



THE UNIVERSITY OF  
NEW MEXICO.

# **Mobile Crisis Team Screening and Assessment Tools and Procedures**

**Prepared by:**

Alex Severson, Ph.D.

Prepared for:

Bernalillo County Department of  
Behavioral Health Services  
Behavioral Health Initiative

## INTRODUCTION

Mobile Crisis Teams (MCTs) are specialist behavioral health teams designed to provide psychiatric emergency care – including crisis assessment, crisis intervention and stabilization, temporary shelter, and appropriate referral services – in naturalistic, non-clinical environments to individuals experiencing acute behavioral health crises. MCT units are usually comprised of a sworn LEO and a licensed clinician<sup>1</sup> who are typically dispatched to crisis sites in response to 911 or crisis hotline calls (Kisely et al., 2010).

MCTs have been increasingly deployed in recent years as an alternative to hospital-based psychiatric emergency services and law enforcement only responses to address the interconnected problems of (1) excessive psychiatric boarding in emergency room departments, (2) understaffed emergency room departments, and (3) inadequate behavioral health crisis training of law enforcement officers (Forchuk et al., 2010; Lancaster 2016). Advocates of the MCT approach to crisis intervention reason that, compared to police-centric models of crisis intervention, MCTs are better equipped to triage, screen, assess, and divert those experiencing acute mental distress to appropriate levels of service because of the comparatively higher level of technical expertise of the MCT members and because mobile crisis services tend to be delivered in disarming non-clinical environments (Scott 2000).

To date, the academic literature on MCTs has explored whether MCTs, when compared to general care models of crisis response and other crisis intervention training (CIT) programs, are better able to improve clinical and behavioral health outcomes among those experiencing behavioral health crises (e.g., reduced suicidality), reduce arrest rates and the amount of time officers typically spend at the scene of the crisis events, and reduce hospitalization burden. While a substantial proportion of early research on MCTs made use of non-experimental descriptive methodologies, more recent, high-quality meta-analytic, experimental, and quasi-experimental research tentatively suggests that MCTs are a cost-effective intervention for reducing future emergency room usage and suicidality (Baess 2005; Kisely et al., 2010; Murphy et al. 2015; Fendrich et al. 2019). However, it is worth highlighting that the existing evidence-base on MCT interventions suffers from some limitations in research design, that the positive effects of MCTs are often moderated by organizational factors and community-level resources, and that effect sizes are sensitive to the specific outcome of interest (Compton et al., 2006; Shapiro et al., 2015; Heyman and McGeough 2018). While there has been considerable research exploring whether MCTs are effective, decidedly less is known about the specific suite of diagnostic tools crisis teams should use for screening and assessment purposes. Furthermore, there is not consistent guidance on the practices and protocols crisis teams should use when administering these tools. In the absence of specific clinical guidelines, we can identify the scope of which scales should be included on the basis of the Substance Abuse and Mental Health Services Administration's (SAMHSA) recently-published *National Guidelines for Behavioral Health Crisis Care Best Practice Toolkit* and can observe which clinical screening and assessment tools have been consistently used across multiple evaluations of MCTs to get a crude sense of which may be appropriate to use (SAMHSA 2020)<sup>2</sup>.

The SAMHSA toolkit does not provide specific recommendations for which screening and assessment instruments should be deployed by MCTs beyond noting the need for crisis teams to use the Level of Care

---

<sup>1</sup> These teams typically consist of a pairing of any two of the following: LEOs, psychiatrists, nurses, social workers, psychologists, mental health technicians, addiction specialists, or peer counselors. MCTs often are multidisciplinary partnerships between police and health interventionists.

<sup>2</sup> The popularity of a given's scale usage does not necessarily reflect its predictive accuracy or usefulness as an analytic tool.

Utilization System (LOCUS) to make disposition decisions and noting the need to collect measurements on suicidality risk given the comparatively high frequency of suicidal ideations among those experiencing mental health crises (Currier, Fisher, and Caine 2010).

Furthermore, the screening and assessment tools mentioned in some of the reviewed implementation and process evaluations vary in the extent to which their psychometric properties are known, the feasibility of their application in crisis situations, and the degree to which they are culturally-sensitive. Additionally, to our understanding, a set of best practices – specifically related to the sequencing of the delivery of particular scales and the timeline over which such scales are administered in the context of a crisis event and subsequent aftercare – has not been formally articulated yet by governing health authorities.

This review is limited to summarizing recent best practice recommendations for MCTs, identifying a few additional best practices informed by other recent areas of research on MCTs, and comments on recent state-level technological innovations in MCT assessment protocols. Additionally, we enumerate a series of screening and assessment scales mentioned frequently in MCT implementation evaluations and comment on the quality of evidence – to the extent that it exists – on the psychometric properties of these scales.

### **MCT BEST PRACTICES RECOMMENDATIONS**

SAMHSA published a document in February 2020, *National Guidelines for Behavioral Health Crisis Care Best Practice Toolkit*, which delineates a set of minimal and best practices for MCTs which we reproduce below verbatim. SAMHSA, for instance, notes that at a minimum level, MCTs must:

- Include a licensed and/or credentialed clinician capable to assessing the needs of individuals within the region of operation
- Respond where the person is (home, work, park, etc.) and not restrict services to select locations within the region or particular days/times
- Connect individuals to facility-based care as needed through warm hand-offs and coordinating transportation when and only if situations warrant transition to other locations (SAMHSA 2020: 18).

