

JOSHUA A. CLELAND, PT, PhD, OCS, FAAOMPT<sup>1</sup> • J. TIMOTHY NOTEBOOM, PT, PhD, SCS<sup>2</sup>  
JULIE M. WHITMAN, PT, DSc, OCS, FAAOMPT<sup>3</sup> • STEPHEN C. ALLISON, PT, PhD<sup>4</sup>

# A Primer on Selected Aspects of Evidence-Based Practice Relating to Questions of Treatment, Part 1: Asking Questions, Finding Evidence, and Determining Validity

**A**s the physical therapy profession continues the transition toward autonomous practice, the emphasis on decreasing practice variation and standardizing care around best practice patterns to maximize clinical outcomes and cost effectiveness will continue to evolve.<sup>10,18</sup> Providing “best care” is linked to the basic concept of evidence-based practice (EBP): integration of clinical expertise with the patient’s values and the best available research evidence to ensure optimal outcomes.<sup>50</sup> Clinical expertise refers to

the clinician’s proficiency and acuity when making judgments and applying clinical skills in the care of individual patients. Finally, another goal of EBP is to improve clinical performance through critical evaluation of the current evidence and the integration of the “best evidence” in the management of individual patients.<sup>15</sup>

Ideally, the best available evidence, integrated into decisions about individual patients, should be based on patient-centered clinical research.<sup>46</sup> Many physical therapists believe that practicing EBP requires too much time for the busy clinician; but, in reality, a purpose of EBP is to improve efficiency in clinical decision

making and to assist clinicians in selecting and applying interventions that will maximize positive patient outcomes.<sup>33</sup>

According to the Guide to Physical Therapist Practice,<sup>7</sup> patient management consists of 5 interrelated elements: examination, evaluation, diagnosis, prognosis, and interventions and outcomes. Data collected during the initial examination should be evaluated and should facilitate decision making regarding management strategies that are most appropriate for the individual patient. The diagnostic process in physical therapist practice has been described in detail elsewhere.<sup>20</sup> Once the diagnostic process surpasses the treatment threshold, or the point in the examination at which the clinician has determined that treatment may begin,<sup>20</sup> the clinician must determine the optimal intervention or combination of interventions needed to maximize patient outcomes. The clinician uses data collected during the examination, along with the diagnosis and patient goals, to determine the patient’s prognosis and likely response to treatment. All elements of patient management as described by the Guide to Physical Therapist Practice<sup>7</sup> relate to components of EBP. However, this clinical commentary will focus on as-

• **SYNOPSIS:** The process of evidence-based practice (EBP) guides clinicians in the integration of individual clinical expertise, patient values and expectations, and the best available evidence. Becoming proficient with this process takes time and consistent practice, but should ultimately lead to improved patient outcomes. The EBP process entails 5 steps: (1) formulating an appropriate question, (2) performing an efficient literature search, (3) critically appraising the best available evidence, (4) applying the best evidence to clinical practice, and (5) assessing outcomes of care. This first commentary in a 2-part series will review

principles relating to steps 1, 2, and 3 of this 5-step model. The purpose of this commentary is to provide a perspective to assist clinicians in formulating foreground questions, searching for the best available evidence, and determining validity of results in studies of interventions for orthopaedic and sports physical therapy. *J Orthop Sports Phys Ther* 2008;38(8):476-484. doi:10.2519/jospt.2008.2722

• **KEY WORDS:** *critical appraisal, physical therapy, treatment effectiveness*

<sup>1</sup>Associate Professor, Department of Physical Therapy, Franklin Pierce University, Concord, NH; Physical Therapist, Rehabilitation Services, Concord Hospital, Concord, NH.

<sup>2</sup>Associate Professor, School of Physical Therapy, Regis University, Denver, CO. <sup>3</sup>Assistant Professor, School of Physical Therapy, Regis University, Denver, CO; Faculty, Regis University Manual Therapy Fellowship, Regis University, Denver, CO. <sup>4</sup>Professor, Rocky Mountain University of Health Professions, Provo, UT; Associate Professor, Baylor University, Waco, TX. Address correspondence to Dr Joshua Cleland, Franklin Pierce College, 5 Chenell Drive, Concord, NH 03301. E-mail: joshcleland@comcast.net

pects of EBP associated with determining appropriate treatment.

In the clinical decision-making process, a certain degree of uncertainty exists with regard to interventions that will most likely maximize the chance of obtaining successful outcome for an individual patient.<sup>14,54</sup> Although the volume and quality of evidence for the efficacy and effectiveness of many commonly used physical therapy interventions is improving, the ability to identify the most appropriate treatment strategy can be a difficult task when faced with varying levels of uncertainty about the validity of a respective study's findings. Efficiency in incorporating EBP into clinical practice specifically related to treatment is a 5-step process: (1) developing an answerable question, (2) identifying the evidence for treatment, (3) critically appraising the evidence (which requires an understanding of research design and statistical principles<sup>33</sup>), (4) incorporating evidence into clinical practice, and (5) evaluating the effectiveness and efficiency with which steps 1 through 4 were carried out when determining an appropriate intervention strategy for the particular patient.<sup>46</sup> The purpose of this clinical commentary will be to provide a perspective of the first 2 steps related to treatment and that part of step 3 related to validity of evidence, with an emphasis on studies of interventions in orthopaedic and sports physical therapy. This commentary is the first of a 2-part series. Part 2 will provide a perspective of principles for interpreting results from evidence for treatment, applying the evidence to patient care, and evaluating proficiency with EBP skills.

