

MEASURING CLINICAL OUTCOMES AND CONDUCTING PROGRAM EVALUATION

By Martha Shumway, Ph.D. and Alicia Boccellari, Ph.D.



“Birds sing after a storm; why shouldn’t people feel as free to
delight in whatever sunlight remains to them?”

–Rose Kennedy

*The UC San Francisco Trauma Recovery Center Model: Removing Barriers to Care and
Transforming Services for Survivors of Violent Crime.* Edited by Stacey Wiggall, LCSW &
Alicia Boccellari, Ph.D. Produced in collaboration with Allen/Loeb Associates
Version 1: March 15, 2017

The UC San Francisco TRC has developed and incorporated an integrated system of standardized forms, data collection procedures, relational databases, and performance improvement processes. It also uses standardized mental health measures to assess a client's symptoms and their response to treatment interventions. The UC San Francisco TRC database enables staff to examine the effectiveness of TRC services at both a programmatic level and an individual client level.

As discussed in the chapter *TRC Assessment and Treatment Planning*, the UC San Francisco TRC uses several standardized measures to assess mental health symptoms and other clinically relevant domains, such as pain, sleep disturbance, and quality of life. Standardized measures are integrated into the initial comprehensive assessment as part of the Multi-Area Review and Trauma History Assessment (MARTHA) and are repeated every eight treatment sessions to evaluate clients' progress and response to treatment. Standardized measurement facilitates comparison of the TRC's client population to other populations and helps clinicians to monitor individual client progress, identify areas to target in treatment, and to select intervention strategies. Clinicians also discuss the standardized measurements with clients and use results to engage clients in shared decision making about their treatment. As noted previously, the ultimate goal is to use the standardized assessment data to develop algorithms to guide selection of evidence-based interventions. This chapter outlines the rationale the UC San Francisco TRC uses to select standardized measures and then describes the domains that are assessed, the measures in use at the UC San Francisco TRC, and other potentially useful measures.

Strategy for Selecting Standardized Measures

The UC San Francisco TRC uses a set of common principles to select standardized measures for routine use. First, there should be published data showing that the measures have strong “psychometric properties.” This means that the measures should be **reliable**, in that they work similarly across people, groups, and time, **valid** in that they measure what they are intended to measure, and **responsive**, in that they are sensitive to change.

Second, measures should have low “cognitive complexity.” Survey researchers have identified several factors that make standardized questions difficult to answer (Krosnick, 1999; Sudman, Bradburn, & Schwarz, 1996; Tourangeau, Rips, & Rasinski, 2000). Questions that are long, use highly sophisticated vocabulary, or are grammatically ambiguous tend to yield poor quality answers. Similarly, measures with many similar items can lead to fatigue that limits the quality and accuracy of answers. Therefore, measures are selected that have short, straightforward questions. Measures are chosen to be as brief as possible while allowing measurement of change over time. Very brief screening measures, with only a few items, typically aren't sensitive to changes over time.

Third, measures with evidence of cross-cultural applicability are preferred because they are likely to be relevant to the TRC's diverse client population. Measures that have already been translated into other languages, particularly Spanish, are ideal.

Fourth, the UC San Francisco TRC uses measures that are available free-of-charge. Many high quality measures have been developed with government support or otherwise developed for public use. As a result, it is not necessary to use limited funding to purchase measures.

PTSD Symptoms

PTSD symptoms are a central focus of mental health treatment for crime victims. To measure PTSD symptoms, the UC San Francisco TRC uses the PTSD Checklist (PCL). The original 17-item PCL, developed to assess PTSD as defined in DSM-IV (Weathers, Litz, Herman, Huska, & Keane, 1993), has been repeatedly shown to be valid, reliable, and highly concordant with clinician-administered diagnostic tools in a wide range of populations (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996; Grubaugh, Elhai, Cusack, Wells, & Freuh, 2007). The 20-item PCL-5 was recently developed to assess PTSD as defined in DSM-5 (Weathers, et al., 2013). Preliminary evidence suggests that it is as reliable and valid as the original PCL [(Blevins, Weathers, Davis, Witte, & Domino, 2015; Bovin, et al., 2015; Wortmann, et al., 2016). The PCL-5 items are all answered on a 5-point response scale that ranges from 0="not at all" to 4="extremely." The PCL-5 is particularly useful because it can be scored to yield a continuous measure of symptom severity and to make a provisional diagnosis of PTSD according to DSM criteria. Separate scores for the four clusters of PTSD symptoms can also be calculated. PCL scores are calculated as simple sums of the responses which makes it possible for clinicians to immediately calculate scores when that is clinically useful. Based on available data, a score of 33 or higher on the PCL-5 is associated with a diagnosis of PTSD. Additional information on the PCL is available from the National Center for PTSD at <http://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp>

