

Single-item predictive validity of the Short-Term Assessment of Risk and Treatability (START) for violent behaviour in outpatient forensic psychiatry

Andreas Paetsch, Titus W. D. P. van Os, Nadine A. C. Troquete & Rob H. S. van den Brink

To cite this article: Andreas Paetsch, Titus W. D. P. van Os, Nadine A. C. Troquete & Rob H. S. van den Brink (2019) Single-item predictive validity of the Short-Term Assessment of Risk and Treatability (START) for violent behaviour in outpatient forensic psychiatry, *The Journal of Forensic Psychiatry & Psychology*, 30:4, 630-641, DOI: [10.1080/14789949.2019.1581831](https://doi.org/10.1080/14789949.2019.1581831)

To link to this article: <https://doi.org/10.1080/14789949.2019.1581831>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 26 Feb 2019.



[Submit your article to this journal](#)



Article views: 605



[View related articles](#)



[View Crossmark data](#)

Single-item predictive validity of the Short-Term Assessment of Risk and Treatability (START) for violent behaviour in outpatient forensic psychiatry

Andreas Paetsch^a, Titus W. D. P. van Os^b, Nadine A. C. Troquete^a
and Rob H. S. van den Brink^a

^aDepartment of Psychiatry, Rob Giel Research Center, University of Groningen, University Medical Center Groningen, The Netherlands; ^bDepartment of Forensic Psychiatry, Mental Health Organisation Friesland, Leeuwarden, The Netherlands

ABSTRACT

The single-item predictive validity of the Short-Term Assessment of Risk and Treatability (START) has not been thoroughly investigated, although this has great clinical relevance for the selection of treatment targets. Furthermore, it remains unclear whether the characteristic START additions of scoring strengths next to vulnerabilities and selecting key items, add incremental predictive validity. Finally, predictive validity has primarily been studied in inpatient settings and included mainly patients with a psychotic disorder. We analysed data from a mixed diagnostic sample of 195 forensic psychiatric outpatients with a 3-month and 170 patients with a 6-month follow-up period, using logistic regression analysis. The occurrence of violent or criminal behaviour was established based on the case manager's recordings in the patient's file. Only 5 of the 20 START items were found to have predictive validity: *Impulse Control*, *Attitudes*, *Material Resources*, *Rule Adherence* and *Conduct*. The last three were the only items for which incremental predictive validity was found with respect to scoring it as a strength and a vulnerability. Selection of key items did not add to the predictive validity. While possibly having therapeutic significance, the scoring of strength next to vulnerability and the selection of key items, may not be beneficial for risk assessment.


ARTICLE HISTORY Received 10 August 2018; Accepted 2 February 2019

KEYWORDS Short-Term Assessment of Risk and Treatability (START); predictive validity; outpatient forensic psychiatry; violence risk assessment; treatment goals

Background

Risk assessment instruments serve two purposes in forensic psychiatry: first, to guide placement decisions and liberties of patients based on an overall

CONTACT Andreas Paetsch  andreaspaetsch@gmx.net

The manuscript contains original work that has not been previously published and is not being submitted for publication elsewhere. All authors contributed significantly to the study and manuscript.  Supplemental data for this article can be accessed [here](#).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

level of risk, and second to indicate relevant treatment targets for risk reduction (Andrews & Bonta, 2010; Douglas & Kropp, 2002). The Short-Term Assessment of Risk and Treatability (START) is a structured professional judgement instrument developed for clinical assessment of short term risks in inpatient and outpatient forensic settings (Nicholls, Brink, Desmarais, Webster, & Martin, 2006; Webster, Nicholls, Martin, Desmarais, & Brink, 2006). It consists of 20 dynamic factors (i.e. areas of clinical, psychological or social functioning), which should be rated both in terms of risk (termed 'vulnerabilities') and protective factors (termed 'strengths') judged present in the current client functioning (Webster, Martin, Brink, Nicholls, & Desmarais, 2009). In addition, the rater is asked to mark the strengths and vulnerabilities considered of key importance for the client's risk (termed 'key strengths' and 'critical vulnerabilities' in the START). Based on these dynamic factor ratings for strength and vulnerability, selected key factors and the patient's risk history, case managers finally estimate overall short-term risks. These can be rated as low, moderate, or high for seven outcomes relevant in clinical forensic practice (violence towards others, self-harm, suicide, unauthorized leave, substance abuse, self-neglect, and being victimized; Webster et al., 2009). For forensic settings and the larger community, the crucial risk estimate concerns future violent behaviour, and the START aims to guide this assessment process.

