

An Evaluation of Alaska Crossings: Comparison of the Client Status Review and the Youth Outcome Questionnaire

Keith C. Russell, Ph.D.

Western Washington University

Harold L. (Lee) Gillis, Ph.D.

Georgia College & State University

Jerrie Dee Harvey, LPC

Alaska Crossings

Abstract

Alaska Crossings (Crossings) is a 63-day open enrollment wilderness treatment program with a solid client base in predominantly Alaskan communities. The goal of this study was to provide quality, useful, and reliable evaluation data of the Crossings Program using the Youth-Outcome Questionnaire (Y-OQ), an alternative outcome monitoring system which is compared to the performance management system developed by the Alaska Division of Human and Social Service's Behavioral Health Division called the Client Status Review (CSR). Several objectives guided this evaluation project, including: a) analyze existing outcome data and other records from current monitoring or outcome evaluations to inform future outcome monitoring efforts; b) use the findings of this analysis to inform key stakeholders and provide recommendations for future outcome monitoring efforts; and c) make recommendations to key stakeholders as to the relative advantages and disadvantages of their utilization of the Client Status Review (CSR) and the Youth Outcome Questionnaire (30.2), including instrument sensitivity to change, outcome relevance given client needs and treatment protocol, and other identified goals. The results showed that when clients entered treatment, their scores averaged 28.76, which was statistically similar to juvenile justice samples reported by Burlingame et al. (2005) and to scores reported by Beckstead et al. (2015) on a sample of Native American/ Native Alaskan youth in residential treatment. As time progressed, average scores dropped to 14.97 at discharge, which indicated statistically significant ($t(64) = -8.847, p < .001$) and clinically significant improvement during this time period (a drop of greater than 10 points in scores). It was concluded that the CSR is a useful tool for assessment purposes and to assess treatment satisfaction, but the Y-OQ was best for tracking treatment outcomes.

Keywords: Outdoor Behavioral Healthcare, Youth-Outcome Questionnaire, Native Alaskan Youth.

AN EVALUATION OF ALASKA CROSSINGS

Alaska Crossings (Crossings) is a 63-day open enrollment wilderness treatment program with a solid client base in predominantly Alaskan communities. Clients are referred to the program by a variety of sources which include: parents, school officials, mental health agencies, judicial systems including parole officers and judges, and word-of-mouth. Crossings works with adolescents (ages 12-17) who represent at least 17 different traditional Alaskan cultures. Expedition staff attend at least 30 days of annual training and must possess the necessary skills to serve the youth with whom they work. The wilderness tripping is conducted in rugged wilderness that presents formidable challenges in logistics and operations including ocean and river based canoe trips in very remote settings.

Crossings works with high-risk youth, focusing on social and emotional learning utilizing intermittent therapy delivered by masters level behavioral health clinicians in the field. The model utilizes psycho-educational groups and a “point-system” that rewards exceptional behavior tied to wilderness living that is tracked throughout the 63-day experience. Alaska Crossings is unique in that staff form a cohesive unit with clientele and work with the same group for 21-day blocks of time during the 63-day experience. The experience is broken into three parts, each staffed by a different guide team. With new guides, the client group is resupplied and typically delivered to a different field setting (e.g., open-ocean to river paddling). There are two types of concurrent groups at Crossings: 1) open programs, where clients rotate in and out of the group as they enter treatment, and 2) closed programs, where the peer group begins the experience and ends the experience together as a cohesive unit. Staff and leaders switch in and out of either type of group on a rotational basis.

This facilitates significant relationship building between the peer group and the staff, and provides the environment for social and emotional learning to occur. Daily groups are held that ask clients to be mindful of their behaviors and feelings, and to reflect on their contributions to the group and wilderness community living. Much of the treatment model reflects a learner-centered, skill building, sensitive approach to working with Native Alaskan youth, who are prone to certain stigmatized perceptions of clinical treatment approaches (Beckstead, Lambert, DuBose, & Linehan, 2015). That said, each client has an individual treatment plan developed by masters level clinicians, that addresses treatment designed to affect oppositional defiant disorder, conduct disorder, and other disorders associated with severely emotionally disturbed youth. Clinical staff also work with field staff throughout the experience and expend significant energy and resources communicating with families about the well-being of clients. The clinical staff is also responsible for communication with youth-care advocates to ensure smooth transitions to post-treatment environments and the development of aftercare plans.

This study had three specific aims: 1) to provide quality, useful, and reliable evaluation data of the Crossings Program using the Youth-Outcome Questionnaire (Y-OQ), 2) determine if Routine Outcome Monitoring (ROM), or the periodic assessment of a client’s progress while the client is in treatment,

AN EVALUATION OF ALASKA CROSSINGS

could be implemented in a wilderness treatment setting, and 3) to compare these results to the performance management system developed by the Alaska Division of Human and Social Service's Behavioral Health Division called the Client Status Review (CSR) (<http://dhss.alaska.gov/dbh/Pages/Performance%20Measures/Default.aspx>). The results were used to provide recommendations to Crossings' about the relative advantages and disadvantages of each instrument. A brief review of the literature on wilderness treatment programs like Crossings, especially as they relate to ROM, situates these findings in its broader context.

Literature Review

Wilderness treatment (WT) programs for adolescents are being increasingly utilized in the United States to treat a variety of disorders, and is gaining acceptance in the medical community as evidenced by increasing insurance coverage (Willie, 2017). Recent estimates suggest that more than 5,000 adolescents attend wilderness therapy programs in the U.S every year (Gass, Gillis, & Russell, 2012). Wilderness expedition models like Crossings use continuous backcountry travel in groups of 8-10 clients for up to 60-days (Bettman & Tucker, 2011). In a recent and comprehensive meta-analysis of WT outcomes, Bettman, Gillis, Speelman, Parry and Case (2016) found differences in the effect sizes regarding studies that reported the use of mental health practitioners when compared to studies that indicated the programs were run by what they termed "therapeutic staff" (not licensed therapists or counselors). The results showed stronger effect sizes relating to locus of control, behavioral measures, and interpersonal skills when a mental health practitioner was present, and when field staff were present, only self-esteem was found to have stronger effects. This is important in the context of Crossings, which only recently shifted from employing therapeutic staff to using licensed clinicians in the field in conjunction with therapeutic staff. The study also highlighted that the training, experience, and roles that masters or doctoral-level trained clinicians play in WT is rarely reported in published studies. It was concluded that WT research needs to more clearly articulate exactly what the therapists' roles are in each program, including how often visits are made to the field, how they are structured, and what types of psychotherapeutic models are being employed. In this study, licensed clinicians check in routinely with the therapeutic staff, work with families, and conduct weekly visits to groups for individual and group-based therapy. At each 21-day interval, a staff exchange takes place that is also facilitated by the clinician. It was decided that at this juncture, the clinician would facilitate a routine outcome monitoring process to assess client progress and to use the information for the individual and group sessions, and in communication with the in-coming and out-going staff groups. Though regarded as an evidence-based practice that shows promising results regardless of treatment model or type, few if any wilderness WT programs are utilizing the process.

AN EVALUATION OF ALASKA CROSSINGS

Routine Outcome Monitoring

Evidenced Based Practice in Psychology (EBPP) as outlined by the American Psychological Association (APA) is a core component of mental health care (Levant et al., 2006). Currently there is increasing call for the use of EBPP from various stake holders in mental health care services ranging from clients, to practitioners, through administrators and regulating boards (Lambert, 2007; Levant et al., 2006). The APA supports several research approaches, including the use of client reported outcomes to examine questions of treatment efficacy, effectiveness and clinical utilization (Levant et al., 2006). During the last 20 years the development of Routine Outcome Monitoring (ROM) has emerged as a method to assess all three of these levels of interest (Anker, Duncan, & Sparks, 2009; Boswell, Kraus, Miller, & Lambert, 2015; Howard, Moras, Brill, Martinovich, & Lutz, 1996).

