



Reliability and Validity of START and LSI-R Assessments Completed in Mental Health Diversion Programs

Evan M. Lowder¹, Sarah L. Desmarais¹, Richard A. Van Dorn², Candalyn B. Rade¹, Robin P. Telford³, & John Pettila³

¹North Carolina State University, ²RTI International, ³University of South Florida



Introduction

More than 500 mental health diversion programs exist across the U.S. (Case et al., 2009). Completion of a risk assessment at the time of referral is essential to effective case management (Desmarais, Van Dorn et al., 2012); however, few studies have examined the reliability and validity of risk assessments completed in these programs.

The Short-Term Assessment of Risk and Treatability (START; Webster et al., 2009) is one instrument being piloted in a mental health jail diversion program (Desmarais, Van Dorn et al., 2012). The START is a structured professional judgment instrument that guides the assessment of short-term risks among adults with behavioral health disorders (Webster et al., 2009). START assessments have demonstrated high inter-rater reliability and validity in psychiatric populations (Desmarais, Nicholls, et al., 2012; Nicholls et al., 2006.). The Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 2001) is a widely-used actuarial instrument for assessing general offending risk. Research supports the validity of LSI-R assessments among adult offenders (Vose et al., 2008).

To our knowledge, no study has examined the validity of START or LSI-R assessments completed on mental health diversion program clients, nor have these two instruments been compared to each other.

The Present Study

We examined the reliability and validity of START and LSI-R assessments completed on mental health diversion program clients. Specifically, we examined: 1) inter-rater reliability of START and LSI-R assessments; 2) associations between START vulnerability and LSI-R total scores, and START general offending risk estimates (piloted in this study) and LSI-R risk classifications; and 3) validity of START and LSI-R total scores and risk estimates in predicting self-reported recidivism.

Methods

Study Context

Data were collected as part of an ongoing evaluation of behavioral health services delivered to clients participating in mental health jail diversion programs operated by the Criminal Mental Health Project (CMHP) in Miami, FL. The CMHP is a court-based initiative that identifies and diverts adults with mental illnesses and co-occurring substance use disorders into community-based treatment.

Sample

The sample includes 94 diversion clients (79.9% male). The average age was 35.31 years ($SD = 14.37$). The sample comprised felony (46.8%) and misdemeanor diversion program clients (53.2%). Half (54.3%) identified as Hispanic or Latino and half identified as African American (45.7%). Participants reported an average of 6.74 jail bookings ($SD = 5.44$) in the past three years and an average of 16.17 ($SD = 15.04$) lifetime jail bookings. Consistent with inclusion criteria, the presenting problems were primarily schizophrenia (36.2%), psychosis NOS (25.5%), and bipolar disorder (20.2%).

Measures

START. The START vulnerability and strength total scores were summed based on vulnerability and strength ratings for 20 START items, respectively. The general offending risk estimate, piloted in this study, categorizes risk as low, moderate, or high based on the assessors structured professional judgment.

LSI-R. The LSI-R total score is a sum of 54 items from which five risk categories are created. Due to sample size restrictions, these were collapsed into three categories: low/low-moderate, moderate, and moderate-high/high risk (Singh et al., 2013).

Recidivism. Measures included *self-reported arrests, incarceration, and any criminal activity* in the past three months. Both binary (yes/no) and continuous measures (count) were used.

Procedures

Semi-structured interviews were conducted at baseline and 3-month follow-up. At baseline, START and LSI-R assessments were completed based on self-report and official records. At follow-up, participants self-reported recidivism in the past three months.

Results

Descriptive Statistics

Table 1 presents the descriptive statistics for risk assessments and recidivism measures. In summary, participants showed few strengths and considerable risks, particularly on the LSI-R total scores. Overall, rates of recidivism were low.

Reliability

Calculated on a subset of 22 participants, inter-rater reliability (IRR) was good for START strength ($ICC_2 = .77$) and vulnerability total scores ($ICC_2 = .78$), and adequate for START general offending risk estimates (63.6% agreement). IRR was good for LSI-R total scores ($ICC_2 = .71$). All $p < .01$.