## STEP 1. DEVELOPING AN ANSWERABLE QUESTION

**T**HE FIRST AND OFTEN MOST DIFFICULT step is the development of a well-built clinical question that facilitates a literature search, ultimately leading to the best evidence available to remove or optimally reduce clinical uncertainty.<sup>44</sup> There are 2 types of clinical

questions: background and foreground. Background questions are developed to enhance knowledge relative to a specific disorder.<sup>46</sup> For example, a clinician may ask "What causes carpal tunnel syndrome?" or "Why do patients develop coronary artery disease?" While these background questions will lead clinicians to information regarding the specific pathology,<sup>46</sup> they usually do not provide the clinician with up-to-date information about optimal treatment options for the patient. Foreground questions of therapy are developed in response to the need to identify evidence regarding the use of a specific intervention in the management of a particular patient.<sup>46</sup> As it is the purpose of this commentary to discuss studies of treatment effectiveness, foreground questions will remain the focus of this section.

Foreground questions of therapy consist of 4 components: (1) a patient or problem, (2) an intervention, (3) a comparison intervention (if relevant), and (4) an outcome.<sup>46</sup> These 4 components may be referred to as PICO (patient, intervention, comparison, outcome). Some examples of foreground questions, including these 4 components, are as follows: (1) In a 38-year-old female with carpal tunnel syndrome, what is the efficacy of exercise and ergonomic interventions compared to no treatment for decreasing pain and disability? or (2) In a 43-year-old female with plantar fasciitis, are custom-fit orthotics more effective than prefabricated orthotics in decreasing plantar foot pain?

## STEP 2. SEARCHING THE LITERATURE

**I**T IS CRUCIAL TO DEVELOP ACCURATE and efficient search strategies when seeking the best available evidence in the literature. Computerized literature searching is an essential skill necessary to efficiently practice EBP.<sup>16</sup> A number of searchable databases exist, including MEDLINE and the Cumulative Index of Nursing and Allied Health Literature

(CINAHL). Search strategies entail using 1 or more key words that may be found in the article's title or abstract. Additionally, some databases use Medical Subject Headings (MeSHs), which are biomedical terms that designate major concepts within the MEDLINE database.<sup>32</sup> Search strategies for a specific MeSH term will reveal articles relevant to that heading and others associated within the respective database. It has been reported that combining MeSH terms and key words yields the most sensitive search results (ability to detect all citations in the database) in MEDLINE, as compared to simply searching by key words or MeSH.<sup>51</sup> PubMed offers several helpful tutorials for using MeSH terms in online searches (<http://www.nlm.nih.gov/bsd/disted/pubmed.html>).

PubMed Clinical Queries<sup>4</sup> is a very helpful and efficient utility available within PubMed (<http://www.ncbi.nlm.nih.gov/sites/entrez>), the public access portal for MEDLINE searches. Key words entered into PubMed Clinical Queries search fields are automatically incorporated into predetermined EBP-compliant search strategies to find the best evidence to answer foreground questions on diagnosis, prognosis, therapy, or etiology/harm. For each search type, users can specify whether narrow (specific) searches or broad (sensitive) searches are desired. Additional search strategies within Clinical Queries target systematic reviews and clinical prediction guides (rules). Search strategies used within Clinical Queries have been systematically tested to filter results based on study design. This approach can substantially reduce time and effort for a busy clinician searching to identify studies of a particular design; but we must be aware that it does not provide an assessment of how well the study was conducted. To illustrate the efficiency of searches using PubMed Clinical Queries, we can compare results obtained with and without the methodologic filters. A PubMed search using the search string "exercise AND patellofemoral" yields 161 hits when used without the filters. How-

**TABLE 1**

**RESULTS OBTAINED ENTERING THE SAME SEARCH STRING INTO THE PUBMED<sup>4</sup> SEARCH ENGINE IN 3 WAYS**

Search Type	Search String	Targeted Study Types	Results (n Hits)
PubMed search without Clinical Queries filters	exercise AND patellofemoral	All	161
Broad, sensitive filter for therapy using PubMed Clinical Queries <sup>3</sup>	(exercise AND patellofemoral) AND ((clinical[Title/Abstract] AND trial[Title/Abstract]) OR clinical trials[MeSH Terms] OR clinical trial[Publication Type] OR random[Title/Abstract] OR random allocation[MeSH Terms] OR therapeutic use[MeSH Subheading])	Cohort, case-control, RCT*	54
Narrow, specific filter for therapy using PubMed Clinical Queries <sup>4</sup>	(exercise AND patellofemoral) AND (randomized controlled trial[Publication Type] OR (randomized[Title/Abstract] AND controlled[Title/Abstract] AND trial[Title/Abstract]))	RCT*	22
Filter for systematic reviews using PubMed Clinical Queries <sup>4</sup>	(exercise and patellofemoral) AND systematic[ <i>sb</i> ]	Systematic review	7

Abbreviation: RCT, randomized controlled trial.

\* Note that search string content other than "exercise AND patellofemoral" was appended automatically in the searches performed with PubMed Clinical Queries.<sup>4</sup>

ever, a Clinical Queries search using the same string yields only 54 hits using the broad, sensitive filter for therapy, but only 22 hits (a more manageable number) using the narrow, specific filter for therapy. Examination of the automated transformations of the simple search string using the 2 search hedges (TABLE 1) reveals that broad searches include lower-quality studies, while narrow searches target higher-level studies. Using these filters a clinician can avoid inefficient searches that yield too many studies of lesser quality, searching first for studies of higher quality when looking for best available evidence.

