Evidence-Based Treatment

27

Nicole M. DeRosa, William E. Sullivan, and Henry S. Roane

Introduction

Individuals with intellectual disabilities have not always had access to appropriate care. In fact, throughout history there is evidence of how individuals with disabilities were shunned and considered a detriment to society. Institutions, which segregated individuals with disabilities from the mainstream public, were commonplace in the late 1800s and early 1900s. The "care" individuals received in these institutions was often inhumane, lacked a focus on the needs of each individual, and was geared toward lessening the burden of individuals with disabilities on the general public. The advent of the civil rights movement sparked a shift in the acceptance and care of individuals with disabilities. The mistreatment of individuals within institutions was revealed and led to improved conditions, as well as eventually the deinstitutionalization of people with disabilities. Movement toward incorporating individuals with disabilities into society and addressing individual needs to improve their quality of life emerged.

An integral part of meeting the needs of individuals with intellectual disabilities involves identifying and implementing effective treatments. Selecting the most appropriate treatment for individuals with intellectual disabilities can be challenging, in part due to the heterogeneity of the population. That is, the unique needs of individuals diagnosed with intellectual disability do not necessarily adhere to a "one-size fits all" approach. Thus, caregivers, practitioners, physicians, and educators are tasked with finding efficient and effective treatments to meet the specific needs of individuals with intellectual disability.

Caregivers (e.g., parents, educators) must be aware of the most effective treatment options in order to provide appropriate care for individuals with intellectual disability. Historically, limited information or guidance has been available regarding the best treatments and how to effectively implement these treatments. However, the evidence-based practice (EBP) movement brought to light the importance of applying research findings into everyday clinical practice. Thus, the feasibility of non-researchers to access the relevant research in order to identify and implement the most appropriate treatments for individuals with intellectual disabilities has not only been made possible but is also now commonplace. As a result of the EBP movement, information regarding evidence-based practices and treatments (more commonly referred to as empirically supported treatments) for a variety of populations, diagnoses, and symptoms is readily disseminated through registries, guides, and journals. The accessibility of this information

N. M. DeRosa (✉) · W. E. Sullivan · H. S. Roane
Department of Pediatrics, SUNY Upstate Medical University, Syracuse, NY, USA
e-mail: derosan@upstate.edu

J. L. Matson (ed.), *Handbook of Intellectual Disabilities*, Autism and Child Psychopathology Series, https://doi.org/10.1007/978-3-030-20843-1_27

allows for individuals with intellectual disabilities to obtain the care they need and deserve.

This chapter will first review the history or evolution of the EBP movement to guide the reader in understanding EBP as it is known today. Next, the definition of EBP will be presented along with a breakdown of the three main components: (1) research evidence, (2) clinical expertise, and (3) patient characteristics, preferences, and values. A brief discussion regarding the differences between EBP and empirically supported treatments (ESTs) is then provided, followed by a detailed section on the evaluation of ESTs for individuals with intellectual disabilities. Finally, the chapter concludes with consideration of barriers to identifying and implementing ESTs, as well as discussion of the future directions of EBP.

Evolution of Evidence-Based Practice

Evidence-based practice (EBP) is a decision-making model rooted within the field of medicine (Sackett, Rosenberg, Muir Gray, Haynes, & Richardson, 1996; Thoma & Eaves, 2015). Although the philosophical foundation of evidence-based medicine can be traced back to Europe in the eighteenth and nineteenth centuries, a more widespread movement toward evidence-based medicine across Europe and North America did not occur until the mid-twentieth century. It wasn't until the early 1990s that the more modern concept of EBP emerged (APA, 2006; Sackett et al., 1996; Zimerman, 2013). The philosophical underpinnings of EBP are closely linked to epidemiology, which focuses on the needs of populations. In the late 1930s, John R. Paul founded clinical epidemiology, which stemmed from the desire to address not only the needs of entire populations but also of individual patients (Zimerman, 2013). That is, the practices of employing measurable theories and approaches to study diseases and other health-related incidents in populations were incorporated into the everyday decision-making model of clinical practice to positively affect individual patient outcomes. The use of scientific models to inform clinical practice is the impetus behind the EBP movement.

The focus on specific patient needs within clinical practice, as well as refinement of the medical decision-making process, proliferated during the 1960s, as described in a review of EBP conducted by Sur and Dahm (2011). Alvan Feinstein, a follower of Paul's who studied mathematics prior to becoming a physician, played an integral role in incorporating basic science into the practice of medicine (Zimerman, 2013). During his involvement in a rheumatic fever epidemiological study, Feinstein identified an over-reliance on clinical expertise alone when determining the presence of relevant symptoms of the disease. This led Feinstein to develop scientific-based criteria for evaluating the symptoms of study participants, which resulted in a more accurate classification of the symptoms of rheumatic fever. In turn, a dramatic decrease in the prevalence of the disease was observed following improved patient outcomes. Furthermore, given Feinstein's criticism of the ability of public health studies to appropriately train physicians in clinical care, he paved the way for improving the value of medical research by linking science and clinical experience. That is, the subjective nature of public health research limited the ability to apply the research findings to improve clinical care. Thus, Feinstein began to modify traditional medical teaching by merging statistical methods, inherent to epidemiology, with clinical reasoning to positively affect patient outcomes.

Around the same time that Feinstein's work was unfolding, events that influenced the movement toward EBP were also occurring in Canada, namely, a shift toward universal healthcare (Zimerman, 2013). One outcome of the change in the Canadian healthcare system during the mid-1960s was the emergence of new medical schools aimed at incorporating basic science into the medical curriculum. Of particular importance was the founding of McMaster University, where the first department of clinical epidemiology and biostatistics emerged, a program heavily influenced by the work of Feinstein who served as a visiting professor during the program's initial years. A key feature of the McMaster program

was the development of a "problem-based learning" (Zimerman, 2013; p. 72) technique in which clinical scenarios were presented in a group format in order to incorporate both the basic sciences and clinical medicine into the teaching experience. Although Feinstein's involvement with the newly developed department was crucial in fostering the novel "problem-based learning" teaching method, much of the development of the curriculum and training techniques that influenced the future of EBP is credited to the department head at the time, David Sackett (Sur & Dahm, 2011).

