7 Rural America and Evidence-Based Assessment: The Potential for a Happy Marriage

Alex Kirk, Rafaella Sale, and Eric A. Youngstrom

Evidence-based medicine (EBM) is a clinical practice that uses the most empirically supported methods and most current evidence in the clinical decision-making process (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). EBM is a patient-focused, self-correcting model, intended to improve the quality of care for clients. Benefits of EBM include the following: (1) allows clinicians to utilize updated research in decision making; (2) improves efficacy and efficiency for treatment provision; (3) decreases the use of ineffective or untested methods; (4) informs both clinician and client about the best treatments by utilizing publicly available information on different treatment routes; and (5) provides an empirical basis for constructing health-care policy (Romana, 2006).

In the field of mental health, this surge in EBM promotion pushes for the application of both evidence-based assessment (EBA; Hunsley & Mash, 2007; Youngstrom, 2013) to enhance treatment guidance and empirically supported treatments (EST; Chambless & Hollon, 1998; Southam-Gerow & Prinstein, 2014) to ensure that the chosen approaches have been empirically validated. This chapter focuses on EBA for rural settings while seeking to integrate the guidelines of EBM with traditional psychological assessment through a client- and test-centered approach. Given the lack of literature geared for rural settings, this chapter partly uses a "lessons learned" approach gained through the authors' real-world experiences while still providing an empirical rationale for EBA.

The benefits of adopting an EBA approach to diagnosis are substantial. Based upon multiple studies examining decision-making skills in diagnosis with real clinicians, obtaining an accurate diagnosis is significantly more likely when working through the EBA approach than when relying solely upon clinician judgment (Ægisdóttir et al., 2006; Dawes, Faust, & Meehl, 1989; Jenkins, Youngstrom, Washburn, & Youngstrom, 2011; Meehl, 1954). Data collected through intake or subsequent sessions are used more efficiently and objectively than when we impressionistically interpret the same information. The EBA approach reduces a clinician's tendency to over-interpret cues of risk in ways that often lead to over-diagnosis of disorders

A. Kirk (✉)
Department of Psychology and Neuroscience, University of Colorado at Boulder, 345 UCB, Boulder, CO 80309, USA
e-mail: alex.kirk@colorado.edu

R. Sale
Department of Educational Psychology, University of Texas at Austin, Austin, TX, USA

E.A. Youngstrom
Department of Psychology, University of North Carolina Chapel Hill, Chapel Hill, NC, USA

© Springer International Publishing AG 2017
K.D. Michael, J.P. Jameson (eds.), *Handbook of Rural School Mental Health*,
DOI 10.1007/978-3-319-64735-7_7

(Croskerry, 2003; Gigerenzer, 2002; Jenkins et al., 2011). Lastly and most importantly for academic and community settings alike, using this type of conscientious approach to diagnosis is not only economical but also more feasible. Once it becomes habit, this method saves already overscheduled clinicians' valuable time (Straus, Glasziou, Richardson, & Haynes, 2011; Youngstrom, Choukas-Bradley, Calhoun, & Jensen-Doss, 2014). For these reasons, the EBA approach is a commonsense model to overcome some of the difficulties faced in rural settings.

Rurality is difficult to define, and often those unfamiliar with rural settings mistakenly believe that all are identical. More individuals are living at or below the poverty level in rural areas, and there is less access to specialized medical and mental health services, including psychological care (Owens, Watabe, & Michael, 2013; Slama, 2004). Instead of these barriers thwarting the eventual existence of EBA within the rural setting, these issues clearly define more reason to employ the method. Rural areas need a simple diagnostic approach that saves money and time, increasing diagnostic validity and, in turn, accurate treatment selection.

An Introduction to Evidence-Based Assessment

EBA seeks to calculate the probability of a diagnosis when forming decisions concerning further assessment or initiation of treatment (Straus et al., 2011). Each client has a single probability of having a given diagnosis lying on a continuum between 0 and 100%. There are two thresholds that divide the continuum into three major zones of clinical action. EBA refers to these thresholds as the wait-test threshold and the test-treatment threshold (Straus et al., 2011; Youngstrom & Frazier, 2013; see Fig. 7.1). Finding the probability of a diagnosis below the wait-test threshold suggests that a diagnosis can be ruled out with confidence, while a probability estimate of a diagnosis above the test-treatment threshold suggests that a diagnosis can be confidently assigned, and treatment tailored to that diagnosis or diagnoses can begin (Straus et al., 2011).

Determining where to set these thresholds requires clinical judgment based on the risks and benefits associated with a diagnosis or lack of a diagnosis (Youngstrom, Jenkins, Jensen-Doss, & Youngstrom, 2012). For example, clinicians may decide that if the probability of a diagnosis is 30% or below, then the odds of a client having the diagnosis in question are sufficiently low to cease further assessment. Thus, 30% would signify the wait-test threshold. Conversely, the clinician may decide that if the probability of a diagnosis is 80% or higher, then the odds of a client having the diagnosis in question are sufficiently high to start treatment. Here, 80% would signify the test-treatment threshold. So long as the client is above 30% and below 80%, further testing may resume. Once the client reaches 30% or lower, or 80% or higher, then the appropriate action can be taken (i.e., discontinue further assessment or initiate treatment, respectively). These numbers are not official suggestions, as they will vary from one diagnosis to

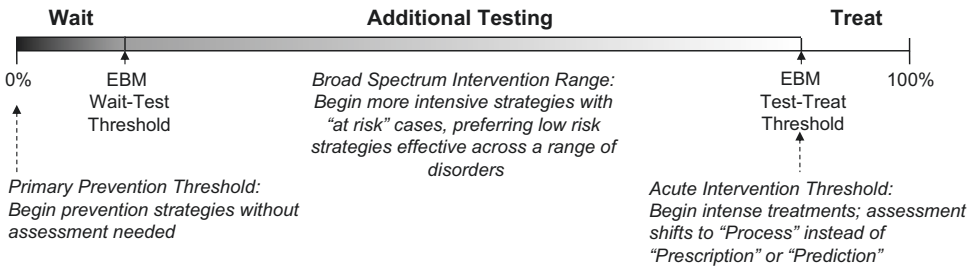

Fig. 7.1 The decision continuum with three major zones of clinical action divided by the wait-test threshold and the test-treat threshold. *Note:* This is not drawn to scale, and determining where the thresholds are to be placed will vary depending on the client

