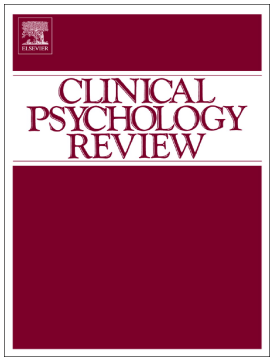

Measurement properties of tools used to assess suicidality in autistic and general population adults: A systematic review

S.A. Cassidy, L. Bradley, E. Bowen, S. Wigham, J. Rodgers

PII: S0272-7358(18)30069-2
DOI: doi:10.1016/j.cpr.2018.05.002
Reference: CPR 1687

To appear in: *Clinical Psychology Review*

Received date: 20 February 2018
Revised date: 27 April 2018
Accepted date: 3 May 2018

Please cite this article as: S.A. Cassidy, L. Bradley, E. Bowen, S. Wigham, J. Rodgers , Measurement properties of tools used to assess suicidality in autistic and general population adults: A systematic review. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Cpr(2017), doi:10.1016/j.cpr.2018.05.002

Measurement properties of tools used to assess suicidality in autistic and general population adults: A Systematic Review.

*Cassidy, S. A.,[1,2,3] Bradley, L.,[1] Bowen, E.,[4] Wigham, S.,[3] and Rodgers, J.[3]

[1]School of Psychology, University of Nottingham, UK.

[2]Centre for Innovative Research across the Life Course, Coventry University, UK.

[3]Institute of Neuroscience, Newcastle University, UK.

[4]Centre for Violence Prevention, Worcester University, UK.

*Corresponding Author: Dr Sarah Cassidy, School of Psychology, University of Nottingham, University Park, Nottingham, NG7 2RD. Telephone: +44 (0) 115 95 13470. Email: Sarah.Cassidy@Nottingham.ac.uk

Abstract

Adults diagnosed with autism are at significantly increased risk of suicidal thoughts, suicidal behaviours and dying by suicide. However, it is unclear whether any validated tools are currently available to effectively assess suicidality in autistic adults in research and clinical practice. This is crucial for understanding and preventing premature death by suicide in this vulnerable group. This two stage systematic review therefore aimed to identify tools used to assess suicidality in autistic and general population adults, evaluate these tools for their appropriateness and measurement properties, and make recommendations for appropriate selection of suicidality assessment tools in research and clinical practice. Three databases were searched (PsycInfo, Medline and Web of Knowledge). Four frequently used suicidality assessment tools were identified, and subsequently rated for quality of the evidence in support of their measurement properties using the COSMIN checklist. Despite studies having explored suicidality in autistic adults, none had utilised a validated tool. Overall, there was lack of evidence in support of suicidality risk assessments successfully predicting future suicide attempts. We recommend adaptations to current suicidality assessment tools and priorities for future research, in order to better conceptualise suicidality and its measurement in autism.

Keywords: Autism Spectrum Condition, Autistic, Suicidality, Suicide, Self-harm, Measurement properties, COSMIN, Systematic review.

Introduction

Adults diagnosed with Autism Spectrum Conditions (ASC, hereafter autistic adults) are at high risk of experiencing suicidality compared to other clinical groups (Hedley and Uljarević, 2018; Zahid and Upthegrove, 2017; Cassidy et al, 2014; Segers and Rawana, 2014; Hannon and Taylor, 2013). Up to 66% of newly diagnosed adults with Asperger Syndrome (ASC without language delay or intellectual disability) reported having contemplated suicide, significantly higher than the UK general population (17%); and 35% reported that they had planned or attempted suicide (Cassidy et al. 2014). In a recent large-scale population study in Sweden, those diagnosed with ASC, without co-occurring ID, were at high risk of dying by suicide compared to the general population (Hirvikoski et al. 2016). However, there are very few studies exploring suicidality in ASC, with no known measures or models yet validated for this group (Cassidy and Rodgers, 2017). Clearly, it is crucial to effectively assess suicidality in autistic adults. However, it is unclear if there are valid tools available to assess suicidality in autistic adults, or whether existing tools need to be adapted for this group.

ASC is characterised by difficulties in socialisation, imagination, communication, narrow obsessive interests, and sensory difficulties (APA, 2013). A number of characteristics of ASC may present challenges for clinicians in accurate identification of suicidality in this group. Self-injurious behaviour is commonly associated with ASC, particularly in the context of challenging behaviour, repetitive behaviours and co-occurring intellectual disability (ID) (see Matson et al. 2007 for a review). However, there is only one study on non-suicidal self-injury (NSSI) in autistic adults (Maddox et al. 2017). No research has yet explored whether autistic people who present with self-injurious behaviours may be experiencing suicidality or NSSI (Hannon and Taylor, 2013). This could therefore increase risk of such behaviours being inaccurately attributed to autism (termed diagnostic overshadowing), with suicidality not necessarily being considered.

Other characteristics of ASC may also affect the utility of current tools, which were designed for non-autistic populations, to accurately detect suicidality in this group. For example, many autistic people tend to interpret information literally (Happe et al. 1995), and experience difficulties in ability to recall what has happened to them in the past, and imagine what may happen to them in the future (Crane, Lind and Bowler, 2013; Lind and Bowler, 2010). These difficulties could affect the ability to answer questions about lifetime suicidality, or future suicidal intent (e.g. "How likely are you to attempt suicide someday?" Suicidal Behaviour Questionnaire – Revised, Osman et al. 2001). Many autistic people also experience difficulty articulating their own internal emotional experiences (termed alexythymia, Bird et al. 2010), which may present difficulties when self-reporting on internal emotional distress in mental health and suicidality assessments (Cassidy et al. 2018; Cassidy et al. 2014).

There may also be unique aspects of suicidality in ASC which may not be captured in traditional tools designed for other populations. For example, many suicide risk assessments enquire about communication of suicide intent to others, which is taken to indicate increased suicide risk in the general population. However, difficulties in communication, and increased chance of being isolated both in terms of access to health services (Raja, 2014) and social connections (Orsmond, Shattuck, Cooper, Sterzing, & Anderson, 2013) among autistic people, may lead to lower endorsement of such items while not necessarily indicating reduced suicide risk. Social isolation and loneliness in autistic people and people with high autistic traits increase risk of suicidality (Pelton and Cassidy, 2017; Hedley et al. 2018), as does lack of tangible social support (Hedley et al. 2017). Therefore it is important for clinicians as part of suicidality assessments to probe for social isolation, loneliness and support needs in autistic people. Checking understanding of questions in suicidality assessments, why the person may not have told others about their suicidality (e.g. I had no

one to tell, I did not consider it important etc.), could also reveal important information regarding risk level.

Given that the presentation of suicidality and cognitive characteristics of ASC may impede effective suicide risk assessment using traditional tools, it is crucial to identify what suicide risk assessments have been utilised in this group, and if none are available, to identify the most robust candidate tools in the general population to adapt. There is a growing body of systematic reviews showing a paucity of research exploring the measurement properties of outcome measures in ASC, which have made important recommendations to improve research and clinical assessment (Cassidy et al. 2018; Hanratty et al. 2015; Wigham and McConachie, 2014; McConachie et al. 2015). These reviews have used a validated research tool developed to assess the methodological quality of studies assessing the measurement properties of health outcome assessment tools: the consensus based standards for the selection of health measurement instruments (COSMIN) (Mokkink et al. 2016; Mokkink et al. 2012; Mokkink et al. 2010). The COSMIN method involves two stages. First, tools used to assess a health outcome in a well-defined population are identified from a systematic search of the literature. Subsequently, the tools used frequently (at least twice), with evidence of validity (i.e. with reference to a previously published study), are searched for using a comprehensive search tool validated for this purpose (Terwee et al. 2009). The quality of the available evidence is subsequently rated using the COSMIN checklist (Mokkink et al. 2016).

It is important to note that tools are not either valid or invalid, but are rather valid for certain purposes or circumstances (Kamphaus & Frick, 2005). The COSMIN checklist allows a systematic assessment of the quality of evidence for and against a range of measurement properties, pooled across studies, thus providing a picture of the strengths and weaknesses of the most frequently used tools in different contexts. This allows us to make evidence

based recommendations on which tools to select for particular clinical and/or research contexts. We therefore utilise this robust method to identify suicidality assessment tools used in autistic and general population adults, with similar age and intellectual ability, in order to draw conclusions about the relative quality of the evidence in each group regarding the measurement properties of these tools. Given that autistic adults have difficulty accessing psychiatric services due to lack of expertise and service provision for mental health in autism (Crane et al. 2018; Raja, 2014), suicidality assessment tools used in screening the general population in research and clinical practice will be particularly useful to adapt for autistic adults. The current study thus focused on identifying suicidality screening tools used in general population screening studies, as opposed to tools primarily used in psychiatric groups. From this synthesis of the available evidence, we subsequently make recommendations for future research and clinical practice aiming to effectively assess suicidality in autistic and non-autistic adults. Given the higher risk of death by suicide in autistic adults, without ID (Hirvikoski et al. 2016), we focused the search on adults without ID.

Review Methods: Stage 1

The protocol for this review is registered within the International Register of Systematic Reviews (Registration number: CRD42016035217), and can be accessed online (http://www.crd.york.ac.uk/PROSPERO/prospero.asp). This systematic review follows the guidelines for Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards (Moher et al. 2015).

Search Strategy

The following electronic bibliographic databases were searched: Medline, Psychinfo and Web of Knowledge. The Cochrane library was also searched to confirm that no other

systematic reviews of the current study topic existed. There were two searches carried out in stage one for suicidality measures used in; a) autistic adults, without co-occurring ID; and b) general population adults, without any co-occurring conditions or ID. The terms for each search strategy are included in table 1. The searches were restricted to peer reviewed articles published in the English language, between 1992 and 22nd January 2018 – when the last searches were run. The current study focused on literature pertaining to ASC without co-occurring ID, which is frequently referred to as Asperger Syndrome (AS). AS was first included as a separate diagnosis in the WHO International Classification of Diseases in 1992, so we focused on studies published after this date, when we expected reference to AS to be more consistent in the literature.