ROM uses session-to-session or pre-determined periodic assessment of a client's progress to track outcomes and inform treatment decisions in a responsive and timely manner while the client is engaged in the treatment process (Boswell et al., 2015; Howard et al., 1996). ROM is especially attractive because it can be used to support psychotherapeutic practice across a range of treatment populations (Anker et al., 2009; Lambert, 2007). Studies are demonstrating its effectiveness in the context of several DSM-5 disorders and child and adolescent mental health care (ex. Anker et al., 2009; Carlier et al., 2012; Shimokawa, Lambert, & Smart, 2010; Sundet, 2012; Timimi, Tetley, Burgoine, & Walker, 2012). The application of ROM to child and adolescent populations is in its infancy; however, preliminary findings are replicating the benefits found in adult populations (Bickman, Kelley, Breda, Regina de Andrade, & Riemer, 2011; Timimi et al., 2012).

Routine outcome monitoring utilizes client reports of distress, collected on standardized measures, as feedback for clinicians detailing behavioral or symptomatic change (Carlier et al., 2012; Shimokawa et al., 2010). A further specification, resulting from the maturation of ROM, includes discussing self-report data with clients in the context of therapy sessions (Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004; Lambert, 2007; Shimokawa et al., 2010; Sundet, 2010). Also, in the case of ROM with child and adolescent populations, reports may come from, and be discussed with, a host of stake holders including parents, caregivers, case workers, teachers, clinicians, and young people themselves (Timimi et al., 2012). In 2005, the APA appointed the Presidential Task Force on Evidence-Based Practice in response to expanding need and expectation for effective integration of research into psychology practice (Levant et al., 2006). The aims of EBPP are to improve quality and cost-effectiveness of psychological services, while also increasing the accountability of providers (Lambert, 2007; Levant et al., 2006). The Task Force defined EPBB as "...the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences..." (Levant et al., 2006, p. 273). This definition parallels ROM; it is a synthesis of current research, clinical

AN EVALUATION OF ALASKA CROSSINGS

practice and patient values. This significant alignment emphasizes ROM's relevance to contemporary practice.

Lambert and colleagues have published extensively on ROM. While several ROM systems and approaches exist, their development of the OQ-45 and the Y-OQ and their associated algorithm are widely used in the field (Anker et al., 2009; Duncan & Shaw, 2012; Lambert & Shimokawa, 2011). Their work clearly outlines early ROM efforts where client outcomes were simply returned to clinicians and their evolution to using these client outcomes as conversation starters with clients. More recently, they have developed intervention supports based on ROM called Clinical Support Tools (CST) that serve as an automated warning system and decision making framework that can assist therapists in identifying clients that are at risk of treatment failure (Lambert, 2007; Shimokawa et al., 2010). Implementation of ROM in other contexts and alternative methods has strived to further validate its use and improve implementation methods, but has also uncovered significant barriers to its use and implementation (Boswell et al., 2015). These include financial and time burdens, administrative misuse resulting in therapist competition, clients completing assessments that are unreliable due to trying to please the therapist, as well as privacy considerations and ethics of data use (Boswell et al., 2015; Lambert, 2007; Sundet, 2013).

ROM systems are now standard practice in mental healthcare settings and have been shown to increase client outcomes and build strong therapeutic alliances with clinicians. Barriers exist to their implementation, but can be overcome with clinician and staff buy-in and strong administrative support, including automated systems like the Y-OQ from OQ Systems that allow clinicians to provide real-time feedback to clients and staff with automated systems. This instrument was adopted by Crossings in the summer of 2015 and is currently in use. This study provides initial results from the analysis of data that was collected between April and November 2015. Crossings Y-OQ data, and the subsequent ability to compare it to several other wilderness treatment studies that utilize the Y-OQ, make it a desirable system for Crossings to implement. This will be compared with the data that the CSR produces, which is limited to a pre-post design, and may not be as sensitive to change or relate directly to the treatment process and culture of Crossings.

Methods

Crossing participants enter the program from rural and urban Alaskan communities and are a population, based on demographic and other information, that would be considered at extreme risk due to socio-economic situations and the lack of appropriate external assets. These include a lack of family support, positive family communication, caring and supportive neighborhoods, parent involvement in schooling, and other community resources (see www.search-institute.org for list of external assets that support youth development, which is also corroborated with CSR data presented below). Clients consented to participation of the data gathering in their admissions process to the program.

AN EVALUATION OF ALASKA CROSSINGS

All data were collected between April 2015 and November 2015. Clients entering treatment were asked to complete the Y-OQ and the CSR. All data were collected by clinicians and staff at Crossings. The data were then stripped of any identifying information and sent, via a secured link, to the evaluator with no identifiers other than a randomly assigned code to ensure confidentiality and anonymity. Y-OQ data was gathered at intake, at three-week intervals during the course of the program, and at discharge from the program. Clients completed the CSR as part of their intake and discharge processes. Demographic data were collected via the CSR. Data were analyzed using the Statistical Package for the Social Sciences (SPSS). The first set of analyses focused on the demographic characteristics of the adolescent clients and included frequency distributions of age, gender, and other demographic variables of interest. The second set of data analyses focused on an assessment of outcomes from a client self-report rating, where admission and discharge scores were calculated and paired sample t-tests were conducted to examine statistical change.

The CSR was analyzed based on the broad domains reported above. Descriptive statistics were first run, and then where possible, paired sampled t-tests were used to examine change in the domain from pre- to post-treatment. Domains that were reasoned to reflect various dimensions of the Y-OQ were then compared to assess the consistency and relative usefulness of the data. Demographic information of the study participants was collected, including age, gender, and ethnicity. Items related to participants' overall perceptions of quality of life are addressed below when examining the descriptive statistics generated from the analysis of the CSR. A brief review of the development and validity of each instrument is warranted to orient the reader to the constructs and domains assessed.

The Youth-Outcome Questionnaire

The YOQ-30 was developed as a brief measure of severity of disturbance in mental health patients under 18 that can be used repeatedly and is sensitive to change (Wells et al., 1996). It provides a total score or global index of behavioral and emotional distress in a child/adolescent's life. The reliability of the YOQ-30 was tested using Cronbach's alpha. The YOQ-30 has a remarkably high internal consistency estimate of .96 across the three samples. Reliability was also high within community and patient samples. Critical items alert clinicians to potential high-risk behaviors (e.g. suicide, substance abuse) and other item analysis provides easy-to-use interpretive indices. A cut off score discriminates between the normal and dysfunctional range and a Reliable Change Index (RCI) is used to determine if the change during treatment is clinically significant. The sensitivity and specificity analysis for the YOQ-30 is based on the cut-off score of 29. ***A score of 29 or higher is in the clinical range; a score below 29 is in the non-clinical or normal range.*** This higher score reflects the tendency of adolescents to under-report their symptomatology as compared to parents and the finding that parents are better sources of data regarding objective behaviors, e.g. oppositional attitude, externalizing behaviors, school failures, etc. Adolescents are considered

AN EVALUATION OF ALASKA CROSSINGS

to be more accurate informants regarding their subjective states, e.g. moods, feelings, etc.

The RCI was derived to determine clinically significant change (Jacobson & Truax, 1991). In order for an individual's score to be considered to have changed reliably for either version of the Y-OQ, the RCI value must be 10 points or greater. The cutoff score and the RCI values enable clinicians to interpret the clinical significance of patient change in treatment. Thus, if a client's score has decreased by ten points or more over the course of treatment, then the change may be characterized as "clinically significant improvement." If the score has decreased by ten or more points and the client's total score is in the "normal range," then the client is considered to have "recovered," in addition to having improved. If a child's score increases by ten or more points, then the child's progress may be characterized as "deteriorated." Normative data on the YOQ-30.2 were drawn from several large-scale samples across the United States (Burlingame et al., 2005) (see Table 1).