Corresponding to the first purpose of risk assessment instruments distinguished above, the predictive validity of the START has almost exclusively been studied for its summary statistics and overall risk estimates (O'Shea & Dickens, 2014). However, an investigation of the predictive validity of the individual items of the START would have great additional clinical relevance for the second purpose of risk assessment instruments discerned. It would enable the identification of crucial factors for violence risk prediction in the studied patient group, and hence also potential treatment targets for the individual patient. Whether a factor is found to be predictive of a violent incident, may depend on the setting in which it is being studied, and thus far, the START has almost exclusively been studied in inpatient forensic settings. In these studies, the START was frequently scored retrospectively by researchers on case files instead of prospectively by the case manager, based on clinical experience with the client (O'Shea & Dickens, 2014). Furthermore, inpatient forensic settings differ greatly from outpatient forensic settings in terms of diagnoses and legal status of the patients, and comprehensiveness of the case manager's view on the client's life (Bouman, de Ruiter, & Schene, 2008, 2009, 2010; Nicholls, Petersen, Brink, & Webster, 2011; Troquete et al., 2013, 2015). Finally, an outpatient setting will provide different opportunities, challenges and temptations for patients than an inpatient setting, and hence the predictive value of risk and protective factors for violent behaviour may differ between settings.

The START has some unique features of scoring. It not only allows a case manager to rate a client's vulnerabilities but also his/her strengths. This is theorized to counteract a negative bias in assessment, which could result in risk overestimations and, ultimately, unwarranted restrictions and detentions (O'Shea, Picchioni, & Dickens, 2016). According to de Ruiter and Nicholls (2011), the consideration of protective factors might also hold advantages for clinical practice, such as improvement of the client-therapist relationship, identification of key areas to foster personal development, and better motivation for treatment in the patient. The possibility for an evaluator to indicate whether an item is especially important for a patient ('key-selection'), is another unique addition of the START to help identify key areas for intervention and risk assessment.

The current study aims to extend available predictive validity studies of the START for violent behaviour, by evaluating the validity of individual items in addition to summary statistics, by doing so in an outpatient instead of inpatient forensic setting, and by evaluating the incremental value of the unique scoring features of the START. This will be done by examining:

(1) the predictive validity of vulnerability and strength scores of individual START-items, (2) the incremental (i.e. additional) predictive validity of strength scores to vulnerability scores and vice versa for individual START-items, and (3) whether the predictive value of START-items for clients for whom the item was selected as a key item is higher than for those for whom it was not selected. To the authors' knowledge, only one study (i.e. O'Shea & Dickens, 2015) of the single-item predictive validity of the START has been conducted to this point.

Method

The present study builds on the same sample and data set as the study by Troquete et al. (2015), where more detailed information regarding data collection and design can be found. Data was collected as part of a randomized controlled trial on the effect of Risk Assessment and Care Evaluation (RACE-study; Troquete et al., 2013). Case managers in the intervention arm were instructed to assess their patients on the START for each treatment plan evaluation, which legally should occur at least once a year. Only the first assessment per patient was used for the current study. The outcome was established as the occurrence of any incident of violent or criminal behaviour by the client over a follow-up period up to 6 months. The study was conducted in an outpatient forensic psychiatric setting, where treatment is offered to individuals who already encountered or are at risk of encountering, the criminal justice system (Troquete et al., 2015). The study was approved by the Dutch Medical Ethical Committee for Mental Healthcare (protocol number NL16215.097.07).