Selection Criteria

We utilised a standardised approach to the selection of studies as in previous COSMIN reviews (e.g. Cassidy et al. 2018). We focused on tools that include more specific (i.e. specifically suicidality as opposed to self-harm or non-suicidal self-injury), and broader (including in depth assessment of suicidality to help gauge risk level) conceptualisations of suicidality than is feasible in single items or subscales. These typically fail to distinguish broader conceptualisations of self-harm from suicidal intent, and lack information on important risk indicators, such as current and lifetime experience, frequency, intensity, intent and access to means. Therefore studies had to focus on a tool specifically assessing suicidality, including assessment of suicidal intent (as opposed to self-harm more generally), clinically defined as in the International Classification of Diseases (ICD-10), and Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Studies which utilised tools with a single suicide related question, item or subscale contained within a larger measure (e.g. Mini-International Neuropsychiatric Interview (MINI) (Hergueta et al. 1998), Structured Clinical Interview for DSM-IV (SCID-II) (First et al. 1997)), and/or without evidence of validity (i.e. by

reference to a previously published study), were excluded. This is necessary to maximise the probability of identified tools having evidence regarding their measurement properties in search two.

We searched for studies utilising tools to assess both prevalence of suicidality (epidemiological/population studies), and assess outcomes (treatment/intervention and longitudinal/cohort studies). To be included studies had to focus on adults aged 18 years and over, without ID. Where the age range was partly outside this, studies were included if 50% or more of the total population studied was over 18 years, and the mean age of the sample was 18 years or above. This ensured that the tools were likely to be appropriate for adults. We excluded articles using tools which had been adapted specifically for another population than ASC or the general population (e.g. for older adults, a particular gender, or a specific culture). This was to ensure that the tool would likely be useful for assessing suicidality in general population adults, as opposed to a specific sub-group of the general population. We included studies using the most up to date version of the tool available, as this is most likely to be used in future research and clinical practice.

General population adult search criteria

Studies were included if data from general population adults, without ID or co-occurring conditions, were presented separately, and comprised at least 50% of the total sample. Any studies including an autistic comparison group were excluded and considered for inclusion in the ASC search.

Autistic adults search criteria

Studies were included if data from autistic adults were presented separately, and if 50% or more of the participants had a diagnosis of ASC.

One reviewer (SC) screened the titles and abstracts of articles for inclusion, and where there was any doubt on whether an article should be carried over to the full text sift, it was included. SC then conducted the full text sift of articles, with any ambiguous papers discussed with LB, EB and JR to reach consensus. All references of included articles were also searched for additional articles to include.

Data Extraction

Data extraction was performed by SC, and 20% of articles independently checked by LB. A data extraction form was adapted from a previously developed form used in similar research (Cassidy et al. 2018; Wigham and McConachie, 2014). Data pertaining to: participant characteristics, tools used, domains captured and study type, were recorded.

Results: Stage 1

ASC

The search for studies using tools to assess suicidality in autistic adults, identified 672 articles which were screened, none of which were retained for analysis (Figure 1). A majority of the studies initially screened and excluded in the ASC search had explored self-injury and challenging behaviour in autistic adults, often with co-occurring ID, as opposed to suicidality – i.e. including intent to end one's own life. Crucially although a limited number of studies had explored suicidality in autistic adults, none had used a validated tool designed to assess suicidality specifically. A majority of studies in both groups searches had utilised a single item designed for the specific study with no evidence of validity, or a single item or subscale contained within a larger mental health (MINI, SCID) or depression (e.g. PHQ-9, BDI) measure. As stated above, the current study focused specific and broader conceptualisations of suicidality than is possible in single items or subscales. Additionally, it is vital that there is evidence of validity of tools (e.g. by reference to a previous study) in the first stage, in order

to identify tools which are likely to meet COSMIN inclusion criteria in the second stage. Hence, no studies of suicidality in ASC were identified which have used a suicidality assessment tool with evidence of validity to consider further in stage two.

General Population

The search for studies using tools to assess suicidality in general population adults identified 1,774 articles which were screened, with 25 retained for analysis (Figure 1). Fourteen different tools were used to assess suicidality in the studies (Appendix A). Self-report questionnaires included: Columbia Suicide Severity Rating Scale (C-SSRS) (Posner et al. 2011), Measure of Episodic Planning of Suicide (MEPOS) (Anestis et al. 2014), Suicide Behaviours Questionnaire Revised (SBQ-R) (Osman et al. 2001), Beck Scale for Suicidal Ideation (BSS) (Beck, Steer, and Ranieri, 1988), Beck Suicide Intent Inventory (BSI) (Beck, Schuyler and Herman, 1974), Depression Severity Index – Suicide Subscale (DSI-SS) (Metalsky and Joiner, 1997), Modified Scale for Suicidal Ideation (MSSI) (Miller et al. 1986), Paykel (Paykel et al. 1974), Sheehan Suicidality Tracking Scale (S-STS) (Coric et al. 2009), Suicide Assessment Scale (SUAS-S) (Stanley, Träskman-Bendz, and Stanley, 1986), Suicide Ideation Scale (SIS) (Rudd, 1989), Suicide Score Scale (SSS) (Innamorati et al. 2008), and the Plutchik Suicide Risk Scale (PSRS) (Aradilla-Herrero et al. 2014). The searches also identified clinician interview versions of the Columbia Suicide Severity Rating Scale (C-SSRS) (Posner et al. 2011), Beck Suicide Intent Inventory (BSI) (Beck, Schuyler, and Herman, 1974), Paykel (Paykel et al. 1974), Sheehan Suicidality Tracking Scale (S-STS) (Coric et al. 2009), and the Self-Injurious Thoughts and Behaviours Interview (SITBI) (Nock et al. 2007). Eight of these tools had each only been used in one study in the general population, without co-morbid conditions (MEPOS; BSI; DSI-SS; MSSI; S-STS; SUAS-S; SIS; and SSS). Therefore these tools were not considered further, as we were interested in tools which had been used frequently (at least twice) in the general population with some evidence of validity, to maximise the

chances of there being evidence available to evaluate using the COSMIN checklist. Hence, four tools (C-SSRS; SBQ-R; BSS; and Paykel) were considered further in stage 2.

Review Methods: Stage 2

The second stage of the review searched for evidence of the measurement properties of the tools identified in stage 1. In order to do this, a comprehensive search was carried out using a methodological filter in PubMed, designed to search for studies assessing the measurement properties of health outcome assessment tools (Terwee et al. 2009). We focused on studies which had explored the measurement properties of the tools in adults (18 years and over), without co-occurring ID. Unlike in stage 1, Adult samples with co-occurring conditions were included, as studies exploring the validity of suicidality assessment tools in the general population may nevertheless be validated in psychiatric samples. Including studies of clinical samples thus provides useful information regarding the contexts the tools may be most useful in research and/or clinical practice.

Data extraction method

Once articles were identified from the search, the methodological quality of each article was assessed using the COSMIN checklist (Consensus based Standards for the selection of health based measurement Instruments) (Monkkink et al. 2016). COSMIN rates the evidence in support of 9 measurement properties on a 4-point scale (from excellent to poor): internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, criterion validity, responsiveness to change, and cross cultural validity. COSMIN implements a 'worst score counts' method, by which an overall rating is assigned to each measurement property based on the lowest score provided. For example, if a study is rated excellent on all criteria related to internal consistency (e.g. Cohen's Kappa was calculated, an

adequate sample size was utilised etc.), but the study failed to check the uni-dimensionality of the scale, then this study would still be rated as 'poor' overall (Monkkink et al. 2010).

The checklists were completed by SC, with 9 (34.6%) of the articles independently rated by SW, both of whom were trained and experienced in using COSMIN. Inter-rater reliability between SC and SW was 73%, similar to previous studies (e.g. Cassidy et al. 2018; Wigham and McConachie, 2014). Disagreements were resolved with discussion and these agreed COSMIN ratings were utilised in the subsequent evidence synthesis.

Evidence Synthesis

The quality of the evidence in support of each measurement property needs to be considered in the context of the studies' findings, in order to gauge the amount of evidence available for or against each measurement property. First, the quantitative findings from each study are given a rating of positive (in support of the property), indeterminate (not possible to deduce whether the evidence is for or against the property), or negative (evidence against the property). For example, criterion validity is considered positive when the study supplies convincing evidence that the criterion used is indeed a gold standard, and the correlation between the outcome measure and the gold standard criterion is greater than 0.7 (De Vet et al. 2011). Subsequently, the quality of the evidence is considered in the context of the studies quantitative findings. *Strong* evidence (+++/---) is defined as one methodologically excellent or several good studies which find consistent evidence for or against a measurement property; *moderate* evidence (++/--) is defined as one methodologically good or several fair studies which find consistent evidence for or against a measurement property; *limited* evidence (+/-) is defined as one methodologically fair study finding evidence for or against a measurement property; *conflicting* evidence (+/-) is where the evidence for or against a measurement property is not consistent between studies; and

indeterminate evidence (?), is where there are only studies of poor methodological quality available for a measurement property (Monkkink et al. 2012).

Results: Stage 2

The PubMed search for studies assessing the measurement properties of suicidality tools used in general population adults identified 218 articles which were screened, 26 of which were retained for analysis (figure 2) (see Appendix B for characteristics of the study populations included in the analysis).

No articles assessing the measurement properties of the Paykel were identified from the search. The methodological quality of the included studies are presented in Table 2 and the collated evidence pertaining to the measurement properties for each tool are presented in Table 3. Many of the articles reported data on differences in scores and normative data, which are important for interpretability (De Vet et al. 2011). However, no studies reported minimal important change or floor or ceiling effects.