Table 1

Self-report normative groups for the YOQ-30.2 total score

Variable Pairs	<i>N</i>	<i>M</i>	<i>SE</i>
Inpatient Treatment	435	68.1	.96
Outpatient	2,297	43.3	.46
Juvenile Justice	719	32.6	.76
Community	1,091	17.3	.43

Client Status Review

The CSR is a self-report instrument that collects information on a persons' quality of life (DBH, 2011). The CSR was first developed in 2001 when the Division of Substance Abuse and the Division of Mental Health were being integrated. A broad group of stakeholders recommended performance measures for the new service system, including the Alaska Screening Tool, the Client Status Review of Life Domains, and the Mental Health Statistics Improvement Program (MHSIP) Consumer Survey. The structure and logic of the CSR were, at the time, consistent with emerging national policy and planning, including the Center for Substance Abuse Treatment (CSAT) and the Center for Mental Health Services (CMHS). The goal was to have the CSR represent several required national outcome measures and federal agencies, like the Substance Abuse Mental Health Services Agency (SAMHSA), which included "quality of life" in their working definition of recovery for mental health and substance abuse populations.

AN EVALUATION OF ALASKA CROSSINGS

In 2011, the CSR was revised to improve the ability to assess change over time. Focus was placed on the scoring methodology and the language used to ask questions, the number (volume) of questions necessary in order to measure change, and alignment with national data requirements. Specific to the scales used to measure change, the original CSR lacked the sensitivity and range to measure change over time. Findings from the initial CSR had most respondents at a level that could be described as “functioning well,” resulting in a lack of sensitivity for the instrument to measure improvement at a later point in time. Analysis of the pilot study demonstrated that the modified scales were successful in resolving this deficiency. The current version of the CSR examines various domains that are reasoned to be reflective of quality of life. The four broad domains are: 1) health (physical, mental, substance use, harm and access to emergency services), 2) safety (legal involvement, domestic violence, and general safety), 3) productive activity (employment/school, other productive activities), and 4) living with dignity. Some researchers have questioned the validity of self-report assessment and the challenge of assessing one’s subjective quality of life. For example, Awad and Voruganti (2000) stated that “by definition, quality of life is a subjective construct that needs to include patients’ self-reports and their subjective judgment. As such, it requires a degree of cognitive ability. Traditionally, clinicians have been suspicious of subjective assessment by patients of treatment outcomes” (p. 178). Despite the concerns regarding the reliability of self-report measures, it is now widely agreed-upon that self-report measures for persons with serious behavioral health conditions are useful both clinically and in performance measurement (e.g. Carlier et al., 2012).

It is reported that the information from the CSR can be used in two ways: 1) as a supplemental screening device to be coupled with the information obtained in the Alaska Screening Tool (AST) to inform the assessment, and 2) as a baseline measure of a persons’ quality of life prior to an assessment and entry into services. This initial CSR can be compared with subsequent CSR’s to monitor change over time. The CSR becomes an outcome instrument that links the result of treatment with the treatment intervention. Examining this claim forms the central focus of this project and supports the rationale to compare the CSR as an outcome and monitoring tool for clients in the Crossings program with the Y-OQ.

Results

A total of 79 clients were included in the study, with an average age of 15.6 years. Clients averaged 58.3 days in treatment, with a minimum of 20 and a maximum of 77 days. It is unknown why some clients were discharged from treatment. Typical reasons identified in the literature from previous studies suggest that treatment was not a good fit for the client or the client became a risk to themselves or others (e.g., see Russell, 2008). The closed programs averaged 61.2 days in treatment ($n = 48$; 60.8%) and the open programs averaged 53.9 days ($n = 31$; 39.2%). The majority of clients were male (male = 70.9%; female = 20.1%) and identified as Native Alaskan, including Haida, Tlingit, Yupik, Inupiat,

AN EVALUATION OF ALASKA CROSSINGS

and Athabascan (48.1%). Other ethnicities identified were Caucasian (38%), Mixed/other (10.1%) or American Indian (3.8%). Demographic information is displayed in Table 2.

Table 2

Demographic information for clients in the Crossings program.

Demographic Information	<i>N</i>	<i>M</i>	<i>SD</i>
Age in years	79	15.61	1.50
Days in Treatment	79	58.34	11.58
Gender	<i>N</i>	%	
Male	56	70.9	
Female	23	29.1	
Total	79	100.0	
Ethnicity	<i>N</i>	%	
Caucasian	30	38.0	
Native Alaskan	38	48.1	
American Indian	3	3.8	
Mixed / Other	8	10.1	
Total	79	100.0	

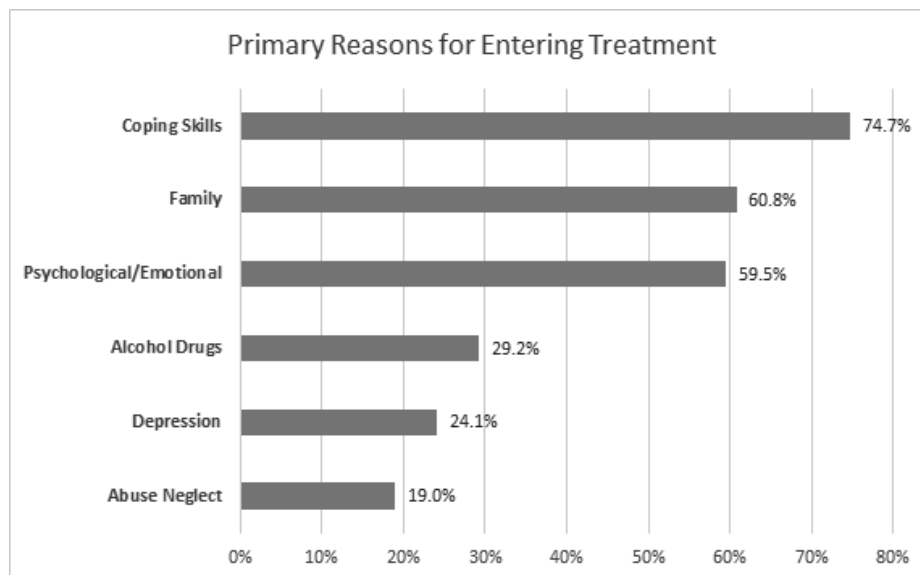
The primary reasons that the youth entered treatment are captured in Figure 1 (below). These primary reasons are from the perspective of the referring clinician, given the fact that each client that is referred to Crossings presented with serious emotional disturbance (SED). Children and youth with SED frequently require and receive services from a variety of agencies that apply different eligibility criteria. Beyond a common SED diagnosis, these young people are quite diverse in terms of their needs and strengths. Almost three-quarters of the clients were referred to Crossings due to troubles coping with daily roles and activities in their lives. This construct would reflect disruptive behavioral disorders, like Attention-Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD), and are referred to as such because affected children tend to disrupt people around them, including family members, school staff, and peers. The next most prevalent referring issue was family related (60.8%), stressing the troubled home and community situations in which many of these youth struggle. Their symptoms may cause family or community problems; or their symptoms may be aggravated by family problems. The variety of living situations in the group ranged from living in private residences with family or relatives, youth correctional facilities, or foster care situations. Almost equal (59.5%) were clients presenting with “psychological or emotional” issues, which could include a wide variety of

AN EVALUATION OF ALASKA CROSSINGS

mental disorders. The next three referral issues were alcohol and drugs (29.2%), depression (24.1%) and histories of physical and sexual abuse (19.0%). When integrated and examined as a whole, this sample reflects youth considered to be at extreme risk.

Figure 1

Primary reasons or issues that were recorded that led to a referral for treatment to the Crossings program



Youth-Outcome Questionnaire

The total score of the Y-OQ is the best index to track global change and has the highest reliability and validity when compared to any of the subscales. According to the manual (Burlingame et al., 2004) the Y-OQ “is meant to be used as a global index or summary score by which a clinician can quickly assess a child’s general functioning relative to normative populations as well as his or her progress in treatment” (p. 4). Figure 1 shows that for clients entering treatment, their scores averaged 28.76, which was statistically similar to juvenile justice samples reported by Burlingame et al. (2005) (see Table 1). Average scores dropped to 14.97 at discharge, which indicated statistically significant ($t(64) = -8.847, p < .001$) and clinically significant improvement during this time period (a drop of greater than 10 points in scores). At three-months, even amid limitations to the data due to common attrition in longitudinal assessments, the scores continued to trend well below community sample levels (indicated by a score of 17 or lower).

