

Review Article

Cite this article: Anderson JK *et al* (2018). A systematic review of effectiveness and cost-effectiveness of school-based identification of children and young people at risk of, or currently experiencing mental health difficulties. *Psychological Medicine* **49**, 9–19. <https://doi.org/10.1017/S0033291718002490>

Received: 13 March 2018
Revised: 31 July 2018
Accepted: 13 August 2018
First published online: 13 September 2018

Key words:

mental health; mental wellbeing; schools; identification; screening; children; young people; students; education

Author for correspondence:

Joanna K. Anderson, E-mail: jpa44@medschl.cam.ac.uk

A systematic review of effectiveness and cost-effectiveness of school-based identification of children and young people at risk of, or currently experiencing mental health difficulties

Joanna K. Anderson¹, Tamsin Ford², Emma Sonesson³, Jo Thompson Coon², Ayla Humphrey³, Morwenna Rogers², Darren Moore⁴, Peter B. Jones³, Emmet Clarke⁵ and Emma Howarth¹

¹NIHR CLAHRC East of England, University of Cambridge, Institute of Public Health, Douglas House, 18 Trumpington Road, Cambridge CB2 8AH, UK; ²University of Exeter Medical School, South Cloisters, St Luke's Campus, Exeter EX1 2LU, UK; ³Department of Psychiatry, University of Cambridge, Herschel Smith Building, Forvie Site, Robinson Way, Cambridge CB2 0SZ, UK; ⁴Graduate School of Education, St Luke's Campus, Exeter EX1 2LU, UK and ⁵Norfolk and Suffolk NHS Foundation Trust, 80 St Stephens Rd, Norwich NR1 3RE, UK

Abstract

Background. Although school-based programmes for the identification of children and young people (CYP) with mental health difficulties (MHD) have the potential to improve short- and long-term outcomes across a range of mental disorders, the evidence-base on the effectiveness of these programmes is underdeveloped. In this systematic review, we sought to identify and synthesise evidence on the effectiveness and cost-effectiveness of school-based methods to identify students experiencing MHD, as measured by accurate identification, referral rates, and service uptake.

Method. Electronic bibliographic databases: MEDLINE, Embase, PsycINFO, ERIC, British Education Index and ASSIA were searched. Comparative studies were included if they assessed the effectiveness or cost-effectiveness of strategies to identify students in formal education aged 3–18 years with MHD, presenting symptoms of mental ill health, or exposed to psychosocial risks that increase the likelihood of developing a MHD.

Results. We identified 27 studies describing 44 unique identification programmes. Only one study was a randomised controlled trial. Most studies evaluated the utility of universal screening programmes; where comparison of identification rates was made, the comparator test varied across studies. The heterogeneity of studies, the absence of randomised studies and poor outcome reporting make for a weak evidence-base that only generate tentative conclusions about the effectiveness of school-based identification programmes.

Conclusions. Well-designed pragmatic trials that include the evaluation of cost-effectiveness and detailed process evaluations are necessary to establish the accuracy of different identification models, as well as their effectiveness in connecting students to appropriate support in real-world settings.

Background

In the UK, one in 10 children and young people (CYP) aged 5–16 years suffers from a psychiatric disorder; many more experience symptoms that, whilst not reaching the threshold of clinical disorder, cause significant distress for CYP and their families (Green *et al.*, 2005). Failure to address mental health difficulties (MHD) early in life affects individuals' long-term functioning and wellbeing, and may also generate significant societal costs related to increased health care usage, unemployment, and antisocial behaviours (Joint Commissioning Panel for Mental Health, 2013).

Less than 35% of CYP with diagnosable MHD are identified (Burns *et al.*, 1995), and only 25% of those with clinically impairing psychiatric disorder receive specialist care (Ford *et al.*, 2007). A small number of studies suggest that parents of CYP with MHD often do not realise that their child may benefit from specialist support (Girio-Herrera *et al.*, 2013). Formal identification can highlight the severity of the child's MHD, and encourage parents to seek professional help. Well-designed programmes to identify CYP with MHD show promise for increasing access to supportive services, and may improve mental health (MH) outcomes if combined with evidence-based interventions (D'Souza *et al.*, 2005; Sayal *et al.*, 2010; Husky *et al.*, 2011; Mitchell *et al.*, 2012).

There is strong international policy consensus that schools are well positioned to play a significant role in the early identification of CYP at risk of mental illness. Systematic school-based approaches detect a greater proportion of CYP with MHD compared with less formal processes (i.e. ad-hoc teacher or parent identification, or self-identification) (Garland, 1995; Eklund *et al.*, 2009; Scott *et al.*, 2009; Dowdy *et al.*, 2013; Kieling *et al.*, 2014). Students identified in school settings are more likely to receive parental and school support, as well as referral and access to MH services (D'Souza *et al.*, 2005; Nemeroff *et al.*, 2008; Sayal *et al.*, 2010; Lyon *et al.*, 2015), and to achieve better long-term MH outcomes, compared with students with MHD identified in community healthcare settings (Ford *et al.*, 2008; Husky *et al.*, 2011; Mitchell *et al.*, 2012). However, teachers do not feel well equipped to perform this role and consistently under-identify early symptoms of various disorders (Caldarella *et al.*, 2008; Bruhn *et al.*, 2014; Cunningham and Suldo, 2014).

The evidence-base on programmes to improve identification of MHD in school settings has not been synthesised. In this paper, we sought to synthesise evidence on the effectiveness and cost-effectiveness of school-based methods to identify CYP at risk of or experiencing MHD. This is a part of a larger systematic review of the effectiveness, harms, feasibility, and acceptability of school-based methods to identify CYP with MHD; findings on harms, feasibility, and acceptability will be published separately in due course. Given that we were specifically interested in the utility of the identification mechanism, effectiveness was defined as (i) rate of accurate identification (i.e. correct identification of cases) of CYP with MHD; (ii) rate of referrals to appropriate supportive services following identification; (iii) uptake of referrals to supportive services. Cost-effectiveness was defined broadly as the outcome of analysis comparing the resources required to deliver an intervention with the health, quality of life or other assumed outcomes achieved by an intervention (Knapp and Iemmi, 2013).

Methods

The protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO; <https://www.crd.york.ac.uk/prospero>), registration number: 42016053084 (amended version dated 18 January 2017).

School-based methods of identification

The literature describes four main models of school-based identification of MHD (Whitney *et al.*, 2011). Universal screening programmes aim to systematically assess all students for risks of MHD using self-, parent-, or teacher-report measures (Whitney *et al.*, 2011). Curriculum-based models, delivered to all students in a year group by a staff member or external person with relevant knowledge, are designed to increase students' knowledge and recognition of common MH problems, and develop skills to address them. (Whitney *et al.*, 2011). Staff in-service models rely on training all members of staff to recognise early signs of MHD and link students deemed to be at-risk with appropriate support. (Whitney *et al.*, 2011). Teacher nomination involves asking a class teacher to identify students in their classroom who exhibit concerning behaviours or symptoms that may indicate the presence of MHD (Cunningham and Suldo, 2014). Additionally, we included traditional identification methods using office disciplinary referrals (ODRs), grade point average, attendance data, and teacher referral to identify students at risk of MHD.

