Journal of
Child Health Care

Assessment tools of immediate risk of self-harm and suicide in children and young people: A scoping review

Journal:	*Journal of Child Health Care*
Manuscript ID	JCHC-2017-0288.R2
Manuscript Type:	Original Article
Keyword:	Scoping review, self-harm, suicide, risk assessment, children

Abstract:	There are increasing numbers of children presenting to paediatric hospital settings in mental health crisis. Typically, non-mental health professionals are responsible for the initial assessment of these children and are required to identify immediate physical and emotional health needs. To ensure the safety of these children, immediate risk of suicide and self-harm should be assessed. However, no standardised assessment tool is used in clinical practice, and for those tools that are used, their validity and reliability is unclear. A scoping review was conducted to identify existing assessment tools of immediate self-harm and suicide risk. Searches of electronic databases and relevant reference lists were undertaken. Twenty-two tools were identified and most assessed acute risk of suicide with only four tools incorporating a self-harm assessment. The tools varied in number of items (4-146), subscales (0-11), and total scores (16-192). Half incorporated Likert scales, and most were completed via self-report. Many tools were subject to limited psychometric testing and no single tool was valid or reliable for use with children presenting in mental health crisis to non-mental health settings. As such, a clinically appropriate, valid and reliable tool that assesses immediate risk of self-harm and suicide in paediatric settings should be developed.

Abstract

There are increasing numbers of children presenting to paediatric hospital settings in mental health crisis. Typically, non-mental health professionals are responsible for the initial assessment of these children and are required to identify immediate physical and emotional health needs. To ensure the safety of these children, immediate risk of suicide and self-harm should be assessed. However, no standardised assessment tool is used in clinical practice, and for those tools that are used, their validity and reliability is unclear. A scoping review was conducted to identify existing assessment tools of immediate self-harm and suicide risk. Searches of electronic databases and relevant reference lists were undertaken. Twenty-two tools were identified and most assessed acute risk of suicide with only four tools incorporating a self-harm assessment. The tools varied in number of items (4-146), subscales (0-11), and total scores (16-192). Half incorporated Likert scales, and most were completed via self-report. Many tools were subject to limited psychometric testing and no single tool was valid or reliable for use with children presenting in mental health crisis to non-mental health settings. As such, a clinically appropriate, valid and reliable tool that assesses immediate risk of self-harm and suicide in paediatric settings should be developed.

Keywords: scoping review, self-harm, suicide, risk assessment, children and young people

Introduction

Globally, the prevalence of mental health problems in children and young people (CYP) is increasing (Merikangas et al., 2009), with estimates of up to 10% of all CYP being clinically diagnosable (Green et al., 2005). These statistics represent a spectrum of conditions, requiring different levels of health care across primary and secondary care service settings.

At the acute end of this spectrum are CYP experiencing mental health crisis. This is defined as a psychiatric emergency involving 'an acute disruption of psychological homeostasis whereby usual coping mechanisms fail and distress and functional impairment' results (Lewis and Roberts, 2001). This may include: extreme anxiety or panic attacks; psychotic episodes (including delusions, hallucinations, paranoia or hearing voices); hypomania or mania; other behaviours that feel out of control; and acts of suicide or self-harm (Mind, 2013).

There are varying definitions of both self-harm and suicide in the literature. However, for the purpose of this article, self-harm has been defined as, an act with nonfatal outcome, in which an individual, irrespective of motivation, initiates a non-habitual behaviour that, without intervention from others, will cause self-harm, or deliberately ingests a substance in excess of the prescribed or generally recognized therapeutic dosage, and which is aimed at realising desired changes (Schmidtke et al, 1996; National Institute for Health and Clinical Excellence, 2013).

Internationally, emergency department (ED) attendance for self-harm and suicidal behaviour is high (Bethell et al., 2013; Rhodes et al., 2014), with literature indicating that CYP presenting to the ED due to self-harm are likely to present again within the first months following a previous presentation (Hulten et al., 2001; Bennardi et al., 2016).Suicide remains prevalent in CYP, with prevalence rates of 5.3 per 100,000 in 15-19 year olds in the United Kingdom (UK) (Office of National Statistics, 2016). Moreover, 145 suicides of young people under the age of 20 were reported in England between 2014-2015 (Rodway, 2016). Mental health crisis is the primary cause of approximately 5% of emergency department attendances (Parsonage et al., 2012) with the most prevalent presenting conditions being self-harm or suicide behaviours. Moreover, in CYP aged 10-19 years in England, suicide prevalence is 4.3 per 100,000, and self-harm is 435.95 per 100,000, with repeat ED attendance becoming increasingly commonplace (Hawton et al., 2012).

In the UK, for CYP presenting to hospital in mental health crisis, initial assessment is often undertaken by non-mental health professionals (paediatricians or children's nurses) within emergency department and paediatric ward settings (Anderson and Standen, 2007). This assessment aims to address immediate physical health needs (Olfson et al., 2005), and identify immediate risks to CYP's safety whilst they await expert assessment by specialist mental health professionals.

Evidence suggests that risk assessments are no more accurate at predicting risk than expert specialist mental health professional clinical judgement in non-acute psychiatric outpatients (Quinlivan et al., 2017). However, acute paediatric care settings present specific differences in utility, focus and context that make the application of an assessment of suicide and self-harm unique. For example, the assessment is usually made by non-mental health experts who may lack specialist knowledge and experience to inform clinical decisions (Crawford et al., 2003). Furthermore, the focus of these assessments is to assess any immediate (i.e. hours or days) risks of self-harm or suicide whilst in receipt of acute paediatric care. Additionally, assessments are performed in time limited circumstances with CYP with potentially dynamic and fluctuating mental health. Therefore to enable implementation of a plan of care where immediate risks can be mitigated, health care professionals require appropriate support and guidance to inform their assessment. In the UK, the National Institute for Health and Care Excellence (NICE) (2004) guidelines advocate CYP who self-harm should be assessed for risk. This assessment should identify the psychiatric illness and its relationship to self-harm, assess personal and social context and any specific factors predicting self-harm, and recognize any significant relationships that may be supportive or represent a threat. Such an assessment would need to consider the relatively immediate risk of self-harm or suicide in order to make time critical risk management decisions. Moreover, it would need to consider the developmental age of the CYP as children can often find verbal expression difficult, especially when in emotional distress (Vatne et al., 2010). Furthermore, the risk assessment should include assessment of previous ED presentations as this represents one of the strongest predictors of future ED repetitions across age and gender in young people (Hawton et al., 2005; Bennardi et al., 2016). Currently however, there is no standardised assessment tool utilised in clinical practice in the UK, and for those that are used; their validity, reliability and acceptability remain questionable.

Aim

There is need for a scoping review exploring the breadth and psychometric properties of existing risk assessment tools of immediate risk of self-harm and suicide in CYP. The aim of this review is to scope the literature for existing assessment tools of immediate risk of self-harm and suicide in CYP and synthesise their characteristics and psychometric properties.

Method

A scoping review method adhering to a published framework (Arksey and O'Malley, 2005) was employed to guide evidence identification, data charting, collating, summarizing and reporting. Scoping reviews offer a transparent and systematic approach to reviewing literature and are particularly useful in research areas with emerging evidence bases and where the research questions go beyond intervention effectiveness (Arkey and O'Malley, 2005; Levac et al, 2010).

This scoping review employed a sequential two-phased approach. Phase 1 identified the assessment tools from the published literature. Phase 2 identified the psychometric testing papers for each assessment tool. Phase 1 and 2 both involved searching four online databases (PubMed, MEDLINE, EMBASE, and PsychINFO) and reference lists of included papers.

Search strategy

Phase 1 searches were conducted in November 2016. Pre-defined search terms and Boolean phrasing were used to identify assessment tools of immediate risk of self-harm or suicide (see Table 1).

<<Insert Table 1 here>>

Phase 2 was conducted in May 2017. The assessment tool names identified from the Phase 1 search were used to search each online bibliographic database to identify the psychometric testing papers for each assessment tool.

For both Phases, the searches were saved and the references extracted into a reference management package (Mendeley™) for duplicate removal, followed by abstract and full text eligibility screening.

Eligibility criteria

Inclusion criteria: (1) an assessment, scale or measure that assesses immediate suicide/self-harm risk; (2) validity/reliability testing of the assessment with CYP (aged 1-18 years); (3) English language publication; (4) full text accessible; (5) peer reviewed journal publication.

Exclusion criteria: (1) validity/reliability tested in adults only; (2) reported only in books/commentaries; (3) assessment is a subscale only; (4) assessment is a structured

interview; (5) accessible as abstract only; (6) unpublished/grey literature; (7) the assessment tool is a screener of previous behaviour as opposed to an assessment tool of potential future behaviour.

Screening, data extraction and analysis

For products of search Phases 1 and 2, each abstract and full text were screened for eligibility by two reviewers independently. Following identification of eligible full texts, data was charted, collated and summarized using the approach outlined by Arksey and O'Malley (2005). This involved one researcher (GMW) extracting data pertaining to the characteristics (including: focus of assessment, number of items, target population, completion and response formats) and psychometric properties (specifically reliability and validity) of the assessment tools into a table with pre-defined headings to ensure standardization of included data. Two researchers then agreed suitability and checked for accuracy (TC, JCM). This charted data was then collated and narratively summarized in relation to the risk assessment tool characteristics, and then their psychometric properties.

Findings

Phase 1 searches revealed 22 eligible full text articles through which 26 risk assessment tools were identified. From these, 20 assessment tools met the eligibility criteria with reasons for exclusion shown in Figure 1. Phase 2 searches revealed 2 further assessment tools which met the eligibility criteria and were subsequently included in the review. The Phase 2 searches also identified 62 papers that tested the reliability and validity of the 22 assessment tools (See Figure 2).

<<Insert Figure 1 and Figure 2 here >>

Overview of risk assessment tool characteristics

<<Insert Table 2 here>>

Most assessment tools assessed immediate risk of suicide only (18/22; 81%), with the remainder (4/22; 18%) incorporating a limited number of self-harm questions (Angelkovska, 2014; Horowitz et al., 2001; Pfeffer, 1986; Reynolds, 1990). The completion format for most of the assessment tools was self-report (13/22; 59%) (Conrad et al., 2009; Cotton and Range, 1996; Cull and Gill, 1982; Horowitz et al., 2001, 2012; Miller et al., 1986; Osman et al.,

1998; Pfeffer et al., 2000; Plutchik et al., 1989; Range and Lewis, 1992a; Reynolds, 1987a, 1987b; Shaffer et al., 2004); with the remainder being clinician report (7/22; 32%) (Beck et al., 1974; Larzelere et al., 2004; Orbach et al., 1984, 1991; Pfeffer, 1986; Posner et al., 2011; Reynolds, 1990); parent-report (1/22; 4.5%) (Angelkovska, 2014); or included provision for self, parent or clinician report (1/22; 4.5%) (Flamarique et al., 2016).

The assessment tools varied in relation to the number of items/questions (range: 4-146); subscales (range: 0 -11); and maximum total score (range: 16-192), with less than half of the assessment tools not reporting total scores (9/22; 41%) (Angelkovska, 2014; Conrad et al., 2009; Cull and Gill, 1982; Horowitz et al., 2001; Larzelere et al., 2004; Plutchik et al., 1989; Range and Lewis, 1992a; Reynolds, 1990; Shaffer et al., 2004). The assessment tools varied in response format with a mixture of Likert only (11/22; 50%)(Beck et al., 1974; Cotton and Range, 1996; Cull and Gill, 1982; Flamarique et al., 2016; Miller et al., 1986; Orbach et al., 1984, 1991; Osman et al., 1998; Range and Lewis, 1992a; Reynolds, 1987b, 1987a); binary only (6/22; 27%)(Conrad et al., 2009; Horowitz et al., 2001, 2012; Larzelere et al., 2004; Pfeffer et al., 2000; Plutchik et al., 1989); mixed response (4/22; 27.5%)(Angelkovska, 2014; Pfeffer, 1986; Posner et al., 2011; Reynolds, 1990); and visual analogue scales (1/22; 4.5%)(Shaffer et al., 2004).