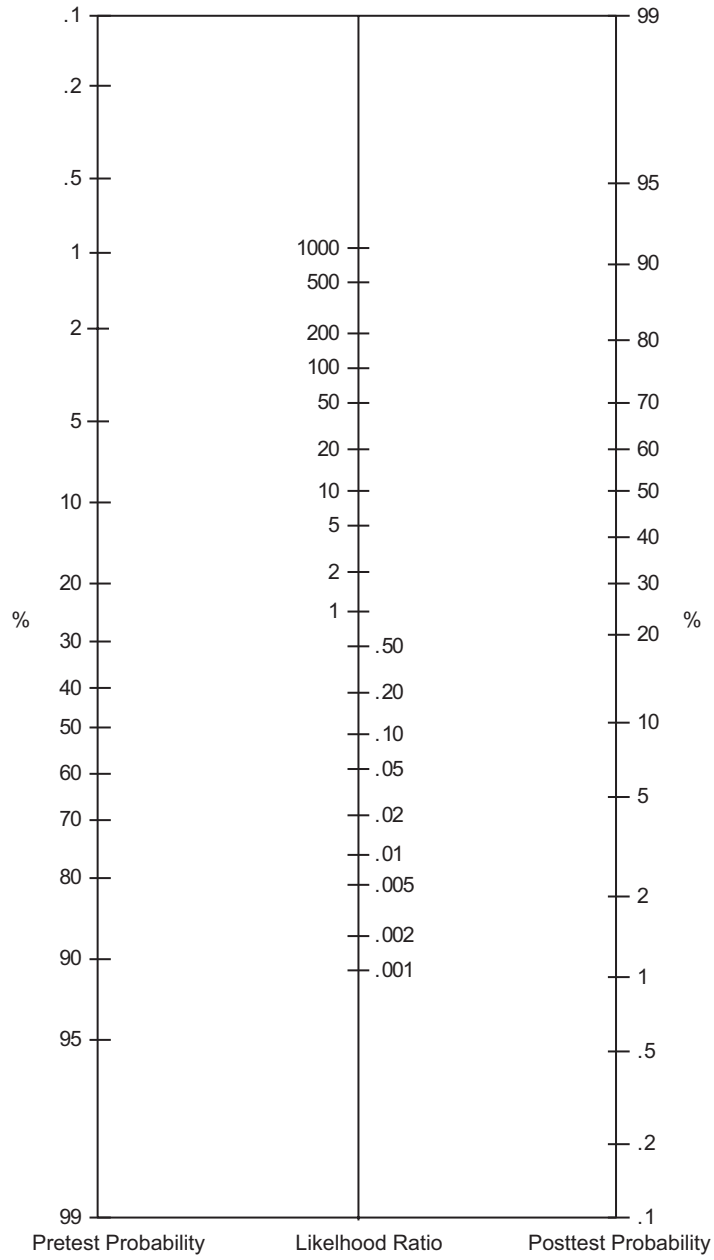

Fig. 7.2 Probability nomogram for combining prior probability with risk factors and test results. See Jenkins et al. (2011) for a detailed example of using the nomogram with a clinical case

the next. In the case of treatments where a risk of harm is minimal (e.g., exposure therapy for panic disorder; Olatunji, Deacon, & Abramowitz, 2009), then the test-treatment threshold may be lower relative to disorders where there is a greater risk of harm (e.g., those requiring high doses of psychiatric drugs that carry significant adverse effect risks).

How does a clinician figure out these probabilities for each case? New approaches take care of all the algebra for us. A paper tool called a "probability nomogram" turns this into an exercise in connecting the dots (Youngstrom, 2013; see Fig. 7.2). The base rate of a disorder within a clinical setting is a good starting probability for a client before we integrate other information we have learned about them (Meehl, 1954). This base rate goes on the left-hand line of the nomogram. Factors such as high scores on diagnostic

measures increase the probability of a specific diagnosis. The change in probability for a diagnosis is known as a diagnostic likelihood ratio (DLR). DLRs go on the middle line of the nomogram. For example, an externalizing *T* score of 81+ on the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) increases the odds of a bipolar disorder diagnosis by 4.3-fold for adolescents when compared to all other presenting problems at an outpatient clinic (Youngstrom et al., 2004). Looking at the nomogram in Fig. 7.2, an individual might have started with a 2% chance of a bipolar diagnosis according to the base rate in the general population (Van Meter, Moreira, & Youngstrom, 2011); we then move across the nomogram by lining the pretest probability at 2%, the likelihood ratio (middle column) at 4.3, and then ending with a revised posttest probability of about 7–8%. With each new source of data for a given client, we revise this probability. Each updated probability can be used as the new pretest probability when more data and subsequent DLRs are attained, highlighting the potential value of multiple sources and reporters of information, such as base rates, risk factors, self-report measures, teacher or parent report measures, and diagnostic interviews. There also are apps for smartphones and web pages that will do the calculations, but the nomogram provides a visual reference that can be engaging to use in session, with the client and clinician exploring "what-if" scenarios and discussing whether more assessment would be helpful for answering the client's questions.

The general model of an EBA battery includes broad measures, brief screeners specific to certain diagnoses, self-report measures, other-report measures (i.e., parents, teachers), and structured interviews (Youngstrom, 2013). The battery promotes the use of multiple sources of information spanning multiple syndromes (De Los Reyes, Thomas, Goodman, & Kundey, 2013). Clinicians are encouraged to adopt a similar approach within their own settings based on local referral patterns, demographic and cultural factors, along with individual client needs.

In order to calculate the most accurate probability of a diagnosis, the application of a step-by-step process within an EBA framework is essential (Youngstrom, 2013). These steps can be viewed in four stages: (1) pre-diagnosis, (2) diagnostic procedures, and (3) treatment planning, (4) all while ensuring that client preferences are taken into account along each stage. Following this process allows for multiple sources of data to be considered to fit within that framework. The overarching benefit of this framework is that it is inexpensive and simple, and requires little additional time over a standard intake protocol while greatly enhancing accuracy (Youngstrom, 2013). Table 7.1 lists the steps for EBA, elaborated upon in greater detail below, estimating the time each step requires and highlighting how a wealth of data can be collected in what many might consider to be a narrow window of time by clinical standards (i.e., between half a session and two sessions, depending on the complexity of the case). A composite case example for "Walter" will be used to illustrate this process at each stage.

Four Overarching Stages of EBA

Walter, a 14-year-old White male, was referred for school-based mental health services by his school counselor 2 months into his ninth-grade fall semester. His school counselor noticed that Walter's grades had dropped significantly from As and Bs in middle school to Ds and Fs soon into high school. Walter appears to have lost all motivation for future plans because of the difficulty concentrating he has experienced in his classes. Walter disclosed to his school counselor that he gets distracted easily in class by what others are doing around him, and that he has begun to feel insecure about participating in class. The school counselor arranges a meeting between the school-based clinician and the parents. The school-based clinician becomes concerned given that Walter's mother was diagnosed with attention-deficit hyperactivity disorder (ADHD) in middle school. Walter's parents give their consent to a psychological evaluation administered through the school.

Table 7.1 Outline of the four stages of EBA

Step	Rationale	Additional time/cost
1. Pre-diagnosis		
Identify common diagnoses	Helps to prepare necessary assessment tools for most common issues	Time: 0 Cost: 0
Know the base rate	Important starting point to anchor evaluations	Time: 0 Cost: 0
Identify risk factors	Risk factors raise "index of suspicion"—enough will elevate to assessment or possible treatment zone	Time: 2–10 min Cost: 0
2. Diagnostic procedures		
Information from broad scales	Broad externalizing/internalizing scales; self-reports; parent reports; teacher reports; low scores are less informative	Time: 0 if already part of routine assessment Cost: 0 if already part of routine assessment
Brief screens for more in-depth information	Brief family history measure adds new information; parent report for specific information pertaining to symptomatology	Time: 5 min Cost: 0—utilize instruments that are in public domain
Get multiple perspectives	Parent reports specific to diagnosis in question and to confirm change in functioning; youth and teacher reports helpful for measuring pervasiveness and motivation for treatment	Time: 5 min per informant Cost: 0—utilize instruments that are in public domain
Intensive assessment for diagnosis in question	Clinical interview focusing on specific presentation; K-SADS, MINI-KID	Time: 30–120 min Cost: 0–$4.00 for applications
3. Treatment planning		
Assessment for treatment planning	Rule out medical conditions and medications; family and school functioning, quality of life, personality, comorbid diagnoses	Time: variable Cost: variable
Process monitoring	Therapy assignments; medication monitoring; SUD ratings for exposure tasks	Time: <5 min per day Cost: 0
Progress and outcome	Repeat assessments with main severity measures—interview and/or parent report most sensitive to treatment effects	Time: 10–40 min Cost: 0
Maintenance	Review critical events and life transitions; assign ongoing therapy tasks	Time: negligible Cost: 0
4. Client preferences	Make a client a part of the process by applying his/her attitudes to adjust wait-test and test-treatment thresholds, and with treatment planning	Time: variable Cost: 0