In addition to electronic search engines, a number of online databases provide clinicians with evidence summaries, as well as quality ratings of the available evidence. A number of evidence sources currently exist including the Australian-based Physiotherapy Evidence Database (PEDro) (<http://www.pedro.fhs.usyd.edu.au/>), McMaster University's Health Information Research Institute and Centre for

EBP (<http://www.bmjupdates.com>), the Cochrane Library ([www.cochranelibrary.com](http://www.cochranelibrary.com)), and the American Physical Therapy Association's (APTA) Hooked on Evidence online database (<http://www.apta.org/hookedonevidence.org>). Hooked on Evidence allows APTA members to perform a quick search on a specific topic and provides detailed description of the current evidence and allows for clinicians to quickly implement evidence into clinical practice.<sup>34</sup> An extensive list of EBP-related databases with advantages and disadvantages of each can be found in the article by MacDermid.<sup>33</sup> Additionally, a number of free online rehabilitation and medical journals and lists of these online journals exist, such as *British Medical Journal* (<http://bmj.bmjournals.com>), BioMed Central (<http://www.biomedcentral.com/bmccomplementalternmed>), free full-text journal listings (eg, <http://www.freemedicaljournals.com>), and Google Scholar, which is often able to find full text (<http://scholar.google.com>). Lastly, Open Door is a new feature

available to APTA members through the APTA web site (<http://www.apta.org/opendoor>). The mission of Open Door is to allow physical therapists easy access to clinical research. This service provides full-text access online to articles directly relevant to physical therapy practice through ProQuest, the Cochrane Library, and CINAHL. Readers are referred to the article by Doig and Simpson<sup>16</sup> for more information on conducting efficient literature searches.

### STEP 3A. CRITICALLY APPRAISING THE LITERATURE: ARE THE RESULTS VALID?

THE EBP METHOD REQUIRES INDIVIDUAL clinicians to make independent professional judgments about the validity and strength of the research, and relevance of the evidence to the clinician's particular patient. This process is based on the premise that the interpretations and conclusions of authors in published studies should not be accepted without close scrutiny by the reader. Fortunately, the EBP approach defines a finite set of key validity issues for consideration and provides methods for making decisions about clinical meaningfulness of treatment effects reported. The critical appraisal process enables a clinician to answer 3 questions<sup>24</sup> once the best evidence is found: (1) Are the results valid? (2) What are the results? and (3) How can I apply the results to patient care? The remainder of this commentary will address the first of these 3 questions. The remaining 2 questions will be addressed in part 2 of this series.

#### Hierarchy of Evidence

When evaluating evidence for effectiveness of an intervention, clinicians often find it helpful to use a system to determine the level of evidence for a particular study. A level of evidence is a label reflecting a study's position on the hierarchy of evidence, providing a rough indication of inherent protections against validity threats, or sources of bias, based

on the study design and methods.<sup>41</sup> After assessing over 100 different systems for rating the strength and quality of evidence, the Agency for Healthcare Research and Quality identified 7 systems that fully address all important domains for a body of evidence.<sup>5</sup> Among these 7 systems is one developed by David Sackett and colleagues, freely accessible from the Centre for Evidence-Based Medicine website.<sup>1</sup> Levels of evidence applicable for studies exploring the efficacy of clinical treatments were extracted from that system and presented with descriptions in **TABLE 2**. In addition to identifying the level of evidence on the hierarchy, therapists must also consider critically appraising the study's overall quality and the study's internal and external validity, prior to implementing the results in clinical practice.

### Internal and External Validity

When critiquing a study regarding treatment, one must consider threats to both internal and external validity. Internal validity relates to elements of research design intended to exert control over extraneous variables that could potentially impact the outcomes of the study, including interactions between patient assignment, competing interventions, history, maturation, and instrumentation.<sup>43</sup> External validity refers to the generalizability of the study's results to actual clinical practice.<sup>17</sup> While validity is not a true dichotomy (ie, either internal or external), some would argue that the most optimal single-study design for determining treatment effectiveness is a randomized controlled design with strong internal validity where a single intervention of interest is being compared to a group receiving either no treatment or a placebo intervention.<sup>19</sup> This type of research design provides strong internal validity by exerting more rigid control over all possible extraneous variables and is the optimal method for determining the efficacy of a particular intervention. However, this study design does not account for a number of variables that

<b>TABLE 2</b>	
<b>LEVELS OF EVIDENCE FOR TREATMENT STUDIES FROM THE OXFORD CENTRE FOR EVIDENCE-BASED MEDICINE*<sup>1</sup></b>	
<b>Level</b>	<b>Therapy, Prevention, Aetiology, Harm</b>
1a	SR (with homogeneity <sup>†</sup> ) of RCTs
1b	Individual RCT (with narrow confidence interval)
1c	All or none <sup>‡</sup>
2a	SR (with homogeneity <sup>†</sup> ) of cohort studies
2b	Individual cohort study (including low-quality RCT [eg, less than 80% follow-up])
2c	"Outcomes" research; ecological studies
3a	SR (with homogeneity <sup>†</sup> ) of case control studies
3b	Individual case control study
4	Case series (and poor-quality cohort and case control studies <sup>§</sup> )
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research, or "first principles"

*Abbreviations: RCT, randomized controlled trial; SR, systematic review.*  
*\* Adapted with permission from Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Daves since November 1998.<sup>1</sup> Source: <http://www.cebm.net/index.aspx?o=1025>*  
<sup>†</sup> *By homogeneity we mean a systematic review that is free of worrisome variations (heterogeneity) in the directions and degrees of results between individual studies. Not all systematic reviews with statistically significant heterogeneity need be worrisome, and not all worrisome heterogeneity need be statistically significant.*  
<sup>‡</sup> *Met when all patients died before the treatment became available, but some now survive on it; or when some patients died before the treatment became available, but none now die on it.*  
<sup>§</sup> *By "poor-quality" cohort study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably blinded), objective way in both exposed and nonexposed individuals and/or failed to identify or appropriately control known confounders and/or failed to carry out a sufficiently long and complete follow-up of patients. By "poor-quality" case-control study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably blinded), objective way in both cases and controls and/or failed to identify or appropriately control known confounders.*

exist in diverse clinical situations including multimodal treatments, skill level of the treating clinicians, compliance with home exercise programs, and competing interventions.<sup>23</sup> Another critique of the controlled trial design is that patients who do not receive any treatment have no expectations as to their potential outcomes; hence the effects of expectation alone could potentially account for any differences observed between patients or groups of patients.<sup>54</sup>