All included assessment tools were psychometrically tested in at least one subsequent testing paper. The Columbia Suicide Severity Rating Scale (Posner et al., 2011) was the most rigorously studied, with 11 subsequent psychometric testing papers (Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013; Flamarique et al., 2016; Horwitz et al., 2015; Kerr et al., 2014; King et al., 2015; Knafo A et al., 2015; Mirkovic et al., 2015; Posner et al., 2011).

Overview of psychometric testing

<<Link to supplementary files Table 3 and Table 4 here>>

Psychometric testing across the assessment tools was undertaken on mixed ethnicities and populations aged 5 to 19 years. It was also undertaken across various settings, including: inpatient hospitals (14/22; 63.6%)(Cotton and Range, 1996; Eltz et al., 2007; Fennig et al., 2005; Ferrara et al., 2012; Grilo et al., 1999; Gutierrez et al., 2000; Knafo A et al., 2015; Koutek et al., 2016; Mcnicholas, 2011; Mieczkowski et al., 1993; Mirkovic et al., 2015; Morano et al., 1993; Ofek et al., 1998; Orbach et al., 1984, 1991, Osman et al., 1994, 2000;

Pettit et al., 2009; Pfeffer et al., 2000; Posner et al., 2011; Range and Lewis, 1992a, 1992b; Romanowicz et al., 2013; Schwartz-Stav et al., 2006; Shaunesey et al., 1993; Spirito et al., 1987, 1996); schools (13/22; 59.1%)(Allison et al., 1995; Angelkovska, 2014; Cotton and Range, 1996; Davis, 1992; Jia et al., 2015; Labelle et al., 2015; Lee, 2011; Mazza, 2000; Mazza and Reynolds, 1999; Miranda et al., 2014; Orbach et al., 1984, 1991, Osman et al., 1994, 1998; Pfeffer et al., 2000; Range and Lewis, 1992a, 1992b; Reynolds, 1990; Reynolds and Mazza, 2001; S.-C. et al., 2008; Shaffer et al., 2004; Wong, 2004); universities (1/22; 4.5%)(Osman et al., 1993); outpatient departments (9/22; 40.9%)(Angelkovska, 2014; Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013; Flamarique et al., 2016; King et al., 1997, 2014; Labelle et al., 2015; Orbach et al., 1984; Range and Lewis, 1992a, 1992b; Rosenberg et al., 2006; Storch et al., 2014); emergency departments (4/22; 18.2%)(Horowitz et al., 2001, 2012; Horwitz et al., 2015; King et al., 2015; Stanley et al., 2013); non-hospital community settings (3/22; 13.6%)(Angelkovska et al., 2012; Gutierrez, 1999; Kerr et al., 2014; Zhang et al., 2014); detention centres (1/22; 4.5%)(Stathis et al., 2008); foster care settings (1/22; 4.5%)(Larzelere et al., 2004); and residential facility/home settings (3/22; 13.6%)(Badura Brack et al., 2012; Larzelere et al., 2004; Larzelere et al., 1996).

Most psychometric testing papers were undertaken in English speaking populations in the United States (US) (45/62; 72.6%). Several were tested in non-English language translation, as follows: Hebrew (4/22; 18.2%)(Fennig et al., 2005; Ofek et al., 1998; Orbach et al., 1991; Schwartz-Stav et al., 2006); Chinese (3/22; 13.6%)(Jia et al., 2015; Wong, 2004; Zhang et al., 2014); Korean (2/22; 9.0%)(Lee, 2011; S.-C. et al., 2008); French (4/22; 18.2%)(Flamarique et al., 2016; Knafo A et al., 2015; Labelle et al., 2015; Mirkovic et al., 2015), German (1/22; 4.5%)(Flamarique et al., 2016); Dutch(1/22; 4.5%)(Flamarique et al., 2016); Italian (2/22; 9.0%)(Ferrara et al., 2012; Flamarique et al., 2016); and Spanish (1/22; 4.5%)(Flamarique et al., 2016).

Face validity

Face validity was tested with varying degree of rigour for five (22.7%) assessment tools (Flamarique et al., 2016; Larzelere et al., 2004; Mieczkowski et al., 1993; Pfeffer et al., 2000; Range and Lewis, 1992a, 1992b). The Suicide Intent Scale (Mieczkowski et al., 1993) reported face validity without description of method or outcome. The Life Orientation Inventory items were reviewed by psychologists and previously suicidal individuals (Range

and Lewis, 1992a, 1992b). The Child Suicide Risk Assessment (Larzelere et al., 2004) items were reviewed by children and it was found that 85% of the items were understood well. The Suicidality Treatment Occurring Paediatrics- Suicidality Assessment Scale (Flamarique et al., 2016) reported child feedback of item comprehension and problems differentiating items, consequently a process of re-wording, sentence shortening and children's suggested examples were incorporated into the scale. The Child-Adolescent Suicidal Potential Index (Pfeffer et al., 2000) was reviewed by psychiatric professionals leading to revision of instructions, items and response formats. Children's suggested changes to wording comprehension were also implemented.

Predictive validity

Predictive validity was tested for 19 assessment tools (86.4%).. Methods of predictive validity were: firstly, assessment score correlations with actual events (such as past, present or future suicide/self-harm thoughts or behaviours) (14/22; 63.6%) (Eltz et al., 2007; Fennig et al., 2005; Ferrara et al., 2012; Flamarique et al., 2016; Gutierrez et al., 2000; Horwitz et al., 2015; Kerr et al., 2014; King et al., 2014, 2015; Koutek et al., 2016; Larzelere et al., 2004; Larzelere et al., 1996; Mieczkowski et al., 1993; Miranda et al., 2014; Osman et al., 2000; Pfeffer et al., 2000; Posner et al., 2011; Reynolds, 1990; Zhang et al., 2014); secondly, sensitivity and specificity (13/22; 59.1%) (Flamarique et al., 2016; Gutierrez et al., 2000; Horowitz et al., 2001, 2012; King et al., 2015; Larzelere et al., 2004; Larzelere et al., 1996; Osman et al., 1994, 2000; Pfeffer et al., 2000; Posner et al., 2011; Shaffer et al., 2004; Stathis et al., 2008; Zhang et al., 2014); and thirdly, the proportion of positive and negative findings that were true positive and true negative results, i.e. Positive Predictive Value (PPV) and Negative Predictive Value (NPV) (8/22; 36.4%) (Horowitz et al., 2001, 2012; Koutek et al., 2016; Larzelere et al., 2004; Larzelere et al., 1996; Shaffer et al., 2004; Zhang et al., 2014). The Columbia Suicide Severity Rating Scale (Posner et al., 2011) had predictive validity most rigorously tested (four psychometric testing papers) (Horwitz et al., 2015; Kerr et al., 2014; King et al., 2015; Posner et al., 2011).

Eight assessment tools (36.4%) consistently predicted suicide/self-harm events (Eltz et al., 2007; Flamarique et al., 2016; Grilo et al., 1999; Gutierrez et al., 2000; Koutek et al., 2016; Larzelere et al., 1996; Miranda et al., 2014; Pfeffer et al., 2000; Reynolds, 1990; Zhang et al., 2014); five (22.7%) predicted suicide/self-harm variably (Fennig et al., 2005; Ferrara et al., 2012; Horwitz et al., 2015; Kerr et al., 2014; King et al., 2014, 2015; Larzelere et al.,

2004; Osman et al., 2000; Posner et al., 2011; Zhang et al., 2014); and one (4.5%) did not predict suicide/self-harm (Mieczkowski et al., 1993).

Sensitivity and Specificity testing across the studies revealed substantial variability suggesting that although these scales were able to identify those at risk they were also likely to classify some individuals' as at risk when they were not. The Columbia Suicide Severity Rating Scale, the Suicide Ideation Questionnaire-Junior and the Suicide Ideation Questionnaire had the highest Sensitivity ratings suggesting they are the most likely to be able to identify those at risk of engaging in suicidal or self-harming behaviour.

Total item PPVs were performed for 6/22 (27.2%) assessment tools (Cull and Gill, 1982; Horowitz et al., 2012; Larzelere et al., 2004; Reynolds, 1987a, 1987b; Shaffer et al., 2004) and NPVs were performed for 3/22 (13.6%) assessment tools (Horowitz et al., 2012; Larzelere et al., 2004; Shaffer et al., 2004). Total item PPVs was variable across studies (range: 8.8-71.3%) (Horowitz et al., 2012; Larzelere et al., 2004; Larzelere et al., 1996; Shaffer et al., 2004; Zhang et al., 2014) as was total item NPVs (range: 13.6-99.7%) (Horowitz et al., 2012; Larzelere et al., 2004; Shaffer et al., 2004). The Suicide Probability Scale (Cull and Gill, 1982) had the lowest PPV (Larzelere et al., 1996) and the Child Suicide Risk Assessment (Larzelere et al., 2004) had the lowest NPV. The Adolescent Suicide Questionnaire (Horowitz et al., 2012) had the highest PPV and NPV.

Convergent validity

Convergent validity, i.e. the degree to which two measures should theoretically correlate, was tested for 19 (86.4%) assessment tools, all of which tested total-item convergent validity. Subscale convergent validity was tested in 10/22 (45.5%) of the assessment tools (Beck et al., 1974; Cotton and Range, 1996; Cull and Gill, 1982; Horowitz et al., 2001; Larzelere et al., 2004; Orbach et al., 1991; Osman et al., 1998; Pfeffer, 1986; Pfeffer et al., 2000; Posner et al., 2011). Correlations between assessments and construct measures were variable. Five assessment tools (22.7%) failed to demonstrate significant correlations between all subscales and construct measures (Cotton and Range, 1996; Ferrara et al., 2012; Gutierrez, 1999; Mieczkowski et al., 1993; Ofek et al., 1998; Orbach et al., 1991; Osman et al., 1994, 2000; Rosenberg et al., 2006; Spirito et al., 1996). Furthermore, four assessment tools failed to correlate total-item scores with some construct measures (18.2%) (Grilo et al., 1999; Pettit et al., 2009; Rosenberg et al., 2006; Storch et al., 2014). The Multi-Attitude Suicide Tendency Scale had convergent validity most rigorously tested in six psychometric testing papers

(Ferrara et al., 2012; Gutierrez, 1999; Orbach et al., 1991; Osman et al., 1994, 2000; Wong, 2004).

Discriminant validity (between groups)

Discriminant validity was tested for 20 (90.1%) assessment tools, of which 16 (72.7%) tested total item subscale discriminant validity (Angelkovska, 2014; Beck et al., 1974; Conrad et al., 2009; Cotton and Range, 1996; Cull and Gill, 1982; Horowitz et al., 2012; Larzelere et al., 2004; Orbach et al., 1991; Osman et al., 1998; Pfeffer et al., 2000; Plutchik et al., 1989; Posner et al., 2011; Range and Lewis, 1992a; Reynolds, 1987a, 1987b; Reynolds, 1990), and nine (40.9%) tested subscale discriminant validity (Conrad et al., 2009; Cull and Gill, 1982; Miller et al., 1986; Orbach et al., 1984, 1991; Osman et al., 1998; Pfeffer et al., 2000; Posner et al., 2011; Reynolds, 1990). Numerous demographic and characteristic domains were also tested, and some assessment tools were consistently able to discriminate between: age (2/22; 9.0%) (Pfeffer et al., 2000; Reynolds, 1990); gender (2/22; 9.0%) (Horwitz et al., 2015; Pfeffer et al., 2000); psychiatric diagnosis (6/22; 27.3%) (Knafo A et al., 2015; Mazza, 2000; Range and Lewis, 1992a, 1992b; Schwartz-Stav et al., 2006; Spirito et al., 1987, 1996); suicide and self-harm status, (10/22; 45.5%) (Grilo et al., 1999; Gutierrez et al., 2000; Horwitz et al., 2015; King et al., 2015; Larzelere et al., 2004; Lee, 2011; Morano et al., 1993; Osman et al., 1998; Range and Lewis, 1992a, 1992b; Reynolds, 1990; Romanowicz et al., 2013; Shaunesey et al., 1993; Spirito et al., 1987, 1996); physical illness status (2/22; 9.0%)(Angelkovska, 2014; Spirito et al., 1996); accidental injury (1/22; 4.5%) (Rosenberg et al., 2006); and family history of suicide (1/22; 4.5%) (Romanowicz et al., 2013). Some assessment tools consistently failed to discriminate for age (3/22; 13.6%) (Romanowicz et al., 2013; Spirito et al., 1987, 1996; Zhang et al., 2014); gender (3/22; 13.6%) (Allison et al., 1995; Grilo et al., 1999; Spirito et al., 1996); psychiatric diagnosis (2/22; 9.0%) (Grilo et al., 1999; Rosenberg et al., 2006); and history of abuse (1/22; 4.5%)(Grilo et al., 1999). The Columbia Suicide Severity Rating Scale(Posner et al., 2011) was most rigorously tested for discriminant validity (eight psychometric testing papers) (Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013; Horwitz et al., 2015; King et al., 2015; Knafo A et al., 2015; Mirkovic et al., 2015).

