

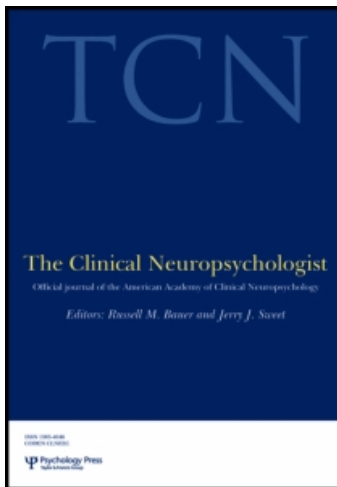
This article was downloaded by: [Chelune, Gordon J.]

On: 27 September 2008

Access details: Access Details: [subscription number 903073914]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## The Clinical Neuropsychologist

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t113721659>

### Evidence-Based Research and Practice in Clinical Neuropsychology

Gordon J. Chelune<sup>a</sup>

<sup>a</sup> Center for Alzheimer's Care, Imaging and Research, Department of Neurology, University of Utah School of Medicine, Salt Lake City, UT, USA

First Published on: 25 September 2008

**To cite this Article** Chelune, Gordon J.(2008)'Evidence-Based Research and Practice in Clinical Neuropsychology',The Clinical Neuropsychologist,

**To link to this Article:** DOI: 10.1080/13854040802360574

**URL:** <http://dx.doi.org/10.1080/13854040802360574>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## EVIDENCE-BASED RESEARCH AND PRACTICE IN CLINICAL NEUROPSYCHOLOGY

**Gordon J. Chelune**

*Center for Alzheimer's Care, Imaging and Research, Department of Neurology,  
University of Utah School of Medicine, Salt Lake City, UT, USA*

*While a definition for evidence-based clinical neuropsychological practice (EBCNP) has yet to emerge, it is likely to integrate the same core features as evidence-based medicine; namely, best research evidence, clinical expertise, and individual patient needs. Given the nascent stage of EBCNP, suggestions are made to advance evidence-based approaches in both research and practice. The common elements are: recognition that clinical outcomes are recorded at the level of the individual; and to be useful, outcomes research must be presented in a way that can be directly applied on a case-by-case basis. Tracking the outcomes of our clinical services in an evidence-based manner that is publicly verifiable will demonstrate the value of neuropsychological services to our patients, our referral sources, and ultimately to payers.*

**Keywords:** Evidence-based practice; Outcomes research; Advocacy.

### INTRODUCTION

The term “evidence-based practice” has become almost ubiquitous in today’s health care environment, yet it remains poorly understood and seldom practiced in clinical neuropsychology. The impetus for evidence-based practice in medicine has its roots in the outcomes movement of the 1980s when it became increasingly apparent to payers and many practitioners that a significant portion of health care expenditures in the United States was being wasted on unproven or ineffective tests and treatments (Horwitz, 1996). As a result, a more value-driven, evidence-based health care system began to emerge in which procedures and treatments were seen as having “value” (i.e., were reimbursable) only if they could be objectively demonstrated to have a positive impact on a patient’s condition in a cost-effective manner. Outcomes accountability and the management of individual patients on the basis of epidemiologic information regarding outcomes became increasingly critical to the practice of medicine (Johnson, 1997). The term “outcomes” here is used in a broad sense to refer to discrete, measurable events that impact a patient’s condition and that can be tracked both in the aggregate on a group level but also, importantly, at the level of the specific individual.

---

Address correspondence to: Gordon J. Chelune, Ph.D., Center for Alzheimer’s Care, Imaging and Research, 650 Komas Drive, Suite 106A, Salt Lake City, UT 84108-1225, USA. E-mail: Gordon.chelune@hsc.utah.edu

Accepted for publication: July 8, 2008.

The term “evidence-based medicine” (EBM) was introduced in 1992 (Evidence-Based Medicine Working Group, 1992) and is defined by Sackett and colleagues (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000) as simply “the integration of best research evidence with clinical expertise and patient values” (p. 1) with the goal to maximize clinical outcomes and quality of life for the patient. According to Sackett et al. (2000) the practice of EBM requires clinicians to focus their assessments by first translating their need for information regarding the management and treatment of a given patient into answerable questions, and then finding the best empirical evidence to answer those questions through a critical appraisal of existing evidence for its validity, impact, and applicability. This appraisal of information must be integrated with the clinician’s expertise within the context of the patient’s unique characteristics and needs.