Suicide Behaviour Questionnaire – Revised (SBQ-R)

Despite evidence of being widely used in general population studies of suicidality, only two studies were found assessing the measurement properties of the SBQ-R in adults. The quality of the evidence in support of hypothesis testing was weak, with one fair study showing significant differences between psychiatric and non-clinical populations in line with hypotheses with large effect (Osman et al. 2001). The quality of the evidence in support of criterion validity was moderate: sensitivity (>.882) and specificity (>.875) were acceptable for successfully differentiating suicidal from non-suicidal individuals, using both the first item of the SBQ-R (Aloba et al. 2017) and total scores (Osman et al. 2001; Aloba et al. 2017). The quality of the evidence for internal consistency was strong, with one excellent study showing acceptable Cronbachs alpha (.8) for the whole scale, confirmed as unidimensional via factor

analysis (Aloba et al. 2017). Evidence for structural validity was also strong, with one excellent study showing support for a one-factor solution (Aloba et al. 2017).

Beck Scale for Suicide Ideation (BSS)

The measurement properties of the BSS have primarily been assessed in psychiatric patient samples, despite being used in many general population studies. The evidence in support of hypothesis testing was mixed. One study showed weak evidence against the BSS predicting future adverse events (e.g. future suicide attempts) (de Beurs, Fokkema and O'Connor, 2016), and one poor study (Cochrane-Brink et al. 2000, due to the small sample size) showed evidence in support of the BSS predicting future adverse events. However, there was moderate evidence for the BSS significantly correlating with other relevant measures and demographics (Esfahani, Hashemi and Alavi 2015; Horon et al. 2012; Kliem et al. 2017; Cochrane-Brink et al. 2000), moderate evidence for factors remaining consistent over time (de Beurs et al. 2015), and strong evidence for the BSS distinguishing subgroups (e.g. suicide attempters vs. non-attempters) (Horon et al. 2012; Healy et al. 2006; Pinninti et al. 2002).

The evidence in support of criterion validity for the BSS was similarly mixed. One excellent study (de Beurs et al. 2016) showed low specificity (.2) but high sensitivity (.95) for the BSS predicting future suicidal behaviour, a good study (Chang and Tan, 2015) showed a poor AUC (<.44) for predicting future adverse events and a poor study (due to small sample size) showed excellent sensitivity (100%) and specificity (90%) for predicting future hospitalisations (Cochran-Brink et al. 2000). However, there was consistent strong evidence for high sensitivity and specificity when distinguishing clinical groups (Horon et al. 2012; Pinninti et al. 2002; Cochrane-Brink et al. 2002) (e.g. hospitalised vs. non-hospitalised; multiple, single or no previous suicide attempts).

The evidence in support of internal consistency for the BSS was strong, with one excellent study showing high internal consistency in a translated version of the BSS (Esfahani et al. 2015). Evidence in support of reliability of the BSS was indeterminate, with two studies of poor methodological quality - observations were not independent (Healy et al. 2006), or sample size was small with patients undergoing treatment between measurements (Pinninti et al. 2002). The evidence in support of structural validity for the BSS was strong with factor analysis supporting a one-factor solution (Esfahani et al. 2015; Steer et al. 1993; de Beurs et al. 2015). One study assessed cross-cultural validity of the BSS in a Tehran sample, but did not perform cognitive interviews. However the quality of the translation was fair (with one back and forward translation), giving overall weak evidence in support of the BSS cross-cultural validity (Esfahani et al. 2015).

Columbia Suicide Severity Rating Scale (C-SSRS)

Most studies of the C-SSRS explored the measurement properties of the clinician interview version (7/11 studies). There was mixed evidence for internal consistency: two studies showed high Cronbach's alpha for the whole measure (Madan et al. 2016, Posner et al. 2011), but not on some subscales (Madan et al. 2016), and another study showed a poor alpha (Al-Halabi et al. 2016). There was mixed evidence for reliability: one excellent study showed a large range of inter-rater reliability (r = .5 - .9) (Youngstrom et al. 2016), but two poor studies with small samples showed high agreement between raters (.9+) (Hesdorffer et al. 2013; Mundt et al. 2010). The evidence in support of structural validity was strong (Madan et al. 2016), as was the evidence in support of hypothesis testing and criterion validity (Madan et al. 2016; Al-Halabi et al. 2016; Horwitz, Czyz and King, 2015; Hesdorffer et al. 2013; Mundt et al. 2010; Posner et al. 2011). It is also important to note that one of these studies rated as 'good', showed that the C-SSRS had acceptable specificity and sensitivity (>.7) for predicting future adverse events 6 months after discharge (Madan et al. 2016).

Evidence for responsivity to change was moderate, with two fair studies (Al-Halabi et al. 2016; Posner et al. 2011).One good study found moderate evidence in support of cross-cultural validity, but cognitive interviews were not conducted (Al-Halabi et al. 2016).

Four studies explored the measurement properties of the C-SSRS self-report version. There was weak evidence against hypothesis testing, with one fair study showing a poor correlation with the S-STS (Sheehan et al. 2014). There was mixed evidence for criterion validity: one good study showed high specificity and sensitivity with clinical assessment (Yiguera et al. 2015); one fair study showed evidence against the measure with poor agreement with the S-STS (25), and another good study showed evidence against the measure with poor prediction of future adverse events (Chang and Tan, 2015).

Self-Injurious Thoughts and Behaviours Interview (SITBI)

One study had explored measurement properties of the translated Spanish version of the SITBI in 150 inpatients (Garcia-Nieto et al. 2013). Evidence for reliability was mixed. Evidence for inter-rater reliability was poor given the small subsample in which this was assessed (n=15), but in support of the measure with near perfect agreement between raters (k = .09 – 1). Evidence for test retest reliability was fair, but against the measure with poor reliability for suicidal gestures and self-harm. Evidence for hypothesis testing was fair, but against the measure with poor agreement with certain measures of similar constructs. Evidence for cross-cultural validity was poor, with only a forward translation carried out.

Discussion

Although research shows high rates of suicidality (Cassidy et al. 2014) and death by suicide in autistic adults (Hirvikoski et al. 2016), it was unknown whether any suicidality assessment tools had been used or validated in this group, or whether a robust tool developed for the general population needed to be adapted. Results from this review show that despite

studies having explored suicidality in autistic adults without ID, no research has yet used a validated suicidality assessment tool in this group. This is consistent with the growing body of COSMIN reviews showing a paucity of validated outcome measures for autistic people (Cassidy et al. 2018; Hanratty et al. 2015; Wigham and McConachie, 2014). These results are an important call to action for the research community, to improve the characterisation of outcomes and their measurement in ASC, in research and clinical practice.

Studies of suicidality in ASC were found to utilise a question generated for use in the specific study, without evidence of validity, or used a single question or brief subscale from a broader mental health measure (e.g. PHQ-9, BDI, MINI, SCID). This may reflect the fact that currently many studies of suicidality in ASC have utilised convenience samples from clinical settings, wider studies and existing databases. This lack of standardised and in depth assessment is problematic. For example, single questions from depression measures such as the PHQ-9 do not distinguish self-harm from suicidal intent, and therefore do not assess suicidality per se. The range of measures, many of which lack evidence of validity, could also explain, at least in part, the wide range of suicidality estimates cited in recent reviews of suicidal ideation (11-66%) and attempts (1-35%) in ASC (Hedley and Uljarevic, 2018). A clear recommendation for future suicidality in ASC research is to start using suicidality assessment tools with high quality evidence in support of their measurement properties, in line with the recommendations of COSMIN (Mokkink et al. 2016). We make recommendations on future selection of such tools based on our synthesis of the available evidence below.

A number of validated suicidality assessment tools have been used frequently in studies of general population adults, without ID or co-occurring conditions; the SBQ-R, C-SSRS, BSS, Paykel and SITBI. Interestingly, no studies were revealed from the comprehensive search that had assessed the measurement properties of the Paykel, despite it being utilised in a number of research studies. The C-SSRS and BSS had also been validated mainly in

psychiatric samples despite being used in a number of general population studies. Importantly, although the evidence for hypothesis testing and criterion validity was mixed for the BSS, this clearly depended on the context in which this tool was used. Specifically, the BSS had strong evidence in support of distinguishing sub-groups (e.g. those who have and have not attempted suicide), but strong evidence against predicting future adverse events (e.g. hospital admissions for suicide attempt). The BSS also had strong evidence for internal consistency, structural validity, and moderate evidence for cross-cultural validity. Hence, the strengths of the BSS lie in distinguishing sub groups in research, but not when predicting future adverse events in clinical practice.

Two versions of the C-SSRS were assessed; the self-report and clinician interview versions. The self-report version has been more recently developed, and therefore fewer studies (4) were available assessing its measurement properties than the clinician interview version (7). For the self-report C-SSRS, there was weak evidence against hypothesis testing, and mixed evidence for criterion validity. Specifically, there was moderate evidence in support of agreement between the C-SSRS self-report and clinician assessment (Yiguerta et al. 2015), but moderate evidence against the C-SSRS self-report predicting future adverse events (Chang and Tan, 2015). Currently, there is not yet enough evidence to recommend this tool for use in research or clinical practice.

However, the clinician interview version of the C-SSRS had evidence in support of a number of measurement properties. The strengths of the measure lie in structural validity, hypothesis testing, criterion validity and responsiveness to change, and weak evidence in support of cross-cultural validity. Importantly, there was moderate evidence in support of the C-SSRS predicting future suicidal behaviour within 6 months of discharge (Madan et al. 2016). There was however mixed evidence for internal consistency and reliability. This suggests that the clinician interview version of the C-SSRS is likely to be most useful in

clinical contexts, to aid clinicians in helping to gauge potential suicide risk as part of a holistic psychosocial assessment, and changes in response to treatment or within clinical trials. However, more research is needed to establish evidence in support of inter-rater agreement, and internal consistency, particularly concerning subscales.

There was only one study that had explored the measurement properties of the SITBI in adults without ID (with one additional validation study in an adolescent sample which was not included) (Garcia-Nieto et al. 2013). Hence there was limited evidence in support of its measurement properties. Notably, the study showed evidence against hypothesis testing with low agreement with measures of similar constructs. Future research needs to establish the measurement properties of this tool.

