. Others make it part of assessment and find that it provides a structure with broad coverage that often leads smoothly into discussion of risk and of particular problems.

Technical Support

CORE-IMS provide high quality support both for the software but more importantly to help services choose how best to use CORE (or other) scores. This support is widely used in the UK, the Netherlands, Norway and Denmark with CORE-IMS support appropriate to different needs in different countries developing as demand emerges. Extensive information about the entire CORE system is available at www.coreims.co.uk.

Psychometric Properties

Psychometric properties are sometimes reported as if they were as fixed properties of measures but are, of course, statistical parameters, i.e. empirical findings from samples generalised to populations. Translations will not show exactly the same psychometric properties as English UK samples (e.g. Chris Evans et al., 2002). So far, reliability and validity in different samples using the English version and translations have been good. Overall internal reliability has been excellent, in the range of .92 to .94. Internal reliability for domain scores varies more across samples but is always acceptable. Test-retest reliability is good but not excessive, as appropriate to a change measure: typical values are from .64 (for the risk domain) to .91 (overall score). Discrimination between clinical and non-clinical samples is always strong and sensitivity to change good. Correlations with other instruments show strong convergent validity and often slight evidence of within-domain correlations being higher than cross-domain correlations (Evans et al., 2002)².

Patient /Client/Clinician Feedback

An excellent collection of accounts of using the CORE-OM and other components of the CORE system (Gray, Penny & Mellor-Clark, John, 2007) is freely available at http://www.coreims.co.uk/site_downloads/CORE-A-Decade-of-Development.pdf. The CORE-OM emerged well from user feedback in the NHS compendium of approved mental health outcome measures (National Institute for Mental Health England, 2008) and a recent user-led review of measures (Crawford et al., 2011).

Unique Features of the Measure

The CORE-OM is not unique: a number of good general measures exist, though few are free to reproduce on paper as the CORE measures are. The empirical literature on general and specific measures shows them all to have high covariance with only small variance specific to diagnosis or problem area. In this convergent field, particular but not unique strengths of the CORE-OM are: its broad domain coverage with risk "flag" items; the existence of well tested short forms, adaptations for families, young people and people with learning difficulties; the mapping of six items to health economic valuation; and the existence of a standard translation protocol with 22 good translations.

Institutional Implementation

As the measure is copyleft there is no single register of uptake. However, over 500 services and over 5000 practitioners use CORE-IMS software. In the UK, services

using the CORE-OM range from primary care medical practitioners and psychological therapists, through a plethora of secondary care psychological therapy services including the Tavistock Clinic, to prisons and high secure hospitals. It is used in Clinical Psychology trainings in Norway and by large provider organisations in the Netherlands and Norway and international use is growing steadily. YP-CORE is one of the measures recommended by CORC (www.corc.uk.net) and the SCORE measures are being fostered by AFT (www.aft.org.uk).

Service comparison was a primary driver; Michael Barkham, Margison, et al. (2001) and Evans, Connell, Barkham, Marshall & Mellor-Clark (2003) are early examples of such work. This work has been extended to specific areas, e.g. university student counselling services (Janice Connell, Barkham & Mellor-Clark, 2007, 2008). Further examples can be found in the list of CORE related publications (see next page).

2 The main criticisms of the psychometric properties of the CORE-OM come from our decision to facilitate reporting of domain scores. Publications and presentations by the CORE team have always recommended using domain scores *only* where there might be a specific clinical or research interest in the domain and we never expected the domains to emerge as cross-sectional factors of variance. Initial analyses (Evans et al., 2002) using principal component analysis showed a large first factor, a second mainly involving the positively keyed items and a third involving mainly the risk items. Criticism that the domains do not fall out as neat factors recurs. We believed we had laid these issues to rest with a hierarchical factor analysis showing second-order general factor and first-order factors of the domains and positively and negatively keying methods factors and we noted that scale quality was satisfactory where the non-risk items are treated as a single scale and risk items as a second is satisfactory (Lyne, Barrett, Evans & Barkham, 2006). However, the fantasy that a 34 item measure could show a neat four factor structure in which well-being, problems, functioning and risk were factorially simple and distinct persists (e.g. Bedford et al., 2010). We are aware of no psychological theory, no psychotherapy theory nor any empirical data from any measure, that has ever suggested that this is the case. A recent development may help though. This work extracts a small set of items from measures to provide a direct translation of scores to QALY (Quality of Life Year) valuations used for health economic (HE) life valuation through time trading tests widely used for HE scales. This work, (Mavranzeouli et al., 2010) which the same team have applied to other problem specific measures with equal success, neatly inverts the usual complaint noting that the CORE-OM's complex factor structure and broad domain coverage gives a good set of anchor items for HE evaluation. Exploration of the CORE-OM in clinical samples led to selection of six items, five covering psychological state and one physical state, that provide HE evaluation. Perhaps this will finally dent the fantasy that domain scores would or could show a clean factor structure.

REFERENCES

An extensive list, currently of 123 publications many with abstracts and some with full text is at www.coreims.co.uk/Downloads_References.html and updated regularly.

- Barkham, Michael, Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., et al. (2001). Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69, 184-196.
- Bedford, A., Watson, R., Lyne, J., Tibbles, J., Davies, F., & Deary, I. J. (2010). Mokken scaling and principal components analyses of the CORE-OM in a large clinical sample. *Clinical Psychology & Psychotherapy*, 17, 51-62. doi:10.1002/cpp.649
- Brooks, M., & Davies, S. (2008). Pathways to participatory research in developing a tool to measure feelings. *British Journal of Learning Disabilities*, 36, 128-133. doi:10.1111/j.1468-3156.2007.00476.x
- Cahill, P., O'Reilly, K., Carr, A., Dooley, B., & Stratton, P. (2010). Validation of a 28-item version of the Systemic Clinical Outcome and Routine Evaluation in an Irish context: the SCORE-28. *Journal of Family Therapy*, 32, 210-231. doi:10.1111/j.1467-6427.2010.00506.x
- Connell, Janice, Barkham, M., & Mellor-Clark, J. (2007). CORE-OM mental health norms of students attending university counselling services benchmarked against an age-matched primary care sample. *British Journal of Guidance and Counselling*, 35(1), 41-57. doi:10.1080/03069880601106781
- Connell, Janice, Barkham, M., & Mellor-Clark, J. (2008). The effectiveness of UK student counselling services: an analysis using the CORE System. *British Journal of Guidance & Counselling*, 36, 1-18. doi:10.1080/03069880701715655
- Crawford, M. J., Robotham, D., Thana, L., Patterson, S., Weaver, T., Barber, R., Wykes, T., et al. (2011). Selecting outcome measures in mental health: the views of service users. *Journal of Mental Health*, 20, 336-346. doi:10.3109/09638237.2011.577114
- Evans, C., Connell, J., Barkham, M., Marshall, C., & Mellor-Clark, J. (2003). Practice-based evidence: Benchmarking NHS primary care counselling services at national and local levels. *Clinical Psychology and Psychotherapy*, 10(6), 374-388.
- Evans, Chris, Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180(JAN.), 51-60.
- Evans, Chris, Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health*, 9, 247-255.
- Gray, Penny, & Mellor-Clark, John. (2007). CORE: a decade of development. Rugby, England: CORE IMS. Retrieved from http://www.coreims.co.uk/site_downloads/CORE-A-Decade-of-Development.pdf
- Lambert, M. (2010). Prevention of treatment failure : the use of measuring, monitoring, and feedback in clinical practice (1st ed.). Washington D.C.: American Psychological Association.
- Leach, C., Lucock, M., Barkham, M., Stiles, W. B., Noble, R., & Iveson, S. (2006). Transforming between Beck Depression Inventory and CORE-OM scores in routine clinical practice. *British journal of clinical psychology*, 45, 153-166.
- Lyne, K. D., Barrett, P., Evans, C., & Barkham, M. (2006). Dimensions of variation on the CORE-OM. *British journal of clinical psychology*, 45, 185-203.
- Mavranezouli, I., Brazier, J. E., Young, T. A., & Barkham, M. (2010). Using Rasch analysis to form plausible health states amenable to valuation: the development of CORE-6D from a measure of common mental health problems (CORE-OM). *Quality of Life Research*, 20, 321-333. doi:10.1007/s11136-010-9768-4
- National Institute for Mental Health England. (2008). Mental Health Outcomes Compendium (p. 84). London: National Institute for Mental Health England. Retrieved from http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_093677.pdf
- Sinclair, A., Barkham, M., Evans, C., Connell, J., & Audin, K. (2005). Rationale and development of a general population well-being measure: Psychometric status of the GP-CORE in a student sample. *British journal of guidance and counselling*, 33, 153-173.
- Stratton, P., Bland, J., Jones, E., & Lask, J. (2010). Developing an indicator of family function and a practicable outcome measure for systemic family and couple therapy: The SCORE. *Journal of Family Therapy*, 32(3), 232-258.
- Twigg, E., Barkham, M., Bewick, B. M., Mulhern, B., Connell, J., & Cooper, M. (2010). The Young Person's CORE: Development of a brief outcome measure for young people. *Counselling and Psychotherapy Research: Linking research with practice*, 9, 160 - 168.
- Wright, F., Bewick, B. M., Barkham, M., House, A. O., & Hill, A. J. (2009). Co-occurrence of self-reported disordered eating and self-harm in UK university students. *British Journal of Clinical Psychology*, 48, 397-410. doi:10.1348/014466509X410343

The BASIS-24 Behavior and Symptom Identification Scale



Thomaskutty B. Idiculla, Ph.D.,
McLean Hospital,
Harvard Medical School

tidiculla@mclean.harvard.edu

Susan V. Eisen, Ph.D., Center for Health Quality, Outcomes & Economic Research (CHQOER), Bedford, MA, VA Hospital, Boston University School of Public Health

The Behavior and Symptom Identification Scale 24 (BASIS-24) is a 24 item patient self-report questionnaire designed to assess treatment outcomes by measuring symptoms and functional difficulties experienced by individuals seeking mental health services. The survey is intended for adults of ages 18 and older who present a broad range of symptoms and problems at all levels of care including inpatient, residential, partial and outpatient settings. The BASIS-24, which has been translated into five languages, can be used across a wide range of therapeutic treatment models. Scores are computed for the overall scale, as well as for six subscales that assess depression and functioning; interpersonal relationships; psychosis, substance abuse; emotional lability, and self-harm. Studies using the questionnaire have shown that its use is associated with an increase in patient satisfaction with care after the domain scores were discussed with the patient as part of developing a treatment plan. In this paper, the authors describe the measure, the procedures related to its use, its psychometric properties and the contexts in which it can be used.

Keywords: Behavior and Symptom Identification Scale; BASIS-24; treatment outcome; progress monitoring; psychotherapy

The Behavior and Symptom Identification Scale 24 (BASIS-24), copyrighted by McLean Hospital, is a 24 item patient self-report questionnaire designed to assess treatment outcomes by measuring symptoms and functional difficulties experienced by individuals seeking mental health services. The original tool, Behavior and Symptom Identification Scale 32 (BASIS-32) was developed in the early 1980s to meet the need for a brief but comprehensive mental health status measure that would be useful in assessing the outcomes of mental health treatment from the consumer's point of view. It is a measure of self-reported difficulty in the major symptom and functioning domains that lead to the need for mental health services (Eisen, Dill & Grob, 1994).

BASIS-24 sought to improve upon BASIS-32 by increasing applicability across diverse populations and improving the reliability and validity of the instrument. Revision of the instrument included (a) review of literature; (b) input from 75 researchers, administrators, clinical providers, and consumers; (c) readability analysis; (d) review of survey question design principles and methods; (e) meeting of the research team to review progress and make suggestions for the revision; (f) drafting of a revised instrument; (g) cognitive testing of the revised instrument; (h) analysis of cognitive test data; (i) further revisions of the instrument; (j) a second round of cognitive testing; (k) analysis of the second round of cognitive testing; and (l) further revisions and construction of the instrument for field-testing. The revised instrument was then tested on over 6,000 individuals who were receiving inpatient or outpatient treatment for mental health or substance abuse (Eisen et al., 2004a).

BASIS-24 is intended for adults, ages 18 and older, and is appropriate for individuals who present a broad spectrum of symptoms and problems at all levels of care including inpatient, residential, partial and outpatient. Furthermore, the instrument can be used across a wide range of therapeutic treatment models. It was designed to assess improvement over time regardless of type of mental health treatment. A BASIS-24 Adolescent Pilot version is also available. Both instruments can be used with multiple administrations throughout treatment and post-treatment follow-up (Eisen et al., 2004b). The adult version of BASIS-24 is available in English, Spanish, Portuguese, French, and Russian.

Domains Assessed

The BASIS-24 survey cuts across diagnoses, recognizing the wide range of symptoms and problems that occur across the diagnostic spectrum. BASIS-24 is designed to measure outcomes for a broad range of treatments and services encompassing many theoretical

orientations. Scores can be computed for the overall BASIS-24, as well as for six subscales: Depression and Functioning, Interpersonal Relationships, Psychosis, Substance Abuse, Emotional Lability, and Self-Harm.

The first of these domains, *Depression and Functioning*, seeks to assess daily/role functioning and depression and anxiety symptoms. The *Interpersonal Relationships* domain evaluates the client's perception of the quality of their interpersonal experiences with family and others. The *Psychosis* domain assesses four symptoms of psychotic disorders, such as hallucinations and delusions. The *Substance Abuse* domain seeks information regarding the client's urge to drink or use drugs as well as possible problems resulting from alcohol and/or drug use. The *Emotional Lability* domain, which measures what the Royal College of Psychiatrists describes as an "excessive emotional response to a minor stimulus," includes three items reflecting mood swings, racing thoughts and feeling short-tempered. Finally, the *Self-Harm* domain measures clients' thoughts about hurting themselves and/or ending their lives.