The American Psychological Association (APA Presidential Task Force on Evidence-Based Practice, 2006) has also adopted a similar definition for evidence-based psychological practice (EBPP), and suggests that “clinical decisions be made in collaboration with the patient, based on the best clinically relevant evidence, and with consideration for the probable costs, benefits, and available resources and options” (p. 285). EBPP is generating considerable interest (see the 2007 special issue of the *Journal of Clinical Psychology*, 63, 607–707), and is believed to have the potential to improve psychological care through increased accountability and quality improvement, enhance financial reimbursement for services, and advance education, training and research (Hunsley, 2007; Kazdin, 2008).

A definition of evidence-based clinical neuropsychological practice (EBCNP) has not yet emerged, but is likely to contain many of the key elements found in EBM and EBPP, namely the integration of clinical expertise, best outcomes research, and the unique characteristics of the patient. However, EBCNP is also likely to involve additional considerations and requirements. Because neuropsychological assessment services are often rendered as consultative services to other professionals who have the primary responsibility for managing the patient’s care, the special concerns and needs of these referral sources must also be integrated into EBCNP. For example, a college student who has undergone a temporal lobectomy for intractable seizures may be referred by his/her epileptologist to determine whether s/he has incurred any significant memory loss as a result of surgery. The same patient may be referred by a guidance counselor at the college to determine whether the patient is capable of successfully returning to school. While the clinician may feel that the patient might benefit from a comprehensive assessment that includes academic achievement testing, EBCNP would suggest that such comprehensive testing would be superfluous and inappropriate for addressing the epileptologist’s concerns, yet be entirely appropriate for answering the guidance counselor’s question. To practice in a cost-effective manner, EBCNP will require practitioners to do only what is necessary to answer the specific questions asked while continuing to use their clinical judgment and expertise to know when to clarify or expand the referral question to explore alternative hypotheses and to meet the needs of the patient.

Embedded in both situations is the need for clinicians to: (1) evaluate presenting problems and complaints and convert these into answerable questions (i.e., “Has the patient incurred a memory deficit as the result of surgery?” or “Is the

patient able to successfully return to his/her academic program?"); (2) gather relevant background information from the patient's history and integrate this with the best available outcomes research specific to these questions; (3) choose a relevant test protocol that will generate meaningful data for the given patient and specific questions; (4) analyze these data within the context of the individual patient; and (5) objectively monitor the outcomes of the evaluation. Indeed, EBCNP is clinical outcomes research in practice; i.e., the scientific method applied at the level of the individual—hypothesis formation, literature review, study design and data collection, analysis, and conclusion. By tracking the outcomes of our EBCNP across patients in a manner that is publicly verifiable, we can demonstrate our "value" to our patients, our referral sources, and to payers (see Prigatano & Morrone-Stupinsky, in press). To make EBCNP the standard of practice in neuropsychology, clinicians and researchers will need to change how they routinely carry out their activities and develop methods of increasing the interactions between research and practice.

## **FOCUSED GOALS**

To advance evidence-based research and practice in neuropsychology, it is this writer's opinion that researchers and clinicians alike need to embrace two simple tenets: (1) clinical outcomes are individual events that are characterized by a change in status, performance, or other objectively defined endpoint; and (2) to be useful in the care of patients, data from outcomes research must be analyzed and packaged in such a way that they can be directly evaluated and "used" by the end-user, namely the clinician (Chelune, 2002b). In the sections to follow, I will offer some suggestions and examples on how investigators can transform "research" into "outcomes research" by augmenting traditional data analyses that are designed for making group inferences with simple base-rate analyses that permit individual case inferences. Next, I will offer several ways in which the EBCNP-minded clinician can better integrate outcomes research into practice and report writing. Although presented sequentially, I believe there is a marriage between clinical practice and outcomes research such that clinical questions should drive outcomes research as much as the latter informs practice.

