

Name of Tool	Short Term Assessment of Risk and Treatability (START)
Category	Violence Risk (Validated)
Author / Publisher	Webster and colleagues
Year	2009

Description

- The START is a 20-item structured professional judgement tool designed to structure regular clinical assessments within inpatient and community contexts.
- The tool is intended to assess, document, communicate and manage risk across diverse settings.
- The 20 items included in the START are drawn from research that have shown these variables to be associated with seven risk estimates/adverse outcomes to individuals with mental health problems and personality dysfunctions, as well as persons who come into conflict with the law.
- Assessors code the items according to two scales presented in the tool: (1) Strength and (2) Vulnerability.
- The items are rated on a 3-point Likert scale from 0 to 2 and can be coded as both a strength and a vulnerability.
- The START includes seven risk estimates which include violence, suicide and self-harm. The risk estimates are derived from the consideration of the ratings from the strength and vulnerability scales.
- The tool was initially designed to capture dynamic vulnerabilities and strengths while generating a framework for periodic assessment of risk to inform clinical progress reviews. It should inform treatment, daily management and decision-making.
- The START is intended for use with adults diagnosed with mental, personality and substance-related disorders. It is relevant to inpatient and community psychiatric, forensic and correctional populations.

Age Appropriateness

16+

Assessor Qualifications

Experienced clinicians from a mental health background. Assessors are required to have participated in relevant training for this tool. It can be completed either by an individual practitioner or jointly by a clinical team via group discussion and reaching a consensus.

Strengths

- The tool considers strengths rather than being purely risk orientated ([Nicholls et al., 2006](#)).
- Collins et al. (2008) found that clinicians deemed START as appropriate, easy to use and clinically useful.
- START is intended for use in both inpatient units and outpatient services.

Empirical Grounding

- The manual claims that the tool is grounded in the HCR-20 and relevant studies of acute violence (Webster et al., 2004). The authors drew upon research from civil psychiatry, forensic psychiatry and corrections reflecting studies from both institutional and community settings.
- The START is a concise clinical guide for the dynamic assessment of short-term (i.e. weeks to months) risk for violence (to self and others) and treatability.

Inter-Rater Reliability

a) UK Research

- [Timmins, Evans and Tully \(2018\)](#) assessed the inter-rater reliability of START across disciplines, recruiting psychiatrists, mental health nurses, psychologists and occupational therapists to rate 20 case items and 7 risk estimates for two test cases. Good to excellent IRR was found for START items; whilst moderate-to-poor IRR was found for risk estimates amongst raters. There were clear differences between disciplines at item levels, highlighting the importance of collaborating as a team when completing risk assessments.

b) International Research

- [Desmarais et al. \(2012a\)](#) found ICCs of .93 for strength scores, .95 for vulnerability scores and .85 for risk estimates respectively.
- [Nicholls et al. \(2006\)](#) - the START attained excellent inter-rater reliability (ICC) in various settings; Psychiatry (.80), Nursing (.88) and Social work (.92).
- [Dickens and O'Shea \(2015\)](#) reported "Inter-rater reliability for coding the SOS (Start Outcome Scale) from progress notes was in the excellent range: Cohen's Kappa ranged from .83 to 1.00, the lowest being for self-neglect and the highest for self-harm and physical aggression."
- [Viljoen et al. \(2011\)](#) - the START strength and vulnerability scale total scores attained good ICC values of .62, and .68.
- [Wilson et al. \(2010\)](#) - found ICCs of .85, .90 and .81 for the strength, vulnerability and risk estimates respectively.
- [Crocker et al. \(2011\)](#) carried out START assessments on 42 individuals at a civil psychiatric unit in Canada. An inter-rater reliability check on six patients six months later found that there was low IRR for total risk score of .38, whereas total strength score was strong at .81.

- A study by [O'Shea, Picchioni and Dickens \(2016\)](#) of 22 adults in a secure mental hospital found that the inter-rater reliability for START items was in the excellent range.