Use and Procedures

BASIS-24 is typically self-administered by the patient and takes between 5 and 15 minutes to complete. The 24 items are written at a fifth- to sixth-grade reading level, which maximizes the number of individuals who are able to complete the questionnaire by themselves. However, it can also be administered via computer, on the telephone, or in an interview. When the BASIS-24 is administered through a structured interview, a clinician, researcher, support staff member, or volunteer reads the items to the respondent and elicits ratings with the help of 8 x 11 laminated

"response cards" on which the rating scale is printed in large letters. Telephone interviews and mailed self-report questionnaires can be used at discharge or termination and at follow-up time points (Eisen et al., 2004a; McLean Hospital, 2006).

Among the 24 items, each has five ordered response options reporting either the level of difficulty experienced (no difficulty to extreme difficulty), or the frequency with which a symptom or problem has occurred (none of the time to all of the time). Respondents answer each question describing behavior and symptoms during the past week. For example, "During the past week, how much of the time did you feel sad or depressed?" BASIS-24 is administered at the beginning of a treatment episode, with repeat assessments obtained at desired intervals to assess change during or following treatment. The 24 questions are scored on a 5-point scale (from 0 to 4) and each subscale and overall mean scores also range from 0 to 4, with 0 being the lowest severity of symptoms and 4 being the highest severity of symptoms. The overall BASIS-24 score is a weighted sum that is computed by multiplying the rating for each question by its weight and totaling the weighted ratings for each question. The score for each subscale is a weighted sum that is computed by multiplying the rating for each question in the subscale by its weight and totaling the weighted ratings for all questions in the subscale (McLean Hospital, 2006).

In order to use BASIS-24, a single-site client must purchase an annual site license for \$US395. Organizations with multiple sites can purchase additional site licenses for \$US95 each. Included with the license fee are a BASIS Instruction Guide, a survey form for photocopying and the scoring procedures and algorithm for the instrument. Each site license allows for an

unlimited number of BASIS-24 administrations. Along with BASIS-24, clients may also purchase WebScore, an optional online scoring and reporting application. The cost of WebScore is based on the estimated number of surveys entered into the online application for each year (McLean Hospital, 2011).

Assessment and Treatment Planning

Evidence-based practice has been trending to include analysis of individual patient responses for real-time intervention and treatment planning in addition to the older models of aggregate benchmarking. Self-report measures are especially important in this new trend of measurement tools as an adjunct to treatment planning as they systematically inform providers about difficulties the patient may not otherwise express in other clinical measures (Newnham & Page, 2010). The use of a self-report tool also allows the patient to take a more active role in treatment planning. Studies using the BASIS-32 have shown an increase in patient satisfaction with care after the domain scores were discussed with the patient as part of developing the treatment plan. The patients specifically reported a greater feeling of involvement in care decisions and respect from the clinicians (Eisen, Dickey & Sederer, 2000). The use of self-report measures in treatment planning can be useful in more focused treatment as well as a better patient-clinician rapport.

Although the BASIS-24 is not designed to replace a comprehensive clinical evaluation, the tool documents the consumers' perspectives on the symptoms and problems that bring them to treatment. It is also simple to incorporate into a clinical evaluation process when already part of a Quality Improvement or outcome

assessment program, allowing it to fulfill both roles in the same administration (Eisen et al., 2004a). Using the same standardized tool for both individual planning and aggregate outcomes assessment also provides consistent metrics between the care objectives and its results. The constructs at the individual assessment level thus match those used to measure overall outcomes and assessment, leading to greater consistency.

The BASIS-24 can be used to identify primary and secondary problems from the individual consumer's perspective (Eisen et al., 2004a). Where BASIS-24 subscales overlap with diagnosis, there is usually a consistency between subscale scores and clinical diagnosis in that consumers diagnosed with depression or anxiety tend to report more difficulty with depression and anxiety than do consumers with other diagnoses (Eisen, Dill & Grob, 1994). Frequently, however, consumers tend to report high levels of difficulty in areas that do not correspond to their diagnosis as well. Problems in interpersonal relationships, managing day-to-day life, and depression often tend to be identified as more difficult than psychotic symptoms for consumers diagnosed with schizophrenia (Eisen et al., 2004a). In these cases, whereas a clinician may see psychosis as the main focus of treatment, the consumer may identify other priority areas for treatment. The BASIS-24 can thus highlight possible high levels of distress not directly symptomatic of primary diagnosis as well as areas in which the diagnosis has had a negative impact on day-to-day functioning. In addition, Eisen and Grob (1982) found that psychiatric outpatients in a rehabilitation program improved significantly in the areas they themselves had identified as goals for treatment, but

did not improve in areas they had not identified, as indicated by both clinician and patient reports. Patient perception is shown to be a predictor of outcomes, and should thus be addressed at the planning stage in order to maximize the effectiveness of treatment. As evidence-based practice becomes more integrative of all facets of care, it is important to use each metric in one's toolbox to its maximum potential.

Technical Support

The process to use BASIS-24 begins with preregistration on www.ebasis.org, where the client can create an account and agree to the terms of a general service agreement. Once payment and completed paperwork, including signed end-user license agreement, has been received, the account will be approved by an eBASIS staff member and clients will be able to begin using BASIS-24. Additionally, large volume users can utilize BASIS-24's optical scanning forms. Clients can complete the survey on these forms and the eBASIS staff will provide scanning services at a current charge of \$US1 per form so that data does not need to be manually entered.

Clients who wish to use BASIS-24 have four levels of service available to them: (a) BASIS-24 license; (b) license and access to WebScore; (c) license, WebScore, and Performance Measurement System Reporting; or (d) consultation. WebScore is an internet-based scoring and reporting tool for the BASIS-24 survey. It is an easy-to-use data entry and reporting application that lets users automatically score the BASIS-24 from a personal computer, download and print survey results, and maintain data for future analysis and reporting. Clients may try a free demo of WebScore for 30 days by signing up at

<https://secure.ebasis.org/basisdemo/login.php>.

WebScore provides both patient-level and aggregate population-level reporting capabilities, which can be sorted by time point, level of care, gender, or age. Reports can also be produced by patient admission or discharge date. Results can be downloaded into a CSV/Excel file that can easily be imported into SPSS, SAS or other statistical software applications. For those clients who choose to utilize WebScore for online scoring, eBASIS Systems ensures the highest standards of confidentiality and security, including compliance with all HIPAA guidelines and requirements. Data is stored in a secure server and individual cases are not identified in any report or aggregate results. All hard-copy patient information is stored in confidential, locked areas and paper surveys are shredded after 3 years, once all data is verified, cleaned, and backed up (McLean Hospital, 2011).

The Performance Measurement System offers custom reports, including change scores for a given quarter, control charts showing month-by-month outcomes, and comparison charts comparing the client's data with national benchmarks. Consultation is available regarding design of an outcome assessment system, data collection, data management, and reporting, and can be customized to meet the organization's specific needs.

Psychometric Properties

Confirmatory factor analysis conducted with BASIS-24 items confirmed the six factors described above under BASIS-24 domains. The Adjusted Goodness of Fit Index (0.81), root mean square error of approximation (0.08), standardized root mean squared residual (0.06), Comparative

Fit Index (0.95), and Non-Normed Fix Index (0.95) all indicate adequate to excellent fit (Eisen et al., 2004b). Internal consistency reliability (Cronbach's alpha) coefficients for the 6 domains ranged from 0.75 to 0.89 for inpatients and from 0.77 to 0.91 for outpatients (Eisen et al., 2004b). When broken down by race-White, African-American, and Latino-Cronbach's alpha coefficients exceeded 0.70 for all domains and for all race/ethnicity groups for both inpatients and outpatients, with one exception: for Latino inpatients, the alpha was 0.66 for the emotional lability domain (Eisen et al., 2006). When broken down by gender, the internal consistencies ranged from 0.73 to 0.89 for males and 0.77 to 0.89 for females (Idiculla, 2008). Test-retest reliability coefficients ranged from 0.81 to 0.96 for inpatients, and 0.89 to 0.96 for outpatients (Eisen et al., 2006).

For both inpatients and outpatients, correlations of the BASIS-24 domain and summary scores with the Mental Component Score of the Short Form (SF)-12 (Ware et al., 1996) ranged from 0.15 to 0.77, and correlations with global ratings of mental health ranged from 0.12 to 0.75. Correlations of the BASIS-24 scores with the Physical Component Score (PCS) of the SF-12 ranged from 0.01 to 0.15 for inpatients, and from 0.06 to 0.28 for outpatients (Eisen et al., 2004b). In a later study comparing racial/ethnic groups, correlations of the BASIS-24 summary score with other self-reported measures of mental health status (MCS, global mental health, and satisfaction with life) ranged from 0.59 to 0.82, for both inpatients and outpatients in each group. Additionally, correlations between the summary score with PCS were consistently lower, ranging from 0.07 to 0.45 (Eisen et al., 2006), indicating that, as expected, BASIS-24 is substantially cor-

related with other measures of mental health, but not with measures of physical health.

Institutional Implementation

BASIS-24 is currently in use in 5 countries in over 200 hospitals, mental health centers, community-based outpatient clinics, schools and managed care organizations. BASIS-24 was previously used for accreditation purposes by The Joint Commission and is approved by the Massachusetts Behavioral Health Partnership for use in clinical outcomes measurement.

For more information about

BASIS-24, please visit

<http://ebasis.org/>

or contact by mail at

McLean BASIS Plus/eBASIS,
MS-112,
McLean Hospital,
115 Mill Street, Belmont,
MA 02478-9106;

by phone at

(617) 855-2424;

or by email at

basisadmin@mclean.harvard.edu.

REFERENCES

- Eisen, S.V., Dickey, B., & Sederer, L.I. (2000). A self-report symptom and problem rating scale to increase inpatients' involvement in treatment. *Psychiatric Services*, 349-353.
- Eisen, S.V., Dill, D.L., & Grob, M.C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry*, 45(3), 242-247.
- Eisen, S.V., Gerena, M., Ranganathan, G., Esch, D., & Idiculla, T. (2006). Reliability and Validity of the BASIS-24 Mental Health Survey for Whites, African-Americans, and Latinos. *The Journal of Behavioral Health Services & Research*, 33, 304-323.
- Eisen, S.V., & Grob, M.C. (1982). Clients' rehabilitation goals and outcome. *Psychological Reports*, 50, 763-767.
- Eisen, S.V., Normand, S.L.T., Belanger, A.J., Gevorkian, S., & Irvin, E.A. (2004a). BASIS-32® and the Revised Behavior and Symptom Identification Scale BASIS-R. In M. Maruish (Ed.) *The Use of Psychological Testing for Treatment Planning and Outcome Assessment, Third Edition* (79-113). Mahwah, NJ: Lawrence Erlbaum.
- Eisen, S.V., Normand, S.L.T., Belanger, A.J., Spiro, A., & Esch, D. (2004b). The Revised Behavior and Symptom Identification Scale (BASIS-R): Reliability and Validity. *Medical Care*, 42, 1230-1241.
- Eisen, S.V., Youngman, D., Grob, M.C., & Dill, D.L. (1992). Alcohol, drugs and psychiatric disorders: A current view of hospitalized adolescents. *Journal of Adolescent Research*, 7(2), 250-265.
- Idiculla, T. (2008). Gender invariance of Behavior and Symptoms Identification Scale factor structure, *International Journal of Diversity* (in progress).
- McLean Hospital. (2006). BASIS-24 Instruction Guide. Belmont, MA: Author.
- McLean Hospital. (2011). McLean Hospital BASISplus/eBASIS. Retrieved from <http://ebasis.org/>
- Newham, E.A., & Page, A.C. (2010). Bridging the gap between best evidence and best practice in mental health. *Clinical Psychology Review*, 30(1), 127-142.
- Ware JE, Kosinski M, Keller S. A 12-item short-form health survey (SF-12): construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;24:220-233.

The Integra/ COMPASS Tracking Assessment System



Robert J. Lueger, Ph.D.,
Creighton University

This paper presents the Integra/COMPASS Tracking Assessment System, which can be used for the ongoing monitoring of patient response to mental health treatment. The system, which is available in four languages, includes 68 items, with 4 items devoted to assessing subjective well-being; 40 items assessing symptoms of anxiety, depression, obsessive-compulsive disorder, physical symptoms, adjustment problems, and post-traumatic stress; and 24 items that assess disabilities in daily living, social, work, intimacy, and development of self. Two shorter forms of the 68-item measure have also been developed. The system can be used to define a probable course of response to treatment which becomes the standard for defining whether the treatment is going as expected, which in turn can enable clinicians to make adjustments given evidence of the patient's response or non-response to intervention. This paper describes the Integra/COMPASS Tracking Assessment System and presents its psychometric properties, as well as the procedures related to its use in diverse professional contexts.

Keywords: Integra/COMPASS Tracking Assessment System; COMPASS; Integra; treatment outcome; progress monitoring; psychotherapy

The COMPASS Tracking Assessment System originated in 1991 as an outcome measure for assessing whether psychotherapy interventions produced measurable change. The first version of the COMPASS was developed using data and measures from the Northwestern Study of Long-term Psychotherapy that was funded by the U.S. National Institutes of Health and conducted by Kenneth Howard, Ph.D. of Northwestern University and David Orlinsky, Ph.D. of the University of Chicago among others. In 1992 Dr. Howard teamed with Peter Brill, M.D., founder of the Integra, Inc., a managed behavioral health care network that operated primarily on the U.S. eastern coast, to publish a slightly revised measure under the title, the *Integra Outpatient Tracking System (IOTA)*. Basic psychometric data were published in a manual (Howard et al., 1992), and data collection began with the Integra, Inc., service delivery system in 1992. By 1996 approximately 16,000 service participants had contributed a measurement on at least one occasion. The IOTA was renamed the COMPASS Tracking Assessment System when the corporate entity, COMPASS Information Services, was formed in 1993.

