Pragmatic Case Studies and Evidence-Based Treatment: Research and Clinical Applications of a Computerized Outcomes Management System

GRANT GRISSOM [a,c] & **JOHN S. LYONS** [b]

[a] President and CEO, Polaris Health Directions, Inc.
[b] Feinberg School of Medicine, Northwestern University
[c] Correspondence concerning this article should be addressed to Grant Grissom, Polaris Health Directions, 446 Lincoln Highway, Fairless Hills, PA 19030
Email: grant@polarishealth.com

ABSTRACT

This article describes an outcomes management system, Polaris-MH, and its potential utility for supporting the Pragmatic Case Study (PCS) model for psychotherapy research proposed by Fishman. Polaris-MH is a computer-based system for assessment of adults in outpatient mental health treatment. It provides for the collection, storage, analysis and reporting of information on a patient's clinical condition and progress. The system is grounded in research on dose-response, the phase model, and expected treatment response. This research foundation is described along with Polaris-MH development, structure, psychometric properties and clinical utility. Polaris-MH features are discussed in relation to the requirements of the PCS model.

Key words: outcomes management; computerized assessment; mental health treatment; pragmatic case study; learning systems; expected treatment response (ETR)

In his advocacy for a paradigm expansion in psychotherapy research, Fishman (2000, 2005) describes limitations of the predominant, *treatment-focused research model*. This model is grounded in group-based efficacy (e.g., Nathan & Gorman, 2002) and effectiveness (e.g., Seligman, 1995) approaches to psychotherapy research. To address these limitations and following Howard, Moras, Brill, and Martinovich (1996), Fishman proposes an additional, *patient-focused* model that examine therapy process and outcome in individual cases. Specifically, he proposes a Pragmatic Case Study (PCS) paradigm that involves the aggregation of individual case studies for which the context, process and outcomes of treatment re systematically recorded; and he has developed this *PCSP* journal as a vehicle for promoting the PCS model.

The PCS model has three essential requirements: (1) development of outcomes constructs that are both grounded in established theory and appropriate to a very broad range of patients and treatment approaches; (2) psychometrically strong measures of each domain; and (3) the ability to collect, store and analyze data for a large number of patients. The successful implementation of PCS requires the construction of automated outcomes management systems that achieve the

requirements outlined above. The purpose of this paper is to describe the development and utility of "Polaris-MH" (formerly, "Treatment Evaluation and Management" or "TEaM"), which is such a system. It not only provides researchers with a measurement tool that is appropriate to efficacy and effectiveness studies, it also is uniquely suited to the PCS model. For clinicians it provides an evidence-based alternative to exclusive reliance on the "empirically supported treatments" paradigm (e.g., Nathan & Gorman, 2002), preserving their option to employ whatever treatment model they consider appropriate, as long as they can document that the treatment works for the clients they serve. (Of course, many clinicians use the literature on empirically supported treatments as a guide for their treatment. However, the patient-focused model of Polaris-MH enables clinicians to go outside this literature if their results equal or exceed those of empirically supported treatments.)

Polaris-MH is a personal-computer-based outcomes management system that includes measures appropriate to monitoring the general effectiveness of adult outpatient psychotherapeutic services (Grissom, Lyons, & Lutz. 2002; Grissom, 2002). In addition, to improve the likelihood that clinicians become involved in the collection and use of data, it has numerous features that facilitate integration of the assessments into routine clinical practice. The use of standardized assessment in routine treatment is critical to the PCS model, which requires a growing database of individual cases with clinically relevant treatment outcome measures. This in turn requires either a major increase in funding for psychotherapy research, or the harvesting of the data available from patients in routine treatment.

POLARIS-MH DESCRIPTION AND DEVELOPMENT

As mentioned, Polaris-MH provides for the collection, clinical use and accumulation of information about patient characteristics, treatments and outcomes. This makes it possible to learn from experience – i.e., to continually refine our understanding of "what works, for whom" (Sperry, Brill, Howard & Grissom, 1996). The system includes:

1) *Hardware*, consisting of the devices used to administer surveys, process and store data, and produce reports.

2) *Software* that provides for questionnaire administration, data processing, data base management, report generation, and various administrative functions.

3) *Questionnaires*:
   - *Patient Intake*: Provides patient-completed, detailed information for treatment planning.
   - *Patient Update*: Concurrently with treatment, provides patient-completed information about the patient's condition, progress, therapeutic bond and satisfaction at regular points over the course of treatment.
   - *Brief Patient Update*: This is a subset of items from the Patient Uptake form. It includes a global mental health status indicator, a measure of the severity of depressive symptoms, and a treatment satisfaction item.

- *Counselor*: Provides for a counselor-completed assessment of the patient's condition, progress and prognosis (e.g., for utilization review).

4) *Procedures* for administering the patient questionnaires.

5) *Reports* that provide information for clinical decision support (individual patient reports) and for outcomes assessment (program level aggregate data).

Web-based and FAX options are also available. The web-based system is functionally identical to the personal-computer-based version.

In a typical implementation patients are told that the therapist wants to provide treatment that addresses the patient's situation and, to do that, asks the patient to complete the Polaris-MH assessments at admission and at regular intervals during treatment. Patients respond to multiple-choice questions using the numeric keys of a personal computer. A detailed report is immediately printed for use in the clinical session.

The development of Polaris-MH has been described in detail elsewhere (Grissom, Lyons & Lutz, 2002). Summarized below are the system's research foundation and psychometric properties. The three key components of the research foundation – the Dosage, Phase and Expected Treatment Response models – are described along with the Polaris-MH measurement domains. The section on psychometric properties presents information relating to the internal consistency reliability (Cronbach Alpha) of Polaris-MH scales and concurrent validity; it also addresses issues relating to integration of the system into routine clinical use, including the interpretation (clinical and statistical significance) of changes in patients' scores during the course of treatment. The concluding section of the paper describes features that make Polaris-MH an appropriate option for PCS applications.

FOUNDATIONS FOR POLARIS-MH DEVELOPMENT: DOSAGE, PHASE AND EXPECTED-TREATMENT-RESPONSE (ETR) MODELS

The late Kenneth Howard's contributions to psychotherapy research earned him numerous prestigious awards, including the Distinguished Contributions to Knowledge award from the American Psychological Association. Three of his research interests formed the scientific foundation for Polaris-MH.

Dosage Model.

From meta-analytic data, Howard and his colleagues developed a model of psychotherapeutic effectiveness that demonstrated a positive relationship between the log of the number of sessions (dose) and the normalized probability of patient improvement (effect) during outpatient adult mental health treatment (Howard, Kopta, Krause & Orlinsky, 1986; Newman &

Howard, 1986). Subsequent dose-effect work has provided evidence for the differential, but lawful, responsiveness to psychotherapy of various symptoms (Kopta, Howard, Lowry & Beutler, 1994); interpersonal problems (Horowitz, Rosenberg, Baer, Ureno & Villasenor, 1988; Maling, Gurtman & Howard, 1995); and diagnoses (Howard et al, 1996; Pilkonis & Frank, 1988). The dosage model describes a pattern of relatively rapid early improvement, with more and more sessions needed to achieve incremental improvement later in treatment.