A paucity of data and research exists on Native American and Native Alaskan youth outcomes using the Y-OQ. The most comparable sample would be the study conducted by Beckstead, Lambert, DuBose, and Linehan (2015).

AN EVALUATION OF ALASKA CROSSINGS

The article examined dialectical behavior therapy (DBT) in a sample of Native American and Native Alaskan youth in a residential treatment setting in Alaska. The study used the Y-OQ SR full 64-item instrument to track changes in outcome (whereas the current study utilized the 30.2 instrument to reduce respondent burden and because of time and resource challenges in a backcountry wilderness environment). The 64-item Y-OQ has a total score of 256, whereas the Y-OQ 30.2 has a total score of 120 (herein referred to as the Y-OQ throughout). Using simple fractional comparisons, the samples can be compared to look for consistency or disparity of intake scores with similar sample populations. The average score at intake reported herein was 28.76, which was 23% of the total score. The effect sizes generated from the clinically and statistically significant change across broad demographics, was for males ($d = 1.31$), females ($d = 1.23$), Caucasian ($d = 1.25$) and Native American/Native Alaskan ($d = 1.26$) large and significant. These scores and effect sizes are comparable to Beckstead et al.'s (2015) sample, which reported an intake score of 50.78 (19% of the total score of 256) with a clinically and statistically significant effect size ($d=1.3$). In discussions with one the developers of the Y-OQ (Gary Burlingame, 2016, personal communication), these comparisons are sound and will be used in helping normalize the instrument to this population.

Figure 2

Y-OQ scores at admission, time 1, time 2, discharge, and 6-month follow-up

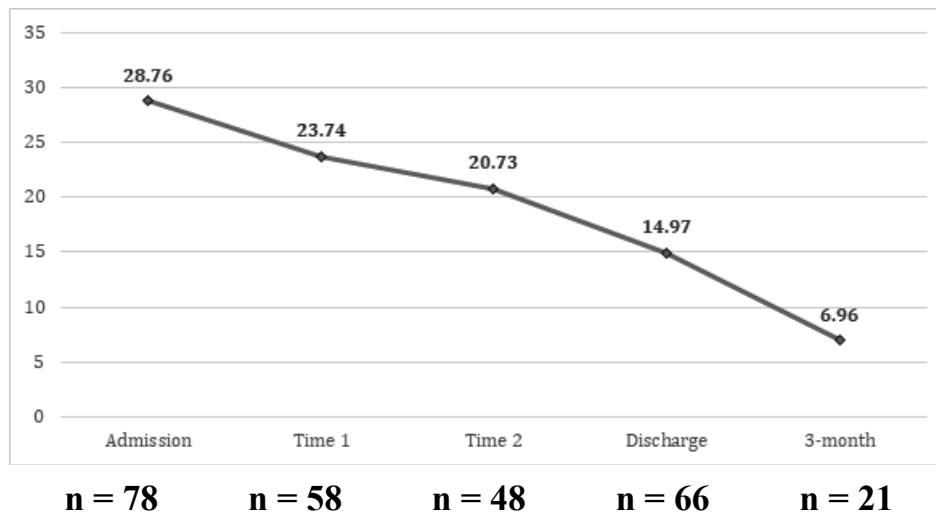


Table 3 shows the relative frequencies tabulated by using cut-scores as markers to get a better idea of the range of clients entering the Crossings program at admission. These scores indicate that almost half of the sample (40.5%) entered treatment with Y-OQ scores that ranged between juvenile justice and inpatient samples, with an average score of 41.56. Another 41.7% of clients entered treatment with scores that ranged from 17 - 31, in the range between the juvenile justice and outpatient treatment samples with an average score in this

AN EVALUATION OF ALASKA CROSSINGS

group of 23.58. The remaining 13 clients scored between 2 and 16, below the cut-score for a community sample. This data suggests that over 80% of clients are presenting with significant symptomology and were exhibiting poor functioning in their lives.

Table 3

Y-OQ scores at admission placed into normed categories of Community, Juvenile Justice and Inpatient samples

Groups	<i>N</i>	<i>%</i>
Scores from 2-16 (Community Sample = 17.3)	13	16.4%
Scores from 17-31 (Juvenile Justice = 32.6)	33	41.8%
Scores from 33-63 (Inpatient Sample = 68)	33	40.5%
Total	79	100%

Table 4 explores the relative differences in Y-OQ scores by gender and ethnicity. Few differences, if any exist between male and female intake scores, and when comparing Caucasian and Native Alaskan participants.

Table 4

Y-OQ score change by gender and ethnicity

	<i>Admission</i>	<i>Discharge</i>	<i>Mean Diff</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>df</i>	<i>d</i>
Gender								
Male	29.64	14.77	14.86	15.31	6.51	.001	43	1.31
Female	30.62	13.57	17.04	12.18	6.41	.001	21	1.23
Ethnicity								
Caucasian	29.76	14.10	15.65	15.11	5.57	.001	28	1.25
Native Alaskan	30.96	14.92	16.04	15.84	5.16	.001	26	1.26

Subscale analysis offers additional insight into client presenting issues as well as areas where treatment is making impacts on client well-being, especially when compared to other domains inherent in the Y-OQ. There are a total of six subscales that comprise the Y-OQ:

1. Somatic (S): This scale assesses change in somatic distress that the child or adolescent may be experiencing. Items address symptoms that are typical presentations, including headaches, dizziness, stomachaches, or troubles

AN EVALUATION OF ALASKA CROSSINGS

related to sleep (score range of 0-12).

2. **Social Isolation (SI):** The purpose of this scale is to assess changes in a child's or adolescent's social isolation. Items address whether or not the child or adolescent has friends, can't keep friends long, or feels as if no one likes him or her (score range of 0-8).
3. **Aggression (A):** This scale seeks to address changes in the level of aggression displayed by children or adolescents toward others. Although aggressiveness is also assessed in the CP scale below, aggressive content found in this scale is more related to physical violence. Items ask respondents whether or not the child or adolescent threatens others, bites, kicks, scratches, hits, or engages in physical fights with adults or peers (score range of 0-12).
4. **Conduct Problems (CP):** This scale assesses change in problematic behaviors that are socially related. Many of the items describe delinquent behaviors that are frequently the cause for bringing a child or adolescent into treatment. Items assess a child's or adolescent's propensity to destroy property, lie, steal, break rules, or disrespect others (score range of 0-24).
5. **Hyperactivity/Distractibility (HD):** This scale assesses change in the child's or adolescent's ability to organize tasks, complete assignments, concentrate and includes items measuring inattention, hyperactivity, and impulsivity. Although many of the items on this scale tap features of specific disorders (e.g., Attention Deficit Hyperactivity Disorder) the scale is not intended to be diagnostic but rather to track areas of change suggested by the literature, focus groups, and hospital records (score range of 0-12).
6. **Depression/Anxiety:** The purpose of this scale is to assess changes in depressive and/or anxiety-related symptoms in children or adolescents. Items assess the degree to which a child or adolescent feels sad, worries they can't get thoughts out of their head, considers suicide, feels withdrawn, can't trust others, or doesn't participate in activities that used to be fun. Since depression and anxiety are frequently correlated in assessment instruments (Burlingame et al., 1995) no attempt was made at differentiating these symptoms (score range of 0-24).

Table 5 highlights the subscales and dimensions, indicated by both the statistical significance (p) and the effect size (Cohen's d) where clients showed the highest symptomology and the most relative improvement as a result of treatment. The greatest gains were made in the Hyperactivity/Distractibility and the Conduct Problem subscales, both with large effect sizes. The subscales with small to medium effect sizes were Social Isolation and Aggression.

