Evidence-Based Assessment of Attention-Deficit/Hyperactivity Disorder: Using Multiple Sources of Information

THOMAS W. FRAZIER, PH.D., AND ERIC A. YOUNGSTROM, PH.D.

The present column illustrates step-by-step the process of acquiring and integrating information according to the recommendations of evidence-based practices (Kraemer, 1992; Sackett et al., 2000). A case example models the process, leading to specific recommendations regarding instruments and strategies for evidence-based assessment (EBA) of attention-deficit/hyperactivity disorder, combined type (ADHD-C). The diagnostic process presented here is consistent with EBA as defined in the text for this series by Guyatt and Rennie (2002). In addition, the column shows how to convert scores using normative data and effect sizes (see Morgan et al., 2005 for information on effect sizes) into useful diagnostic information about individual cases. Because ADHD, predominantly inattentive type is thought by some experts to be a very different disorder (Milich et al., 2001), with a distinct cognitive profile (Hinshaw et al., 2002), this article is focused on ADHD-C.

POLLY (A COMPOSITE CASE)

Polly, age 11, was referred for evaluation to an outpatient neuropsychological service for problems with concentration and “out of control” behavior.

Her parents were concerned that Polly’s behavior may represent ADHD because her older brother was recently diagnosed with ADHD-C. According to the referral source, her parents reported “excessive motor activity” and difficulty focusing on her homework. Her teachers have often commented on her difficulties with concentration. However, the parents would prefer she not receive medication unless she “has ADHD.”

Table 1 presents recommendations for initiating an evidence-based approach to diagnostic assessment of ADHD. It is possible to have either too little or too much assessment. If there is too little assessment, then treatment decisions are made with inadequate information, increasing the risk of misdiagnosis. However, too much assessment adds time and expense from redundant or unnecessary tests and actually may result in less accurate clinical decisions (Kraemer, 1992). In an evidence-based framework, the test/no test threshold defines the probability of the disorder at which further evaluation is recommended and provides a rational approach for determining when to collect additional assessment information. The treatment threshold defines the probability of the disorder at which treatment is recommended (Guyatt and Rennie, 2002).

What information have we collected already (step 1)? First, we know that the parents complained of Polly’s difficulties with concentration and excessive motor activity. The parents’ remarks suggest that Polly is likely exhibiting symptoms from both the inattention symptom cluster and from the hyperactivity/impulsivity cluster.

Second, the parents report a positive family history of ADHD-C. How much does a first-degree relative with ADHD increase Polly’s probability of having ADHD? To answer this question, we first need to determine the initial risk that Polly has ADHD, namely, the base rate of ADHD in this clinical setting.
<table>
<thead>
<tr>
<th>Step</th>
<th>How Implemented</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identify whether presentation is consistent with ADHD</td>
<td>Are there symptoms of inattention, hyperactivity, or impulsivity as part of the chief complaint? What are the most difficult differential diagnoses? – Bipolar disorder (mood fluctuation and irritability) or other neurological disorders (e.g., traumatic brain injury)</td>
</tr>
<tr>
<td>2. Establish local base rate of ADHD</td>
<td>Identify published rate with most similar sample in terms of participant characteristics and referral/recruitment patterns Directly estimate from own medical records (but carefully consider threats to the validity of these diagnoses and any unique aspects of specific clinical setting)</td>
</tr>
<tr>
<td>3. Determine whether other risk factors such as family history are useful</td>
<td>MEDLINE or PsycINFO search: “attention-deficit-disorder-with-hyperactivity” AND “risk factor” (MeSH terms) Look for published likelihood ratios</td>
</tr>
<tr>
<td>4. Find likelihood ratios associated with risk factor</td>
<td>Convert sensitivity and specificity into likelihood ratios for positive and negative test results Convert effect sizes into likelihood ratio estimates Convert normative data for test into sensitivity and specificity estimates (which in turn can yield a likelihood ratio): Norms for the ADHD group can yield sensitivity (what percentage of ADHD sample scores this high or higher? Subtract the percentile for the next lower score from 100% to obtain the sensitivity) Norms for the non-ADHD comparison group can yield the specificity (e.g., what percentage of the non-ADHD sample scores this high or higher? Subtract the percentile for the next lower score from 100% to obtain the specificity)</td>
</tr>
<tr>
<td>5. Identify relevant parent report instruments</td>
<td>MEDLINE or PsycINFO search: “attention-deficit-disorder-with-hyperactivity” AND “sensitivity and specificity” (MeSH term) AND “parent report OR subjective report”</td>
</tr>
<tr>
<td>6. Compare relevant instruments and identify likelihood ratio</td>
<td>Compare published information to existing criteria for evaluating the quality of the evidence (Bossuyt et al., 2003) Convert sensitivity and specificity(^a) or well-established effect size(^b) into estimate of likelihood ratio for positive and negative test results</td>
</tr>
<tr>
<td>7. Identify and compare teacher report instruments</td>
<td>MEDLINE or PsycINFO search: “attention-deficit-disorder-with-hyperactivity” AND “teacher-report” Comparisons should be done using the same criteria as above for parent report instruments</td>
</tr>
<tr>
<td>8. Use nomogram or spreadsheet formula</td>
<td>Keep copies of nomogram or spreadsheet file at offices Consider “premarking” nomogram or include information in spreadsheet to indicate base rate estimates for ADHD at your setting Consider “premarking” nomogram or include information in spreadsheet to indicate likelihood ratios associated with risk factors (family history) or test scores</td>
</tr>
<tr>
<td>9. Regularly review and update tool kit</td>
<td>Periodically repeat MEDLINE searches (or PsycINFO) “ADHD or ADD or attention-deficit-disorder-with-hyperactivity” AND “sensitivity and specificity” “ADHD or ADD or attention-deficit-disorder-with-hyperactivity” AND “risk factors or family history” Adopt new tests with better norms, better criterion diagnosis, and/or better diagnostic efficiency</td>
</tr>
</tbody>
</table>