Intimately associated with issues of validity is the relative importance of efficacy versus effectiveness study approaches to research. Studies using an efficacy approach are designed to investigate the benefits of an intervention under ideal and highly controlled conditions. Although this design typically minimizes threats to internal validity, the generalizability to clinical practice (external validity) may be less ideal. In contrast, studies

using an effectiveness or pragmatic approach seek to examine the outcomes of interventions under circumstances that more closely approximate clinical reality, including less standardized multimodal treatment protocols applied to more heterogeneous patient populations. Although there may be a current bias towards efficacy, a growing trend exists in the importance of studies of effectiveness in evidence-based guideline development.<sup>19</sup>

Although it is sometimes feasible to have a true control (no treatment) group<sup>28,36,37</sup> or a placebo group,<sup>11,26</sup> or both,<sup>30</sup> in trials of orthopaedic and sports physical therapy, there may be situations when it is considered unethical to withhold treatment. In such circumstances, the preferred pragmatic design is to compare the intervention of interest to another intervention considered a standard of practice in the physical therapy profes-

sion.<sup>19</sup> While the latter design sacrifices some internal validity, studies with this design<sup>21,39</sup> do not subject the patients to placebo treatments and exhibit increased external validity, allowing for greater generalizability to everyday clinical practice.

Clinical trials examining treatment effectiveness will often trade off between strong internal and strong external validity. For example, Childs and colleagues<sup>9</sup> examined the effectiveness of thrust manipulation in patients who satisfied a clinical prediction rule. The researchers exerted rigid control over the interventions delivered to assure that all patients received a standardized lumbopelvic thrust manipulation and exercise program. Conversely Deyle and colleagues<sup>13</sup> utilized a pragmatic approach to treatment in their clinical trial, which examined the effectiveness of manual therapy and exercise for the management of knee pain secondary to osteoarthritis. Clinicians selected the regions and specific type of manual intervention the patients would receive based upon their individual clinical examinations.

After formulating an appropriate foreground question and performing an effective search, one may discover that no systematic reviews of randomized controlled trials (RCTs) or single RCTs have been conducted on the topic of interest. In these situations, clinical decisions must be based on evidence from sources lower on the evidence hierarchy. While nonrandomized study designs are reported to provide much weaker evidence than randomized designs,<sup>25</sup> it should be recognized that RCTs are not always necessary, especially when the treatment effects are dramatic and readily recognizable,<sup>14</sup> or when rapid changes occur in chronic conditions with well-documented natural history. For example, the natural history of adhesive capsulitis of the shoulder may involve symptoms and impairments for up to 2 years.<sup>52</sup> However, in a prospective single-group study, Placzek and colleagues<sup>42</sup> performed translational glenohumeral thrust manipulation on 31 patients with adhesive capsulitis of the

shoulder. The results demonstrated rapid and dramatic improvements in range of motion, pain, and function. Based on the immediate improvements in this population with a known natural history, one can more readily accept a cause-and-effect relationship between the intervention and the outcomes. However, in circumstances where a great deal of uncertainty exists between alternative interventions, the RCT is still the best method to resolve these uncertainties.<sup>14</sup>

### Randomization and Baseline Homogeneity of Groups

Randomization should theoretically ensure that each group of subjects is similar at baseline so that no extraneous variables (such as known and unknown prognostic factors) compete with the intervention to explain observed outcomes. Extraneous variables that could potentially affect outcomes in studies of treatment effectiveness include patient, age, race, gender, symptom duration, condition severity, comorbidities, intellectual status, motivation, and treatment expectations. Although randomization should ideally produce observed homogenous groups at baseline, there is always a chance, particularly with small samples, that groups may be dissimilar in important known and unknown prognostic factors, which may affect group homogeneity. For this reason a reader performing a critical appraisal must independently judge the extent to which groups are similar in key prognostic factors. This task can usually be accomplished by inspecting values in the table reporting baseline patient characteristics. The implication of baseline dissimilarities for overall appraisal of the evidence will depend on whether the specific dissimilar attribute constitutes a prognostic factor for the outcome being studied. For example, if gender does not pose a competing explanation for why groups might have different posttreatment outcome scores, dissimilar group-wise proportions of men and women at baseline will not seriously affect the interpretability of comparisons between

groups. An analysis of covariance (ANCOVA) provides an adjusted comparison between groups, in which important covariates are identified and used to make statistical adjustments to posttreatment group mean scores for outcomes of interest. The ANCOVA may produce results that are easier to interpret and may have more precision than an unadjusted analysis of variance.