AN EVALUATION OF ALASKA CROSSINGS

Table 5

Pairwise t-tests for each subscale examining differences between pre and post treatment scores for all Crossings clients

Subscale	<i>Admission</i>	<i>Discharge</i>	<i>Mean Diff</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>df</i>	<i>d</i>
Somatic	4.77	2.78	1.98	3.06	5.221	.000	64	.76
Social Isolation	1.42	.78	.631	1.79	2.827	.006	64	.36
Aggression	2.11	.80	1.30	1.73	6.089	.000	64	.37
Conduct Problems	7.98	3.51	4.47	4.86	7.420	.000	64	1.08
Hyperactivity/Distractibility	5.55	2.48	3.07	2.85	8.698	.000	64	1.22
Depression/Anxiety	8.12	4.03	4.09	4.81	6.853	.000	64	.98

Examining Follow-up Y-OQ Scores as an Indicator of Treatment Effectiveness

To examine whether treatment outcomes noted above were maintained by clients after treatment, Y-OQ scores assessed at the three-month follow-up period were analyzed and compared to scores at discharge. Though attrition is always an issue, this sub-sample appears representative of the overall sample, as the scores at admission and discharge were similar for this group when compared to the overall sample scores (see Table 7). The challenges of collecting data at follow-up periods are discussed at length in the literature (see Russell, 2008), but in this sample, the issue is partly due to waiting for data to be returned by clients and their families. Table 7 shows that clients have continued to do well psychologically, emotionally, and behaviorally at the three-month follow-up period, as evidenced by Y-OQ scores.

Table 7

Scores at admissions, discharge and 3-month follow-up for the group that had been assessed at follow-up compared to the group with no assessment at 3-months to date

	<i>N</i>	<i>Admission</i>	<i>Discharge</i>	<i>Three-month Follow-up</i>
Three-month Follow-up Group	21	30.24	16.68	12.68
Remaining Sample	57	28.21	14.11	-

Client Status Review

Examining the claim that the CSR can be used as an outcome tracking instrument forms the central focus of this project and supports the rationale to compare the CSR as an outcome and monitoring tool for clients in the Crossings program with the Y-OQ. The CSR contains four domains that were used to

AN EVALUATION OF ALASKA CROSSINGS

examine the impact of treatment and program impact: 1) Health (Physical, Mental, Substance Use, Harm, and Access to Emergency Services), 2) Safety (Legal Involvement, Domestic Violence, and General Safety), 3) Productive Activity (Employment/School, Other Productive Activities), and 4) Living with Dignity. An initial issue with the CSR and the focus of questions that are being asked is the reference point that the clients have to use for the assessment device. Crossings is a 63-day residential treatment program. Therefore, if clients are being asked the degree to which they engaged in certain behaviors in the last 30 days, many of the questions seem inappropriate, out of context, or redundant. For example, questions 7-9, assessing the number of times they have used alcoholic beverages and/or marijuana, would appear impertinent to this population at this time. However, at the follow-up period they would be very relevant. Thus, assessing program impact for this construct (substance use) or others (legal involvement) appear problematic. Asking clients to complete these items could cause confusion or respondent burden and could affect reliability of the items and those subscales.

Of note in the interpretation of these findings (at least descriptively) is the reference point for clients when considering these questions. The CSR asks clients how many days in the last 30 days that he/she had felt, did, or acted in a certain manner. This would appear extremely challenging to answer this in a reliable manner. Another issue that was noted is how to interpret the rather vague and subjective meaning of the response options (e.g. “not good”). A third issue is that many of the questions are assessing multiple dimensions within the same question (e.g., Question 2 asks: “How many days during the past 30 days was your mental health (including depression and/or problems with emotions, behavior, or thinking) not good?”). For example, behavior may be fine, as a person may be doing what is asked and functioning well at work or school, but may not be doing well emotionally. How would a respondent reconcile these incongruences within these individual questions? Moreover, how accurate could these assessments be when recall is over the past 30 days? (see Bradburn, Rips, & Shevell, 1987; Hammersley, 1994; Shiffman, Stone, & Hufford, 2008).

CSR: Outcome Dimensions. Table 8 presents the results of a series of t-tests that explored the differences for each of the clients on the outcome dimensions described above. The variables were, in essence, the number of times that a specific construct had occurred over that time period (30 days for most constructs and seven days for healthy eating and exercise). Two constructs showed significant differences from admission to discharge: 1) the number of days a client reported that their mental health was “not good” over the last 30 days ($t(75) = 5.53, p < .001$), and 2) the number of days that clients indicated in the past 30 days that their physical health kept them from doing activities they would otherwise not be able to do ($t(75) = 1.98, p = .051$). All of the other constructs, which would be considered “outcome indicators” showed no significant change as a result of treatment. The mental health improvement outcome, though promisingly indicating statistically significant change, speaks little to the programmatic and clinical leadership team because of the issues with the construct noted above. For example, the question asks about “mental health,

AN EVALUATION OF ALASKA CROSSINGS

including depression, and/or problems with emotions, behaviors, or thinking.” An obvious question would be: In what area did the client make improvement, as they are in essence separate questions? The physical health question is more clear and illustrative for programs to be more certain that treatment did indeed improve physical health in this area.

Table 8

Pairwise t-test comparisons of health outcome variables contained in the CSR

Pair	<i>Mean</i>	<i>S.D</i>	<i>S.E. Mean</i>	<i>Lower</i>	<i>Upper</i>	<i>t</i>	<i>df</i>	<i>p</i>
Admission Physical Health – Discharge Physical Health (# past 30 days)	-.750	3.87	.444	-1.635	.135	-1.689	75	.095
Admission Mental Health – Discharge Mental Health (# past 30 days)	7.15	11.67	1.337	4.494	9.82	5.352	75	.001*
Admission Health Not Do – Discharge Physical Health Not Do (# past 30 days)	1.46	6.42	.737	-.007	2.92	1.982	75	.051*
<i>Admission Suicide Thoughts – Discharge Suicide Thoughts (# past 30 days)</i>	.145	2.07	.238	-.330	.619	.608	75	.545
<i>Admission EMS Service Use – Discharge EMS Service Use (# past 30 days)</i>	-.079	.560	.064	-.207	.049	-1.229	75	.223
Admission Exer. Past 7 Days – Discharge Exer. Past 7 Days	-1.90	2.96	.341	-2.58	-1.22	-5.588	75	.341
Admission Fruit Past 7 Days – Discharge Fruit Past 7 Days	-.947	4.05	.465	-1.87	-.021	-2.038	75	.465
Admission Veggies Past 7 Days – Discharge Fruit Past 7 Days	-2.18	4.10	.470	-3.12	-1.24	-4.643	75	.470

*Significantly different at the $p < .05$ level

Italics indicates variables where no change or negative change from admission to discharge indicates a positive outcome.

Table 9 shows outcomes associated with substance use and indicate that two areas showed statistically significant improvement as a result of treatment: 1) the number of days in the last 30 days in which the client consumed four or more drinks ($t(75) = 3.73, p < .001$), and 2) the number of days in the last 30 days in which the client used tobacco ($t(75) = .608, p < .000$). For the first item,

AN EVALUATION OF ALASKA CROSSINGS

it would be assumed that the client would have reduced their use of alcohol as a result of treatment because they were in a wilderness-based treatment program where no drugs and/or alcohol were available. In addition, it is curious as to why the other variables assessing alcohol and marijuana use were not consistent with this finding for the same reason. This may be due to measurement error and poor reliability of the items as noted above. Crossings participants are not allowed to use tobacco while in the program, raising additional concerns.