In 1997, the 84-item COMPASS for Primary Care (COMPASS-PC) was developed in collaboration with Bristol-Myers Squibb for assessment of patient response to medication and behavioral treatments in primary care settings. Within a year, however, Bristol-Myers Squibb reorganized its behavioral health division, and support for network-wide implementation was shelved. In 1999, the remnants of the COMPASS service delivery system were sold to a new management group using the older name of Integra, Inc., and the rights to the outcomes measurement system were included in the purchase. The outpatient tracking system continued to be used in the Integra, Inc., network under the title COMPASS(R) for at least another five years, and was converted to web-based delivery in 2000. A scan of the professional literature reveals few or no publications on the COMPASS-PC or the COMPASS(R) after 2003, but the Integra/COMPASS, which contains 68 items, continues to be used widely. It is available in English, Spanish, German, and Italian languages.

Domains Assessed

The Integra/COMPASS outpatient tracking system was developed from an *a priori* conceptual theory, the Phase Model, which characterizes change during psychotherapy as consisting of three distinct and sequential phases. The Phase Model of psychotherapeutic change (Howard, Lueger, et al. 1993) postulates that patients first

are *remoralized* to become hopeful, symptoms then go into remission, and finally problems in functioning are *rehabilitated* or more adaptive skills are learned. The Phase Model was introduced to provide a guide to *what changes* when in the course of a psychotherapeutic treatment. Empirical support for the Phase Model generally has been positive (Callahan et al., 2006; Hilsenroth et al., 2001; Lueger, 1995; Lueger, 2010; Stulz & Lutz, 2007). Approximately 63% of psychotherapy patients/clients show a Phase Model characteristic pattern of change during psychotherapy (Stulz & Lutz, 2007). For patients who respond positively during psychotherapy, about 42% show the pattern of remoralization preceding remission of symptoms, and remission of symptoms preceding rehabilitation of functioning (Lueger, 2010) as predicted by the Phase Model, and only 12-16% of those who manifest reliable change are not consistent with the predictions. Cumulative patient improvement curves indicate that approximately 50% of all patients who begin psychotherapy remoralize by session 6, 50% show symptoms improvement by session 12, and 50% show improvement in functioning by session 24-27 (Lueger, 2002).

The Integra/COMPASS measure consists of 68 items, with 4 items devoted to assessing Subjective Well-Being (Remoralization), forty items assessing symptoms of anxiety, depression, obsessive-compulsive disorder, physical symptoms, adjustment problems, and post-traumatic stress. The Life Functioning portion of the measure consists of 24 items that assess disabilities in daily living, social, work, intimacy, and self development. The Mental Health Index has served as the focal variable for most of the outcomes work related to the COMPASS. Intake

norms on the Integra/COMPASS were obtained from approximately 7,000 patients seeking psychotherapy in the first session of treatment. Two non-patient samples totaling approximately 700 adults were used to obtain comparison norms. The responses from first-session patients were normalized and converted to T-scores with a mean of 50 and a standard deviation of 10. Higher scores indicated greater behavioral/psychological health.

Short forms of the 68-item measure have been developed. A 12-item short form of the Integra/COMPASS measure was constructed and psychometrically validated to more economically track patient outcomes, although this brief form has been seldom used in clinical practice. Like most brief forms, the 12-item short form seems to measure a single construct. A 35-item version of the full 68-item Integra/COMPASS measure was developed by focusing on items that had demonstrated sensitivity to change in psychotherapy (Lueger, 2010).

Assessment and Treatment

Planning

A distinctive feature of the Integra/COMPASS system is its use of clinical characteristics (distress level, severity, chronicity, previous treatment history, expectation of improvement) to define a probable course of response to treatment. This expected course, or "Expected Treatment Response," becomes the standard for defining whether the treatment is going as well as, less than, or better than expected. This approach to outcomes standards takes into account the wide range of differences, response styles, and levels of change that characterize individuals experiencing psychotherapeutic interventions. For some patients, the goal may be to pre-

vent deterioration, for others the goal might be to manage chronic problems more effectively (but not to a symptom free state), and for still others, the goal may be a return to a normal or non-patient state. By using the known courses of treatment for similar patients, there is a normalized, empirical reference or benchmark for an individual patient in treatment.

Empirical studies with large data sets of former patients have identified clinical variables for predicting individual responses to treatment (Lutz et al., 1999; Lueger et al., 2001; Stulz & Lutz, 2007) uses random regression or hierarchical linear modeling, and Nearest Neighbor grouping techniques. Knowing the initial scores on the components of the phase model and the status on the clinical characteristics, the pattern of change can be represented as a slope of change (modeled with either linear or log-linear assumptions). Patients can be categorized or grouped as most probably belonging to a set of patients who had similar clinical characteristics prior to treatment. The patients of a group will have similar patterns or slopes of change. Most of the work to date has used the Mental Health Index (MHI) as the outcome variable to be modeled in the expected course. Confidence boundaries can be built around the expected course of response by identifying the percentage of individuals (top and bottom 25th percentiles) at a particular session of treatment. A score outside the confidence boundaries is either a better or a worse than expected outcome.

Early in its development, the potential of using Integra/COMPASS data as feedback during the treatment ("smart system") was recognized. Feedback about patient progress enables clinicians to make adjustments given evidence of the

patient's response or non-response. Feedback also can enhance the confidence of a therapist that the treatment is on track. Given that approximately 14% of completed treatments end as failures, the greatest payoff in using feedback seems to be that of preventing these failure treatment outcomes. Research (Lueger et al., 2001) has focused on predictors of treatment failure, and has identified several potential indicators using the overall measure of functioning, the MHI. When the self-reported overall functioning is very positive and the clinician-rated functioning is low, there is a high probability of later treatment failure. About two-thirds to three-quarters of patients who have two successive tracking measurements in the less-than-expected category later become failures. Non-changes on Phase Model components also predict treatment failure. Half of the patients who do not remoralize by the fourth session, and 80% who do not improve in symptoms by the twelfth session will not improve by the end of treatment.

Feedback to the therapist follows the phase model of change. Lacking evidence of improvement in remoralization, the therapist is encouraged to attend to instilling hope, building the relationship, normalizing the experience of the patient, and increasing the confidence of the patient in treatment. Lacking evidence of improvement in symptoms, the therapist is encouraged to refocus on attainable, specific, short-term treatment goals related to elements of the symptoms, or possibly to add adjunctive treatments (e.g., medication management) or to increase the frequency of the treatment sessions.

Use and procedures/Technical Support

The Integra/COMPASS measure was originally within the private domain and subject to costs per administration. Since the demise of the supporting organization, Integra Incorporated, the measure has been used within private mental health service clinics as part of treatment outcome assessment. The technical support originally made available through Integra Incorporated a decade ago no longer is available, and clinics generally have developed their own web-based technical support. The extent of the use of the original 68-item measure is unknown, but several behavioral health clinics in Minnesota—Center for Life Counseling, Midwest Center for Personal and Family Development—use the measure. The original measure, scales for scoring, and original norms, as well as the shorter 35-item scale are available from this author.

Psychometric Properties

For the Integra/COMPASS measure, internal consistency measures range from .79 for the four-item Subjective Well-Being subscale, to .95 for the Symptoms subscale, .93 for the Life Functioning subscale, and .87 for the global scale (Mental Health Index) using normalized scores on each of the three domains. For the 35-item short form of the measure, factor analyses reveal that this shorter form retains the components of the Phase Model, and has high intra-measure reliability ($\alpha = .94$).

Institutional implementation

The first ten years of using the COMPASS Tracking Assessment System included inevitable resistance to incorporating outcomes measures into a naturalistic treatment process. The demands of evidence of outcome as a condition of reimbursement brought by managed care processes have changed the nature of that resistance. Nonetheless, a premium is placed on the most time efficient method of obtaining that evidence. Thus shorter forms have emerged from the longer forms of most outcomes measures. Also, significant advances in electronic personal devices in the past ten years have provided multiple platforms (cell phones, smart phones, iPads, and computers) for completing these briefer forms. This has eliminated the need for paper management, and has enabled immediate scoring of the completed forms. These platforms also permit form completion away from the site of the service delivery. For intermittent sessions, weekly readings can be taken using abbreviated forms even when the patient does not come to the clinic for treatment. However, issues of confidentiality are raised in these contexts.

The science and technology of outpatient tracking and assessment have greatly advanced over the past 20 years. The Integra/COMPASS Tracking Assessment System was an early pioneer of these efforts to identify a suitable outcome measure and to develop rules of use that related to treatment goals. The clinically adjusted expected treatment response has been a relatively unique feature of the COMPASS system. The accepted use of tracking progress feedback has altered the way that psychotherapeutic interventions are delivered, and has increased the acceptance of benchmarked outcome standards.

REFERENCES

- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69, 2, 184-196.
- Callahan, J.L., Swift, J.K., & Hyman, M.T. (2006). Test of the phase model of psychotherapy in a training clinic. *Psychological Services*, 3, 129-136.
- Grissom, G.R., Lyon, J.S., & Lutz, W. (2002). Standing on the shoulders of a giant: Development of an outcome management system based on the dose model and phase model of psychotherapy. *Psychotherapy Research*, 12 (4), 397-412.
- Hilsenroth, M.J., Ackerman, S.J., & Blagys, M.D. (2001). Evaluating the phase model of change during short-term psychodynamic psychotherapy. *Psychotherapy Research*, 11 (1), 29-47.
- Howard, K.I., Brill, P., Lueger, R.J., & O'Mahoney, M. (1992, 1993, 1995). *The Integra Outpatient Tracking Assessment*. Radnor, PA: Integra, Inc.
- Howard, K.I., Lueger, R.J., Maling, M., & Martinovich, A. (1992, June). *A phase model of psychotherapy outcomes*. Paper presented at the annual meeting of the Society for Psychotherapy Research, Berkeley, CA.
- Kopta, S., & Lowrey, J. (2002). Psychometric evaluation of the Behavioral Health Questionnaire-20: A brief instrument for assessing global mental health and the three phases of psychotherapy. *Psychotherapy Research*, 12, 413-426.
- Lueger, R.J. (1995). A phase model of psychotherapy outcome. *Psychotherapeut*, 40, 267-278.
- Lueger, R.J. (2002, June). *The phase model of psychotherapy outcomes: Memorial to Ken Howard*. Paper presented at the international meeting of the Society for Psychotherapy Research, Santa Barbara, CA.
- Lueger, R.J. (2010, June). *Testing predictions of the Phase Model: a comparison of therapist and patient report*. Paper presented at the international meeting of the Society for Psychotherapy Research, Asilomar, CA.
- Lueger, R.J., Howard, K.I., Martinovich, Z., Lutz, W., Anderson, E.A., & Grissom, G. (2001). Assessing treatment progress of individual patients with expected treatment response models. *Journal of Consulting and Clinical Psychology*, 69, 150-158.
- Lutz, W., Martinovich, Z., & Howard, K.I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology*, 67, 571-577.
- Stulz, N., & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: The phase model revisited. *Journal of Clinical Psychology*, 63 (9), 817-833.

The Outcome Questionnaire-45



Michael J. Lambert, Ph.D.,
Brigham Young University

michael_lambert@byu.edu

The Outcome Questionnaire-45 (OQ-45) is a 45 item self-report scale which can be used to estimate client disturbance at the outset and over the course of treatment. The scale provides an index of mental health functioning for adults 18 years of age or older, which can be compared to inpatient, community mental health, outpatient, employee assistance program, college counseling center, and normal population data. Scores on the measure can also be referenced against expected treatment responses thus serving as a benchmark of treatment success on a session-by-session basis. Furthermore, the OQ, which is available in over 30 languages, provides cut-scores for reliable change and recovery as markers for gauging treatment success and possible termination of services. In this paper, the author presents the scale, including how it is administered, the domains it assesses and its psychometric properties, and describes how it can be used to improve treatment delivery in clinical practice.

Keywords: Outcome Questionnaire; OQ-45; treatment outcome; progress monitoring; psychotherapy

The Outcome Questionnaire-45 (OQ-45) is a 45 item self-report scale used to estimate client degree of disturbance at the outset and over the course of treatment. It provides an index of mental health functioning for adults 18 years of age or older. It was originally developed for use in managed care as a means of measuring the outcomes of treatment and enhancing them. A person who takes the measure is compared to inpatient, community mental health, outpatient, employee assistance program, college counseling center, and normal populations. Scores on the measure are referenced against expected treatment responses based on the progress of 12,000 treated individuals across the United States. These data provide a bench mark of success on a session-by-session basis in order to identify treatment non-responders, and clients at-risk for negative outcomes. It also provides cut-scores for reliable change and recovery as markers for gauging treatment success and possible termination of services. It has been translated into more than 30 languages other than English, including French. It requires reading ability at the 6th grade level. Copyright to the measure is held by OQ Measures, LLC; 2171 Lake Street, Salt Lake City, UT 84106; www.oqmeasures.com; email: office@oqmeasures.com.

Domains Assessed

Because almost all adults who enter treatment experience symptoms of anxiety and depression, half of the items of the OQ-45 measure core aspects of these disorders, or what may be called symptomatic distress or subjective discomfort. Because a satisfactory quality of life and the wellbeing depends on positive interpersonal functioning, a quarter of the items assess disturbance in interpersonal relationships with intimate others. The final quarter of the items assess functioning in social roles such as work, school, homemaking and leisure activities. Effective functioning in social roles has important consequences for society as well as individuals. Nine items are worded in a positive manner in an attempt to tap into wellbeing in addition to psychopathology. These items are a part of the other three domains, as are substance abuse and suicide screening items. The clinician report highlights substance abuse, suicidal ideation, and anger when scores reach a critical level. The OQ-45 can be used regardless of the type of psychotherapy, mode of psychotherapy, or medication intervention. It is atheoretical in nature and serves as a mental health vital sign or lab test to be used by clinicians to manage illness by quantifying the patient's current mental health functioning. There is also a brief version of the OQ-45, the OQ-30 which does not include subscales.