Phase Model

A central challenge in the development of outcomes management systems has been the need to decide what should be measured. Efforts to establish a core battery of measures resulted in very lengthy assessments, in the hope of accommodating the needs of various schools of therapy. Batteries proposed by the lead author to clinical practitioners during the late 1980s were routinely rejected with the same two comments: It is too long to be practical as a clinical tool; and it is essential that you add …(the practitioner's favorite measure). This seemed an insurmountable obstacle.

Howard and his colleagues at Northwestern University were at that time conducting research that would provide a solution. The phase model (Howard, Lueger, Maling & Martinovich, 1993) specified three dimensions of patient change that appear necessary to treatment success, regardless of theoretical therapeutic approach. It extended and interpreted the dosage model. The phase model proposes three progressive, sequential phases of the psychotherapeutic recovery process: *re-moralization*, the enhancement of well-being; *remediation*, the achievement of symptomatic relief; and *rehabilitation*, the reduction of troublesome, maladaptive behaviors that interfere with life functioning (e.g. functioning in areas such as family relationships and work).

The phase model suggests that the decelerating curve of improvement for a patient can be attributed to the increasing difficulty of treatment goals as they change over the course of treatment. The researchers enrolled 471 patients seen by 86 therapists at the Northwestern Memorial Hospital Institute of Psychiatry. Self-report assessments were completed at intake (N=471), session 2 (N=184), session 4 (N=157) and session 17 (N=74). Measures at each update assessment for each construct (subjective well-being, symptomatic distress and life functioning) were converted into dichotomous (improved-not improved) scores, using a method based upon the Jacobson & Truax (1991) reliable change index. Analysis of the 2X2 cross-classification tables generated from these dichotomous measures indicated that functioning was unlikely to improve unless there had been a remediation of symptoms, and that this was in turn unlikely unless preceded by improved subjective well-being (Howard et al, 1993). Re-moralization (improvement in measures of well-being) usually improves within the first sessions; symptom remediation typically requires 12-16 sessions. The time required for rehabilitation of life functioning – establishing new ways of dealing with various aspects of life – is quite variable, depending upon the degree and area of disability (Sperry et al, 1996; Kopta et al, op.cit.).

Both the dosage and phase models rely on group data to provide outcome information for an average client in outpatient therapy. However, patterns of improvement for individuals vary around this general trend due to pretreatment clinical characteristics, therapeutic bond, and other patient variables (Barkham, Stiles & Shapiro, 1993; Kadera, Lambert, & Andrews, 1996; Krause, Howard, & Lutz, 1998; Martinovich, 1998).

Since the models are trans-theoretical (they do not reflect a specific "school" of therapy) they lend themselves well to the development of an Expected Treatment Response (ETR) model. An ETR model that estimates the rate and type of treatment progress expected after adjusting for relevant patient characteristics, independent of the specific therapeutic approach, offers an option for the empirical validation of models of therapy. A treatment approach would be validated if it produced results at least as favorable as the outcomes projected by the ETR model, as described in the following section.

(Note that the method which "adjusts" relevant patient characteristics involves fitting ETR curves for subgroups by a method developed by Lutz, Martinovich, and Howard (1999) rather than the traditional approach of statistically adjusting through the use of analyses of covariance.)

Expected Treatment Response (ETR) Model

Expected Treatment Response (ETR) research aims at projecting an individual patient's likely response to treatment based upon information provided by the patient and/or the clinician (Howard, Moras, Brill, & Martinovich, 1996). The model assumes an underlying log-linear course of recovery in treatment for each patient, as described in the dosage model. The ETR model utilizes a hierarchical linear modeling strategy (Bryk, & Raudenbush, 1992) to depict the patient's behavioral health status over treatment as a log-linear function of session number. Specifically, it uses pretreatment clinical characteristics (e.g., severity, chronicity, previous treatment, and treatment expectation) to predict the patient's expected response over the course of his or her treatment. On the basis of the results of such an 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/her intake information is available (Howard, Moras, Brill, & Martinovich, 1996). Furthermore, 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.

WHAT POLARIS-MH MEASURES

The structure, logic, and internal consistency reliabilities of the Polaris-MH measures are outlined in Table 1. As indicated, Polaris-MH assesses a patient's (a) *subjective well-being* (SWB); (c) the severity of patient *symptoms* (S) associated with the most common disorders treated in outpatient settings; (c) the impact of the patient's psychological problems upon the patient's life (*functional disability*, FD); and (d) *therapeutic bond/satisfaction* with treatment. Polaris-MH tracks these measures, and the patient's Behavioral Health Status (BHS) -- a

composite score derived from SWB, S and FD – to indicate the patient's total response to treatment. The three components are readily accepted by clinicians, operating from all major therapeutic models, as being central to clinical decisions and outcomes assessment.

The Polaris-MH measurement domains correspond to the three phases of therapeutic progress described by Phase Theory: remoralization (SWB), symptom remediation (S) and rehabilitation (FD). The S scale is a composite of subscale scores, with each subscale corresponding to a disorder commonly treated in outpatient settings. The FD scale includes three subscales based upon the guidelines of the U.S. Social Security Administration. All scales and subscales are listed in Table 1.

Finally, a broad range of items is included either for their clinical utility, or their relationship to clinical outcomes (to provide for case mix adjustment). These include demographic and treatment history content; screens for general health problems, substance abuse, psychosis and bipolar disorder; and assessment of strengths (resilience, meaning or purpose in life), treatment motivation, treatment satisfaction, and the therapeutic bond.

Counselor Questionnaire (Optional). The counselor questionnaire is designed for programs that intend to use the system for utilization review or clinical supervision, or to improve their ability to predict a patient's response to treatment. The form is very brief, requiring five minutes to complete at intake and two minutes later in treatment. The counselor may provide DSM codes, case severity indicators, an assessment of the patient's motivation for treatment, severity of symptoms and functional disability, progress, and prognosis.

PSYCHOMETRIC PROPERTIES OF POLARIS-MH

Reliability

An assessment system designed for outcomes management must be administered upon admission and re-administered periodically during treatment. It must provide the clinical detail and scope necessary to inform clinical decisions. Scales must include enough items to ensure adequate internal consistency reliability. But the assessment must not be so lengthy that it imposes an undue burden upon patients. Polaris-MH uses response-adaptive ("branching") logic to achieve an effective balance between these two competing goals. The development process that led to achieving this aim – balancing satisfactory scale reliabilities and acceptable questionnaire length – has been documented (Grissom, Lyons, & Lutz, 2002). As shown in Table 1, scale reliabilities are all above 0.75. In term of internal reliability, these values support that the scales are reliable.

Concurrent Validity

The scales most often used to evaluate treatment progress are those corresponding to the three phases of change generally (Subjective Well Being, Symptoms, and Functional Disability – see Table 1), and to Depression, specifically. The composite Behavioral Health Status score is useful as a global indicator of patient severity and progress. As shown in Table 2, these scales evidence good concurrent validity in relation to existing scales covering the same general domains.