### Concealment of Allocation to Groups

Even when randomization procedures are followed, bias from investigators influencing subject enrollment and group composition can threaten validity if allocation to groups is not concealed from those enrolling subjects in the study.<sup>49</sup> Concealment of group allocation is typically accomplished by first obtaining informed consent and enrolling a new subject into a clinical trial, and only then opening a sealed envelope obtained from a locked filing cabinet to reveal group assignment. Readers performing a critical appraisal should look for language in a published RCT reflecting these or similar methods for concealing group allocation. Interestingly, despite strong rationale for concealment of group allocation, a study of 2297 RCTs in the PEDro database revealed that only 16% of these studies reported concealment of allocation.<sup>38</sup>

### Blinding

In an attempt to minimize the effect of rater or subject bias, studies use various blinding schemes. There are 4 categories of study participants who should ideally be blinded to group assignment: (1) patients, (2) treating clinicians, (3) data collectors, and (4) data analysts.<sup>48</sup> Although it is usually feasible to blind those from all 4 categories in a pharmaceutical study, this is usually not possible in studies of physical therapy interventions. Physical therapists are usually aware of the treatment they are delivering (rater bias); blinding the patient with sham interventions may be difficult or impossible.<sup>29</sup> Additionally, most current Institutional Review Boards require that patients are aware of

all of the possible interventions they may receive as part of the informed consent process, which provides another barrier to complete patient blinding. However, the person measuring outcomes in physical therapy trials can almost always be blinded to group assignment in order to minimize rater bias. Authors should reveal this antibias protection with clear language, such as, “An investigator, who was blinded to the treatment condition... performed this measurement.”<sup>53</sup> Nevertheless, Moseley et al<sup>38</sup> found that only 5% of studies in the PEDro database reported using blinded outcome assessors. Therefore, a reader performing a critical appraisal must decide whether blinding occurred and, if not, how serious a threat to validity is posed by this problem. The implication of nonblinding for overall appraisal of the evidence will depend on the context and particulars of the study. For example, self-report outcome tools (eg, Oswestry Scale, WOMAC scale, etc) are not as readily subject to rater bias, even when the outcomes assessor is not blinded to group assignment.

### **Completeness of Follow-up: Intention-to-Treat Principle**

The authors should report the reasons for any patient dropouts from the study and identify any patients who were lost to follow-up.<sup>6</sup> It is important for the clinician to know if the patient withdrew from the study due to full resolution of symptoms, for reasons unrelated to the study, or because the person experienced a worsening in status that was directly or potentially related to the examination or treatment program provided by the study protocol. When subjects are lost to follow-up, it may still be possible to include data from all subjects in the final data set using an intention-to-treat (ITT) approach, which has been used in a variety of recently published studies.<sup>8,12,55</sup> Although several imputation methods are possible, the most common approach to substitute a value for a missing data point with continuous scale outcomes is to carry the last known value forward to any subsequent times

in the dataset that were missed because the subject dropped out of the study. This form of ITT analysis allows subject data that otherwise would have been removed to be used in the final analysis.

Results of the ITT analysis are often compared to results from a separate analysis based only on the subjects who completed all aspects of the study (a per-protocol analysis or completer analysis). A per-protocol analysis will usually overestimate the benefit of treatment if there are dropouts in a study.<sup>46</sup> If significant results favor the treatment group with both per-protocol and ITT analyses, this strengthens the findings and suggests that the dropout rate did not threaten validity of the results. On the other hand, if significant benefit is found with a per-protocol analysis, but not with the ITT analysis, this seriously undermines the initial findings.

An ITT analysis is not always necessary to comply with the intention-to-treat principle. For example, Herrington and Al-Sherhi<sup>28</sup> randomized 45 subjects with patellofemoral pain into 3 treatment groups. At the end of the 6-week intervention, all 45 patients had complied with treatment regimens or control conditions and were available for posttreatment outcomes measurement. Therefore the intention-to-treat principle was satisfied without an ITT analysis.

A reader performing a critical appraisal must decide whether an ITT analysis was indicated, whether one was performed, and, if not, how serious a threat to validity is posed by this problem. For example, if the RCT is a “negative trial” in which no statistically significant benefit of treatment was found, failure to perform an ITT analysis may not threaten the validity of the results. Savolainen et al<sup>47</sup> examined the effectiveness of active or passive treatment in the management of neck and shoulder pain. The researchers initially randomized 75 subjects to receive either an active range-of-motion exercise program or thrust manipulations directed at the thoracic spine. At the time of the 6-month follow-up, 34 subjects

(45%) had dropped from the study and their data were not available to include in the analysis. While the authors state that the dropouts were related to work pressure, impossible timetables, etc, the high dropout rate in the absence of an ITT analysis casts considerable doubt on the results of this particular study. Therefore, clinicians should be very cautious about applying the results of such studies in their clinical practice.

Dichotomous outcomes (presence versus absence of some clinically important end point) require a different form of imputation in an ITT analysis. The most common method of imputation for these outcomes is the worst-case-scenario analysis.<sup>46</sup> This entails assigning a bad outcome to every dropout patient in the treatment group; every dropout patient from the control (or comparison) group is assigned a good outcome. This is a conservative approach that makes it more difficult to find a statistically significant difference favoring the treatment group. As with continuous scale outcomes, if the dropout rate is high and no ITT analysis was performed for dichotomous outcomes, any statistically significant findings in favor of the treatment are suspect.

Although some authors<sup>50</sup> suggest a rule of thumb for maximum acceptable dropout rate (20% is common), others<sup>24</sup> state that rules of thumb are misleading. Guyatt et al<sup>24</sup> recommend, instead, that authors perform a worst-case analysis, as described above for dichotomous outcomes: if results favoring the treatment group are still statistically significant with this conservative approach, then dropout rates exceeding 20% are not threats to the validity of the results.