Internal consistency

Internal consistency was tested for 17/22 (77.3%) assessment tools (Angelkovska, 2014; Beck et al., 1974; Cull and Gill, 1982; Flamarique et al., 2016; Horowitz et al., 2012;

Larzelere et al., 2004; Miller et al., 1986; Orbach et al., 1991; Osman et al., 1998; Pfeffer et al., 2000; Pfeffer, 1986; Plutchik et al., 1989; Posner et al., 2011; Range and Lewis, 1992a; Reynolds, 1987a, 1987b; Reynolds, 1990). Total-item internal consistency (range: α=0.60-0.99) was higher and less variable overall than subscale internal consistency (range: 0.38 to 0.95). Therefore, when taken as a whole, the scales demonstrate better internal consistency and less fluctuation than when exploring between subscales. The Suicidal Ideation Questionnaire- Junior Version (Reynolds, 1987a) achieved the highest internal consistency (r=0.99) (Gutierrez, 1999). The Multi-Attitude Suicide Tendency Scale (Orbach et al., 1991) was most rigorously tested for internal consistency (five psychometric testing papers) (Gutierrez, 1999; Orbach et al., 1991; Osman et al., 1994, 2000; Wong, 2004).

Inter-rater reliability

Inter-rater reliability was tested for 4/22 (18.2%) assessment tools (Flamarique et al., 2016; Pfeffer, 1986; Posner et al., 2011; Reynolds, 1990). These assessment tools were subjected to total item inter-rater reliability (2/22; 9.0%) (Flamarique et al., 2016; Reynolds, 1990) and subscale inter-rater reliability (2/22; 9.0%) (Fennig et al., 2005; Kerr et al., 2014; Ofek et al., 1998). Total item inter-rater analyses revealed variable correlations (range: 0.47 to 0.99) as did the Subscale inter-rater analyses (range: 0.40 to 0.97). The Suicide Behaviour Interview had the highest inter-rater reliability (ICC= 0.99) (Reynolds, 1990) and the Suicidality Treatment Occuring Paediatrics- Suicidality Assessment Scale had the lowest (r=0.47)(Flamarique et al., 2016). The Child-Adolescent Suicide Potential Scale (Pfeffer, 1986) had inter-rater reliability most rigorously tested (two psychometric testing papers) (Fennig et al., 2005; Ofek et al., 1998).

Test re-test reliability

Test re-test reliability was tested for 7/22 (31.8%) assessment tools, demonstrating variable reliability (range: r=0.32 to 0.92) (Cull and Gill, 1982; Orbach et al., 1984; Pfeffer, 1986; Pfeffer et al., 2000; Range and Lewis, 1992a; Reynolds, 1987a; Shaffer et al., 2004). Three (14%) assessment tools reported subscale test-re-test reliability, with less variability (r=0.39-0.78), suggesting these scales have some ability to remain consistent over time (Ofek et al., 1998; Orbach et al., 1984; Pfeffer et al., 2000). The Suicide Probability Scale and Life Orientation Inventory (r= 0.92) (Larzelere et al., 1996; Range and Lewis, 1992a, 1992b) had the highest test re-test reliability and the Columbia Suicide Screen had the lowest (r=0.32)

(Shaffer et al., 2004). 1/22 (4.5%) The Fairy Tales Test assessment tool (Orbach et al., 1984) failed to achieve test re-test reliability for all questions.

DISCUSSION

The assessment tools included in this review varied in length, response and scoring format, age ranges and degree of psychometric testing. Most assessments were tested across broad age ranges, and may be criticised as lacking developmental sensitivity. The SIQ and the SIQ-JR however were exceptions, having undergone age- based revisions/adaptations. Some measures of suicide risk incorporated risk items relating to self-harm. No measure assessed risk of self-harm in isolation. Most assessment tools were tested only in the USA and primarily with inpatients, in contrast to cross-cultural psychometric guidelines (Beaton et al., 2000). Few papers reported language translations and none reported cultural adaptations. Most assessment tools were originally developed in the English language, but few reported psychometric testing in UK populations, suggesting limited applicability in acute paediatric settings in this region. As such, it is understandable that UK guidelines do not promote the use of any one assessment tool to safely manage immediate risk of self-harm or suicide to inform clinical decisions in acute paediatric settings (Horowitz et al., 2014).

Across the included tools, internal consistency and test-re-test reliability was generally moderate to good, suggesting that many are constructed of items that are likely to measure the same construct (i.e. risk of suicide) and that the tools are able to produce similar scores when tested over a number of time points, respectively. Test re-test reliability was, however, variable across many of the studies and may be due to suicide/self-harm risk being sensitive to change.

Only four assessments (Flamarique et al., 2016; Pfeffer, 1986; Posner et al., 2011; Reynolds, 1990) investigated inter-rater reliability, thus we have little evidence that the current assessment tools provide consistent results across different raters. Moreover, for those tools for which this testing was undertaken, it appears the majority were tested with raters (i.e. clinician, self, parent) with limited scientific or clinical justification.

Although face validity is considered the weakest validity test (Devon et al., 2007), it is typically considered a pre-requisite before performing other validity/reliability tests (Devon et al., 2007). However, few studies tested it, and those that had, lacked strong methodological report, thus reducing the potential usefulness for the tools, and limiting the ability to replicate procedures (Schulz et al., 2010). Moreover, there appears to be limited consideration to the developmental issues within the tools included in this review. As such,

considering the substantial differences in cognitive ability, perception and understanding between younger children and those closer to 18 years of age, the current tools appear unable to provide accurate representation of potential risk for CYP across the age range.

This review highlights that the majority of previous assessment tools of immediate risk of self-harm and/or suicide have not been tested to levels recommended by psychometric guidelines (Devon et al., 2007). Moreover, several of these assessment tools demonstrate inconsistent validity and reliability ratings across different testing studies. Additionally, cut off values denoting high risk scores are sparsely defined thus limiting their clinical utility, as such values can be a useful adjunct to suicde risk assessment in non-psychiatric emergency settings (Cochrane-Brink et al, 2000). Several assessment tools were only tested in one subsequent psychometric testing paper, highlighting limited testing across the majority of the assessment tools. An exception is the Columbia Suicide Severity Rating Scale which generally performed well across multiple psychometric domains and has been used to monitor medication safety in clinical practice (Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013).

The findings from this scoping review stem from an extensive, transparent search of the literature and provide a summary of the characteristics, and ratings of reliability and validity of assessments tools of immediate self and suicide risk in CYP. This is a scoping review however, and as such it cannot be concluded with certainty that additional risk assessment tools have not been developed and psychometrically tested. Moreover, use of only the terms 'self-harm' and 'deliberate self-harm' (and not the other close terms) in the search strategy represents a potential limitation, as other additional studies may have identified the behaviour using alternative terminology.

However, the review has identified key gaps and deficits including limited immediate self-harm risk assessment tools for CYP, limited psychometric testing of current assessment tools in specific contexts and regions, and no one assessment tool having been fully validated in an inpatient paediatric setting.

Thus there are clear implications for clinical practice as currently there appears to be no suicide/self-harm assessment tool validated for use in inpatient paediatric settings whereby immediate risk (i.e. within hours of the assessment) need to be taken place. As a result, health care professionals working within paediatric inpatient settings have to resort to using

their own clinical judgement (which may be based on minimal experience and training) or a risk assessment framework/tool that has not been developed for the specific needs of this population/setting. Consequently, this may lead to an inaccurate assessment of risk potentially resulting in either over or under estimation of risk rating, and subsequent inappropriate safety management strategies being utilised.

Considering the increasing prevalence of mental health problems in children and young people, and the paucity in existing risk assessments outlined here, future research should be focused on the development of a clinically appropriate, psychometrically tested assessment tool of immediate risk of self-harm and suicide behaviour for children and young people. This assessment tool could then be used to support safety management decisions across acute paediatric care settings.

Contributorship

TC and JCM designed the scoping review protocol. GMW wrote the initial draft of the manuscript. TC, AA and JCM revised, edited and finalised the manuscript. TC, JCM and GMW conducted the database searches, confirmed included papers, and agreed data extraction. All authors agreed the final manuscript for publication.

Declaration of Conflicting interests

The authors declared no potential conflicts of interests in consideration this research publication.

Funding

This review was completed through funding provided by Nottinghamshire Clinical Commissioning Group and Nottingham University Hospitals NHS Trust sponsorship. The views represented are the views of the authors alone and do not necessarily represent the views of the Nottinghamshire Clinical Commissioning Group or Nottingham University Hospitals NHS Trust.

References

Allison S, Pearce C, Martin G, et al. (1995) Parental influence, pessimism and adolescent suicidality. *Archives of Suicide Research* 1(4): 229–242.

Anderson M and Standen PJ (2007) Attitudes towards suicide among nurses and doctors working with children and young people who self-harm. *Journal of Psychiatric and Mental Health Nursing* 14(5): 470–477.

Angelkovska A (2014) The Development and validation of an instrument to identify risk of self-harm in children. Available from: http://research-repository.uwa.edu.au/files/3236726/Angelkovska_Agni_2007.pdf.

Angelkovska A, Houghton S and Hopkins S (2012) Differential profiles of risk of self-harm among clinically referred primary school aged children. *School Psychology International* 33(6): 646–660.

Arksey H and O'Malley L (2005) Scoping studies: Framework, Towards a methodological. *International of Social Research Methodology* 8(1): 19–32.

Atkinson SD, Prakash A, Zhang Q, et al. (2014) A Double-Blind Efficacy and Safety Study of Duloxetine Flexible Dosing in Children and Adolescents with Major Depressive Disorder. *Journal of Child and Adolescent Psychopharmacology* 24(4): 180–189.

Badura Brack A, Huefner JC, Handwerk ML, et al. (2012) The impact of abuse and gender on psychopathology, behavioral disturbance, and psychotropic medication count for youth in residential treatment. *American Journal of Orthopsychiatry* 82(4): 562–572.

Beaton DE, Bombardier C, Guillemin F, et al. (2000) Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine* 25(24): 3186–3191.

Beck AT, Schuyler D and Herman I (1974) Development of suicidal intent scales. In: Beck A, Resnik H, and Lettieri D (eds), *The Prediction of Suicide*, Oxford, England: Charles Press Publishers, pp. 45–56.

Bennardi, M., McMahon, E., Corcoran, P. et al. (2016) Risk of repeated self-harm and associated factors in children, adolescents and young adults. *BMC psychiatry*, 16(1), p.421-433

Bethell J, Bondy SJ, Wendy Lou WY, et al. (2013) Emergency department presentations for self-harm among Ontario youth. *Canadian Journal of Public Health*, Canada 104(2): e124-30.

Centers for Disease Control and Prevention (2017) Definitions: Self-directed Violence. *Centers for Disease Control and Prevention*.

Cicchetti D V. (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* 6(4): 284–290.