There were only two studies exploring the measurement properties of the SBQ-R, despite being used in a number of general population studies of suicidality. Despite this, there was strong evidence in support of internal consistency and structural validity, moderate evidence in support of criterion validity, and weak evidence in support of hypothesis testing. In particular, the SBQ-R showed evidence for high sensitivity and specificity for distinguishing sub-groups using the first item (Aloba et al. 2017) and total scores (Osman et al. 2001; Aloba et al. 2017). Notably, the SBQ-R is the briefest tool out those identified in this review (with 1-4 items), does not carry a cost to use, and has comparable quality of evidence in support of a range of measurement properties compared to the other scales which are longer and carry a cost (C-SSRS and BSS). Hence, the SBQ-R could be particularly useful for future research.

In summary, the current study revealed strong consistent evidence across three frequently used suicidality assessment tools (BSS, C-SSRS and SBQ-R), for reliably distinguishing sub-groups (e.g. those who have or have not attempted suicide in the past). However, there were relatively few studies exploring an important component of criterion

validity for suicidality assessment tools – prediction of future adverse events (e.g. future suicidal behaviour, future hospitalisations or emergency department visits). Research has suggested that suicidality assessment tools on the whole are poor predictors of future attempts, many perform worse than patient or clinical assessment, and may therefore be a waste of valuable resources (Quinlivan et al. 2017; Quinlivan et al. 2016). The current study adds useful evidence to this debate, as it is the first to use a validated research tool (COSMIN), to synthesise the quality of the evidence for a range of measurement properties, across a number of studies. On the basis of our synthesis of the available evidence, results suggest that certain tools (i.e. C-SSRS interview) may have greater utility in predicting future adverse events than others (e.g. BSS). Results also suggest that designs which assess criterion validity on the basis of distinguishing sub-groups may over-estimate diagnostic accuracy of a tool. This is consistent with previous research (Lijmer et al. 1999), which recommends the use of cohort studies in assessing the usefulness of suicidality assessment tools.

Future Research

No studies have yet utilised any of the suicidality assessment tools that have been developed for and widely used in the general population, in autistic adults. As discussed above, the characteristics of ASC, and differing presentation of suicidality in this group, could all affect the utility of these tools. A first step would be to explore the content validity of these existing tools through focus groups and cognitive interviews, to inform adaptations, prior to exploring other measurement properties of these tools. COSMIN criteria stipulates that excellent studies should compare the performance of adapted to original measures (Mokkink et al. 2016). We also recommend comparing the performance of measures between ASC and general population groups, to ascertain whether measurement properties of tools are similarly robust in autistic and non-autistic populations. For example, if a

measure designed for the general population does not adequately capture a health outcome in ASC, then we would expect a different factor structure, lower internal consistency, and criterion validity compared to the general population, which should then improve for the adapted version (see Cassidy et al. 2018).

Out of the tools identified and evaluated in this review, the SBQ-R is a free brief measure with only 4 questions, with comparable evidence in support of a range of psychometric measures compared to longer and more expensive tools (BSS and C-SSRS). It is therefore a promising potential candidate tool to begin exploring suicidality in ASC in research now, as an important stop gap before validated tools become available. Items one and two of the SBQ-R focus on suicidal thoughts and behaviours over one's lifetime and in the past year, with clear definitions, e.g. "rarely (1 time)". This could be potentially useful for assessing presence of suicidal thoughts and behaviours in autistic adults. However, items three and four will likely require adaptations for autistic adults which importantly provide more information on risk level. For example, autistic people's communication difficulties may mean they are less likely to have had spontaneously communicated their suicidal intent to others in the past (item three), despite high risk. Additionally, literal interpretation and difficulties in imagination and abstract future thinking in ASC may lead to difficulties interpreting and responding to the final question (item four) about likelihood of attempting suicide in the future.

A crucial aspect of exploring validity of suicidality assessment tools, are whether these are useful to clinicians in gauging risk of future suicide attempts. However, few studies have explored this crucial aspect of criterion validity. Hence, it is critical that future studies assessing criterion validity of suicidality assessment tools in autistic and general populations not only rely on distinguishing sub-groups, which over-estimate diagnostic accuracy of a tool. Rather, cohort studies are needed to assess whether current and adapted suicidality

assessment tools can predict future suicidal behaviour significantly more accurately than clinician opinion or patient self-report.

Strengths and Limitations

A key strength was using a rigorous method (COSMIN) to systematically identify and evaluate relevant studies. However, following this strict method meant that some tools were excluded from the analysis, such as single items or subscales from broader mental health measures. As suicidality in ASC is such a new area of research, it could be argued that adopting such rigorous methods might have led us to overlook other relevant data which could indicate the usefulness of one tool over another. However, we were interested in more specific and broader conceptualisations of suicidality than is feasible in single questions or subscales. We also focused on tools which had been used frequently in general population adults, without ID, or co-occurring conditions, rather than including measures only used in psychiatric groups, as these tools were more likely to be useful for a range of non-clinical and clinical groups, and in a range of clinical and research contexts. Our search was also limited by focusing only on studies in English, due to lack of translation resources, and data extraction was also conducted only in part by two independent reviewers. Although COSMIN is a validated research tool, there is a certain level of subjectivity in rating each article. However, there was good agreement between raters in the current study (73%), similar to previous COSMIN reviews (e.g. Cassidy et al. 2018; Wigham and McConachie, 2014).

Conclusion

This is the first systematic review to use a robust research tool (COSMIN) to synthesise the evidence regarding the assessment of suicidality in autistic and general population adults without ID. Although a growing number of studies are beginning to assess suicidality in

autistic adults, none have yet used a validated suicidality assessment tool, and there are currently no validated suicidality assessment tools available for this group. Three robust suicidality assessment tools were identified which have been used frequently in general population studies. Future ASC studies must begin to use and explore the measurement properties of such robust tools designed for the general population. Our research group are currently undertaking this research in order to better characterise suicidality and its measurement in ASC.

Funding Source

This work was supported by an Economic and Social Research Council Research Grant [ES/N000501/1].

Author Biography

Dr Sarah Cassidy holds a PhD in Psychology from the University of Nottingham, UK, where she is currently an Assistant Professor. Dr Cassidy currently leads a research program aiming to understand and prevent mental health problems, self-harm and suicidality in autistic adults (Mental Health in Autism). Dr Cassidy's seminal work was the first the show significantly increased risk of suicidal thoughts and suicidal behaviours in newly diagnosed adults with Aspergers Syndrome, compared with the general population and other clinical groups.

Author Disclosure Statements

Role of funding sources

This research was supported by the Economic and Social Research Council, [ES/N000501/1]. The funder had no role in the study design, collection, analysis or interpretation of data, writing of the report, or the decision to submit the article for publication.

Contributions

SC designed the study, wrote the search protocol, conducted the searches, data extraction, and evidence synthesis. LB assisted with data extraction, and critical revision of the search protocol. EB assisted with study design, and critical revision of the search protocol. SW assisted with data extraction, and evidence synthesis. JR assisted with study design, and critical revision of the search protocol. SC wrote the first draft of the manuscript and all authors contributed to and have approved the final manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest.

References

Al-Halabí, S., Sáiz, P. A., Burón, P., Garrido, M., Benabarre, A., Jiménez, E., ... & Muñiz, J. (2016). Validation of a Spanish version of the Columbia-Suicide Severity Rating Scale (C-SSRS). *Revista de Psiquiatría y Salud Mental (English Edition)*, *9*(3), 134-142.

Aloba, O., Ojeleye, O., & Aloba, T. (2017). The psychometric characteristics of the 4-item Suicidal Behaviors Questionnaire-Revised (SBQ-R) as a screening tool in a non-clinical sample of Nigerian university students. *Asian journal of psychiatry*, *26*, 46-51.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub.

Anestis, M. D., Pennings, S. M., & Williams, T. J. (2014). Preliminary results from an examination of episodic planning in suicidal behavior. *Crisis, 35*, 186-192.

Aradilla-Herrero, A., Tomás-Sábado, J., & Gómez-Benito, J. (2014). Associations between emotional intelligence, depression and suicide risk in nursing students. *Nurse Education Today*, *34*(4), 520-525.

Beck, A. T., Schuyler, D., & Herman, I. (1974). *Development of suicidal intent scales*. Charles Press Publishers.

Beck, A. T., Steer, R. A., & Ranieri, W. F. (1988). Scale for suicide ideation: Psychometric properties of a self-report version. *Journal of clinical psychology*, *44*(4), 499-505.

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, *133*(5), 1515-1525.

Cassidy, S. A., Bradley, L., Bowen, E., Wigham, S., & Rodgers, J. (2018). Measurement properties of tools used to assess depression in adults with and without autism spectrum conditions: A systematic review. *Autism Research*.

Cassidy, S., & Rodgers, J. (2017). Understanding and prevention of suicide in autism. *The Lancet Psychiatry, 4*(6), e11.

Cassidy, S., Bradley, L., Shaw, R., and Baron-Cohen, S. (Submitted). Risk markers for suicidality in autistic adults.

Cassidy, S., Bradley, P., Robinson, J., Allison, C., McHugh, M., & Baron-Cohen, S. (2014). Suicidal ideation and suicide plans or attempts in adults with Asperger's syndrome attending a specialist diagnostic clinic: a clinical cohort study. *The Lancet Psychiatry, 1*(2), 142-147.

Chang, B. P., & Tan, T. M. (2015). Suicide screening tools and their association with near-term adverse events in the ED. *The American journal of emergency medicine*, *33*(11), 1680-1683.

Cochrane-Brink, K. A., Lofchy, J. S., & Sakinofsky, I. (2000). Clinical rating scales in suicide risk assessment1. *General Hospital Psychiatry*, *22*(6), 445-451.

Coric, V., Stock, E. G., Pultz, J., Marcus, R., & Sheehan, D. V. (2009). Sheehan Suicidality Tracking Scale (Sheehan-STS): preliminary results from a multicenter clinical trial in generalized anxiety disorder. *Psychiatry (Edgmont)*, *6*(1), 26.

Crane, L., Adams, F., Harper, G., Welch, J., & Pellicano, E. (2018). 'Something needs to change': Mental health experiences of young autistic adults in England. *Autism*, 1362361318757048.

Crane, L., Lind, S. E., & Bowler, D. M. (2013). Remembering the past and imagining the future in autism spectrum disorder. *Memory*, *21*(2), 157-166.