Trained in Public Health, Sackett saw a need for clinical practice to be driven by science in order to better affect the health outcomes of individuals; however, he recognized that most epidemiological training programs addressed public health issues in a manner that was not readily accessible to physicians (Thoma & Eaves, 2015). At the time, clinical practice predominantly involved patient observation and physician opinion. That is, medical programs were not training physicians to be consumers of epidemiological research such that they could apply those findings to their patients. Thus, during his time as head of the Department of Clinical Epidemiology and Biostatistics, Sackett focused on structuring the curriculum such that issues related to public health and practices rooted in science (e.g., biostatistics) were readily incorporated into physician training.

In addition to a novel epidemiological curriculum that encompassed the problem-based learning approach, Sackett and colleagues published a series of articles to assist physicians with effectively accessing information from the literature that could then be applied at the bedside (Thoma & Eaves, 2015). These articles were published in the *Canadian Medical Association Journal (CMAJ)* in the early 1980s and eventually titled "Readers' Guides" (Zimerman, 2013). Following the initial publication of the Readers' Guides, Sackett and colleagues continued to disseminate information on the problem-based learning technique through workshops and additional publications. This laid the groundwork for the future of EBP (i.e., linking science and clinical medicine).

Although Sackett was pivotal in the movement toward EBP, it was Gordon Guyatt, a young faculty member at McMaster, who is officially credited with coining the term "evidence-based medicine" (Thoma & Eaves, 2015; Zimerman, 2013). Furthermore, Guyatt was instrumental in continuing the work of Sackett throughout the 1990s and further disseminating the methods of the newly termed concept of evidence-based medicine. During his time at McMaster, Guyatt further cultivated evidence-based teaching methods, as well as developed a medical residency program based on these methods (Zimerman, 2013). Additionally, Guyatt and colleagues partnered with US academic institutions in the early 1990s to form a workgroup aimed at assessing the state of evidence-based medicine at that time. Conclusions of the workgroup suggested that although Sackett's Readers' Guides affected the bedside practice of medical professionals, limitations of the work existed. More specifically, the workgroup noted that the Readers' Guides included a heavy focus on the quality of evidence but lacked focus on the application of evidence to individual patient needs. That is, guidance was needed on how to appropriately apply the evidence to specific clinical cases, particularly when the relevant evidence may be limited (Sur & Dahm, 2011).

The decision to republish Sackett's Readers' Guides into the *Journal of the American Medical Association* (*JAMA*) in the early 1990s allowed for the opportunity to address the limitations noted by the workgroup. An updated, two-part series of the Guides was developed that included a section on evaluation of clinical measurements, edited by Sackett, and a new section, "Users' Guides to the Medical Literature," headed by Guyatt. This update on the Readers' Guides more heavily focused on the application of evidence into the daily practice of clinical medicine (Zimerman, 2013). The first publication of the new series appeared in the pages of *JAMA* in 1992 and publication continued for 8 years, concluding with a total of 32 articles. During the initial years following the first publication, the evidence-based medicine literature was predominantly maintained by McMaster researchers. However, beginning in 1995 the number of outside researchers who

published using the phrase "evidence-based medicine" surpassed that of McMaster researchers, indicating that these methods had successfully infiltrated the field of medicine (Zimerman, 2013). Furthermore, a PsychInfo search conducted by DiGennaro-Reed and Reed (2008) displayed a notable increase in publications using the term "evidence-based practice" beginning in the early 2000s demonstrating the breadth of the evidence-based movement beyond the field of medicine.

Although, the evidence-based movement is largely credited to the field of medicine, education and psychology are also known to have strong roots in an evidence-based approach to assessment and treatment. Of particular significance is Lightner Witmer's establishment of the first psychological clinic at the University of Pennsylvania in 1896, which is known as one of the first institutions to conduct experimentation on the implementation of various teaching strategies (APA, 2006; DiGennaro-Reed & Reed, 2008; Fagan & Wise, 2000). Additionally, in 1908 Witmer founded the journal *The Psychological Clinic*, which focused on clinical services and disabled children (Fagan & Wise, 2000). Likewise, in the late nineteenth and early twentieth century, schools began to adopt an evidence-based approach to improving students' success in the classroom through the work of G. Stanley Hall, the founder of the American Psychological Association (APA). Hall's work influenced the development of the first clinic facility to function in a public school, the Department of Scientific Pedagogy and Child Study (Fagan & Wise, 2000). Furthermore, in the 1960s, during the time in which the evidence-based medicine movement was taking off, it became increasingly common for researchers in psychology and education to conduct treatment comparisons to identify the most effective strategies to improve client and student outcomes (Biglan & Ogden, 2008). Thereafter, research in education and psychology proliferated on effective treatments for a variety of problems (e.g., mental health issues, classroom behavior, drug use) across populations (e.g., children, adolescents). Thus, history suggests that the fields of education and psychology were frontrunners in the EBP movement (APA, 2006).

Although education has a long-standing history in the EBP movement, recent developments have been significantly impactful in improving EBP within the field of education, notably the Council for Exceptional Children (CEC) *Standards for Evidence-Based Practices in Special Education* (2014). The CEC standards provide quality indicators and criteria to assist educators with evaluating the quality of research studies to aid in developing and refining teaching practices. Additionally, the What Works Clearinghouse (WWC), funded by the Institute for Educational Sciences within the US Department of Education, is a research database that systematically evaluates school-based academic and behavioral interventions in order to guide the practice of educators. The availability of resources offered by the CEC and WWC is critical, particularly given federal mandates regarding instructional strategies. That is, the *Individuals with Disabilities Education Improvement Act* (IDEA, 2004) and the *No Child Left Behind Act* (2001) both specify that instructional strategies must be selected and designed using "scientifically based research." Therefore, with the help of initiatives like the CEC and WWC, educators can be equipped with treatments that have been scientifically shown to produce favorable outcomes within the school setting.