Cochrane-Brink, KA., Lofchy, JS. and Sakinofsky, I. (2000) Clinical rating scales in suicide risk assessment11Information about the rating scales can be obtained from the author. *General Hospital Psychiatry*. 22(6): 445-451.

Conrad AK, Jacoby AM, Jobes DA, et al. (2009) A psychometric investigation of the Suicide Status Form II with a psychiatric inpatient sample. *Suicide & life-threatening behavior* 39(3): 307–20.

Cotton CR and Range LM (1996) Suicidality, hopelessness, and attitudes toward life and death in clinical and nonclinical adolescents. *Death Studies*, Taylor & Francis Group 20(6): 601–610.

Cotton CR, Peters DK and Range LM (1995) Psychometric properties of the suicidal behaviors questionaire. *Death Studies* 19(4): 391–397.

Crawford T, Geraghty W, Street K, et al. (2003) Staff knowledge and attitudes towards deliberate self-harm in adolescents. *Journal of Adolescence* 26(5): 619–629.

Cull JG and Gill WS (1982) Suicide Probability Scale. Western Psychological Services, Los Angeles, CA, Los Angeles, CA.: WPS.

Davis JM (1992) Suicidal Ideation Questionnaire. Journal of Psychoeducational Assessment, Odessa, FL: Psychological Assessment Resources.

Devon HA, Block ME, Moyle-Wright P, et al. (2007) A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship* 39(2): 155–164.

Eltz M, Evans AS, Celio M, et al. (2007) Suicide probability scale and its utility with adolescent psychiatric patients. *Child Psychiatry and Human Development* 38(1): 17–29.

Emslie GJ, Prakash A, Zhang Q, et al. (2014) A Double-Blind Efficacy and Safety Study of Duloxetine Fixed Doses in Children and Adolescents with Major Depressive Disorder. *Journal of Child and Adolescent Psychopharmacology* 24(4): 170–179.

Emslie GJ, Wells TG, Prakash A, et al. (2015) Acute and Longer-Term Safety Results from a Pooled Analysis of Duloxetine Studies for the Treatment of Children and Adolescents with Major Depressive Disorder. *Journal of Child and Adolescent Psychopharmacology* 25(4): 293–305.

Fennig S, Geva K, Zalsman G, et al. (2005) Effect of gender on suicide attempters versus nonattempters in an adolescent inpatient unit. *Comprehensive Psychiatry* 46(2): 90–97.

Ferrara M, Terrinoni A and Williams R (2012) Non-suicidal self-injury (Nssi) in adolescent inpatients: assessing personality features and attitude toward death. *Child and Adolescent Psychiatry and Mental Health* 6(1): 12.

Findling RL, Robb A and Bose A (2013) Escitalopram in the Treatment of Adolescent Depression: A Randomized, Double-Blind, Placebo-Controlled Extension Trial. *Journal of Child and Adolescent Psychopharmacology* 23(7): 468–480.

Flamarique I, Santosh P, Zuddas A, et al. (2016) Development and psychometric properties of the Suicidality: Treatment Occurring in Paediatrics (STOP) Suicidality Assessment Scale (STOP-SAS) in children and adolescents. *BMC Pediatrics* 16(1): 213.

Green H, McGinnity Á, Meltzer H, et al. (2005) Mental health of children and young people in Great Britain. Available from: http://www.esds.ac.uk/doc/5269/mrdoc/pdf/5269technicalreport.pdf.

Grilo CM, Sanislow C, Fehon DC, et al. (1999) Psychological and behavioral functioning in adolescent psychiatric inpatients who report histories of childhood abuse. *Am J Psychiatry* 156(4): 538–543.

Gutierrez PM (1999) Suicidality in parentally bereaved adolescents. *Death Studies* 23(4): 359–370.

Gutierrez PM, Osman A, Kopper BA, et al. (2000) Why Young People Do Not Kill Themselves: The Reasons for Living Inventory for Adolescents. *Journal of Clinical Child Psychology* 29(2): 177–187.

Hawton, K. and James, A., 2005. ABC of adolescence: Suicide and deliberate self harm in young people. *British Medical Journal,* 330(7496), p.891-894.

Hawton K, Bergen H, Kapur N, et al. (2012) Repetition of self-harm and suicide following self-harm in children and adolescents: Findings from the Multicentre Study of Self-harm in England. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 53(12): 1212–1219.

Horowitz LM, Wang PS, Koocher GP, et al. (2001) Detecting Suicide Risk in a Pediatric Emergency Department: Development of a Brief Screening Tool. *Pediatrics* 107(5): 1133–1137.

Horowitz LM, Bridge JA, Teach SJ, et al. (2012) Ask Suicide-Screening Questions (ASQ). *Archives of Pediatrics & Adolescent Medicine*, American Medical Association 166(12): 1170.

Horowitz LM, Bridge JA, Pao M, et al. (2014) Screening youth for suicide risk in medical settings: Time to ask questions. *American Journal of Preventive Medicine* 47(3 SUPPL. 2): S170–S175.

Horwitz AG, Czyz EK and King CA (2015) Predicting Future Suicide Attempts Among Adolescent and Emerging Adult Psychiatric Emergency Patients. *Journal of Clinical Child and Adolescent Psychology* 44(5): 751–761.

Hulten, A., Jiang, G.X., Wasserman, D., et al. (2001) Repetition of attempted suicide among teenagers in Europe: frequency, timing and risk factors. *European Child & Adolescent Psychiatry*, 10(3), pp.161-169.

Jia C-X, Wang X-T and Zhao Z-T (2015) Psychometric Properties of the Suicidal Ideation Questionnaire among Chinese High School Students. *OMEGA - Journal of Death and Dying* 70(2): 195–207.

Kerr DCR, Gibson B, Leve LD, et al. (2014) Young adult follow-up of adolescent girls in juvenile justice using the Columbia Suicide Severity Rating Scale. *Suicide and Life-Threatening Behavior* 44(2): 113–129.

King CA, Katz SH, Ghaziuddin N, et al. (1997) Diagnosis and assessment of depression and suicidality using de NIMH Diagnosis Interview Schedule for Children (DISC 2.3). *Journal of Abnormal Child Psychology* 29: 173–181.

King CA, Jiang Q, Czyz EK, et al. (2014) Suicidal ideation of psychiatrically hospitalized adolescents has one-year predictive validity for suicide attempts in girls only. *Journal of Abnormal Child Psychology* 42(3): 467–477.

King CA, Berona J, Czyz E, et al. (2015) Identifying Adolescents at Highly Elevated Risk for Suicidal Behavior in the Emergency Department. *Journal of Child and Adolescent Psychopharmacology* 25(2): 100–108.

Knafo A, Guile JM, Breton JJ, et al. (2015) Coping strategies associated with suicidal. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie* 60(2 Suppl 1): S46–S54.

Koutek J, Kocourkova J and Dudova I (2016) Suicidal behavior and self-harm in girls with eating disorders. *Neuropsychiatric Disease and Treatment* 12: 787–793.

Labelle R, Breton J-J, Berthiaume C, et al. (2015) Psychometric properties of three measures of protective factors for depression and suicidal behaviour among adolescents. *Canadian journal of psychiatry. Revue canadienne de psychiatrie* 60(2 Suppl 1): S16-26.

Larzelere RE, Andersen JJ, Ringle JL, et al. (2004) The child suicide risk assessment: A screening measure of suicide risk in pre-adolescents. *Death Studies* 28(9): 809–827.

Larzelere RE, Smith GL, Batenhorst LM, et al. (1996) Predictive Validity of the Suicide Probability Scale among Adolescents in Group Home Treatment. *Journal of the American Academy of Child & Adolescent Psychiatry* 35(2): 166–172.

Lee SY (2011) Reasons for living and their moderating effects on korean adolescents' suicidal ideation. *Death Studies* 35(8): 711–728.

Levac, D., Colquhoun, H. and O'Brien, K. K. (2010) Scoping studies: advancing the methodology. *Implementation Science.* 5(1):69.

Lewis S and Roberts AR (2001) Crisis assessment tools: The good, the bad, and the available. *Brief Treatment and Crisis Intervention* 1: 17–28.

Mazza JJ (2000) The relationship between posttraumatic stress symptomatology and suicidal behavior in school-based adolescents. *Suicide & life-threatening behavior* 30(2): 91–103.

Mazza JJ and Reynolds WM (1999) Exposure to violence in young inner-city adolescents: Relationships with suicidal ideation, depression, and PTSD symptomatology. *Journal of Abnormal Child Psychology* 27(3): 203–213.

Mcnicholas LF (2011) Poster Abstracts from the AAAP 21st Annual Meeting and Symposium. *American Journal on Addictions* 20(4): 391–392.

Merikangas KR, Nakamura EF and Kessler RC (2009) Epidemiology of mental disorders in children and adolescents. *Dialogues.Clin.Neurosci.* 11(1): 7–20.

Mieczkowski TA, Sweeney JA, Haas GL, et al. (1993) Factor Composition of the Suicide Intent Scale. *Suicide and Life-Threatening Behavior* 23(1): 37–45.

Miller IW, Norman WH, Bishop SB, et al. (1986) The Modified Scale for Suicidal Ideation: Reliability and validity. *Journal of Consulting and Clinical Psychology*, American Psychological Association 54(5): 724–725.

Mind (2013) What is a mental health crisis? *Mental Health Crises*. Available from: https://www.mind.org.uk/information-support/guides-to-support-and-services/crisis-services/#.WUj6-a-Wzcs.

Miranda R, Ortin A, Scott M, et al. (2014) Characteristics of suicidal ideation that predict the transition to future suicide attempts in adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 55(11): 1288–1296.

Mirkovic B., Labelle R., Guilé J-M., et al. (2015) Coping skills among adolescent suicide attempters: results of a multisite study. *Canadian journal of psychiatry. Revue canadienne de psychiatrie* 60(2): S37–S45.

Morano C, Cisler R and Lemerond J (1993) Risk Factors for Adolescent Suicidal Behavior: Loss, Insufficient Familial Support, and Hopelessness Morano, Christopher D;Cisler, Ron

A;Lemerond, John Adolescence; Winter 1993; 28, 112; ProQuest pg. 851. *Adolescence* 28(112): 851.

National Institute for Health and Clinical Excellence (2004) The short-term physical and psychological management and secondary prevention of self-harm in primary and secondary care. Manchester.

National Institute for Health and Clinical Excellence (2013) *Self Harm: Quality Standard 34.* London, UK.

Ofek H, Weizman T and Apter A (1998) The child suicide potential scale: Inter-rater reliability and validity in israeli in-patient adolescents. *Israel Journal of Psychiatry and Related Sciences* 35(4): 253–261.

Office of National Statistics (ONS) (2016). Suicides in the UK: 2016 registrations. London: HMSO. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/suicidesintheunitedkingdom/2016registrations

Olfson M, Gameroff MJ, Marcus SC, et al. (2005) Emergency Treatment of Young People Following Deliberate Self-harm. *Archives of General Psychiatry* 62(10): 1122.

Orbach I, Feshbach S, Carlson GA, et al. (1984) Attitudes toward life and death in suicidal, normal, and chronically ill children: An extended replication. *Journal of Consulting and Clinical Psychology*, American Psychological Association 52(6): 1020–1027.

Orbach I, Milstein I, Har-Even D, et al. (1991) A Multi-Attitude Suicide Tendency Scale for adolescents. *Psychological Assessment*, American Psychological Association 3(3): 398–404.

Osman A, Barrios FX, Grittmann LR, et al. (1993) The multi-attitude suicide tendency scale: Psychometric characteristics in an american sample. *Journal of Clinical Psychology* 49(5): 701–708.

Osman A, Barrios FX, Panak WF, et al. (1994) Validation of the multi???attitude suicide tendency scale in adolescent samples. *Journal of Clinical Psychology* 50(6): 847–855.

Osman A, Kopper BA and Linehan MM (1998) The Reasons for Living Inventory for Adolescents (RFL-A): Development and psychometric properties. *Journal of Clinical Psychology* 54(8): 1063–1078.