## **FUTURE EFFORTS**

### **The research perspective**

Several years ago Glenn Smith (2002) posed an intriguing question for outcome researchers: "is the outcome of interest how many or how much?" (p. 432). Essentially, what is the goal of our clinical research—to compare groups on the basis of how much they differ on a given variable or how frequently they differ? To be maximally informative for clinicians, research results must be applicable in the case of the individual patient. Historically, neuropsychological research has focused on "how much," and has employed group designs with null hypothesis significance testing (NHST) methods. While it is useful to test whether the mean performance of a patient group with a condition of interest (COI) reliably differs

from a contrast group (normal controls or other relevant patient groups) on some performance parameter, reports that solely present mean group differences with associated  $p$ -values are not particularly useful in clinical practice. At a minimum, the American Psychological Association (Wilkinson & Task Force on Statistical Inference, 1999) and others (Zakzanis, 2001) have urged researchers to present effect sizes for primary outcomes as well as overlap statistics (i.e., amount of test measure overlap between groups) to better characterize the magnitude of the group differences. While statistically savvy EBCNP clinicians may be able to use these data to derive individual effect sizes for their patients' performances, such data still are difficult to apply to the individual patient being evaluated.

It is argued that researchers could easily supplement their primary NHST analyses with additional base-rate data for their groups and perhaps incorporate simple epidemiological tests such as the odds ratio (Bieliauskas, Fastenau, Lacy, & Roper, 1997) to translate neuropsychological test scores into outcomes data that can be applied at the level of the individual and better inform the clinical process. For example, it may be of interest to know that patients with intractable seizures who have high baseline memory performances show greater decrements on average on repeat memory testing following left temporal lobectomies than those with lower baseline scores (Chelune, Naugle, Luders, & Awad, 1991). However, this information is difficult to apply at the level of the individual prospective surgical candidate with a given memory score of 96. Rather, it is far more useful for the neuropsychologist and the referring epileptologist when counseling this prospective surgical candidate to know there is outcomes research suggesting patients with memory scores  $\geq 90$  at baseline have an estimated relative risk of an acquired post-surgical memory deficit that is 4.9 times higher than those with scores below 90 (Chelune & Najm, 2001). To advance EBCNP research and practice, investigators are encouraged to consider the following three suggestions.

(1) *Define neuropsychological outcomes in a manner that can be applied.* Most neuropsychological studies report mean performance differences between groups or associations between cognitive variables and other measures of interest (e.g., biomarkers). Since clinical outcomes are recorded as individual events (Chelune, 2002b), such group data are difficult to interpret at the level of the individual without converting the findings into operational terms. For example, it has been reported that patients with Alzheimer's disease (AD) can be distinguished from patients with frontal-temporal dementia (FTD) by greater discrepancies between letter and semantic fluency (Rascovsky, Salmon, Hansen, Thal, & Galasko, 2007). To be most useful in clinical practice, the end-user of this information would need to know the "size" of the discrepancy that best distinguishes the two groups; that is, the investigators would need to determine a "magnitude" of discrepancy that represents a positive versus negative outcome or indicator of AD. The choice of cutoff is somewhat arbitrary and could be adjusted to be maximally sensitive and/or specific to AD (see Sackett et al., 2000, for a discussion of SpIn and SnOut cutoffs). Nonetheless, conversion of research findings into such discrete outcomes would allow the EBCNP practitioner to directly apply the information when evaluating the individual patient. Converting group data into discrete outcomes also provides the

basis for reporting base-rate information; that is, moving from “how much” to “how many” (Smith, 2002).

(2) *Report base-rate information.* If researchers are willing to supplement their reports by defining discrete “outcomes,” it becomes a simple matter of then reporting base-rate information about the groups of interest. In the example above (Rascovsky et al., 2007), it would be helpful to the clinician to know what percentage of cases in the AD and FTD groups had positive discrepancy findings as this information could be combined with other relevant research (e.g., data concerning differences in executive functions, discrepancies between recall and retrieval in these groups, etc.) to increase the accuracy of the EBCNP practitioner’s evidence-based inferences on a case-by-case basis. As discussed in the next section, base-rate data also provide the basis for some simple, but powerful, statistical tests often used in evidence-based medicine.