Since the paper describes studies that evaluated the accuracy of identification of suicide risk, it is important to note there is a consensus that, although identification of suicide risk yields a high number of false positive results, the harm of these inaccurate identifications is outweighed by the benefit of prevention for future suicides among those whose risk is correctly identified (Carter *et al.*, 2017). The result of recent systematic review showed that pooled positive predictive value of clinical instruments used to assess suicide risk is around 5.5%, which suggests that majority of individuals who screen positive will, in fact, not attempt suicide (Carter *et al.*, 2017).

Inclusion and exclusion criteria

Comparative studies were included if they assessed the effectiveness or cost-effectiveness of strategies to identify students in formal education aged 3–18 years (1) with a MHD, (2) presenting symptoms of mental ill health, or (3) exposed to psychosocial risks that increase the likelihood of developing a MHD.

Studies that focused on the identification of global and specific learning disabilities were excluded.

We included studies published in any year comparing the effectiveness of different identification models within the same group, and studies in which the accuracy of identification was verified by a subsequent clinical evaluation, or compared with existing MH diagnoses.

Search strategy

Electronic bibliographic databases: MEDLINE and Embase via OvidSP; PsycINFO, ERIC, and British Education Index via EBSCOhost; and ASSIA via ProQuest were searched in May and June 2017 and again in July 2018. The search strategy combined terms for identification and school settings with terms for MH. Search terms were generated by examining the terminology used in key publications in the field, identifying synonyms, and discussing with experts in school-based MH research. The search terms were combined with standard MeSH terms for the MEDLINE database, Emtree terms for Embase, Thesaurus terms for ERIC, British Education Index and ASSIA, and Subject Headings for the PsycINFO database. Supplementary search methods included forward and backward citation search, and hand-searching CYP MH journals. The MEDLINE search strategy is shown in an online Supplementary Table S1.

Selection of studies

Two independent reviewers selected studies in three stages: (1) all titles were examined to remove obviously irrelevant reports; (2) abstracts of remaining studies were examined against inclusion/exclusion criteria; (3) full-texts of remaining reports were examined for compliance with inclusion/exclusion criteria. We resolved disagreements by referral to another research team member.

Data extraction

The fields of the extraction tables were piloted and refined using three randomly selected studies included in the review. Two researchers independently extracted data from included studies. We extracted the following information: first author, year of publication, and country where the study was conducted, study design, study aims, school level, informants, identification

measures, description of an identification programme, characteristic of a sample, and findings. In a separate table, we listed programmes' components (online Supplementary Table S4). Results were compared and disagreements were resolved by referral to another research team member.

Study appraisal

We appraised the quality of included studies with the Effective Public Health Practice Project (EPHPP) Quality Assessment Tool for Quantitative Studies (Armijo-Olivo *et al.*, 2012), which has been deemed suitable to use systematic reviews of effectiveness (Deeks *et al.*, 2003). The tool includes six quality components: selection bias, study design, confounders, blinding, data collection and drop-out rated against set criteria as strong, moderate, or weak. Two researchers independently conducted quality appraisal judging each study against criteria listed for each quality component; results were compared and disagreements were resolved by referral to another team member.

Synthesis of results

Due to high heterogeneity of study designs, interventions, and outcome measures, it was not appropriate to conduct a meta-analysis. We provided a numerical account of evidence and narrative synthesis of evidence-guided by the framework for systematic reviews developed by Popay *et al.* (2006). This framework comprises four iterative stages: developing the theory of change, preliminarily synthesising of findings, exploring relationships in the data, and assessing the robustness of syntheses. We described findings separately for each research questions, as well as an overall summary and conclusions (Popay *et al.*, 2006).

Findings

Twenty-seven studies were included in the final review Fig. 1 outlines the study selection process.

Characteristics of included studies are shown in Table 1. Studies covered a total of 44 unique identification programmes. Publication dates suggest increasing interest in this area over the last two decades, but it should be noted that nearly all evidence comes from the USA. Nearly half of the studies were cross-sectional ($n = 13$), followed by comparison group ($n = 8$) and cohort analytic studies ($n = 4$). There was only one case-control study and one randomised controlled trial (RCT). Most focussed on secondary school settings ($n = 16$) and identification of behavioural and emotional problems. Nine studies evaluated universal screening models; remaining studies compared universal screening with teacher nomination ($n = 12$), traditional identification methods ($n = 3$) and staff in-service training ($n = 1$). One cost-effectiveness study compared universal screening, staff in-service training, and curriculum-based models. Detailed characteristics of studies are presented in an online Supplementary Table S2.

Quality of included studies

As shown in Table 2, nearly a quarter of included studies were rated weak on selection bias, lacking sufficient description of recruitment procedures and representativeness of the sample. Nearly half of the studies failed to report withdrawals and attrition. All but one study was rated strong for data collection, having utilised standardised and validated measures.

Rates of accurate identification

Findings from all studies are described in an online Supplementary Table S3. Section 'Universal screening programmes' describes studies that evaluated the effectiveness of a single identification model (universal screening programmes); subsequent sections describe studies that compared the effectiveness of universal screening and other identification models.

(1) Universal screening programmes

Eight studies of universal screening programmes reported on rates of identification (Tisher, 1995; Jones *et al.*, 2002; Gould *et al.*, 2009; Robinson *et al.*, 2010; Husky *et al.*, 2011; Morey *et al.*, 2015; Hilt *et al.*, 2018). In six studies, positive screening results were verified by subsequent clinical interview conducted by MH professionals (Gould *et al.*, 2009; Robinson *et al.*, 2010; Husky *et al.*, 2011; Morey *et al.*, 2015; Hilt *et al.*, 2018), or an existing diagnosis of MHD (Tisher, 1995), giving a reliable rate of false positives.

Depression and anxiety: Studies that utilised student-report screening measures and subsequent clinical interview, found that 45–63% of secondary school students were identified as being at high-risk for depression (Robinson *et al.*, 2010; Morey *et al.*, 2015). Teacher-completed universal screening accurately distinguished between students with and without clinical depression diagnoses; students currently treated for depression scored significantly higher compared with their non-diagnosed counterparts ($p < 0.0001$) (Tisher, 1995).