Osman a, Gilpin a R, Panak WF, et al. (2000) The multi-attitude suicide tendency scale: further validation with adolescent psychiatric inpatients. *Suicide & life-threatening behavior* 30(4): 377–85.

Parsonage M, Fossey M and Tutty C (2012) *Liaison psychiatry in the modern NHS*. Available from: http://www.centreformentalhealth.org.uk/publications/liaison_psychiatry.aspx?ID=665.

Pettit JW, Garza MJ, Grover KE, et al. (2009) Factor structure and psychometric properties of the modified scale for suicidal ideation among suicidal youth. *Depression and Anxiety* 26(8): 769–774.

Pfeffer CR (1986) Child suicide potential scales. In: *The suicidal child.*, New York.: Guilford Press., pp. 173–203.

Pfeffer CR, Zuckerman S and Plutchik R (1979) Suicidal behavior in latency age children: An empirical study. *Journal of the American Academy of Child Psychiatry* 19(1): 679–692.

Pfeffer CR, Plutchik R, Mizruchi MS, et al. (1986) Suicidal behavior in child psychiatric inpatients and outpatients and in nonpatients. *American Journal of Psychiatry*, 143(6): 733–738.

Pfeffer CR, Jiang H and Kakuma T (2000) Child–Adolescent Suicidal Potential Index (CASPI): A screen for risk for early onset suicidal behavior. *Psychological Assessment*, American Psychological Association 12(3): 304–318.

Plutchik R, Van Praag HM, Conte HR, et al. (1989) Correlates of suicide and violence risk: I. The suicide risk measure. *Comprehensive Psychiatry* 30(4): 296–302.

Posner K, Brown G and Stanley B (2011) Columbia-Suicide everity Rating Scale: initial validity and internal consistency finding from three multisite studies with adolescents and adults. *Am. J. Psychiatry* 168(December): 1266–1277.

Quinlivan L, Cooper J, Meehan D, et al. (2017) Predictive accuracy of risk scales following self-harm: Multicentre, prospective cohort study. *British Journal of Psychiatry* 210(6): 429–436.

Range LM and Lewis LS (1992a) Life Orientation Inventory: A method of assessing suicide risk. Pro-Ed. Journal of Psychoeducational Assessment, Austin, Texas.

Range LM and Lewis LS (1992b) Life Orientation Inventory: A method of assessing suicide risk. *Journal of Psychoeducational Assessment*, 10(3): 296–297.

Reynolds W and Mazza J (2001) Assessment of suicidal ideation in inner-city children and young adolescents: reliability and validity of the Suicidal Ideation Questionnaire-JR. *School Psychology Review* 30: 376–384.

Reynolds WM (1987a) *Suicidal ideation questionnaire-junior*. *Psychological Assessment Resources*, Odessa, FL: Psychological Assessment Resources.

Reynolds WM (1987b) *Suicide Ideation Questionnaire.* Odessa, FL: Psychological Assessment Resources.

Reynolds William M. (1990) Development of a semistructured clinical interview for suicidal behaviors in adolescents. *Psychological Assessment*, American Psychological Association 2(4): 382–390.

Reynolds W M (1990) *Suicide Behaviors Interview*. Odessa, FL: Psychological Assessment Resources.

Rhodes AE, Bethell J, Carlisle C, et al. (2014) Time trends in suicide-related behaviours in girls and boys. *Canadian Journal of Psychiatry*, United States 59(3): 152–159.

Rodway, C., Tham, S. G., Ibrahim, S., Turnbull, P., Windfuhr, K., Shaw, J., Kapur, N. and Appleby, L. (2016) Suicide in children and young people in England: a consecutive case series. *Lancet Psychiatry.* 3(8):.751-759.

Romanowicz M, O'Connor SS, Schak KM, et al. (2013) Use of the Suicide Status Form-II to investigate correlates of suicide risk factors in psychiatrically hospitalized children and adolescents. *Journal of Affective Disorders* 151(2): 467–473.

Rosenberg L, Robert R, Thomas C, et al. (2006) Assessing potential suicide risk at young adults burned as children. *Journal of Burn Care & Research* 27(6): 779–785.

Cho, S.C., Kim, J.W., Choi, H.J., et al. (2008) Associations between symptoms of attention deficit hyperactivity disorder, depression, and suicide in Korean female adolescents. *Depression and Anxiety* 25(11): E142–E146.

Schmidtke, A., Bille-Brahe, U., DeLeo, D., et al. (1996) Attempted suicide in Europe: rates, trends and sociodemographic characteristics of suicide attempters during the period 1989-1992. Results of the WHO/EURO Multicentre Study on Parasuicide. *Acta Psychiatr Scand* 93: 327-338.

Schulz KF, Altman DG and Moher D (2010) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Ann Int Med* 152(1): 18.

Schwartz-Stav O, Apter A and Zalsman G (2006) Depression, suicidal behavior and insight in adolescents with schizophrenia. *European Child and Adolescent Psychiatry* 15(6): 352–359.

Shaffer D, Wilcox H and Lucas CP (1996) The development of a screening instrument for teens at risk for suicide. In: *Meeting of the Academy of Child and Adolescent Psychiatry*, New York.

Shaffer D, Scott M, Wilcox H, et al. (2004) The Columbia SuicideScreen: Validity and Reliability of a Screen for Youth Suicide and Depression. *Journal of the American Academy of Child & Adolescent Psychiatry* 43(1): 71–79.

Shaunesey K, Cohen JL, Plummer B, et al. (1993) Suicidality in hospitalized adolescents: relationship to prior abuse. *Am J Orthopsychiatry* 63(1): 113–119.

Spirito A, Stark L and Fristad M (1987) Adolescent suicide attempters hospitalized on a pediatric unit. *Journal of Pediatric Psychology* 12(2): 171–189.

Spirito A, Sterling CM, Donaldson DL, et al. (1996) Factor Analysis of the Suicide Intent Scale With Adolescent Suicide Attempters. *Journal of Personality Assessment* 67(1): 90–101.

Stanley IH, Snyder D, Westen S, et al. (2013) Self-Reported Recent Life Stressors and Risk of Suicide in Pediatric Emergency Department Patients. *Clinical pediatric emergency medicine* 14(1): 35–40.

Stathis S, Litchfield B, Letters P, et al. (2008) A comparative assessment of suicide risk for young people in youth detention. *Archives of suicide research : official journal of the International Academy for Suicide Research* 12(1): 62–6.

Storch EA, Bussing R, Jacob ML, et al. (2014) Frequency and Correlates of Suicidal Ideation in Pediatric Obsessive–Compulsive Disorder. *Child Psychiatry and Human Development* 46(1): 75–83.

Vatne TM, Finset A, Ornes K, et al. (2010) Application of the Verona Coding Definitions of Emotional Sequences (VR-CoDES) on a pediatric data set. *Patient Education and Counseling* 80(3): 399–404.

World Health Organisation (2014) Preventing suicide: A global imperative. Geneva: World Health Organisation

Wong WS (2004) Attitudes toward life and death among chinese adolescents: The chinese version of the multi-attitude suicide tendency scale. *Death Studies* 28(2): 91–110.

Zhang Y, Yip PS and Fu K (2014) Validation of the Chinese version of the Reynolds' suicidal ideation questionnaire: psychometric properties and its short version. *Health and Quality of Life Outcomes* 12(1): 33.

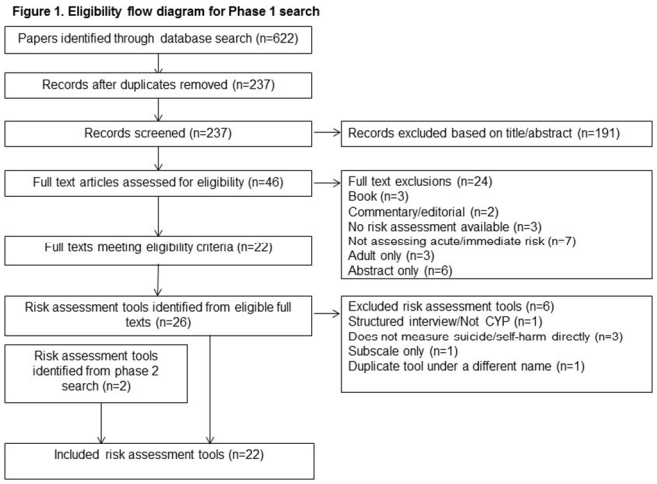

Figure 1: Eligibility flow diagram for Phase 1 search

338x190mm (96 x 96 DPI)

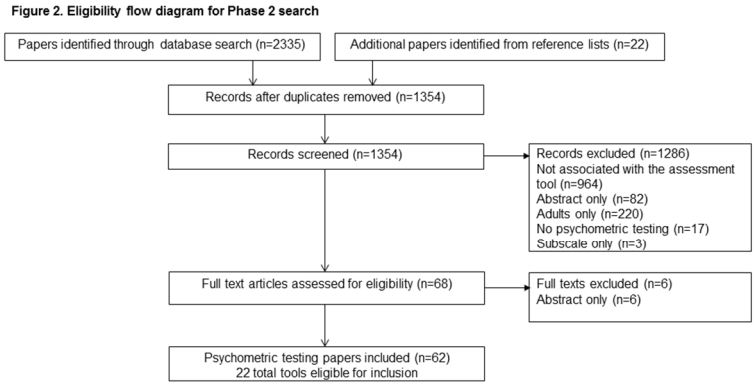

Figure 2: Eligibility flow diagram for Phase 2 search

338x190mm (96 x 96 DPI)

Population:	Adolescen* OR Young people OR Child* OR Teenagers		
Intervention:	Suicid* OR suicidal behaviour OR suicide attempt OR suicidal ideation OR self harm* OR self-harm* OR deliberate self-harm	AND	Risk OR Assessment OR Screening

Table 1: Phase 1 search terms

Table 2: Assessment tool/scale characteristics

Scale/tool	Focus of assessment	No. of items/ subscales	Population tested (age; gender; ethnicity; diagnosis; setting; country)	Completion format	Response format
Adolescent Suicide Questionnaire (ASQ)(Horowitz et al., 2012)	Risk of suicidal behaviour	4 items	10-21 years olds; male/females; mixed ethnicity; psychiatric, medical/surgical(Horowitz et al., 2012), young offenders(Stathis et al., 2008), Students(Allison et al., 1995); emergency department(Horowitz et al., 2012), detention centre(Stathis et al., 2008), school(Allison et al., 1995); USA(Horowitz et al., 2012), Australia(Allison et al., 1995; Stathis et al., 2008)	Self-report	Binary Yes/No/No response (missing data). Scores: 1-0 Total score: 0-4 Cut off: score of 5
Child-Adolescent Suicide Potential Scale (CSPS)(Pfeffer, 1986)	Risk of suicidal behaviour	146 items 11 subscales	5-12 year olds; mixed ethnicity; psychiatric(Fennig et al., 2005; Ofek et al., 1998; Schwartz-Stav et al., 2006); inpatients(Fennig et al., 2005; Ofek et al., 1998; Schwartz-Stav et al., 2006); Israel(Fennig et al., 2005; Ofek et al., 1998; Schwartz-Stav et al., 2006)	Clinician-rated	Semi-structured interview Binary: Yes/No (Scores: 1-0) 5 point likert scale (Scores: 1-5) 3 point likert scale (Scores: 1-3)
Columbia Suicide Screen (CSS)(Shaffer et al., 2004)	Risk of suicidal behaviour and ideation	11 items	11-18 year olds; male/female; mixed ethnicity; suicidal(Miranda et al., 2014; Shaffer et al., 2004) and non-suicidal(Shaffer et al., 2004); Schools(Miranda et al., 2014; Shaffer et al., 2004); USA(Miranda et al., 2014; Shaffer et al., 2004)	Self-report	5 point visual analog scale Scores 1-5 Total score: 1-55
Columbia Suicide Severity Rating Scale (C-SSRS)(Posner et al., 2011)	Severity and intensity of suicide risk	19 items 4 Subscales	11-17 year olds; males/females; mixed ethnicity; suicide attempters(Knafo A et al., 2015; Mirkovic et al., 2015; Posner et al., 2011), delinquents(Kerr et al., 2014), psychiatric(Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013; Flamarique et al., 2016; Horwitz et al., 2015; King et al., 2015; Knafo A et al., 2015; Posner et al., 2011), medicated(Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013; Posner et al., 2011), non-psychiatric(King et al., 2015);	Clinician-rated	Semi-structured clinical interview Ordinal scales Binary yes/no scales 3 subscales (score: 0-5) 1 subscale (score: 0-6) 1 subscale (score 0-2); 3 open ended Total score: 0-44.