1. Pre-diagnosis. The first stage might be viewed as a preparation stage, before any actual diagnosis begins. What are the frequent clinical issues in Walter's geographic and demographic region? Are there relevant social issues, such as general or specific substance abuse or multigenerational issues of psychopathology? Once these clinical issues and the variables surrounding them are identified, it is important to determine whether the EBA tools necessary for these common issues are available (Youngstrom, 2013). There are reviews that critically evaluate different psychological instruments and determine the evidence for their reliability and validity (Hunsley & Mash, 2008). Within this first stage it is important to note that few clinical settings will be able to maintain the resources needed to address all possible diagnoses. Prioritize the common issues, as roughly 80% of cases in most clinics will present with the same approximate 20% of clinical issues (Youngstrom, 2013). Thus, the goal ought to be maintaining the most updated tools for these more

common diagnoses. Once the most common clinical issues are identified, then your clinical setting can be more prepared to address them. Measures for rare conditions can get added to the tool kit as time, resources, and need dictate.

Base rates. Next, find benchmark base rates for conditions at clinics similar to yours (Youngstrom et al., 2012). This rate gives Walter's clinician a sense of the starting point in pursuing the proper diagnosis for Walter. If the clinician knew nothing else about him, then the base rate would be the most accurate guess as to how likely it is that he meets the criteria for the diagnosis in question (Meehl, 1954). The hypothesized diagnosis for Walter is ADHD, and so the school-based clinician will start the diagnostic process by attaining a base rate for ADHD. In the general population, the lifetime prevalence rate for moderate-severe ADHD for adolescents is ~4% (Merikangas et al., 2010); therefore, this is an appropriate starting base rate to use in examination of Walter's profile.

Determining base rates in rural settings within the scientific literature can be difficult given the underrepresentation of rural mental health research. Fortunately, prevalence rates generally tend to be similar across demographic settings (Kessler et al., 1994). Should a clinician be unable to find an accurate published base rate estimate for a specific diagnosis within his/her client population, Plan B would be to estimate the rate oneself. Dividing the number of cases with a specific diagnosis by the total number of cases reviewed provides the base rate estimate (Youngstrom et al., 2014). For example, should a clinician see 75 clients in a year, 3 of whom are diagnosed with bipolar disorder, then the base rate of bipolar disorder in this setting would be 4% (i.e., 3/75). This process takes some time to set up, but the long-term benefits of having these base rates are substantial, and updating them regularly would require simply revising the number of cases meeting a diagnosis and the number of cases reviewed in total. These rates will provide the starting point for further assessment and subsequent treatment guidance. It is important to find base rates not just for one's geographic location, but also for their actual treatment setting.

Finding base rate estimates in the literature can be done by searching various scientific databases (e.g., Medline, *PsycINFO*, PubMed, Google Scholar). Medline and *PsycINFO* are high-quality sources for psychological research, but subscriptions are expensive. PubMed and Google Scholar are free resources one can access from anywhere with an Internet connection. Before initiating a search, it would be wise to be accustomed with the syntax used with each search engine (i.e., whether quotations are required, or using "and" or "or" within the search phrase). If one is attempting to find base rates for depression, then a search consisting of the key words "depression" and "prevalence" or "epidemiology" may yield results (Youngstrom et al., 2012). Further, clinicians can change the search terms to include populations of interest (e.g., "rural") or settings of interest (e.g., "community") to narrow the results (Table 7.2).

Risk factors. Certain risk factors or clinical signs for various disorders encourage further diagnostic hypotheses after finding a base rate for the diagnosis in question (Morrison, 2006). Different risk factors can be more indicative of certain disorders, though it is important to con-

Table 7.2 Base rates for various diagnoses in different settings

Diagnosis	Setting	Base rate
Depression	Rural sample	6.1% (1)
Generalized anxiety	Rural childhood sample	1.7% (2)
Separation anxiety	Rural childhood sample	3.5% (2)
Social anxiety	General adolescent sample	1.3% (3)
Bipolar I/II	General adolescent sample	2.6% (3)
ADHD	General adolescent sample	4.2% (3)
Oppositional defiant disorder	General adolescent sample	6.5% (3)
Conduct disorder	General adolescent sample	2.2% (3)
Any substance-use disorder	General adolescent sample	11.4% (3)

Note: (1) Probst et al. (2006); (2) Costello et al. (1996); (3) Merikangas et al. (2010)

sider contextual factors such as family history of mental disorders or abuse. Generally speaking, risk factors may provide greater insight as far as what disorder the clinician ought to consider assessing. For example, clients presenting for disruptive behaviors may exhibit symptoms shared by multiple diagnoses, such as ADHD and bipolar disorder (Biederman, Klein, Pine, & Klein, 1998). However, in the event that such clients present with early-onset depression or a decreased need for sleep, then these risk factors might warrant further testing for bipolar disorder given the relative frequency that these are warning signs for bipolar disorder while having less association with ADHD (Youngstrom et al., 2012).

Clinicians can use scientific databases to determine relevant risk factors for different disorders. Some examples might include a family history, or characteristic symptom clusters. A clinician searching for risk factors associated with a diagnosis may again check the databases mentioned when discussing base rates by searching for "ADHD" (or any diagnosis of interest) and "risk factor." If searching for a specific risk factor, such as heritability, then a search consisting of "depression," "offspring," and "heritability," may narrow results specifically to heritable risk factors (Youngstrom et al., 2012). To identify a broader set of risk factors, remove the search terms relevant to heritability to expand the search. Web pages curated by the National Institute of Mental Health or the Centers for Disease Control often pull together relevant information, but the depth of coverage is uneven across disorders. Browsers let us bookmark our favorite pages and save our favorite searches, making it fast to update or tailor them (e.g., swapping "anxiety" for "ADHD"). It is an activity that fits easily in spaces created by "no shows" and cancellations, and we can share the best information with our colleagues.

Looking at Walter's possible risk factors, we learned earlier that Walter's biological mother was diagnosed with ADHD in middle school. Previous research suggests that the DLR associated with an immediate family member carrying an ADHD diagnosis is 4–5 (Faraone, Biederman, & Friedman, 2000; Faraone et al., 2000). As such, Walter's risk, starting at 4.2% as a result of the base rate of ADHD in a general adolescent population, can now be calculated to be at 11% using the nomogram in Fig. 7.2.

This first stage is not so much a task to do for a single client, but rather a way of reconfiguring our practice to work more efficiently. It may require a large amount of time up front in order to obtain a list of common reasons for referral, base rates, necessary assessment tools, and risk factors. Having these readily accessible saves an inordinate amount of time long-term as clinicians will not have to continually research these variables for every new referral, but rather every few years as such information is updated.