Sample characteristics and follow-up period

All case managers and clients of the intervention arm of the RACE-study were included in the present study. Of 29 case managers (29% psychologists, 25% occupational therapists, 18% nurses, 17% specialists providing solely forensic psychiatric home care) 55% were female, with on average 7 years (SD = 6) of work experience in forensic settings. Scores on the START of 195 clients with a follow-up period of 3 months and 170 with a follow-up period of 6 months were available for the present analyses. These clients stem from an original sample of 310 clients (94% male) who presented with personality disorders (69%), substance-related disorders (38%), impulse control disorders (27%), mood disorders (21%), paraphilia (20%), and psychotic disorders (7%) (Troquete et al., 2015). Clients mostly committed violent offences (56%), property offences (37%), or sexual offences (32%). The majority of clients were treated either voluntarily (55%) or were under probation (28%) (Troquete et al., 2015).

The analyses of this study have been conducted for follow up periods of 3 and 6 months, but will only be reported for the 6 months follow up period and compared to the 3 months follow up period (reported as supplemental material at *URL*). Even though a follow-up period of 3 months is advised in the START manual (Webster et al., 2009), Troquete et al. (2015) showed that the number of incidents within 3 months was too low to provide sufficient power for a predictive validity study (in the present study 11% of the clients had an incident during 3 months follow-up versus 20% during 6 months follow-up).

Risk assessment

All case managers took part in the official training for the application of the Dutch version of the START. The 20 dynamic strength and vulnerability ratings were scored as 0 (*absent*), 1 (*possibly present*), or 2 (*present*). Key vulnerabilities and strengths could be marked by the case manager (key/not-key) and final risk estimates were given (1 = *low risk*; 2 = *medium risk*; 3 = *high risk*) for the 7 outcomes described above. From these risk estimates, only the predictive validity of the estimate for violence towards others will be reported here, because it corresponds most closely to the outcome measured. Strength and vulnerability sum scores were calculated by summing the strength and vulnerability scores on the 20 individual items.

In accordance with the START manual (Webster et al., 2009), case managers also scored all patients included in the RACE-study on the 10 historical factors of the Historical, Clinical and Risk management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997). This was done at baseline, independent of and several months prior to the first START assessment. The contribution of historical factors to violence prediction will not be subject of the current study.

Outcome assessment

The measured outcome was the occurrence of violent or criminal behaviour during the follow-up period. Violent behaviour was defined as intimidating or seriously threatening aggression or intentional behaviour that could potentially lead to physical harm of a person or animal. Criminal behaviour extends this to stalking, drug dealing, driving without a licence or under the influence of alcohol or drugs, possession of an illegal weapon or child pornography, theft, exhibitionism, and vandalism. All of these behaviours or outcomes were included, as they would be serious indicators for the need for further treatment (Troquete et al., 2015). Case managers recorded incidents that could potentially satisfy the above definitions on a standard form in the client's case file. Inclusion as an incident of violent or criminal behaviour was determined through consensus between three experts in outpatient forensic psychiatry, who were blind to the risk assessments. The outcome was coded as either the absence (= 0) or occurrence of one or more incidents (= 1) of violent or criminal behaviour during the follow-up period. For the sake of convenience, we will refer to incidents of violent or criminal behaviour as 'violent incidents'.

Analyses

All research questions are examined by logistic regression analysis. First, the predictive value of an individual vulnerability or strength score is assessed by the univariate Odds Ratio (OR) between that score and the occurrence of a violent incident during follow-up. To facilitate comparison with other studies, the area under the curve (AUC) values based on Receiver Operating Characteristic (ROC) analysis, will also be provided. For comparison with the individual items, the OR and AUC will also be reported for the strength and vulnerability sum scores and for the risk estimate for violence towards others. Second, the incremental predictive values of vulnerability and strength scores on a particular item are studied by entering both scores into the logistic regression model and controlling their ORs for each other. Finally, the difference in predictive value between patients for whom the item was selected as key item versus those for whom it was not, is examined by the interaction between the selection as key and the OR for their individual vulnerability or strength scores. All the above effects are tested for statistical significance, using an alpha of .05.

Results

All results for the START assessments of 170 clients with a follow-up period of 6 months can be found in [Table 1](#). Results for the vulnerability scores are shown on the left side of the table; results for the strength scores on the

right. On each side, the first three columns concern the first research question on the predictive validity of individual START-items, the next two columns the second question on incremental predictive value of strength and vulnerability scores to each other, and the last column the third question on the predictive value of selecting items as 'key' items. Below the findings for the three research questions are described successively.