### **Equivalent Experience of Groups, Apart from Treatment of Interest**

It is possible to introduce bias into a study of treatment if there are important between-group differences in the overall patient experience, aside from the treatment itself. For example, if one group receives more time with treating therapists

# CLINICAL COMMENTARY

or receives cointerventions in addition to the intended treatment, this disparity can present a competing explanation for any observed benefits. For this reason, investigators often try to structure study protocols to minimize any unnecessary between-group differences in overall experience during the study, other than the treatment(s) of interest. For example, in a study comparing general exercise plus trunk muscle stabilization to general exercise only for patients with recurrent low back pain, Koumantakis et al<sup>31</sup> took careful precautions to ensure that both groups were treated the same, except for the difference of interest. Patients in both groups warmed up with identical regimens prior to targeted exercise performance. The same treating therapist conducted the exercise sessions for both groups. Both groups received the same frequency and duration of exercise. Subjects in both groups kept exercise diaries and were asked to repeat the exercises at home. Finally, all subjects in the study received the same informational booklet about low back pain management. Readers performing critical appraisal need to decide whether any between-group disparities in the overall research experience constitute unintended beneficial cointerventions for the treatment group, and, if so, whether this problem threatens validity of the results.<sup>50</sup>

## Validity Scores Available Online

Clinicians may gain some insight into relative overall validity for published trials for treatments relevant to physical and occupational therapists by searching the PEDro<sup>3</sup> and OTSeeker<sup>2</sup> databases. Both databases use the PEDro scale to rate overall quality of clinical trials based on adherence to the principles of validity discussed above. The PEDro scale ranges from 0 to 10, with 1 point assigned for adequate protection against each of 10 validity threats (TABLE 3). The PEDro scale has been shown to have fair to good interrater reliability (ICC<sub>1,1</sub> = 0.68; 95% confidence interval [CI]: 0.57 to 0.76).<sup>35</sup> PEDro scale scores posted online are independently

TABLE 3	ELEMENTS OF THE PEDRO SCALE <sup>3</sup> FOR RANDOMIZED CONTROLLED TRIALS*
<ol style="list-style-type: none"> <li>1. Eligibility criteria were specified</li> <li>2. Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received)</li> <li>3. Allocation was concealed</li> <li>4. The groups were similar at baseline regarding the most important prognostic indicators</li> <li>5. There was blinding of all subjects</li> <li>6. There was blinding of all therapists who administered the therapy</li> <li>7. There was blinding of all assessors who measured at least one key outcome</li> <li>8. Measures of at least 1 key outcome were obtained from more than 85% of the subjects initially allocated to groups</li> <li>9. All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least 1 key outcome was analysed by "intention to treat"</li> <li>10. The results of between-group statistical comparisons are reported for at least 1 key outcome</li> <li>11. The study provides both point measures and measures of variability for at least 1 key outcome</li> </ol>	
<p>* Although the PEDro scale includes all 11 items listed, only items 2 through 11 are used for the PEDro scale score, which ranges from 0 (no validity protections satisfied) to 10 (all validity issues satisfied). One point is awarded for each validity issue satisfied (only items 2 through 11). Adapted with permission from the Physiotherapy Evidence Database (PEDro).</p>	

confirmed when annotated as such. A PEDro search for "Pilates AND low back pain" yields 3 clinical trials.<sup>3</sup> Judging by PEDro scale scores alone, seeking best evidence would suggest initial preference for the study by Rydeard et al<sup>45</sup> (PEDro scale score of 8) over the study by Gladwell et al<sup>22</sup> (PEDro scale score of 5). Although a convenient and freely available resource to get a quick indication of validity for many trials, consulting PEDro scale scores does not obviate the need for independent professional judgments regarding validity threats as part of the critical appraisal process.

## Validity of Systematic Reviews

Systematic Reviews are conducted by employing explicit methods for exhaustive searching and selective inclusion of original studies for analysis based on specified methodologic criteria. Systematic reviews of treatment studies can be performed for RCTs, cohort studies, or case control studies (TABLE 2). Readers must take care to distinguish systematic reviews from unsystematic "literature reviews" in which authors survey published literature without explicit search criteria or without specified selection criteria for studies to include in the review. These

unsystematic reviews are more common in older literature and may be considered expert opinion (level 5 evidence), because conclusions by authors are subject to multiple forms of bias.

Oxman et al<sup>40</sup> suggest 4 key validity issues that one should consider when critically appraising a systematic review: (1) authors should address a clinical foreground question that is explicit and sufficiently narrow in scope; (2) the search for relevant studies should be detailed, exhaustive, and fully revealed; (3) authors should use and report explicit criteria for assessing methodologic quality of studies considered for inclusion or exclusion in the review; (4) adequate reliability between 2 or more assessors should be reported for decisions about which studies to include, quality of included studies, and data extracted from original studies.

## Validity of Clinical Practice Guidelines

Clinical practice guidelines are another form of synthesized evidence wherein broader cultural, societal, and patient interest considerations are integrated with the best available evidence. Although the quality and completeness of practice guidelines can vary, the best guidelines are created by panels of experts repre-

senting a spectrum of constituencies, using EBP principles to generate explicit grades of recommendations supported by specified levels of evidence. Perhaps the best single online resource for clinical practice guidelines is located at the National Guideline Clearinghouse ([www.guideline.gov](http://www.guideline.gov)) sponsored and maintained by the Agency for Healthcare Research and Quality. A simple search of that site in July 2008 using the words “physical therapy” yielded 1163 guidelines.

Guyatt et al<sup>24</sup> suggest 4 key validity issues for consideration when critically appraising a clinical practice guideline: (1) recommendations should broadly consider all relevant patient groups, treatment options, and outcomes; (2) recommendations should be linked to the best available evidence; (3) values and preferences should be explicitly linked to treatment outcomes; (4) recommendations should be accompanied by grades of recommendation that indicate strength of associated evidence. The AGREE Collaboration (<http://www.agreecollaboration.org>) has recently developed an instrument to assess the quality of clinical guidelines.

## SUMMARY

**S**UCCESSFUL INTEGRATION OF individual clinical expertise, patient values and expectations, and the best available evidence requires familiarity and skill with the EBP process. Formulating an appropriate question, performing an efficient literature search, critically appraising the best available evidence relative to treatment, applying the best evidence to clinical practice, and ultimately assuring proficiency with the process will ultimately lead to improved care for our patients. Developing proficiency in the 5-step EBP process requires strong dedication and effort, and can be quite challenging initially. However, as with any skill attainment, the process gets easier and faster with practice and experience.