One significant contribution that the field of psychology had on the EBP movement was APA's development of best practice guidelines. In 1992, APA created a joint task force among the Board of Scientific Affairs, the Board of Professional Affairs, and the Committee for the Advancement of Professional Practice, which resulted in the publication of the *Template for Developing Guidelines*: *Treatments for Mental Disorders and Psychosocial Aspects of Physical Disorders* (APA, 1995, 2006). The goal of the "Template" was to have structured criteria for establishing best practice guidelines that included reference to both relevant research and clinical expertise. The "Template" was later replaced with the *Criteria for Evaluating Treatment Guidelines* in 2002 (APA, 2006). The Task Force on Promotion and

Dissemination of Psychological Procedures emerged from Division 12 (Clinical Psychology) of APA in 1993, which developed criteria for identifying empirically supported treatments. This came in response to many psychological treatments being viewed as inferior to medical treatments (e.g., pharmacology), despite proven effectiveness of certain therapies (i.e., psychotherapy) dating back to the 1970s.

Although the development of treatment criteria brought about increased awareness regarding the effectiveness of psychological treatments for specific disorders, many found the criteria to be too restrictive given an emphasis on randomized controlled trials (RCTs) and treatment manuals (APA, 2006). Thus, psychologists within specific disciplines (e.g., counseling) began to develop additional guidelines that also accounted for patient variables to better inform treatment development and application within clinical practice. Additionally, the 2005 APA Presidential Task Force on Evidence-Based Practice was established in an effort to identify global practice standards to benefit all disciplines of psychology, as well as to help guide policymakers who had a newfound interest in EBP.

In the early 2000s, healthcare systems and healthcare policy began to readily adopt the notion of EBP. To date, various federal, state, and local organizations and agencies incorporate guidelines and/or mandates that specify the need for EBP. For example, several states have enacted an autism insurance mandate that provides coverage for several medical and behavioral treatments (e.g., applied behavior analysis), such that individuals with autism are able to access appropriate, EBP-derived care. Additionally, APA Task Forces continue to work toward the development and refinement of criteria for EBP across psychology disciplines (APA, 2006). Similar developments are observed across several other fields, including behavioral medicine, education, and behavior analysis (Council for Exceptional Children, 2006; Davidson, Trudeau, Ockene, Orleans, & Kaplan, 2003; DiGennaro-Reed & Reed, 2008). A major impetus for incorporating EBP across disciplines, organizations, and agencies is to maximize consumer (e.g., patients, students) outcomes, as well as to ensure physicians, psychologists, teachers, and the like are able to not only access the relevant research but also assess and evaluate it and most importantly put it into practice.

Definition of Evidence-Based Practice

The main purpose of EBP is to narrow the research-to-practice gap to ensure that treatment is not only effective for a given individual but also cost-efficient (Spencer, Detrich, & Slocum, 2012). However, variations in wording and perhaps meaning of what EBP actually entails can be found across disciplines, the impact of which will be discussed later in this chapter. Despite this, there are three general variables that encompass EBP: (1) research evidence, (2) clinical expertise, and (3) patient characteristics. For example, the Institute of Medicine (2001) defines EBP as "the integration of best research evidence with clinical expertise and patient values" (p. 147; adapted from Sackett, Strauss, Richardson, Rosenberg, & Haynes, 2000). Similarly, the 2005 APA Presidential Task Force defines EBP as the "integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (p. 273). Of significance is the emphasis on research evidence rather than expert opinion alone. However, it is also important to note that EBP does not emphasize research evidence over expert opinion but instead stresses a collaboration between the two. That is, EBP encompasses professional opinion and clinical practice that is guided by research to best meet the needs of individual consumers (Cook, Tankersley, & Harjusola-Webb, 2008). Below we discuss each of the three areas of EBP.

Research Evidence. In order to consider a treatment or other clinical practice to be effective, it must be backed by a strong research base. The APA Presidential Task Force on Evidence-Based Practice indicates that "a sizeable body of evidence drawn from a variety of research designs and methodologies attests to the effectiveness of

psychological practices" and furthermore states that "evidence derived from clinically relevant research on psychological practices should be based on systematic reviews, reasonable effect sizes, statistical and clinical significance, and a body of supporting evidence" (APA, 2006, p. 284). However, determining the strength of the relevant research may not always be straightforward.

Most researchers and practitioners would agree that a large amount of scientific evidence is necessary in order to be confident that a particular practice is effective and appropriate for a given individual, but the specific amount or type of evidence needed is not always agreed upon. For example, RCTs are commonly referred to as the "gold standard"; however, this type of research design may not always be the most appropriate given a specific research question or analysis. Accordingly, a number of differing criteria for determining the level of evidence available for a given treatment or practice have been developed and will be discussed in detail later in this chapter. Nonetheless, what is vital to EBP is a critical appraisal of the research.

Clinical Expertise. Clinical expertise encompasses a practitioner's theoretical understanding of the relevant topic, clinical training, knowledge and understanding of the relevant research, self-evaluation, and ongoing education and training (APA, 2006). Sackett et al. (1996) described clinical expertise as the "proficiency and judgment" practitioners obtain through their education and training (p. 71). Through clinical expertise, the relevant research can be effectively applied to individual patients given an understanding of the clinical needs of that patient. Although clinical practice should be guided by the relevant research, at times a particular evidence-based treatment targeting a specific concern may not match the client's unique needs or be feasible to implement in the relevant treatment setting. For example, for a student with an intellectual disability that has verbal outbursts in the classroom maintained by teacher attention, evidence would suggest that withholding attention when the student shouts out and providing attention contingent on an appropriate alternative behavior (e.g., the child raises his/her hand and waits to be called on) can reduce the student's vocal outbursts. However, in a classroom setting, it may be very difficult, if not impossible, to ignore the student's vocal outbursts given the impact this may have on the learning of peers. Therefore, clinical expertise may come into play such that appropriate modifications are made to the treatment to best meet the needs of the student while accounting for variables related to the treatment setting (e.g., classroom).

Furthermore, at times there may simply be an insufficient research base related to the specific needs of a given patient or consumer, and thus practitioners must rely on expert opinion and practice principles guided by science. For instance, suppose there are two separate treatments that have been shown to produce favorable outcomes for a particular concern across multiple empirical studies. However, within the literature there are *no* studies that have directly compared these treatments. In a case such as this, the practitioner must rely on their own clinical expertise to select the most appropriate treatment for their client.