With respect to the predictive validity of individual items, the vulnerability scores on the items *Impulse Control*, *Material Resources*, *Attitudes*, *Rule Adherence* and *Conduct*, and the START Vulnerability sum score were found to be valid univariate predictors for the occurrence of a violent incident during the 6 months following the START assessment. Concerning the strength scores, the items *Attitudes*, *Rule Adherence* and *Conduct*, and the Strength sum score show univariate predictive validity. In addition, the risk estimate for violence towards others was predictive for violent incidents (OR = 2.48; 95%CI: 1.30–4.74; $p < .01$; AUC = .63; $p = .02$; not shown in Table 1). All these relationships were in the expected direction of an increased risk in case of higher vulnerability or risk-estimate scores and a decreased risk for higher strength scores.

The incremental predictive power of strength to vulnerability scoring was only found for the items *Conduct* and *Rule Adherence* ($p = .05$) and of vulnerability to strength scoring for *Material Resources*. Only for these items, scoring strengths next to vulnerabilities or vice versa improved violence prediction over the next six month. Again, all of these findings were in the expected direction. No incremental predictive validity was found for the START sum scores.

Finally, only for the strength scoring on the item *Relationships* a marginally significant difference was found in the predictive value for clients for whom this item was selected as key factor versus clients for whom it was not. Furthermore, this difference was in the unexpected direction. Only for clients for whom strength in *Relationships* was not selected as a key factor, higher strength scores were associated with lower risk for violent incidents (OR = 0.44; 95%CI: 0.20–0.97; $p = .04$). However, for clients for whom this item was considered to be of key importance, higher strength scores tended to be related to higher risk for violent incidents, although this association did not reach statistical significance (OR = 2.90; 95%CI: 0.51–16.65; $p = .23$). Hence, the selection of items as 'key' for the individual does not seem to contribute to the predictive validity of the START.

When we compare the above findings for the 6-month follow-up period with those for the 3-month period (presented as supplemental material at URL), as expected, less significant relationships were found for the 3-month follow-up, due to a lack of power related to a small number of incidents. However, the ORs for the significant relationships reported above differed no more than 25% from the corresponding ORs for 3-months, with the exception of the vulnerability score for *Impulse Control* (OR = 1.11 for



Table 1. Predictive validity of individual START items over a 6 months follow-up period (N = 170).

Vulnerability				Strength			
Incremental predictive value to Strength score				Incremental predictive value to Vulnerability score			
OR	95%CI	AUC	P	OR	95%CI	AUC	P
Interaction with Key selection [#]				Interaction with Key selection [#]			
Item				Item			
1.45	0.74; 2.83	.56	.65	0.64	0.34; 1.20	.55	.20
1.51	0.81; 2.79	.54	.80	0.65	0.35; 1.21	.55	.05
1.05	0.61; 1.81	.51	.55	0.91	0.56; 1.48	.51	1.00
1.37	0.83; 2.26	.55	1.00	0.65	0.39; 1.08	.58	.63
1.13	0.58; 2.21	.48	.63	0.89	0.46; 1.72	.52	.86
0.83	0.46; 1.52	.48	1.00 [§]	1.07	0.60; 1.91	.51	1.00 [§]
1.18	0.60; 2.33	.53	.79	0.74	0.35; 1.58	.55	1.00 [§]
1.41	0.87; 2.28	.59	.66	0.73	0.45; 1.17	.56	.79
2.14*	1.14; 3.99	.57	.37	0.58	0.30; 1.12	.57	.98
1.51	0.90; 2.52	.57	1.00	0.81	0.47; 1.38	.55	.22
1.49	0.86; 2.59	.52	.54	0.75	0.43; 1.32	.51	.17
2.11**	1.23; 3.59	.63*	.99	0.62	0.36; 1.09	.61	.47
2.21*	1.15; 4.22	.66*	.11	0.45*	0.24; 0.85	.67*	.82
1.69	0.88; 3.28	.56	.72	0.70	0.41; 1.22	.57	1.00
2.01*	1.07; 3.78	.63*	1.00	0.42**	0.23; 0.78	.66*	1.00
2.50*	1.25; 5.01	.64*	1.00	0.35**	0.18; 0.70	.69**	1.00
0.77	0.42; 1.41	.45	.16	1.16	0.64; 2.13	.54	.67
1.10	0.62; 1.94	.50	.65	0.71	0.39; 1.28	.57	1.00
1.90	0.97; 3.69	.61	.35	0.65	0.33; 1.27	.60	1.00
1.36	0.76; 2.45	.54	1.00 [§]	0.78	0.44; 1.37	.58	1.00
1.58*	1.08; 2.30	.60	NA	0.65*	0.44; 0.95	.63*	NA