Behavioural and socioemotional problems: One study considered the utility of teacher- and parent-completed universal screening by the examination of long-term outcomes for children identified in kindergarten as at high-risk of behavioural and socioemotional problems (Jones *et al.*, 2002). Children identified as high-risk were significantly more likely, in the next 6 years, to receive professional outpatient and inpatient MH services ($p < 0.05$), take medication for MH reasons ($p < 0.01$), and receive special education services or MH-related school counselling ($p < 0.01$), compared with low-risk children (Jones *et al.*, 2002).

Risk of suicide: Student-report universal screening identified 317 students out of 2342 screened (13%) as being at risk of suicide. Subsequent clinical interview, however, indicated that 43% of these outcomes were false positives, (Gould *et al.*, 2009), which indicates over-identification as we anticipated. In contrast, following completion of a universal screening measure, all students in two other studies, regardless of risk-status, participated in a brief interview with a school counsellor. Interview outcomes indicated that screening yielded around 20% false negative results, which suggests that this method is likely to miss a significant number of students needing support (Husky *et al.*, 2011; Hilt *et al.*, 2018).

(2) Universal screening programmes *v.* teacher nomination

Twelve studies compared identification rates from universal screening and school staff nomination models (Tisher, 1995; Auger, 2000; 2004; Campbell, 2004; Dwyer *et al.*, 2006; Eklund *et al.*, 2009; Scott *et al.*, 2009; Dowdy *et al.*, 2013; Cunningham and Suldo, 2014; Kieling *et al.*, 2014; Sweeney *et al.*, 2015; Kilgus *et al.*, 2018). In four studies positive identification outcomes were verified by subsequent clinical interview (Auger, 2000; 2004; Scott *et al.*, 2009; Kieling *et al.*, 2014; Sweeney *et al.*, 2015), while remaining studies reported rates of overlap in identification between screening and staff nomination.

Depression and anxiety: Findings from a study that employed a multi-stage model of universal screening and a clinical interview

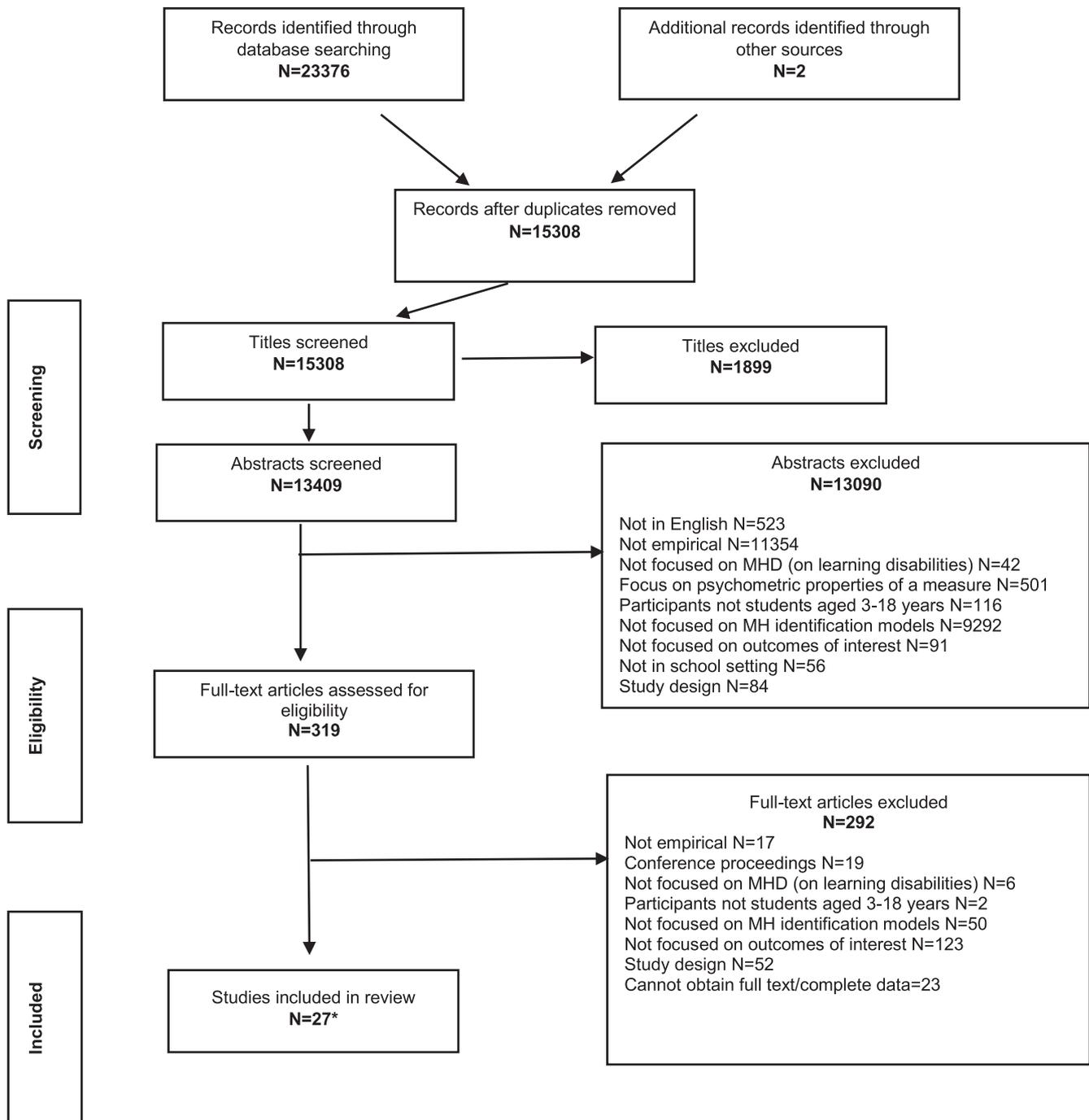


Fig. 1. Study selection and exclusion flow diagram.

to identify students with depression showed that this method produced a high number of false-positive results (up to 90%). By comparison, teacher nomination yielded a false positive rate of nearly 70% (Auger, 2000; 2004). Universal screening for social anxiety disorder (SAD) yielded fewer false-positives (20%), with only 12% of subsequently diagnosed students identified by teachers (Sweeney *et al.*, 2015); seven students with a final SAD diagnosis were identified solely by teacher nomination. Other evidence suggests that teachers correctly nominate 41–68% of students who screen positive for depression and/or anxiety, but since tested models did not include a clinical interview, the rates of false

positive and negative results for each method cannot be determined (Campbell, 2004; Cunningham and Suldo, 2014).