There are many reliable and valid standardized measures of PTSD symptoms. However, not all established measures have been updated to measure PTSD as defined by DSM-5. Many measures have more questions than the PCL, which increases both administration time and clients' cognitive burden. Some measures, for example the Davidson Trauma Scale (DTS), separate the measurement of symptom frequency from the measurement of distress associated with symptoms. This additional information could be clinically useful; however the DTS is not available free of charge and the parallel sets of questions increase administration time.

Complex PTSD

People who experience childhood abuse and/or repeated traumas in adulthood can experience what is often termed "Complex PTSD" (CPTSD), or "Disorders of Extreme Stress," which involves chronic affect dysregulation and interpersonal difficulties in addition to the symptoms of PTSD (Herman, 1997; van der Kolk, Roth, Pelcovitz, Sunday, & Spinazzola, 2005). The concept of CPTSD has been controversial and no practical standardized measures have been validated. The one standardized measure available, the 45-item Structured Interview for Disorders of Extreme Stress (Pelcovitz, et al., 1997), has not been well-validated and is time-consuming and potentially distressing to recently traumatized persons. To meet the need for a practical standardized measure of CPTSD, researchers working with the UC San Francisco TRC have developed an initial 23-item measure that is currently being tested. The goal is to reduce the number of items to produce a more manageable and useful measure. As part of the formulation of a CPTSD diagnostic category for ICD-11, a similar, 24-item questionnaire has been developed, but is not yet publicly available.

Lifetime Trauma History

In planning treatment for a recently traumatized crime survivor, it is important to understand the individual's trauma history. Individuals who have experienced many prior traumas will respond differently, and have different treatment needs, than individuals who have not experienced prior traumas (Herman, 1997; van der Kolk, Roth, Pelcovitz, Sunday, & Spinazzola, 2005). Childhood trauma is of particular interest because it has been repeatedly shown to be associated with revictimization (Coid, et al., 2001; Dong, et al., 2004) and adverse health and mental health consequences over the life course (Felitti, et al., 1998). There are many published trauma history checklists. The UC San Francisco TRC has used different checklists in the past and now uses the Trauma History Screen (THS) (Carlson, et al., 2011) because the THS describes 14 key types of trauma very succinctly. To get a basic understanding of each client's trauma history without causing the client to become distressed or overwhelmed, only the initial checklist portion of the THS is used. The detailed follow-up questions about each type of trauma are not used. The clinician can follow-up on details as clinically indicated over the course of treatment.

Traumatic Brain Injury

Traumatic brain injuries (TBI) can have lingering effects that can impact individuals' response to traumatic events and treatment for traumatic stress. The UC San Francisco TRC screens for prior TBI using 3 yes-no questions adapted from the Ohio State University TBI Identification Method Short Form (Corrigan & Bogner, 2007). The questions are:

1. In your lifetime, have you ever injured your head?
2. (if yes to 1.) Were you "knocked out" or did you lose consciousness after (this injury / any of these injuries)?
3. (if yes to 1.) Were you dazed or did you have any gaps in your memory after (this injury / any of these injuries)?

Yes answers to 1 and 2 indicate the strong likelihood of a TBI; yes answers to 1 and 3 indicate a possible TBI.