Use and Procedures

Ideally, the OQ-45 is administered online, via handheld devices, or personal computer (it can be administered and scored via hard copy as well). It takes about 5-10 minutes of patient time to rate all of the 45 items, typically prior to the treatment session. Each item is answered on a five point scale according to the patient's recollection of the preceding week—on a scale from "almost always" to "never". This allows the clinician to get a quick overview of functioning that would take an excessive amount of time if based on a clinical interview. Software (OQ-Analyst) scores the measure, graphs the results in relation to earlier administrations, and in relation to normative functioning and expected treatment response based on other individuals who have the same initial level of disturbance. All this is accomplished and available on the therapist's computer in about one second after completion of the 45th item.

It takes therapists about 18 seconds to access the client's report on his or her computer or a clinician report and client report can be printed and delivered to individuals as a hard copy. Cumulative data from multiple administration of the measure can be housed on a self-supported server (or personal computer) or through OQMeasures. The OQ-45 is part of a larger Outcome Measurement System—the OQ-Analyst, which includes measures of child functioning, the Brief Psychiatric Rating Scale, and the Assessment for Signal Clients, a clinical support tool used to guide problem solving with failing cases. Users typically prefer to keep the data collected in house, but data stored outside a clinic is encrypted and HIPPA compliant methods are used to protect confidentiality of the data. As noted above, automated methods of administrating,

scoring, and creating clinical reports make the work load for clinicians and support staff minimal. If a therapist, instead of a third party such as a receptionist is in charge of handing out a handheld device, this takes an additional 30 to 60 seconds of time.

The OQ-Analyst software system is available for \$US200 per clinician per year. This is based on an average of 200 clients per year per clinician, at \$1 per client per year. The cost includes unlimited administrations for each client along with scoring, alerts, and progress profiling. A fully hosted Web-based system is available using a prorated cost model for organizations that serve more than 1,000 patients per year and employ more than 50 clinicians. The initial start-up costs for the hosted system average about \$3 per patient per year, which includes software and hardware, and yearly costs thereafter are under \$1 per patient per year. Both cost models deliver an unlimited number of administrations per patient providing an incentive to repeatedly track patients at no additional cost (i.e., there are no per-administration charges).

Assessment and Treatment

Planning

Although the OQ-45 is designed to track changes in mental health functioning over time, scores on the subscales can be used to determine which areas of functioning are most problematic for the patient and these can be imbedded in a treatment plan and tracked over time. Patient progress graphs can be cut and pasted into clinical records. In addition, OQ-reports estimate the number of sessions needed for a patient to return to a state of normal functioning or at least achieve reliable change. A major advantage of the OQ-45 is underlying valid algorithms that

predict treatment failure and provide alerts to clinicians if the patient is predicted to have a negative treatment outcome. Because clinicians are confident in their ability to recognize and predict treatment failure, but fail to do so, a psychological test that can perform this task is an invaluable addition to routine decision making and care. Between 85-100% of treatment failures can be identified by the OQ-45 before they leave treatment and often within 3-5 sessions after entering treatment. When clinicians are alerted to potential treatment failure they tend to retain patients in treatment longer. Deterioration rates are reduced by 1/3 to 2/3, while improvement rates double for these difficult patients. Tracking treatment response and feedback to clinicians reduces the number of sessions used by clients who never go off-track, and improves their final outcome.

The OQ-Analyst can be set up to provide clinical information to interdisciplinary teams who are also working with each specific patient. The clinical administrator enters all providers within a system of care, a primary, and secondary providers are then selected from the list. Thus all members of a team who are given access by administrators through a password, including clinical managers and supervisors, can be given access to progress reports and alert notifications for specific patients who are being treated by team members. This shared information allows for all the team to be aware of no, or negative treatment response and settle on coordinated efforts to turn the course of treatment in a positive direction. The OQ-Analyst also produces reports, summed across all patients or patient subgroups (such as substance-abuse patients, males vs. females, program A versus program B) for the purpose of comparing

benchmarks across similar services or for particular clinicians.

Technical Support

Technical support is available for \$150 per year. Technical manuals for the instruments and user guides on implementation of the OQ-Analyst are included at no additional charge and are regularly updated as new research supports modifications. NREPP rated the degree to which the OQ-Analyst is ready for dissemination at 3.9 on a 4-point scale. Users choose to host the OQ-Analyst themselves or have OQMeasures host it. In order to use the OQ-Analyst IT, support is sometimes required. And on occasion clinical questions arise. This support comes through www.oqmeasures.com or office@oqmeasures.com or toll free at 1-888-647-2673 or specifically from individuals who are in contact with users: Individuals include: Tameisha Hastings, Marketing and Sales, Tameisha.Hastings@OQMeasures.com; Sue A. Jenkins, Executive Officer, Sue.Jenkins@OQMeasures.com, Amy, IT Manager, Amy@OQMeasures.com. Clinical inquires and research questions are directed to: Michael J. Lambert, Ph.D., Mike.Lambert@OQMeasures.com; Gary M. Burlingame, Ph.D., Gary.Burlingame@OQMeasures.com

Training can be provided upon request and ranges from trainings lasting from two hours to all day workshops. Training can be accomplished on site or via phone conference with internet connection. Developers of the OQ-Analyst & Lanark Systems work closely with OQMeasures and have made modifications of the software at the request of large

users. This is very helpful to insure that the system becomes adapted to the emerging needs of agencies or if they want OQ-Measures to host their system.

Psychometric Properties

Since its development in the early 1990's dozens of studies have been published on the psychometric properties of the OQ-45. It has high internal consistency (.90), test re-test reliability (.84 over 3-weeks), and concurrent validity with scales such as the Symptom Checklist-90 and BDI, with coefficients hovering in the mid .80s. Factor analytic studies support the presence of an overall distress factor with three subordinate factors consistent with the subscales. Most items, the subscales, and the Total Score are sensitive to the effects of interventions while remaining stable in untreated individuals. This is the most important psychometric characteristic of the OQ-45 since it is used to monitor change in patients in treatment. The items in the OQ-45 have been examined over time in both patients and individuals who are disturbed, but not in psychological treatments or using psychoactive medications.

Patient/Client/Clinician Feedback

The strength of the OQ-45 is the extensive published evidence on the degree to which providing feedback to clinicians and patients based on the OQ-45 alert system maximizes patient outcome and reduces treatment failure. Seven RCT's have been published and two more have been completed showing that the feedback and problem-solving tools delivered to therapists work in a variety of routine care settings from university counseling centers,

outpatient substance abuse, to inpatient eating disorder treatment. The OQ-45 has been judged by the National Registry of Evidence-based Programs & Practices in the United States (NREPP) as an evidence-based practice based on the weight of evidence derived from experimental studies in routine care settings.

Qualitative research has been completed indicating that clients do not mind taking the OQ-45 on a weekly basis. This is particularly true if staff have a positive opinion about the measure, the patient is provided with a rationale for taking the measure (such as it is like monitoring blood pressure in order to manage it, or, it is a way of finding out how you are doing), and clinicians provide some feedback, indicating they are aware that the patients has completed it. Patients are very pleased with the clinical reports that they receive (if clinicians choose to share them). Providers are less positive initially than clients, especially if use of the measure is forced on them by administrators. After a time even the most resistant clinicians come to enjoy the feedback graphs of patient progress and find ways to use the information provided in treatment planning and risk assessment. From the clinicians point of view there is no efficient way to gather information about patient functioning across the range of factors that make up their mental health. Thus they come to see it as a quick check on areas of functioning that may not be the focus of a particular treatment session.

Institutional Implementation

There are hundreds of clinics, institutions, and individuals across North America and the world who are licensed users of the OQ-45 and related instruments for child assessment contained within the OQ-Analyst. It is widely used and clearly the dominant outcome instrument in training clinics serving clinical psychology graduate trainees in the USA and Australia. Use in Training clinics allows for evaluation of patient progress as well as trainee performance over time.

REFERENCES

- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. M., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying patients at risk for treatment failure. *Journal of Clinical Psychology: In Session*, 61(2), 155-163.
- Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, & feedback in clinical practice*. Washington, DC: APA Press.
- Lambert, M. J., Hansen, N. B., & Harmon, S. C. (2010). The OQ-45 system: Development and practical applications in health care settings. In M. Barkham, G. Hardy, & J. Mellor-Clark (Eds.) *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 141-154). New York: Wiley-Blackwell.
- Lambert, M. J., Kahler, M., Harmon, C., Burlingame, G. M., & Shimokawa, K. (2011). *Administration & scoring manual for the Outcome Questionnaire-45.2*. Salt Lake City: OQMeasures.
- Okiishi, J., Lambert, M. J., Eggett, D., Nielsen, S. L., Dayton, D. D., & Vermeersch, D. A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients' psychotherapy outcome. *Journal of Clinical Psychology*, 62, 1157-1172.
- Shimokawa, K., Lambert, M. J., & Smart, D. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic & mega-analytic review of a psychotherapy quality assurance program. *Journal of Consulting & Clinical Psychology*, 78, 298-311.
- Vermeersch, D. A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome Questionnaire: Item sensitivity to changes in counseling center clients. *Journal of Counseling Psychology*, 51, 38-49.

The Outcome Rating Scale (ORS) and the Session Rating Scale (SRS)



Scott D. Miller, Ph.D.,
International Center for Clinical Excellence,
Chicago

scottdmiller@talkingcure.com

Susanne Bargmann, International Center
for Clinical Excellence, Chicago

The Outcome and Session Rating Scales are brief measures which can be used to track client functioning and the quality of the therapeutic alliance over the course of psychotherapy. Versions of the scales are available for adults, children, adolescents and groups in 18 different languages. The Outcome Rating Scale (ORS) is designed to assess the individual, interpersonal, and social functioning of the client, whereas the Session Rating Scale (SRS) assesses the quality of the relational bond between the client and therapist, the degree of agreement between the client and clinician regarding treatment goals, and their agreement regarding the methods and approach employed in care. The tools do not require that practitioners adhere to a particular model or approach and administering and scoring the measures is simple and straightforward. In this paper, the authors present the two measures, discuss the domains they assess, and describe how they can be used in routine clinical practice to aid in service plan development. In addition, the authors discuss the psychometric properties of the scales and describe the resources available to clinicians who wish to use them.

Keywords: Outcome Rating Scale; ORS; Session Rating Scale; SRS; treatment outcome; progress monitoring; psychotherapy

The Outcome and Session Rating Scales (ORS and SRS) are brief measures for tracking client functioning and the quality of the therapeutic alliance. Each instrument takes less than a minute for consumers to complete and for clinicians to score and interpret. Both scales were developed in clinical settings where longer, research-oriented measures had been in use and deemed impractical for routine use. Versions of the ORS and SRS are available for adults, children, adolescents and groups in 18 different languages, including French. Individual clinicians may download the scales free-of-charge after registering online at: <http://www.scottdmiller.com/?q=node/6>. A significant and growing body of research shows the scales to be valid, reliable, and feasible for assessing progress and the alliance across a wide range of consumers and presenting concerns.

Domains Assessed

The ORS is designed to assess the individual, interpersonal, and social functioning of the consumer. On the other hand, the SRS assesses three elements of the alliance, including: (1) the quality of the relational bond; (2) the degree of agreement between consumer and clinician regarding goals; and (3) consumer and clinician agreement regarding the methods and approach employed in care. The tools neither assume nor require that practitioners adhere to a particular model or approach. Instead, clinicians from any background or discipline may solicit feedback from consumers regarding the working relationship and outcome of care and use the resulting information to inform and tailor service delivery. Routinely monitoring of progress and the quality of the relationship is not only consistent with but also operationalizes the American Psychological Association's definition of evidence-based practice, which includes, "the integration of the best available research... and monitoring of patient progress (and of changes in the patient's circumstances—e.g., job loss, major illness) that may suggest the need to adjust the treatment... e.g., problems in the therapeutic relationship or in the implementation of the goals of the treatment" (APA, 2006, p. 273, 276-277).

Use and Procedures

Administering and scoring the measures is simple and straightforward. The ORS is given at the beginning of the session. The scale asks consumers of therapeutic services to think back over the prior week (or since the last visit) and place a hash mark (or "x") on four different lines, each representing a different area of functioning (e.g., individual, interpersonal, social, and overall wellbeing). The

SRS, by contrast, is completed at the end of each visit. Here again, the consumer places a hash mark on four different lines, each corresponding to a different and important quality of the therapeutic alliance (e.g., relationship, goals and tasks, approach and method, and overall). On both measures, the lines are ten centimeters in length. Scoring is a simple matter of determining the distance in centimeters (to the nearest millimeter) between the left pole and the client's hash mark on each individual item and then adding the four numbers together to obtain the total.

In addition to hand scoring, several computer-based applications are also available which can simplify and expedite the process of administering, scoring, and aggregating data from the ORS and SRS. As just one example, consider the web-based application, www.fit-outcomes.com. Briefly, the system organizes treatment outcome and therapeutic alliance data, and compares the scores to the expected treatment response (ETR) of the client. Importantly, the client and therapist receive feedback in real time, indicating whether treatment is on or off track. Additionally, the system aggregates outcome and alliance data across episodes of care, thereby providing clinicians and agencies with an overall measure of effectiveness as well as the ability to compare the outcomes of individual clinicians and programs. With regard to privacy and security, all data entered into [fit-outcomes.com](http://www.fit-outcomes.com) is first anonymized and then encrypted according to current international standards.