Given clinical expertise encompasses not only the relevant research but also practices based on opinions, experiences, and an understanding of the patient or consumer, it is important that practitioners are aware of the limits of their knowledge, as well as their own biases when making clinical decisions (APA, 2006). One manner in which clinical expertise can be evaluated is through examining the efficacy and efficiency of clinical decision-making (e.g., patient outcomes; Sackett et al., 1996). Although relying solely on clinical expertise can produce outdated practices that negatively impact the consumer, clinical practice in the absence of clinical expertise (i.e., relying only research evidence) may result in the unique needs of individual consumers not being reliably met (Sackett et al., 1996).

Patient Characteristics. Consideration of patient characteristics entails more than the target symptoms or problem at hand. Evidence-based practice also considers patient or consumer preferences, values, sociocultural context, and strengths, as these variables are likely to affect treatment outcomes (APA, 2006). Thus, by taking patient characteristics into account, the prac-

titioner is best equipped to apply the relevant research, in conjunction with expert opinion, to best meet the needs of the patient rather than the disorder, symptom, or concern. For example, suppose a contingency has been arranged for a patient with an intellectual disability to increase their physical activity such that for every mile the patient walks, they can earn a specific reinforcer. The effectiveness of this contingency ultimately relies on how motivated the patient is to obtain the programmed reinforcer. One way to increase the chances that the contingency will result in the desired behavior change (i.e., increase the patient's physical activity) is to assess the patient's preference for items they could earn as a reinforcer (e.g., Fisher et al., 1992). By doing so, the patient's most preferred item can be programmed into treatment. Accordingly, contingent access to a highly preferred reinforcer will likely increase the patient's physical activity to a greater extent than earning access to a low-preferred reinforcer.

Along these lines, Swift, Callahan, and Vollmer (2011) conducted a meta-analysis (described later in this chapter) that compiled data from 35 studies comparing outcomes of patients in which preference was incorporated into treatment or not. Findings suggested that for those patients in which preferences were accounted for, significantly better outcomes (d = 0.31) were obtained. This study highlights the benefits of incorporating patient preference into treatment and aligns with EBP. For further discussion of incorporating patient characteristics as a component of EBP, please refer to APA Presidential Task Force on Evidence-Based Practice (2006).

Evidence-Based Practice and Empirically Supported Treatments

Evidence-based practice is a comprehensive, global approach to patient or consumer care. In other words, it is a decision-making model that seeks to answer a specific question as it relates to an individual (APA, 2006; Schlosser & Sigafoos, 2008). Thus, it is important to point out that EBP differs from empirically supported treatments (EST), which are treatments with known effectiveness for a given population and/or problem. Schlosser and Sigafoos (2008) define ESTs as treatments that have "obtained a certain threshold of research evidence" (p. 61). Although significant to EBP, ESTs are one component of the decision-making process. That is, when employing EBP, practitioners will select treatments that (a) have a substantial research base (i.e., ESTs), (b) best meet the individual's needs based on the practitioner's clinical expertise, and (c) account for the individual's characteristics and preferences.

Evaluating Empirically Supported Treatments

As previously noted, EST are but one component of the EBP decision-making process. In order for a treatment to be empirically supported, it is imperative that studies investigating the efficacy of the particular treatment are carefully evaluated. Stated differently, the efficacy of any treatment must be documented in carefully controlled empirical research that demonstrates the treatment in question is responsible for the beneficial effects observed, rather than other confounding variables or chance (see Campbell & Stanley, 1963; Kazdin, 1992). Within the literature, a number of criteria have been put forth to classify treatments as evidence-based (Chambless & Hollon, 1998; Kratochwill & Stoiber, 2002; Odom et al., 2005; WWC, 2017). Common to these criteria is assessing the quality of experimental studies employing randomized controlled trials (RCTs), single-case experimental designs (SCEDs), and meta-analyses. Any particular treatment may have an accumulation of evidence supporting its efficacy; however, if that evidence is derived from low-quality sources, the confidence a clinician has in recommending that treatment might be truncated. In the following section, we introduce the aforementioned experimental methodologies (i.e., RCTs, SCEDs, and meta-analyses), the various quality assessments linked

to each design, and general criteria outlining the amount of research required for a treatment to be considered an EST.

Randomized Controlled Trials

Randomized controlled trials (RCTs) are experimental group designs that are largely regarded as the "gold standard" of clinical research (Schulz, Altman, & Moher, 2010a, 2010b; Sturmey, 2014). In RCTs, participants are randomly assigned to treatment and control groups. This random assignment is done to ensure that any confounding variables that may influence the outcome of the study are equally dispersed across groups. In the treatment group, participants receive the treatment under investigation (i.e., independent variable), whereas in the control group, participants receive no treatment, an alternative treatment, or a placebo. Then, at specified times during the study, quantifiable data are collected on the participant variable (i.e., dependent variables) and compared across groups. In general, RCTs are highly regarded in clinical research because (1) they carefully control for confounding variables that may threaten internal validity and (2) large groups of participants undergo random assignment, which enhances external validity. In other words, treatment effects derived from a well-designed RCT can be attributed to the treatment under investigation and extended to the larger population. A detailed analysis of RCTs and other group designs go beyond the purpose of this chapter but are available from other sources (Machin & Fayers, 2010; Torgerson & Torgerson, 2008).

Although RCTs have a sound methodological foundation and can produce empirical evidence regarding the effectiveness of a particular treatment, not all RCTs are conducted with the same level of experimental rigor (Grossman & Mackenzie, 2005). Therefore, there are a number of methodological considerations that need to be accounted for when conducting and reporting a RCT, some of which may be difficult to achieve. For instance, Nichol, Bailey, and Cooper (2010) outlined a number of challenges with employing RCTs: (1) unclear hypotheses and multiple objectives, (2) poor selection of endpoints, (3) inappropriate subject selection criteria, (4) non-clinically relevant or feasible treatments, (5) inadequate randomization, stratification, and blinding, (6) insufficient sample size, (7) failure to use intention-to-treat analysis, and (8) failure to anticipate common practical constraints. Although we will not detail each of these potential issues, it is important to acknowledge that even though RCTs are highly regarded in terms of experimental design, the quality of an RCT can vary widely and dramatically impact the level of evidence provided.