The 18-item General Well-Being Scale (GWS; Dupuy, 1977) is a self-report questionnaire that measures a broad range of satisfaction with self and/or quality of life. It is a composite measure of depressive affect, anxiety, stress, physical well-being, and sense of emotional control. The GWS has normative data on a national sample of nearly 7000 adults and has demonstrated good reliability and validity.

The Positive and Negative Affect Scale (PANAS; Watson & Tellegen, 1985) consists of a 10-item scale of positive affect and a 10-item scale of negative affect. PANAS is based on the observation that in a number of studies of self-reported mood a Positive and a Negative Affect factor consistently emerged as the first twoVarimax rotated dimensions in orthogonal factor analyses.

The Beck Depression Inventory (BDI) is a self-report measure designed to assess the intensity of depression in psychiatric patients as well as to detect depression in normal populations. The 21 items are rated on four-point, anchored response alternatives. The BDI has good psychometric properties. Beck, Steer, and Garbin (1988) summarized 25 years of research with the BDI and reported a range of the internal consistency across various subject populations between .73 and .92. The test-retest coefficients ranged from .48 to .86, depending on the time intervals for retesting and sample characteristics.

The Center for Epidemiology Studies Depression Scale (CES-D; Markush & Favero, 1973) is a self-report measure with each of the 20 items being rated on a four-point scale for severity of depressive symptoms over the previous week. It is used primarily as a screen for depressive symptoms in psychiatric, general medical, and community samples. The scale has good internal consistency, test-retest reliability and validity.

The Symptom Checklist-90R (SCL-90R; Derogatis, 1977) is a widely used self-report measure of psychiatric symptoms. The response categories of the SCL-90R inquire into how much distress each symptom has caused. The SCL-90R yields nine factorially-derived scores as well as three summary scores. The symptom scales are: Somatization, Obsessive-Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism. The SCL-90R has norms from over 900 non-patients and consists of 90 items with 6-13 items per scale. The scales have good internal consistency and test-retest reliability.

The Social Adjustment Scale (SAS) is a measure of social functioning (Weissman & Bothwell, 1976). It surveys seven instrumental and expressive role performance areas: work; social and leisure activities; relationships with extended family; marital role and parental role; family unit role and economic role. The items are rated on a five-point scale, where higher scores indicate more impairment. The reported psychometric properties have been adequate for this measure.

Finally, the Outcome Questionnaire is a 45-item self-report measure of behavioral health that includes scales for Symptom Distress, Interpersonal Relations, and Social Role (OQ-45, Lambert et al., 1996). The internal consistency of the Symptom Distress and Total Score are excellent (above .90). The internal consistency of the Interpersonal Relations and Social Role Scales are acceptable (.70–.75).

Polaris-MH Clinical Utility

A very large number of case studies must be constructed if the Pragmatic Case Study (PCS) model is to become a viable alternative to traditional research models. This could be best accomplished if de-identified, quantitative case study data for patients treated in clinical (i.e., non-research) settings were made available to PCS researchers. Ideally, the assessments that provide patient and outcomes data would be integrated into routine treatment, so that the cases would be representative and varied and the database could grow quickly. But integration will occur only if the assessments are perceived by the clinic's stakeholders -- patients, therapists, clinical supervisors, administrators, care managers and payers -- as useful to those stakeholders' various goals. Polaris-MH is thus designed to be useful to all of these stakeholders (Grissom, 2002), providing numerous features that facilitate integration into routine treatment.

Treatment Integration – Polaris-MH in Everyday Use

The full Polaris-MH system, as outlined in Table 1, is designed for use throughout treatment. Upon admission, the patient completes an initial assessment using the numeric keys on a computer. Computer literacy is not required. A clinical report is printed and available immediately for review with the patient (see sample in Figure 1). The report summarizes information about the patient's treatment history, motivation and assets for treatment (strengths) as well as clinical information – severity of the patient symptoms associated with each of the disorders listed in Table 1, and their Subjective Well Being, Symptoms, Functional Disability, and Behavioral Health Status scores. All scores are reported as percentiles, based upon norms for adult outpatients. A screen for inauthentic responding, based upon the amount of time that the patient required to complete the assessment and the number of inconsistent response patterns, alerts the clinician to potentially invalid assessments. Finally, Mental Health and Health Screens (see Table 1) for suicidality, dangerousness, chemical dependency, bipolar disorder and psychosis call the clinician's attention to potentially serious conditions.

The clinic may designate when the update assessments are to be done, and the system will keep track of each patient, printing lists (broken down by therapist) of patients who are due to complete the update assessment. The clinic may also specify the conditions (e.g., patient severity or length of time since last full update) that will determine whether a brief or full update assessment is administered. The system will automatically administer the correct assessment when the patient signs on. Update intervals of three or four weeks are often employed: the interval 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). However, the availability of the brief update assessment makes it feasible to administer at every session, which is operationally ideal.

An update report (see sample in Figure 2) is printed immediately after completion of the assessment, and is available for discussion during the treatment session. The update report indicates the patient's satisfaction with treatment, therapeutic bond, and compliance with medication (when applicable). Clinical data include trend lines for the BHS and Depression scale scores, showing the changes that have occurred since admission. The same data are presented for all other clinical scales in tabular form. Areas within the symptoms and functioning domains of continuing difficulty are listed, along with areas where significant improvement has been achieved. Screening data are reported as on the initial assessment.

Polaris-MH reports will be enhanced in early 2006 with the addition of information on expected treatment response. Figure 3 shows how the ETR curve is used clinically for a sample client. The ETR curve is generated on the basis of the following at initial assessment: (a) the patient's Behavioral Health Status score, and (b) the patient's characteristics, such as treatment history, motivation, and expectation for improvement. It is displayed on the clinical report, together with a Failure Boundary (see below). As update assessments are completed, the patient's progress (change in the Behavioral Health Status score) is added to the ETR curve and Failure Boundary. The ETR curve in effect indicates the rate and amount of improvement that would normally be achieved by patients with similar characteristics and a similar 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 the 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 (Lutz, Martinovich & Howard, 1999).

Determining the Statistical and Clinical Significance of Change.

In routine clinical use a therapist or care manager can monitor a patient's progress by noting the trend in patient severity scores on any of the Polaris-MH scales: Subjective Well Being, Symptoms and its seven subscales, Functional Disability and its three subscales, and the composite Behavioral Health Status score. Two questions are of particular importance when evaluating a change in a measure of the patient's clinical status: Is it likely that the change reflects improvement (or deterioration) in the patient's condition, or might it simply reflect measurement error? Is the change clinically, as well as statistically, meaningful?

The Reliable Change Index (RCI) introduced by Jacobson and Truax (1991) can be used to answer the first question. For example, the RCI for a change ($BHS^2 - BHS^1$) in a patient's BHS score measured at two points during treatment is:

$$RCI = (BHS^2 - BHS^1)/SDdif = 1.96$$

To have confidence (at the $p<.05$ level) that the difference in scores (BHS^2-BHS^1) reflects an actual change in the patient's condition, the difference would have to be at least 1.96x(SDdiff), where SDdiff is a function of the pooled variance (SD) and the reliability (r) of BHS:

$$SDdiff = \sqrt{2 \times (SD \times SQRT(1-r))^2}$$

In practice, the RCI is most readily calculated using T-scores, which are stored for each scale in the system's database and easily exported into the Excel or SPSS software programs. Since the SD for T-scores is always 10, the only parameter that must be obtained to evaluate RCI is the scale's reliability (r).