Depression

Depression is common following traumatic events and co-occurs with PTSD. The UC San Francisco TRC uses the PHQ9 (Kroenke, Spitzer, & Williams, 2009) to measure depression. The nine PHQ items are directly linked to DSM diagnostic criteria. All questions use the same response scale that ranges from 0=not at all to 3=nearly every day. Responses can be easily summed by the clinician as needed. The PHQ9 has been shown to be valid and reliable across a range of diverse populations and settings (Becker, Al Zaid, & Al Faris, 2002; Wulsin, Somoza, & Heck, 2002), as well as sensitive to change over time (Lowe, Kroenke, Herzog, & Grafe, 2004). Validated cut-points distinguishing five levels of depression severity have been established. The PHQ9 and its administration and publication manual are readily available for download in multiple languages at: <http://www.phqscreeners.com>.

Sleep Disturbance

A growing body of evidence documents a strong association between sleep disturbance and persistence of PTSD symptoms (Brownlow, Harb, & Ross, 2015; Gilbert, Kark, Gehrman, & Bogdanova, 2015). The UC San Francisco TRC uses the 4-item Sleep Disturbance scale developed as part of the NIH's PROMIS (Patient-Reported Outcomes Measurement Information System) initiative. This is the shortest of the multiple PROMIS Sleep Disturbance measures that have been rigorously developed using item-response theory methods (Yu, et al., 2012). In UC San Francisco TRC data, the 4-item version was correlated with longer versions at $r < .9$; therefore the 4-item version is used to reduce respondent burden and save time. The PROMIS Sleep Disturbance measure is much shorter and easier to follow than other well-established sleep measures, such as the widely-used Pittsburgh Sleep Quality Index (PSQI) (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). However, it is still somewhat problematic for clients with limited literacy, limited English proficiency, or cognitive difficulties because it includes "reversed" items in which the response scale has a different meaning for different questions. PROMIS measures are available at: <http://www.healthmeasures.net>.

Pain

Emerging evidence suggests a bidirectional connection between chronic pain and PTSD symptoms. The UC San Francisco TRC uses the recently developed 3-item PEG pain measure (Krebs, et al., 2009) to assess pain. The PEG is based on the longer, well validated and widely used pain scale Brief Pain Inventory (BPI) (Tan, Jensen, Thornby, & Shanti, 2004). The three questions use similar 0 to 10 point scales to measure pain severity, the extent to which pain interferes with enjoyment of life and the extent to which pain interferes with general activity.

Quality of Life

To obtain a more general more general measure of clients' well-being that goes beyond symptoms, the UC San Francisco TRC uses the 26-item WHOQOL-BREF as measure of quality of life (WHOQOL Group, 1998) The WHOQOL-BREF was multinationally developed to assess QOL related to physical health, psychological well-being, social relationships and environment. It is sensitive to changes associated with mental health treatment (Carta, et al., 2008; Chand, Mattoo, & Sharon, 2004).

Other Diagnoses

In addition to the structured measures discussed above, the UC San Francisco TRC uses symptom checklists linked to DSM criteria that have been tailored over time to meet the needs of the clinic, such as Panic Disorder, Bipolar Disorder, and Psychotic Disorders. The M.I.N.I. (Sheehan et al., 1997) is a well-validated diagnostic tool that takes a similar approach. Information on current ways to use the M.I.N.I. can be found at: <http://www.medical-outcomes.com/index/mini>.

Service Delivery and Program Evaluation

Along with utilizing standardized measures to track each client's response to treatment, UC San Francisco TRC collects additional data that is used to evaluate program and clinical effectiveness.

The use of standardized referral and productivity forms allows the TRC to track staff allocations, staff productivity, and client flow, including the number of clients served, units of service, and the types of service provided. It also allows for the collection of individual client data including demographics, crime type, and other variables from referral information.

REFERRAL FORM. This form contains information about which agency is making the referral, basic client demographics, and the type of violent crime that is necessitating the referral. It also contains information as to whether the client showed for the initial intake evaluation. Data from this form can be used to generate a “Recruitment Report” which allows the TRC to keep track of the number of referrals received, broken down by the variables of age, gender, ethnicity, and type of crime. It also allows staff to look at client attrition occurring between the point of referral and the scheduled intake, and helps to guide initial outreach attempts.