In terms of best practices, SAMHSA notes the need for MCTs to:

- Incorporate peers within the mobile crisis team
- Respond without law enforcement accompaniment unless special circumstances warrant inclusion in order to support true justice system diversion
- Implement real-time GPS technology in partnership with the region’s crisis call center hub to support efficient connection to needed resources and tracking of engagement
- Schedule outpatient follow-up appointments in a manner synonymous with a warm handoff in order to support connection to ongoing care (SAMHSA 2020: 18).

These recommendations are partially consistent with the existing evidence-base on MCTs and crisis care. For instance, while ISR’s previous literature review evaluated how various co-response models influence behavioral health outcomes (Alej and Pierotti 2016), to our understanding, to date, no studies have evaluated the effectiveness of peer inclusion in MCTs despite the common recommendation for peer

inclusion in MCTs advised by governing agencies. Advocacy for the use of peers on MCTs occasionally appeals to the more extensive literature on peer case management. However, the evidence-base on the effectiveness of the use of peers is often more equivocal than is presumed given theoretical and methodological deficits in previous studies (see, for instance, Bellamy, Schmutte, and Davidson (2017)).

The recommendation that MCTs should try to respond to crisis events without law enforcement is consistent with tentative findings from studies suggesting that MCT users prefer non-uniformed LEOs and other interventionists because they are concerned about the possibility of escalation by, and stigmatization of, the police (Boscarato et al. 2014; Daggenvoorde, Gijsman, and Goossens 2018; Lamanna et al. 2018) though it is worth noting that research which directly compares co-response models is limited in scope and may be confounded by wide variations in model fidelity across comparison sites (Puntis et al. 2018; Park et al. 2019) and that there are understandably some crisis scenarios (e.g., those which present high risks of violence) where law enforcement involvement is necessary to ensure public safety. While there have not been any studies which evaluate the effectiveness of GPS technology in the context of crisis management, the use of GPS technology seems reasonable to fold into existing practice to more efficiently triangulate and expedite care delivery, barring prohibitive costs to adopting such technologies and technological barriers to the technology's implementation (e.g., poor service areas). The use of warm-handoff outpatient follow-ups is similarly consistent with the evidence-base (Goldman et al., 2020; Young et al. 2020).

A review of process, implementation, and outcome evaluations of MCTs suggest some additional approaches MCTs might consider going forward. For instance, one common recommendation articulated by users of crisis services and their families is the need for rapid initial response (e.g., < 1 hour) by MCT teams (Morant et al., 2017; Bailey et al., 2018). A second common recommendation is that services should be available 24/7 and not tethered to particular catchment areas. A third common recommendation is to supplement MCTs with 40-hour CIT trainings for police officers largely centered on de-escalation and communication techniques. However, a recent review paper suggests that there is not sizable supportive empirical evidence to justify CIT's widespread adoption given a host of issues with previous studies including selection biases, issues with self-reported data, wide variations in training design and delivery, and logistical issues with implementing CIT in smaller departments (Booth et al., 2017; Peterson and Densley 2018). To overcome some of these issues with CIT training, other training models - specifically, the R-Model - have recently been developed which synthesize some of the better elements of CIT training, Mental Health First Aid, Crisis Prevention Institute (CPI) training, Integrating Communications, Assessment and Tactics training, and the Police-Mental Health Collaboration toolkit and hold potential promise going forward (Peterson, Densley, and Erickson 2020). Finally, more recent state-level innovations in MCTs include the increasing proliferation of crisis mobile applications - such as [Suicide Safe](#) and [Disaster Distress](#) - which serve as diagnostic tools first responders can use to simplify and expedite existing assessment and referral processes (Gaziel-Yablowitz and Schwartz 2018).

## **SCREENING AND ASSESSMENT SCALES**

### ***Considerations in Crisis Screening and Assessment***

When deciding which specific screening and assessment tools should be administered in the field, it is worth being cognizant of a few common issues surrounding scale administration which are unique to crisis ecologies. As Bonyng and Thurber (2008) note, individuals experiencing acute mental distress are pressed for time and typically less motivated to complete a series of long-form behavioral health scales

and more inclined to impression-manage when they do which can lead to considerable response falsification. To this point, Myer and Conte (2006) note: "Simple reliance on diagnostic models, standardized tests, or intake protocols can mislead clinicians in these situations. A model specifically designed for crises is needed to guide the assessment process" (965). To this point, recent research suggests that having individuals reflect on traumatic events can generate considerable biases in the answers they provide to mental health screeners (Russell and Russell 2019). Thus, MCTs should aim to not overly rely on self-reported instruments for screening and assessment purposes and supplement these measures, when possible, with clinical ratings made from observations by trained clinicians. Further, it is important to consider how the unique landscape of crisis environments might complicate screening and assessment administration and contribute to accented assessment burden. As Myer and Conte (2006) note:

Assessment in crisis intervention is different from that in other types of assessment with respect to goals, process, relation to treatment, and type of information gathered. Recognition of these differences is essential for clinicians who are involved in providing crisis intervention. Although skills learned for other types of assessment are useful, clinicians cannot rely on these in crises. At times the skills used in other types of assessment can actually hinder and prevent effective crisis intervention. For example, clinicians must often be prepared to provide treatment within the first 5 minutes of contact. Clinicians must have assessed clients' reactions well enough to have begun the intervention process. The luxury of writing a report and getting results from standardized tests is not practical in crisis situations. Clinicians need to be trained to adapt skills and increase their effectiveness in crisis assessment (2006: 966).