Generally, the clinician is most often concerned with the question: "Can I have confidence that the patient's improved score suggests actual clinical improvement?" Table 3 shows the change in T-scores required (one-tailed test) for 90% confidence ($p<.10$) and 95% confidence ($p<.05$), for each scale.

For any change in a scale score, the greater the scale's reliability (r), the less likely it is that the change is due to measurement error. Internal consistency reliabilities for each scale are provided in Table 1. Internal consistency reliability is the optimal reliability estimate for transient internal states. Internal states are not subject to observation, therefore, ratings by anyone other than the individual are impossible. Transient experience can change over relatively short periods of time so test-retest reliabilities can underestimate the reliability of the measure. Thus we use the internal consistency reliability as the reliability estimate (r) in calculating RCI. This differs from the approach used by Jacobson and Truax (1991), who recommended the use of test-retest reliability. However, the initial phase of the model of change utilizes remoralization, which is known to change rapidly as a result of events in the individual's life (e.g., affirmation, setting up an appointment, and/or a success at home or work). Test-retest reliability generally requires a period of one to two weeks in order to reduce the likelihood that the respondent simply remembers his/her previous response. Thus for the initial phase of response, test-retest reliability might underestimate the actual reliability of this relatively easily influenced subjective state. Therefore, instead of using different reliability estimates for different phases of the model, at this stage in the model development we chose the consistent use of internal consistency reliability for the Polaris-MH system. Internal consistency reliabilities appear comparable for symptom and functioning scales with all scales above 0.75 (Table 1). The impact of using different reliability estimates over different phases is an area for future investigation. Understanding these potentially complex relationships of reliability to phase could be particularly important for scales in which the expectation is that the majority of respondents move into non-clinical ranges of a

measure. If scores in the non-clinical range have restricted ranges, internal consistency may decrease as a statistical artifact.

Lueger, Martinovich, Anderson, Howard, Lutz, & Grissom (2001) have described how the second question – whether a patient's progress is clinically as well as statistically significant – can be addressed in one of two ways. The first involves classification of patients as treatment "responders", "non-responders" or "successes" depending upon whether the patient has shown reliable change ("responder") and/or a severity score that is more typical of non-patients than of patients ("success").

The cut-off score that best discriminated patient versus non-patient severity ranges for the scale studied by Lueger et al. was established by first assessing both adult outpatients (N=6,591) and a community sample (N=493), and then determining the level at which a patient's score is closer to the mean of the community sample (non-patients) than to that of the patient sample (Sperry et al., 1996, p.96). That scale, the Mental Health Index (MHI), was similar in construction and components to the Subjective Well Being, Symptoms, and Functional Disability scales of the Polaris-MH BHS scale (see Table 1). While BHS scores for a community sample have not been obtained, it is likely that the "normal range" for the BHS is similar to that of the MHI; it would be appropriate to consider patients to be in the normal range ("successes") if their BHS T-score were 60 or higher (percentile score 84 or higher).

There are methodological difficulties and significant expense involved in constructing a sample truly representative of non-patients. Polaris-MH provides for an alternative method for defining "treatment success," which we believe makes the collection of data from non-patients unnecessary.

The second method proposed by Leuger et al. capitalizes upon the predictive strength of the ETR. Progress during treatment, and the outcome of treatment, can be evaluated by comparing a patient's actual treatment response to the ETR. In effect, this method addresses the question: "Has the patient made as much progress as would be expected, after adjusting for initial severity and other patient characteristics related to outcome?" Note that this approach is methodologically more sophisticated than the first in that it controls for patient characteristics.

The ETR provides an evidence-based criterion for "success" that does not require collection of data for non-patient samples. Treatment "success" can be defined using the RCI criterion applied to the difference between the ETR score and the patient's actual BHS score. If the patient's score is at least as high as the ETR score (for the period of treatment that has been completed) less the measurement error in T-score points (see Table 3), the treatment is a "success." In sum, this approach is equivalent to defining treatment "success" as progress that is not statistically less favorable than empirically determined expected success, after adjusting for patient factors at intake.

Polaris-MH Features and PCS Requirements

The quantitative requirements for Pragmatic Case Study research include typing of patients upon admission, and measurement of clinical change during treatment. The patient characteristics that were included in the Polaris assessment were selected for their empirically based utility for case mix adjustment. They are therefore well suited to the typing requirement: the assessment covers characteristics that are believed to be related to treatment response. The clinical change measure component scales – Social Well Being, Symptoms, and Functional Disability – are psychometrically sound and grounded in the phase model.

Fishman has indicated how such a system might contribute to the PCS paradigm:

> By employing standardized input and outcome measures for each case…norms could be inductively established for superior, average and inferior outcomes relative to a particular type of case….If a [specific] case deviated from the average expected outcome …process information would help to differentiate whether the deviation from average goal attainment was due to exceptional therapy, or to special [circumstances]. In a parallel way, in instances of inferior outcomes, the process information would help to differentiate whether the deviation from average goal attainment was due to problematic therapy, or to special constraints, obstacles, and complexities of the case. In both instances, as the case database developed, there would be more and more cases of a particular type with "superior" or "inferior" outcomes, allowing for cross-case analyses of factors and themes to provide guidelines for improving the overall practice of therapy….(Fishman, 1999, 225-226)

To the extent that contextual information -- patient, setting, and treatment process information that the therapist/researcher believes to be relevant to the outcome of the case – can be standardized and rated by the therapist, the Polaris-MH Counselor questionnaire (see Table 1) could provide an effective tool to capture these data and link it to the case record.

Accumulation of individual case studies will allow for continual refinement of the predictive (ETR) models embedded in Polaris-MH, which in turn will enhance the PCS method of assigning cases to "superior," "average," and "inferior" outcomes categories.

In conclusion, we suggest that Polaris-MH can be a vehicle to overcome the most significant obstacle to the PCS model, the need for "a very large number of cases, which include both failures and successes of treatment…spread over thousands of therapists…" (Seligman, 2000, p.2).

ATTRACTIONS OF THE POLARIS-MH MODEL

There are two factors that could motivate "thousands of therapists" to use Polaris-MH in their practices. First, it provides the sort of outcomes data that is increasingly required by

managed care organizations (MCOs). Magellan Health Services, the world's largest MCO, began introducing the system to providers in Massachusetts in 2004. Kaiser Permanente has used a Polaris outcomes system designed for addictions medicine for several years in its southern California region. Second, since the system harvests the data from every treatment episode and provides it in a form (i.e., as an ETR) that can easily be used to guide treatment planning and management, Polaris-MH represents a highly effective and efficient form of evidence-based treatment.