PRODUCTIVITY FORM. This form is filled out by the Clinicians for each service they provide in order to capture all client encounters. It includes the type of service provided (i.e., individual therapy, group therapy, case management), the type of evidence-based treatment provided, the location of the service (at TRC, at the client’s home, or in the community) and the length of time it took to provide the service. Clinicians also enter the amount of Direct time (face-to-face) or Indirect time (on the phone with a client or engaged in case management on a client’s behalf), in addition to noting the time it took to complete any service-related documentation (progress notes, intake assessment reports, etc.) This data can then be used to generate a “Productivity Report,” which captures the total number of clients seen throughout the TRC, and the number and types of services that were provided. It also allows for tracking which clients completed treatment and which dropped out. Service data can be broken down by demographic information, type of crime, or Clinician. This tool is helpful for reviewing Clinician productivity and caseload size.

NEEDS ASSESSMENT, BARRIERS TO CARE AND LINKAGE TO SERVICES. In an effort to help guide case management interventions and to evaluate our effectiveness in providing case management services, Clinicians complete a Needs Assessment and Barriers to Care form when a client begins services, at the 8th session, and at the close of treatment. Based on the clinical interview and using a 4-point scale that ranges from 0 (None) to 3 (Extensive Need), the Clinician rates a client’s level of need in 13 domains. These include: financial needs, legal assistance, transportation, housing, and other domains where clients may have service needs. The Clinician also indicates any existing barriers to care, such as language barriers, or difficulty leaving the house to access services. At every 8th session, the Clinician indicates whether case management services were offered to the client to meet their needs, and whether TRC provided the service or helped the client link to a resource outside of TRC.

SERVICE SATISFACTION SURVEY. In order to evaluate and monitor the program’s responsiveness to clients, a Service Satisfaction Survey is given at every 8th session and at the closing of treatment. The survey is composed of 13 items that use a 5-point scale,

and includes items such as, “How satisfied were you with the progress made towards your goals for treatment?”, and “How much were your cultural values and/or language preferences respected during your visits?”

Summary

A guiding principle of the TRC model is accountability to the clients we serve and to our funding agencies. This is accomplished through the collection of data on both clinical outcomes and services provided, so that the program is evaluated for clinical effectiveness and fiscal responsibility. An effective program evaluation can be used to help maintain current funding and to acquire new funding sources. Data can also be used to advocate for system-wide policy change that removes barriers to care for underserved survivors of violent crime.

References

- Becker, S., Al Zaid, K., & Al Faris, E (2002). Screening for somatization and depression in Saudi Arabia: a validation study of the PHQ in primary care. *International Journal of Psychiatry in Medicine*, 32, 271-283.
- Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C.A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, 8, 669-673.
- Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, 28, 489-498.
- Bovin, M. J., Marx, B. P., Weathers, F. W., Gallagher, M. W., Rodriguez, P., Schnurr, P. P., & Keane, T. M. (2015). Psychometric properties of the PTSD Checklist for Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (PCL-5) in veterans. *Psychological Assessment*. Advance online publication.
- Brownlow, J. A., Harb, G. C., & Ross, R. J. (2015). Treatment of sleep disturbances in post-traumatic stress disorder: A review of the literature. *Current Psychiatry Reports*, 17, 41.
- Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28, 193–213.

Carlson, E. B., Smith, S. R., Palmieri, P. A., Dalenberg, C., Ruzek, J. I., Kimerling, R., ... Spain, D. A. (2011). Development and validation of a brief self-report measure of trauma exposure: The Trauma History Screen. *Psychological Assessment, 23*, 463-477.

Carta, M. G., Hardoy, M. C., Pilu, A., Sorba, M., Floris, A. L., Mannu, F. A., ... Salvi, M. (2008). Improving physical quality of life with group physical activity in the adjunctive treatment of major depressive disorder. *Clinical Practice and Epidemiology in Mental Health, 4*, 1.

Chand, P. K., Mattoo, S. K., & Sharan, P. (2004). Quality of life and its correlates in patients with bipolar disorder stabilized on lithium prophylaxis. *Psychiatry and Clinical Neurosciences, 58*, 311-318.

Coid, J., Petruckevitch, A., Feder, G., Chung, W., Richardson, J., & Moorey, S. (2001). Relation between childhood sexual and physical abuse and risk of revictimization in women: A cross-sectional survey. *Lancet, 358*, 450-454.