Behavioural and socioemotional problems: Seven studies compared identification rates of students with behavioural and socioemotional problems that used universal screening and nomination models (Tyne and Flynn, 1981; Garland, 1995; Dwyer *et al.*, 2006; Eklund *et al.*, 2009; Dowdy *et al.*, 2013; Kieling *et al.*, 2014; Kilgus *et al.*, 2018). Only one study verified positive results of screening with subsequent clinical evaluation (Kieling *et al.*, 2014); remaining studies reported rates of overlaps between model outcomes. Evidence suggests that student-report

Table 1. Characteristics of included studies

Criterion	Characteristic	No of studies (<i>n</i> total = 27)
Year	Before 1990	1
	1991–2000	5
	2001–2010	9
	After 2010	12
Country	USA	21
	Ireland	2
	Australia	3
	Brazil	1
Study design	RCT	1
	Case-control	1
	Comparison group	8
	Cohort analytic	4
	Cross-sectional	13
Identification model ^a	Universal screening	26
	Teacher/ school staff nomination	11
	Traditional school identification	5
	Curriculum-based	1
	In-service training	2
Informants ^a	Student	16
	Teacher/other school staff	17
	Parent	4
Condition	Depression and anxiety	7
	Behavioural and socioemotional problems	12
	Risk of suicide	6
	Substance abuse	1
School level ^b	Preschool	3
	Primary school	11
	Secondary school	16
Parent consent ^a	Yes	30
	No	0
	ND	12
Student assent ^a	Yes	12
	No	0
	ND	30
Follow-up clinical evaluation ^a	Yes	16
	No	29
	ND	9
Informing students/parents about results ^a	Yes	5

(Continued)

Table 1. (Continued.)

Criterion	Characteristic	No of studies (<i>n</i> total = 27)
	No	0
	ND	37
Referral/ recommendations following identification ^a	Yes	9
	No	1
	ND	32

ND, not defined

^aIn total 44 identification programmes are described in 24 studies. Some studies describe more than one programme thus some characteristics are reported multiple times for one study^bSome studies are conducted in multiple schools, at different school levels

universal screening identifies at least twice as many at-risk students as teacher nomination (Garland, 1995; Eklund *et al.*, 2009; Dowdy *et al.*, 2013). Teachers identify 10–30% of students identified by a universal screener (Garland, 1995; Dwyer *et al.*, 2006; Eklund *et al.*, 2009; Dowdy *et al.*, 2013). They are more likely to nominate students who have more severe difficulties (Garland, 1995), and more ODRs (Dowdy *et al.*, 2013). However, combining universal screening and nomination did not increase the accuracy of identification compared with universal screening alone (Kilgus *et al.*, 2018). One study found that a parent-completed universal screener more accurately identified students subsequently diagnosed with internalising disorders, compared with a teacher-completed measure (30–46% and 26–34%, respectively). In contrast, teachers' positive global judgement about children's risk of developing MHD better predicts future externalising problems, compared with parent's judgement (Dwyer *et al.*, 2006). Findings from another study suggest the agreement between results of peer-report universal screening and teacher nomination increases with students' age from 19% in 3rd grade (7–8 years old) to 55% in 5th grade (10–11 years old) (Tyne and Flynn, 1981), perhaps because older students can more accurately judge others' behaviours.

ADHD: Only one study focussed on the identification of children with ADHD. Identification results were verified by the full clinical assessment that suggested very low levels of agreement between teacher-completed screening and simple nomination ($p < 0.0002$) (Kieling *et al.*, 2014). Seventeen out of 18 children with clinically-confirmed ADHD diagnoses were identified by at least one screening measure (Kieling *et al.*, 2014), while the agreement between nomination and the final diagnosis was significantly higher for negative cases than positive cases.

Risk of suicide: One study compared the results of student-report universal screening for suicide risk and school staff nomination, with outcomes verified by subsequent clinical interview (Scott *et al.*, 2009). MH professionals correctly nominated twice as many students as did administrative staff, with an accurate nomination rate of 36%, whereas screening correctly identified 63% of at-risk students. Screening yielded a 9% false-positive rate compared with 24% produced by staff nomination. Both methods combined produced only 5% false positives.

(3) Universal screening programmes *v.* traditional school identification methods

Three studies compared the accuracy of universal screening and traditional identification methods used by schools (i.e.

Table 2. Quality of included studies – EPHPP tool

1st author (year); study design	Selection bias	Study design	Confounders	Blinding	Data collection	Drop out
<i>Rates of accurate identification</i>						
Universal screening						
Depression and anxiety						
Morey (2015); comparison group	weak	weak	NA	NA	strong	strong
Robinson (2010); cohort analytic	weak	moderate	NA	NA	strong	moderate
Tisher (1995); case-control	moderate	moderate	NA	NA	strong	strong
Behavioural and socioemotional problems						
Forness (1998); comparison group	moderate	moderate	NA	NA	strong	weak
Jones (2002); cohort analytic	moderate	moderate	NA	NA	strong	strong
Risk of suicide						
Gould (2009); cohort analytic	moderate	moderate	NA	NA	strong	moderate
Hilt (2018); cross-sectional	Moderate	Weak	NA	NA	Strong	weak
Husky (2011); RCT	moderate	strong	strong	moderate	strong	strong
<i>Universal Screening v. Nomination</i>						
Depression and anxiety						
Auger (2000, 2004); cross-sectional	weak	weak	NA	NA	strong	weak
Cunningham (2014); cross-sectional	weak	weak	NA	NA	strong	weak
Sweeney (2015); cross-sectional	strong	weak	NA	NA	strong	strong
Campbell (2004); comparison group	moderate	weak	NA	NA	strong	weak
Behavioural and socioemotional problems						
Dowdy (2013); cross-sectional	moderate	weak	NA	NA	weak	strong
Dwyer (2006); cohort analytic	moderate	moderate	NA	NA	strong	weak
Eklund (2009); comparison group	moderate	weak	NA	NA	strong	weak
Garland (1995); cross-sectional	moderate	weak	NA	NA	strong	moderate
Kieling (2014); comparison group	moderate	weak	NA	NA	strong	strong
Kilgus (2018) cross-sectional	moderate	weak	NA	NA	strong	moderate
Tyne (1981); cross-sectional	weak	weak	NA	NA	strong	weak
Risk of suicide						
Scott (2009); cross sectional	moderate	weak	NA	NA	strong	moderate
1st author (year); study design	Selection bias	Study design	Confounders	Blinding	Data collection	Drop out
<i>Universal screening v. traditional identification methods</i>						
Behavioural and socioemotional problems						
Eklund (2014); cross Sectional	moderate	weak	NA	NA	strong	weak
Hallfors (2006); Comparison group	moderate	moderate	NA	NA	strong	weak
Naser (2014); Cross-sectional	moderate	weak	NA	NA	weak	moderate
Universal screening v. staff in-service training						
Substance abuse						
McLaughlin (1993); Comparison group	moderate	weak	NA	NA	strong	strong
Rates of referral and uptake						
Cotter (2015); Comparison group	moderate	weak	NA	NA	strong	weak
Also see: Gould (2009), Hilt (2018), Husky (2011)						
Cost-effectiveness						
Burke (2013); cross-sectional	weak	weak	NA	NA	strong	weak

ODRs, grade point average, attendance data, and teacher referral) (Hallfors *et al.*, 2006; Eklund and Dowdy, 2014; Naser, 2014). Neither study verified outcomes by clinical assessment, reporting only rates of overlaps between methods.