Thus, there exists a tension between the need to acquire exhaustive evaluation-relevant information from clients in real time and the need to be sensitive to individual crisis psychologies. Despite the absence of evidence speaking to the ecological validity and reliability of most commonly used behavioral health scales in crisis applications, most reported process and outcome evaluations of MCTs in the past decade surveyed for this review indicate the widespread use of such scales among MCTs. Finally, MCTs should jointly consider who is delivering the screening and assessment instruments and who these instruments are being delivered to. Some specific scales, such as LOCUS, should only be administered by licensed professionals with sufficient training. Other scales, such as PHQ-9, can be administered by paraprofessionals in MCTs who have little to no training. Thus, it is important to consider the technical expertise and certifications of those on the MCT primarily for instruments which depend upon clinical discretion or client observation. Relatedly, it is worth considering the cultural sensitivity of different scales: some commonly used behavioral health scales have been translated and psychometrically validated in multiple languages whereas others have not (Arnold and Matus 2000). Some scales are only intended for use in adults; other scales can be appropriately used for both adults and adolescents. For these reasons, it is important to be mindful of this set of contingencies when it comes to screening and assessment practices in crisis situations.

### ***Triage Scales***

Triage scales are the first measurement tool used in MCTs, can help clinicians categorize the severity of clinical presentations, and are an important precondition for coordinating appropriate crisis service responses and referrals. Because crisis teams are often mobilized by different actors such as suicide hotline operators or 911 dispatchers and these actors often have variant levels of triage training and skill,

MCTs can sometimes be inappropriately dispatched to scenarios where they are not needed or, alternatively, can be underutilized due to a lack of referrals (Landinez 2019). Thus, there exists a need to deploy standardized crisis triaging scales across those who initialize contact with MCTs within the same integrated behavioral health network.

The initial stage of screening MCT calls often involves dispatcher assessments regarding the level of severity of the crisis situation and a determination as to whether the crisis in question warrants medical or police response. To our knowledge, there have not been studies in the United States evaluating the reliability of specific triage screening tools, though early evidence from Canada suggests that the use of such screening tools improves nurse satisfaction and some client-side outcomes (Broadbent, Jarman, and Berk 2004). Two commonly used scales identified in the literature on triage scales are (a) the Australian Mental Health Triage Scale (Department of Health Victoria 2009) and the adapted UK Mental Health Triage Scale (UK-MHTS) (Sands et al., 2016). Both scales have higher ecological validity than some of the earlier crisis triage scales, such as the Crisis Triage Rating System (Bengelsdorf et al. 1984; Turner and Turner 1991), in part because of the greater variety of disposition options provided though reliability has not been assessed in the United States context. The development and use of crisis triaging scales could serve as an initial starting point for reducing existing inefficiencies in and complications with initial referrals to MCTs. We provide a framework for collecting and reporting co-response triage models in Appendix A.

### ***Mental Health Screening and Assessment***

Once MCTs arrive on scene and once crisis stabilization interventions have been successfully deployed, MCTs should begin the screening and assessment process. The SAMHSA toolkit provides some general guidelines on what specifically should be screened and assessed for in the context of MCT assessments. Specifically, SAMHSA noted that MCTs should collect data on:

- Causes leading to the crisis event; including psychiatric, substance abuse, social, familial, legal factors and substance use
- Safety and risk for the individual and others involved; including an explicit assessment of suicide risk
- Strengths and resources of the person experiencing the crisis, as well as those of family members and other natural supports
- Recent inpatient hospitalizations and/or any current relationship with a mental health provider
- Medications prescribed as well as information on the individual's compliance with the medication regimen
- Medical history as it may relate to the crisis (SAMHSA 2020: 20)

Consistent with these recommendations, we provide a list of some of the scales MCTs have used in recent evaluations which capture these recommendations and which seem to have desirable psychometric properties in terms of validity, reliability, sensitivity, and specificity<sup>3</sup>. We provide a more detailed table of these scales and their respective psychometric properties in Appendix B.

---

<sup>3</sup> Validity of a screening test refers to its ability to discriminate between an individual with a problem and one without such a problem. Reliability refers to the ability of a measure to produce consistent results. Sensitivity refers to the accuracy of the test in identifying a problem. Specificity is the accuracy of the test in identifying individuals who do not have a problem.

- Columbia Suicide Severity Rating Scale (Posner et al., 2011)
- Suicide Assessment Checklist (Rogers 1994)
- Scale for Suicide Ideation (SSI) (Beck, Kovacs, and Weissman 1979)
- The UCLA Post-Traumatic Stress Index (Rodriguez, Steinberg, and Pynoos 1998)
- The Structured Assessment of Violence Risk in Youth (Borum, Bartel, and Forth 2002)
- The Global Appraisal of Individual Needs-Short Screener (Dennis et al. 2006)
- The Ohio Scales (Ogles et al., 2001)
- HEADS-ED Scale (Cappelli et al 2012; Cappelli et al. 2020)
- Global Assessment Scale (GAS) (Endicott et al. 1976)
- PHQ – 9 (Kroenke, Spitzer, and Williams 2001)
- Client Satisfaction Questionnaire (CSQ-8) (Attkisson and Zwick 1982)
- Manchester Short Assessment of Quality of Life (MANSA) (Priebe et al. 1999)

As noted earlier, there are a series of issues associated with an approach to crisis screening and assessment dependent upon clinicians on the MCT using multiple diagnostic instruments. First, there are a series of practical issues associated with giving clients a series of different behavioral health scales. A number of stand-alone scales measuring related concepts (e.g., cognition; depression) can introduce unnecessary redundancies in measurement which can increase assessment burden on the client side (Ebesutani et al., 2012) and potentially generate statistical bias across scales. Discretion on the clinician-side over which specific scales to use in a given crisis scenario invites a level of subjectivity in scale administration which can result in the inconsistent application of scales across MCTs. With multiple scales in use, it is unclear to what extent, if any, modifying the sequence of scale delivery biases subsequent client responses, raising the specter of potential ordering effects (McFarland 1981). Finally, there may be licensing and cost issues associated with scale use as well as exclusion criteria which limit the populations for which such scales can be used (e.g., some of the scales listed above only apply to individuals aged 0-18).