It should be noted that Michael Lambert and his associates have developed a Pragmatic-Case-Study-based system with a similar logic and structure to Polaris (e.g., Lambert & Hawkins, 2004; Lambert, 2005; Lambert, Harmon, Slade, Whipple, & Hawkins, 2005). They have shown that a PCS system can enable therapists to improve treatment outcomes by reviewing assessment results with patients, both increasing the number of patients who achieved clinically significant and reliable change, and reducing the number of patients who experienced deterioration during treatment (Lambert, Hansen & Finch, 2001). We look forward to following and learning from the developing results of this model as we continue to maintain a constructive competition with it. We believe that having two, parallel PCS efforts will strength the contribution of each paradigm to the larger psychotherapy field.

The major competing paradigm for establishing accountability in psychotherapy practice is the "empirically supported treatment" or "EST" model (Chambless & Ollendick, 2001). Developed within the "treatment-focused research" tradition mentioned at the beginning of this paper, ESTs employ group efficacy studies to identify treatments that show superiority over control conditions (e.g., "treatment as usual" or "waiting list") for particular types of patients. While this approach has persuasive and thoughtful advocates in the scientific community (e.g., Nathan & Gorman, 2002), it is not without its critics (e.g., Henry, 1998; Wampold, 1997; Westen, Novotny, & Thompson-Brenner, 2004), and it is less clinician-friendly than the PCS-based, Polaris-MH system (Fishman, 2005; Peterson, 1991).

One of the advantages of the Polaris-MH approach is that while an EST may be an acceptable response to the question, "What treatment works for a patient with this diagnosis?", the Polaris-MH and ETR technology can answer the question: "Is this treatment working for this particular patient?" We suggest that the strongest form of evidence-based treatment is that which can be shown to be effective for a specific patient, when the patient's condition, treatment history and other characteristics are taken into consideration. Polaris-MH provides a tool that any therapist, using any therapeutic modality, can use to document the effectiveness of his/her services. We believe it will be embraced because it preserves the therapist's prerogative to employ whatever treatment strategy seems most appropriate, by enabling the therapist to provide evidence from many thousands of cases (i.e., patient progress in relation to ETR) that the treatment is effective.

While there is not space to discuss a systematic comparison of the EST model and a PCS model like Polaris-MH, each of these models contributes in its own, complementary way to

improving the practice of psychotherapy (Fishman, 1999, 2005). Moreover, having two contrasting models for pursuing the goal of evidence-based accountability in psychotherapy practice increases the opportunities for mental health delivery systems to provide effective services to their patients. This is the conclusion reached by a Presidential Task Force of the American Psychological Association (APA), which developed a model of "Evidence-Based Practice in Psychology" (EBPP; American Psychological Association, 2005). This model was approved a year ago as official policy by the organization. The EBPP model encourages both treatment-focused approaches to outcome studies, like EST research, and patient-focused approaches, like PCS research. In setting forth "some of the most pressing research needs" to fully achieve the goals of the EBPP paradigm, the Task Force Report includes the following, for which the Polaris-MH system is ideally suited:

- Developing well-normed measures that clinicians can use to quantify their diagnostic judgments, measure therapeutic progress over time, and assess the therapeutic process . . .

- Providing clinicians with real-time patient feedback to benchmark progress in treatment and [with] clinical support tools to adjust treatment as needed (American Psychological Association, 2005, p. 16).

Our team developing Polaris-MH believes it is important to locate its work within the broader field of psychotherapy research, and we find it very encouraging that the present thinking of the American Psychological Association is supportive of patient-focused research models like Polaris-MH.

REFERENCES

American Psychological Association (2005). Report of the 2005 Presidential Task Force on Evidence Based Practice. Washington, D.C.: Author. Retrieved August 2005 from http://www.apa.org/practice/ebpreport.pdf

Barkham, M., Stiles, W., & Shapiro, D. (1993). The shape of change in psychotherapy: Longitudinal assessment of personal problems. *Journal of Consulting and Clinical Psychology, 61,* 667-677.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years later. Clinical Psychology Review, 8, 77-100.

Bryk, A., & Raudenbush, S. (1992). *Hierarchical linear models: applications and data analysis methods.* Newbury Park: Sage Publications.

Chambless, D.L., & Ollendick, T.H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52,* 685-716.

Derogatis, R. L. (1977). SCL-90: Administration and procedures manual-I for the r(evised) version. Baltimore: Clinical Psychometrics Research.

Dupuy, H. J. (1977). A current validational study of the NCHS general well-being schedule (DHEW Publication No. HRA 78-1347). Hyattsville, MD: National Center for Health Statistics, US Department of Health, Education, and Welfare.

Fishman, D. (1999). *The case for pragmatic psychology.* New York: NYU Press.

Fishman, D. (2000). Transcending the efficacy versus effectiveness research debate: Proposal for a new, electronic "Journal of Pragmatic Case Studies." *Prevention & Treatment, 3,* Article 8. Available: http://journals.apa.org/prevention/volume3/pre0030008a.html

Fishman, D. B. (2005). Editor's Introduction to PCSP - From single case to database: A new method for enhancing psychotherapy practice. *Pragmatic Case Studies in Psychotherapy* [Online], Vol. 1(1), Article 2. Available: http://hdl.rutgers.edu/1782.1/pcsp1.1.47

Grissom, G. (2002). *Polaris-MH: Product overview and key features.* Fairless Hills, PA: Polaris Health Directions

Grissom, G., Lyons, 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. *Journal of Psychotherapy Research, 12,* 397-412.

Henry, W. (1998). Science, politics and the politics of science: The use and misuse of empirically validated treatments. *Psychotherapy Research, 8,* 126-140.

Horowitz, L., Rosenberg, S., Baer, B., Ureno, G. & Villasenor, V. (1988). Inventory of personal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56,* 885-892.

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.

Jacobson, N.S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Kadera, S., Lambert, M. & Andrews, A. (1996). How much therapy is really enough: A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice and Research, 5,* 132-151.

Kopta, S.M., Howard, K.I., Lowry, J.L. & Beutler,L.E. (1994). Patterns of symptomatic recovery in time-unlimited psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 1009-1016.

Krause, M., Howard, K., & Lutz, W. (1998). Exploring individual change. *Journal of Consulting and Clinical Psychology, 66,* 838-845.

Lambert, M.J. (2005). Emerging methods for providing clinicians with timely feedback on treatment effectiveness: An introduction. *Journal of Clinical Psychology, 61,* 41-144.

Lambert, M.J., Hansen, N., & Finch, A. (2001). Patient focused research: using treatment outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology,* 69, 159-172.

Lambert, M. J., Hansen, N. B., Umpress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W. (1996). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire)*. Stevenson, MD: American Professional Credentialing Services LLC.

Lambert, M.J., Harmon, C., Slade, K., Whipple, J.L., & Hawkins, E.J. (2005). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology*, *61*, 165-174.

Lambert, M., & Hawkins, E.J. (2004). Measuring Outcome in Professional Practice: Considerations in Selecting and Using Brief Outcome Instruments. Professional Psychology: Research & Practice. 35(5), Oct 2004, 492-499.

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.