Given the number of potential issues with conducting RCTs, it is crucial that studies be adequately described to allow for objective quality assessment. That is, there must be clear descriptions of the methodologies employed and findings obtained in order to assess the quality of the study and determine the level of evidence it provides for a particular treatment. Without such technical information, many flaws in a study may be masked or overlooked by reviewers. To ensure that RCTs receive adequate and objective quality assessment, it was necessary to standardize the manner in which RCTs are reported.

Accordingly, the Consolidated Standards of Reporting Trials (CONSORT standards) were developed and described by Schulz, Altman, Moher, and the CONSORT group in 2010. The primary aim of the CONSORT standards is to ensure that adequate information is reported to allow for a quality evaluation. A total of 25 standards are included in the CONSORT criteria. As an example, Standard 5 states that treatments for each group are described with sufficient detail to allow for replication, including how and when the treatments were administered. We will not be discussing each standard here but want to highlight the importance of these standards as they provide guidelines for how to report RCTs, allowing for quality assessment. Ultimately, this provides an objective basis to determine whether or not a treatment should be considered "evidence based."

Within the educational literature, and as an example of RCT quality assessment, WWC

(2017) developed criteria for evaluating individual studies. First, studies are screened to determine if the study is relevant to the literature base on the treatment in question. Second, the measurement system employed in the study is examined to ensure the researchers selected and used an appropriate (i.e., reliable and valid) measurement system. Third, the design of the study (e.g., RCT) and sample (e.g., participants; sample size) used in the study are assessed. Fourth, baseline data (if obtained) are examined looking for equivalence across treatment and control group for the dependent variables. Then finally, the analyses performed (e.g., statistical analyses) are assessed for appropriateness. Please see *WWC Study Review Guide, Group Design Studies* (2018) for a more detailed description of the evaluation process.

Single-Case Experimental Designs

Single-case experimental designs (SCEDs) have a long history of scientific merit and provide empirical evidence at the level of the individual. Although a thorough discussion of SCEDs is beyond the scope of this chapter, a number of detailed discussions are available elsewhere (DeRosa, Sullivan, Roane, & Kadey, in press; Kratochwill & Levin, 2014; Perone & Hursh, 2013). Like RCTs, SCEDs are experimental in nature. The goal of any SCED is to demonstrate a causal relationship between the independent variable (e.g., treatment) and dependent variable (e.g., behavior targeted for change). This causal, or functional, relation relies on the demonstration of experimental control. When the introduction of the independent variable occasions a change in the dependent variable, and the effect is reliably replicated within and/or across participants, experimental control is demonstrated. By making both within- and between-subject comparisons, threats to internal validity are controlled for, while systematic replication of the effects bolsters external validity (Martella, Nelson, & Marchand-Martella, 1999).

Like all experimental designs, the quality of a SCED also varies and requires quality assessment to determine the level of evidence provided. To assess the level of evidence presented in SCEDs, Horner et al. (2005) established specific criteria. That is, for a treatment to be considered evidence-based derived from SCEDs, five standards need to be met: (1) the treatment is operationally defined, (2) the setting in which the treatment is to be used is defined, (3) the treatment is implemented with fidelity, (4) demonstration of a functional relation between the treatment and a change in behavior, and (5) the experimental effects are replicated across a sufficient number of studies, researchers, and participants. In general, Horner and colleagues suggested that in order for a treatment to be evidence based, a series of experimental studies need to be conducted that demonstrate an effect by the treatment on behavior under specified conditions (i.e., setting and procedures). For a more detailed discussion of these standards please see Horner et al. (2005) or WWC (2017) for SCED standards used in the educational literature.

Meta-analyses

Meta-analyses are an integral part of EBP as they offer an approach to aggregate the results of multiple studies and synthesize their collective findings. Meta-analyses are based on a family of statistical procedures that combine data from multiple studies and calculate average effect sizes. An effect size is a statistical metric that estimates the magnitude of effect the independent variable has on the dependent variable. Stated differently, an effect size estimates how much of an effect the treatment had, rather than simply did it have an effect greater than what would be expected by chance (i.e., *p*-value).

Meta-analytic strategies have been developed for both RCTs and SCEDs. In the case of RCTs, the most commonly used effect size is Cohen's *d,* which standardizes differences between group means (i.e., treatment group vs. control group) expressed as a *z*-score. However, for SCED, Cohen's *d* is inappropriate as it relies on group differences, which SCEDs inherently do not provide and because SCED data do not meet many

of the basic assumptions needed to conduct parametric analyses (e.g., homoscedasticity; Burns, 2012). For these reasons researchers have developed a family of nonparametric effect sizes for use with single-case data, for example, percentage of nonoverlapping data (Scruggs, Mastropieri, & Castro, 1987), percentage of all nonoverlapping data (Parker, Hagan-Burke, & Vannest, 2007), and nonoverlap of all pairs (Parker & Vannest, 2009), to name a few. Similar to RCTs, meta-analytic researchers utilize these nonparametric effect sizes to synthesize the results of numerous studies simultaneously. Thus, meta-analyses are well-suited for questions regarding a particular treatment's evidence base as they can simultaneously synthesize and analyze results from numerous studies employing various types of experimental designs. For a comprehensive review of meta-analytic strategies, please see Faith, Allison, and Gorman (Faith, Allison, & Gorman, 1996) and Schmidt and Hunter (Schmidt & Hunter, 2014).

One method for assessing the quality of individual studies, which may be included in a meta-analysis, is the *Scientific Merit Rating Scale* (SMRS), developed by the National Center on Autism's *National Standards Project* (National Autism Center, 2009). The SMRS checklist uses a 5-point rating scale to evaluate study quality across five domains. First, the research design is considered in terms of the degree to which experimental control is demonstrated and design employed. Second, a rating is given regarding the extent to which reliable and valid measurement of the dependent variable took place. Third, procedural integrity is reviewed to determine if the independent variable was implemented as intended. Fourth, participant characteristics are reviewed to determine if the treatment was evaluated within the appropriate population. And fifth, a rating is given on generalization or the extent to which the study attempted to objectively demonstrate that treatment effects occur across time, setting, stimuli, and responses.