2. Diagnostic procedures. This second stage can be seen as the initiation of more active diagnostic procedures. Broad-scale measurement tools (e.g., Behavior Assessment System for Children, 2nd Edition [BASC-2]; Reynolds & Kamphaus, 2004) are useful in capturing a broad array of symptomatology within an individual such as Walter. Although these measures may not be highly specific to detecting specific disorders, they are useful in refining the probability of a specific diagnosis, and they can be decisive at ruling out some otherwise common possibilities. Broadly speaking, examining the internalizing and externalizing subscales helps to narrow a tentative diagnosis. For example, a high CBCL externalizing scale score will increase the odds of a diagnosis of pediatric bipolar disorder 3- or 4-fold, while a low externalizing scale score decreases the odds 20-fold (Youngstrom et al., 2004).

The benefit of utilizing broad-scale information is not solely to refine tentative diagnoses, but also to incorporate new DLRs to update the probability of the diagnosis in question more objectively. One common way of calculating a DLR uses the sensitivity and specificity of a measure (Straus et al., 2011). The *sensitivity* refers to the percentage of cases who have the disorder in question that would be correctly classified by the measure, while *specificity* refers to the percentage of cases who do not have the disorder in question and would be correctly classified as not having the disorder by the measure (Youngstrom & Frazier, 2013). These sensitivity and specific-

ity percentages can be used to calculate a DLR simply by taking the sensitivity and dividing it by the "false alarm rate" (i.e., 1—specificity). This results in a value that has a multiplicative effect on the odds of a diagnosis. DLRs below 1 suggest a reduction in the probability of a diagnosis, while a value above 1 will increase the probability of a diagnosis.

If Walter meets the criteria for ADHD, he is expected to receive high ratings on diagnostic scales devoted to attention problems and/or hyperactivity. A subscale on the BASC-2 parent report assesses for attention problems. Research has shown that a *T* score of 59.5 or higher on the attention problems subscale of the BASC-2 parent report differentiates children with ADHD from those without any disorders with a sensitivity of 93.3% and a specificity of 93.5% (Ostrander, Weinfurt, Yarnold, & August, 1998). As such, the DLR would be calculated as .933/[1 − .935] = 14.35. Thus, children with parents who indicate a *T* score of 59.5 or higher on the attention problems subscale of the BASC-2 parent report are at a 14.35-fold risk of an ADHD diagnosis relative to those without any evidence of pathology. Adding on to previous stage, this ratio would be applied to the currently existing DLR stemming from all previous steps, to then calculate a newly revised DLR taking all this data into account. Assuming that the parent form of the BASC-2 filled out by Walter's parents reported a *T* score of 60, this DLR of 14.35 can now be applied to his previous diagnostic probability of 11%. Using Fig. 7.2, Walter's revised diagnostic probability for ADHD is now between 60 and 65%.

Finding studies that report DLRs is uncommon, so it is typically a productive strategy to seek studies that report sensitivity and specificity values in order to calculate DLR by hand given the ease of calculation. Searching for such studies might include a search containing key words such as "ADHD" (or any disorder of interest) and "sensitivity and specificity" (Youngstrom et al., 2012). It is important to determine the assessment score relevant to the sensitivity and specificity percentages as sensitivity and specificity, and subsequently DLRs, shift based on scores (Tables 7.3 and 7.4). Further, it is imperative to determine whether the measure used in the study is differentiating between individuals with the specified diagnosis from those with no traces of pathology, or if the measure is differentiating between individuals with the specified diagnosis from those with some other pathology. Finally, when searching for studies reporting sensitivity and specificity values, it is important that the sample for the study contains at least ten individuals with the diagnosis and ten individuals without (Kraemer, 1992).

Brief screens. After using broad measurement tools to narrow the tentative diagnoses, the clinician ought to be gathering brief screening instruments that focus on the leading clinical hypotheses for Walter (Youngstrom et al., 2012). Looking at symptoms that might be most associated with certain disorders helps refine diagnostic probability estimates (e.g., grandiosity indicative of bipolar symptomatology). While these symptoms may not always create the greatest level of dysfunction, the goal is to find symptoms that will weaken the case for certain diagnoses while possibly strengthening others. For example, if Walter does not experience patterns of decreased need for sleep paired with excitability, the probability of the issue at hand being pediatric bipolar disorder, a diagnosis that often competes with ADHD, decreases (Youngstrom et al., 2012). The results from these brief screening devices tend to outperform the broad scales as these briefer measures will be more focused on the symptom patterns of interest. For example, Table 7.3 reports the DLRs associated with mood disorders for the CBCL—a broad scale, while Table 7.4 reports the DLRs associated with the Kutcher Adolescent Depression Scale (KADS; Brooks, 2004)—narrowly focused on depression. As a rule of thumb, comparing these DLRs suggests that the KADS may be more useful than the CBCL in detecting depression given the higher DLR (see Youngstrom, 2014 for more detail about methods for comparing tools).

In choosing a short screener to administer to Walter, the clinician chooses the Strengths and Difficulties Questionnaire (SDQ; Goodman,

Table 7.3 Diagnostic likelihood ratios for various diagnoses from broad assessments

Diagnosis	Setting	Measure	Score	DLR
ADHD	General sample	CPRSR:L (ADHD subscale)	70 (*T* score)	1.85 (1)
ADHD	General outpatient setting	BASC-2 parent form (attention problems subscale)	59.5+ (*T* score)	14.35 (2)
Anxiety (any)	Academic clinic; urban community health center	CBCL (internalizing subscale)	71–77; 78+ (*T* scores)	1.51; 2.03 (3)
Anxiety (any)	Academic clinic; urban community health center	YSR (internalizing subscale)	66–72; 73+ (*T* scores)	1.64; 2.35 (3)

Note: (1) Frazier and Youngstrom (2006); (2) Ostrander et al. (1998); (3) Van Meter et al. (2014); (4) Youngstrom (2013)

Table 7.4 Diagnostic likelihood ratios for various diagnoses from brief screening assessments

Diagnosis	Setting	Measure	Score	DLR
ADHD	General population	SDQ, hyperactivity subscale	≥6 (Raw score)	7.3 (1)
Conduct disorder/oppositional defiant disorder	General population	SDQ, conduct subscale	≥4 (Raw score)	9.1 (1)
Mania	Urban university clinics	CMRS-10, parent report	≥10 (Raw score) on CMRS	4.94 (2)
Mania	Urban university clinics	PGBI-10 M, parent report	Report 6 different DLRs; 18+ is highest	7.25 (3)
Depression	Urban/rural municipality	KADS	6; 10 (Raw scores)	3.2; 8.9 (4)
Depression	Group Health Research Institute	PHQ-9	11+ (Raw score)	3.97 (5)
Generalized anxiety	General population	GAD-7	8; 12; 15 (Raw scores)	3.8; 6.5; 8.7 (6)
Social anxiety	Midwest public schools	MASC, social anxiety subscale	8.5; 13.5 (Raw scores)	1.9; 3.43 (7)
Social anxiety	Midwest public schools	SAS-A	50 (Raw score)	2.52 (8)
Social anxiety	Midwest public schools	SPAI-C	18 (Raw score)	3.55 (8)

Note: (1) He et al. (2013); (2) Henry, Pavuluri, Youngstrom, and Birmaher, 2008; (3) Youngstrom, Frazier, Findling, and Calabrese (2008); (4) LeBlanc, Almudevar, Brooks, and Kutcher (2002); (5) Richardson et al. (2010); (6) Spitzer, Kroenke, Williams, and Löwe (2006); (7) Anderson et al. (2009); (8) Inderbitzen-Nolan et al. (2004)

1997). After calculating the results, the clinician finds that Walter scored a 7 on the hyperactivity subscale, which yields a DLR of 7.3 for ADHD (He, Burstein, Schmitz, & Merikangas, 2013). Using Fig. 7.2, we can revise Walter's odds of having ADHD from the previous probability of 60–65%, to just over 90%.