Assessment and Treatment

Planning

Soliciting clinically meaningful feedback requires more than administering two scales, the ORS and SRS or otherwise. Clinicians must work at creating an atmosphere where consumers feel free to rate their experience of the process and outcome of services: (1) without fear of retribution; and (2) with a hope of having an impact on the nature and quality of services delivered. Beyond displaying an attitude of openness and receptivity, creating a "culture of feedback" involves taking time to introduce the measures in a thoughtful and thorough manner. Providing a rationale for using the tools is critical, as is including a description of how the feedback will be used to guide service delivery (e.g., enabling the therapist to catch and repair alliance breaches, prevent dropout, correct deviations from optimal treatment experiences, etc). With regard to interpreting the ORS, low scores correspond to a poor sense of well-being (or high level of distress). Note that the average ORS intake score in outpatient mental health settings is between 18 and 19. Over time, whatever the initial score, the number should increase in response to services offered. A lack of movement, deterioration, or seemingly random pattern of scores is cause for concern and should be discussed with the client at the time of service delivery. Between 25-33% of people completing the measure will fall above a total score of 25 at intake—a number known as the cutoff, or the dividing line between a clinical and non-clinical population (Miller & Duncan, 2000, 2004). The most common reason for such a score is that the consumer has been mandated into treatment. Another is that the person desires help for

a very specific problem—one that does not impact the overall quality of life or functioning, but is troubling nonetheless. Less frequent causes for a high initial ORS include: (1) high functioning people who want therapy for growth, self-actualization, and optimizing performance; and (2) people who may have difficulties reading and writing or who have not understood the meaning or purpose of the tool. With regard to the latter, it should be noted that a validated oral version of the ORS is available and can be administered. Research and experience document that consumers scoring above 25 at intake are at a heightened risk for deterioration. Therefore, care should be taken to clarify the wishes of the person in treatment. In order to maintain engagement, the best approach is a cautious one. In particular, using the least invasive and intensive methods needed to resolve the problem at hand.

With regard to interpreting the SRS, research to date shows that the majority of clients score relatively high. Thus, the cutoff on the measure is 36. It is important to keep in mind that a high score (36+) does not necessarily confirm the presence of a strong alliance. The best response to a high score is thanking the consumer and remaining open to the possibility of feedback in the future. Scores that fall at or below 36 are considered "cause for concern" and should be discussed prior to ending the visit. Single-point decreases in SRS scores from session to session have also been found to be associated with poorer outcomes at termination—even when the total score consistently falls above 36—and should therefore be addressed in the session

(Miller, Hubble & Duncan, 2007). Interestingly, there is growing evidence that the process of responding to a client's negative feedback, even about an aspect of therapy that may seem relatively trivial, can contribute to the strength of the therapeutic alliance and set in place a strong foundation for future work. There is also evidence that the most effective therapists elicit more negative feedback from their clients. Whatever the circumstance, openness and transparency are central to successfully eliciting meaningful feedback on the SRS.

Technical Support

An international, online community is available to support the use of the scales for informing, evaluating, and improving the quality of behavioral healthcare. Membership in the International Center for Clinical Excellence (ICCE) is free-of-charge, open to clinicians from all disciplines and approaches, and no selling or promotion of products or particular treatment approaches is allowed. The site features hundreds of discussion groups, articles, and how-to videos in many different languages. Members also have access to the "Get Answers" feature to obtain specific help quickly from community members. Certified trainers and associates are available for consultation and training. To register, go to: www.centerforclinicalexcellence.com.

A series of six manuals are available that cover the most important information for practitioners and agencies implementing the ORS and SRS are available (International Center for Clinical Excellence FIT

Manuals Development Team, 2011a,b, c, d,e,f [<http://www.scottdmiller.com/?q=node/5>]). The manuals are written in clear, practical, step-by-step, and easy-to-understand language and cover:

- (1) the empirical foundation;
- (2) basics of administration, scoring, and interpretation;
- (3) use of the measures in supervision;
- (4) aggregation and interpretation of data generated by the ORS and SRS;
- (5) application of the ORS and SRS with special populations; and
- (6) implementing the measures in agencies and systems of care.

As mentioned previously, several computer and web-based applications are available for administering, scoring, interpreting, and aggregating data from the ORS and SRS. The most current information about such applications can be found online at: <http://www.scottdmiller.com/?q=node/6>.

Psychometric Properties

The ORS has been shown to be sensitive to change among those receiving behavioral health services. Numerous studies have documented concurrent, discriminative, criterion-related, and predictive validity, test-retest reliability, and internal-consistency reliability for the ORS (e.g., Anker, Duncan & Sparks, 2009; Bringhurst, Watson, Miller & Duncan, 2006; Campbell & Hemsley, 2009; Duncan, Miller, Reynolds, Brown & Johnson, 2003; Duncan, Sparks, Miller, Bohanske & Claud, 2006; Miller, Duncan, Brown, Sparks & Claud, 2003;

Reese, Norsworthy & Rowlands, 2009). The SRS has been shown to assess the qualities of the alliance as first defined by Bordin (1976). Numerous studies have documented the concurrent validity, test-retest reliability, and internal consistency of the SRS (e.g., Duncan et al. 2003, Miller, Duncan, Brown et al. 2003). Several randomized clinical trials have documented the significant impact that both measures have on the outcome of and retention in treatment (e.g., Anker et al., 2009; Miller et al., 2006; Reese et al., 2009).

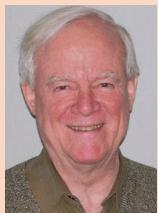
Institutional Implementation

Worldwide, there are currently 30,000+ registered individual practitioners, and 100's of licensed agencies and treatment settings using the scales. Since 2009, the membership of the International Center of Clinical Excellence (ICCE) has grown exponentially. The ICCE community is where most users receive training and support in the use of the measures. Each year, the ICCE conducts two intensive training events: (1) the "Advanced Intensive"; and (2) the "Training of Trainers" course. Attendance at both trainings, submission of a sample training video, and passing the "core competency" exam enable participants to become ICCE Certified Trainers. Currently, the ICCE has "Certified Trainers" available for consultation in the USA, Canada, Australia, New Zealand, Western and Eastern Europe.

REFERENCES

- Anker, M., Duncan, B., Sparks, J. (2009). Using client feedback to improve couple therapy outcomes: an RCT in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, 77, 693-704.
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, 61(4), 271-285.
- Bordin, E.S. (1976). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy*, 16, 252-260.
- Bringhurst, D. L., Watson, C. S., Miller, S. D., & Duncan, B. L. (2006). The reliability and validity of the outcome rating scale: A replication study of a brief clinical measure. *Journal of Brief Therapy*, 5(1), 23-29.
- Campbell, A., & Hemsley, S. (2009). Outcome rating scale and session rating scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist*, 13, 1-9.
- Duncan, B. L., Miller, S. D., Sparks, J. A., Reynolds, L. R., Brown, J., Johnson, L. D. (2003). The session rating scale: Preliminary psychometric properties of a "working alliance" inventory. *Journal of Brief Therapy*, 3(1), 3-11.
- Duncan, B. L., Sparks, J. S., Miller, S. D., Bohanske, R., Claud, D. (2006). Giving youth a voice: A preliminary study of the reliability and validity of a brief outcome measure for children, adolescents, and caretakers. *Journal of Brief Therapy*, 5, 71-87.
- International Center for Clinical Excellence FIT Manuals Development Team. (2011a). Manual 1: What works in therapy: A primer. Chicago, IL: ICCE Press.
- International Center for Clinical Excellence FIT Manuals Development Team. (2011b). Manual 2: Feedback Informed Clinical Work: The Basics. Chicago, IL: ICCE Press.
- International Center for Clinical Excellence FIT Manuals Development Team. (2011c). Manual 3: Feedback Informed Supervision. Chicago, IL: ICCE Press.
- International Center for Clinical Excellence FIT Manuals Development Team. (2011d). Manual 4: Documenting Change: A Primer on Measurement, Analysis, and Reporting. Chicago, IL: ICCE Press.
- International Center for Clinical Excellence FIT Manuals Development Team. (2011e). Manual 5: Feedback Informed Clinical work; Advanced Applications. Chicago, IL: ICCE Press.
- International Center for Clinical Excellence FIT Manuals Development Team. (2011f). Manual 6: Implementing Feedback-Informed Work in Agencies and Systems of Care. Chicago, IL: ICCE Press.
- Miller, S. D., & Duncan, B. L. (2000, 2004). The Outcome and Session Rating Scales: Administration and Scoring Manual. Chicago, IL: ITC.
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, 2(2), 91-100.
- Miller, S.D., Duncan, B.L., Sorrell, R., Brown, G.S., & Chalk, M.B. (2006). Using outcome to inform therapy practice. *Journal of Brief Therapy*, 5(1), 5-22.
- Miller, S.D., Hubble, M.A., & Duncan, B.L. (2007). Supershinks: Learning from the Field's Most Effective Practitioners. *Psychotherapy Networker*, 31(6), 26-35, 56.
- Reese, R.J., Norsworthy, L.A., & Rowlands, S.R. (2009). Does a continuous feedback system improve psychotherapy outcome. *Psychotherapy: Theory, Research, Practice, Training*, 46, 418-431.

The Polaris-MH



Grant Grissom, Ph.D.,
Polaris Health Directions

grant@polarishealth.com

Tina Harralson, Ph.D., Polaris Health Directions
Jesse Nankin, MA, Polaris Health Directions

Polaris-MH is a Web-based system designed to plan, manage, and improve adult outpatient mental health treatment. It is the product of more than a decade of research and development by internationally respected researchers working in collaboration with clinicians, patients, IT professionals, utilization review professionals, provider system administrators, and managed care executives (3,12). Polaris-MH was developed with funding from the National Institutes of Health. It combines a strong scientific foundation with state-of-the-art technology to help providers and health care organizations improve the quality and cost-effectiveness of their services, and document treatment outcomes for payers and accreditation organizations. It is designed for multiple administrations during treatment and provides a uniquely comprehensive assessment of clinical problems, patient strengths, and progress. It addresses the needs of various stakeholders: patients, clinicians, clinical managers/administrators, case managers/utilization reviewers and payers. For clinicians and case managers, it provides an evidence-based answer to the question, "Is this treatment working for this patient?" For administrators, quality improvement professionals and payers, the system provides multivariate severity adjustment of program outcomes for identification of best practices, and "apples-to-apples" evaluation of program effectiveness, controlling for initial severity and other patient characteristics.

The system can be used with adults 18 years of age or older in outpatient mental health treatment, including those with co-occurring substance use disorders. Aside from English, it is available in a number of languages, including French. Literacy at the 6th grade reading level is required.

Domains Assessed

Polaris-MH is grounded in basic research on psychotherapy process and outcome: Phase, Dose-Response and Expected Treatment Response (ETR) models. *The Dose-Response Model* (4) of psychotherapeutic impact describes a positive relationship between therapeutic dose and rate of clinical improvement, and a pattern of relatively rapid early improvement with more and more sessions needed to achieve incremental improvement later in treatment (a pattern of diminishing returns). *The Phase Model* (5) extended and interpreted the dose-response model by proposing three progressive sequential phases of the psychotherapeutic recovery process: (a) remoralization—the enhancement of well-being; (b) remediation—the achievement of symptomatic relief; and (c) rehabilitation—the reduction of troublesome, maladaptive behaviors that interfere with life functioning. The phase model suggests that the decelerating curve of improvement for a patient

Polaris-MH is a Web-based system designed to plan, manage, and improve adult outpatient mental health treatments. The system, which is available in a number of languages, can be used with adults 18 years of age or older in outpatient mental health settings, including those with co-occurring substance use disorders. The system measures domains, which are pan-theoretical, that correspond to the three phases of the therapeutic progress established through Phase Theory: Subjective Well-Being (remoralization), Symptoms (remediation) and Functional Disability (rehabilitation). It includes five major components: (1) a web-based patient self-report assessment (initial and update assessments); (2) real-time reports; (3) a data management module that allows clinical staff to securely log in and access data and reports; (4) an aggregate reporting module that provides online access to customizable aggregate reports; and (5) documentation and training materials. In this paper, the authors describe the measure, the procedures related to its use, its psychometric properties and the clinical contexts in which it can be used.

Keywords: Polaris-MH; Polaris; treatment outcome; progress monitoring; psychotherapy

can be attributed to the increasing difficulty of treatment goals as they change (e.g., from symptom remediation to improved functioning) over the course of treatment. *The Expected Treatment Response (ETR) Model* (6,7,8) uses pretreatment clinical characteristics (e.g., severity, chronicity, previous treatment, treatment expectation) to predict the patient's expected response during the course of treatment. Using individualized growth curve analysis for a large sample of outpatients in psychotherapy, a single patient's course of treatment can be predicted as soon as his or her intake information is available. Ongoing therapeutic effectiveness can be assessed for a single patient by tracking the patient's actual progress in comparison to his or her expected progress.

The Polaris-MH measurement domains, which are pan-theoretical, correspond to the three phases of the therapeutic progress established through Phase Theory: Subjective Well-Being (remoralization), Symptoms (remediation) and Functional Disability (rehabilitation). These domains are readily accepted by clinicians as being central to clinical decisions and outcomes assessment, irrespective of the therapeutic model used. These domains provide the framework for constructive dialogue between clinicians and utilization review staff. Subjective Well-Being is a single scale. The Symptoms scale is a composite of seven subscale scores; each subscale corresponds to a disorder commonly treated in outpatient settings (Depression, Anxiety, PTSD, Panic, OCD, Phobia, Somatization). Functional Disability is a composite of three subscales (Social, Vocational, Personal). A measure of overall clinical status, Behavioral Health Status (BHS), is a composite of the three core measures. Additional screens and

scales extend clinical and prognostic usefulness. In keeping with the basic design principle, the scales have pan-theoretical utility and assess alcohol and drug severity and resilience. Screens include medical health problems, use of psychoactive medications, psychosis and bipolar disorder. Strengths (resilience, meaning or purpose in life) and items relating to treatment motivation, treatment satisfaction and the therapeutic bond are also assessed. Customization for the Canadian Forces included addition of the PCL-C (PTSD Checklist, Civilian Version) developed by the U.S. Department of Veterans Affairs (10).

Use and Procedures

Polaris-MH consists of five major components: (1) Web-based patient self-report assessments (initial and update assessments); (2) Real-time reports; (3) Data management module that allows clinical staff to securely log in and access data and reports; (4) Aggregate reporting module that provides online access to customizable aggregate reports; and (5) Documentation and training materials. Polaris-MH is typically delivered as a service through the Web from Polaris's secure data facilities, though some customers choose to install the Web-delivery system on their own networks and manage the entire system themselves.