*Note: ADHD = attention-deficit/hyperactivity disorder.
\(^a\) The likelihood ratio associated with a positive test result is equal to sensitivity/(false alarm rate), or sensitivity/(1 − specificity). The likelihood ratio for a negative test result is equal to (1 − sensitivity)/specificity. See Guyatt and Rennie (2002) for more details.
\(^b\) Estimation of likelihood ratios from well-established effect sizes was described in the text. However, this approach should be seen as a last resort if no sensitivity and specificity data or published likelihood ratios are available.
BASE RATE OF ADHD (Step 2)

Base rate estimates of ADHD vary widely depending on setting and methodology. Epidemiological estimates of ADHD have ranged from 3% to 12% (American Academy of Pediatrics, 2000; American Psychiatric Association, 1994). For this reason, a practitioner may calculate the base rate of ADHD in his or her setting using data collected for previous patients. However, local idiosyncratic diagnostic conceptualizations and referral patterns can both be issues. For this reason, clinicians should compare local base rate estimates to published estimates from similar settings. If widely discrepant results are obtained, then a solid rationale for the discrepancy, such as referral bias, is needed before using the calculated local base rate as the first step in a diagnostic process.

Using the same search strategy as described in the column concerning pediatric bipolar disorder (Youngstrom and Duax, 2005) and the search terms “ADHD” AND “prevalence OR base rate OR epidemiological” AND “review” yields 39 hits in Medline and 58 in PsychINFO, with more than 20 estimates of ADHD prevalence. The reported base rates range from 0.03 to 0.74. The higher estimates are from specialized ADHD clinics and the lower estimates are from epidemiological studies; none of the reported rates are from specialized neuropsychology clinics. In the absence of reported base rates from neuropsychology clinics, let us use data from our own clinic. The calculated local base rate is 0.35. This high rate makes sense because outpatient psychological and neuropsychological clinics have enriched samples as a result of referral sources weeding out unlikely cases.

However, the confidence with which we can use this local base rate is lower than if we had found strong support in the literature for a 0.35 base rate in clinical settings. To show the effect of changes in base rate on the final result, let us also do two separate calculations. First, we can use 0.03, the low end of epidemiological estimates, as a base rate. Second, we can use 0.19, midway between 0.03 and 0.35.

COMBINING THE BASE RATE AND FAMILY HISTORY INFORMATION (Steps 3 and 4)