Behavioural and socioemotional problems: Traditional identification based on teacher-referral and academic performance identified less than 40% of children who screened positive for internalising or externalising disorders (Eklund and Dowdy, 2014). Of kindergarten children identified by teacher-completed universal screening as being high-risk, less than 40% were identified by traditional methods during the first year of primary school (Forness *et al.*, 1998). However, a substantial number of children identified by schools' normal procedures were assigned to a different diagnostic category than the one indicated by the results of screening using validated, standardised measures (Forness *et al.*, 1998).

Substance abuse: A study that compared the outcomes of universal screening and traditional methods of identifying substance-abusing students based on student's GPA, attendance record, and teacher referrals yielded equivocal results (Hallfors *et al.*, 2006). In one sample of students, high-risk of substance abuse indicated by student-report screening was associated with low GPA, while in other sample, low attendance and teacher referral, but not GPA, were strong predictors of substance abuse.

(4) Universal screening programmes *v.* staff in-service programmes

Substance abuse: Results of one study suggest that attending in-service training improves teachers' ability to correctly nominate students identified by student-report universal screening as being at-risk of substance abuse (McLaughlin *et al.*, 1993). Teachers who completed the training more accurately identified students who were experimenting with, and regularly using drugs and alcohol, compared with their colleagues who had not attended the training, thereby reducing the gap in identification rates between the two methods ($p < 0.001$).

Rates of referrals and service uptake

Although a number of studies indicated that a referral was made for students identified as being at-risk, only three studies reported numbers and uptake of referrals to specialist support (Gould *et al.*, 2009; Husky *et al.*, 2011; Hilt *et al.*, 2018). All three studies evaluated universal screening for risk of suicide, although referral processes and services offered varied by programme. Of 317 students identified as at-risk of suicide by student-report screening in the Gould *et al.* (2009) study, 182 (57%) were deemed to require additional support following a second clinical stage interview (Gould *et al.*, 2009). Referrals were made for 147 students (of whom 29 were already receiving MH services) reporting severe suicidality; the remaining 35 were given a list of local providers without a specific referral. The uptake of follow-up recommendations was 70.3%. Uptake did not differ between students who received a specialist referral or list of providers; those who were not currently receiving services were significantly more likely to follow-up with the referral compared with those already in treatment. Overall, 24% of the new service users had their first appointment within a month of the screening. Within 6 months, 52% attended their first appointment, and within a year, 70% had successfully accessed a MH care provider (Gould *et al.*, 2009). Of 2022 students participating in a universal screening for suicide risk programme, 444 students were determined to be in need of MH services following screening and clinical interview (Hilt *et al.*, 2018). Of those identified as being at-risk, 77% were not

currently in treatment. The majority (89%) were referred to community services, and those remaining received referrals to school services. Case-management confirmed that 50.2% of referred students attended one or more appointments; 22.5% completed three or more appointments.

Of the 2488 students included in the Husky *et al.* (2011) study, universal screening and subsequent clinical interview identified 299 (12%) students as at-risk of suicide (Husky *et al.*, 2011). Based on current suicidal ideation as assessed by clinical interview, past suicide attempts, and current MH treatment status, 128 (43% of those identified) students received a referral to school-based MH services only, 78 (26%) to community-based MH services only, and 93 (31%) to both school and community-based services. Of those referred, 76% had at least one appointment with a MH provider and 56% received minimally adequate treatment defined as three or more appointments or any number if terminated by provider's recommendation. Among the 221 students referred to school-based services, 80% attended at least one appointment, 71.3% of whom received minimally adequate treatment. Of 171 students referred to community-based services, 42% received at least one visit, 68% of whom received minimally adequate treatment.

Another study reported the uptake of a clinical interview following a positive student-report screen for suicide risk (Cotter *et al.*, 2015). Of 516 students invited for a follow-up assessment, 37% attended. Recent suicide attempt, high levels of depressive, anxiety or emotional symptoms, hyperactivity/inattention, peer relationship problems, and functional impairment increased the likelihood of attending a follow-up interview.

Cost-effectiveness

Only one study compared the cost-effectiveness of different methods of identifying suicide risk (Burke *et al.*, 2013), and concluded that universal screening is more cost-effective (in terms of improving quality-adjusted life years – i.e. function of quality and length of life), than curriculum-based or in-service training programmes. The study utilised data from a sample of 11 100 adolescents from 168 schools across 10 European Union countries, so although the findings may accurately represent average cost-effectiveness of suicide screening across the EU countries, results may differ by country and world region. This represents a gap in the research literature.

Discussion

Summary of findings

We identified 27 studies with a total of 26 256 participants that analysed the effectiveness of school-based MHD identification programmes. None of the studies was UK-based. Only one study used a randomised design. Most studies evaluated the utility of universal screening but programmes differed in format and outcomes; where comparison of identification rates was made, the comparator test varied across studies. Whilst the purported aim of many programmes was to increase the rate of MH support among children and young people, only two studies reported referral and uptake data.

Overall, the heterogeneity of studies, the absence of randomised studies and poor outcome reporting make for a weak evidence-base that only generate tentative conclusions about the effectiveness of school-based identification programmes.

Summary of effects of interventions

Some evidence suggests that overall, universal screening may be the most effective method of identification; however, the rate of false-positive results yielded by this method is high (Auger, 2000; 2004; Husky *et al.*, 2011), so the expectations of teachers, pupils, and parents would need to be managed accordingly. Some findings indicate that multistage models are more accurate (Scott *et al.*, 2009; Morey *et al.*, 2015; Sweeney *et al.*, 2015); however two studies reported that a single assessment with a universal screening measure is sufficient to accurately identify high-risk individuals, and additional assessments and informants do not improve accuracy (Dowdy *et al.*, 2016; Kilgus *et al.*, 2018). Teacher nomination yields a higher number of false negative results than universal screening (Campbell, 2004; Dwyer *et al.*, 2006; Eklund *et al.*, 2009; Dowdy *et al.*, 2013; Cunningham and Suldo, 2014). Teachers are most likely to nominate high-risk students, while those who are at-risk but without obvious signs of MHD are often overlooked in ad-hoc identification procedures (Ollendick *et al.*, 1990; Auger, 2004). Limited evidence suggests that staff in-service training and curriculum-based programmes improve identification of MHD (McLaughlin *et al.*, 1993; Robinson *et al.*, 2010); however, costs associated with programme delivery make them less feasible than universal screening (Burke *et al.*, 2013). Combining universal screening and staff nomination shows promise for increasing accuracy of identification (Gould *et al.*, 2009; Scott *et al.*, 2009), although this proposition requires testing using randomised designs.