De Beurs, D. P., Fokkema, M., & O'Connor, R. C. (2016). Optimizing the assessment of suicidal behavior: the application of curtailment techniques. *Journal of affective disorders*, *196*, 218-224.

de Beurs, D. P., Fokkema, M., de Groot, M. H., de Keijser, J., & Kerkhof, A. J. (2015). Longitudinal measurement invariance of the Beck Scale for Suicide Ideation. *Psychiatry research*, *225*(3), 368-373.

De Vet, H.C.W., Terwee, C.B., Mokkink, L.B., & Knol, D.L. (2011). *Measurement in medicine. A practical guide.* Cambridge: Cambridge University Press.

Esfahani, M., Hashemi, Y., & Alavi, K. (2015). Psychometric assessment of beck scale for suicidal ideation (BSSI) in general population in Tehran. *Medical journal of the Islamic Republic of Iran*, *29*, 268.

First, M. B., Gibbon, M., Spitzer, R. L., & Benjamin, L. S. (1997). *User's guide for the structured clinical interview for DSM-IV axis II personality disorders: SCID-II*. American Psychiatric Pub.

García-Nieto, R., Blasco-Fontecilla, H., Yepes, M. P., & Baca-García, E. (2013). Translation and validation of the "Self-Injurious Thoughts and Behaviours Interview" in a Spanish population with suicidal behaviour. *Revista de Psiquiatría y Salud Mental (English Edition)*, *6*(3), 101-108.

Hannon, G., & Taylor, E. P. (2013). Suicidal behaviour in adolescents and young adults with ASD: Findings from a systematic review. *Clinical psychology review, 33*(8), 1197-1204.

Hanratty, J., Livingstone, N., Robalino, S., Terwee, C. B., Glod, M., Oono, I. P., … McConachie, H. *(2015).* Systematic review of the measurement properties of tools used to measure behaviour problems in young children with autism. *PloS one, 10,* e0144649.

Happé, F.G. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbol, 10*, 275–295.

Healy, D. J., Barry, K., Blow, F., Welsh, D., & Milner, K. K. (2006). Routine use of the Beck Scale for Suicide Ideation in a psychiatric emergency department. *General hospital psychiatry*, *28*(4), 323-329.

Hedley, D., & Uljarević, M. (2018). Systematic Review of Suicide in Autism Spectrum Disorder: Current Trends and Implications. *Current Developmental Disorders Reports*, 1-12.

Hedley, D., Uljarević, M., Wilmot, M., Richdale, A., & Dissanayake, C. (2018). Understanding depression and thoughts of self-harm in autism: A potential mechanism involving loneliness. *Research in Autism Spectrum Disorders, 46*, 1-7.

Hedley, D., Uljarević, M., Wilmot, M., Richdale, A., & Dissanayake, C. (2017). Brief report: Social support, depression and suicidal ideation in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*(11), 3669-3677.

Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IVand ICD-10. *Journal of Clinical Psychiatry*, *59*(Suppl 20), 2233.

Hesdorffer, D. C., French, J. A., Posner, K., DiVentura, B., Pollard, J. R., Sperling, M. R., ... & Kanner, A. M. (2013). Suicidal ideation and behavior screening in intractable focal epilepsy eligible for drug trials. *Epilepsia*, *54*(5), 879-887.

Hirvikoski, T., Mittendorfer-Rutz, E., Boman, M., Larsson, H., Lichtenstein, P., & Bölte, S. (2016). Premature mortality in autism spectrum disorder. *The British Journal of Psychiatry, 208*(3), 232-238.

Horon, R., McManus, T., Schmollinger, J., Barr, T., & Jimenez, M. (2013). A study of the use and interpretation of standardized suicide risk assessment: Measures within a psychiatrically hospitalized correctional population. *Suicide and Life-Threatening Behavior*, *43*(1), 17-38.

Horwitz, A. G., Czyz, E. K., & King, C. A. (2015). Predicting future suicide attempts among adolescent and emerging adult psychiatric emergency patients. *Journal of Clinical Child & Adolescent Psychology*, *44*(5), 751-761.

Innamorati, M., Pompili, M., Lester, D., Tatarelli, R., & Girardi, P. (2008). Recreational drug use and suicidality among Italian young adults. *Journal of addictive diseases*, *27*(4), 51-59.

Kamphaus, R. W., & Frick, P. J. (2005). *Clinical assessment of child and adolescent personality and behavior*. Springer Science & Business Media.

Kliem, S., Lohmann, A., Mößle, T., & Brähler, E. (2017). German Beck Scale for Suicide Ideation (BSS): psychometric properties from a representative population survey. *BMC psychiatry*, *17*(1), 389.

Lijmer, J. G., Mol, B. W., Heisterkamp, S., Bonsel, G. J., Prins, M. H., van der Meulen, J. H., & Bossuyt, P. M. (1999). Empirical evidence of design-related bias in studies of diagnostic tests. *Jama*, *282*(11), 1061-1066.

Lind, S. E., & Bowler, D. M. (2010). Episodic memory and episodic future thinking in adults with autism. *Journal of abnormal psychology, 119*(4), 896.

Madan, A., Frueh, B. C., Allen, J. G., Ellis, T. E., Rufino, K. A., Oldham, J. M., & Fowler, J. C. (2016). Psychometric Reevaluation of the Columbia-Suicide Severity Rating Scale: Findings From a Prospective, Inpatient Cohort of Severely Mentally Ill Adults. *The Journal of clinical psychiatry*, *77*(7), e867-73.

Maddox, B. B., Trubanova, A., & White, S. W. (2017). Untended wounds: Non-suicidal self-injury in adults with autism spectrum disorder. *Autism, 21*(4), 412-422.

Matson, J. L., & Nebel-Schwalm, M. (2007). Assessing challenging behaviors in children with autism spectrum disorders: A review. *Research in Developmental Disabilities, 28*(6), 567-579.

McConachie, H., Parr, J. R., Glod, M., Hanratty, J., Livingstone, N., Oono, I. P., ... & Garland, D. (2015). Systematic review of tools to measure outcomes for young children with autism spectrum disorder.

Metalsky, G. I., & Joiner, T. E. (1997). The hopelessness depression symptom questionnaire. *Cognitive Therapy and Research*, *21*(3), 359-384.

Miller, I. W., Norman, W. H., Bishop, S. B., & Dow, M. G. (1986). The Modified Scale for Suicidal Ideation: reliability and validity. *Journal of consulting and clinical psychology*, *54*(5), 724.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews, 4*(1), 1.

Mokkink, L. B., de Vet, H., Prinsen, C. A., Patrick, D., Alonso, J., Bouter, L. M., & Terwee, C. B. (2016). COSMIN checklist 2.0 for assessing the methodological quality of studies on the measurement properties of Patient-Reported Outcome Measures. *Quality of Life Research, 25*, 21-21.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., & Stratford, P.W. (2012). COSMIN checklist manual. VU University Medical Center, Department of Epidemiology and Biostatistics (www.cosmin.nl).

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., … De Vet, H.C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research, 19*, 539–549.

Mundt, J. C., Greist, J. H., Gelenberg, A. J., Katzelnick, D. J., Jefferson, J. W., & Modell, J. G. (2010). Feasibility and validation of a computer-automated Columbia-Suicide Severity Rating Scale using interactive voice response technology. *Journal of psychiatric research, 44*(16), 1224-1228.

Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-Injurious Thoughts and Behaviors Interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment*, *19*(3), 309-317.

Orsmond, G. I., Shattuck, P. T., Cooper, B. P., Sterzing, P. R., & Anderson, K. A. (2013). Social participation among young adults with an autism spectrum disorder. *Journal of autism and developmental disorders*, *43*(11), 2710-2719.

Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment*, *8*(4), 443-454.

Paykel, E. S., Myers, J. K., Lindenthal, J. J., & Tanner, J. (1974). Suicidal feelings in the general population: a prevalence study. *The British Journal of Psychiatry*, *124*(582), 460-469.

Pelton, M. K., & Cassidy, S. A. (2017). Are autistic traits associated with suicidality? A test of the interpersonal-psychological theory of suicide in a non-clinical young adult sample. *Autism research*, *10*(11), 1891-1904.

Pinninti, N., Steer, R. A., Rissmiller, D. J., Nelson, S., & Beck, A. T. (2002). Use of the Beck Scale for Suicide Ideation with psychiatric inpatients diagnosed with schizophrenia, schizoaffective, or bipolar disorders. *Behaviour research and therapy*, *40*(9), 1071-1079.

Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., ... & Mann, J. J. (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.

Quinlivan, L., Cooper, J., Davies, L., Hawton, K., Gunnell, D., & Kapur, N. (2016). Which are the most useful scales for predicting repeat self-harm? A systematic review evaluating risk scales using measures of diagnostic accuracy. *BMJ open*, *6*(2), e009297.

Quinlivan, L., Cooper, J., Meehan, D., Longson, D., Potokar, J., Hulme, T., ... & Page, L. (2017). Predictive accuracy of risk scales following self-harm: multicentre, prospective cohort study. *The British Journal of Psychiatry*, *210*(6), 429-436.

Raja, M. (2014). Suicide risk in adults with Asperger's syndrome. *The Lancet Psychiatry*, *1*(2), 99-101.

Rudd, M. D. (1989). The prevalence of suicidal ideation among college students. *Suicide and Life-Threatening Behavior*, *19*(2), 173-183.

Segers, M., & Rawana, J. (2014). What do we know about suicidality in autism spectrum disorders? A systematic review. *Autism Research, 7*(4), 507-521.

Sheehan, D. V., Alphs, L. D., Mao, L., Li, Q., May, R. S., Bruer, E. H., ... & Williamson, D. J. (2014). Comparative validation of the S-STS, the ISST-Plus, and the C–SSRS for assessing the suicidal thinking and behavior FDA 2012 suicidality categories. *Innovations in clinical neuroscience*, *11*(9-10), 32.

Stanley, B., Träskman-Bendz, L., & Stanley, M. (1986). The suicide assessment scale: a scale evaluating change in suicidal behavior. *Psychopharmacology bulletin*, *22*(1), 200.