To combine base rate and family history information, we first determine the relative value of family history of ADHD for altering the probability of diagnosis (steps 3 and 4). To do this, we perform a similar search to that described above using the terms “ADHD” AND “family history OR offspring.” The search generates 128 hits in PsychINFO and 72 in Medline. From this search, we focus on data from two recent, large sample studies of male and female ADHD probands (Faraone et al., 2000a, b). Large sample studies are likely to provide more stable estimates of the impact of familial risk, and recent studies are less susceptible to changes in diagnostic definitions and assessment instruments over time. The results of these studies indicate an approximate four- to fivefold increase in the likelihood of ADHD when a first-degree relative has ADHD. Using a nomogram (Fig. 1), we locate the base rate of 0.35 in the left column and use a ruler to line up this estimate with the likelihood ratio of 4.0 (derived as a conservative estimate of the risk associated with positive family history of ADHD in a first-degree relative) in the second column. This procedure yields a revised probability of approximately 0.65, read from the third column (actual probability using the formula of 0.68). The base rate has a substantial impact on the posterior probability. For example, if we use 3%, the epidemiological rate, as the base rate, then the same family history is associated with a probability of only 0.11 that the child has ADHD. The probability increases to 48% if the base rate started at 19%. These posterior probability estimates are positive.
predictive values, indicating the likelihood that an individual has ADHD. Note that the difference in base rates has led to substantial differences in posterior probability. Furthermore, we suggest that none of these posterior probability values justify treatment because the parents said they wanted to be “confident” of the diagnosis. Therefore, let us set the treatment threshold at 0.90 to represent the parents’ perspectives concerning risks and benefits of treatment. In most cases, the treatment threshold may be set somewhat lower because of the minimal risk of treatment (American Academy of Pediatrics, 2000) and the fact that a stimulant medication trial, unlike an operation, is readily reversible. The posterior probabilities are also far too high for most clinicians or parents to consider ADHD “ruled out.” In other words, the risk of ADHD is too high to ignore, but not high enough for these parents to warrant initiation of treatment without gathering further diagnostic information. What do we do now? What other data can lower the probability of diagnosis below the test/no test threshold or raise the probability above the treatment threshold?

**PARENT REPORT (Steps 5 and 6)**

In consensus statements (American Academy of Child and Adolescent Psychiatry, 2002; American Academy of Pediatrics, 2000), the use of ADHD-specific parent rating scales is recommended when evaluating children for ADHD. Thus, the next steps (steps 5, 6, and 7) of this EBA are to locate, compare, and use information from relevant parent report instruments. The consensus statements’ strong assertion that parent report measures are useful for making a diagnosis of ADHD is based on a comparison of several parent report EBA and ADHD diagnosis 4 instruments and the extremely high validity for the discrimination of ADHD from non-ADHD children reported for most of these instruments. For example, the Conners Parent Rating Scale-Revised: Long Form (CPRSR:L).

The ADHD subscale has been found to have an extremely large effect size \(d = 2.3–3.1\), depending on the data set) in discriminating ADHD from non-ADHD children in a community sample. This effect size indicates almost no overlap between ADHD and control groups.

Unfortunately, such huge effect sizes are unrealistically large in our situation with Polly, in which we need to distinguish referred individuals who actually have ADHD from referred individuals who do not. Fortunately, the Conners manual also reports the results for a comparison group of individuals who were labeled by a psychologist or psychiatrist as having emotional problems. The effect size comparing the true ADHD group and the emotional problems group is much more modest \(d = 0.67\). Note that for most clinical decisions, the important discrimination is not between healthy nonreferred individuals and individuals with ADHD, but rather between patients referred with other emotional, cognitive, or behavioral problems and patients who meet criteria for ADHD.

We therefore use the more modest and clinically realistic effect size of 0.67 to calculate the likelihood ratio for Polly. Her parents rated her 2.0 SD above the mean of the normal control group on the CPRS-R:L ADHD subscale based on age- and sex-adjusted norms \(T = 70; T\) scores have SD = 10\). This score puts her approximately 0.3 SD below the mean of the clinical ADHD group \(mean T\) score = 73\) and 0.4 SD above the mean of the emotional problems control group \(emotional\ control\ mean\ T\) score = 66\). The test norms can be used to estimate the sensitivity and specificity associated with Polly’s score.

The sensitivity of a test at a defined score is the percentage of cases with ADHD obtaining that score or higher. According to the Conners parent norms, a score of 70 is lower than average for ADHD cases. Thus, at this score, it is relatively easy for ADHD cases to clear this threshold. For example, if the parents of 100 children with ADHD completed the test, then test norms indicate that a mean of 63 cases would obtain a score of 70 or higher, resulting in a sensitivity of 63\%. Similarly, if one looked at scores for 100 cases with emotional problems, then only 34 would score at or above \(T = 70\), and 66 would score below, yielding a specificity of 66\%. Some manuals report the percentiles associated with each score; otherwise percentiles can be estimated by converting to standard \(Z\) scores and estimating the percentage of the normal curve falling beyond the test score. Figure 2 illustrates this process visually. Using this example, the score of \(T = 70\) defines both the sensitivity and specificity of the test, and these numbers can be used to compute the likelihood ratio \([LR = sensitivity/(1 - specificity)]\). Using these values, we estimate a likelihood ratio of 1.85 = (63/100)/(34/100). This likelihood ratio indicates that scores of 70 or higher are nearly twice as likely to be seen in
youths with ADHD versus individuals with emotional problems; scores in this range almost double the odds that a given case has ADHD. These calculations are available on the Journal’s Web site (www.jaacap.com) via the Article Plus feature.