Few studies focused on identification of pre- and primary school children. It is vital to identify children with MHD as early as possible since evidence shows that presence MHD in children as young as three years old can impact future outcomes across multiple domains including education, employment, substance use, criminal activity, and physical and MH (Jones *et al.*, 2015). Half of MHD is evident by the age of 14, with the even earlier onset of anxiety and impulse control disorders. There is, therefore, a strong case for developing methods of identification for use in primary school settings (Jones, 2013).

Very few studies reported rates of service referral and uptake following identification. Given that MH services are already overwhelmed, commissioners, and service providers may be concerned that school-based identification, and universal screening programmes in particular (which yield a significant number of false positive results), will add unwarranted pressure to already struggling services. Conversely, evidence suggests some children have subclinical levels of psychopathology and will benefit from specialist support (Ford T *et al.*, 2005).

Few studies explicitly set out to assess adverse events or harms associated with identification. Since it is recognised that the identification process may cause distress, especially in high-risk students (Robinson *et al.*, 2011), all studies should assess negative consequences associated with identification.

In general, the description of programmes was poor, with key details such as methods for obtaining consent omitted. Poor reporting of interventions is ubiquitous, and is in part explained by the word limits imposed on authors for papers published in peer-reviewed journals (Hoffmann *et al.*, 2014; Maggin and Johnson, 2015). Nevertheless, without adequate description, it is difficult, if not impossible, to compare trials. The mechanisms by which interventions were hypothesised to lead to change were also rarely reported. This, in combination with poor programme description, means we were unable to identify and define

the role of programme components in the causal pathway leading to benefit, no effect or harm. We also note that there was poor attention to broader contextual factors that may influence programme implementation and outcomes, Intervention development, modelling, feasibility and pilot studies, along with trials of effectiveness, need to theorise and evaluate the contextual conditions necessary for intervention mechanisms to be activated. If there is to be any hope of identifying and scaling promising programmes, then concerted effort is needed to articulate, test and refine programme theories underpinning these complex interventions so as to make explicit how individual study components and contextual factors interact to generate desired outcomes (Wells *et al.*, 2012; Fletcher *et al.*, 2016; Howarth *et al.*, 2016).

Commissioners and practitioners call for more interventions to be tested real-life settings, since the focus on internal validity and creating optimal conditions can significantly limit external relevance and impede dissemination (Bowen *et al.*, 2009). More economic evaluations of identification programmes are required to inform the resource allocation to achieve the best value for money.

Quality of the evidence

More than half of included studies were rated weak in terms of study design, and documentation of withdrawals and drop-outs. Only one RCTs was identified, despite recommendations for trials that focus on both outcomes and processes (Oakley *et al.*, 2006). Nearly a quarter had not sufficiently described sample selection and recruitment procedures, which raises questions about the generalisability of results. Most included studies compared outcomes of two different identification models in the same sample of students. Authors draw conclusions about models' accuracy based on overlaps between their results, thereby assuming that if students are identified by two independent models, then the outcome is most likely correct. Few studies that verified outcomes with subsequent clinical interview further assessed students initially identified as being at-risk, thereby failing to account for false negative results. Studies evaluating the effectiveness of identification models need to include an established, reliable method of verifying both positive and negative results, to minimise the risk of harms that may result from the failure to identify children who have MHD, as well as the over-diagnosis of MHD among children without MH problems (Cohen *et al.*, 2016).

Strengths and limitations

To our knowledge, this is the first attempt to synthesise evidence for the effectiveness of school-based identification models. The inclusion criteria were designed to encompass all existing identification models. The broad scope in terms of study design offered a more comprehensive and realistic understanding of the state of school-based identification than if we excluded ad-hoc identification methods. Second, the review included all age groups, from pre-school to secondary school, which allowed for comparison across school levels. Third, the review did not place any restriction on the type of MH condition, which allowed for cross-condition comparisons. Finally, in addition to exploring the effectiveness of the different models, we also examined referral and uptake rates. This is important because school-based identification models do not end at screening, and understanding the subsequent pathways to care is essential.

Notwithstanding, we acknowledge several limitations. First, the review only included studies published in English. Second, while we generally view our broad inclusion criteria as positive, the lack of exclusion based on study design led to the inclusion of several methodologically weak studies. While we kept a broad scope in terms of MHD, we did not include neurodevelopmental conditions or learning disabilities, which may be closely linked to MH problems. Finally, the quality and heterogeneity of included studies precluded meta-analysis or any other statistical summary. Future reviewers may seek to broaden the aims of the present review through the inclusion of these conditions.

Conclusions

This first comprehensive systematic review of the effectiveness of school-based models of identifying children at risk of, or experiencing MHD shows that the current evidence-base is very limited and does not support the recommendation of any particular model. Well-designed pragmatic trials that include the evaluation of cost-effectiveness and detailed process evaluations are necessary to establish the accuracy of different models, as well as effectiveness in connecting pupils to appropriate support in real-world settings.

Recommendations

- (1) Precise rates of false positive and false negative results yielded by different identification methods need to be established through a reliable method of outcome verification (i.e. clinical assessment by MH professionals or standardised diagnostic assessment with or without clinical review)
- (2) Research to establish which identification models work for younger children, including those under 5 years of age, are particularly needed.
- (3) Studies are needed to evaluate and report the uptake of supportive services following positive identification to estimate additional demand on MH services.
- (4) Detailed descriptions of evaluated identification programmes highlighting their 'core components' should be an essential part of every study, to ensure effective implementation and optimal outcomes once a programme is rolled-out.
- (5) Effectiveness trials including process evaluation components (from identification to treatment) are needed to establish which models most accurately identify which conditions, and which external factors may influence programme outcomes. Identification models need to be tested in real life conditions to ensure that they are sustainable beyond the duration of a research project. Cost-effectiveness is an essential and currently under-studied component of this work of effectiveness trials.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0033291718002490>

Author contributions. J. Anderson led the systematic review including the development of methodology, literature searches, selection of studies for inclusion, data extraction, quality rating, synthesis of findings and writing manuscript drafts. E. Howarth and E. Sonesson acted as second reviewers for selection of studies and data extraction, reviewed all manuscript versions, and made substantial revisions. E. Howarth made substantial contributions to writing introduction and discussion, particularly sections regarding implications for practice and directions for future research. E. Sonesson acted as the second rater for the quality of studies appraisal. T. Ford advised on the

scope of the review, particularly defining the search terms and specifying the inclusion/exclusion criteria, reviewed, and reviewed all manuscript versions. J. Thompson Coon and M. Rogers advised on methodology, particularly the development of search strategy, inclusion/exclusion criteria, and data extraction, and reviewed all manuscript versions, with a particular focus on methods. A. Humphrey reviewed all manuscript versions and contributed to writing the discussion. D. Moore and P. Jones reviewed all manuscript versions. E. Clarke, as a second reviewer, contributed to a screening of titles and abstracts. All authors read and approved the final version of the manuscript.