Another method to assess the quality of studies included in meta-analyses was developed by Guyatt et al. (2008) termed the "grading of recommendation assessment, development, and evaluation" (GRADE) system. The GRADE system first assesses the quality of study design, with RCTs starting as "high evidence" and observational studies or SCEDs as "low evidence." From there, five factors may result in a study's quality of evidence being downgraded: (1) study limitations, (2) inconsistency of results, (3) indirectness of evidence, (4) imprecision, and (5) publication bias. In opposition, there are three factors that may improve the quality of evidence provided by a particular study: (1) large magnitude of effect, (2) dose response, and (3) confounding variables likely minimize the effect.

After individual studies have been assessed, the GRADE system applies a broad classification system to the entire body of work on a particular treatment across four levels: "(1) High quality—further research is very unlikely to change confidence in the estimate of effect, (2) Moderate quality—Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate, (3) Low quality—Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate, and (4) Very low quality—Any estimate of effect is uncertain" (Guyatt et al., 2008, p. 926). This leads to an overall rating of the quality of evidence provided across studies related to a particular treatment. The GRADE system then outlines how to decide whether the treatment in question should be recommended based on (a) quality of evidence, (b) cost/benefit analyses, and (c) patient or client values and preferences. For a more complete discussion, please refer to Guyatt et al.

Amount of Evidence Required for EBP

Although high-quality empirical research is needed in order for a treatment to be considered evidence based, one perfectly conducted study does not suffice. For a treatment to be considered evidence based, there not only has to be one quality empirical investigation demonstrating its effectiveness, there must be multiple experimental demonstrations confirming the treatments efficacy.

Horner et al. (2005) proposed that five well-designed single-case studies, conducted by at least three different research groups, across three different geographical locations, with a total of at least 20 participants, are needed to establish a treatment as evidence based. The WWC (2017) also set forth criteria to categorize treatments based on the amount of evidence in the literature. More specifically, the "extent of evidence" for a particular treatment is based on the number and sizes of the studies. A "small" amount of evidence is defined as either one study, one school, or less than 350 participants. Whereas a "medium to large" amount of evidence consists of more than one study, more than one school, and at least 350 participants. The purpose of criteria like these is used to give an indication to the practitioner of how generalizable the research findings are for a given treatment.

Similarly, the National Autism Center (2009) developed a classification system used to describe the strength of evidence available for a particular treatment. This classification system divided the level of evidence provided for a particular treatment into four categories: established, emerging, unestablished, and ineffective/harmful. An established treatment, based on the National Autism Center's criteria, has been documented in either two group experiments with at least 12 participants that display no conflicting results or three group experiments or six, single-case experiments with at least 18 participants with no more than one study that found conflicting results. Additionally, all studies need to have received high ratings (i.e., 3 or greater) on the SMRS to be considered established treatment. Table 27.1 provides an outline of the National Autism Center's criteria, and Table 27.2 provides an example of how these criteria have been applied to behavioral treatments for individuals with autism.

Barriers to Evidence-Based Practice

Despite the breadth of EBP, selecting an appropriate treatment can be met with challenges for several reasons. Differences in terminology used across disciplines, agencies, and organizations are one barrier to an evidence-based approach. Terms that are used interchangeably include

Table 27.1 A summary of the four levels of evidence from the National Autism Center's (2009) *National Standards Project*

Level of evidence	Definition
Established	1. Two group experiments or four single-case experiments with at least 12 participants without conflicting results; three group experiments or six single-case experiments with a minimum of 18 participants with no more than one study presenting conflicting results 2. Had SMRS scores of 3 or greater 3. Reported beneficial treatment effects for specific targets 4. These results may be supplemented by other lower-quality studies
Emerging	1. One group experiment or two single-case experiments with a minimum of six participants without conflicting results 2. Has SMRS scores of 2 3. Reported beneficial treatment effects on one dependent variable for a specific target 4. These studies may be supplemented by those with higher or lower SMRS scores
Unestablished	1. May or may not be based on research 2. Had beneficial effects reported on poorly controlled studies with SMRS scores of 0 or 1 3. Have claims based on testimonials, opinions, or speculation 4. Were ineffective, unknown, or adverse treatment effects based on poorly controlled studies
Ineffective/ harmful	1. Had two group experiments or four single-case experiments with at least 12 participants without conflicting results; three group experiments or six single-case experiments with a minimum of 18 participants with no more than one study presenting conflicting results 2. Had SMRS scores of at least 3 3. No beneficial treatment effects for one dependent variable for a specific target or had adverse treatment effects on dependent variable

Adapted from the National Autism Center (2009).

Table 27.2 A summary of treatments for children with ASD under the age of 21 that meet three levels of evidence (Established, Emerging, and Unestablished) from the National Autism Center's (2015) *National Standards Project*

Level of evidence	Treatments
Established	1. Behavioral treatments 2. Cognitive behavior treatment package 3. Comprehensive behavioral treatment for young children 4. Language training (production) 5. Modeling 6. Natural teaching strategies 7. Parent training 8. Peer training package 9. Pivotal response training 10. Schedules 11. Scripting 12. Self-management 13. Social skills package 14. Story-based treatment
Emerging	1. Augmentative and alternative communication devices 2. Developmental relationship-based treatment 3. Exercise 4. Functional communication training 5. Imitation-based treatment 6. Initiation training 7. Language training (production and understanding) 8. Massage therapy 9. Multicomponent package 10. Music therapy 11. Picture exchange communication system 12. Reductive package 13. Sign instruction 14. Social communication intervention 15. Structured teaching 16. Technology-based treatment 17. Theory of mind training
Unestablished	1. Animal-assisted therapy 2. Auditory integration training 3. Concept mapping 4. DIR/floor time 5. Facilitated communication 6. Gluten-free/casein-free diet 7. Movement-based treatment 8. SENSE Theatre treatment 9. Sensory treatment package 10. Shock therapy 11. Social behavioral learning strategy 12. Social cognition treatment 13. Social thinking treatment

Adapted from the National Autism Center (2015).