(3) *Provide contingency table analyses.* Conventional NHST results can be supplemented by information about outcomes and their base rates by casting outcomes data into a simple  $2 \times 2$  contingency table similar to the one depicted in Figure 1. Additional clinically useful information about the patient groups and test operating characteristics can be derived using this contingency table format for analyses and ultimately used by clinicians to inform the individual decision-making process (Ivnik et al., 2001; Sackett et al., 2000). Table 1 summarizes a number of these statistics, and provides the reader with their simple calculation formulas that can be computed by hand or with a calculator (no computer is necessary).

Of the various statistics listed, several deserve special attention for their ability to inform clinical practice. *Sensitivity* and *Specificity* provide the clinician with important information about how sensitive the cutoff is for accurately identifying “true positive” cases with the condition of interest (COI) and how specific it is for identifying “true negative” cases without the COI. Most studies attempt to maximize both sensitivity and specificity, although clinically this may not be the most useful approach. For example, Sackett et al. (2000) suggest that in some instances we may want cutoff scores that are highly sensitive to the COI even though some false positives (normals) may be included such that a negative result almost always rules the case out (SnOut: high sensitivity and negative result, rules case out). In other situations we may want a cutoff that maximizes specificity even if some

|             |     | Condition of Interest |                       |
|-------------|-----|-----------------------|-----------------------|
|             |     | Yes                   | No                    |
| Test Factor | Yes | True Positive (TP) A  | False Positive (FP) B |
|             | No  | False Negative (FN) C | True Negative (TN) D  |

**Figure 1** Contingency table layout depicting the outcomes for the condition of interest and the results of the test factor.

patients with the COI are included (false negatives) such that an abnormal test finding almost always rules the case in (SpIn: high specificity plus positive result, rules case in).

In retrospective studies, the *Odds Ratio* (OR) provides an estimate of *relative risk* by comparing the odds of having the COI when the test factor is positive to the odds of having the COI when the test factor is negative. For example, in studies of patients with lung cancer (the COI), a history of smoking (the test factor) is associated with a higher relative risk of the disease. The odds ratio is sample specific, but provides useful information when sample is large and/or when the clinician's patient appears to be drawn from a sample similar to that described in the research. The *Likelihood Ratio* (LR) is related to the odds ratio but describes the likelihood of having the COI given the specific sensitivity and specificity of the cutoff score. By varying the cutoff score and the associated sensitivity and specificity values, the researcher can provide estimates of likelihood of having the COI for different test values.

Information about the *Pre- and Post-Test Odds* is likely to be extremely valuable for EBCNP practice, as a comparison of the two describes the efficacy of the test measure in the diagnostic process. For example, if prior to any testing a patient has a pre-test odds of having the COI of .20 (a one in five chance), but after testing the post-test odds is now .60 (three in five), our ability to correctly identify the COI has increased threefold. Data such as these are not only important for practitioners, but also provide a basis for cost-effectiveness analyses important for third party payers.

**Example application.** To illustrate how contingency table analyses can supplement traditional NHST approaches, data are presented from a study (Chelune & Stone, 2005) designed to examine processing speed deficits among patients with relapsing remitting (RRMS) and secondary progressive (SPMS)

**Table 1** Contingency table statistics, calculation formulas, and results from the application example

| Contingency table statistic                   | Calculation formula         | Example |
|---|-----------------------------|---------|
| Prevalence of the condition of interest (COI) | Cells (A + C)/N             | 20.8%   |
| Percentage of impaired test results           | (A + B)/N                   | 58.7%   |
| Percentage of normal test results             | (C + D)/N                   | 41.3%   |
| Overall correct hit rate                      | A + D/N                     | 56.4%   |
| Sensitivity (% of true positives)             | A/(A + C)                   | 86.1%   |
| Specificity (% of true negatives)             | D/(B + D)                   | 48.5%   |
| Positive predictive power                     | A/(A + B)                   | 30.5%   |
| Negative predictive power                     | D/(C + D)                   | 93.0%   |
| Odds ratio (estimated relative risk)          | (A*D)/(B*C)                 | 5.8     |
| Likelihood ratio (LR)                         | Sensitivity/(1-Specificity) | 1.67    |
| Pre-test odds                                 | Prevalence/(1-Prevalence)   | .26     |
| Post-test odds                                | Pre-test Odds*LR            | .43     |