Table 9

CSR substance use indicator change from admission to discharge

Pair	Mean	S.D	S.E. Mean	Lower	Upper	t	df	p
Admission # of Days Drink Alcohol – Discharge # of Days Alcohol (Last 30 days)	.250	1.93	.222	-.192	.692	1.127	75	.263
Admission # of Days 4+ Drinks Alcohol – Discharge # of Days 4+ Drinks Alcohol (Last 30 days)	3.17	7.416	.851	1.47	4.86	3.728	75	.001
Admission Marijuana or Illegal Drug– Discharge Marijuana or Illegal Drug (Last 30 days)	.105	.531	.061	-.016	.227	1.728	75	.088
Admission Tobacco Use– Discharge Tobacco Use (Last 30 days)	.145	2.07	.238	-.330	.619	.608	75	.001
Admission Smoked 20 Day– Discharge Smoked 20 Day (Last 30 days)	-.079	.560	.064	-.207	.049	-1.229	75	.159

*Significantly different at the $p < .05$ level

CSR: Protective Factors and Treatment Services. Table 10 illustrates client satisfaction at admission and discharge with a variety of protective factors associated with their lives, including their living situation, ability to support their needs, their safety, their family and friends, and spirituality and meaning in life. Only one item was slightly below 5.0 (5.0 = “Satisfied”) and that was the item relating to their family situation. This is consistent with the fact that many of them have been referred to Crossings because of “family issues.” The highest rated item was for “safety in the home where they sleep,” which was 6.42 at admission and 6.26 at discharge (indicating “pleased”). In general, participants in this sample were satisfied or pleased with the various protective factors in their lives at admission and discharge, and the scores were very stable. None of the items were statistically different from admission to discharge indicating their

AN EVALUATION OF ALASKA CROSSINGS

relative satisfaction with these various protective factors.

Table 10

Client assessment of relative satisfaction with protective factors related to their lives at admission and discharge (based on a scale of 1-Terrible to 7-Delighted)

<i>Subscale</i>	<i>Admission</i>			<i>Discharge</i>		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Housing	79	5.66	1.29	76	5.69	1.37
Support basic needs	79	6.15	.988	76	6.11	.873
Safety in home	79	6.42	.886	76	6.25	1.04
Safety outside of home	79	5.72	1.31	76	5.86	1.16
People in life support them	79	5.78	1.47	76	6.00	1.07
Friendships	79	6.00	1.34	76	6.08	1.00
Family situation	79	4.95	1.57	76	5.25	1.39
Spirituality and meaning in life	79	5.29	1.23	76	5.77	1.08
Life in general	79	5.51	1.32	76	5.67	1.19

Table 11 shows the descriptive scores for items that asked the clients at discharge to rate their relative satisfaction with the treatment services they received and the degree to which the services helped them handle daily life, get along with other people, cope with issues in their lives, and their overall quality of life. All items were over 6.0, which indicated that the clients were relatively “pleased” with the services they received and they are better off as a result of the Crossing program and treatment process. The highest rated item was for the overall quality of life item ($M = 6.37$).

Conclusions and Comparisons of the Y-OQ and the CSR

There are several conclusions that can be made when comparing the Y-OQ and the CSR from the perspective of the assessment of client well-being and when using each instrument in tracking treatment outcome. This comparison should be taken in the context of the written purpose and intent for which each instrument is used. According to Burlingame et al. (2004), the Y-OQ measures the treatment progress for children and adolescents receiving any form of mental health treatment including psychoactive medications. In contrast to traditional diagnostic measures oriented to the measurement of psychopathology, the Y-OQ family of measures was specifically constructed to assess the occurrence of observed behavior or symptom change. The instrument is completed at admission into treatment to establish a baseline level of severity for symptom distress and at later sessions or time periods to track the child’s progress. The psychometric calculations from the normative database permit determination of the client’s symptom distress similarity at each measurement interval with several normative

AN EVALUATION OF ALASKA CROSSINGS

populations, including inpatient, outpatient, and community samples. Utilizing cut-scores and a RCI, clinicians, parents, guardians, clients and administrators can determine if and when the client’s symptom distress has entered the normal range. The information gleaned from the CSR can be used in two ways: 1) as a supplemental screening device to be coupled with the information obtained in the AST to inform the assessment, and 2) as a baseline measure of a persons’ quality of life prior to an assessment and entry into services. This initial CSR can be compared with subsequent CSR’s to monitor change over time, thus making the CSR an outcome instrument that links the result of treatment with the treatment intervention. The current version of the CSR examines various domains that are reasoned to be reflective of quality of life. The four broad domains are: 1) Health (Physical, Mental, Substance Use, Harm, and Access to Emergency Services), 2) Safety (Legal Involvement, Domestic Violence, and General Safety), 3) Productive Activity (Employment/School, Other Productive Activities), and 4) Living with Dignity.

Table 11

Client assessment of relative satisfaction with treatment services and the degree to which these services helped them in their lives (Based on a scale of 1-Terrible to 7-Delighted)

<i>Subscale</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Treated with respect	76	6.20	.817
Provided information about their rights	76	6.30	.766
Helped to choose their treatment goals	76	6.21	1.08
Could ask questions about treatment process	76	6.16	.910
Able to receive services that were needed	79	6.12	.909
Because of services received:			
Better able to handle daily life	76	6.30	.895
Getting along better with other people	76	6.32	.734
Able to cope when things go wrong	76	6.24	.814
Quality of life has improved	76	6.37	.846

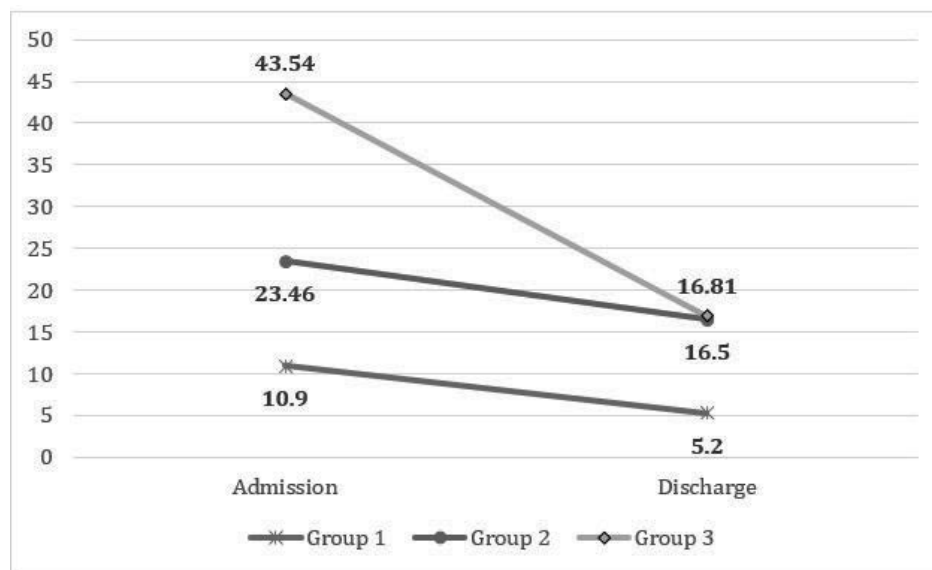
The Y-OQ was a useful global index and summary score by which staff and clinicians at Crossings could quickly assess a client’s general functioning relative to normative populations as well the client’s individual progress in treatment at various points in time throughout treatment. In this manner, staff are conducting ROM which improves clinical outcomes, increases client motivation, and is now considered to be a “best practice” in psychotherapy (Wompold, 2015). The data shows that when clients entered treatment, their scores averaged 28.76, which was statistically similar to juvenile justice samples reported by Burlingame et al. (2005) in Table 1 and similar in pathology to scores reported by Beckstead et al. (2015) on a sample of Native American/Native Alaskan youth in residential treatment. As time progressed, average scores dropped to 14.97 at discharge,

AN EVALUATION OF ALASKA CROSSINGS

which indicated statistically significant ($t(64) = -8.847, p < .001$) and clinically significant improvement during this time period (a drop of greater than 10 points in scores). By taking assessments at various time points during treatment, clinicians and staff can assess the degree to which treatment is having a positive or deleterious effect on the client. In addition, as staff change over from periodic shifts in the field, these assessments can be used to discuss client progress and other clinical notes that would help entering staff become more aware of clients' progress towards treatment goals.