Acknowledgements. This paper presents independent research funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research & Care (CLAHRC) East of England, at Cambridgeshire and Peterborough NHS Foundation Trust, and CLAHRC South West. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Conflict of interests. None

References

- Armijo Olivo S, Stiles CR, Hagen NA, Biondo PD and Cummings GG (2012) Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration risk of bias tool and the effective public health practice project quality assessment tool: methodological research. *Journal of Evaluation in Clinical Practice* **18**, 12–18.
- Auger RW (2000) An examination of teachers' ability to identify middle school students with depressive characteristics. In *Education* University of Iowa: University of Iowa.
- Auger RW (2004) The accuracy of teacher reports in the identification of middle school students With depressive symptomatology. *Psychology in the Schools* **41**, 379–389.
- Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, Bakken S, Kaplan CP, Squiers L, Fabrizio C and Fernandez M (2009) How we design feasibility studies. *American Journal of Preventive Medicine* **36**, 452–457.
- Bruhn AL, Woods-Groves S and Huddle S (2014) A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education & Treatment of Children* **37**, 611–634.
- Burke LA, Wasserman D, Carli V, Corcoran P, Keeley H, Balazs J, Bobes J, Apter A, Brunner R, Cosman D, Haring C, Pierre Kahn J, Marusic D, Postuvan V, Saiz P and Varik A (2013) A cost effectiveness analysis of four arms of a school-based mental health intervention in Europe. *European Child and Adolescent Psychiatry* **1**, S198–S199.
- Burns BJ, Costello EJ, Angold A, Tweed D, Stangl D, Farmer EM and Erkanli A (1995) Children's mental health service use across service sectors. *Health Affairs* **14**, 147–159.
- Caldarella P, Young EL, Richardson MJ, Young BJ and Young KR (2008) Validation of the systematic screening for behavior disorders in middle and junior high school. *Journal of Emotional and Behavioral Disorders* **16**, 105–117.
- Campbell MA (2004) Identification of 'at-risk' students for prevention and early intervention programs in secondary schools. *Australian Journal of Guidance and Counselling* **14**, 65–77.
- Carter G, Milner A, McGill K, Pirkis J, Kapur N and Spittal MJ (2017) Predicting suicidal behaviours using clinical instruments: systematic review and meta-analysis of positive predictive values for risk scales. *The British Journal of Psychiatry* **210**, 387–395.
- Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, De Vet HC and Bossuyt PM (2016) STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* **6**, e012799.
- Cotter P, Kaess M, Corcoran P, Parzer P, Brunner R, Keeley H, Carli V, Wasserman C, Hoven C, Sarchiapone M, Apter A, Balazs J, Bobes J, Cosman D, Haring C, Kahn JP, Resch F, Postuvan V, Varnik A and Wasserman D (2015) Help-seeking behaviour following school-based screening for current suicidality among European adolescents. *Social Psychiatry and Psychiatric Epidemiology* **50**, 973–982.