Corrigan, J. D., & Bogner, J. A. (2007). Initial reliability and validity of the OSU TBI Identification Method. *Journal of Head Trauma Rehabilitation, 22*, 318-329.

Dong, M., Anda, R. F., Felitti, V. J., Dube, S. R., Williamson, D. F., Thompson, T. J., ... Giles, W. H. (2004). The interrelatedness of multiple forms of childhood abuse, neglect, and household dysfunction. *Childhood Abuse & Neglect, 28*, 771-784.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., ... Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine, 14*, 245-258.

Gilbert, K.S., Kark, S. M., Gehrman, P., Bogdanova, Y. (2015). Sleep disturbances, TBI and PTSD: Implications for treatment and recovery. *Clinical Psychology Review, 40*, 195-212.

Grubaugh, A. L., Elhai, J. D., Cusack, K. J., Wells, C., & Frueh, B. C. (2007) Screening for PTSD in public-sector mental health settings: The diagnostic utility of the PTSD Checklist. *Depression and Anxiety, 24*, 124-129.

Herman, J. (1997). *Trauma and Recovery: The aftermath of violence--from domestic abuse to political terror*. New York: Basic Books.

Kennedy, R. (n.d.) Goodreads.com. Retrieved November 8, 2016 from Goodreads website: <http://www.goodreads.com/quotes/251632-birds-sing-after-a-storm-why-shouldn-t-people-feel-as>

Krebs, E. E., Lorenz, K. A., Bair, M. J., Damush, T. M., Wu, J., Sutherland, J. M., ... Kroenke, K. (2009). Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. *Journal of General Internal Medicine, 24*, 733-738.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606-613.

Krosnick, J. (1999). Survey research. *Annual Review of Psychology 50*, 537-567.

Lowe, B., Kroenke, K., Herzog, W., & Grafe, K. (2004). Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders, 81*, 61-66.

Pelcovitz, D., van der Kolk, B., Roth, S., Mandel, F., Kaplan, S., & Resick, P. (1997). Development of a criteria set and a Structured Interview for Disorders of Extreme Stress (SIDES). *Journal of Traumatic Stress, 10*, 3-16.

Sheehan D., Lecrubier Y., Harnett-Sheehan K., Janavs J., Weiller E., Bonara L., Keskiner A, Schinka J, Knapp E., Sheehan M., Dunbar G. (1997). Reliability and validity of the M.I.N.I. International Neuropsychiatric Interview (M.I.N.I.): According to the SCID-P. *European Psychiatry, 12*, 232-241.

Sudman, S., Bradburn, N., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.

Tan, G., Jensen, M. P., Thornby, J. I., & Shanti, B. F. (2004) Validation of the Brief Pain Inventory for chronic nonmalignant pain. *The Journal of Pain, 5*, 133–137.

Tourangeau, R., Rips, L. J., & Raskinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.

van der Kolk, B. A., Roth, S., Pelcovitz, D., Sunday, S., & Spinazzola, J. (2005) Disorders of extreme stress: The empirical foundation of a complex adaptation to trauma. *Journal of Traumatic Stress, 18*, 389-399.

Weathers, F., Litz, B., Herman, D., Huska, J., & Keane, T. (1993) The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility. Presented at the International Society for Traumatic Stress, San Antonio, TX, 1993.

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The PTSD Checklist for DSM-5 (PCL-5). National Center for PTSD at www.ptsd.va.gov.

WHOQOL Group. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*, 28, 551-558.

Wortmann, J. H., Jordan, A. H., Weathers, F. W., Resick, P. A., Dondanville, K. A., Hall-Clark, B., ... Litz, B. T. (2016). Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. *Psychological Assessment*. Advance online publication.

Wulsin, L., Somoza, E, & Heck, J. (2002). The feasibility of using the Spanish PHQ-9 to screen for depression in primary care in Honduras. *Primary Care Companion to the Journal of Clinical Psychiatry*, 4, 191-195.

Yu, L., Buysse , D. J., Germain , A., Moul, D. E., Stover, A. , Dodds, N. E. , Pilkonis, P. A. (2012). Development of short forms from the PROMIS™ Sleep Disturbance and Sleep-Related Impairment item banks. *Behavioral Sleep Medicine*, 10, 6-24.