Steer, R. A., Rissmiller, D. J., Ranieri, W. F., & Beck, A. T. (1993). Dimensions of suicidal ideation in psychiatric inpatients. *Behaviour research and therapy*, *31*(2), 229-236.

Terwee, C.B., Jansma, E.P., Riphagen, I.I., & de Vet, H.C.W. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research, 18*, 1115–1123.

Viguera, A. C., Milano, N., Laurel, R., Thompson, N. R., Griffith, S. D., Baldessarini, R. J., & Katzan, I. L. (2015). Comparison of electronic screening for suicidal risk with the Patient Health Questionnaire Item 9 and the Columbia Suicide Severity Rating Scale in an outpatient psychiatric clinic. *Psychosomatics*, *56*(5), 460-469.

Wigham, S., & McConachie, H. (2014). Systematic review of the properties of tools used to measure outcomes in anxiety intervention studies for children with autism spectrum disorders. *PLoS One, 9*, e85268.

Youngstrom, E. A., Hameed, A., Mitchell, M. A., Van Meter, A. R., Freeman, A. J., Algorta, G. P., ... & Meyer, R. E. (2015). Direct comparison of the psychometric properties of multiple interview and patient-rated assessments of suicidal ideation and behavior in an adult psychiatric inpatient sample. *Journal of clinical psychiatry*, *76*(12), 1676-1682.

Zahid, S., & Upthegrove, R. (2017). Suicidality in Autistic Spectrum Disorders. *Crisis.*

Figure Captions:

Figure 1: Results of Search One

Figure 2: Results of Search Two

Tables

Table 1: Stage 1 review search terms.

1. **(general population or population sample or community sample or national* survey or household* survey or non referred or non clinical or population screen*)**
2. **(ASC or ASD or Asperg* or Autis* or high functioning or pervasive developmental disorder* or PDD or HFA)**
3. **(adult*)**
4. **(assess* or tool or treatment outcome or measur* or scale or quotient or inventory or instrument)**
5. **(suicid* or self harm or self inj* or parasuicide or suicide attempts or attempted suicide)**
6. **randomised controlled trial or randomized controlled trial**
7. **random***
8. **comparative stud***
9. **prospective stud***
10. **intervention**
11. **treatment effectiveness evaluation or treatment response or treatment study**
12. **epidemiolog***
13. **prevalence**
14. General Population Search (**6 or 7 or 8 or 9 or 10 or 11 or 12 or 13) and (1 and 3 and 4 and 5)**
15. Autism Spectrum Condition Search (**6 or 7 or 8 or 9 or 10 or 11 or 12 or 13) and (2 and 3 and 4 and 5)**
16. **limit 14 and 15 to English Language; 1992 – current; age 18 years +**

Table 2: Methodological quality of studies included in the qualitative synthesis.

Tool	Article	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity	Criterion validity	Responsiveness
SBQ-R	Osman et al. (2001).	poor					fair		fair	
SBQ-R	Aloba et al. (2017).	excellent				excellent			good	
BSS	De Beurs Fokkema and O'Connor (2016).						fair		excellent	
BSS	Esfahani, Hashemi and Alavi (2015).	excellent				good	fair	fair		
BSS	Chang and Tan (2015).								good	
BSS	Steer et al.					good				

	(1993).									
BSS	De Beurs et al. (2015).					excellent	good			
BSS	Ayuh (2008).	fair				fair		fair		
BSS	Horon et al. (2012).						good		good	
BSS	Chioqueta and Stile (2006).	poor					fair	poor		
BSS	Healy et al. (2006).		poor				good			
BSS	Holden and DeLisle (2005).					good				
BSS	Pinninti et al. (2002).	poor	poor				fair		good	
BSS	Cochrane-Brink et al.						poor		poor	

	(2000).									
BSS	Kliem et al. (2017).	poor					fair			
C-SSRS (self-report electronic version)	Mundt et al. (2013).								good	
C-SSRS (self-report)	Viguera et al (2015).								good	
C-SSRS (self-report)	Sheehan et al. (2014).						fair		fair	
C-SSRS (self-report)	Chang and Tan (2015).								good	
C-SSRS (interview)	Madan et al. (2016).	good				good	fair		good	
C-SSRS (interview)	Al-Halabi et al. (2016).	good				good	fair	fair		fair

C-SSRS (interview)	Youngstrom et al. (2015).	poor	fair				fair		good	
C-SSRS (interview)	Horwitz, Czyz and King (2015).						good		good	
C-SSRS (interview)	Hesdorffer et al. (2013).		poor				fair			
C-SSRS (interview)	Mundt et al. (2010).		poor				poor			
C-SSRS (interview)	Posner et al. (2011).	poor					fair		good	fair
SITBI	Garcia-Nieto et al. (2013).		Poor/ fair				fair	poor		

Table 3: Collated evidence of measurement properties for each tool.

Measure	**Version**	**Measurement Properties**								**Interpretability**
		Internal Consistency	Reliability	Content Validity	Structural Validity	Hypothesis Testing	Criterion Validity	Responsiveness	Cross-cultural validity	Differences in scores between groups
SBQ-R	V2	+++			+++	+	++			Y
BSS	-	+++*	?		+++	+/-	+/-		++	Y
C-SSRS	Self-report					-	+/-			
	Interview	+/-	+/-		+++	++	+++	++	+	
SITBI	-		+/-			-			?	

*denotes evidence from translated version only; ? indeterminant evidence for or against a measurement property; +/- mixed evidence for and against a measurement property; + consistent evidence in support of a measurement property (+ weak, ++ moderate, +++ strong evidence) evidence; - consistent evidence against a measurement property (- weak, -- moderate, --- strong evidence).

Appendices:

Appendix A: Characteristics of suicidality tools identified in search one.

Measure	Version	Year published	Aim of tool	Number of items	Subscales	Response options (e.g. 4 point scale, yes/no etc.)	Format (e.g. self-report questionnaire, interview etc.)	Used in which references?	
								ASC	Gen Pop
Columbia Suicide Severity Rating Scale (C-SSRS)	Interview Self-report	2011 2009	Measure of suicide risk in research and clinical practise.	6 6	N/A	Yes/No Yes/No	Clinician administered interview Self-report questionnaire	N/A	DeVyld et al (2015) Childress et al. (2010)
Measure of episodic planning of suicide (MEPOS)	V1	2014	Assesses frequency and characteristics of prior suicide attempts	4	N/A	Yes/No and single response follow up questions	Self-report questionnaire	N/A	Anestis et al. (2014)
Suicide Behaviours Questionnaire Revised	V2	2001	Measure of suicide risk	4	N/A	5/6 point scales	Self-report questionnaire	N/A	Campose and Holden (2014);

(SBQ-R)									Campos, Besser and Blatt (2013); Wagner et al. (2013).
Beck Scale for Suicidal Ideation (BSS)	V2	1991	Assesses suicidal ideation and behaviours	21	1-19 current suicidal ideation, 20-21 past suicide attempts	Yes/No (check this!)	Self-report questionnaire	N/A	Spiker et al. (2012); Van Orden et al. (2012); Cleary, Nixon and Fitzgerald, (2007); Lane, Cheref and Miranda (2016); Zuromski et al (2017).
Beck Suicide Intent Inventory	V1	1974	Assesses risk of suicidal attempts in patients who have	15	1-8 objective circumstances surrounding suicide	Scale 1-3 options	Clinician administered and self-report questionnaire	N/A	Moran et al. (2012).

			attempted suicide		attempt, 9-15 self-report questions surrounding suicide attempt				
Depression Severity Index – Suicide Subscale (DSI-SS)	V1	1997	Identify frequency and intensity of current suicidal ideation and impulses in the past 2 weeks	4	N/A	Scale 1-4 options	Self-report questionnaire	N/A	Cukrowicz et al. (2009).
Modified Scale for Suicidal Ideation (MSSI)	V1	1986	Assess presence and severity of current suicidal ideation in the past 2 weeks.	18	N/A	Scale 1-4 options	Self report questionnaire	N/A	Bagge et al. (2014).
Paykel	V1	1974	Assess lifetime and current experience of suicidal ideation and	5	N/A	Scale 1-4	Clinician administered interview or self-report questionnaire	N/A	Jonson et al (2012); Rancāns et al. (2003);

			behaviours						Renberg, (2001).
Sheehan Suicidality Tracking Scale (S-STS)	V10	2009	Tracks spontaneous and treatment emergent suicidal ideation and behaviours	8	Self-injury, self-harm, suicidal ideation and suicide attempts.	Yes/No or scale 1-4	Clinician administered or self report rating scale	N/A	Preti et al. (2013).
Suicide Assessment Scale (SUAS-S)	V1	2006	Assesses signs and symptoms related to suicidality	20	N/A	Scale 0-4	Self report questionnaire	N/A	Zhang et al. (2012).
Suicide Ideation Scale (SIS)	V1	1989	Measured a continuum of suicidal thoughts and attempts in clinical and non clinical samples	10	N/A	Scale 1-5	Self report questionnaire	N/A	Chu et al. (2008).
Suicide Score Scale (SSS)	V1	2008	Designed to obtain information about life time and past	12	N/A	Yes/No	Self report questionnaire	N/A	Innamorati et al. (2008).

			year previous suicidal ideation, planning or attempts						
Self-Injurious Thoughts and Behaviours Interview (SITBI)	V1	2007	A structured interview that assesses the presence, frequency, and characteristics of a wide range of self-injurious thoughts and behaviours, including suicidal ideation, suicide plans, suicide gestures, suicide attempts, and non-suicidal self-injury (NSSI).	169 items across 5 modules	5 subscales: (a) suicidal ideation; (b) suicide plans; (c) suicide gestures; (d) suicide attempts; and (e) non-suicidal self-injury.	Yes/No and scales 1-4	Interview	N/A	Dhingra, Boduszek and Klonsky (2016); Mortier et al. (2017a); Mortier et al. (2017b).