The system is designed for use throughout treatment. It assesses factors known to predict engagement in treatment and clinical outcomes. The data synthesized on the system's intake, update and severity-adjusted, aggregate-level reports can guide decision-making from the individual to the organizational level. A typical implementation, which can be adapted to specific provider requirements, consists of one or more patient-accessible

computers in a provider's facility for administering assessments, and a printer in a secure location for printing assessment reports. A broadband Internet connection is required, but no additional equipment (servers, etc.) is needed. The patient-accessible computers may be configured as "kiosks" on which only the Polaris-MH application is accessible, or they may be multi-purpose machines with Polaris-MH as one available option.

Development of Polaris MH involved extensive interviews with clinicians and support staff. As a result, it includes numerous features that facilitate its integration into routine clinic procedures. The patient completes an initial assessment, usually prior to their first appointment. Remote Access enables patients to complete assessments from any computer with Internet access, including their home computer. Polaris-MH does not require prior computer experience; patients use only the number keys and "Enter," and do not need the mouse. Patients can "pause" the assessment at any point. If they log on within 48 hours the assessment will resume where they left off. This enables a patient to complete the assessment after their session with the clinician. Mean completion time is 14 minutes (see next page for a briefer version, the Emotional Vital Signs). The unassisted completion rate in outpatient mental health settings is more than 95%. Patients are typically asked to complete assessments prior to meeting with their clinician. This ensures that the time required to complete the assessment does not reduce the clinician's time with the patient, and it makes the results of the assessment available at the start of the session.

As soon as the assessment is completed, all scores are automatically calculated and the clinical report is gener-

ated and available for review with the patient. The system may be configured so the clinical report prints automatically, or it can be easily accessed online. The report presents all the measures and clinical items described above in an easy-to-use "lab report" format. All scores are reported as percentiles, based upon norms for mental health outpatients. Not all patients respond authentically to all questions on an assessment, due to literacy limitations, cognitive impairment, deliberate "faking" and other factors. Polaris-MH automatically conducts a series of checks and alerts clinicians with a note on the report if there is evidence of inauthentic responding. Licenses for the use of Polaris-MH start at \$US750 per year. It is available for free for researchers. Please contact Polaris sales for additional information: sales@polarishealth.com.

An abbreviated version of Polaris-MH (Polaris-EVS – "Emotional Vital Signs") is available in both computer and paper/fax formats. The EVS was developed to accommodate programs where it is impractical for all patients to complete an assessment using a computer so Polaris developed a one-page fax form as an alternative. EVS preserves as much of the clinical content and functionality of the Polaris-MH system as possible while substantially reducing its length; the EVS requires six minutes to complete. Additional information on EVS can be obtained from the Polaris website (www.polarishealth.com).

Assessment and Treatment Planning

The broad scope of the clinical report supports treatment planning and joint patient-clinician identification of treatment goals. Presentation of both problem areas and strengths enables the clinician to draw

upon the patient's assets and identify specific areas of concern when making decisions about the course of a patient's treatment. The clinic may designate when to give the update assessments, and the type of update to be administered. A Full Update provides scores for all the scales assessed at Intake, requiring about eight minutes to complete. A Brief Update provides scores for all core scales, requiring about four minutes to complete. The system keeps track of each patient, and provides a number of tools to manage when patients are due for update assessments, and which type they should complete. Update intervals of three or four weeks are often employed: this period is long enough so that measurable improvement is likely to occur, and the program is assured of a final assessment within a few weeks of termination (necessary to the evaluation of treatment outcomes). The update report indicates the patient's satisfaction with treatment, therapeutic bond, and compliance with medication (when applicable); a trend line of BHS and Depression scale scores show the changes that have occurred since admission; and a table of change scores for scales and subscales of the symptoms and functioning domains indicate areas of improvement or of continuing difficulty.

The update report also includes an ETR curve (described above). It is displayed on the update report, together with the patient's actual BHS score and a Failure Boundary. The ETR curve indicates the rate and amount of improvement that would normally be achieved by patients with similar characteristics and initial severity score. By comparing the patient's progress with the ETR the clinician, (or clinical supervisor, care manager, or patient) can readily determine whether treatment is "working" as well as expected. A patient

score that falls below the Failure Boundary suggests, with 75% certainty, that the treatment outcome will not be favorable (8).

Technical Support

Complete documentation of the system, assessment and reports are provided. A *User Manual* and *Clinical Reference Guide* are also provided. Polaris can conduct training sessions (either in-person or through Web meetings) and supports train-the-trainer approaches. Polaris provides end-user technical support during normal business hours.

Psychometric Properties

Polaris-MH is normed for adults in outpatient mental health treatment. Psychometric properties have been documented across a broad range of behavioral and medical patient populations. They include internal consistency reliability; face, construct, criterion, concurrent and predictive validity; and sensitivity to change. Findings are published elsewhere (2,3) and summarized here. The internal consistency (Coefficient Alpha) of all scales is acceptable ($r>.70$) for use in individual patient monitoring. Reliabilities of the core scales are all in the Good-Excellent range (.80-1.0): Subjective Well-Being (SWB) - .86; Symptoms/S - .93; and Functional Disability/FD - .81; BHS - .83. The majority of subscales have internal consistency in the good-excellent range as well, ranging from .75 to .91. Internal consistency for the strength subscale (Resilience) is .80.

All Polaris-MH items are face valid, and directly relevant to the evaluation of a patient's condition and treatment monitoring. Staff report strong patient acceptance; patients view the comprehensive assessment as indicative of the clinician's commitment to providing

the best possible care. In more than 60,000 patient assessments there have been no reports of concern about the appropriateness of the questions. Construct validity of all scales is strong. Items for the symptom scales were constructed by re-casting symptoms from the *Diagnostic and Statistical Manual, Fourth Edition* (1) for self-report. Items of the Functional Disability subscales were constructed from the U.S. Social Security disability guidelines. Concurrent validity (with the Social Adjustment Scale, Global Severity Index, OQ-45, and General Well Being Scale) of the core scales is also strong. Sensitivity to change for 792 adults in outpatient mental health treatment for the core scales, after an average of 37 days (s.d.=28 days) in treatment was also assessed. Findings are consistent with the Phase Model, with effect sizes of .32 for functioning, .51 for

total symptoms, .61 for behavioral health status, and .75 for subjective wellbeing.

Institutional Implementation

Polaris-MH is uniquely suited to the requirements of interdisciplinary case management. Its pan-theoretical construction avoids the limitations inherent in measures that are based upon a specific "school" of therapy. It tracks changes in patients' feelings of well-being, symptoms and functioning—issues of primary concern to all mental health providers. A Summary Screen enables supervisors, case managers and clinicians to quickly review the clinician's caseload to identify cases for intensive review, e.g. due to suicidal risk, drug abuse or poor progress. In consideration of psychiatrists, Polaris-MH contains items regarding medication compliance. The measures for strengths and social

support reflect the holistic approach often used by social workers. Review of the patient's progress in relation to ETR is useful to all involved in a case by enabling clinicians, clinical supervisors and case managers to identify and conduct a detailed review of cases that are at risk for poor outcomes or are candidates for (successful) termination.

Polaris MH is used in diverse mental health treatment settings, including Kaiser Permanente Psychiatric Clinics in southern California, the Canadian Forces behavioral health treatment facility at Stadacona (N.S.), the Gosnold treatment network in Massachusetts (specializes in dually-diagnosed patients), the University of Wisconsin (Madison) Clinical Training Program and by individual practitioners.

REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC.
- Grissom, G. and Lyons, J. (2006). Pragmatic Case Studies and Evidence-Based Treatment: Research and Clinical Applications of a Computerized Outcomes Management System. *Pragmatic Case Studies in Psychotherapy* Vol 2 (3), 1-28.
- Grissom, G., Lyons, J. & Lutz, W. (2002): Standing on the shoulders of a giant: development of an outcome management system based on the dose model and phase model of psychotherapy. *Journal of Psychotherapy Research*, 12, 397-412.
- Howard, K.I., Kopta, S.M., Krause, M.S. & Orlinsky, D.E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist*, 41, 159-164.
- Howard, K., Lueger, R., Maling, M. & Martinovich, Z. (1993). A phase model of psychotherapy: Causal mediation of outcome. *Journal of Consulting and Clinical Psychology*, 61, 678-685.
- Howard, K.I., Moras, K., Brill, P.L., & Martinovich, Z. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*. 51, 1059-1064.
- Lueger, R., Martinovich, Z., Anderson, E., Howard, K., Lutz, W. & Grissom, G. (2001). Assessing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology*, 69, 150-158.
- Lutz, W., Martinovich, Z., & Howard, K. I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology*, 67, 571-577.
- McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Grissom, G., Pettinati, H., and Argeriou, M. (1992). Fifth Edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment*, Vol.9.
- Ruggiero, K.J., Del Ben, K., Scotti, J.R. & Rabalais, A.E. (2003). Psychometric Properties of the PTSD Checklist—Civilian Version. *Journal of Traumatic Stress*, Volume 16, Number 5, 495-502.
- Sangstrand, Scott B. (2001) Outcomes Management: A Case Study of Implementation-Strategy and Experience. Behavioral Outcomes and Guidelines Sourcebook, New York: Faulkner & Gray, 312-320.
- Sperry, L., Brill, P., Howard, K. I., & Grissom, G. (1996). Treatment outcomes in psychotherapy and psychiatric interventions. New York: Bruner/ Mazel.

The PSYCHLOPS (Psychological Outcome Profiles)



Mark Ashworth, Ph.D., DM, FRCGP, MRCP,
Department of Primary Care and Public Health
Sciences, King's College London

mark.ashworth@kcl.ac.uk

Maria Kordowicz, Department of Primary Care and Public Health Sciences, King's College London

Peter Schofield, Department of Primary Care and Public Health Sciences, King's College London

The PSYCHLOPS, or Psychological Outcome Profiles, is a mental health outcome measure which can be used as a before and after measure for determining outcomes following therapy, and as a repeated measures instrument used to track patient progress over the course of therapy. PSYCHLOPS aims to refocus outcome measurement away from professionally determined domains and towards a patient centered definition of outcomes. This emphasis on the patient perspective is intended to capture items of greatest personal significance rather than imposing an external frame of reference to interpret psychological distress. The measure, which is available in 6 languages, is designed for use in the context of primary care psychotherapy; with the development of a version for children, the PSYCHLOPS is suitable for anyone from the age of five years and upwards. The scale is self-administered and self-completed, and assesses three domains, including client problems, function and wellbeing. In this paper, the authors describe the scale, including how it is administered and scored, and the domains it assesses. Its use in clinical practice is also discussed.

Keywords: PSYCHLOPS; Psychological Outcome Profiles; treatment outcome; progress monitoring; psychotherapy

PSYCHLOPS, or Psychological Outcome Profiles, has been designed as a mental health outcome measure. Originally intended as a before and after measure, to be used for determining outcomes following therapy, it has now developed into a repeated measures instrument used to track progress throughout a course of therapy. PSYCHLOPS aims to refocus outcome measurement away from professionally determined domains and towards a patient centred definition of outcomes. This emphasis on the patient perspective is intended to capture items of greatest personal significance rather than imposing an external frame of reference to interpret psychological distress.

The development of PSYCHLOPS started in 1999 with the search by primary care therapists for an outcome instrument that captured aspects of recovery which appeared to be missed by conventional instruments. They reported patients who, during the course of therapy, appeared to have resolved many of the issues for which they were originally referred and yet outcome measurement had failed to capture this recovery. In response, an idiographic instrument was designed which contained questions asking patients to describe their problems in their own words using freetext response boxes. Idiographic instruments are well described in the literature (Donnelly & Carswell, 2002) but are generally used in secondary care and require assistance to complete either from the therapist or by someone trained in the specific requirements of the instrument. In contrast, PSYCHLOPS was developed with the intention of being the first, easy-to-use, self-administered idiographic measure. The end product was a one-page questionnaire.

Two principles guided the development of PSYCHLOPS. User-involvement was a key feature in its development and was provided by the UK patient organisation, Depression Alliance. Since the instrument was intended to be patient centred, it also had to be patient friendly. The wording was scrutinised by the Plain English Campaign and, after various revisions, the instrument was awarded the 'Crystal Mark' in recognition of the clarity of its language. An attractive design was also a feature of the patient-friendly approach to development and a simple format was devised with coloured banding used to highlight each question and colour differences to distinguish each version.

PSYCHLOPS was piloted then launched in 2004 as a pre-therapy and post-therapy mental health outcome instrument. Following validation studies, the instrument went through several stages of refinement to both wording and scoring, and a new during-therapy version was introduced in 2010 (Version 5). The intention of this version was that, after a period of change derived

through the iterative process of validation, it should be the definitive version of PSYCHLOPS, remaining unchanged for a minimum of five years. Stability over a longer period was intended to promote international collaborations and longer term studies.

The measure is designed for use in the context of primary care psychotherapy. Validation studies have excluded those with literacy problems, although therapist assisted completion would be possible in this situation. In 2011, a children's version of PSYCHLOPS was launched: PSYCHLOPS Kids. This is a shortened version of PSYCHLOPS and uses emoticon faces rather than tickboxes to elicit scores. With the development of 'PSYCHLOPS Kids', the instrument is suitable for anyone from the age of five years and upwards.

PSYCHLOPS is not intended for use as a diagnostic instrument and can therefore be used with patients experiencing a wide variety of mental health problems without being confined to those fulfilling single disease-based diagnostic criteria. This broad spectrum of distress is typical of the sort of mental health problems encountered in primary care. PSYCHLOPS has been developed in English but is also available in French, Spanish, Dutch, Polish and Arabic.