The average Y-OQ score at admission was 28.76, and over 40% of the clients entering treatment had symptomology consistent with inpatient samples, indicated by scores between 33 and 63. This information is useful to communicate with staff in which areas or domains these clients are struggling and highlight the value of critical item indicators (e.g., suicidality and self-harm) and other elements of the Y-OQ that offer information that could be useful for clinicians and staff. When examining treatment progress for these three groups, the group reflecting inpatient sample scores (Group 3; $n = 33, m = 43.74$) improved to the same level (a Y-OQ score of ~ 16 at discharge) as the group indicated by the juvenile justice sample scores (Group 2; $n = 33, m = 23.46$). This is an important finding and conclusion in that the treatment process appears to be effective for more seriously symptomatic clients (see Figure 3).

Figure 3



Clients placed into categories of symptom severity based on Y-OQ score at admission illustrated by admission and discharge score change during treatment

There were no apparent differences across male or female clients nor the two ethnicity groups. Similar scores at admission and discharge were noted as well as similar change scores for each demographic. The greatest gains when examining Y-OQ subscales were in Hyperactivity/Distractibility (Effect size d

AN EVALUATION OF ALASKA CROSSINGS

= 1.22), Conduct Problems (Effect size $d = 1.08$), Depression/Anxiety (Effect size $d = .98$), and Somatic Issues (Effect size $d = .76$). These changes correspond to the reasons clients were referred to treatment (see Figure 1) suggesting that the outcomes indicated by the Y-OQ and the treatment process at Crossings are effectuating change in areas important to the clients, their referral sources and their families. The Y-OQ appears to be an effective tool in assessing the subtlety of these changes. Though the data were incomplete, three-month assessments indicated that clients continued to do well at follow-up and outcomes were being maintained from treatment. Follow-up assessment should be continued, with potential demographic information being collected to help interpret outcomes, including use of aftercare, living situation, and other moderators of client well-being post-program.

The CSR appears to be most effective as a pretreatment assessment tool for certain domains and the associated demographic information could be useful to help clinicians and staff better understand a client's history and the amount of protective factors present or not in their lives. For example, the questions asking about housing, safety, etc. can provide important context when considering treatment goals. That said, some of the data generated from this sample didn't seem to connect with the reasons that clients were in treatment. For Question 18: "Family situation," the lowest rated item at admission for this sample was $M = 4.95$, which is "Mixed" according to the reference point. Yet, for this sample, the majority (60.8%) were referred to treatment because of family issues. It would seem logical that the scores at admission would be lower than "Mixed," perhaps "Unhappy" or "Unsatisfied." Given this context, it should also be noted that all other scores at admission and discharge across all of the items in Question 18 were above 5.0, indicating satisfaction, with scores ranging from $M = 5.29$ at admission for "Spiritual and meaning in life" to $M = 6.42$ at admission for "Safety in home." It would be illustrative to compare these scores with other sample populations because at first glance, the descriptive statistics in these domains appear to be relatively standard for an adolescent population in a community sample. These items were not useful for any type of outcome indication, because it would be hard to infer that treatment could have altered these domains, and which was also corroborated by statistical analysis indicating no significant change from admission to discharge on any of the items.

Question 19 (Treatment satisfaction section) is an important and ethically responsible assessment domain, asking clients to assess the treatment services they have received only to be completed at discharge and reflect traditional treatment satisfaction instruments used in addictions and other mental health settings. The clients were, on average, pleased with their services and were treated with respect at Crossings. The items in Question 19 asking clients to assess the degree to which the services had helped them in their daily lives ("better able to handle life," "get along with other people," "cope when things go wrong," and "improved quality of life") correspond to a post-treatment only assessment, which have been shown to be unreliable in the literature due to "post-treatment or experience euphoria."

AN EVALUATION OF ALASKA CROSSINGS

The CSR data offered mixed results across the various domains it portends to assess and in its usefulness as an outcome tool. This conclusion stems from only the data that was generated from this assessment and is not meant to be generalized to other populations and settings. The first issue to note is the reference period that many of the questions were asking clients to consider. Asking an adolescent how many days in the last 30 days that he/she had felt, did, or acted in a certain manner seems to be problematic. It would seem to be extremely challenging for anyone to answer this question accurately. Typical reference periods to assess physical, mental, or emotional health cover seven-day periods at most. Another issue with the CSR as an assessment and outcome device is the rubric used to assess many of the key quality of life items. The respondent is asked to assess whether these items were “not good,” across questions related to physical or psychological health. Though all self-report assessments are subjective and open to consideration, having a single reference point of “not good” as opposed to the more reliable and frequently used multiple point (3-, 4-, 5-, 7-point) Likert scaling appears problematic.

Finally, another issue uncovered when examining the CSR is that many of the items or constructs that are being assessed are asking multiple questions within a single item, thus making the referent confusing or misleading. For example, one of the key questions asked the respondent to assess their mental health, described as “emotions, behaviors, or thinking; taking care of yourself, work, or recreation.” These are all very different dimensions of how someone may be doing and are typically subdivided to provide more accurate assessment of client well-being (see Y-OQ subscales). How a respondent would address these incongruences within these questions is unclear, lending their accuracy questionable, especially when we consider the issues raised above about “good” and “not good” and the recall period of 30 days.

In making recommendations to Crossings and other programs that utilize instrumentation or assessments that are required by their funders or stakeholders, some important issues could be considered when considering these findings. The CSR was useful in highlighting areas where clients were at risk and determining if treatment was warranted for this highly vulnerable population. The Y-OQ was a useful global index and outcome monitoring tool that is an important component of routine outcome monitoring, which is quickly becoming the standard of care (see Wompold, 2015) and should be continued. Continuing the routine outcome monitoring at Crossings can increase therapeutic communication between clinical and field staff and is a useful tool in helping establish treatment plans and revising and adapting the treatment process to meet client needs. This practice could be adopted by other programs considering implementing ROM. Elements of the CSR that assessed satisfaction of treatment services at discharge should be continued because they appear to be an important and ethically responsible assessment domain. Asking clients to assess the treatment services they have received (perhaps only to be completed at discharge) reflect traditional treatment satisfaction instruments used in addictions and other mental health settings.

AN EVALUATION OF ALASKA CROSSINGS

Limitations of the study include the relatively small sample size used in analysis, and the fact that no control group was used. Results from this study should not be generalized to other treatment contexts or populations. In addition, though all attempts were made to have data collection efforts be similar for each group and client, there are invariably deviations from this due to unforeseen conditions at admission and discharge, or from scheduling and other field-based anomalies when collecting within program data. Despite these limitations, the following conclusions emerged from this study, which include those developed from analysis of the demographic, Y-OQ and CSR data.

AN EVALUATION OF ALASKA CROSSINGS

References

- Alaska Division of Human and Social Service's Behavioral Health Division (2017). *Client status review*. Retrieved from <http://dhss.alaska.gov/dbh/Pages/Performance%MEasures/Default.aspx>
- 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*(4), 693–704. doi:10.1037/a0016062
- Awad, A. G., & Voruganti, L. N. (2000). Intervention research in psychosis: Issues related to the assessment of quality of life. *Schizophrenia Bulletin, 26*(3), 557-564. doi:10.1093/oxfordjournals.schbul.a033477
- Beckstead, D. J., Lambert, M. J., DuBose A. P., & Linehan, M. (2015). Dialectical behavior therapy with American Indian/Alaska Native adolescents diagnosed with substance use disorders: Combining an evidence-based treatment with cultural, traditional, and spiritual beliefs. *Addictive Behaviors, 51*, 84-87. doi:10.1016/j.addbeh.2015.07.018
- Bettmann, J. E., Gillis, H. L., Speelman, E. A., Parry, K. J., & Case, J. M. (2016). A meta-analysis of wilderness therapy outcomes for private pay clients. *Journal of Child and Family Studies, 25*(9), 2659-2673. doi:10.1007/s10826-016-0439-0
- Bettmann, J. E., & Tucker, A. (2011). Shifts in attachment relationships: A study of adolescents in wilderness treatment. *Child and Youth Care Forum, 40*, 499-519. doi:10.1007/s10566-011-9146-6
- Bickman, L., Kelley, S. D., Breda, C., Regina de Andrade, A., & 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. doi:10.1176/appi.ps.002052011
- Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research, 25*(1), 6-19. doi:10.1080/10503307.2013.817696
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science, 236*, 157–161. doi:10.1126/science.3563494
- Burlingame, G. M., Wells, M. G., & Lambert, M. J. (1995). *The Youth Outcome Questionnaire*. Stevenson, MD: American Professional Credentialing Services.