Validation History

General Predictive Accuracy



a) UK Research

- [Braithwaite et al. \(2010\)](#) suggested there was partial support for the predictive validity of the instrument. Both the strength and vulnerability scales significantly predicted aggression against others and suicidality (AUC= .65 for each scale and behaviour). AUCs of .67 and .63 were generated for substance abuse in the strength and vulnerability scales respectively. Neither scale, however, significantly predicted the occurrence of self-harm, suicidality, self-neglect or victimisation (AUCs ranging from .52 - .58).

- [Gray et al. \(2011\)](#) tested the START in a limited population study of 51 mentally disordered patients. The SPJ scores were able to predict violence to others, verbal aggression, self-harm and victimisation (AUCs of .65, .70, .86 and .67 respectively). The strength and risk scores varied in their ability to predict certain behaviours. The strength scores were poor predictors for all behaviours bar self-harm (AUC= .61), with an AUC range of .21-.47. The risk scores were better predictors with an AUC range of .60-.74 for all behaviours; the only exception to this is for self-harm which generated an AUC of .48.

b) International Research

- [Crocker et al. \(2011\)](#) found that whilst START total risk scores showed good predictive accuracy in relation to physical aggression for periods of 1 and 3 months (AUC ranging from .65-.77), they were not as accurate for the long-term of 6 to 12 months. Individuals displaying physical and property aggression had higher risk and lower strength scores on the START.

- [O'Shea, Picchioni and Dickens \(2016\)](#) found that the inclusion of strengths improved the predictive accuracy of the START tool. The percentage of cases correctly classified increased from 0.6% to 4.4%. The specific risk estimates scale showed increased predictive accuracy over both the vulnerability and strength scales, showing moderate to large predictive accuracy for all behaviours (AUCs range from .640-.783), bar self-neglect (AUC of .546).

	<ul style="list-style-type: none"> • In a study examining aggression data retrieved from institutional records, START strength and vulnerability total scores predicted all forms of aggression, bar physical aggression towards objects, Moderate to large effect sizes were generated for any aggression, verbal aggression and physical behaviours (others) with AUCs ranging from .65-.90. For physical aggression against objects, an AUC of .62 was generated in the strength total score (Cartwright et al., 2018). • de Vogel, Bruggeman and Lancel (2019) coded file information for 78 female forensic psychiatric patients using a number of structured professional judgement tools. The START Vulnerability scores showed moderate and large predictive accuracy for all recidivism in medium and long term follow-ups (AUCs of 0.748 and 0.698 respectively), as well as for violent recidivism (AUCs of .697 and .704 for medium and long term respectively).
--	--

Validation History	
Applicability: Females	
a) UK Research	<ul style="list-style-type: none"> • O'Shea and Dickens (2015) found START was a stronger predictor of aggression and self-harm in women than men. • Quinn et al. (2013) found significant predictive validity for adverse incidents at the one month time point and this then diminished over time. Females were rated as having significantly less strengths and more risks than males.
b) International Research	<ul style="list-style-type: none"> • Viljoen et al. (2011) - in a 3-year follow-up in a sample of female forensic patients, the START strength and vulnerability scores showed moderate to large AUCs at .70 and .80 respectively The results show the START scales were predictive of successful reintegration into the community (defined as the absence of readmission to hospital and the presence of an absolute discharge decision) in a sample of female forensic patients.

Validation History	
Applicability: Ethnic Minorities	
No Empirical Evidence Available	

Validation History

Applicability: Mental Disorders

a) UK Research

- [Gray et al. \(2011\)](#) - the vulnerability scale was moderately predictive of violence to others (AUC =.68). The strength scale had a significant negative correlation with violence to others (r=-.42) and a corresponding AUC value of .21. Low scores on the strength scale were, thus, predictive of violence.

- Predictive validity of START was evident when [Marriott et al. \(2017\)](#) administered the tool to 527 inpatients within a secure mental health facility in the United Kingdom.

- [Quinn et al. \(2013\)](#) discovered that START scores were capable of distinguishing between those with mental disorders at the various stages of their care pathways.

- [Alderman, Major and Brooks \(2016\)](#) used the START to examine 4559 aggression recordings related to 76 patients with an acquired brain injury. The START risk of violence was classed as low and high for 50% and 13.7% of the sample respectively; suggesting the need for specific tools to be developed for use in patients with ABI.

b) International Research

- [Crocker et al. \(2011\)](#) carried out a longitudinal study, which indicated that START was well integrated into a Canadian unit's administrative activities.