To overcome this suite of issues, a recent article by Hirdes et al. (2020) proposes the use of an integrated mental health scale which incorporates aspects of some of the most commonly used mental health screeners and assessments. Hirdes et al. (2020) discuss the interRAI Emergency Screener for Psychiatry (ESP), a comparatively brief 30-minute screening assessment specifically designed for use by MCTs during crisis situations and which has been validated cross-nationally. The authors note that the ESP can be used at the time of the crisis event, can be used for both adults and children, and that the full-form interRAI Community Mental Health (CMH) assessment - from which the ESP is derived - can be used once a patient has stabilized as well as when they are discharged. From a psychometric perspective, the various ESP scales have exhibited substantial levels of inter-rater reliability, internal consistency, face validity, content validity, convergent validity, criterion validity, and, importantly, predictive validity including studies where these assessment tools were used to predict re-hospitalization rates and suicidality (Perlman, Hirdes, and Vigod 2015; Neufeld et al., 2015; Hirdes et al. 2020). The use of an integrative screening and assessment tool specifically developed for use by MCTs could theoretically reduce the likelihood that users report assessment burden when surveyed soon after a crisis episode but the choice to use an integrative scale must be evaluated against (a) the comparative ease of access and licensing costs of other behavioral health scales and (b) the training required for administration of specific scales.

### ***Referrals and Dispositions***

Finally, the SAMHSA toolkit notes the need for MCTs to use the Level of Care Utilization System (LOCUS) to assess the severity of service-recipient crisis conditions and needs (SAMHSA 2020). LOCUS has been widely used across states and, in early tests, exhibited some degree of predictive validity (Sowers, George, and Thompson 1999). However, it is worth noting that there have not been many studies assessing other forms of reliability or validity of LOCUS despite wide usage as a disposition tool among mental health practitioners (though see Kimura, Yagi, and Yoshizumi 2013 and Thurber et al. 2018 for some preliminary studies).

## REFERENCES

- Alei, Angela, and Caroline Pierotti. 2016. "Literature Review: Crisis Services (Mobile Crisis Teams)." *Institute for Social Research*.
- Arnold, Bill R., and Yolanda E. Matus. 2000. "Test Translation and Cultural Equivalence Methodologies for Use with Diverse Populations." *Handbook of Multicultural Mental Health*. Academic Press: 121-136.

- Attkisson, C. Clifford, and Rebecca Zwick. 1982. "The Client Satisfaction Questionnaire: Psychometric Properties and Correlations with Service Utilization and Psychotherapy Outcome." *Evaluation and Program Planning* 5(3): 233-237.
- Baess, E. 2005. "Integrated Mobile Crisis Response Team (IMCRT): Review of Pairing Police with Mental Health Outreach Services." Victoria, BC: Vancouver Island Health Authority.
- Bailey, Katie, et al. 2018. "Barriers and Facilitators to Implementing an Urban Co-Responding Police-Mental Health Team." *Health Justice* 6(1): 1-12.
- Beck, Aaron T., Maria Kovacs, and Arlene Weissman. 1979. "Assessment of Suicidal Intention: The Scale for Suicide Ideation." *Journal of Consulting and Clinical Psychology* 47(2): 343-352.
- Bellamy, Chyrell, Timothy Schmutte, and Larry Davidson. 2017. "An Update on the Growing Evidence Base for Peer Support." *Mental Health and Social Inclusion*: 161-167.
- Bengelsdorf, Herbert, et al. 1984. "A Crisis Triage Rating Scale: Brief Dispositional Assessment of Patients at Risk for Hospitalization." *Journal of Nervous and Mental Disease* 72(7): 424-430.
- Bonyng, Eugene R., and Steven Thurber. 2008. "Development of Clinical Ratings for Crisis Assessment in Community Mental Health." *Brief Treatment and Crisis Intervention* 8(4): 304-312.
- Booth, Alison, et al. 2017. "Mental Health Training Programmes for Non-Mental Health Trained Professionals Coming into Contact with People with Mental Ill Health: A Systematic Review of Effectiveness." *BMC psychiatry* 17(196).
- Borum, R., P. Bartel, and A. Forth. 2002. *Manual for the Structured Assessment for Violence Risk in Youth (SAVRY): Consultation Version*. Tampa, Florida: University of South Florida, Florida Mental Health Institute.
- Boscarato, Kara, et al. 2014. "Consumer Experience of Formal Crisis-Response Services and Preferred Methods of Crisis Intervention." *International Journal of Mental Health Nursing* 23(4): 287-295.
- Broadbent, Marc, Heather Jarman, and Michael Berk. 2004. "Emergency Department Mental Health Triage Scales Improve Outcomes." *Journal of Evaluation in Clinical Practice* 10(1): 57-62.
- Cappelli, Mario, et al. 2012. "The HEADS-ED: A Rapid Mental Health Screening Tool for Pediatric Patients in the Emergency Department." *Pediatrics* 130(2): 321-327.
- Cappelli, Mario, et al. 2020. "The HEADS-ED: Evaluating the Clinical Use of a Brief, Action-Oriented, Pediatric Mental Health Screening Tool." *Pediatric Emergency Care* 36(1): 9-15.
- Compton, Michael, et al. 2008. "A Comprehensive Review of Extant Research on Crisis Intervention Team (CIT) Programs." *Journal of the American Academy of Psychiatry and the Law Online* 36(1): 47-55.
- Currier, Glenn W., Susan G. Fisher, and Eric D. Caine. 2010. "Mobile Crisis Team Intervention to Enhance Linkage of Discharged Suicidal Emergency Department Patients to Outpatient Psychiatric Services: A Randomized Controlled Trial." *Academic Emergency Medicine* 17(1): 36-43.
- Daggenvoorde, Thea H., Harm J. Gijsman, and Peter Goossens. 2018. "Emergency Care in Case of Acute Psychotic and/or Manic Symptoms: Lived Experiences of Patients and Their Families with the First