Plutchik Suicide Risk Scale (PSRS)	V1	1989	Measures the degree to which an individual reveals characteristics similar to those of a suicide prototype.	26 items	N/A	Yes/No	Self-report	N/A	Pereira-Morales et al. (2017).

Appendix B: Characteristics of study populations included in the qualitative synthesis.

Measure	Article	Study population/ sample	Study type (prospective, case-control etc.)	Mean age (SD) years; range	Total N	Male n, female n.	Country
SBQ-R	Osman et al. (2001).	Psychiatric patients	Case-control	15.63/15.56 (0.98) 14-17 years	120	65 male, 55 female	US
		High school adolescent		16.51/16.47 (1.33/1.14) 14-18 years	138	72 male, 66 female	
		Psychiatric patients		32.14/33.47 (7.43/8.79) age range not reported	120	65 male, 55 female	
		Undergraduate Psychology student		21.19/20.97 (2.98/2.91) age range not reported	135	69 male, 66 female	
SBQ-R	Aloba et al. (2017).	Adult undergraduate students	Cross-sectional	22.51 (2.94) 18–31 years	536	272 (50.7%) male, 263 (49.3%) female	Nigeria

BSS	De Beurs Fokkema and O'Connor (2016).	Psychiatric patients	Longitudinal	33/38 (13.2/13.8) age range not provided	366	158 male, 208 female	UK
BSS	Esfahani, Hashemi and Alavi (2015).	General population	Cross-sectional	27 (9.5)18-70 years	535	138 male, 397 female	Tehran
BSS / C-SSRS (self report)	Chang and Tan (2015).	Psychiatric patients	Prospective	36.4 age range 20-47 years	50	22 male, 28 female	US
BSS	De Beurs et al. (2015).	Psychiatric patients	Longitudinal	43 (15) age range not provided	872 (at baseline)	415 male, 457 female	Netherlands
BSS	Horon et al. (2013).	Psychiatric Patients	Cross-sectional	37, age range 19-75	342	342 male	US
BSS	Ayub (2008).	General population adolescents and young adults	Cross-sectional	20.06, (2.39) 17-25 years.	904	442 female	Pakistan
BSS	Chioqueta and Stiles (2006).	University students	Cross-sectional	21.46 (3.63) 17-44 years.	314	71 male, 243 female	Norway
BSS	Healy et al. (2006)	Psychiatric patients	Cross-sectional	37.51 (13.52), 18-37 years.	735	319 male, 413 female	US
BSS	Holden and DeLisle (2005).	Psychology students and community sample suicide	Cross-sectional	24.39 (11.47), 17-68 years.	134	15 male, 119 female	Canada

		attempters					
BSS	Pinninti et al (2002).	Psychiatric patients	Cross-sectional	41.68 (4.91), age 18+	130	64 male, 66 female	US
BSS	Cochrane-Brink et al (2000).	Psychiatric patients	Cross-sectional	34.7 (10.9), age 18+	55	31 male, 24 female	Canada
BSS	Steer et al. (1993).	Psychiatric patients	Cross-sectional	38.28 (14.14), age range not provided	330	154 male, 176 female	US
BSS	Kliem et al. (2017).	General population adults identified as high suicide risk on BSS	Cross-sectional	49.7 (17.83), age 18+	112	53% male	Germany
C-SSRS	Madan et al. (2016).	Inpatient cohort of mentally ill adults	Prospective	35.2 (+/- 14.7 years)	1055	540 male, 515 female	US
C-SSRS	Al-Halabi et al. (2016).	Psychiatric patients	Cross-sectional	46.93	467	Not specified	Spain
C-SSRS	Youngstrom et al. (2015).	Adult inpatients	Cross-sectional	38.5 (+/- 12.4 years)	199	57% female	US
C-SSRS	Viguera et al. (2015).	Psychiatric patients	Cross-sectional	43.7 (+/- 14.9 years), age range 18-94 years	1416	533 male, 883 female	US
C-SSRS	Sheehan et al.	Adults with self-injurious	Cross sectional	39.9 (15), age range 19-73	40	44.4% male	US

	(2014).	behaviour		years			
C-SSRS	Horwitz, Czyz and King (2015).	Psychiatric patients	Longitudinal	19.38 (2.9), age range 15-24 years.	473	220 male, 253 female	US
C-SSRS	Mundt et al. (2013).	Clinical samples from treatment trials	Prospective	Not reported	3776	Not reported	US
C-SSRS	Hesdorffer et al. (2013).	Treatment resistant focal epilepsy	Cross-sectional	41.2 (31.2), age range 18-70 years	208	71 male, 137 female	US
C-SSRS	Mundt et al. (2010).	Psychiatric patients General population controls	Case-control	30.9 years, age range 18-57 years 32.9 year, age range 24-40 years	10 10	4 male, 6 female 1 male, 9 female	US
C-SSRS	Posner et al. (2011).	Adolescent suicide attempters Depressed adolescents Psychiatric patients	Treatment study Medication efficacy trial – longitudinal Cross	12-18 years 11-17 years Over 18 years	124 312 237		US

			sectional study				
SITBI	Garcia-Nieto et al. (2013).	Psychiatric patients	Cross sectional	43.3 years (10.3), no age range given	150	84 (56) female	Spain

Appendix C: References identified from search one (Appendix A).

Anestis, M. D., Kleiman, E. M., Lavender, J. M., Tull, M. T., & Gratz, K. L. (2014). The pursuit of death versus escape from negative affect: An examination of the nature of the relationship between emotion dysregulation and both suicidal behavior and non-suicidal self-injury. *Comprehensive psychiatry*, *55*(8), 1820-1830.

Bagge, C. L., Lamis, D. A., Nadorff, M., & Osman, A. (2014). Relations between hopelessness, depressive symptoms and suicidality: mediation by reasons for living. *Journal of Clinical Psychology*, *70*(1), 18-31.

Campos, R. C., & Holden, R. R. (2014). Suicide risk in a Portuguese non-clinical sample of adults. *The European Journal of Psychiatry*, *28*(4), 230-241.

Campos, R. C., Besser, A., & Blatt, S. J. (2013). Recollections of parental rejection, self-criticism and depression in suicidality. *Archives of Suicide Research*, *17*(1), 58-74.

Childress, A. C., & Berry, S. A. (2010). The single-dose pharmacokinetics of NWP06, a novel extended-release methylphenidate oral suspension. *Postgraduate medicine*, *122*(5), 35-41.

Chioqueta, A. P., & Stiles, T. C. (2006). Psychometric properties of the Beck Scale for Suicide Ideation: A Norwegian study with university students. *Nordic journal of psychiatry*, *60*(5), 400-404.

Chu, J., Floyd, R., Diep, H., Pardo, S., Goldblum, P., & Bongar, B. (2013). A tool for the culturally competent assessment of suicide: The Cultural Assessment of Risk for Suicide (CARS) Measure. *Psychological assessment*, *25*(2), 424.

Cleary, A., Nixon, E., & Fitzgerald, M. (2007). Psychological health and well-being among young Irish adults. *Irish journal of psychological medicine*, *24*(4), 139-144.

Cukrowicz, K. C., Smith, P. N., Hohmeister, H. C., & Joiner, T. E. (2009). The moderation of an early intervention program for anxiety and depression by specific psychological symptoms. *Journal of clinical psychology*, *65*(4), 337-351.

DeVylder, J. E., Jahn, D. R., Doherty, T., Wilson, C. S., Wilcox, H. C., Schiffman, J., & Hilimire, M. R. (2015). Social and psychological contributions to the co-occurrence of sub-threshold psychotic experiences and suicidal behavior. *Social psychiatry and psychiatric epidemiology*, *50*(12), 1819-1830.

Dhingra, K., Boduszek, D., & Klonsky, E. D. (2016). Empirically derived subgroups of self-injurious thoughts and behavior: Application of latent class analysis. *Suicide and life-threatening behavior*, *46*(4), 486-499.

Innamorati, M., Pompili, M., Lester, D., Tatarelli, R., & Girardi, P. (2008). Recreational drug use and suicidality among Italian young adults. *Journal of addictive diseases*, *27*(4), 51-59.

Jonson, M., Skoog, I., Marlow, T., Fässberg, M. M., & Waern, M. (2012). Anxiety symptoms and suicidal feelings in a population sample of 70-year-olds without dementia. *International psychogeriatrics*, *24*(11), 1865-1871.

Lane, R., Cheref, S., & Miranda, R. (2016). Ethnic differences in suicidal ideation and its correlates among South Asian American emerging adults. *Asian American journal of psychology*, *7*(2), 120.

Moran, P., Coffey, C., Romaniuk, H., Olsson, C., Borschmann, R., Carlin, J. B., & Patton, G. C. (2012). The natural history of self-harm from adolescence to young adulthood: a population-based cohort study. *The Lancet*, *379*(9812), 236-243.

Mortier, P., Demyttenaere, K., Auerbach, R. P., Cuijpers, P., Green, J. G., Kiekens, G., ... & Bruffaerts, R. (2017a). First onset of suicidal thoughts and behaviours in college. *Journal of affective disorders*, *207*, 291-299.

Mortier, P., Kiekens, G., Auerbach, R. P., Cuijpers, P., Demyttenaere, K., Green, J. G., ... & Bruffaerts, R. (2017b). A risk algorithm for the persistence of suicidal thoughts and behaviors during college. *The Journal of clinical psychiatry*, *78*(7), e828.

Pereira-Morales, A. J., Adan, A., Camargo, A., & Forero, D. A. (2017). Substance use and suicide risk in a sample of young Colombian adults: An exploration of psychosocial factors. *The American journal on addictions*, *26*(4), 388-394.

Preti, A., Sheehan, D. V., Coric, V., Distinto, M., Pitanti, M., Vacca, I., ... & Petretto, D. R. (2013). Sheehan suicidality tracking scale (S-STS): reliability, convergent and discriminative validity in young Italian adults. *Comprehensive psychiatry*, *54*(7), 842-849.

Rancāns, E., Lapiņš, J., Renberg, E. S., & Jacobsson, L. (2003). Self-reported suicidal and help seeking behaviours in the general population in Latvia. *Social psychiatry and psychiatric epidemiology*, *38*(1), 18-26.