			outpatient(Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013; Flamarique et al., 2016), inpatient(Knafo A et al., 2015; Mirkovic et al., 2015; Posner et al., 2011); emergency department(Horwitz et al., 2015; King et al., 2015), community(Kerr et al., 2014); Spain(Flamarique et al., 2016), UK(Flamarique et al., 2016), Italy(Flamarique et al., 2016), Germany(Atkinson et al., 2014; Emslie et al., 2015; Flamarique et al., 2016), Netherlands(Flamarique et al., 2016), USA(Atkinson et al., 2014; Emslie et al., 2014, 2015; Findling et al., 2013; Horwitz et al., 2015; Kerr et al., 2014; King et al., 2015; Posner et al., 2011), South Africa(Atkinson et al., 2014; Emslie et al., 2015), Mexico(Emslie et al., 2014, 2015), Canada(Emslie et al., 2014, 2015), Argentina(Emslie et al., 2014, 2015), Estonia(Atkinson et al., 2014; Emslie et al., 2015), Finland(Atkinson et al., 2014; Emslie et al., 2015), Russia(Atkinson et al., 2014; Emslie et al., 2015), Slovakia(Atkinson et al., 2014; Emslie et al., 2015), Ukraine(Atkinson et al., 2014; Emslie et al., 2015), France(Atkinson et al., 2014; Emslie et al., 2015; Flamarique et al., 2016; Knafo A et al., 2015; Mirkovic et al., 2015).		
Fairy Tales Test (FT)(Orbach et al., 1984)	Intensities of Attitudes towards Life and Death	4 items	6-12 year olds; males/females; mixed ethnicity; suicidal(Orbach et al., 1984), chronically ill(Orbach et al., 1984), no health issues (Orbach et al., 1984); outpatients(Orbach et al., 1984), inpatients(Orbach et al., 1984), school(Orbach et al., 1984); USA(Orbach et al., 1984)	Clinician-rated	4 point likert scale Scores 0-3 with a half-point interval Total score: 0-12
Life Orientation Inventory (LOI)(Range and Lewis, 1992a)	Intentions of suicide	113 items 6 subscales	13-18 years old; mixed ethnicity; Low proportion Native Americans(Range and Lewis, 1992a, 1992b), 6th grade reading level(Range and Lewis, 1992a, 1992b), normal(Range and Lewis, 1992a, 1992b), students(Range and Lewis, 1992a, 1992b); outpatients(Range and Lewis,	Self-report	4 point likert scale

			1992a, 1992b), psychiatric inpatients(Range and Lewis, 1992a, 1992b), schools(Range and Lewis, 1992a, 1992b); USA(Range and Lewis, 1992a, 1992b)		
Multi-Attitude Suicide Tendency Scale (MAST)(Orbach et al., 1991)	Suicidal tendencies and conflicting attitudes related to suicidality	30 items 4 Subscales	11-18 years old; males/females; mixed ethnicity; students(Orbach et al., 1991; Osman et al., 1993, 1994; Wong, 2004), non-referred(Osman et al., 1994), suicidal(Orbach et al., 1991), non suicidal(Osman et al., 1993; Wong, 2004), suicide ideators(Osman et al., 1993; Wong, 2004), suicide attempters(Wong, 2004), non-suicidal self injurers(Ferrara et al., 2012), psychiatric(Ferrara et al., 2012; Orbach et al., 1991; Osman et al., 1994, 2000); normative(Orbach et al., 1991), parentally bereaved(Gutierrez, 1999); inpatients(Ferrara et al., 2012; Orbach et al., 1991; Osman et al., 1994, 2000); schools(Orbach et al., 1991; Osman et al., 1994; Wong, 2004), university(Osman et al., 1993), community(Gutierrez, 1999); USA(Gutierrez, 1999; Osman et al., 1993, 1994, 2000); Israel(Orbach et al., 1991); China(Wong, 2004), Italy(Ferrara et al., 2012)	Clinician-rated	5 point likert scale Scored 1-5 Total score: 1-150
Reasons for living Inventory for adolescents (RFL-A)(Osman et al., 1998)	Adaptive reasons for not ending own life.	32 items 5 Subscales	13-19 year olds; males/females; mixed ethnicity; students(Labelle et al., 2015; Lee, 2011; Osman et al., 1998), psychiatric(Gutierrez et al., 2000; Labelle et al., 2015; Osman et al., 1998), suicide attempters(Gutierrez et al., 2000; Osman et al., 1998), non-suicidal(Gutierrez et al., 2000; Osman et al., 1998); inpatients(Gutierrez et al., 2000), outpatients(Labelle et al., 2015), school(Labelle et al., 2015; Lee, 2011; Osman et al., 1998), college(Osman et al., 1998); USA(Gutierrez et al., 2000; Osman et al., 1998), Korea(Lee, 2011), French(Labelle et al., 2015).	Self-report	6 point likert scale Scores 1-6 Total score: 1-192 Cut-off: 4.63
Risk of Suicide Questionnaire 14 item (RSQ)(Horowitz	Suicide risk	14 items	11 to 16 year olds; males/females; mixed ethnicity; psychiatric(Horowitz et al., 2001); emergency department(Horowitz et al., 2001); USA(Horowitz et al., 2001)	Self-report	Binary: Yes/No or No response (missing data).

et al., 2001)					
Modified Scale for Suicide Ideation (M-SSI)(Miller et al., 1986)	Suicide risk	18 items 3 Subscales	13-18 year olds, male/female, mixed ethnicity; suicide attempters, suicide ideators, psychiatric; inpatients; USA(Pettit et al., 2009)	Self-report	4 point likert scale Scores: 0-3 Total score: 0-54.
Self-Harm Risk Assessment for Children (SHRAC)(Angelkovska, 2014)	Self-harm risk	33 items	6-12 year olds; males/females; mixed ethnicity; behavioural/learning problems(Angelkovska et al., 2012), non-psychiatric(Angelkovska et al., 2012), non-clinical(Angelkovska et al., 2012), students(Angelkovska, 2014; Angelkovska et al., 2012); externalizers(Angelkovska et al., 2012), Internalizers(Angelkovska et al., 2012), Combined internalizers/ externalizers(Angelkovska et al., 2012), self-harmers(Angelkovska, 2014), self-harm ideators(Angelkovska, 2014), clinic referred(Angelkovska, 2014), non-referred community comparisons(Angelkovska, 2014); outpatient(Angelkovska, 2014), community(Angelkovska et al., 2012), schools(Angelkovska, 2014); Australia(Angelkovska, 2014; Angelkovska et al., 2012)	Parent-rated	4 point likert scale Scores: 1-4 Binary: Yes/No
Suicidal Behaviours Questionnaire for Children (SBQ-C)(Cotton and Range, 1996)	Suicide risk	4 items	15-18 years old; males/females; mixed ethnicity; students, psychiatric; school, inpatients; USA(Cotton and Range, 1996)	Self-report	2 point, 4 point and 6 point likert scales. Total score: 0-16
Suicidal Ideation Questionnaire – Junior Version (SIQ-JR)(Reynolds, 1987a)	Frequency and severity of suicidal ideation	15 items	11-18 year olds; males/females; mixed ethnicity; students(Mazza, 2000; Mazza and Reynolds, 1999; William M. Reynolds, 1990; Reynolds and Mazza, 2001), psychiatric(King et al., 1997, 2014; Mazza, 2000; Storch et al., 2014), suicide ideators(King et al., 2014), parentally bereaved(Gutierrez, 1999); inpatients(King et al.,	Self-report	7 point likert scale continuum Scores: 0-6 Total score: 0-90 Cut off: 31

			1997, 2014), school(Mazza, 2000; Mazza and Reynolds, 1999; William M. Reynolds, 1990; Reynolds and Mazza, 2001), outpatients(Storch et al., 2014), community(Gutierrez, 1999; Zhang et al., 2014); USA(Gutierrez, 1999; King et al., 1997, 2014; Mazza, 2000; Mazza and Reynolds, 1999; William M. Reynolds, 1990; Reynolds and Mazza, 2001; Storch et al., 2014), China(Zhang et al., 2014)		
Suicidal Ideation Questionnaire (SIQ)(Reynolds, 1987b)	Frequency and severity of suicidal ideation	30 items	13-19 years; males/females; mixed ethnicity; students(Davis, 1992; Jia et al., 2015; S.-C. et al., 2008), suicide attempters(Shaunesey et al., 1993; Spirito et al., 1987)(Spirito et al., 1996), suicide ideators(Shaunesey et al., 1993), non-suicidal controls(Shaunesey et al., 1993) medical/surgical(Spirito et al., 1987; Stanley et al., 2013), psychiatric(Shaunesey et al., 1993; Stanley et al., 2013); Inpatient(Shaunesey et al., 1993; Spirito et al., 1987, 1996), emegency department(Stanley et al., 2013), school(Davis, 1992; Jia et al., 2015; S.-C. et al., 2008), community(Zhang et al., 2014); UK, USA(Davis, 1992; Shaunesey et al., 1993; Spirito et al., 1987, 1996; Stanley et al., 2013), China(Jia et al., 2015; Zhang et al., 2014), Korea(S.-C. et al., 2008)	Self-report	7 point likert scale continuum Scores of 0-6 Total score 0-180. Cut off: 41
Suicide Behaviour Interview (SBI)(William M. Reynolds, 1990)	Suicide risk	22 items 2 Subscales	12-19 year olds; males/females; mixed ethnicity; psychiatric; inpatients; USA(Mieczkowski et al., 1993)	Clinician-rated	Semi-structured interview 0-2 or 0-4 point scale Scored to half a point (0.5). 2 open ended questions
Suicide Intent Scale (SIS)(Beck et al., 1974)	Suicide intent	20 items 2 Subscales	Children/Adolescents (ages not defined); males/females; mixed ethnicity; psychiatric(Mieczkowski et al., 1993; Morano et al., 1993; Spirito et al., 1996), non-suicide attempters(Morano et al., 1993), suicide attempters(Mieczkowski et al., 1993; Morano et al., 1993; Spirito et al., 1996); Inpatient(Mieczkowski et al., 1993; Morano et	Clinician-rated	3 item Likert scale Scores: 1-3 Total score: 1-60.