Domains Assessed

Three domains are included in PSYCHLOPS: Problems (2 questions); Function (1 question); Wellbeing (1 question). The underlying Problem-Function-Wellbeing domains are derived from a pan-theoretical model which describes an empirical sequence of causality; psychological problems which then trigger deficits in

functional capacity which in turn triggers diminished wellbeing. In parallel with its applicability to a broad range of mental health problems, the measure is applicable to a breadth of talking therapies and may be used before, during and after any type of psychological intervention.

Use and Procedures

The Problem and Function domains of PSYCHLOPS elicit freetext responses which are then scored by the therapist on an ordinal 6-point scale (ranging from a score of zero to five). If the patient only reports one Problem, rather than two, then the score is pro-rated (doubled) such that the maximum possible score for the Problem domain remains at ten. The Wellbeing domain is a nomothetic measure (omitting a freetext component), again scored zero to five. Thus the score range, derived from the sum of each domain, is from zero to 20.

In the pre-therapy version of PSYCHLOPS, patients are asked to describe their main Problem (in a freetext box) and to score it. In subsequent during-therapy and post-therapy versions, the therapist transcribes the freetext description of the original Problem, and the patient is asked to re-score the original Problem (the original score is not disclosed). The same process is followed for the other freetext questions: the second Problem and Function. The Wellbeing score is simply scored on the scale numbered zero to five, each time the instrument is administered.

PSYCHLOPS is self-administered and self-completed with the proviso that the therapist transcribes the freetext sections from the pre-therapy version to all subsequent during-therapy versions and the post-therapy version. PSYCHLOPS is an outcome measure, designed to measure

change, and the score reflects its purpose. The actual change is simply the during-therapy or post-therapy total score subtracted from the pre-therapy score. Interpreting this change requires calculation of the Effect Size for a sample. The Effect Size is calculated by dividing the change score by the standard deviation of the pre-therapy score. By using this method, change is 'standardised' and the greater the pre-therapy score variability, the greater the pre-therapy standard deviation and the less the overall Effect Size. Effect Size values greater than 0.8 are generally considered large in health service research (Kazis *et al.*, 1989). In common with all idiographic instruments, there is no population norm since the baseline score is a measure of items which differ between each person and is not strictly comparable between individuals.

All data are stored with the therapist and instrument completion is based on hard copies of the questionnaire, not on-line copies. There is no on-line version. Data collection is not centralised. The simplicity of score calculation means that a score can be calculated immediately upon completion of the instrument. Progress, or otherwise, can be charted on a zero to twenty scale. Ideally, this score would be available at the start of each talking therapy session.

There is a charge for use of PSYCHLOPS. Specimen copies may be viewed on the website. Actual copies are available on CD-ROM and cost £40 (\$CA65) for individual therapists, £100 (\$CA160) for small organizations and £250 (\$CA400) for larger health service organizations (employing over 100 people). There is no annual fee and there is no limit to usage.

Assessment and Treatment

Planning

Therapists have described the usefulness of pre-therapy freetext information reported in the Problem and Function domains of PSYCHLOPS (Ashworth et al., 2005a). This information can be triangulated with referral information to the therapist from other health professionals such as general practitioners, and provides a focus for therapy from the outset, acting as a tool to instigate therapeutic work. The during-therapy versions elicit information on new problems arising during the course of therapy, adding to the information available to the therapist. Although research evidence on the importance of this information is not yet available, it would seem intuitive to suggest that new issues described by patients on PSYCHLOPS would need to be addressed by therapists during the talking therapy process. There are no population norm data for PSYCHLOPS, in common with all idiographic instruments (Lacasse et al., 1999; Donnelly & Carswell, 2002).

Technical Support

Background information about PSYCHLOPS is available from the website: www.psychlops.org.uk. The website provides links to background literature, validation studies, the scoring system, latest developments and an email address for further information. Upon purchase, an information pack is mailed out, providing further in-depth information on instrument usage.

Psychometric Properties

Internal reliability has been tested by calculating Cronbach's alpha for the three domain scores in PSYCHLOPS. Three studies have so far reported internal reliability data based on alpha scores: 0.79 pre-therapy and 0.87 post therapy (Ashworth et al., 2005b); 0.75 pre-therapy and 0.83 post therapy (Ashworth et al., 2008); and 0.81 pre-therapy, 0.85 during therapy and 0.88 post therapy (Czachowski et al., 2011). One study has reported test-retest reliability, based on a survey of students in higher education. The test-retest intra-class correlation coefficient was 0.70 (Evans et al., 2010). Sensitivity to change has been reported as Effect Size in three studies: 1.53 (95% CI 1.30 to 1.76) (Ashworth et al., 2005b); 1.61 (95% CI 1.41, 1.80) (Ashworth et al., 2008); 3.1 (95% CI 2.7, 3.4) (Czachowski et al., 2011)

The first two studies were conducted in the setting of psychotherapy offered within the context of primary care in the UK. The latter study was conducted in a Polish setting and the brevity of the psychotherapy programme (three sessions of CBT conducted by GPs with a special interest in CBT) may have contributed to the high observed Effect Size.

Convergent validity of the measure has been reported in two studies: comparison with CORE-OM (Clinical Outcomes in Routine Evaluation – Outcome Measure) revealed a Spearman's *rho* of 0.61, pre- and post therapy data combined (Ashworth et al., 2005b); comparison with HADS (Hospital Anxiety Depression Scale) showed a *rho* of 0.47 pre-therapy and 0.63 post-therapy (Ashworth et al., 2008).

Patient/Client/Clinician Feedback

Based on the findings of a qualitative study of the views of therapists (Ashworth, et al., 2005a), PSYCHLOPS was perceived as complementing the information derived from conventional quantitative instruments, with its qualitative information being of particular interest to therapists, contributing to the therapist-patient interaction. Therapists reported that it was a 'therapist friendly' instrument and likely to increase acceptance and uptake of outcome measures.

Institutional Implementation

PSYCHLOPS is a generic instrument, designed within the context of primary care but not aligned to any one primary care discipline nor psychotherapy discipline. Its findings are of relevance to all primary care health professionals involved in the care of patients with mental health problems. The UK Department of Health (DH) has produced a list of DH 'approved' mental health outcome measures which was published in an 'Outcomes Compendium' in 2009. PSYCHLOPS is included in the list of approved measures: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_093316

PSYCHLOPS is also included in the international Quality of Life Instruments database: www.proqolid.org.

PSYCHLOPS combines both quantitative and qualitative information. On-going studies will report on both qualitative analysis of patient reported data and more detailed quantitative analyses in a variety of international contexts.

If you would like to know more about PSYCHLOPS or would like to purchase a copy, please go to the website: www.psychlops.org.uk or contact the PSYCHLOPS research manager, Marilyn Peters: marilyn.peters@kcl.ac.uk

REFERENCES

- Ashworth M, Robinson S, Godfrey E, Parmentier H, Shepherd M, Christey J, Wright K, Matthews V. The experiences of therapists using a new client-centred psychometric instrument, 'PSYCHLOPS' ('Psychological Outcome Profiles'). *Counselling and Psychotherapy Res* 2005;5:37-41.
- Ashworth M, Robinson S, Godfrey E, Shepherd M, Evans C, Seed P, Parmentier H, Tylee A. Measuring mental health outcomes in primary care: the psychometric properties of a new patient-generated outcome measure, PSYCHLOPS (Psychological Outcome Profiles). *Primary Care Mental Health* 2005;3:261-270.
- Ashworth M, Evans C, Clement S. Measuring psychological outcomes after cognitive behaviour therapy in primary care: a comparison between a new patient-generated measure, 'PSYCHLOPS' (Psychological Outcome Profiles) and 'HADS' (Hospital Anxiety Depression Scale). *Journal of Mental Health* 2008;1-9 iFirst article.
- Czachowski S, Seed P, Schofield P, Ashworth M. Measuring psychological change during cognitive behaviour therapy in primary care: a Polish study using 'PSYCHLOPS' (Psychological Outcome Profiles). *PLoS ONE* 2011; 6(12): e27378 doi:10.1371/journal.pone.0027378
- Donnelly, C. & Carswell, A. (2002) Individualized outcome measures: a review of the literature. *Canadian Journal of Occupational Therapy*, 69, 84-94.
- Evans C, Ashworth M, Peters M. Are problems prevalent and stable in non-clinical populations? Problems and test-retest stability of a patient-generated measure, 'PSYCHLOPS' (Psychological Outcome Profiles), in a non-clinical student sample. *British Journal of Guidance and Counselling* 2010;38:431-439.
- Kazis, L., Anderson, J. & Meenan, R. (1989) Effect sizes for interpreting changes in health status. *Medical Care*, 27, S178-189.
- Lacasse, Y., Wong, E. & Guyatt, G. (1999) Individualising questionnaires. In: Joyce CRB, O'Boyle CA, McGee H, eds. *Individual quality of life. Approaches to conceptualisation and assessment*. Harwood Academic Publishers: Amsterdam, 87-103.

The Schwartz Outcome Scale – 10 (SOS)-10



Mark A. Blais, Psy.D.,
Department of Psychiatry,
Massachusetts General Hospital
and Harvard Medical School

Mblais@partners.org

The Schwartz Outcome Scale – 10 (SOS-10) was developed to fill the need for an outcome measure suitable for all patients, all treatments, and all levels of care. More specifically, it was designed to monitor outcomes, at both the individual and aggregate level, across a wide range of adult mental health services. The SOS-10 is suitable for individuals ages 17 and up, and is available in a number of languages. It assesses a number of key areas, including life satisfaction, interpersonal effectiveness, positive self-appraisal, optimism, and the absence of psychiatric symptoms. SOS-10 scores can be used to rapidly identify a patient's level of emotional distress or psychological dysfunction, which in turn is valuable in clarifying the intensity of services needed and in treatment planning. This paper describes the development of the scale, how it is administered and scored, its psychometric properties, and the domains it assesses. Its use in clinical practice is also discussed.

Keywords: Schwartz Outcome Scale – 10; SOS-10; treatment outcome; progress monitoring; psychotherapy

In the late 1990's leadership in the Department of Psychiatry at Massachusetts General Hospital identified the need for "*an outcome measure suitable for all patients, all treatments, and all levels of care*". The Schwartz Outcome Scale (SOS-10; Blais et al., 1999) was created to fill that need. The SOS-10 is a unique broadband low burden measure developed to monitor outcomes, at both the individual and aggregate level, across a wide range of adult mental health services. The distinctiveness of the SOS-10 derives from its method of development. Rather than relying on theory, symptoms or existing instruments, construction of the SOS-10 was guided by insights obtained from a diverse group of senior clinicians and patients. Specifically, interviews conducted with senior psychologists, psychiatrists and a neurosurgeon along with patient focus groups were used to discover the changes that occurred (excluding symptoms) with successful treatment. The interviews and focus group discussions were transcribed and reviewed for common themes. Common themes were used to generate an initial item pool. Empirical evaluation and refinement identified 20 well performing items and Rasch analysis was employed to reduce the scale to its final 10-item version (see Blais et al., 1999 for a detailed description of the development process).

The SOS-10 is suitable for individuals ages 17 and up. It has been formally translated into French, Czech, and Spanish. Chinese and Italian translations are also available. Recent promising efforts to extend the use of the scale downward into adolescent populations are also in process.

Domains Assessed

The SOS-10 is a measure of psychological health and well-being. Psychological health is conceived of as an overarching construct that encompasses life satisfaction, interpersonal effectiveness, positive self-appraisal, optimism, and the absence of psychiatric symptoms.

Use and Procedures

Patients are asked to rate how they have been doing over the last week on 10 items using a 0 (Never) to 6 (All or nearly all the time) scale. The SOS-10 is scored by summing the numerical ratings for each item. This process creates a total score ranging from 0 to 60 with higher scores representing greater psychological health well-being and lower scores indicating emotional distress and poorer psychological health. While the SOS-10 has no validity scale, scores at the extreme ends of the range (0 or 60) are rare (occurring

less than 2 percent of the time) and are therefore considered invalid. The scale can be scored with up to two missing items by using a mean score imputation to generate a total score. The SOS-10 can be administered in traditional paper-and-pencil format or electronically (score equivalence has been demonstrated for web based administration). It is recommended that patients complete the scale prior to a treatment appointment. This way the clinician can determine whether the SOS-10 was completed and is valid, and review the total score for clinical implications prior to the session.

The SOS-10 has been widely adapted as a program level treatment evaluation tool and many programs have contributed data to the SOS-10 interpretive database. Presently, the database contains intake SOS-10 scores for 8,056 outpatients and 5,541 inpatients. As a measure of psychological health and well-being, the SOS-10 is also attractive to non-clinical researchers. As a result our data base also contains SOS-10 scores for 2,000 non-patients.

Although the SOS-10 is a proprietary instrument, the scale is made available free of charge for practitioners, researchers and non-profit healthcare organizations.

Assessment and Treatment Planning

Owen and Imel (2010) outline a rationale and a practice friendly procedure for incorporating the SOS-10 into ongoing clinical care. The availability of non-patient reference data is valuable as it allows for calculation of both a Reliable Change Index and Clinically Significant Improvement. The ability to apply more sophisticated treatment effectiveness analyses

greatly enhances the information obtained from TAU outcome measurement programs and increases the comparability of findings across studies (see Blais et al., 2011). SOS-10 scores can also be used to rapidly identify a patient's level of emotional distress or psychological dysfunction. Drawing on data from over 8,000 outpatients the following distress ranges may prove helpful markers: *Minimal* (59-40), *Mild* (39-33), *Moderate* (32-23) and *Severe* (22-1). Accurately identifying a patient's level of distress at the outset of treatment can help clarify the intensity of services needed, i.e. weekly individual psychotherapy, multiple sessions per week or multiple forms of treatments. In this way routine use of the SOS-10 can aid treatment planning. Furthermore, as clinicians become familiar with the tool, the use of severity ranges can provide easily recognized reference points for multidisciplinary communication. Lastly, because SOS-10 items are not directly related to psychiatric symptoms, reviewing unique responses to individual items with patients can afford a non-threatening avenue for discussing personal strengths and weaknesses.