- Cunningham JM and Suldo SM (2014) Accuracy of teachers in identifying elementary school students who report at-risk levels of anxiety and depression. *School Mental Health* 6, 237–250.
- D'Souza CM, Forman SF and Austin SB (2005) Follow-up evaluation of a high school eating disorders screening program: knowledge, awareness and self-referral. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine* 36, 208–213.
- Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, Petticrew M and Altman DG (2003) Evaluating non-randomised intervention studies. *Health Technology Assessment* 7, 3–173.
- Dowdy E, Doane K, Eklund K and Dever BV (2013) A comparison of teacher nomination and screening to identify behavioral and emotional risk within a sample of underrepresented students. *Journal of Emotional and Behavioral Disorders* 21, 127–137.
- Dowdy E, Dever BV, Raines TC and Moffa K (2016) A preliminary investigation into the added value of multiple gates and informants in universal screening for behavioral and emotional risk. *Journal of Applied School Psychology* 32, 178–198.
- Dwyer SB, Nicholson JM and Battistutta D (2006) Parent and teacher identification of children at risk of developing internalizing or externalizing mental health problems: a comparison of screening methods. *Prevention Science* 7, 343–357.
- Eklund K and Dowdy E (2014) Screening for behavioral and emotional risk versus traditional school identification methods. *School Mental Health* 6, 40–49.
- Eklund K, Renshaw TL, Dowdy E, Jimerson SR, Hart SR, Jones CN and Earhart J (2009) Early identification of behavioral and emotional problems in youth: universal screening versus teacher-referral identification. *California School Psychologist* 14, 89–95.
- Fletcher A, Jamal F, Moore G, Evans RE, Murphy S and Bonell C (2016) Realist complex intervention science: applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. *Evaluation & Research in Education* 22, 286–303.
- Ford T, Sayal K and Meltzer H and Goodman R (2005) Parental concerns about their child's emotions and behaviour and referral to specialist services: general population survey. *Bmj* December 15, 1435–1436.
- Ford T, Hamilton H, Meltzer H and Goodman R (2007) Child mental health is everybody's business: the prevalence of contact with public sector services by type of disorder among British school children in a three-year period. *Child and Adolescent Mental Health* 12, 13–20.
- Ford T, Hamilton H, Meltzer H and Goodman R (2008) Predictors of service use for mental health problems among British schoolchildren. *Child and Adolescent Mental Health* 13, 32–40.
- Forness SR, Cluett SE, Ramey CT, Ramey SL, Zima BT, Hsu C, Kavale KA and MacMillan DL (1998) Special education identification of head start children with emotional and behavioral disorders in second grade. *Journal of Emotional and Behavioral Disorders* 6, 194–204.
- Garland AF (1995) Pathways to adolescent mental health services: adolescent help-seeking and teacher identification and referral. In *Faculty of the Graduate School*. US: Yale University.
- Girio-Herrera E, Owens JS and Langberg JM (2013) Perceived barriers to help-seeking among parents of at-risk kindergarteners in rural communities. *Journal of Clinical Child and Adolescent Psychology* 42, 68–77.
- Gould MS, Marrocco FA, Hoagwood K, Kleinman M, Amakawa L and Altschuler E (2009) Service use by at-risk youths after school-based suicide screening. *Journal of the American Academy of Child and Adolescent Psychiatry* 48, 1193–1201.
- Green H, McGinnity A, Meltzer H, Ford T and Goodman R (2005) *Mental Health of Children and Young People in Great Britain, 2004*. Palgrave Macmillan: Basingstoke.
- Halfors D, Cho H, Brodish PH, Flewelling R and Khatapoush S (2006) Identifying high school students "at risk" for substance use and other behavioral problems: implications for prevention. *Substance Use and Misuse* 41, 1–15.
- Hilt LM, Tuschner RF, Salentine C, Torcasso G and Nelson KR (2018) Development and initial psychometrics of a school-based screening program to prevent adolescent suicide. *Practice Innovations* 3, 1–17.
- Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, Altman DG, Barbour V, Macdonald H, Johnston M and Lamb SE (2014) Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj* 348, g1687.
- Howarth E, Devers K, Moore G, O' Cathain A and Dixon-Woods M (2016) Contextual issues and qualitative research. In *Challenges, Solutions and Future Directions in the Evaluation of Service Innovations in Health Care and Public Health*. *Health Services and Delivery Research* 4(16), pp. 1–36.
- Husky MM, Kaplan A, McGuire L, Flynn L, Chrostowski C and Olfson M (2011) Identifying adolescents at risk through voluntary school-based mental health screening. *Journal of Adolescence* 34, 505–511.
- Joint Commissioning Panel for Mental Health (2013) Guidance for commissioners of children and adolescent mental health services.
- Jones P (2013) Adult mental health disorders and their age at onset. *The British Journal of Psychiatry* 202, s5–s10.
- Jones D, Dodge KA, Foster EM and Nix R (2002) Early identification of children at risk for costly mental health service use. *Prevention Science: the Official Journal of the Society for Prevention Research* 3, 247–256.
- Jones DE, Greenberg M and Crowley M (2015) Early social-emotional functioning and public health: the relationship between kindergarten social competence and future wellness. *American Journal of Public Health* 105, 2283–2290.
- Kieling RR, Kieling C, Aguiar AP, Costa AC, Dorneles BV and Rohde LA (2014) Searching for the best approach to assess teachers' perception of inattention and hyperactivity problems at school. *European Child & Adolescent Psychiatry* 23, 451–459.
- Kilgus S, Taylor C, Van MW and Sims W (2018) Diagnostic accuracy of a universal screening multiple gating procedure: a replication study. *School Psychology Quarterly*.
- Knapp M and Iemmi V (2013) The economic case for better mental health. In *Annual Report of the Chief Medical Officer 2013, Public Mental Health Priorities: Investing in the Evidence*. Department of Health, pp. 147–156.
- Lyon AR, Maras MA, Pate CM, Igusa T and Vander Stoep A (2015) Modeling the impact of school-based universal depression screening on additional service capacity needs: a system dynamics approach. *Administration and Policy in Mental Health and Mental Health Services Research* 43, 168–188.
- Maggin DM and Johnson AH (2015) The reporting of core program components: an overlooked barrier for moving research into practice. *Preventing school failure. Alternative Education for Children and Youth* 59, 73–82.
- McLaughlin RJ, Holcomb JD, Jibaja-Rusth ML and Webb JA (1993) Teacher ratings of student risk for substance use as a function of specialized training. *Journal of Drug Education* 23, 83–95.
- Mitchell SG, Gryczynski J, Gonzales A, Moseley A, Peterson T, O'Grady KE and Schwartz RP (2012) Screening, brief intervention, and referral to treatment (SBIRT) for substance use in a school-based program: services and outcomes. *The American Journal on Addictions/American Academy of Psychiatrists in Alcoholism And Addictions* 21, S5–S13.
- Morey ME, Arora P and Stark KD (2015) Multiple-stage screening of youth depression in schools. *Psychology in the Schools* 52, 800–814.
- Naser S (2014) *Evaluating the Utility of the Behavioral and Emotional Screening System (BESS) as a School-Based Universal Screening Tool*. Tulane University School of Science and Engineering.
- Nemeroff R, Levitt JM, Faul L, Wonpat-Borja A, Bufford S, Setterberg S and Jensen PS (2008) Establishing ongoing, early identification programs for mental health problems in our schools: a feasibility study. *Journal of the American Academy of Child & Adolescent Psychiatry* 47, 328–338.
- Oakley A, Strange V, Bonell C, Allen E and Stephenson J (2006) Process evaluation in randomised controlled trials of complex interventions. *Bmj* 332, 413–416.
- Ollendick TH, Greene RW, Weist MD and Oswald DP (1990) The predictive validity of teacher nominations: a five-year followup of at-risk youth. *Journal of Abnormal Child Psychology* 18, 699–713.
- Popay J, Roberts H, Sowden AJ, Petticrew M, Arai L, Rodgers M, Britten N, Roen K and Duffy S (2006) Guidance on the conduct of narrative synthesis in systematic reviews. *A Product from the ESRC Methods Programme Version 1*, b92.

- Robinson J, Gook S, Yuen HP, Hughes A, Dodd S, Bapat S, Schwass W, McGorry P and Yung AR** (2010) Depression education and identification in schools: an Australian-based study. *School Mental Health* **2**, 13–22.
- Robinson J, Yuen HP, Martin C, Hughes A, Baksheev GN, Dodd S, Bapat S, Schwass W, McGorry P and Yung AR** (2011) Does screening high school students for psychological distress, deliberate self-harm, or suicidal ideation cause distress - and is it acceptable? *Crisis* **32**, 254–263.
- Sayal K, Owen V, White K, Merrell C, Tymms P and Taylor E** (2010) Impact of early school-based screening and intervention programs for ADHD on children's outcomes and access to services: follow-up of a school-based trial at age 10 years. *Archives of Pediatrics & Adolescent Medicine* **164**, 462–469, 8p.
- Scott MA, Wilcox HC, Schonfeld IS, Davies M, Hicks RC, Turner JB and Shaffer D** (2009) School-based screening to identify at-risk students not already known to school professionals: the Columbia suicide screen. *American Journal of Public Health* **99**, 334–339.
- Sweeney C, Warner CM, Brice C, Stewart C, Ryan J, Loeb KL and McGrath RE** (2015) Identification of social anxiety in schools: the utility of a two-step screening process. *Contemporary School Psychology* **19**, 268–275.
- Tisher M** (1995) Teachers' assessments of prepubertal childhood depression. *Australian Journal of Psychology* **47**, 93–96.
- Tyne TF and Flynn JT** (1981) Teacher nominations and peer evaluations in the identification of socioemotional at-risk students. *Exceptional Children* **48**, 66–68.
- Wells M, Williams B, Treweek S, Coyle J and Taylor J** (2012) Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials* **13**, 95.
- Whitney SD, Renner LM, Pate CM and Jacobs KA** (2011) Principals' perceptions of benefits and barriers to school-based suicide prevention programs. *Children and Youth Services Review* **33**, 869–877.