#Excluding cases in which no Key Vulnerabilities were selected (remaining N = 162)

¶ Excluding cases in which no Key Strengths were selected (remaining N = 159)

*No incident or incident perfectly identified by Vulnerability or Strength score in group for which this item was selected as a key item

* p < .05; ** p < .01; [§] p = .05

3-months vs. 2.14 for 6-months) and the risk estimate for violence towards others (OR = 1.48 for 3-months vs. OR = 2.48 for 6-months)

Discussion

Only 5 of the 20 START items – either scored on vulnerability or strength or both – were found to have predictive validity for the occurrence of violent incidents in the following 6 months, as did the START sum scores and final risk estimate for violence towards others. The predictive items were: *Impulse Control*, *Attitudes*, *Material Resources*, *Rule Adherence* and *Conduct*. The last three were also the only items for which incremental predictive validity was found for the unique START feature of scoring both strength and vulnerability on an item. No predictive validity was found for the other unique START feature of indicating items considered to be of particular importance for the violence risk of a client (key-selection).

The only other study which examined the predictive validity of individual items of the START (O'Shea & Dickens, 2015), only showed partially overlapping results to our study. This study examined gender differences in the prediction of aggressive incidents in an inpatient setting and found strength or vulnerability scores on 9 items of the START to be predictive of any aggression in either men or women (O'Shea & Dickens, 2015). For men, this included the items *Impulse Control*, *Rule Adherence* and *Conduct*, which we also found to be predictive in our predominantly male sample. Our study, however, does not corroborate their surprising conclusion that strength scores would be more potent predictors of aggression than vulnerability scores, and neither does a recent meta-analysis of the predictive value of summary protective scores (O'Shea & Dickens, 2016).

Our finding that only five START items were predictive of violent incidents, does not necessarily signify that the other 15 items are irrelevant. For instance, risk assessment by the case manager who subsequently also provided the treatment deemed necessary may have reduced the predictive value of the START, because successful treatment may have prevented incidents to occur. Taking this line of reasoning one step further, it may well be that the items found to be significant are not only crucial for prediction but are also the most refractory and difficult to change. *Impulse Control* refers to acting before thinking and ignoring the consequences of one's actions. *Attitudes* refer to lying, difficulty understanding why others feel sad and feeling that what you want is more important than what other people want. *Rule Adherence* denotes finding it hard to see the point of rules or to obey them, and resisting checks, such as drug tests. *Conduct* implies behaving unpleasantly towards others, bullying or frightening others, insulting others, stealing or destroying property, and sexually harassing others. Taken together, these factors refer to serious deficits in the development of

conscience and self-regulation, which have been shown to have long-term predictive significance for violent behaviour and in particular for encounters with the police (Sijtsema, Kretschmer, & van Os, 2015). These factors highly overlap with the antisocial personality traits and antisocial cognitions of the 'Big Four' criminogenic factors identified by Andrews and Bonta (2010). Treatments for poor impulse control and aggression do exist and primarily comprise medication and cognitive behavioural therapy, but the underlying deficits in the development of conscience and self-regulation prove rather refractory (see, for example, Hornsveld, Kraaimaat, Muris, Zwets, & Kanters, 2015). We, therefore, hypothesize that the START factors found to be predictive for violent behaviour in the current study and the 'Big Four' criminogenic factors may well stand out because they are in fact the most treatment-resistant deficits in patient functioning.