This first commentary in a 2-part se-

ries reviewed principles relating to formulating foreground questions, searching for the best available evidence for treatment effectiveness, and determining validity of results in studies of interventions for orthopaedic and sports physical therapy. Part 2 of this series will assist readers in interpreting results, applying results to patient care, and evaluating proficiency with EBP skills. ●

## REFERENCES

- Centre for Evidence-Based Medicine. Available at: <http://www.cebm.net>. Accessed July 7, 1995.
- OT Seeker. Available at: <http://www.otseeker.com>. Accessed August 6, 2002.
- Physiotherapy Evidence Database (PEDro). Available at: <http://www.pedro.fhs.usyd.edu.au>. Accessed August 6, 2007.
- PubMed Clinical Queries. Available at: <http://www.ncbi.nlm.nih.gov/entrez/query/static/clinical.shtml>. Accessed July 17, 2007.
- Agency for Healthcare Research and Quality. Systems to rate the strength of scientific evidence. *Evidence Report/Technology Assessment*. 2002;47:1-11.
- Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663-694.
- American Physical Therapy Association. Guide to Physical Therapist Practice. Second Edition. *Phys Ther*. 2001;81:9-746.
- Bisset L, Beller E, Jull G, Brooks P, Darnell R, Vicenzino B. Mobilisation with movement and exercise, corticosteroid injection, or wait and see for tennis elbow: randomised trial. *BMJ*. 2006;333:939. <http://dx.doi.org/10.1136/bmj.38961.584653.AE>
- Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med*. 2004;141:920-928.
- Childs JD, Whitman JM. Advancing physical therapy practice: the accountable practitioner. *J Orthop Sports Phys Ther*. 2005;35:624-627. <http://dx.doi.org/10.2519/jospt.2005.0110>
- Cleland JA, Childs JD, McRae M, Palmer JA, Stowell T. Immediate effects of thoracic manipulation in patients with neck pain: a randomized clinical trial. *Man Ther*. 2005;10:127-135. <http://dx.doi.org/10.1016/j.math.2004.08.005>
- Deyle GD, Allison SC, Matekel RL, et al. Physical therapy treatment effectiveness for osteoarthritis of the knee: a randomized comparison of supervised clinical exercise and manual therapy procedures versus a home exercise program. *Phys Ther*. 2005;85:1301-1317.
- Deyle GD, Henderson NE, Matekel RL, Ryder MG, Garber MB, Allison SC. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. A randomized, controlled trial. *Ann Intern Med*. 2000;132:173-181.
- Djulbegovic B, Frohlich A, Bennett CL. Acting on imperfect evidence: How much regret are we ready to accept? *J Clin Oncol*. 2005;23:6822-6825. <http://dx.doi.org/10.1200/JCO.2005.06.007>
- Djulbegovic B, Hozo I, Lyman GH. Linking evidence-based medicine therapeutic summary measures to clinical decision analysis. *Med-GenMed*. 2000;2:E6.
- Doig GS, Simpson F. Efficient literature searching: a core skill for the practice of evidence-based medicine. *Intensive Care Med*. 2003;29:2119-2127. <http://dx.doi.org/10.1007/s00134-003-1942-5>
- Domholdt E. *Physical Therapy Research*. 2nd ed. Philadelphia, PA: W.B. Saunders Company; 2000.
- Fritz J, Flynn TW. Autonomy in physical therapy: less is more. *J Orthop Sports Phys Ther*. 2005;35:696-698.
- Fritz JM, Cleland J. Effectiveness versus efficacy: more than a debate over language. *J Orthop Sports Phys Ther*. 2003;33:163-165.
- Fritz JM, Wainner RS. Examining diagnostic tests: an evidence-based perspective. *Phys Ther*. 2001;81:1546-1564.
- Gerber JP, Marcus RL, Dibble LE, Greis PE, Burks RT, Lastayo PC. Safety, feasibility, and efficacy of negative work exercise via eccentric muscle activity following anterior cruciate ligament reconstruction. *J Orthop Sports Phys Ther*. 2007;37:10-18. <http://dx.doi.org/10.2519/jospt.2007.2362>
- Gladwell V, Head S, Hagger M, Beneke R. Does a program of pilates improve chronic non-specific low back pain? *J Sport Rehab*. 2006;15:338-350.
- Green LW. Public health asks of systems science: to advance our evidence-based practice, can you help us get more practice-based evidence? *Am J Public Health*. 2006;96:406-409. <http://dx.doi.org/10.2105/AJPH.2005.066035>
- Guyatt G, Hayward R, Richardson S, et al. Moving from evidence to action. In: Guyatt G, Rennie D, eds. *User's Guide to the Medical Literature: Essentials of Evidence-Based Practice*. Chicago, IL: American Medical Association; 2002.
- Guyatt G, Rennie D. *User's Guide to the Medical Literature: Essentials of Evidence-Based Practice*. Chicago, IL: American Medical Association; 2002.
- Hall T, Chan HT, Christensen L, Odenthal B, Wells C, Robinson K. Efficacy of a C1-C2 self-sustained natural apophyseal glide (SNAG) in the management of cervicogenic headache. *J Orthop Sports Phys Ther*. 2007;37:100-107. <http://dx.doi.org/10.2519/jospt.2007.2379>
- Herbert RD. Randomisation in clinical trials. *Aust J Physiother*. 2005;51:58-60.
- Herrington L, Al-Sherhi A. A controlled trial of weight-bearing versus non-weight-bearing exer-