Maling, M., Gurtman, M. & Howard, K. (1995). The response of interpersonal problems to varying doses of psychotherapy. *Psychotherapy Research, 5*, 63-75.

Markush, R. E., & Favero, R. V. (1973). Epidemiologic assessment of stressful life events, depressed mood, and psychophysiological symptoms: A preliminary report. In B.S. Dohrenwend (Ed.), *Stressful life events: Their nature and effects* (pp. 171-190). New York: Wiley.

Martinovich, Z. (1998). *Evaluating a phase model of psychotherapy outcome: an application of hierarchical logistic modeling.* Unpublished doctoral dissertation. Northwestern University.

Nathan, P.E., & Gorman, J.M. (2002). *A guide to treatments that work, 2nd ed.* New York: Oxford University Press.

Newman, F.L., & Howard, K.I. (1986). Therapeutic effort, treatment outcome, and national health policy. *American Psychologist*, *41*, 181-187.

Peterson, D. (1991). Connection and disconnection of research and practice in the education of professional psychologists. *American Psychologist,* 40, 441-451.

Pilkonis, P. & Frank, E. (1988). Personality pathology in recurrent depression: Nature, prevalence and relationship to treatment response. *American Journal of Psychiatry, 145*, 435-441.

Seligman, 1995, effectiveness Seligman, M.E.P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist*, *50*, 965- 974.

Seligman, M. (2000). Comment on Fishman's "Transcending the efficacy versus effectiveness reaserach debate." *Prevention & Treatment,* Vol 3, Article 11, posted may 3, 2000. Available: http://journals.apa.org/prevention/volume3/pre0030011c.html

Sperry, L., Brill, P., Howard, K. I., & Grissom, G. (1996). *Treatment outcomes in psychotherapy and psychiatric interventions.* New York: Bruner/ Mazel.

Wampold, B. (1997). Methodological problems in identifying efficacious psychotherapies. *Psychotherapy Research,* 7, 21-43.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*, 219-235.

Westen, D., Novotny, C. & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631-663.

Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. Archives of General Psychiatry, 33, 1111-1115.

Table 1 – Polaris-MH Questionnaires and Scales

SCALES			PHASE IN HOWARD MODEL	QUESTIONNAIRES			
Scale or Item Cluster Name	Number of Items	Alpha*		Patient Intake	Patient Update	Brief Patient Update	Counselor
Subjective well Being (SWB)	3	.86	Remorali-zation	X	X		
Symptoms (S)	27	.93	Symptom remedia-	X	X		
a. Depression	8	.88	tion	X	X	X	
b. Anxiety	3	.84		X	X		
c. Phobia	3	.81		X	X		
d. Obsessive-Compulsive	3	.79		X	X		
e. Somatization	4	.76		X	X		
f. Post Traumatic Stress Disorder	3	.75		X	X		
g. Panic Disorder	3	.81		X	X		
Functional Disability (FD)	13	.81	Rehabilita-tion	X	X		
a. Social	4	.91		X	X		
b. Vocational	4	.89		X	X		
c. Personal	4	.87		X	X		
Behavioral Health Status (BHS: sum of SWB, S, and FD above)	43	.83	Patient's total response to treatment.	X	X	Estimated by 19 of the BHS items)	
Demographic and Treatment History	12	-	Severity of initial condition variables	X			
Mental Health and Health Screens, e.g., for general health problems, substance abuse, psychosis, bipolar disorder, and suicidality	18	-	Severity of initial condition variables	X	11 of the 18 items		
Assessment of Patient's Treatment Strengths: resilience, meaning or purpose in life	11	-	Severity of initial condition variables	X			

Table 1 – Polaris-MH Questionnaires and Scales, continued

SCALES			PHASE IN HOWARD MODEL	QUESTIONNAIRES			
Scale Name	Number of Items	Alpha*		Patient Intake	Patient Update	Brief Patient Update	Counselor
Treatment Motivation	6	--	Severity of Initial condition variables	X			
Treatment Progress/ Satisfaction	2	--	Monitoring of treatment		X	1 of the 2 items	
Therapeutic Bond	4	--	Monitoring of treatment		X		
Counselor Questionnaire: DSM codes, case severity indicators, an assessment of the patient's motivation for treatment, severity of symptoms and functional disability, progress, and prognosis.	19	--	Utilization review; clinical supervision and/or to predict a patient's response to treatment				X
Inauthentic Responding (see text)	--	--	--	X	X	X	

*Internal consistency reliability of the scale. These reliabilities are based on representative outpatient samples ranging from 200 to 400, as reported in Grissom, Lyons, & Lutz, 2002. Alpha is shown for all scale scores.

Table 2 -- Concurrent Validity of Selected Scales

(N=200)

Polaris-MH Scale	Comparison Scale	Correlation (r)*
Subjective Well-Being	General Well-Being Scale (GWS)	.72
Subjective Well-Being	Positive and Negative Affect Scale (PANAS)	.69
Depression	Beck Depression Inventory (BDI)	.80
Depression	Center for Epidemiology Studies Depression Scale (CES-D)	.87
Symptoms (Total)	Symptom Checklist-90R (SCL-90R)	.88
Functional Disability	Social Adjustment Scale(SAS)	.58
Behavioral Health Status	Outcome Questionnaire (OQ-45) (N=126)	.87

* All are significant at $p < .001$

Table 3 -- Change in T-Scores Required for Statistically Significant Improvement

Scale	p<.10*	p<.05*
Subjective well Being	7	9
Symptoms (total)	4	5
Depression	7	9
Anxiety	8	10
Phobia	8	11
Obsessive-Compulsive	9	11
Somatization	9	12
Post Traumatic Stress Disorder	10	12
Panic Disorder	8	11
Functional Disability (total)	8	11
Social	6	7
Vocational	6	8
Personal	7	9
Behavioral Health Status	8	10

* One-tailed probabilities

Figure 1: Polaris-MH Initial Assessment Report

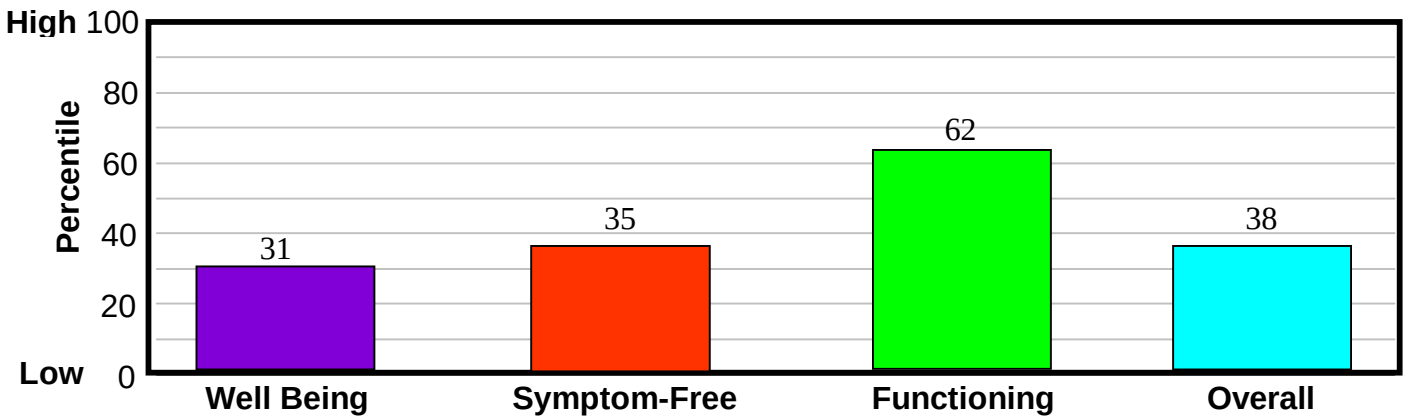
NOTE: Higher Percentile Scores Indicate Healthier Functioning