Alternatively, only a small number of the START items may have been found to be predictive in the current study, because the START is not tailored to the needs of the specific setting of this study. The START was developed from – and has almost exclusively been tested in – studies in inpatient forensic settings (Nicholls et al., 2006; O'Shea & Dickens, 2014; Webster et al., 2006). These studies differ from the current study in some important respects. First and foremost, the inpatient study samples predominantly consisted of patients with a psychotic disorder (O'Shea & Dickens, 2014), while the majority of patients in the current study had personality disorders and only 7% a psychotic disorder. Second, in the majority of previous studies, the risk assessment was conducted by researchers on client files (O'Shea & Dickens, 2014) instead of the clinician based on direct experience with the client, which is the intended use of the START (Webster et al., 2009). Third, outpatient and inpatient settings differ in the frequency of treatment contacts and the comprehensiveness of the case manager's view of the client life, and hence in information available for risk assessment (Troquete et al., 2015). Finally, our outcome measure covered a wider range of clinically relevant outcomes than most other studies, including incidents of criminal as well as violent behaviour. These differences between the current and previous studies may restrict the generalizability of findings between studies.

For clinicians, it is important to know which factors are predictive for violent incidents in the patient group they are working with and how a particular patient is doing on these factors. This information is essential for appropriate care planning and selection of treatment targets tailored to the needs of the patient, which was distinguished above as the second purpose of risk assessment instruments. The current study provides the required information on a group level for patients treated in an outpatient forensic setting. A worrisome issue, however, is that this study also suggests that clinicians do not succeed in selecting the vulnerability and strength

factors of particular importance for a patient. No effect was found for this particular feature of the START, i.e., selection of 'key' items. Whether this is specific for clinicians working in an outpatient setting remains to be tested. Inviting clients to assess their own risk and protective factors for violence and to select their own 'key' factors may be helpful here, both to involve them in shared care planning and to improve the predictive value of violence risk assessment (van den Brink et al., 2015).

Limitations and strengths

The statistical power of this study may be compromised due to the small number of clients with a violent incident during the follow-up periods: 34 [20%] over 6 months and 22 [11%] over 3 months. Furthermore, the power to establish interactions between selection as a key item and predictive value of that item was especially affected, due to the often limited number of key selections per item (mean for vulnerabilities: 25.9 [16%]; mean for strengths: 24.3 [15%]). A general limitation for risk assessment in an outpatient setting compared to an inpatient setting is that case managers only have a partial and time-limited view on the patient's functioning (Troquete et al., 2015), as noted above. Finally, the outcome measure (violent incidents) was assessed from case notes made by the same person (i.e. the case manager) who also assessed the client's risk.

The present study is the first systematic investigation of the predictive validity of individual items of the START. The predictive validity was studied in clinical practice, with real-time scoring of patient functioning by the client's case manager and prospective follow-up for incidents, instead of the usual retrospective scoring of client files by researchers. Moreover, next, to the relatively large sample of clients studied (N = 170), it is one of few studies on the predictive validity of the START, which has been conducted in outpatient forensic psychiatry.

Conclusions

The five START items identified in this study appear to be of particular importance for the violence risk of forensic psychiatric outpatients. These items are: *Impulse Control*, *Material Resources*, *Attitudes*, *Rule Adherence* and *Conduct*. The scoring of strength next to vulnerability, and particularly the selection of key items, may not be useful to improve assessment of violence risk at group level in this setting. Nonetheless, these features may have therapeutic significance, such as drawing attention to positive aspects of client functioning and fostering the therapeutic alliance and patient motivation for treatment, which are without a doubt worthwhile considerations in the clinical application of risk assessment. However, these therapeutic claims require testing (O'Shea & Dickens, 2016; Troquete et al., 2013; van den Brink et al., 2015).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the ZonMw, The Netherlands Organisation for Health Research and Development under Grant [number 100 003 023]. Further funding was provided by the participating mental health organisations Lentis, Drenthe, and Friesland, and the University Center of Psychiatry of the University Medical Center Groningen.