Searches for brief screeners specific to certain diagnoses can use a similar strategy to the one above. Searching for "sensitivity and specificity" specific to diagnoses of interest is the best way to do this, again with the search engines listed in previous steps (Youngstrom et al., 2012). It may be possible to include specific tools of interest within the search terms, though again this will narrow the possible pool of results. Calculating a DLR here would be no different than how it is done for broad-scale measures.

Multiple perspectives. Obtaining multiple perspectives can help to clarify the likelihood of a diagnosis (De Los Reyes et al., 2013). Parent-report measures about Walter's symptoms may add additional important information regarding the frequency and severity of the dysfunction related to reported symptoms (Carlson & Youngstrom, 2003). Teacher-report measures can inform about the pervasiveness of the Walter's symptoms into school contexts (Achenbach & Rescorla, 2001; McDermott, 1995). Specifically, the teacher can help to determine whether the Walter's symptoms are home specific or whether they occur in multiple contexts. However, a plan must be established in cases where there is disagreement among raters, as cross-informant agreement tends to be modest at best (Achenbach, McConaughy, & Howell, 1987). Savvy clinicians consider the contexts under which parent report might be more credible than self-report (Youngstrom et al., 2011). Oftentimes, making this decision may require deferring to what the literature has to say about more specific diagnoses. For example, self-report measures for internalizing disorders (i.e., depression, anxiety) are often seen as more accurate (Rothen et al., 2009), while parent or teacher reports for externalizing disorders may be more favorable, particularly for externalizing behaviors which may involve low insight, such as bipolar disorder (Pini, Dell'Osso, & Amador, 2001; Dell'Osso et al., 2002). Having less formal education or misusing substances may reduce the credibility of parent report (Youngstrom et al., 2011).

The data integration process requires caution, as it can sometimes either inflate or greatly undermine the true probability of a diagnosis. For instance, having Walter's parent fill out multiple measures, all of which are assessing the same symptoms, could inflate the estimated probability of an actual diagnosis if all of these measures were used in revising the probability of a diagnosis (Youngstrom et al., 2012). This is easy to see when using multiple, redundant measures, as the clinician would essentially be counting the same symptoms multiple times from the same reporter, treating them as though they are separate when revising the DLR.

Finding studies that look at parent- or teacher-report measures involves similar search terms to those seeking either broad-scale measurement tools or diagnosis-specific instruments (Steps 4 and 5). If a clinician knows of specific teacher or parent forms of interest, then those terms can be included within the search to try and find those specific instruments. Further, including broad terms such as "parent" or "parent report" may be of use.

Similar to the first stage, finding studies which report sensitivity and specificity figures for different measures and having them readily available save clinicians a great deal of work in the future despite requiring an up-front investment of time. A clinician can create a list of DLRs associated with different measures (e.g., Tables 7.3 and 7.4), and, upon scoring the various measures given to a client or other informants, determine whether any of the results will yield a DLR. Again this highlights that, once these numbers (i.e., base rates, risk factors, DLRs for various measures) are obtained, the EBA process moves fluidly and adds little to no additional time beyond a typical intake while greatly enhancing accuracy.

Intensive assessment. At this point, the clinician ought to be seeking to finalize the updated tentative diagnosis with more rigorous assessment measures. As such, the use of a structured clinical interview may be seen as a more individualized, intensive diagnostic assessment tool. Such interviews will help the clinician to rule out diagnoses that do not fit Walter's presenting symptoms which are primarily externalizing in nature while strengthening the case for diagnoses consistent with reported symptoms. Structured interviews with sound psychometric properties include the Mini-International Neuropsychiatric Interview for Children and Adolescents (MINI-KID; Sheehan et al., 1998) and Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS; Kaufman et al., 1997). These interviews are separated into different diagnostic sections, allowing the clinician to only assess specific diagnostic syndromes (e.g., anxiety disorders, mood disorders), though it would be considered more thorough and complete to assess for all diagnoses given the risk of comorbidity,

as is common for child and adolescent psychopathology (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003).

Often clinicians shy from using structured interviews for fear that clients dislike them, rapport will be damaged, reimbursement will be impossible, or professional autonomy will be constrained. Interestingly, all of these concerns lack empirical validity, as client surveys suggest that they appreciate the thoroughness of structured interviews without feeling a loss in rapport (Suppiger et al., 2009), Medicaid and other insurance companies will reimburse for diagnostic interviews (Youngstrom et al., 2014), and semi-structured interviews still allow for clinical judgment supported by thorough data (Axelson et al., 2003; Kaufman et al., 1997).

Walter is given a MINI-KID to confirm ADHD while ruling out other possible diagnoses. The results lend more support to the proposed diagnosis of ADHD while ruling out other possible competing diagnoses (e.g., bipolar disorder). At this point Walter has roughly a 90% probability of having ADHD, in addition to convergent support from the MINI-KID. Based on the test-treatment threshold chosen by Walter's parent, this probability may be sufficient to begin treatment with the confidence that Walter is in fact struggling with ADHD.

3. Treatment planning. After establishing a diagnostic formulation, EBA shifts to measuring Walter's functioning as well as factors that could moderate treatment selection or response. Such information includes medical history, past or current use of medications, comorbidities, and academic and social functioning. Any potential medical rule-outs need to be clarified at this point assuming that they have not been addressed previously (Youngstrom et al., 2012). Relational and systemic factors such as family functioning and presence of high-expressed emotion will often alter treatment approach (Cicchetti & Toth, 1998; McClellan, Kowatch, & Findling, 2007). Measures tailored to address functionality and quality of life can help the clinician to tailor treatment not solely towards symptom reduction, but also towards greater functioning amidst the presence of symptomatology. One such measure is the KINDL-R (Bullinger et al., 2008).

The rural setting itself may be a moderator of validity for both assessment and treatment. Rural areas have been known to exhibit heightened stigma against mental health services and diagnoses of psychopathology in general (Owens et al., 2013). Placing a diagnostic label on a person seeking services may not be necessary for treatment gains. Thoughtful clinicians use personal judgment informed by research on mental health stigma in regard to this issue, and seek supervision if necessary when deciding whether and how to share diagnostic information with the client or other involved parties.

Pharmacological treatment in addition to psychosocial intervention will often be a part of the treatment plan, especially with more chronic and debilitating syndromes. If this is the case, be prepared as a clinician to act as a "stop-gap" entity, providing treatment while the client is placed on a waiting list for psychiatric care. The availability of psychiatrists within rural settings, particularly those specialized for children and adolescents, is astoundingly low (Gamm, Stone, & Pittman, 2008). Complicating matters and prescription abuse and misuse can be significantly greater in rural settings compared to more urban areas (Anderson, Neuwirth, Lenardson, & Hartley, 2013). Clients may feel as though medication is the only option for symptom relief; therefore, extra attention spent in early intervention while planning for the most impairing symptoms may set up hope and motivation for the treatment process.