			al., 1993; Spirito et al., 1996); USA(Mieczkowski et al., 1993; Morano et al., 1993; Spirito et al., 1996)		
Suicide Probability Scale (SPS) (Cull and Gill, 1982)	Suicide risk	36 items 4 subscales	>14 years old; males/females; mixed ethnicity; psychiatric(Eltz et al., 2007; Rosenberg et al., 2006), suicide attempters(Larzelere et al., 1996), suicide non-attempters(Larzelere et al., 1996), burn survivors(Rosenberg et al., 2006), physical abuse(Badura Brack et al., 2012), sexual abuse(Badura Brack et al., 2012), physical and sexual abuse(Badura Brack et al., 2012), no abuse(Badura Brack et al., 2012), medicated(Badura Brack et al., 2012); inpatient(Eltz et al., 2007), outpatients(Rosenberg et al., 2006), residential facility(Badura Brack et al., 2012; Larzelere et al., 1996); USA(Badura Brack et al., 2012; Eltz et al., 2007; Larzelere et al., 1996; Rosenberg et al., 2006)	Self-report	4 point likert scale
Suicide Risk Scale (SRS)(Plutchik et al., 1989)	Suicide risk	15 items 4 Subscales	12 to 18 year olds; male/female; mixed ethnicity; psychiatric, abused, Non-abused, depressed; inpatients; USA(Grilo et al., 1999)	Self-report	Binary: True/False
Child-Adolescent Suicidal Potential Index (CASPI)(Pfeffer et al., 2000)	Suicide risk	30 items 3 Subscales	6-17 year olds; males/females; mixed ethnicity; psychiatric(Koutek et al., 2016; Pfeffer et al., 2000), students(Pfeffer et al., 2000), eating disorders(Koutek et al., 2016), suicide attempters(Pfeffer et al., 2000), suicide ideators(Pfeffer et al., 2000), assaulters/ideators(Pfeffer et al., 2000); inpatients(Koutek et al., 2016; Pfeffer et al., 2000); schools(Pfeffer et al., 2000); USA(Pfeffer et al., 2000), Czech Republic(Koutek et al., 2016)	Self-report	Binary: Yes/No Scores: 1-0 Total score: 0-30 Cut off: 11
Child Suicide Risk Assessment (CSRA)(Larzelere et al., 2004)	Suicide risk	20 items 3 Subscales	Age 6-12 year olds; females/males; mixed ethnicity; suicide attempters and non-attempters(Larzelere et al., 2004); residential facility(Larzelere et al., 2004), foster	Clinician-rated	Binary: Yes/No Cut-off score: 8

			care(Larzelere et al., 2004); USA(Larzelere et al., 2004)		
Suicide Status Form-II (SSF-II)(Conrad et al., 2009)	Suicide risk and frequency of suicidal ideation	6 items	8-18 year olds; males/females; mixed ethnicity; psychiatric(Mcnicholas, 2011; Romanowicz et al., 2013), suicide attempters(Romanowicz et al., 2013), suicide ideators(Romanowicz et al., 2013); inpatient(Mcnicholas, 2011; Romanowicz et al., 2013); USA(Mcnicholas, 2011; Romanowicz et al., 2013)	Self-report	5 point likert scale Scores 1-5 2 Binary Yes/No
Suicidality Treatment Occuring Paediatrics-Suicidality Assessment Scale (STOP-SAS; Flamarique et al., 2016)(Flamarique et al., 2016)	Suicide risk	14 items-Children 19 items - adolescent, parents, clinician	8-18 year olds; males/females; mixed ethinicity; psychiatric(Flamarique et al., 2016), medicated(Flamarique et al., 2016); outpatients(Flamarique et al., 2016); Spain(Flamarique et al., 2016), UK(Flamarique et al., 2016), Italy(Flamarique et al., 2016), France(Flamarique et al., 2016), Germany(Flamarique et al., 2016), Netherlands(Flamarique et al., 2016)	Self-report; parent- report; clinician - report	6 point likert scale Adolescents, Parents, Clinicians: 0-5 scores Total score: 0-95 4 point likert scale Children: 0-3 scores Total score: 0-42

Table 3: Overview of reliability testing

Tool/scale		Reliability type					
		Internal Consistency		Test re-test		Inter-rater	
		Undertaken	α- Score	Undertaken	Effect Size	Undertaken	Effect Size
Adolescent Suicide Questionnaire (ASQ) Horowitz et al (2012)(Allison et al., 1995; Horowitz et al., 2012; Stathis et al., 2008)	**Item Total**	✓	Good	×		×	
	Subscale	×		×		×	
Child-Adolescent Suicide Potential Scale (CSPS) – (Pfeffer et al., 1979; 1986; 2000)(Fennig et al., 2005; Ofek et al., 1998; Pfeffer, 1986; Pfeffer et al., 1979, 1986; Schwartz-Stav et al., 2006)	**Item Total**	✓	Excellent	×		×	
	Subscale	✓	Unacceptable-Acceptable	✓	5/11 subscales: Medium- Large*	✓	7/11 subscales: Medium-.Large*
Columbia Suicide Screen (CSS) - Shaffer et al., 1996, 2004(Miranda et al., 2014; Shaffer et al., 1996, 2004)	**Item Total**	×		✓	Medium	×	
	Subscale	×		×		×	
Columbia Suicide Severity Rating Scale (C-SSRS)- Posner et al, 2011(Atkinson et al., 2014; Emslie et al., 2015; Findling et al., 2013; Flamarique et al., 2016; Horwitz et al., 2015; Kerr et al., 2014; King et al., 2015; Knafo A et al., 2015; Mirkovic et al., 2015; Posner et al., 2011)	**Item Total**	×		×		✓	Small-Large
	Subscale	✓	1/11 subscales: Acceptable*	×		×	
Fairy Tales Test (FT; Orbach et al., 1984)(Orbach et al., 1984)	**Item Total**	×		×		×	
	Subscale	×		✓	Medium-Large*	×	
Life Orientation Inventory (LOI; Kowalchuk & King 1988)(Range and Lewis, 1992a, 1992b)	**Item Total**	✓	Excellent	✓	Large	×	
	Subscale	✓	Poor-Good	×		×	
Multi-Attitude Suicide Tendency Scale (MAST; Orbach et al., 1991)(Ferrara et al., 2012; Gutierrez, 1999; Orbach et al., 1991; Osman et al., 1993, 1994, 2000; Wong, 2004)	**Item Total**	✓	Questionable	×		×	
	Subscale	✓	Unacceptable-Excellent	×		×	
Reasons for Living Inventory for adolescents (RFL-A) - Osman et al., 1998(Gutierrez et al., 2000; Labelle et al., 2015; Lee, 2011; Osman et al., 1998)	**Item Total**	✓	Excellent	×		×	
	Subscale	✓	Good- Excellent	×		×	
Risk of Suicide Questionnaire 14 item (RSQ; Horowitz, et al., 2001)(Horowitz et al., 2001)	**Item Total**	×		×		×	
	Subscale	×		×		×	
Modified Scale for Suicide Ideation (M-SSI) Miller et al	**Item Total**	✓	Good	×		×	

(1986)(Miller et al., 1986; Pettit et al., 2009)	**Subscale**	×		×		×	
Self-Harm Risk Assessment for Children (SHRAC) - Angelkovska, 2008 (Angelkovska, 2014; Angelkovska et al., 2012)	**Item Total**	×		×		×	
	Subscale	✓	Excellent	×		×	
Suicidal Behaviours Questionnaire for Children (SBQ-C; Cotton and Range, 1993)(Cotton et al., 1995; Cotton and Range, 1996)	**Item Total**	×		×		×	
	Subscale	×		×		×	
Suicidal Ideation Questionnaire – Junior Version (SIQ-JR; Reynolds and Mazza, 1999)(Gutierrez, 1999; King et al., 1997, 2014, 2015; Mazza, 2000; Mazza and Reynolds, 1999; Reynolds, 1987a; William M. Reynolds, 1990; Reynolds and Mazza, 2001; Storch et al., 2014; Zhang et al., 2014)	**Item Total**	✓	Good- Excellent	✓	Large	×	
	Subscale	✓	Poor-Good	×		×	
Suicidal Ideation Questionnaire (SIQ; Reynolds, 1987,1988)(Davis, 1992; Horowitz et al., 2001; Jia et al., 2015; Reynolds, 1987b; S.-C. et al., 2008; Shaunesey et al., 1993; Spirito et al., 1987, 1996; Stanley et al., 2013; Zhang et al., 2014)	**Item Total**	✓	Excellent	×		×	
	Subscale	×		×		×	
Suicide Behaviour Interview (SBI) Reynolds (1990)(W M Reynolds, 1990; William M. Reynolds, 1990)	**Item Total**	✓	Questionable	×		✓	Large
	Subscale	×		×		×	
Suicide Intent Scale (SIS) Beck , Schuyler, & Herman (1974)(Beck et al., 1974; Mieczkowski et al., 1993; Morano et al., 1993; Spirito et al., 1996)	**Item Total**	✓	Good	×		×	
	Subscale	✓	Acceptable-Excellent	×		×	
Suicide Probability Scale (SPS) Cull and Gill (1882/1988) (Badura Brack et al., 2012; Cull and Gill, 1982; Eltz et al., 2007; Larzelere et al., 1996; Rosenberg et al., 2006)	**Item Total**	✓	Excellent	✓	Large	×	
	Subscale	×		×		×	
Suicide Risk Scale Plutchik R et al (1989)(Grilo et al., 1999; Plutchik et al., 1989)	**Item Total**	✓	Acceptable	×		×	
	Subscale	×		×		×	
Child-Adolescent Suicidal Potential Index (CASPI; Pfeffer, Jiang, & Kakuma, 2000)(Koutek et al., 2016; Pfeffer et al., 2000)	**Item Total**	✓	Excellent	✓	Large	×	
	Subscale	✓	Acceptable -Good	✓	Medium-Large	×	
Child Suicide Risk Assessment (CSRA) (Larzelere et al, 2004)(Larzelere et al., 2004)	**Item Total**	✓	Questionable	×		×	
	Subscale	✓	Unacceptable - Acceptable	×		×	
Suicide Status Form-II (SSF-II)	**Item Total**	×		×		×	

Conrad et al (2009)(Conrad et al., 2009; Mcnicholas, 2011; Romanowicz et al., 2013)	**Subscale**	×		×		×	
Suicidality Treatment Occuring Paediatrics- Suicidality Assessment Scale (STOP-SAS) (Flamarique et al, 2016)(Flamarique et al., 2016)	**Item Total**	✓	Excellent	×		✓	Medium-Large
	Subscale	×		×		×	

Table 4: Overview of validity testing

Tool/Scale		Validity type												
		Face	Predictive				Discriminant				Convergent		Divergent
			Sensitivity (Se) Specificity (Sp)	PPV/ NPV	Event		Gender	Age	Clinical Group		Measure	Result	
					Type	Result			Type	Result			
Adolescent Suicide Questionnaire (ASQ) (Allison et al., 1995; Horowitz et al., 2012; Stathis et al., 2008)	**Item Total**	×	Se: Large Sp: Medium-Large	PPV: Small-Medium NPV: Large			No difference**				Care & Protection scales PBI MHATODS	Small-Medium	
	Subscale												
Child-Adolescent Suicide Potential Scale (CSPS) (Fennig et al., 2005; Ofek et al., 1998; Pfeffer, 1986; Pfeffer et al., 1979, 1986; Schwartz-Stav et al., 2006)	**Item Total**	×									CDSS, BDI, CCL, HS, SRS, PANNS Negative, SAUMD	Medium-Large	Large
	Subscale				Suicide attempt	2/11 subscales * (P<0.05)	2/11 subscales (P<0.05)*		Suicide attempters/non-attempters. Schizophrenics with/without post psychotic depression, major depression.	2/11 subscales (P<0.05)* 9/11 subscales (P<0.05)*	OAS, SRS, BDI, STAI, ICS, LSI repression	Small-Medium*	
Columbia Suicide Screen (CSS) (Miranda et al., 2014; Shaffer et al., 1996, 2004)	**Item Total**	×	Se: Medium Sp: Large	PPV: Small NPV: Large	Past/future suicide attempt. Suicide ideation/behaviour	P=<0.01-0.01 Identified 81.1%							
	Subscale		Se: Small-Large Sp: Large										
Columbia Suicide Severity Rating Scale (C-SSRS) (Atkinson et al., 2014; Emslie et al., 2015; Findling et al., 2013; Flamarique et al., 2016; Horwitz et al., 2015; Kerr et al., 2014; King et al., 2015; Knafo A et al., 2015; Mirkovic et al., 2015; Posner et al., 2011)	**Item Total**	×	Se: Medium-Large Sp: Small-Large		Future, actual, interrupted, aborted suicide attempts	P<0.05-0.001*			Presence/absence: Suicide behaviour Borderline PD Productive coping Fluoxetine/Duloxetine/ Placebo Escitalopram/Placebo	Beh: P<0.05 BPD: P<0.001 PC: P<0.01 FDP: P<0.05* EP: No difference**	SSI, SIQ-JR, CSHF, MADRS, BDI suicide ideation, STOP-SAS	Medium-Large	
	Subscale						2/4 subscales (P<0.05)*		Suicide attempters/non-attempters	Attempt: 2/4 subscales (P<0.001)*	Suicide items: BDI, MADRS Other items: BDI, MADRS	Suicide: Medium-Large Other: Small-Medium	
Fairy Tales Test (FT)	**Item Total**	×											

