Screening efficiency of the child behavior checklist and strengths and difficulties questionnaire: a systematic review
Warnick EM, Bracken MB, Kasl S

CRD summary
This review concluded that the Child Behaviour Checklist and Strengths and Difficulties Questionnaire were efficient scales for the identification of psychiatric disorders in young people. The review suffered from a number of limitations in terms of literature search, review methods and analysis and reporting of results. The conclusions were not supported by the results presented.

Authors' objectives
To determine the accuracy of the caretaker-report Child Behaviour Checklist and Strengths and Difficulties Questionnaire for the detection of psychiatric disorders in young people in community and clinical settings.

Searching
MEDLINE, PsycINFO, EMBASE, CINAHL and Cochrane Central Register of Controlled Trials were searched from inception to November 2006. Search terms were not reported. References of retrieved studies were screened. Only full-text English-language publications were included. Abstracts, book chapters, dissertations and theses were excluded.

Study selection
Studies that assessed the accuracy of the Child Behaviour Checklist or Strengths and Difficulties Questionnaire in young people (age five to 18 years) who sought mental health services at a mental health clinic or who were part of community-based research samples were eligible for inclusion. Questionnaires had to be completed by the young person's care taker. Studies had to include either a structured interview or clinician-based diagnosis as the reference standard. Studies of children with learning disabilities, mental retardation, pervasive developmental disorders or dementia were excluded. It appeared but was not explicitly stated that studies had to report sufficient data to allow calculation of sensitivity and specificity.

The age of children in the included studies ranged from five to 18 years. The proportion of boys ranged from 45% to 100%. Some studies provided data on overall problems and others looked at internalising and/or externalising behaviours. Studies used thresholds that ranged from 60 to 70 on the Child Behaviour Checklist scale to define disordered children. Thresholds were poorly defined for studies of the Strengths and Difficulties Questionnaire. Studies were conducted in community and clinic settings. A variety of reference standards were used, most commonly Diagnostic Interview Schedule for Children and criteria from the Diagnostic and Statistical Manual.

One reviewer screened abstracts for inclusion. A random sample of studies was checked by a second reviewer.

Assessment of study quality
Studies were assessed for methodological quality according to the following criteria: index test and reference standard administered in same manner regardless of diagnostic status; blind comparison to the reference standard; use of an appropriate group of patients; and previous validation in independent sample of patients. The authors did not state how many reviewers performed the quality assessment.

Data extraction
Data were extracted on sensitivity, specificity and likelihood ratios, together with their 95% confidence intervals (CIs). If these were not reported, the required data were calculated by the primary reviewer. Data were extracted by one reviewer. A random sample was checked by a second reviewer.

Methods of synthesis
Positive likelihood ratios were pooled using fixed-effect and random-effects models. Data on pooled sensitivity and specificity were also reported; it was unclear how these were calculated. Data were pooled separately for each scale and
for the diagnostic domains addressed (overall presence of disorder, disruptive behaviours, depression/anxiety and attention deficit hyperactivity disorder /inattention). Heterogeneity was assessed visually from forest plots and statistically from Q and I² statistics. Subgroup analysis determined the accuracy of the scales in different settings and demographic groups based on the following criteria: gender, age (<11 years versus 11+ years), race/ethnicity and setting (clinical compared with community).

Results of the review
Thirty two studies were included in the review: 29 assessed the Child Behaviour Checklist (n=25,006); and three assessed the Strengths and Difficulties Questionnaire (n=10,424). Participation rates ranged from 30% to 95%. Most studies collected scale and reference standard results for all children. In most studies Child Behaviour Checklist scores and comparison diagnoses were generated blindly. In nine studies of the Child Behaviour Checklist, data were collected via a multi-stage screening process and it was unclear whether Child Behaviour Checklist results were known when making the diagnosis.

The Child Behaviour Checklist had a sensitivity of 66% (95% CI 60% to 73%) and specificity of 83% (95% CI 81% to 85%) for total problems. Pooled sensitivity was similar for mood/anxiety (59%) and disruptive behaviour (61%) and higher for attention (71%). Estimates of specificity were similar across the diagnostic domains (72% to 79%).

The Strengths and Difficulties Questionnaire was less sensitive (pooled sensitivity 49%, 95% CI 46% to 51%), but more specific (pooled specificity 93%, 95% CI 92% to 94%) than the Child Behaviour Checklist. Sensitivity was lower for mood/anxiety (54%) than for disruptive behaviour (75%) or attention (73%). Estimates of specificity were similar across the diagnostic domains (85% to 91%).

Results of the sensitivity analysis were reported for setting, but insufficient data were reported to allow a meaningful interpretation of results. There was insufficient data to conduct any other of the proposed subgroup analyses. Data on heterogeneity were only presented for likelihood ratios, which were generally found to be homogeneous, but the analysis focused on positive likelihood ratios with no data on negative likelihood ratios and so were not included here.

Authors' conclusions
Both scales were efficient tools for the identification of psychiatric disorders in young people.

CRD commentary
The review addressed a focused question. Inclusion criteria were defined in terms of population, index test and reference standard. It appeared that some restriction was made based on outcome data, but this was not clearly stated. The literature search appeared adequate for published papers, but details of the search strategy (and whether this incorporated a diagnostic filter) were not reported. The review was restricted to published English-language studies, so there was a possibility of language and publication biases. No appropriate steps were taken to minimise bias and errors in the selection of studies or extraction of data. Study quality was assessed using some appropriate criteria, but the results were poorly reported and not considered in the analysis. The analysis and reporting of results suffered from a number of limitations. The text focused on reporting pooled positive likelihood ratios. (Likelihood ratios can be a helpful method of understanding results, but if such measures are used, both positive and negative likelihood ratios should be reported and these should be calculated from summary sensitivity and specificity, rather than pooling likelihood ratios reported in individual papers.) A secondary analysis based on sensitivity and specificity was reported in tables, but it was unclear how these data were pooled and so these should be interpreted with caution. Some subgroup analysis was conducted, but the results were reported only in terms of ranges in positive likelihood ratios, which made them impossible to interpret. The accuracy of the scales reported in the review did not appear sufficiently high to justify the authors' conclusion that these were efficient tools. Limitations in the literature search, review methodology and analysis and reporting of results meant that the findings of this review should be interpreted with extreme caution.

Implications of the review for practice and research
Practice: The authors stated that this review supported the use of the Child Behaviour Checklist and Strengths and Difficulties Questionnaire via caretaker-report in clinical and community samples.

Research: The authors stated that additional research was needed to determine if there was a difference between the
two scales.

**Funding**
Not stated.

**Bibliographic details**

**Original Paper URL**
http://onlinelibrary.wiley.com/journal/120125959/abstract

**Indexing Status**
Subject indexing assigned by CRD

**MeSH**
Child; Child Behavior Disorders /diagnosis; Humans; Questionnaires /standards; Sensitivity and Specificity

**AccessionNumber**
12009100364

**Date bibliographic record published**
07/04/2009

**Date abstract record published**
04/11/2009

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.