Assuming that connections have not already been made, clinicians in a rural setting ought to use this step to create connections within the local community. Given the lack of mental health resources in rural settings, knowledge of the existing or surrounding resources (e.g., safe houses, food banks), nonprofit organizations (e.g., boys and girls clubs), religious support, in addition to other main leaders within the community is imperative for comprehensive care (Owens et al., 2013). Further steps within the EBA framework assume that treatment is under way.

Process monitoring. Once treatment has been initiated, the clinician is encouraged to shift from treatment implementation to treatment monitoring, tracking progress being made towards the client's goals. Ongoing assessment of symptom severity can help the clinician objectively examine the progress being made towards the decided outcome. Measuring changes in severity and defining treatment targets are associated with better therapeutic outcomes (Finn & Tonsager, 1997; Lambert et al., 2006; Poston & Hanson, 2010). Brief screeners or checklists specific to different diagnoses are sensitive to treatment effects (West, Celio, Henry, & Pavuluri, 2011; Youngstrom et al., 2013), and/or brief screeners with broad coverage in order to determine functionality across multiple symptom syndromes, such as the YOQ-30. Progress monitoring helps define the starting point for therapy, along with both short- and long-term goals.

The outcome battery used to measure treatment gains can be brief because it concentrates on the key diagnoses and dimensions of the case formulation and can omit the conditions that were initially in contention but ruled out. The battery should assess current symptom severity for the principal diagnosis. Combining information from multiple sources, which might include parent or teacher scales, will help to gain greater insight into the generalizability of improvement across multiple contexts (Youngstrom et al., 2013). It is rarely necessary to readminister a semi-structured interview to confirm the loss of a diagnosis outside of a research protocol. On the other hand, if a client fails to make expected gains, or if they show new problems, then the assessment strategies from earlier steps could be helpful in revising the diagnoses and formulation.

Progress and outcome. At the end of treatment, measures should be readministered to determine the level of progress over the course of therapy. The outcome battery usually will be a shorter version of the initial, pretreatment battery, and this step can look similar to process monitoring. It should assess current symptom severity for the principal diagnosis, along with related syndromes. Similar to before, combining information from multiple sources will help to gain greater insight into the generalizability of improvement across multiple contexts (Youngstrom et al., 2013).

Using a Reliable Change Index (RCI; Jacobson & Truax, 1991) is especially helpful in tracking progress between pre- and posttreatment assessment scores. The RCI is a statistical method to determine whether a client has experienced meaningful change over the course of treatment. In order for a client to be classified as experiencing meaningful change, two criteria must be met: (1) the client must begin treatment with a level of symptomatology within a clinical range and then end treatment in a subclinical range, and (2) the amount of change that occurs must be sufficiently high to suggest that it cannot be attributed to random fluctuations in symptoms (Jacobson & Truax, 1991). For example, a score of 29 on the YOQ-30 has been established as the clinical cutoff, where scores 29 or higher are deemed clinical, and those below 29 are subclinical (Burlingame et al., 2004). Further, the amount of change on the YOQ-30 that must occur between pre- and posttreatment scores to qualify as reliable change is 10 (Burlingame et al., 2004). To illustrate RCI in the case of the YOQ-30, a client who starts treatment with a score of 40 and ends treatment with a score of 20 would have shown meaningful change as treatment started with a clinical score and ended with a subclinical score, and the amount of change between pre- and posttreatment scores is greater than 10. In this case, the client would be considered "recovered" (Jacobson & Truax, 1991). If this same client were to end with a score of 30 (i.e., a clinical score), then the client would have still exhibited a sufficiently large amount of change in scores where we could confidently say that they are not due solely to random fluctuation. In this case, the client would be "improved" and not yet "recovered" as the posttreatment score remains clinical (Jacobson & Truax, 1991). RCI can be calculated for most measures, often based on the data within the manual (see Jacobson & Truax, 1991 for a more in-depth explanation on how to calculate RCI).

Maintenance. Relapse prevention is a common goal for many disorders. As such, the final step of the EBA system focuses on preventing relapse via long-term monitoring of treatment gains and environmental triggers. This might entail specific tasks the individual can complete on a semi-regular basis, or it might incorporate occasional follow-up sessions, or a combination of both (Youngstrom et al., 2012). For example, generalizing exposure tasks over long periods of time and across multiple contexts helps to offset spontaneous recovery or renewal effects, even if these tasks are initiated after the client reaches subclinical symptomatology (Craske et al., 2008). As such, even after a client can be considered in remission for an anxiety disorder, it would be wise to schedule monthly or bimonthly sessions to continue with exposure tasks over a longer period of time, generalized to new contexts.

4. Client preferences. Each patient will seek treatment with preconceived attitudes which will influence engagement in the assessment and therapeutic process. These attitudes and preferences can be used to adjust the wait-test and test-treatment thresholds. For example, parents of a child who want to be certain of a bipolar disorder diagnosis before initiating a pharmacological regimen may suggest a higher test-treatment threshold, requiring that the estimate of probability for such a diagnosis be 80% or higher. Conversely, a client suspected of having depression who is eager to initiate treatment may suggest a lower test-treatment threshold, requiring that the estimate of probability for such a diagnosis be 60%. In both cases, the client's attitudes help to formulate a cost-benefit analysis which the clinician can then apply to all data attained from the EBA process. After the most accurate DLR is calculated for a given client, both clinician and client can determine together whether the estimated probability of a diagnosis is sufficiently high to initiate treatment, along with the type of treatment that would be pursued, or whether the client would prefer to postpone any psychological treatment.

Early on in the process, Walter's parents decide to set the wait-test threshold at 20%, and the test-treatment threshold at 80%. As such, assessment can continue until the school-based clinician reaches a probability for a diagnosis that is at or outside of this range, where 20% is a sufficiently low probability where Walter's parents are comfortable with him no longer pursuing assessment, and 80% is sufficiently high where they can feel comfortable that he probably has ADHD and treatment can begin.

Conclusion

Evidence-based assessment provides a framework that combines the best available evidence with the individual needs of the client. Once this framework is installed in a clinical practice, it adds little or no time or expense to working with clients, and over the long run it actually can lead to substantial savings in terms of avoiding unnecessary testing, preventing over-diagnosis, misdiagnosis, and missed diagnoses, leading to better treatment matching and ultimately better outcomes. Using an EBA approach in some ways is a natural fit with work in rural mental health, because the challenges of working in rural settings require a pragmatic attitude that repeatedly asks not just "What works?" but also "What could work here, for this person, with the available resources?" The principles and habits of EBA are practical and client centered in a way that is well suited to the rapidly changing realities of work in rural settings.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232. doi:10.1037/0033-2909.101.2.213

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., … Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, *34*, 341–382. doi:10.1177/0011000005285875

Anderson, E. R., Jordan, J. A., Smith, A. J., & Inderbitzen-Nolan, H. M. (2009). An examination of the MASC social anxiety scale in a non-referred sample of adolescents. *Journal of Anxiety Disorders*, *23*, 1098–1105.