Polaris HEALTH DIRECTIONS

Client Name: Christopher Smith **Date of Birth**: 6/11/1971 **Assessment Date** 10/3/2003
Counselor: Counselor **Time to Complete:** 7:50

Screens	Negative	Positive
Inauthentic responding	No	
Dangerousness to self or others		See page 3
Evidence of possible serious disturbance	No	
Evidence of possible chemical dependency	No	
Psychological Treatment History	**Absent**	**Present**
Previously hospitalized	No	
Most recent hospitalization	No	
Previous counseling or psychotherapy		4 or more episodes
Total time in counseling or psychotherapy		6 months - 1 year
Ease of relating to prior therapist(s)		Easy
Benefited from prior treatment		Moderately
Motivation and Attitude	Agree	**Disagree**
Needs professional help	Agree	
Has a lot to lose	Agree	
Confident that treatment can help		Slightly disagree
Problems caused by other people		Disagree
Treatment will be a hardship (expense, time, etc)	Slightly agree	
Who encouraged client to seek treatment?	Friends	

How Does The Client Compare With Other People In Treatment?



Polaris-MH Initial Assessment Report

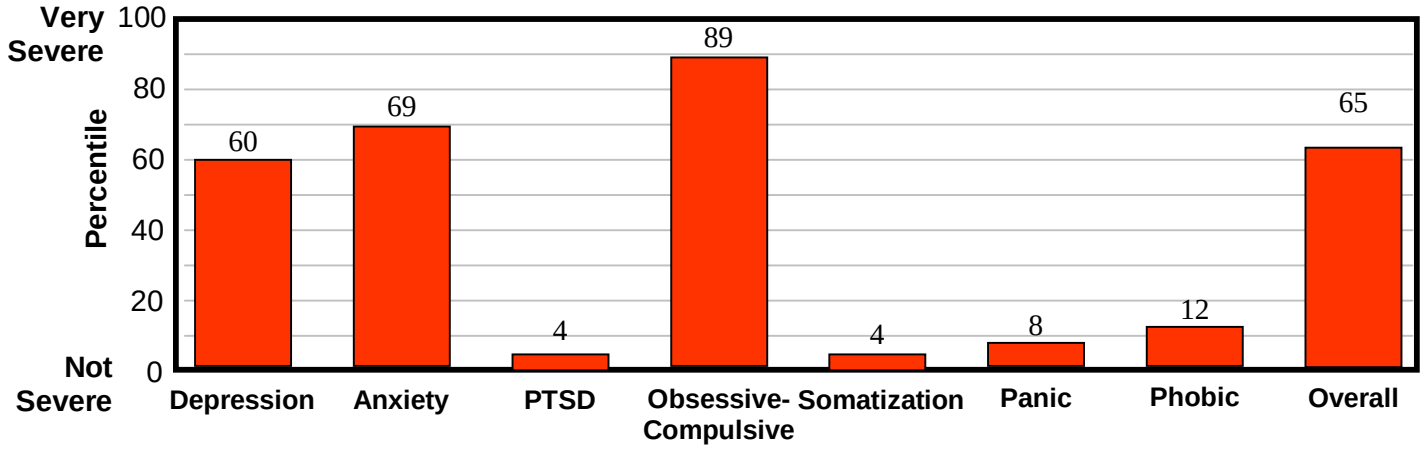
Polaris HEALTH DIRECTIONS

ristopher Smith
unselor 1

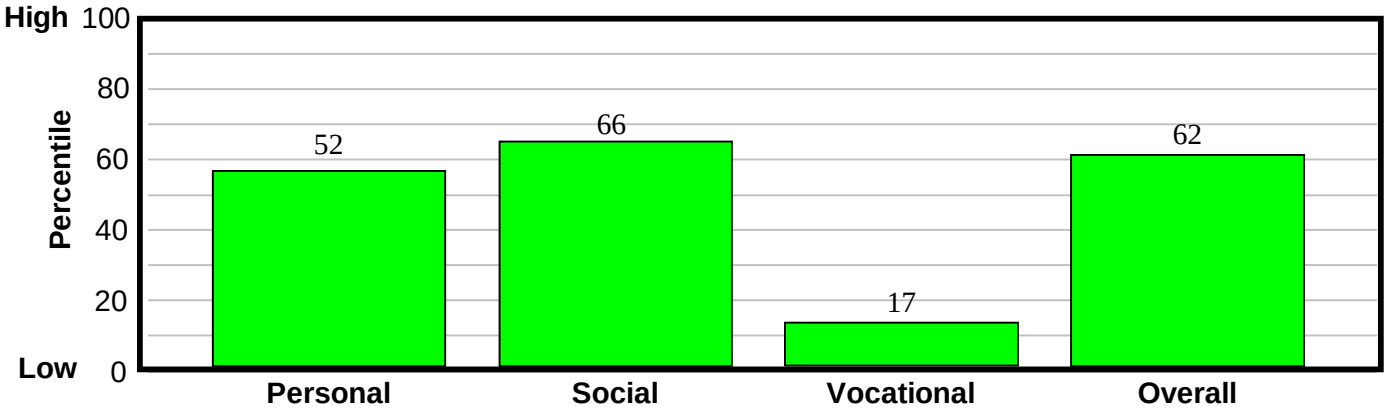
Date of Birth: 6/11/1971
Time to Complete: 7:50

Assessment Date 10/3/2003

Symptoms In Comparison With Other People In Treatment



Functioning In Comparison With Other People In Treatment



Treatment Assets/Strengths	Agree	Disagree
People in my life ***who love me*** very much	**Slightly agree**	
People in my life ***I love*** very much	**Agree**	
Find strength in religious/spiritual practices		**Strongly disagree**
When I have problems I go to people who can help me	**Strongly agree**	
Can think of more than one way to deal with a problem		**Disagree**
Able to bounce back when things go wrong		**Disagree**
Daily routine provides opportunities to do meaningful things		**Disagree**
Usually remain hopeful in the face of hardship		**Strongly disagree**
Something good can come out of my negative experiences		**Slightly disagree**
There is little purpose or meaning to my life	**Strongly agree**	
I have more difficulty than most people adapting to change	**Strongly agree**	

Polaris HEALTH DIRECTIONS

Polaris-MH Initial Assessment Report

Client Name: Christopher Smith	**Date of Birth**: 6/11/1953	**Assessment Date** 10/3/2003
Counselor: Counselor 1	**Time to Complete:** 7:50	

Information Relating to Positive Screen(s)

Danger to self or others

The client reports feeling like **harming himself / herself** during the past month.