Letters A, B, C, D refer to the cells in the contingency table (see Figure 1) starting in upper left-hand corner (i.e., true positives identified by a positive test factor) and proceeding in a clockwise fashion.  $N = \text{Cells } A + B + C + D$ .

multiple sclerosis. The impetus for this study came from referring physicians' clinical need to determine when the course of their patients with RRMS had progressed into a SPMS stage, which would signal a need to adjust the patients' medication regimes. For a given patient the referral question was frequently posed as "my patient with RRMS is complaining of increased cognitive problems, but has a stable physical exam—has the patient's course become secondary progressive?" Review of the literature suggests that cognitive processing speed deficits are common in multiple sclerosis, but particularly frequent in SPMS (Chelune, Stott, & Pinkston, 2008; Rao, Leo, Bernardin, & Unverzagt, 1991). Operationally, the hypothesis became "Do deficits in processing speed help distinguish patients with RRMS from those with SPMS?"

To empirically answer this question we retrospectively examined the performances of 274 patients with RRMS and 72 SPMS on the Wechsler Adult Intelligence Scale-III (The Psychological Corporation, 2002) Processing Speed Index (PSI), Trails B, and the Rao version of the Paced Auditory Serial Addition Test (PASAT) (Strauss, Sherman, & Spreen, 2006). While all three measures were moderately intercorrelated, they were entered into a series of stepwise discriminate function analyses (DFA) to identify which test or combination of tests best discriminated the patient groups. Results indicated that all three tests were significant independent predictors (canonical correlations from .304 to .383), but the PSI variable had the highest canonical correlation (.383) and was the only variable retained in the multivariate DFA. When a linear group contrast was performed, the RRMS group was found to have significantly higher ( $p < .0001$ ) demographically corrected PSI T-scores ( $M = 39.7$ ;  $SD = 10.8$ ) than the SPMS group ( $M = 29.0$ ;  $SD = 9.5$ ). The amount of variance ( $\eta^2$ ) accounted for by the group effect was .146, and Cohen's  $d$  was 1.02 reflecting a large effect size but still associated with a .46 overlap (OL%) between the distributions of the two groups.

From the above analyses it is clear that the RRMS group performed better than the SPMS group on measures of processing speed and that the PSI index clearly differentiated the two groups. However, these group findings are difficult to apply when the clinician must answer "Does *my* patient have SPMS?" To transform these findings into results that can be applied at the level of the individual, we cast our data into the  $2 \times 2$  table depicted in Figure 2, and

| Test<br>Factor | Condition of Interest |          |
|----------------|-----------------------|----------|
|                | SPMS                  | RRMS     |
| $T_c \leq 40$  | 62<br>A               | 141<br>B |
| $T_c > 40$     | 10<br>C               | 133<br>D |

**Figure 2** Contingency table showing the frequency of patients with secondary progressive multiple sclerosis (SPMS) and relapsing remitting multiple sclerosis (RRMS) who had Processing Speed Index (PSI) scores above and below a demographically corrected T-score ( $T_c$ ) of 40.

performed a series of contingency table analyses, with the results summarized in the “Example” column of Table 1. The first step was to choose a PSI cutoff score to define our Test Factor. Following the guidelines suggested by Heaton and colleagues (Heaton, Taylor, & Manly, 2003), we chose demographically corrected T-scores (Tc) one standard deviation or greater below the mean ( $Tc \leq 40$ ; approximately the bottom 15<sup>th</sup> percentile) to reflect *impairment*. The resulting sensitivity of this cutoff was 86.1%, whereas the specificity was 48.5%. Thus, PSI scores  $\leq Tc 40$  were able to identify most of the patients with SPMS in this sample but also misclassified many RRMS cases. Nonetheless, the OR for this cutoff was 5.8, indicating that patients with PSI scores  $\leq Tc 40$  were nearly six times more likely to have SPMS than RRMS. Comparing the pre- (.26) and post-test odds (.43), we see that using PSI scores  $\leq Tc 40$  to classify patients as having SPMS nearly doubled our chances of correctly identifying these patients. Finally, if we compute a *Receiver Operating Curve* (ROC) to plot the sensitivities and specificities for predicting SPMS group membership associated with different PSI scores, we can derive likelihood ratios (LR) for each level of PSI performance. For example, the LR associated with a Tc of 40 is about 1.7, whereas Tc scores of 36 yield an LR of 2.2, and Tc scores of 30 have an LR of 3.4 in this sample. Thus, while the post-test odds using a cutoff of  $Tc \leq 40$  is .43, they are .88 when using a cutoff of  $Tc \leq 30$ . Plotting LR values against different test values is likely to be particularly helpful to the clinician in EBNP practice.