Interventions of a Mobile Crisis Team: A Phenomenological Study." *Perspectives in Psychiatric Care* 54(4): 462-468.

Dennis, Michael L., Ya-Fen Chan, and Rodney R. Funk. 2006. "Development and Validation of the GAIN Short Screener (GSS) for Internalizing, Externalizing and Substance Use Disorders and Crime/Violence Problems among Adolescents and Adults." *The American Journal on Addictions* 15: 80-91.

Department of Health Victoria. 2009. Victorian Mental Health Triage Scale. Available from: <https://www2.health.vic.gov.au/about/publications/policiesandguidelines/triage-scale-mental-healthservices>

Ebesutani, Chad, et al. 2012. "A Transportable Assessment Protocol for Prescribing Youth Psychosocial Treatments in Real-World Settings: Reducing Assessment Burden via Self-Report Scales." *Psychological Assessment* 24(1): 141-155.

Endicott, Jean, et al. 1976. "The Global Assessment Scale: A Procedure for Measuring Overall Severity of Psychiatric Disturbance." *Archives of General Psychiatry* 33(6): 766-771.

Fendrich, Michael, et al. 2019. "Impact of Mobile Crisis Services on Emergency Department Use Among Youths with Behavioral Health Service Needs." *Psychiatric Services* 70(10): 881-887.

Forchuk, Cheryl, et al. 2010. "Psychiatric Crisis Services in Three Communities." *Canadian Journal of Community Mental Health* 29(5): 73-86.

Gaziel-Yablowitz, Michal, and David G. Schwartz. 2018. "A Review and Assessment Framework for Mobile-Based Emergency Intervention Apps." *ACM Computing Surveys (CSUR)* 51(1): 1-32.

Goldman, Matthew L., et al. 2020. "A Novel Continuum-Based Framework for Translating Behavioral Health Integration to Primary Care Settings." *Translational Behavioral Medicine* 10(3): 580-589.

Heyman, Inga, and Emma McGeough. 2018. "Cross-Disciplinary Partnerships Between Police and Health Services for Mental Health Care." *Journal of Psychiatric and Mental Health Nursing* 25(5): 283-284.

Hirdes, John P., et al. 2020. "The interRAI suite of Mental Health Assessment Instruments: An Integrated System for The Continuum of Care." *Frontiers in Psychiatry* 10: 926.

Kimura, Taku, Fukashi Yagi, and Akira Yoshizumi. 2013. "Application of Level of Care Utilization System for Psychiatric and Addiction Services (LOCUS) to Psychiatric Practice in Japan: A Preliminary Assessment of Validity and Sensitivity to Change." *Community Mental Health Journal* 49(4): 477-491.

Kisely, Stephen, et al. 2010. "A Controlled Before-and-After Evaluation of a Mobile Crisis Partnership Between Mental Health and Police Services in Nova Scotia." *The Canadian Journal of Psychiatry* 55(10): 662-668.

Kroenke, Kurt, Robert L. Spitzer, and Janet BW Williams. 2001. "The PHQ-9: Validity of a Brief Depression Severity Measure." *Journal of General Internal Medicine* 16(9): 606-613.

Lamanna, Denise, et al. 2018. "Co-Responding Police–Mental Health Programmes: Service User Experiences and Outcomes in a Large Urban Centre." *International Journal of Mental Health Nursing* 27(2): 891-900.