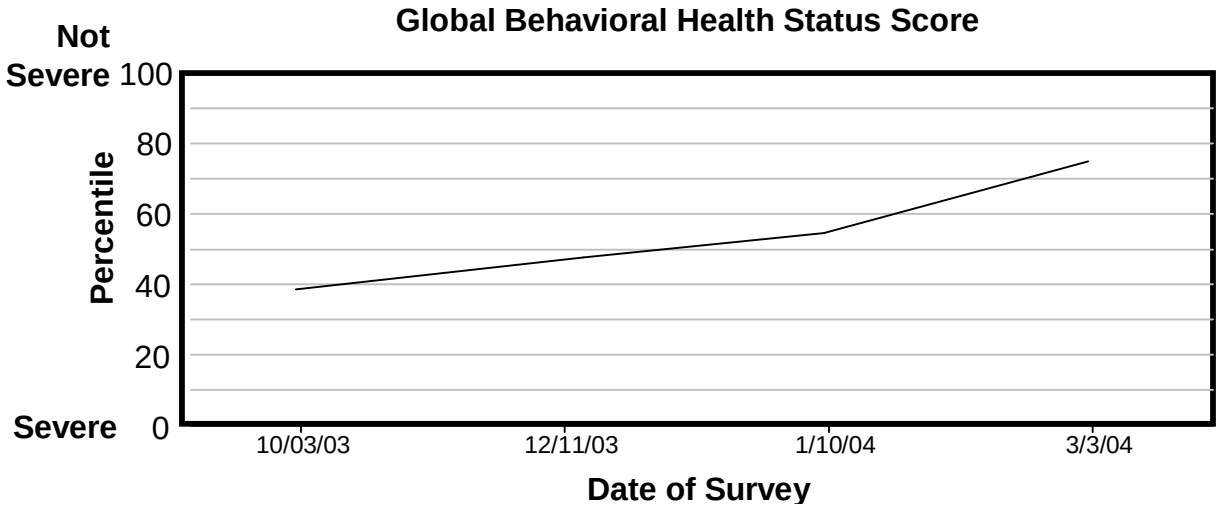
The client reports thoughts of **ending his / her life** during the past month.

Polaris HEALTH DIRECTIONS

Figure 2: Polaris-MH Client Progress Report

Client Name: Christopher Smith **Date of Birth**: 6/11/1971 **Assessment Date** 3/3/2004
Counselor: Counselor **Time to Complete:** 8:32

Screens	Negative	Positive
Inauthentic responding	No	
Dangerousness to self or others	No	
Evidence of possible serious disturbance	No	
Evidence of possible chemical dependency	No	
Therapeutic Bond -- Satisfaction	Agree	**Disagree**
Therapist understands how client thinks/feels	Mostly	
Client feels able to talk about what is on mind	Very much	
Client feels accepted and respected by therapist	Very well	
Therapist has been helpful	Very helpful	
Client assessment of progress so far	Well	
Satisfaction with treatment	Somewhat satisfied	
Medication: Client **has** a CNS med prescription		Yes
Is the client taking the med?		Yes
Are side effects a problem?		Somewhat helpful
How much is the med helping?		

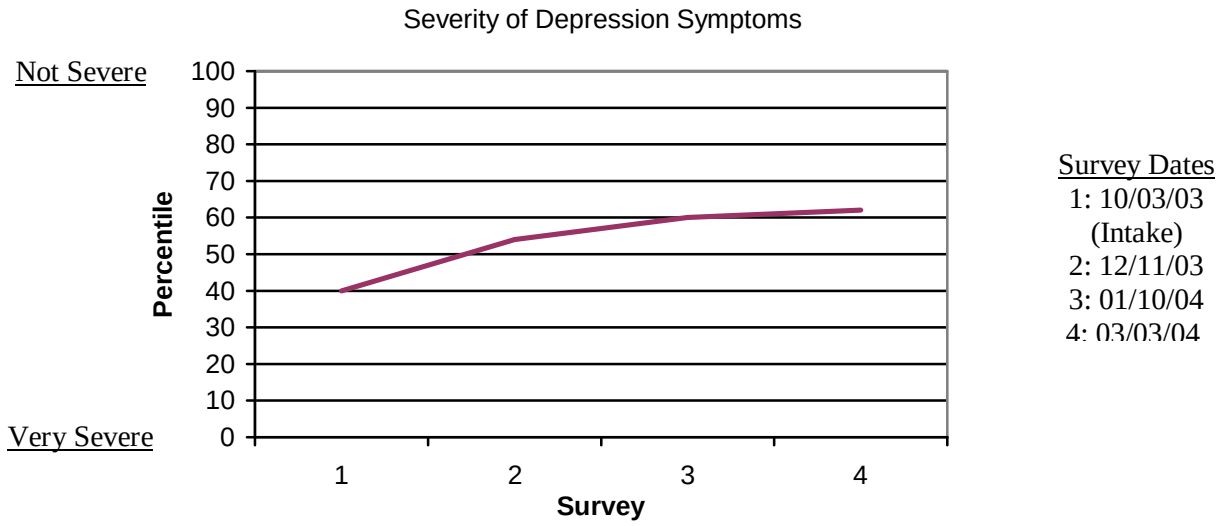


Information Relating to Positive Screen(s)
N\A

Polaris-MH Client Progress Report

Client: ristopher Smith **Date of Birth**: 6/11/1971 **Assessment Date** 3/3/2004
Counselor: Counselor **Time to Complete:** 8:32

Has the severity of depression changed?



How have the client's symptoms, functioning and feelings of well being changed during therapy?

Percentile Score: High is Favorable

Scale	Intake 10/03/03	Previous 01/10/04	Current 03/03/04	Change Since Intake
Subjective Well Being	31	60	63	32
Symptom-Free	35	44	78	43
Depression	40	60	62	22
Anxiety	31	41	61	30
Phobia	88	88	88	0
Obsessive-Compulsive	11	28	53	42
Somatization	96	96	96	0
Panic	92	92	92	0
PTSD	96	96	96	0
Functioning	62	68	80	18
Personal	52	74	74	22
Social	66	70	72	6
Vocational	17	33	20	3
Behavioral Health Status (Global Score)	38	52	77	39

Polaris-MH Client Progress Report

Client Name: Christopher Smith **Date of Birth**: 6/11/1971 **Assessment Date** 3/3/2004
Counselor: Counselor **Time to Complete:** 8:32

In what areas has there been marked change?

Areas Significantly Improved	**Intake**	**Current**
Trouble concentrating	Almost all the time	Some of the time
Felt tense or anxious	Almost all the time	Some of the time
Feeling hopeless/pessimistic	Often	Never or rarely

Areas That Are Significantly Worse	**Intake**	**Current**
Getting along with co-workers	Fairly well	Very poorly

What are the current areas of serious concern?

Getting along with co-workers

Getting along with supervisor

Figure 3: Sample Behavioral Health Status (BHS) Scores Over Time