Convergent and Divergent Validity

Results showed significant associations between START vulnerability and LSI-R total scores, $r(91) = .23, p = .025$. Agreement between START and LSI-R risk classifications was low ($\kappa = .10, p = .152$): LSI-R assessments identified more participants as moderate/high risk than START assessments (91.4% vs. 75.3%). The relationship between START strength and LSI-R total scores was not significant, $p = .272$.

Table 1. Descriptive Statistics

RISK ASSESSMENTS		
Total Scores	M	SD
START Vulnerability (N = 93)	19.10	6.29
START Strength (N = 93)	14.74	7.33
LSI-R (N = 94)	30.62	4.83
Risk Estimates		
START (N = 93)	n	%
Low	23	24.5
Moderate	50	53.2
High	20	21.3
LSI-R (N = 94)		
Low/Low-Moderate	8	8.5
Moderate	64	68.1
Moderate-High/High	22	23.4
RECIDIVISM		
Continuous Measures	M	SD
Arrests (N = 58)	0.22	0.53
Days Incarcerated (N = 59)	4.87	12.96
Days of Criminal Activity (N = 59)	0.25	1.09
Dichotomous Measures (yes = 1)	n	%
Arrest (N = 59)	12	12.8
Incarceration (N = 59)	14	14.9
Criminal Activity (N = 59)	15	16.0

Predictive Validity

Tables 2 and 3 present the findings of the predictive validity analyses for the continuous and dichotomous measures, respectively. Results showed that START strength total scores, START and LSI-R risk estimates trended toward fair to moderate associations with number of arrests and days incarcerated (see Table 2).

Table 2. Associations Between LSI-R and START Assessments and Continuous Recidivism Measures

	TOTAL SCORES						RISK ESTIMATES			
	LSI-R		START Vulnerability		START Strength		LSI-R		START	
	β	R^2	β	R^2	β	R^2	β	R^2	β	R^2
Arrests	.09	.01	.03	<.01	-.22 [‡]	.05	.23 [‡]	.05	.22 [‡]	.05
Incarceration	.12	.02	.10	.01	-.26 [‡]	.07	.25 [‡]	.06	.26 [‡]	.06
Any Criminal Activity	-.07	<.01	.02	<.01	-.13	.02	.07	<.01	.23 [‡]	.05

Notes. N = 59. * $p < .05$. ** $p < .01$. *** $p < .001$. ‡ $p < .10$.

START strength total scores and LSI-R risk estimates were significantly associated with any incarceration or criminal activity at follow-up (see Table 3). Additionally, START strength total scores and LSI-R risk estimates trended toward moderate associations with any arrest at follow-up (see Table 3).

Table 3. Associations Between LSI-R and START Assessments and Dichotomous Recidivism Measures

	TOTAL SCORES						RISK ESTIMATES	
	LSI-R		START Vulnerability		START Strength		LSI-R	START
	AUC (SE)	95% CI	AUC (SE)	95% CI	AUC (SE)	95% CI	$\chi^2(2)$	$\chi^2(2)$
Arrests	.60 (.10)	.40-.79	.50 (.09)	.31-.68	.68 [‡] (.07)	.54-.82	5.84 [‡]	0.79
Incarceration	.61 (.10)	.42-.80	.47 (.09)	.29-.65	.73* (.07)	.60-.86	7.26*	1.61
Any Criminal Activity	.58 (.10)	.40-.77	.46 (.09)	.28-.63	.72* (.06)	.60-.85	6.14*	0.93

Notes. AUC = area under the curve. CI = confidence interval. N = 58. * $p < .05$. ** $p < .01$. *** $p < .001$. ‡ $p < .10$. START Strength scores were reversed scored for ROC analyses.

Discussion

Findings provide some support for the reliability and validity of LSI-R and START assessments completed on mental health jail diversion program clients. Additionally, they provide some evidence for the predictive validity of the START general offending risk estimate. Observed differences in validity may be due to differences between the instruments; for example, the START is designed to assess short-term risk and the LSI-R, longer-term risk. Findings suggest that protective factors, included in the START but not LSI-R, may be particularly relevant to recidivism among 'frequent flyers' in the criminal justice system. However, results may be limited by our small sample size at follow-up, low base rates, and short follow-up timeframe.

Acknowledgments

This project is funded by the Bristol-Meyers Squibb Foundation. We thank the 11th Judicial District Criminal Mental Health Project (CMHP) for its cooperation and support.