evidence-based practice, best practice, research-based practice, empirically supported treatments, empirically validated treatment, and scientifically supported treatments (DiGennaro-Reed & Reed, 2008). More specifically, many disciplines (e.g., psychology, medicine, public health) embrace EBP as a decision-making model that incorporates several components to providing appropriate care to consumers, including the identification and selection of ESTs. However, in other disciplines (e.g., education), EBP typically refers to specific interventions or instructional

Table 27.3 A summary of criteria for empirically supported treatments (Chambless & Hollon, 1998)

Level of evidence	Definition
Efficacious	1. Two between-group design experiments (RCTs), showing treatment is superior (based on statistical significance) to another treatment or placebo 2. Three single-case design experiments ($n > 9$), demonstrating experimental control, and compared to another treatment 3. Experimental procedures detailed 4. Participant characteristics specified 5. Effects demonstrated by at least two different research groups
Possibly efficacious	1. Two experiments showing treatment is superior (based on statistical significance) to a wait-list control 2. All experiments conducted by one research group 3. Single-case design experiment(s) ($n < 3$)

Adapted from Tolin, McKay, Forman, Klonsky, & Thombs (Tolin, McKay, Forman, Klonsky, & Thombs, 2015)

Table 27.4 A summary of treatments and level of evidence for the treatment of aggression for individuals with intellectual disabilities

Level of evidence	Treatments
Evidence-based treatment	**Behavioral treatments** • Functional communication training • Differential reinforcement procedures • Function-based reinforcement; mixed-treatment packages
Lacking sufficient evidence	**Behavioral treatments** • Antecedent manipulations and changes in context • Noncontingent reinforcement alone • Antecedent exercise • Response blocking • Contingent positive punishment • Self-monitoring **Other treatments** • Mindfulness • Teaching family model • Vibroacoustic music • Aromatherapy
Inconclusive evidence	Cognitive-behavioral therapy

The level of evidence is based on Chambless and Hollon's (1998) criteria

Adapted from Healy, O., Lydon, S. and Murray, C. (Healy, Lydon, & Murray, 2014). Aggressive Behavior. In Evidence-Based Practice and Intellectual Disabilities (eds P. Sturmey and R. Didden). doi:https://doi.org/10.1002/9781118326077.ch5

strategies that have a strong empirical base. Thus, the differences in terminology regarding EBP can pose challenges to obtaining appropriate care given a multidisciplinary team approach.

Variation in criteria, regarding what constitutes a treatment as empirically supported, is another challenge that interferes with accessing appropriate treatment for individuals with intellectual disabilities. That is, while the research base for a given treatment may be classified as efficacious or established given one set of criteria, evaluation given a different set of criteria may suggest that the treatment is not empirically supported. The National Autism Center's criteria (Table 27.1) and examples of application of the criteria (Table 27.2) were discussed earlier in this chapter. Tables 27.3 and 27.4 are provided for comparative purposes where Table 27.3 outlines Chambless and Hollon's (1998) criteria for categorizing "efficacious" and "possibly efficacious treatments" and Table 27.4 provides an example of the application of these criteria. Additionally, databases outlining ESTs may not be readily available to practitioners and educators, and resources, particularly funding, may be limited, thus restricting the ability to incorporate ESTs into every day practice (DiGennaro-Reed & Reed, 2008).

Finally, the lack of evidence on the implementation of EBP poses a significant barrier to accessing appropriate care. That is, evaluation of treatment integrity and progress monitoring are considered a part of EBP; however, there is limited guidance regarding how often and in what manner these measures should be taken to ensure the decision-making model is effective (Detrich, 2008; DiGennaro-Reed & Reed, 2008). Thus, the future direction of EBP for individuals with intellectual disabilities is evaluating the model itself to understand if efficient and effective outcomes are in fact being achieved.

Summary

The history of EBP has been quite impactful in improving outcomes for patients and students with intellectual disabilities. The EBP movement was integral in shifting the general perception and understanding of individuals with intellectual disabilities, such that these individuals can now more readily access the treatments needed to enhance their quality of life. Furthermore, rigorous treatment evaluation criteria support the relevance of obtaining and implementing research-based findings in the daily care of individuals with intellectual disabilities.

Despite the long-standing history of EBP and improved care for individuals across disciplines, limitations exist in the ability to consistently and effectively access needed treatments within and across populations. Thus, continued work is needed to ensure the effective implementation of EBP, as well as to refine the current model to best meet the needs of individuals with various presenting symptomology. Although the push toward EBP became prevalent in the 1960s, with the ideas being more readily disseminated into literature and practice in the 1990s, EBP has only become *essential* within the field of medicine, education, psychology, and other health-related disciplines (e.g., speech, public health) within the past two decades. Thus, although the work of Sackett and other pioneers of the EBP movement cannot be discounted, it is now integral to move beyond an understanding and basic application and begin to focus on the implementation and effectiveness of EBP for individuals with intellectual disabilities.

References

American Psychological Association. (1995). *Template for developing guidelines: Interventions for mental disorders and psychosocial aspects of physical disorders*. Washington, DC: Author.

American Psychological Association. (2006). Presidential task force on evidence-based practice: Evidence-based practice in psychology. *American Psychologist, 61*, 271–285.

Biglan, A., & Ogden, T. (2008). The evolution of evidence-based practice. *European Journal of Behavior Analysis., 9*, 81–95.

Burns, M. K. (2012). Meta-analysis of single-case design research: Introduction to the special issue. *Journal of Behavioral Education, 21*, 175–184.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. In *Handbook of research on teaching*. Chicago, IL: Rand McNally.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*(1), 7–18.

Cook, B. G., Tankersley, M., & Harjusola-Webb, S. (2008). Evidence-based special education and professional wisdom: Putting it all together. *Intervention in School and Clinic, 44*, 105–111. https://doi.org/10.1177/1053451208321566

Council for Exceptional Children. (2006). *Evidence-based professional practices proposal*. Arlington, VA: Author.

Council for Exceptional Children. (2014). *Standards for evidence-based practices in special education*. Arlington, VA: Author.

Davidson, K. W., Trudeau, K. J., Ockene, J. K., Orleans, J. K., & Kaplan, R. M. (2003). A primer on evidence-based review systems and their implications for behavioral medicine. *Annals of Behavioral Medicine, 28*, 226–238.