Renberg, E. S. (2001). Self-reported life-weariness, death-wishes, suicidal ideation, suicidal plans and suicide attempts in general population surveys in the north of Sweden 1986 and 1996. *Social Psychiatry and Psychiatric Epidemiology*, *36*(9), 429-436.

Van Orden, K. A., Cukrowicz, K. C., Witte, T. K., & Joiner Jr, T. E. (2012). Thwarted belongingness and perceived burdensomeness: Construct validity and psychometric properties of the Interpersonal Needs Questionnaire. *Psychological assessment*, *24*(1), 197.

van Spijker, B. A., Majo, M. C., Smit, F., van Straten, A., & Kerkhof, A. J. (2012). Reducing suicidal ideation: Cost-effectiveness analysis of a randomized controlled trial of unguided web-based self-help. *Journal of medical Internet research*, *14*(5).

Wagner, B., Klinitzke, G., Brähler, E., & Kersting, A. (2013). Extreme obesity is associated with suicidal behavior and suicide attempts in adults: results of a population-based representative sample. *Depression and anxiety*, *30*(10), 975-981.

Zhang, Y., Träskman-Bendz, L., Janelidze, S., Langenberg, P., Saleh, A., Constantine, N., ... & Postolache, T. T. (2012). Toxoplasma gondii immunoglobulin G antibodies and nonfatal suicidal self-directed violence. *The Journal of clinical psychiatry*, *73*(8), 1069-1076.

Zuromski, K. L., Cero, I., Witte, T. K., & Zeng, P. (2017). The quadratic relationship between body mass index and suicide ideation: a nonlinear analysis of indirect effects. *Suicide and life-threatening behavior*, *47*(2), 155-167.

Highlights

- Suicidality is highly prevalent in autism compared to the general population
- It is unknown whether there are validated tools to assess suicidality in autism
- Four tools are assessed for their appropriateness and measurement properties
- No suicidality assessment tool has been used or validated in an autistic population
- Recommendations are made to adapt currently available tools for autistic people

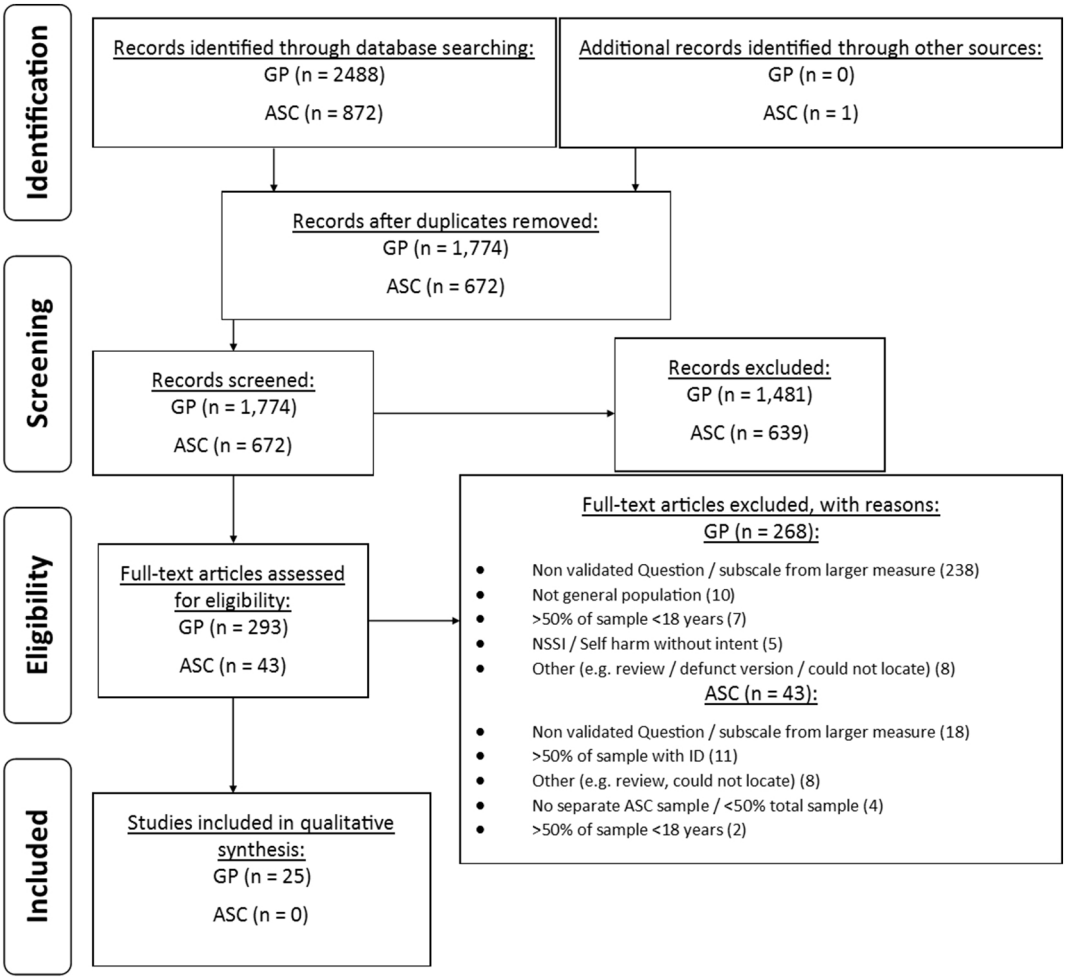
Identification
Screening
Eligibility
Included
Records identified through database searching:
GP (n = 2488)
ASC (n = 872)
Additional records identified through other sources:
GP (n = 0)
ASC (n = 1)
Records after duplicates removed:
GP (n = 1,774)
ASC (n = 672)
Records screened:
GP (n = 1,774)
ASC (n = 672)
Records excluded:
GP (n = 1,481)
ASC (n = 639)
Full-text articles assessed for eligibility:
GP (n = 293)
ASC (n = 43)
Full-text articles excluded, with reasons:
GP (n = 268):
Non validated Question / subscale from larger measure (238)
Not general population (10)
>50% of sample <18 years (7)
NSSI / Self harm without intent (5)
Other (e.g. review / defunct version / could not locate) (8)
ASC (n = 43):
Non validated Question / subscale from larger measure (18)
>50% of sample with ID (11)
Other (e.g. review, could not locate) (8)
No separate ASC sample / <50% total sample (4)
>50% of sample <18 years (2)
Studies included in qualitative synthesis:
GP (n = 25)
ASC (n = 0)

Figure 1

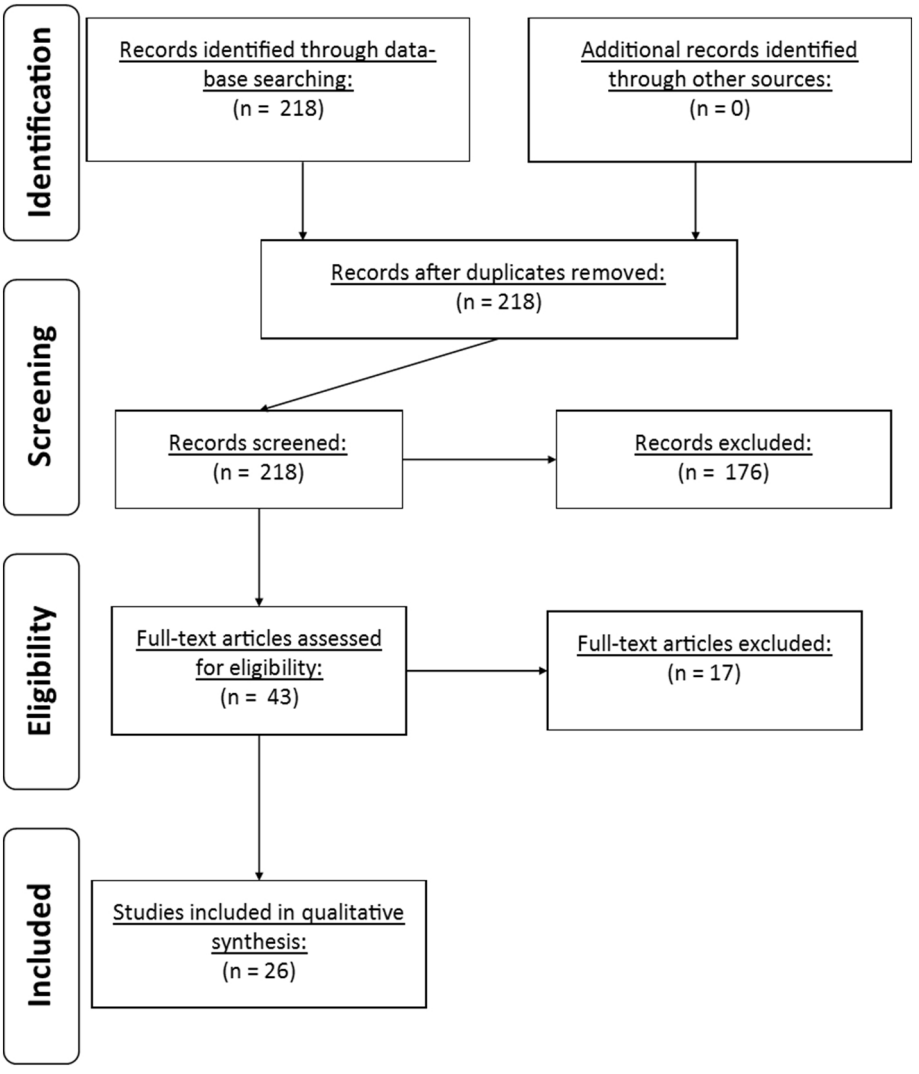
Identification
Records identified through data-base searching: (n = 218)
Additional records identified through other sources: (n = 0)
Records after duplicates removed: (n = 218)
Screening
Records screened: (n = 218)
Records excluded: (n = 176)
Eligibility
Full-text articles assessed for eligibility: (n = 43)
Full-text articles excluded: (n = 17)
Included
Studies included in qualitative synthesis: (n = 26)

Figure 2