AN EVALUATION OF ALASKA CROSSINGS

- Burlingame, G. M., Wells, G. M., Lambert, M. J., & Cox, J. (2004). The Youth Outcome Questionnaire. In Maruish, M. (Ed). *The use of psychological tests for treatment planning and outcome assessment* (3rd ED.). Lawrence Erlbaum.
- Burlingame, G.M., Wells, M.G., Lambert, M.J., Cox, J., Latkowski, M., & Justice, D. (2005). *Administration and scoring manual for the Youth Outcome Questionnaire (YOQ.2.2)*. Salt Lake City, UT: American Professional Credentialing Services.
- Carlier, I. V. E., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van Der Wee, N. J. a, & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: Evidence and theory. *Journal of Evaluation in Clinical Practice*, *18*(1), 104– 110. doi:10.1111/j.1365-2753.2010.01543.x
- Christensen, N. E. (2008). Effects of wilderness therapy on motivation and cognitive, emotional, and behavioral variables in adolescents (Doctoral dissertation). University of Kansas, Lawrence, KS.
- Crawford, S. L., Johannes, C. B., & Stellato, R. K. (2002). Assessment of digit preference in self-reported year at menopause: Choice of an appropriate reference distribution. *American Journal of Epidemiology*, *156*, 676 – 683. doi:10.1093/aje/kwf059
- DeMille, S. M., Comart, C., & Tucker, A. (2014). Body composition changes in an outdoor behavioral healthcare program. *Ecopsychology*, *6*(3), 174-182.
- Duncan, B. L. (2012). The Partners for Outcome Management System (PCOMS): The heart and soul of change project. *Canadian Psychology/Psychologie Canadienne*, *53*(2), 93. doi:10.1037/a0027762
- Evidence-based practice in psychology. (2006). *American Psychologist*, *61*(4), 271-285. doi:10.1037/0003-066X.61.4.271
- Gass, M.A., Gillis, H.L., & Russell, K.C. (2012). *Adventure therapy: Theory, research, & practice*. New York, NY: Routledge.
- Hagan, J. D. (2002). An alternative therapy for the behaviorally challenged youth: The efficacy of wilderness therapy programs (Doctoral dissertation). The University of Toledo, OH.
- Hammersley, R. (1994). A digest of memory phenomena for addiction research. *Addiction*, *89*, 283–293. doi:10.1111/j.1360-0443.1994.tb00890.x
- Hawkins, E. J., Lambert, M. J., Vermeersch, D. a., Slade, K. L., & Tuttle, K. C. (2004). The therapeutic effects of providing patient progress information to therapists and patients. *Psychotherapy Research*, *14*(3), 308–327. doi:10.1093/ptr/kph027

AN EVALUATION OF ALASKA CROSSINGS

- Hessel, P. A. (1986). Terminal digit preference in blood pressure measurements: Effects on epidemiological associations. *International Journal of Epidemiology*, *15*(1), 122–125. doi:10.1093/ije/15.1.122
- Hoag, M.J., Massey, K.E., Roberts, S.D., & Logan, P. (2013). Efficacy of wilderness therapy for young adults: A first look. *Residential Treatment for Children & Youth*, *30*, 294-305. doi:10.1080/0886571x.2013.852452
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy. *American Psychologist*, *51*(10), 1059–1064. doi:10.1037/0003-066X.51.10.1059
- 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*(1), 12. doi:10.1037/0022-006X.59.1.12
- Lambert, M. J. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, *17*(1), 1–14. doi: 10.1080/10503300601032506
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy*, *48*(1), 72–79. doi:10.1037/a0022238
- Levant, R. F., Barlow, D. H., David-, K. W., Hagglund, K. J., Hollon, S. D., Johnson, J. D., ... (2006). APA presidential task force on evidence-based practice: Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271–285. <http://doi.org/10.1037/0003-066X.61.4.271>
- McPherson, R.S., Hoelscher, D.M., Alexander, M., Scanlon, K.S., Serdula, M.K. (2000). Dietary assessment methods among school-aged children: Validity and reliability. *Preventive Medicine*, *31*(2) S11–S33. doi:10.1006/pmed.2000.0631
- Norcross, J.C. (2002). Empirically supported therapy relationships. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 3-16). New York: Oxford University Press.
- Roberts, S. D., Stroud, D., Hoag, M. J. and Combs, K. M. (2016). Outdoor behavioral health care: Client and treatment characteristics effects on young adult outcomes. *Journal of Experiential Education* *39*(3), 288-302. doi:10.1177/1053825916655445
- Russell, K.C., Gillis, H.L. & Lewis, T.G. (2008). A five-year follow-up of a survey of north American outdoor behavioral healthcare programs. *Journal of Experiential Education*, *31*(1), 55-77.

AN EVALUATION OF ALASKA CROSSINGS

- Shiffman, S., Stone, A. A., & Hufford, M. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1–32. doi:10.1146/annurev.clinpsy.3.022806.091415
- 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*(3), 298–311. doi:10.1037/a0019247
- Shryock, H. S., & Siegel, J. S. (1976). *Methods and materials of demography*. New York: Academic Press.
- State of Alaska Department of Health and Social Services-Division of Behavioral Health (DBH). (2011). *Alaska Screening Tool FY2011 and initial client status review FY2011: Supporting clinical decision making and program performance measurement*. Retrieved from <http://dhss.alaska.gov/dbh/Pages/Performance%20Measures/Default.aspx>
- Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2008). *Mental health service use among youths aged 12 to 17: 2005 and 2006*. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.
- Sundet, R. (2010). Therapeutic collaboration and formalized feedback: Using perspectives from Vygotsky and Bakhtin to shed light on practices in a family therapy unit. *Clinical Child Psychology and Psychiatry, 15*(1), 81–95. doi:10.1177/1359104509341449
- Sundet, R. (2012). Therapist perspectives on the use of feedback on process and outcome: Patient-focused research in practice. *Canadian Psychology/Psychologie Canadienne, 53*(2), 122–130. doi:10.1037/a0027776
- Sundet, R. (2013). Postmodern-oriented practices and patient-focused research: Possibilities and hazards. *Australian and New Zealand Journal of Family Therapy, 33*(04), 299–308. doi:10.1017/aft.2012.38
- Timimi, S., Tetley, D., Burgoine, W., & Walker, G. (2012). Outcome Orientated Child and Adolescent Mental Health Services (OO-CAMHS): A whole service model. *Clinical Child Psychology and Psychiatry, 18*(2), 169–184. doi:10.1177/1359104512444118
- Wampold, B. E. (2015). Routine outcome monitoring: Coming of age—With the usual developmental challenges. *Psychotherapy, 52*(4), 458–462.

AN EVALUATION OF ALASKA CROSSINGS

Wells, M. G., Burlingame, G., Lambert, M. J., Hoag, M., & Hope, C. (1996). Conceptualization and measurement of patient change during psychotherapy: Development of the Outcome Questionnaire and Youth Outcome Questionnaire. *Psychotherapy: Theory, Research, Practice, Training*, 33(2), 275-283. doi:10.1037/0033-3204.33.2.275

Willie, J. (2017). Harvard Pilgrim sued over wilderness therapy coverage. *Pension and Benefits Daily*. Retrieved from <https://www.bna.com/harvard-pilgrim-sued-n73014450999/?amp=true>.