DeRosa, N. M., Sullivan, W. E., Roane, H. S., & Kadey, H. J. (in press). Single-case experimental designs. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (2nd ed.). New York, NY: Guilford.

Detrich, R. (2008). Evidence-based, empirically supported, or best practice? A guide for the scientist-practitioner. In J. K. Luiselli, D. C. Russo, W. P. Christian, & S. Wilczynski (Eds.), *Effective practices for children with autism: Educational and behavior support interventions that work*. New York: Oxford Press.

DiGennaro-Reed, F. D., & Reed, D. D. (2008). Towards an understanding of evidence-based practice. *Journal of Early and Intensive Behavior Treatment., 5*, 20–29.

Fagan, T. K., & Wise, P. S. (2000). *School psychology: Past, present, and future* (2nd ed.). Bethesda, MD: National Association of School Psychologists.

Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Groman (Eds.), Design and analysis of single-case research (pp. 267–300). New York, NY: Psychology Press.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*(2), 491–498.

Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: Gold standard, or merely standard? *Perspectives in Biology and Medicine, 48*(4), 516–534. Johns Hopkins University Press. Retrieved July 18, 2018, from Project MUSE database.

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating

quality of evidence and strength of recommendations. *British Medical Journal (Clinical Research Ed.), 336*(7650), 924–926.

Healy, O., Lydon, S., & Murray, C. (2014). Aggressive behavior. In P. Sturmey & R. Didden (Eds.), *Evidence-based practice and intellectual disabilities*. New York: John Wiley & Sons. https://doi.org/10.1002/9781118326077.ch5

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179.

Individuals with Disabilities Education Improvement Act of 2004, H.R. 1350, 108th Congress.

Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academies Press.

Kazdin, A. E. (1992). *Methodological issues & strategies in clinical research* (1st ed.). Washington, DC: American Psychological Association.

Kratochwill, T. R., & Levin, J. R. (2014). In T. R. Kratochwill, J. R. Levin, T. R. Kratochwill, & J. R. Levin (Eds.),. Introduction: An overview of single-case treatment research *Single-case treatment research: Methodological and statistical advances*. Washington, DC: American Psychological Association. https://doi.org/10.1037/14376-001

Kratochwill, T. R., & Stoiber, K. C. (2002). Evidence-based treatments in school psychology: Conceptual foundations of the procedural and coding manual of division 16 and the society for the study of school psychology task force. *School Psychology Quarterly, 17*(4), 341–389. https://doi.org/10.1521/scpq.17.4.341.20872

Machin, D., & Fayers, P. M. (2010). *Randomized clinical trials: Design, practice and reporting*. New York: John Wiley & Sons.

Martella, R. C., Nelson, J. R., & Marchand-Martella, N. E. (1999). *Research methods: Learning to become a critical research consumer*. Madison, WI: Allyn & Bacon.

National Autism Center. (2009). *National standards project: Phase 1*. Randolph, MA: Author.

National Autism Center. (2015). *National standards project: Phase 2*. Randolph, MA: Author.

Nichol, A. D., Bailey, M., & Cooper, D. J. (2010). Challenging issues in randomised controlled trials. *Injury, 41*, S20–S23.

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*(2), 137–148. https://doi.org/10.1177/001440290507100201

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *The Journal of Special Education, 40*, 194–204.

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357–367.

Perone, M., & Hursh, D. E. (2013). Single-case experimental designs. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, K. A. Lattal, G. J. Madden, et al. (Eds.), *APA handbook of behavior analysis* (Methods and principles) (Vol. 1, pp. 107–126). Washington, DC: American Psychological Association. https://doi.org/10.1037/13937-005

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

Sackett, D., Strauss, S., Richardson, W., Rosenberg, W., & Haynes, R. (2000). *Evidence-based medicine: How to practice and teach EBM* (2nd ed.). Edinburgh: Churchill Livingstone.

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage Publications.

Schulz, K. F., Altman, D. G., & Moher, D. (2010a). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*(1), 18. https://doi.org/10.1186/1741-7015-8-18

Scruggs, T. E., Mastropieri, M. A., & Castro, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remediate and Special Education, 8*, 24–33.

Schlosser, R. W., & Sigafoos, J. (2008). Identifying "evidence-based practice" versus "empirically supported treatment". *Evidence-Based Communication Assessment and Intervention, 2*, 61–62.

Schulz, K. F., Altman, D. G., & Moher, D. (2010b). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *BMJ, 23*(340), c332.

Spencer, T. D., Detrich, R., & Slocum, T. A. (2012). Evidence-based practice: A framework for making effective decisions. *Education and Treatment of Children, 35*, 127–151.

Sturmey, P. (2014). Evidence-based practice an introduction. In P. Sturmey & R. Didden (Eds.), *Evidence-based practice and intellectual disabilities* (pp. 3–28). New York: John Wiley & Sons.

Sur, R. L., & Dahm, P. (2011). History of evidence-based medicine. *Indian Journal of Urology., 27*, 487–489. https://doi.org/10.4103/0970-1591.91438

Swift, J. K., Callahan, J. L., & Vollmer, B. M. (2011). Preferences. In J. Norcross (Ed.), *Psychotherapy relationships that work: Evidence based responsiveness* (2nd ed., pp. 301–315). New York: Oxford University Press.

Thoma, A., & Eaves, F. F. (2015). A brief history of evidence-based medicine (EBM) and the contributions of Dr. David Sackett. *Aesthetic Surgery Journal., 35*, 261–263. https://doi.org/10.1093/asj/sjv130

Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D., & Thombs, B. D. (2015). Empirically supported treat-

ment: Recommendations for a new model. *Clinical Psychology: Science and Practice, 22*(4), 317–338. https://doi.org/10.1111/cpsp.12122

Torgerson, D. J., & Torgerson, C. J. (2008). *Designing randomised trials in health, education and the social sciences: An introduction*. New York: Springer.

What Works Clearinghouse. (2017). *Standards handbook, version 4.0*. Washington, DC: US Department of Education.

What Works Clearinghouse. (2018). Study review guide. In *Group design studies*. Washington, DC: US Department of Education.

Zimerman, A. L. (2013). Evidence-based medicine: A short history of a modern medical movement. *American Medical Association Journal of Ethics., 15*, 71–76.