References

- Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH: Anderson Publishing Co.
- Bouman, Y. H. A., de Ruiter, C., & Schene, A. H. (2008). Quality of life of violent and sexual offenders in community-based forensic psychiatric treatment. *Journal of Forensic Psychiatry & Psychology, 19*(4), 484–501.
- Bouman, Y. H. A., de Ruiter, C., & Schene, A. H. (2009). Recent life events and subjective well-being of personality disordered forensic outpatients. *International Journal of Law and Psychiatry, 32*(6), 348–354.
- Bouman, Y. H. A., de Ruiter, C., & Schene, A. H. (2010). Social ties and short-term self-reported delinquent behaviour of personality disordered forensic outpatients. *Legal and Criminological Psychology, 15*(2), 357–372.
- de Ruiter, C., & Nicholls, T. L. (2011). Protective factors in forensic mental health: A new frontier. *The International Journal of Forensic Mental Health, 10*(3), 160–170.
- Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior, 29*(5), 617–658.
- Hornsveld, R. H., Kraaimaat, F. W., Muris, P., Zwets, A. J., & Kanters, T. (2015). Aggression replacement training for violent young men in a forensic psychiatric outpatient clinic. *Journal of Interpersonal Violence, 30*(18), 3174–3191.
- Nicholls, T. L., Brink, J., Desmarais, S. L., Webster, C. D., & Martin, M. (2006). The Short-Term Assessment of Risk and Treatability (START): A prospective validation study in a forensic psychiatric sample. *Assessment, 13*(3), 313–327.
- Nicholls, T. L., Petersen, K. L., Brink, J., & Webster, C. (2011). A clinical and risk profile of forensic psychiatric patients: Treatment team STARTs in a Canadian service. *The International Journal of Forensic Mental Health, 10*(3), 187–199.
- O'Shea, L. E., & Dickens, G. L. (2014). Short-Term Assessment of Risk and Treatability (START): Systematic review and meta-analysis. *Psychological Assessment, 26*(3), 990–1002.
- O'Shea, L. E., & Dickens, G. L. (2015). Predictive validity of the Short-Term Assessment of Risk and Treatability (START) for aggression and self-harm in a secure mental health service: Gender differences. *International Journal of Forensic Mental Health, 14*, 132–146.
- O'Shea, L. E., & Dickens, G. L. (2016). Performance of protective factors assessment in risk prediction for adults: Systematic review and meta-analysis. *Clinical Psychology: Science and Practice, 2*(23), 126–138.

- O'Shea, L. E., Picchioni, M. M., & Dickens, G. L. (2016). The predictive validity of the Short-Term Assessment of Risk and Treatability (START) for multiple adverse outcomes in a secure psychiatric inpatient setting. *Assessment, 23*(2), 150–162.
- Sijtsema, J. J., Kretschmer, T., & van Os, T. (2015). The structured assessment of violence risk in youth in a large community sample of young adult males and females: The TRAILS study. *Psychological Assessment, 27*(2), 669–677.
- Troquete, N. A. C., van den Brink, R., Beintema, H., Mulder, T., van OS, T. W., Schoevers, R. A., & Wiersma, D. (2015). Predictive validity of the short-term assessment of risk and treatability for violent behavior in outpatient forensic psychiatric patients. *Psychological Assessment, 27*(2), 377–391.
- Troquete, N. A. C., van Den Brink, R., Beintema, H., Mulder, T., van Os, T. W. D. P., Schoevers, R. A., & Wiersma, D. (2013). Risk assessment and shared care planning in out-patient forensic psychiatry: Cluster randomised controlled trial. *The British Journal of Psychiatry, 202*(5), 365–371.
- van den Brink, R. H. S., Troquete, N. A. C., Beintema, H., Mulder, T., van Os, T. W. D. P., Schoevers, R. A., & Wiersma, D. (2015). Risk assessment by client and case manager for shared decision making in out-patient forensic psychiatry. *BMC Psychiatry, 15*, 120.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. (1997). *HCR-20. Assessing the risk of violence* (version 2). Burnaby: Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.
- Webster, C. D., Martin, M., Brink, J., Nicholls, T. L., & Desmarais, S. L. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START) (version 1.1)*. Coquitlam, BC: British Columbia Mental Health & Addiction Services.
- Webster, C. D., Nicholls, T. L., Martin, M., Desmarais, S. L., & Brink, J. (2006). Short-Term Assessment of Risk and Treatability (START): The case for a new structured professional judgment scheme. *Behavioral Sciences & the Law, 24*(6), 747–766.