Psychometric Properties

The SOS-10 has outstanding psychometric properties. Its internal consistency in published studies has ranged from 0.84 to 0.96. The test-retest reliability for the scale is also strong, with studies reporting retest correlations of 0.86 and 0.87. In addition, no meaningful age or gender effects have been reported. Multiple studies both in the original English and in translations have found the SOS-10 to be uni-factorial. Factor invariance has also been shown across samples (patients & non-patients), and measurement points (pre & post

treatments). The accumulated research also supports the construct validity of the SOS-10 as a broad measure of psychological functioning (Blais et al., 1999; Haggerty et al., 2009; Young et al., 2004). The SOS-10 correlates significantly and in the predicted direction with measures of psychiatric symptom severity (-0.67), alexithymia (-0.58), hopelessness (-0.66), negative affect (-0.72), self-esteem (0.81), satisfaction with life (0.78), positive affect (0.67) and physical functioning (0.36). SOS-10 is also significantly related to measures of the normal personality (Big Five Traits). The SOS-10 correlates significantly with the Outcomes Questionnaire-45 (OQ-45; Lambert et al., 1996). The SOS-10 is strongly correlated with OQ-45 total score (-0.84), and with its subscales. Together these findings demonstrate the breadth of the SOS-10 and offer solid evidence of its construct validity.

The SOS-10 has also demonstrated sensitivity to change for a wide variety of treatment modalities and may be especially sensitive to detecting early treatment change (Hilsenroth et al., 2001). The SOS-10 has been employed as an outcomes measure in studies of Psychodynamic Psychotherapy, Dialectical Behavior Therapy, residential treatment for refractory Obsessive Compulsive Disorder, Inpatient psychiatric treatment as usual and inpatient substance abuse treatment as usual. A study by Blais et al. (2010) demonstrated the utility of the SOS-10 as a common outcome measure for evaluating treatment as usual across a large diverse outpatient psychiatric practice.

Institutional Implementation

The SOS-10 is currently used as a common outcome measure for all adult psychiatry services provided within the Partners

Healthcare System. Partners Healthcare includes the majority of hospitals and community health clinics associated with Harvard Medical School. Many other psychiatric hospitals, community mental health centers and college counseling centers across the United States have been granted permission to use the SOS-10, as have a number of treatment facilities in the United Kingdom. It has also been licensed for use by managed care organizations.

REFERENCES

- Blais, M. A., Lenderking, W. R. Baer, L., deLorell, A., Peets, K., Leahy, L., & Burns, C. (1999). Development and initial validation of a brief mental health outcome measure. *Journal of Personality Assessment, 73*, 359-373.
- Blais, M. A., Sinclair, S., Baity, M., Worth, J., Weiss, A., Ball, L., et al. (2011). Measuring outcomes in adult outpatient psychiatry *Clinical Psychology & Psychotherapy* published online 17 Mar 2011 (DOI: 10.1002/cpp.749).
- Haggerty, G., Blake, M., Naraine, M., Siefert, C., & Blais, M. (2010). Construct validity of the Schwartz Outcome scale-10: Comparisons to interpersonal distress, adult attachment, alexithymia, the five factor model, romantic relationship length and ratings of childhood memories. *Clinical Psychology and Psychotherapy, 17*(1) 44-50.
- Hilsenroth, M., Ackerman, S., & Blagys, M. (2001). Evaluating the Phase Model of Change During Short-Term Psychodynamic Psychotherapy. *Psychotherapy Research, 11*, 29-47.
- Lambert, M., Hansen, N., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G., & Reinsinger, C. (1996). *Administration and Scoring Manual for the Outcome Questionnaire (OQ-45.2)*. Stevenson, MD: American Professional Credentialing Services.
- Owens, J., and Imel, Z. (2010). Rating scales in psychotherapy practice. In L. Baer & M Blais (Eds.), *Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health*. New York, NY: Humana Press, pages 257-270.
- Young, J. L., Waehler, C. A., Laux, J. M., McDaniel, P. S., & Hilsenroth, M. J. (2003). Four studies extending the utility of the Schwartz Outcome Scale (SOS-10). *Journal of Personality Assessment, 80*, 130-138.

The Treatment Outcome Package (TOP)



David R. Kraus, Ph.D.,
Outcome Referrals, Inc.

dkraus@outcomereferrals.com

This paper presents the Treatment Outcome Package (TOP), which is designed to serve as a clinically useful assessment and outcome battery for all levels of behavioral healthcare, including traditional psychotherapy. The TOP, which is available in nine languages, is a-theoretical and assesses twelve clinical and functional domains. Depending on the age of the identified patient these domains can include: quality of life, substance abuse, depression; panic/anxiety; psychosis, mania, suicidality, violence, work and school, social dysfunction, sexual functioning, sleep disorders, eating disorders, conduct disorders, sexual aggression, separation anxiety, attention deficits, assertiveness, bladder control and psychological strengths. The TOP provides clinicians with real-time clinical reports through the use of alerts to off-track treatment that is likely to end in deteriorated outcomes. Scoring is compared to non-clinical norms so that the level of pathology can easily be tracked. This paper describes the TOP and presents its psychometric properties, as well as the procedures related to its use in diverse clinical contexts.

Keywords: Treatment Outcome Package, TOP; treatment outcome; progress monitoring; psychotherapy

The Treatment Outcome Package (TOP) was designed to meet the objectives of the Core Battery Conference (Horowitz, Lambert & Strupp, 1997) and developed to serve as a clinically useful assessment and outcome battery for all levels of behavioral healthcare (Kraus, Seligman & Jordan, 2005). Initial versions of the tool began with 250 items derived from DSM-IV symptoms and refined through extensive exploratory and confirmatory factor analytic work for adult, adolescent and child populations (e.g., Kraus, Boswell, Wright, Castonguay & Pincus, 2010). TOP is available in English, Spanish, German, Dutch, Portuguese, Chinese, Vietnamese, Haitian, and Cape Verdean. In addition, a French version is in process.

Domains Assessed

TOP is a-theoretical and assesses twelve clinical and functional domains that include the following (depending on the age version chosen):

- Quality of Life
- Substance Abuse
- Depression
- Panic/Anxiety
- Psychosis
- Mania
- Suicidality
- Violence
- Work/School
- Social Dysfunction
- Sexual Functioning
- Sleep Disorders
- Eating Disorders
- Conduct Disorder
- Sexual Aggression
- Separation Anxiety
- Attention Deficits (ADHD)
- Assertiveness
- Bladder Control
- Psychological Strengths

Use and Procedures

TOP is provided as a free service through WellnessCheck.net, including free scoring and real-time clinical reporting. Online manuals and videos provide help with administration and scoring. Methods are also available that send customized links to patients so that clinicians do not need to administer questionnaires in the

office and patients can complete questionnaires (ideally once per month) at home or work. Daily reminders are provided via email until the patient completes the agreed upon questionnaire or decides to withdraw. Privacy is protected by the use of identifiers that only the provider knows how to link back to an individual and all direct identifiers to an individual (name, address, etc.) are not collected or stored. Potentially identifiable information is encrypted and kept in non-linkable systems so that patient anonymity is protected (Kraus & Horan, 1998).

Assessment and Treatment

Planning

Real-time clinical reports are designed by clinicians to maximize clinical relevance (Kraus, Wolfe & Castonguay, 2006) through the use of alerts to off-track treatment that is likely to end in deteriorated outcomes and/or expensive psychiatric hospitalization. Scoring is compared to general population (non-clinical) norms so that the level of pathology can easily be tracked graphically for the past 20 TOP administrations. Reviewing the detailed clinical reports throughout treatment enhance the therapeutic alliance, helps patients to reveal important clinical information, aids in treatment plan reviews and setting goals and priorities.

Monthly aggregate reports are provided for each clinician, service (e.g., a partial hospitalization program) and each agency. These reports highlight areas of strength and weakness and are benchmarked, risk adjusted, and mapped against similar professionals across the

world (or local regions) based on a growing database of over a million patients (Kraus & Castonguay, 2010). TOP domains have been linked to evidence-based practices and principles and reporting structures facilitate rapid improvement of substandard treatment benchmarks (Kraus, Wolfe & Castonguay, 2006; Adelman, 2005, 2006, 2007, 2008). These systems are used to identify the inherent strengths in any given provider population with 96% of all providers demonstrating reliable proficiency in treating at least one major symptom cluster (Kraus, Castonguay, Boswell, Nordberg & Hayes, 2011).

Technical Support

Customer support is provided through toll-free telephone numbers and on-line resources, manuals and videos. Since [WellnessCheck.net](#) scores, reports, stores, and benchmarks all data, there is no need for scoring manuals or procedures.

Psychometrics Properties

Numerous studies (Kraus, Seligman & Jordan, 2005; Kraus & Castonguay, 2010; Kraus, Boswell, Wright, Castonguay & Pincus, 2010) using confirmatory factor analysis with data collected from 19,801 participants in 383 facilities (5 split samples) have shown the construct validity of the TOP, with a Goodness-of-Fit Index above .95, a Comparative Fit Index of .95, a Non-Normed Fit Index of .94, and a Root Mean Square Error of Approximation of .035. Studies have also shown that the TOP has good internal consistency, ranging from .53 to .93, and strong test-retest reliability ranging from .76 to .94.

The discriminant validity of the TOP is excellent, with 92% of consumers showing pathological scores and TOP scores demonstrating an ability to discriminate between subjects who are in treatment and those who are not. Convergent validity has been demonstrated with a number of well established measures, including the Beck Depression Inventory (BDI), the Minnesota Multiphasic Personality Inventory (MMPI), the Brief Symptom Inventory (BSI), the BASIS-32, the SF-36, the Child Behavioral Checklist (CBCL), and UNCOPE. For example, the depression scale of the TOP was highly correlated ($r=.91$) with the BDI. Norms for the TOP were established using large samples of over 1 million participants for the clinical norms and of 2,000 participants for the general population norms.

Patient/Client/Clinician Feedback

Each completed TOP is centrally processed, scored and returned with a two- or three-page feedback report. This feedback provides alerts as to whether treatment is on track and delivers a checklist of evidence-based tasks that should be completed or considered in order to increase the chances that treatment will end successfully. In addition, clinicians are alerted if the patient is at high risk of being hospitalized within the next six months for expensive psychiatric or substance abuse treatment (McAleavy, Nordberg, Kraus & Castonguay, in press). The second page of the TOP feedback report is designed to be copied and given to the client as patient-level feedback (Youn, Kraus & Castonguay, in press).

Institutional Implementation

TOP has been used by more than 40,000 clinicians, and thousands of organizations including self-insured employers, health plans, hospitals, community mental health centers and provider networks to document and improve the quality of care. Each TOP domain is tied to libraries of evidence-based practices and principles that include scorecards, risk-adjusted benchmarking and improvement reports all designed to deliver roadmaps for innovative quality improvement strategies that have won TOP customers the highest awards for quality improvement (Adelman, 2005, 2006, 2007, 2008).

REFERENCES

- Adelman, R. et. al, (2005). Reducing adolescent clients' anger in a residential substance abuse treatment facility. *Journal on Quality and Patient Safety*, 31, (6), 325-327.
- Adelman, R. (2006). The Angry Adolescent & Constructivist REBT. In Cummins, P. (Ed.), *Working with Anger: a Practical Constructivist Approach*, London: John Wiley & Sons.
- Adelman, R. (2007). Reducing Anger in Adolescents: An REBT Approach, Center City, Minnesota: Hazelden Press.
- Adelman, R. (2008). Methods of Reconstruction with Adolescent Substance Abusers: Combining REBT and Constructivism. In Raskin, J.D. & Bridges, S. (Ed.), *Studies in Meaning 3: Constructivist Psychotherapy in the Real World*, New York: Pace University Press.
- Horowitz, L.M., Lambert, M.J., & Strupp, H.H. (Eds.). (1997). Measuring patient change in mood, anxiety, and personality disorders: Toward a core battery. Washington, DC: American Psychological Association Press.
- Kraus, D., Boswell, J., Wright, A., Castonguay, L., & Pincus, A., (2010). Factor Structure of the Treatment Outcome Package for Children. *Journal of Clinical Psychology*, 66, 627-640.
- Kraus, D.R., & Castonguay, L.G. (2010). TOP: Development & use in naturalistic settings. In M. Barkham, G. Hardy, & J. Mellor-Clark (Eds.), *A CORE approach to delivering practice-based evidence in counseling and the psychological therapies*. London: Wiley Press.
- Kraus, D.R., Castonguay, L.G., Boswell, J.F., Nordberg, S.S., & Hayes, J.A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research*, 21, 267-276.
- Kraus, D.R., Castonguay, L.G., Hayes, J.A. & Barber, J.P. (2010) The Empirically Supported Therapist: All Clinicians Have Strengths and Weaknesses. *Psychologie Quebec*, 1, 12-15.
- Kraus, D. R., & Horan, F. P. (1998). Protecting Client Confidentiality and Improving Provider Relations Through a New Form of Managed Care: Collaborative Care Management. The 1998 Behavioral Outcomes and Guidelines; Faulkner and Gray: New York.
- Kraus, D.R., Seligman, D., & Jordan, J.R., (2005). Validation of a behavioral health treatment outcome and assessment tool designed for naturalistic settings: The treatment outcome package. *Journal of Clinical Psychology*, 61, 285-314.
- Kraus, D., & Wolfe, A. & Castonguay, L.G., (2006). The outcome assistant: A kinder philosophy to the management of outcome. *Psychotherapy Bulletin*, 41, 23-31.
- McAleavey, A.A., Nordberg, S.S., Kraus, D.R., Castonguay, L.G., (in press) Errors in treatment outcome monitoring: Implications of multidimensional and general measurements for real-world psychotherapy. *Canadian Psychology*.
- Youn, S.J., Kraus, D.R., Castonguay, L.G. (in press) The Treatment Outcome Package: Facilitating practice and clinically relevant research. *Psychotherapy*.