Anderson, N. J., Neuwirth, S. J., Lenardson, J. D., & Hartley, D. (2013). *Patterns of care for rural and urban children with mental health problems* (Research Report No. 49). Retrieved from Maine Rural Health Research Center: http://www.ruralhealthresearch.org

Axelson, D., Birmaher, B. J., Brent, D., Wassick, S., Hoover, C., Bridge, J., & Ryan, N. (2003). A preliminary study of the kiddie schedule for affective disorders and schizophrenia for school-age children mania rating scale for children and adolescents. *Journal of Child and Adolescent Psychopharmacology*, *13*, 463–470. doi:10.1089/104454603322724850

Biederman, J., Klein, R. G., Pine, D. S., & Klein, D. F. (1998). Resolved: Mania is mistaken for ADHD in prepubertal children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 1096–1099. doi:10.1097/00004583-199810000-00020

Brooks, S. (2004). The Kuthcher adolescent depression scale (KADS). *Child and Adolescent Psychopharmacology News*, *9*(5), 4–6.

Bullinger, M., Brütt, A. L., Erhart, M., Ravens-Sieberer, U., & BELLA Study Group. (2008). Psychometric properties of the KINDL-R questionnaire: Results of the BELLA study. *European Child and Adolescent Psychiatry*, *17*, 125–132. doi:10.1007/s00787-008-1014

Burlingame, G. M., Dunn, T., Hill, M., Cox, J., Wells, M. G., Lambert, M. J., & Brown, G. S. (2004). *Administration and scoring manual for the Y-OQ-30.2: Youth*. Wilmington, DE: American Professional Credentialing Services.

Carlson, G. A., & Youngstrom, E. A. (2003). Clinical implications of pervasive manic symptoms in children. *Biological Psychiatry*, *53*, 1050–1058. doi:10.1016/S0006-3223(03)00068-4

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7–18. doi:10.1037/0022-006X.66.1.7

Cicchetti, D., & Toth, S. L. (1998). The development of depression in children and adolescents. *American Psychologist*, *53*, 221–241. doi:10.1037/0003-066X.53.2.221

Costello, E. J., Angold, A., Burns, B. J., Stangl, D. K., Tweed, D. L., Erkanli, A., & Worthman, C. M. (1996). The great smoky mountains study of youth. *Archives of General Psychiatry*, *53*, 1129–1136. doi:10.1001/archpsyc.1996.01830120067012

Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*, *60*, 837–844. doi:10.1001/archpsyc.60.8.837

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, *46*, 5–27. doi:10.1016/j.brat.2007.10.003

Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine*, *78*, 775–780. doi:10.1097/00001888-200308000-00003

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674. doi:10.1126/science.2648573

Dell'Osso, L., Pini, S., Cassano, G. B., Mastrocinque, C., Seckinger, R. A., Saettoni, M., … Amador, X. F. (2002). Insight into illness in patients with mania, mixed mania, bipolar depression and major depression with psychotic features. *Bipolar Disorder*, *4*(5), 315–322.

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, *9*, 123–149. doi:10.1146/annurev-clinpsy-050212-185617

Faraone, S. V., Biederman, J., & Friedman, B. H. (2000). Validity of the DSM-IV subtypes of attention-deficit/hyperactivity disorder: A family study perspective. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(3), 300–307.

Faraone, S. V., Biederman, J., Mick, E., Williamson, S., Wilens, T., Spencer, T., … Zallen, B. (2000). Family study of girls with attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *157*, 1077–1083.

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment*, *9*, 374–385. doi:10.1037/1040-3590.9.4.374

Frazier, T. W., & Youngstrom, E. A. (2006). Evidence-based assessment of attention-deficit/hyperactivity disorder: Using multiple sources of information. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 614–620. doi:10.1097/01.chi.0000196597.09103.25

Gamm, L., Stone, S., & Pittman, S. (2008). *Mental health and mental disorders: A rural challenge: A literature review. Rural healthy people 2010* (Vol. 2). College Station, TX: The Texas A & M University System Health Science Center, School of Rural Public Health, Southwest rural Health Research Center.

Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you*. New York, NY: Simon and Schuster.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

He, J., Burstein, M., Schmitz, A., & Merikangas, K. R. (2013). The strengths and difficulties questionnaire (SDQ): The factor structure and scale validation in U.S. adolescents. *Journal of Abnormal Child Psychology*, *41*, 583–595. doi:10.1007/s10802-012-9696-6

Henry, D. B., Pavuluri, M. N., Youngstrom, E., & Birmaher, B. (2008). Accuracy of brief and full forms of the Child Mania Rating Scale. *Journal*

of Clinical Psychology, *64*, 368–381. doi:10.1002/jclp.20464

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, *3*, 29–51. doi:10.1146/annurev.clinpsy.3.022806.091419

Hunsley, J., & Mash, E. J. (2008). *A guide to assessments that work*. New York, NY: Oxford University Press.

Inderbitzen-Nolan, H., Davies, C. A., & McKeon, N. D. (2004). Investigating the construct validity of the SPAI-C: Comparing the sensitivity and specificity of the SPAI-C and the SAS-A. *Journal of Anxiety Disorders*, *18*(4), 546–560.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19. doi:10.1037/10109-042

Jenkins, M. M., Youngstrom, E. A., Washburn, J. J., & Youngstrom, J. K. (2011). Evidence-based strategies improve assessment of pediatric bipolar disorder by community practitioners. *Professional Psychology: Research and Practice*, *42*, 121–129. doi:10.1037/a0022506

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., … Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 980–988. doi:10.1016/j.jpsychires.2008.10.003

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., … Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry*, *51*(1), 8–19. doi:10.1001/archpsyc.1994.03950010008002

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. Newbury Park, CA: Sage Publications.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2006). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, *10*, 288–301. doi:10.1093/clipsy.bpg025

LeBlanc, J. C., Almudevar, A., Brooks, A. S., & Kutcher, S. (2002). Screening for adolescent depression: Comparison of the Kutcher Adolescent Depression Scale with the Beck depression inventory. *Journal of Child and Adolescent Psychopharmacology*, *12*, 113–126. doi:10.1089/104454602760219153

McClellan, J., Kowatch, R., & Findling, R. L. (2007). Practice parameter for the assessment and treatment of children and adolescents with bipolar disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 107–125. doi:10.1097/01.chi.0000242240.69678.c4

McDermott, P. A. (1995). Sex, race, class, and other demographics as explanations for children's ability and adjustment: A national appraisal. *Journal of School Psychology*, *33*, 75–91. doi:10.1016/0022-4405(94)00025-4

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication—Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(10), 980–989.

Morrison, J. (2006). *Diagnosis made easier: Principles and techniques for mental health clinicians*. New York, NY: Guilford Press.

Olatunji, B. O., Deacon, B. J., & Abramowitz, J. S. (2009). The cruelest cure? Ethical issues in the implementation of exposure-based treatments. *Cognitive and Behavioral Practice*, *16*(2), 172–180. doi:10.1016/j.cbpra.2008.07.003

Ostrander, R., Weinfurt, K. P., Yarnold, P. R., & August, G. J. (1998). Diagnosing attention deficit disorders with the behavioral assessment system for children and the child behavior checklist: Test and construct validity analyses using optimal discriminant classification trees. *Journal of Consulting and Clinical Psychology*, *66*, 660–672. doi:10.1037/0022-006X.66.4.660

Owens, J. S., Watabe, Y., & Michael, K. D. (2013). Culturally responsive school mental health in rural communities. In C. S. Clauss-Ehlers et al. (Eds.), *Handbook of culturally responsive school mental health: Advancing research, training, practice, and policy* (pp. 31–42). New York: Springer.

