Assessment of Depression in Children and Adolescents

Laurie B. Gray, MD, Abigail Dubin-Rhodin, Ronald A. Weller, MD, and Elizabeth B. Weller, MD

Corresponding author
Laurie B. Gray, MD
University of Pennsylvania, 3440 Market Street, Suite 200, Philadelphia, PA 19104, USA.
E-mail: laurie.gray@uphs.upenn.edu

Current Psychiatry Reports 2009, 11:106–113
Current Medicine Group LLC ISSN 1523-3812
Copyright © 2009 by Current Medicine Group LLC

Depression assessment instruments are valuable tools in the treatment of children and adolescents. Available instruments include diagnostic interviews, self-administered rating scales, and observer-rated scales. To select an appropriate instrument, the user must define the goal of the assessment and then identify instruments with the properties that match this goal. This article discusses how to choose an assessment instrument and gives an overview of currently available depression assessment instruments. Important considerations include how and by whom an instrument is administered, what kind of data are obtained by the instrument, and the instrument’s validity and reliability.

Introduction

Properties of assessment instruments
Other factors also should be considered in choosing an assessment tool. These include how and by whom an instrument is administered, what kind of data are obtained by the instrument, and the instrument’s validity and reliability.

Administration
Important considerations include who administers the instrument and how long it takes to complete. Self-administered assessments require the least amount of labor for the professional. However, live interviewers can mitigate problems with literacy or motivation. Instruments that require clinical judgment (such as evaluation of affect) may provide the richest information, but they require an experienced clinician.

Data
Important questions to ask regarding data generated by an instrument include the following: Can the instrument diagnose depression using DSM-IV criteria? Are there normative data for the instrument? Are items scored as positive or negative, or does the instrument measure a range of severity? An instrument that measures a range of severity may be necessary to evaluate changes in symptoms over time.

Validity and reliability
A detailed review of reliability and validity is available elsewhere [1]. To be used in children and adolescents, instruments must have established reliability and validity in children and adolescents. It is not appropriate to use instruments studied only in adults.

Reliability refers to interrater reliability, which measures whether different raters achieve similar results, and test–retest reliability, which measures whether an instrument administered twice to the same individual produces similar results. Validity includes content, concurrent, and discriminant validity. Content validity refers to the extent to which the instrument measures a certain construct. As there is no gold standard laboratory test or radiologic study that defines depression, content validity in depression assessments is theory driven and thus is subject to interpretation and revision. To gather information about content validity, researchers
analyze face validity, which refers to whether an instrument appears to measure what it claims to measure, and internal consistency, which measures whether the individual items in an assessment correlate with the primary result of the assessment. Ultimately, the best way to examine content validity is to conduct more research and to improve on the theory guiding the construct. Concurrent validity measures whether the results of an assessment correlate with a related construct. For example, a depression assessment should correlate with global assessment of functioning, and both scores should show improvement after interventions with known efficacy. Discriminant validity measures whether an assessment instrument can discriminate between two different groups (eg, between children in the community and children seen in a depression clinic). Unfortunately, many instruments fail to discriminate depression from anxiety or general psychological stress. Depending on the instrument’s goals, this can be a major or minor problem.

Selecting an assessment instrument
In this review, we discuss three classes of assessment instruments: structured or semistructured interviews, self-administered rating scales, and observer-rated scales.

Diagnostic Interviews
Data
Diagnostic interviews systematically collect information by having an interviewer ask questions and then record an individual’s answers. The most useful interviews (and all the interviews reviewed here) apply DSM criteria to make diagnoses. However, such diagnoses should be considered preliminary. Although most interviews attempt to consider severity and time course in determining whether a syndrome meets DSM criteria, these interviews do not replace a thorough evaluation by an experienced clinician.

Administration
The difference between structured and semistructured interviews relates to the type of interchanges allowed between the interviewer and the informant. Less structure allows interviewers to deviate from the “script,” develop their own probes, and clarify questions. More structure requires stricter adherence to the script. Although the degree of structure should be considered on a continuum, in general, the Diagnostic Interview Schedule for Children (DISC), the Children’s Interview for Psychiatric Syndromes (ChIPS), and the Diagnostic Interview for Children and Adolescents (DICA) are more structured, whereas the Child and Adolescent Psychiatric Assessment (CAPA) and the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS) are less structured. Benefits of a less structured design are the potential for gathering richer and more accurate information, as interviewers are allowed to engage in more of a dialogue with the informants. Benefits of a more structured design are that less training and experience are required of the interviewer (Table 1), and the interviewer has less of a chance of influencing the informant or misinterpreting the informant’s response.

<table>
<thead>
<tr>
<th>Interview</th>
<th>Age of informant</th>
<th>Time period covered</th>
<th>Interviewer characteristics (training)</th>
<th>Administration time</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPA [7]</td>
<td>9–17 y (versions for ages 18 y and older and 3–6 y also available)</td>
<td>Past 3 mo</td>
<td>Bachelor’s degree (1–2 wk classroom study and 1–2 wk practice)</td>
<td>About 60 min per informant + 65 min for coding and data entry</td>
</tr>
<tr>
<td>DICA [6]</td>
<td>Child version: 6–12 y; adolescent version: 13–18 y</td>
<td>Lifetime</td>
<td>Trained lay interviewers (2–4 wk training)</td>
<td>1–2 h total</td>
</tr>
<tr>
<td>K-SADS [9]</td>
<td>6–18 y</td>
<td>K-SADS-E: lifetime and current; K-SADS-P/L: lifetime and current; K-SADS-P: past 12 mo and current</td>
<td>Trained clinician</td>
<td>2.5–3 h total</td>
</tr>
<tr>
<td>ChIPS [14,39]</td>
<td>6–18 y</td>
<td>Not specified—if syndrome is endorsed, informants are questioned for timing based on DSM criteria</td>
<td>Trained clinician, familiar with DSM and/or familiar with ChIPS</td>
<td>49 min for inpatients, 30 min for outpatients, 21 min for community sample</td>
</tr>
<tr>
<td>DISC [4]</td>
<td>Parents of youth 6–17 y; youth 9–17 y</td>
<td>Past 12 mo and past 4 wk + optional whole-life module</td>
<td>Trained lay interviewer (2–3 d training)</td>
<td>70 min per informant</td>
</tr>
</tbody>
</table>

CAPA—Child and Adolescent Psychiatric Assessment; ChIPS—Children’s Interview for Psychiatric Syndromes; DICA—Diagnostic Interview for Children and Adolescents; DISC—Diagnostic Interview Schedule for Children; K-SADS—Schedule for Affective Disorders and Schizophrenia for School-Age Children (E—epidemiologic; P—present state; P/L—present and lifetime).