- Lancaster, Andrew. 2016. "Evidence for Joint Police and Mental Health Responses for People in Mental Health Crisis." *Mental Health Practice* 19(10): 20-26.
- Landinez, Rosa. 2019. *Underutilization of Mobile Crisis Team Services from Lack of Referrals*. Diss. Chatham University.
- Morant, Nicola, et al. 2017. "Crisis Resolution and Home Treatment: Stakeholders' Views on Critical Ingredients and Implementation in England." *BMC Psychiatry* 17(1): 254-267.
- Murphy, Suzanne M., et al. 2015. "Crisis Intervention for People with Severe Mental Illnesses." *Cochrane Database of Systematic Reviews* 12.
- McFarland, Sam G. 1981. "Effects of Question Order on Survey Responses." *Public Opinion Quarterly* 45(2): 208-215.
- Neufeld, Eva, et al. 2015. "A Longitudinal Examination of Rural Status and Suicide Risk." *Healthcare Management Forum*. 28(4).
- Ogles, Benjamin M., et al. 2001. "The Ohio Scales: Practical Outcome Assessment." *Journal of Child and Family Studies* 10(2): 199-212.
- Park, Alice, et al. 2019. "Models of Mental Health Triage for Individuals Coming to the Attention of the Police Who May Be Experiencing Mental Health Crisis: A Scoping Review." *Policing: A Journal of Policy and Practice* 1-37.
- Perlman, Christopher M., John P. Hirdes, and Simone Vigod. 2015. "Psychiatric Rehospitalization: Development of a Person-Level Indicator for Care Planning and Quality Assurance." *The Primary Care Companion for CNS Disorders* 17(4).
- Peterson, Jillian, and James Densley. 2018. "Is Crisis Intervention Team (CIT) Training Evidence- Based Practice? A Systematic Review." *Journal of Crime and Justice* 41(5): 521-534.
- Peterson, Jillian, James Densley, and Gina Erickson. 2020. "Evaluation of 'The R-Model' Crisis Intervention De-escalation Training for Law Enforcement." *The Police Journal* 93(4): 271-289.
- Posner, Kelly, et al. 2011. "The Columbia–Suicide Severity Rating Scale: Initial Validity and Internal Consistency Findings from Three Multisite Studies with Adolescents and Adults." *American Journal of Psychiatry* 168(12): 1266-1277.
- Priebe, Stefan, et al. 1999. "Application and Results of the Manchester Short Assessment of Quality of Life (MANSA)." *International Journal of Social Psychiatry* 45(1): 7-12.
- Puntis, Stephen, et al. 2018. "A Systematic Review of Co-Responder Models of Police Mental Health 'Street' Triage." *BMC Psychiatry* 18(1): 1-11.
- Rodriguez N, Steinberg A, Pynoos RS. 1998. *UCLA Post Traumatic Stress Disorder Reaction Index for DSM-IV, Child, Adolescent, and Parent Versions*. Los Angeles, CA: UCLA Trauma Psychiatry Service.
- Rogers, James R., Ralph A. Alexander, and Linda Mezydlo Subich. 1994. "Development and Psychometric Analysis of the Suicide Assessment Checklist." *Journal of Mental Health Counseling* 16(3): 352–368

- Russell, Dale W., and Cristel Antonia Russell. 2019. "The Stress Bias in Mental Health Reporting: Death Anxiety Biases Mental Health Self-Assessments Amongst Deployed Soldiers." *Psychological Services*.
- Sands, Natisha, et al. 2016. "Development and Interrater Reliability of the UK Mental Health Triage Scale." *International Journal of Mental Health Nursing* 25(4): 330-336.
- Scott, Roger L. 2000. "Evaluation of a Mobile Crisis Program: Effectiveness, Efficiency, and Consumer Satisfaction." *Psychiatric Services* 51(9): 1153-1156.
- Shapiro, Gilla K., et al. 2015. "Co-responding Police-Mental Health Programs: Review." *Administration and Policy in Mental Health and Mental Health Services Research* 42(5): 606-620.
- Sowers, Wesley, Charles George, and Kenneth Thompson. 1999. "Level of Care Utilization System for Psychiatric and Addiction Services (LOCUS): A Preliminary Assessment of Reliability and Validity." *Community Mental Health Journal* 35(6): 545-563.
- Thurber, Steven, et al. 2018. "The Relationship Between the INTERMED Patient Complexity Instrument and Level of Care Utilisation System (LOCUS)." *International Journal of Psychiatry in Clinical Practice* 22(1): 80-82.
- Turner, P. M., and T. J. Turner. 1991. "Validation of the Crisis Triage Rating Scale for Psychiatric Emergencies." *The Canadian Journal of Psychiatry* 36(9): 651-654.
- Young, Nicholas D., et al. 2020. "Warm Handoff, or Cold Shoulder? An Analysis of Handoffs for Primary Care Behavioral Health Consultation on Patient Engagement and Systems Utilization." *Clinical Practice in Pediatric Psychology* 8(3): 241-246.
- Vanderploeg, Jeffrey J., et al. 2016. "Mobile Crisis Services for Children and Families: Advancing a Community-Based Model in Connecticut." *Children and Youth Services Review* 71: 103-109.

## Appendix A: A Suggested Framework for Collecting and Reporting Co-Response Models<sup>4</sup>

**Table 3** A suggested framework for collecting and reporting co-response triage models

Topic	Variable to report	Description of variable and reason for reporting
Identify as co-response triage service	Name of service	Give the name of the service to enable grouping of models and their comparison.
	Identify as co-response model of triage	Include a sentence in the service description to signpost to readers and researchers that the service is a co-response model of police mental health triage.
Model characteristics	Model of co-response	Define the model of how the mental health professional assists the police officer during the incident. <i>Ride-along</i> models are those in which the police officer (PO) and mental health (MH) worker attend the incident together in a vehicle. <i>Ride-separate</i> models are when PO and MH worker arrive at the incident separately. <i>Telephone Support</i> models are those in which the MH worker provides assistance via a telephone or radio. Services that provide a combination of these models should specify when each model is used.
	Method of referral to co-response team	Describe how crisis incidents are referred to the triage team (e.g. emergency response, direct from police officers, publicly available direct phone line), and from whom triage can take referrals (e.g. police officers, other emergency services, mental health services, the public, etc.)
	Timing of response	Describe whether the co-response team acts as a <i>first responder</i> , where the team can be referred to and attend incidents prior to any other police involvement, or <i>second responder</i> , where the team is referred following an initial police response.
	Team staffing and roles	Outline the team composition and responsibilities
	Days and hours of operation	Describe the working hours of members of the co-response team. Be explicit if hours differ for different types of response, or members of response.
	Team Location	Describe where each member of the team is located during a usual triage shift.
	Vehicle involved in triage response (if any)	If applicable, describe how the co-response team attend the triage incident (eg. in a marked police car, unmarked car, ambulance, etc.)
Local context	Geography and population	Describe the geographic elements (eg. rural versus city, large area vs small area) and local population to improve readers' understanding of the local context in to which the triage service is placed.
	Local mental health provision and linked partner agencies	Describe local mental health provision and agencies working closely with the triage team (if any) to improve understanding of where triage fits within crisis services and wider mental health provision