Using the nomogram to combine the posterior probability of the previous step (0.68) with the likelihood ratio for the CPARS-ADHD index, we find that the new posterior probability is 0.80. For the two lower base rate estimates (0.03 and 0.19), posterior probability values become 0.19 and 0.63. Although this approach is clearly inferior to more sophisticated statistical approaches (e.g., directly deriving nonparametric estimates of likelihood ratios or sensitivity and specificity values), this approach is far simpler and more feasible in a clinical setting. Yet it is an improvement over simply assuming that a test result is equivalent to a diagnosis (e.g., beginning a stimulant trial with all of the cases scoring a 70 or above on a Conners scale). It also is an improvement over applying inconsistent and unreliable approaches to patient information derived from unstructured interviews.

At this point, Polly’s results indicate that, if our original base rate estimate is correct, we may be at a point at which we can discontinue gathering information and begin treatment because her probability of ADHD is near the treatment threshold. However, Polly’s parents wanted as much confidence in the diagnosis as possible, and we ourselves want to make sure that our original base rate estimate is not exaggerating her risk. Therefore, we decide to gather additional information.

TEACHER REPORT (Step 7)

Teacher report is recommended as a complement to parent report in the assessment of ADHD (Power et al., 2001). The advantages of teacher report are that it facilitates identification of symptoms and impairment in two settings, provides more detailed information about behaviors that may be more reliably observed in a school setting, and also permits the clinician to examine the influence of reporter biases operating in subjective report ratings (Pelham et al., 2005). Furthermore, teachers often have broad experience with comparing one child to another and observe the child in a high-demand setting. In contrast, using multiple subscales from a single information source is less likely to be useful because of the often high correlations between subscales. For example, the CPRS-R:L ADHD and Cognitive Problems subscales correlate highly ($r = 0.86$), whereas the CPRS-R:L and CTRS-R:L ADHD subscales correlate more modestly ($r = 0.49$).

Using both parent and teacher checklists provides a clinical shortcut because they obtain diagnostic information highly similar to structured interviews (for a detailed review, see Pelham et al., 2005). Specifically, studies of incremental validity have indicated that structured interviews are not likely to provide incremental validity to parent and teacher reports in the detection of ADHD (Power et al., 2001; Wolraich et al., 2003). Thus, checklists are useful initial measures because they are easily obtained, require little clinician time, and connect closely to results obtained from structured interviews.

Teacher report via brief ADHD-specific rating scales has been shown to have good discrimination between ADHD and non-ADHD groups (American Academy of Pediatrics, 2000). In fact, the Conners Teacher Rating Scale Revised: Long Form ADHD subscale shows similarly large separation of ADHD and normal control groups ($d = 1.7–3.4$, depending on the data set) to that seen with the CPRS-R:L. However, as was seen with the CPRS-R:L, less discrimination is achieved when individuals with ADHD are compared to individuals with emotional problems ($d = 0.64$), the more interesting clinical comparison. Using this information and Polly’s scores, we can compute a new posterior probability. Assuming that Polly scores 2 SD above the normal control mean ($T = 70$), data from the Conners’ manual (Conners, 1997) indicates that her score would be approximately 0.32 SD above the mean of the ADHD group (ADHD mean score $T = 67$) and 0.96 SD above the emotional problems group (emotional problems mean score $T = 61$).

Fig. 2 Obtaining diagnostic efficiency statistics from normative data and effect sizes.
Thus, if 100 individuals with ADHD and 100 individuals without ADHD were given the test, only approximately 17 individuals with emotional problems would have scored at or above the level of our patient and approximately 37 individuals with ADHD would have scored at or above this level. Using these numbers, one can easily compute the likelihood ratio for this particular score, 2.18 = [(37/100)/(17/100)]. Combining this likelihood ratio with the posterior probability derived by combining the base rate, family history, and parent report information, we find that it is 90% likely that the child has ADHD. Using the lower base rate estimates (0.03 and 0.19), we derive posterior probabilities of 0.34 and 0.80. However, it should be noted that these values overestimate the true posterior probability because of the moderate correlation between parent and teacher report. With this caveat in mind, these calculations suggest that for all but the lowest base rate estimate, the combination of family history and parent and teacher checklist data are compelling. A clinician may seriously consider initiating treatment, particularly if the treatment threshold is moderate or low and the child’s functioning is impaired.