Pini, S., Dell'Osso, L., & Amador, X. F. (2001). Insight into illness in schizophrenia, schizoaffective disorder, and mood disorders with psychotic features. *American Journal of Psychiatry*, *158*, 122–125. doi:10.1176/appi.ajp.158.1.122

Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment*, *22*, 203–212. doi:10.1037/a0018679

Probst, J. C., Laditka, S. B., Moore, C. G., Harun, N., Powell, M. P., & Baxley, E. G. (2006). Rural-urban differences in depression prevalence: Implications for family medicine. *Family Medicine*, *38*, 653–660.

Reynolds, C. R., & Kamphaus, R. W. (2004). *BASC-2: Behavioral assessment system for children manual* (2nd ed.). Circle Pines, MN: AGS Publishing.

Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., … Katon, W. (2010). Evaluation of the Patient Health Questionnaire (PHQ-9) for detecting major depression among adolescents. *Pediatrics*, *126*, 1117–1123. doi:10.1542/peds.2009-2712

Romana, H. W. (2006). Is evidence-based medicine patient-centered and is patient-centered care evidence-based? *Health Services Research*, *41*, 1–8. doi:10.1111/j.1475-6773.2006.00504.x

Rothen, S., Vandeleur, C. L., Lustenberger, Y., Jeanprêtre, N., Ayer, E., Gamma, F., … Preisig, M. (2009). Parent-child agreement and prevalence estimates of diagnoses in childhood: Direct interview versus family history method. *International Journal Of Methods In Psychiatric Research*, *18*, 96–109. doi:10.1002/mpr.281

Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, *312*, 71–72. doi:10.1136/bmj.312.7023.71

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*, 35–57.

Slama, K. (2004). Rural culture is a diversity issue. *Minnesota Psychologist*, *53*(1), 9–12.

Southam-Gerow, M. A., & Prinstein, M. J. (2014). Evidence base updates: The evolution of the evaluation of psychological treatments for children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *43*(1), 1–6. doi:10.1080/15374416.2013.855128

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*, 1092–1097. doi:10.1001/archinte.166.10.1092

Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2011). *Evidence-based medicine: How to practice and teach EBM* (4th ed.). New York: Churchill Livingstone.

Suppiger, A., In-Albon, T., Hendriksen, S., Hermann, E., Margraf, J., & Schneider, S. (2009). Acceptance of structured diagnostic interviews for mental disorders in clinical practice and research settings. *Behavior Therapy*, *40*, 272–279.

Van Meter, A., Youngstrom, E., Youngstrom, J. K., Ollendick, T., Demeter, C., & Findling, R. L. (2014). Clinical decision making about child and adolescent anxiety disorders using the Achenbach System of Empirically Based Assessment. *Journal of Clinical Child and Adolescent Psychology*, *43*(4), 552–565. doi:10.1080/15374416.2014.883930

Van Meter, A. R., Moreira, A. L., & Youngstrom, E. A. (2011). Meta-analysis of epidemiologic studies of pediatric bipolar disorder. *Journal of Clinical Psychiatry*, *72*, 1250–1256. doi:10.4088/JCP.10m06290

West, A. E., Celio, C. I., Henry, D. B., & Pavuluri, M. N. (2011). Child mania rating scale-parent version: A valid measure of symptom change due to pharmacotherapy. *Journal of Affective Disorders*, *128*, 112–119. doi:10.1016/j.jad.2010.06.013

Youngstrom, E. A. (2013). Future directions in psychological assessment: Combining evidence-based medicine innovations with psychology's historical strengths to enhance utility. *Journal of Clinical Child and Adolescent Psychology*, *42*, 139–159. doi:10.1080/15374416.2012.736358

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*, *39*(2), 204–221. doi:10.1093/jpepsy/jst062

Youngstrom, E. A., Choukas-Bradley, S., Calhoun, C. D., & Jensen-Doss, A. (2014). Clinical guide to the evidence-based assessment approach to diagnosis and treatment. *Cognitive and Behavioral Practice*. doi:10.1016/j.cbpra.2013.12.005

Youngstrom, E. A., Findling, R. L., Calabrese, J. R., Gracious, B. L., Demeter, C., Bedoya, D. D., & Price, M. (2004). Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5–17 years. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*, 847–858. doi:10.1097/01.chi.0000125091.35109.1e

Youngstrom, E. A., & Frazier, T. W. (2013). Evidence-based strategies for the assessment of children and adolescents: Measuring prediction, prescription, and process. In D. J. Miklowitz, W. E. Craighead & L. Craighead (Eds.), *Developmental psycopathology* (2nd ed., pp. 36-79). New York, NY: Wiley.

Youngstrom, E. A., Frazier, T. W., Findling, R. L., & Calabrese, J. R. (2008). Developing a ten item short form of the Parent General Behavior Inventory to assess for juvenile mania and hypomania. *Journal of Clinical Psychiatry*, *69*, 831–839. doi:10.4088/JCP.v69n0517

Youngstrom, E. A., Jenkins, M. M., Jensen-Doss, A., & Youngstrom, J. K. (2012). Evidence-based assessment strategies for pediatric bipolar disorder. *Israel Journal of Psychiatry and Related Sciences*, *49*, 15–27.

Youngstrom, E. A., Youngstrom, J. K., Freeman, A. J., De Los Reyes, A., Feeny, N. C., & Findling, R. L. (2011). Informants are not all equal: Predictors and correlates of clinician judgments about caregiver and youth credibility. *Journal of Child and Adolescent Psychopharmacology*, *21*, 407–415. doi:10.1089/cap.2011.0032

Youngstrom, E. A., Zhao, J., Mankoski, R., Forbes, R. A., Marcus, R. M., Carson, W., … Findling, R. L. (2013). Clinical significance of treatment effects with aripiprazole versus placebo in a study of manic or mixed episodes associated with pediatric bipolar I disorder. *Journal of Child and Adolescent Psychopharmacology*, *23*, 72–79. doi:10.1089/cap.2012.0024

Alex Kirk is earning his Ph.D. at the University of Colorado Boulder in both Clinical Psychology and Neuroscience. His past work has looked at dose–response effects of CBT for adolescents, clinical outcomes of CBT, and evidence-based practice in mental

health. Currently, Mr. Kirk is investigating the interaction between behavioral mechanisms of anxiety disorders and inflammation, assessment and impact of safety behaviors, and dissemination/implementation issues. Mr. Kirk's current clinical foci involve both anxiety disorders and medical populations.

Rafaella Sale is earning her Ph.D. at the University of Texas at Austin in Educational Psychology, neuropsychology track. Currently, Ms. Sale is a member of the Leveraging Evidence and Advancing Practice for youth mental health services (LEAP) lab, examining processes within clinical supervision that support the adoption and integrity of evidence-based practice. At this time, her clinical interests include the impact of acute brain injury on learning and school reentry.

Dr. Eric A. Youngstrom is passionate about evidence-based assessment as a way of improving diagnosis and treatment. He specializes in work with bipolar disorder across the life cycle, as well as changing the way research is done and packaged to support clinical decisions and improve outcomes. He loves traveling, exercising, and taking lots of pictures with his family.