<sup>4</sup> Table 3 taken from Puntis et al. (2018: 9).

## Appendix B: List of Assessment Tools

Psychosocial Tool	Number of Items and Format	Administration and Scoring Time	Psychometric Notes
<p><i>Columbia Suicide Severity Rating Scale (C-SSRS) (Posner et al. 2011)</i></p>	<p>6 items (Yes/No Format)</p>	<p>N/A</p>	<p><a href="#"><u>The C-SSRS's validity relative to other measures of suicidal ideation and behavior and the internal consistency of its intensity of ideation subscale were analyzed in three multisite studies: a treatment study of adolescent suicide attempters (N=124); a medication efficacy trial with depressed adolescents (N=312); and a study of adults presenting to an emergency department for psychiatric reasons (N=237).</u></a></p> <p><a href="#"><u>The C-SSRS demonstrated good convergent and divergent validity with other multi-informant suicidal ideation and behavior scales and had high sensitivity and specificity for suicidal behavior classifications compared with another behavior scale and an independent suicide evaluation board. Both the ideation and behavior subscales were sensitive to change over time. The intensity of ideation subscale demonstrated moderate to strong internal consistency. In the adolescent suicide attempters study, worst-point lifetime suicidal ideation on the C-SSRS predicted suicide attempts during the study, whereas the Scale for Suicide Ideation did not. Participants with the two highest levels of ideation severity (intent or intent with plan) at baseline had higher odds for attempting suicide during the study.</u></a></p>

			<p><a href="#">These findings suggest that the C-SSRS is suitable for assessment of suicidal ideation and behavior in clinical and research settings.</a></p> <p>Available in 103 different languages.</p>
<p><i>Suicide Assessment Checklist (Rogers 1994)</i></p>	21 items	N/A	<p><a href="#">Analog interrater reliability (RLB) estimates ranged from .83 to .84, and an analog 4-wk test-retest design resulted in an RLB estimate of .82.</a></p> <p><a href="#">Internal consistency reliability for rated items of the SAC was .87. Supporting construct-related validity, total score differences were found in the expected directions as a function of referral reason. Convergent validity was based on observed correlations between selected SAC items and conceptually similar items on the Beck Depression Inventory (A. T. Beck, 1970). Supporting criterion-related validity, total score differences in the expected directions were found as a function of disposition setting.</a></p>
<p><i>Scale of Suicide Ideation (SSI) (Beck, Kovacs, and Weissman 1979)</i></p>	19 items	5-10 minutes	<p><a href="#">Cronbach's <math>\alpha</math> for the whole SSI was 0.95. The SSI total score differentiated patients and controls, and increased statistically significantly in classes with increasing severity of suicidality derived from the suicidality items of the K-SADS-PL diagnostic interview. Varimax-rotated principal component analysis of the SSI items yielded three theoretically coherent factors suggesting construct validity. Area under the receiver operating characteristic (ROC) curve was 0.84 for the whole sample and 0.80 for the patient sample. The optimal cutoff threshold for the SSI total score was 3/4 yielding sensitivity of 75% and specificity of 88.9% in this population.</a></p>

<i>The UCLA Post-Traumatic Stress Index (Rodriguez, Steinberg, and Pynoos 1998)</i>	20-22 items	20-30 minutes	Good test-retest coefficient of 0.84. A cutoff of 37 provides 0.93 sensitivity and 0.87 specificity.
<i>The Structured Assessment of Violence Risk in Youth (SAVRY) (Borum, Bartel, and Forth 2002)</i>	24 items	10-15 minutes	<a href="#">The predictive validity of the SAVRY was moderate for both violent and general recidivism, but both showed incremental validity in predicting outcomes compared with the PCL Psychopathy Checklist: Youth Version.</a>  Intended use for youth aged 12-18 years.
<i>The Global Appraisal of Individual Needs – Short Screener (GAIN-SS) (Dennis et al. 2006)</i>	16 items	5 minutes	<a href="#">In this project, the internal consistency, validity, and optimal screening cutoff scores were examined in 95 adolescents, most of whom were receiving treatment, and suffered from internalizing (52%; n = 49), externalizing (66%, n = 63), and co-occurring substance use disorders (55%, n = 56). Results indicated adequate internal consistency and overall and subscale construct validity. Receiver operating characteristics (ROC) analyses revealed that the GSS substance use disorder subscale had adequate sensitivity (88%) and specificity (89%). Psychiatric disorder subscales performed less well.</a>
<i>The Ohio Youth Problems, Functioning, and Satisfaction Scales (Ohio Scales) (Ogles et al. 2001)</i>	76 items	N/A	<a href="#">The authors explored the convergent validity of the Ohio Scales by examining the relationship between the scales and subscales' scores generated by parents, agency workers, and youth of the Ohio Scales and DSM—IV diagnoses in a sample of 3,569 youth ages 5 to 18 across parent, agency worker, and youth reports. The Ohio Scales demonstrated evidence of convergent validity when predicted relationships between the scales and diagnoses were examined. The Ohio Scales also showed</a>