- cises for patellofemoral pain. *J Orthop Sports Phys Ther.* 2007;37:155-160. <http://dx.doi.org/10.2519/jospt.2007.2433>
29. Horn SD, DeJong G, Ryser DK, Veazie PJ, Teraoka J. Another look at observational studies in rehabilitation research: going beyond the holy grail of the randomized controlled trial. *Arch Phys Med Rehabil.* 2005;86:S8-S15. <http://dx.doi.org/10.1016/j.apmr.2005.08.116>
  30. Hyland MR, Webber-Gaffney A, Cohen L, Lichtman PT. Randomized controlled trial of calcaneal taping, sham taping, and plantar fascia stretching for the short-term management of plantar heel pain. *J Orthop Sports Phys Ther.* 2006;36:364-371. <http://dx.doi.org/10.2519/jospt.2006.2078>
  31. Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus general exercise versus general exercise only: randomized controlled trial of patients with recurrent low back pain. *Phys Ther.* 2005;85:209-225.
  32. Lowe HJ, Barnett GO. Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches. *JAMA.* 1994;271:1103-1108.
  33. MacDermid JC. An introduction to evidence-based practice for hand therapists. *J Hand Ther.* 2004;17:105-117. <http://dx.doi.org/10.1197/j.jht.2004.02.001>
  34. Maher CG, Sherrington C, Elkins M, Herbert RD, Moseley AM. Challenges for evidence-based physical therapy: accessing and interpreting high-quality evidence on therapy. *Phys Ther.* 2004;84:644-654.
  35. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther.* 2003;83:713-721.
  36. McClure P, Balaicuis J, Heiland D, Broersma ME, Thorndike CK, Wood A. A randomized controlled comparison of stretching procedures for posterior shoulder tightness. *J Orthop Sports Phys Ther.* 2007;37:108-114. <http://dx.doi.org/10.2519/jospt.2007.2337>
  37. Moiler K, Hall T, Robinson K. The role of fibular tape in the prevention of ankle injury in basketball: A pilot study. *J Orthop Sports Phys Ther.* 2006;36:661-668. <http://dx.doi.org/10.2519/jospt.2006.2259>
  38. Moseley AM, Herbert RD, Sherrington C, Maher CG. Evidence for physiotherapy practice: a survey of the Physiotherapy Evidence Database (PEDro). *Aust J Physiother.* 2002;48:43-49.
  39. O'Leary S, Jull G, Kim M, Vicenzino B. Specificity in retraining craniocervical flexor muscle performance. *J Orthop Sports Phys Ther.* 2007;37:3-9. <http://dx.doi.org/10.2519/jospt.2007.2237>
  40. Oxman A, Guyatt G, Rennie D. Summarizing the evidence. In: Guyatt G, Rennie D, eds. *User's Guide to the Medical Literature: Essentials of Evidence-Based Practice.* Chicago, IL: American Medical Association; 2002:
  41. Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. *J Epidemiol Community Health.* 2003;57:527-529.
  42. Placzek JD, Roubal PJ, Freeman DC, Kulig K, Nasser S, Pagett BT. Long-term effectiveness of translational manipulation for adhesive capsulitis. *Clin Orthop Relat Res.* 1998;181-191.
  43. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 2nd ed. Upper Saddle River, NJ: Prentice Hall Health; 2000.
  44. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club.* 1995;123:A12-13.
  45. Rydeard R, Leger A, Smith D. Pilates-based therapeutic exercise: effect on subjects with nonspecific chronic low back pain and functional disability: a randomized controlled trial. *J Orthop Sports Phys Ther.* 2006;36:472-484. <http://dx.doi.org/10.2519/jospt.2008.2669>
  46. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM.* 2nd ed. London, UK: Harcourt Publishers Ltd; 2000.
  47. Savolainen A, Ahlberg J, Nummila H, Nissinen M. Active or passive treatment for neck-shoulder pain in occupational health care? A randomized controlled trial. *Occup Med (Lond).* 2004;54:422-424. <http://dx.doi.org/10.1093/occmed/kqh070>
  48. Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. *Ann Intern Med.* 2002;136:254-259.
  49. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273:408-412.
  50. Straus SE, Richardson WS, Glasziou P, Haynes RB. *Evidence-based Medicine.* 3rd ed. Edinburgh, UK: Elsevier/Churchill Livingstone; 2005.
  51. van der Weijden T, CJ IJ, Dinant GJ, van Duijn NP, de Vet R, Buntinx F. Identifying relevant diagnostic studies in MEDLINE. The diagnostic value of the erythrocyte sedimentation rate (ESR) and dipstick as an example. *Fam Pract.* 1997;14:204-208.
  52. Vermeulen HM, Obermann WR, Burger BJ, Kok GJ, Rozing PM, van Den Ende CH. End-range mobilization techniques in adhesive capsulitis of the shoulder joint: A multiple-subject case report. *Phys Ther.* 2000;80:1204-1213.
  53. Vicenzino B, Branjerdporn M, Teys P, Jordan K. Initial changes in posterior talar glide and dorsiflexion of the ankle after mobilization with movement in individuals with recurrent ankle sprain. *J Orthop Sports Phys Ther.* 2006;36:464-471. <http://dx.doi.org/10.2519/jospt.2006.2265>
  54. Wade DT. Randomized clinical trials in Clinical Rehabilitation. *Clin Rehabil.* 2005;19:233-236.
  55. Whitman JM, Flynn TW, Childs JD, et al. A comparison between two physical therapy treatment programs for patients with lumbar spinal stenosis: a randomized clinical trial. *Spine.* 2006;31:2541-2549. <http://dx.doi.org/10.1097/01.brs.0000241136.98159.8c>



**MORE INFORMATION**  
[WWW.JOSPT.ORG](http://WWW.JOSPT.ORG)