**COMMENTARY**

In most cases, a Guyatt and Rennie (2002)–style EBA approach will result in considerable savings in clinician and patient evaluation time, eliminating the need for unnecessarily lengthy and expensive testing batteries (Kraemer, 1992). It will also provide a more accurate and informed diagnosis and/or basis for treatment initiation. At present, EBA of ADHD is hampered by the lack of published likelihood ratios or sensitivity and specificity data from large samples. Unfortunately, the paucity of this information has probably contributed to the extreme variability of clinical assessment practices, from full neuropsychological assessments to brief parent or patient interview. As the above examples indicate, it is possible for motivated practitioners to use published data to estimate likelihood ratios for existing tests. The effort invested to put together an EBA approach will offer returns in all future cases in which ADHD may be a treatment concern. Clearly, it would be helpful for researchers to publish likelihood ratios to facilitate use of these methods, which have become widespread in other areas of medicine (Guyatt and Rennie, 2002).

The above discussion should be viewed as one approach to EBA of ADHD. There are only a handful of studies reporting enough data to compute effect sizes comparing individuals with ADHD to individuals referred for emotional or psychiatric problems. It is possible that other information sources, such as cognitive test data, may provide better incremental validity when combined with base rate, family history, and parent-report data, although there is considerable disagreement about this issue (American Academy of Pediatrics, 2000; Pelham et al., 2005). We also note that family history data may not be useful or available in all cases, both because often the patient is the first to present with attention problems in the family and parents may have had undiagnosed ADHD for many years. Therefore, clinicians should view the steps described in Table 1 as a flexible guide and skip or add steps as needed, acquiring only the information that may be relevant to a particular case.

The focus of the next column is how best to measure Polly’s treatment response.

**REFERENCES**


**Disclosure:** Dr. Youngstrom is co-investigator on investigator-initiated research grants sponsored by Abbott Laboratories and AstraZeneca Pharmaceuticals; he is the statistical expert for both protocols. Dr. Youngstrom also consults with Otsuka Pharmaceuticals about assessment of pediatric bipolar disorder. Dr. Frazier has no financial relationships to disclose.
**Injury Prevention Advice in Top-Selling Parenting Books**

Wanda M. Hunter, MPH, Samah Helou, MPH, Gitanjali Saluja, MA, PhD; Carol W. Runyan, MPH, PhD; Tamera Coyne-Beasley, MD, MPH

*Objective:* Parenting books are a commonly used source of information on how to keep children and adolescents safe from injuries, the leading cause of death and disability for children aged 1 to 18 years. The content and the quality of the messages contained in these books have not been evaluated formally. The objective of this study was to determine the quantity and the quality of injury prevention messages contained in popular parenting books.

*Methods:* Top-selling parenting books for 2 major booksellers were reviewed to determine the presence and the accuracy of injury prevention messages as compared with those recommended by the American Academy of Pediatrics (AAP) through The Injury Prevention Program (TIPP) for younger children, aged 0 to 12 years, and the American Medical Association (AMA) through its Parent Package for the safety of adolescents.

*Results:* Forty-six parenting books were reviewed, including 41 with messages related to younger children and 19 with messages related to adolescents. These books varied widely with regard to the number of injury prevention messages included. Although some books covered the great majority of TIPP messages for parents of young children, others included very few. In the case of books that address safety for adolescents, no book had more than half of the messages recommended by the AMA. Prevention of burns and motor vehicle injury were the most commonly addressed injury prevention topics in the books focused on younger children, whereas gun safety was the most prevalent injury prevention topic in books that focused on adolescents. Books that were authored by physicians addressed more of the recommended topics and messages than books that were written by authors from other professional backgrounds. The quality of messages was good, i.e., consistent with the advice given by the AAP and the AMA. In only a few cases, the parenting books gave injury prevention advice that was inconsistent with recommendations.

*Conclusions:* Overall, books on parenting adolescents are less likely to contain injury prevention messages than those that address younger children. However, the most frequent injury prevention messages for parents of adolescents describe strategies to prevent firearm injury, a leading cause of death for children in this age group. More emphasis should be placed on prevention of motor vehicle injuries, especially as relates to adolescents. Pediatricians and primary care physicians need to be aware of the strengths and weaknesses of parenting manuals in providing adequate guidance related to injury prevention. *Pediatrics* 2005;116: 1080–1088.