			<a href="#">evidence of differentiating among broad diagnostic categories.</a>
<i>HEADS-ED Scale (Cappelli et al. 2012; Cappelli et al. 2020)</i>	7 items	<3 minutes	<p><a href="#">Initial findings indicate that HEADS-ED is psychometrically sound with evidence of criterion, concurrent and predictive validity, and interrater reliability. The HEADS-ED was correlated with a comprehensive clinician rating of mental health strengths and needs (CANS-MH 3.0) as well as ratings of depression by youths using the Children’s Depression Inventory (CDI). The study also supported the predictive validity of the tool. The total score from the HEADS-ED indicated meaningfully and statistically different mean scores for patients who were referred for admission to an inpatient psychiatric unit (above the 75th percentile) and those who were referred for consultation (above the 50th percentile).</a></p> <p><a href="#">Using an algorithm of a total HEADS-ED score of greater than 7 and a suicidal risk factor of 2, we determined a sensitivity of 81.8 and a specificity of 87 for predicting admission.</a></p>
<i>Global Assessment Scale (GAS) (Endicott et al. 1976)</i>	N/A	N/A	<p><a href="#">The Global Assessment Scale (GAS) is a rating scale for evaluating the overall functioning of a subject during a specified time period on a continuum from psychological or psychiatric sickness to health.</a></p> <p><a href="#">In five studies encompassing the range of populations to which measures of overall severity of illness are likely to be applied, the GAS was found to have good reliability. GAS ratings were found to have greater sensitivity to change over time than did other ratings of overall severity or specific symptom dimensions. Former inpatients in the</a></p>

			<p><a href="#">community with a GAS rating below 40 had a higher probability of readmission to the hospital than did patients with higher GAS scores.</a></p> <p><a href="#">The relative simplicity, reliability, and validity of the GAS suggests that it would be useful in a wide variety of clinical and research settings.</a></p> <p><a href="#">Results showed such team GAS ratings to have good reliability. As for validity, team GAS predicted the outcomes of court hearings at two stages of the civil commitment process and showed construct validity in its relevant correlations with the Psychotic Inpatient Profile. A predicted association between team GAS and a ward atmosphere measure was not obtained. Taken as a whole, the results support the use of team GAS ratings in inpatient settings.</a></p>
<i>PHQ – 9 (Kroenke, Spitzer, and Williams 2001)</i>	9 items (Self-Report)	< 5 minutes to administer; < 3 minutes to score	Excellent internal reliability and test-retest reliability. Sensitivity and specificity of the PHQ-9 for diagnosing major depression were 74% and 91%, respectively, with a score of 10 or higher. For the PHQ-9 a score of 10 or higher detected more cases of major depression than the PHQ determination of major depression originally described by Spitzer et al. in 1999.
<i>Client Satisfaction Questionnaire (CSQ – 8) (Attkisson and Zwick 1982)</i>	8 items (Self-Report)	N/A	<a href="#">The CSQ – 8 scale was strongly correlated with the Treatment Perceptions Questionnaire, TPQ, suggesting high concurrent validity. However, while the TPQ was normally distributed, the CSQ-8 was highly negatively skewed. Significant</a>

			<p><a href="#">associations were found between the CSQ-8 and cross-sectional process measures.</a> Results suggest that that CSQ-8 is an <a href="#">appropriate measure to be used in residential substance abuse treatment settings.</a> CSQ-8 is available in over 30 languages including English and Spanish.</p>
<p><i>Manchester Short Assessment of Quality of Life (MANSA) (Priebe et al. 1999)</i></p>	16 items (Self-Report)	N/A	<p>Correlations between subjective quality of life scores on MANSA and LQLP were all 0.83 or higher (0.94 for the satisfaction mean score). Cronbach's alpha for satisfaction ratings was 0.74, and association with psychopathology was in line with results for LQLP as reported in the literature.</p>
<p><i>interRAI Emergency Screener for Psychiatry (ESP)</i></p>	141 items (Subset of MH and CMH)	30 minutes	<p><a href="#">In a 12-nation study of inter-rater reliability of five different interRAI instruments (MH, LTCF, HC, Palliative Care, and interRAI Post-Acute Care), the mean weighted kappa for the core items common to all instruments was 0.75 and the kappa of the specialized mental health items was 0.64. Both results suggest “substantial” inter-rater reliability.</a></p> <p><a href="#">Arguably the most important (and difficult) form of validity to establish for an assessment system is predictive validity. Presumably, the ultimate purpose of assessment is to guide interventions that will have an impact on a future clinical trajectory of change. This approach was used extensively to identify triggering rules for interRAI's mental health care planning protocols (see discussion below). Examples of publications reporting on this type of validity include studies of inpatient violence, re-hospitalization; and suicide behaviors.</a></p>

			Target Audience: Adults aged 18 years and older. Children version developed.
--	--	--	--

Center for Applied Research and Analysis  
INSTITUTE FOR SOCIAL RESEARCH  
UNIVERSITY OF NEW MEXICO  
November 2020