- [Wilson et al. \(2010\)](#) - in a 12-month follow up, the strength and vulnerability total scores and the final risk estimates significantly predicted any aggressive acts with AUCs ranging from .82 to .89.

- [Chu et al. \(2011\)](#) - in a 1-month follow period, the START vulnerability total scores attained high AUC values in predicting inpatient aggression (.76), interpersonal violence (.78) and verbal threat (.77). Similarly the strength total scores predicted inpatient aggression (.71) and interpersonal violence (.75) but not verbal threat.

- [Braithwaite et al. \(2010\)](#) - the vulnerability scale significantly predicted physical aggression against others (AUC = .66) in a 2-year follow-up period.

- [Nicholls et al. \(2006\)](#) - START generated moderate to high AUC values for a broad range of aggressive behaviours in a psychiatric hospital: verbal aggression against others (.72), physical aggression against objects

(.67), physical aggression against others (.70) and sexual inappropriateness.

Contribution to Risk Practice

- The START has the ability to create awareness of risk factors and strengths presented by the individual. Findings from previous research also suggest that it may be useful for distinguishing between types of patients ([Nicholls et al., 2006](#); [Quinn et al., 2013](#)).
- The START includes dynamic factors and the strengths of individuals, which could inform offence analyses and risk formulations.
- The use of the START can help identify factors that are important targets for treatment, intervention and management planning.
- Repeated assessments using the START can aid assessors in monitoring changes in risk level and identify necessary changes in risk management strategies.
- The START can aid the assessor in examining potential improvement/deterioration in identified risk, responsivity and protective factors which, in turn, can also inform risk management strategies. Further, the tool allows for other harmful scenarios to be considered for individuals, e.g. suicide, substance abuse, self-harm and self-neglect ([O'Shea, Picchioni and Dickens, 2016](#)).
- [O'Shea and Dickens \(2015\)](#) reported: "The study provides limited support for the START by demonstrating the predictive validity of its specific risk estimates for substance abuse and unauthorised leave. High negative predictive values suggest the tool may be of most utility in screening out low risk individuals from unnecessary restrictive interventions; very low positive predictive values suggest caution before implementing restrictive interventions in those rated at elevated risk."
- START is routinely used within forensic mental health populations in the United Kingdom and is recommended by the *Department of Health*.
- Staff members at a forensic high secure unit in Norway were surveyed about the START. It was felt by 68% of respondents that the existing and potential needs of patients were covered by the tool. Moreover, 73% agreed that using the START tool contributed to a more systematic risk assessment and management process ([Kroppan et al., 2011](#)).

Other Considerations

- [Doyle et al. \(2008\)](#) reported uncertainty over time frame in which risk and strengths are applied from a survey conducted with users of the START.
- [Dickens and O'Shea \(2015\)](#) suggested for lower risk patients assessment at 3 month intervals was appropriate. For those with elevated risk rating more frequent assessments were warranted.
- The START can be completed by a single clinician or by the patient's multi-disciplinary team.
- Fewer validation studies have been conducted on samples that consist solely of female patients and patients of other ethnic backgrounds.
- An electronic START Integrated Treatment Plan (START ITP) has been developed and is being pilot tested in Canada (Leech, personal communication, January 2013).
- Research is ongoing for the START and its use in different settings (e.g. jail diversion programs, [Desmarais et al., 2012a](#)).
- An abbreviated manual is available for use with adolescents (Short Term Assessment of Risk and Treatability: Adolescent Version; START-AV) (Nicholls et al. 2010) and the full manual is in preparation by Dr. Viljoen and colleagues. Pilot investigations and other studies have been conducted on the START-AV (see [Desmarais et al., 2012b](#); [Viljoen et al., 2012](#)).

- Those using the START tool are to consider any indicators that there are threats of harm that are real, enactable, acute and targeted. Assessors should be mindful of T.H.R.E.A.T in emergency situations where a comprehensive review of the evidence is not possible. ([O'Shea, Picchioni and Dickens, 2016](#)).
- Potential limitations of the START tool are it may be too general for certain patients or groups of patients (e.g. those with learning disabilities) ([Kroppan et al., 2017](#)).