(Orbach et al., 1984)	**Subscale**								Suicidal/Normal/Chronically ill	2/4 items (P<0.05)* 1/4 items no difference for suicidal**			Subscale: Small-Medium
Life Orientation Inventory (LOI) (Range and Lewis, 1992a, 1992b)	**Item Total**	✓							Controls/Depressed/ Serious suicide risk/ Very serious suicide risk	No P-Value reported	ISO-30, BDI, HA, HSC, Suicidality Index	Medium-Large	
	Subscale												Subscale: Small-Large
Multi-Attitude Suicide Tendency Scale (MAST) (Ferrara et al., 2012; Gutierrez, 1999; Orbach et al., 1991; Osman et al., 1993, 1994, 2000; Wong, 2004)	**Item Total**	×	Se: Large Sp: Large				P<0.05*		Suicide ideators/attempters Suicide attempters/non-attempters Inpatients/Outpatients	SI/A: P<0.05-0.01 SA/NA: No difference ** I/O: P<0.05	DSRS, SBQ subscales	Large	Accounts for 52.31% total variance
	Subscale		3/4 Subscales * Se: Large, Sp: Medium		Past suicide attempts Self-harming behaviours	PSA: 2/4 Subscales (P<0.05)* SHB: 2/4 subscales (P0.05)*	2/4 subscales (P<0.05)*		Suicide attempters/control Suicide ideators/control Suicidal/psychiatric/control Suicidal/psychiatric/students	SA/C: 3/4 subscales (P<0.05)* SI/C: 1/4 subscales (P<0.05)* S/P/C:3/4 subscales (P<0.01)* S/P/S: 3/4 subscales (P< 0.02)*	CDI, ISP, SBQ, PHCS, SPS, GSI, BRFL-A, MMPI-A, SIQ-JR, PPI positive perceptions, RADS, MCSD	All subscales: Medium-Large 3/4 subscales: Small-Large* 2/4 subscales: Large* 1/4 subscales: Medium-Large*	Subscale: Small-Large (accounts for 4.5-53.8% variance) Items: Medium-Large (accounts for 9-87% variance)
Reasons for Living Inventory for adolescents (RFL-A) (Gutierrez et al., 2000; Labelle et al., 2015; Lee, 2011; Osman et al., 1998)	**Item Total**	×			Past suicide status Suicidality	P < .001 Better than BHS	No difference**		High/low suicide risk School non-suicidal/inpatient suicidal/psychiatric non-suicidal Non-suicidal /First attempters/Attempters	H/L SR: P<0.01 SNS/IS/PNS: P<0.05 NS/FA/A: P<0.001	SPS, SBQ, PHCS, MMPI-A indices	Small-Large	Subscale: Large
	Subscale		3/5 subscales * Se: Medium Sp: Large				1/5 subscales (P< 0.05)*	1/5 subscales (P< 0.05)			BDI-II, BHS, CDI, SIQ, KSPSA	5/5 subscales: Small-Large	Subscale: Small-Large (accounts for 64.8-73.23% variance) Items: Large
Risk of Suicide Questionnaire (RSQ) (Horowitz et al., 2001)	**Item Total**	×											
	Subscale		Se: Medium-Large Sp: Small-Large	PPV: Small-Large NPV: Medium-Large							SIQ	Small-Medium	Items: Large
Modified Scale for Suicide Ideation (M-	**Item Total**	×									SBQ-R, BIS, BDI, BHS, SIS	Medium-Large*	

SSI) (Miller et al., 1986; Pettit et al., 2009)	**Subscale**								With/Without Disruptive Behaviour Disorder Other diagnoses/number of diagnoses	W/W DBD: 2/2 subscales P< 0.05 OD/ND: no difference **			Subscale: Large Subscale to total score: Large Subscale to item: Small-Large (accounts for 45.91% item variance) Items: Medium -Large
Self-Harm Risk Assessment for Children (SHRAC) (Angelkovska, 2014; Angelkovska et al., 2012)	**Item Total**	×							Clinic referred/community Executive function impairment/non-impairment Internalisers/externalisers/neither/combined	CR/C: P<0.002 EFI/NI: P<0.01 I/E/N/C: P<0.01	Hyperactive-impulsive symptoms	Medium	
	Subscale												Item affectivity: Medium-Large Item discrimination: Small-Large
Suicidal Behaviours Questionnaire for Children (SBQ-C) (Cotton et al., 1995; Cotton and Range, 1996)	**Item Total**	×							High/Low risk suicide	73.8% agreement	MAST-A, HSC	Small-Large	
	Subscale										HSC	3/4 Subscales: Small-Large*	
Suicidal Ideation Questionnaire – Junior Version (SIQ-JR) (Gutierrez, 1999; King et al., 1997, 2014, 2015; Mazza, 2000; Mazza and Reynolds, 1999; Reynolds, 1987a; William M. Reynolds, 1990; Reynolds and Mazza, 2001; Storch et al., 2014; Zhang et al., 2014)	**Item Total**	×	Se: Large Sp: Medium	PPV: Small	Past /future suicide attempt. Suicide ideation	Past: Better than CES-D, BHS, DASS (P< 0.05). Future: P<0.05* Ideation: Predicts (P<0.001) but not better.	P<0.001 *	No difference**	High risk /Average risk PTSD Suicidal/non-suicidal**	HRP/ARP: P<0.001 S/NS: no difference**	APS-PTS, EVQ, CDRS, MASC, CES-D, BHS, DASS, SBI. COIS-C, COIS-P, CY-BOCS, SNAP IV, YMRS-P**.	Small-Large*	Large
	Subscale												Subscale: Large (accounts for 67.76% item variance)
Suicidal Ideation Questionnaire (SIQ) (Davis, 1992; Horowitz et al., 2001; Jia et al., 2015; Reynolds, 1987b; S.-C. et al., 2008; Shaunesey et al., 1993; Spirito et al., 1987,	**Item Total**	×	Se: Large Sp: Medium-Large	PPV: Small	Past suicide attempt. Suicide ideation	Past: Better than CES-D, BHS, DASS (P< 0.05). Ideation: Predicts (P<0.001) but not better.	P<0.05*	No difference**	Suicide attempters with acute/chronic psychiatric problems Suicide attempters/ideators /controls	All P's < 0.01	RSQ, RSQ recent life stressors, CES-D, BHS, DASS, DSD, CASS short, CDI, SBI, SIS.	Small-Large	Large

1996; Stanley et al., 2013; Zhang et al., 2014)	**Subscale**												Items: Large (accounts for 67.91-75.84% variance)
Suicide Behaviour Interview (SBI) (W M Reynolds, 1990; William M. Reynolds, 1990)	**Item Total**	×			Past suicide attempt & how recent	P<0.001	P<0.05	P<0.01	Past suicide attempt/non-attempt	P<0.001	SIQ, SIQ-JR, RADS	Medium-Large	
	Subscale						4/18 items: P<0.003 *						Subscale-item: Medium-Large (accounts for 61.6% variance) Items: Large
Suicide Intent Scale (SIS) (Beck et al., 1974; Mieczkowski et al., 1993; Morano et al., 1993; Spirito et al., 1996)	**Item Total**	✓					No difference**	No difference**	Medically/psychiatrically hospitalized suicide attempters Suicide Attempter/non-attempter	All P's <0.01	RADS, HSC, SIQ	Small	
	Subscale				Number of past suicide attempts	0/2 Subscales **					BHS, SSI, RADS, HSC, SIQ GAS, HHDRS**	3/3 Subscales: Medium 2/2 subscales: Medium-Large 1/2 subscales: Small* 1/3 subscales: Small* 2/3 subscales: Medium* No relationship**	Items: Small-Large Subscale: Medium (account for 43.7% variance, between 5.6-31% each) Item-total: Small-Large Item-subscale: Small-Large Subscale-total: Large
Suicide Probability Scale (SPS) (Badura Brack et al., 2012; Cull and Gill, 1982; Eltz et al., 2007; Larzelere et al., 1996; Rosenberg et al., 2006)	**Item Total**	×	Se: Small Sp: Large	PPV: Small	Re-admission from suicide behaviour Future suicide attempt, verbalisation, self-destructive behaviour	Re-admission: P<0.05 Future: P<0.01			Suicide attempters/non-attempters Substance abuse disorder/without	SA/NSA: P<0.05 BS/NBS: P<0.05 SAD/W: No difference**	RADS, HSC, STAXI, 16PF.	Small-Large*	
	Subscale						13/16 items P<0.01 3/4 subscales P<0.05		Suicide attempters/non-attempters Physical/sexual abuse/both/none Burn/non-burn survivors	3/4 Subscales P<0.05* 1/4 subscales physical & both P<0.05*	1 FES subscale	2/4 subscales: Small-Medium*	Males subscale-item: Large (accounts for 58.99-73.52% variance) Females subscale-item: Medium-Large (accounts for 55.31-58.79% variance)

Suicide Risk Scale (Grilo et al., 1999; Plutchik et al., 1989)	**Item Total**	×			Suicide risk	P < .0000	No difference**		Non-suicidal/one suicide attempt/multiple attempts Childhood abuse/none and depressed	NS/0SA/MSA: Mean difference: 2 CA/NCAD: No difference**	Abused Children: BDI, HSC, PFAV, DEQ-A Self-criticism factor. Non-abused depressed children: BDI, HSC, DEQ-A Self-criticism factor. Both: ICS, AAIS, DEQ-A dependency factor**	AC: Small-Large NADC: Medium-Large B: no relationship**	
	Subscale												
Child-Adolescent Suicidal Potential Index (CASPI) (Koutek et al., 2016; Pfeffer et al., 2000)	**Item Total**	✓	Se: Medium Sp: Medium	CDI -ve: PPV: 22.2-26.2% CDI +ve: PPV: 78.1-82.8%	Past suicide attempt. Suicide behaviour	P<0.0001 Better than CDI P<0.05.	P<0.05	P< 0.05	Past suicide attempts /ideas/without	P<0.0001	CDI, R-CMAS, HSC.	Medium-Large	
	Subscale		Se: Medium Sp: Medium						Past suicide attempts/ideas Assaultive acts/ideas/none	PSA/I: 1/3 subscales P<0.0001 AA/AI/N: 2/3 subscales P<0.0001	CDI, R-CMAS, HSC	Small-Large	Subscale-item: None-Large (accounts for 37% variance)
Child Suicide Risk Assessment (CSRA) (Larzelere et al., 2004)	**Item Total**	✓	Se: Large Sp: Large	PPV: Medium NPV: Small					Suicide Attempters/non-attempters	Mean difference: 5	4 suicide criterion items CSRA	Medium	
	Subscale				Past suicide attempt Composite suicide risk	12/20 items (P<0.05)*					4 suicide criterion items CSRA	Small-Medium	Subscales: Account for 38.3% variance (range: 9.52-17.4%) Item-total: Small-Medium Subscale-Item: Small-Large
Suicide Status Form-II (SSF-II) (Conrad et al., 2009; Mcnicholas, 2011; Romanowicz et al., 2013)	**Item Total**	×					P<0.05*	No difference**	Primary depression/without /substance misuse Recent suicide attempters/suicide family history/suicide ideation admissions	PD/W: P<0.05 SM: No difference** RSA/SFH/SIA: P<0.05	CRAFFT	No relationship score. P<0.005	

	Subscale						4/5 subscales (no P-value reported)*	5/5 subscales (no P-value reported)					
Suicidality Treatment Occuring Paediatrics-Suicidality Assessment Scale (STOP-SAS) (Flamarique et al., 2016)	**Item Total**	✓	Se: Large Sp: Large		Suicide risk	$P < 0.001$					CSSR-S	Large	
	Subscale												