In summary, supplementary contingency table analyses can enhance our group findings by providing us information that can be directly applied at the level of the individual to address the question “Does my patient have SPMS?” For example, we can now say that a patient with MS who obtains a PSI score of  $Tc = 36$  within the context of the referral question of “Does s/he have SPMS?” is nearly six times more likely to have SPMS than RRMS. Further, the individual risk of SPMS associated with a PSI Tc-score of 36 is more than twice that of those scoring above this level.

## THE CLINICAL PERSPECTIVE

Just as investigators need to provide more information about individual outcomes in their studies to facilitate EBCNP, it is incumbent on practitioners to use this information in their evaluations of the individual patient to make EBCNP the standard of care and a financially viable service within today’s value-drive health care system. All too often neuropsychological reports provide lengthy descriptions of a patient’s test results and then, as if magically, come to a conclusion that the “overall test results are consistent with ...” a post-lobectomy memory deficit, traumatic brain-injury, secondary progressive multiple sclerosis, et cetera. But are these conclusions based on objective evidence, and if so, on what basis? Except perhaps in forensic cases or complex cases where multiple hypotheses are being considered, practitioners are seldom prepared to answer these questions. If we are to practice EBCNP, we must understand and begin to use outcomes research and base-rate information to guide our practice and shape our reports for consumers. This can be accomplished by following three basic practice guidelines inherent in all evidence-based practice (Sackett et al., 2000).

(1) *Convert referral questions into answerable questions.* As noted by Sackett and colleagues (2000), the first step in evidence-based practice is to convert our “need to know” into answerable questions. Rarely are we asked to simply describe a patient’s cognitive skills. Rather, our consumers typically want to know “Does this patient have a memory deficit?” “Has this patient suffered a traumatic brain injury?” “Can this patient return to work?” “Has this patient’s MS become secondary progressive?” “Has this patient’s dementia progressed?” and the like. To answer these questions the clinician must define a priori what will constitute a positive (“yes”) or negative (“no”) outcome. This will, in part, be based on the patient’s unique history and circumstances and the clinician’s expertise, but also will require knowledge of the relevant outcomes literature related to the question, the choice assessment tools best suited to answering the question, and the specific pattern or magnitude of test findings needed to constitute a positive versus negative decision. Only by operationalizing the criteria for our conclusions and making these publicly verifiable (i.e., explicit in our reports) can we examine and track their accuracy and their potential benefit to the patient in an evidence-based manner.

(2) *Use base-rate information when reporting test results.* One way to objectively define an outcome is to place it within the context of base-rate expectations. A base rate simply refers to how frequently an event, condition, or test finding occurs in a specified population, which may be the general population, a carefully screened “normal” population, or a patient population with a documented disease or disorder. It is important for the clinician to know the characteristics of the population from which the base-rate data are derived. It is also important to recognize that some base rates are derived from data that are normally distributed, such as the IQ and factor scores on the Wechsler Adult Intelligence Scale, where the scores reflect standardized deviations from the population mean. Other test measures, such as the recognition scores on the Wechsler Memory Scale-III, are asymmetric and not normally distributed, and base rates are best reported as a cumulative percentage (Tulsky, Chiaravalloti, Palmer, & Chelune, 2003). There is nothing about the base rates themselves that defines a given score as “abnormal” or “impaired,” merely whether it is a common or rare finding in the reference population. However, if a patient’s specific test result is rare in the general population (or reference group) but outcomes research demonstrates that it occurs frequently in a target population of interest, we can infer with some level of probability that the patient more likely belongs to the target population. In the absence of outcomes research to guide us, we can still identify the finding as unusual (i.e., rare), if not impaired.

Most contemporary neuropsychological tests have normative information about the base rates of various test performances in a “normal” population, and in some cases for small patient samples. Also, increasingly these normative data are adjusted for relevant demographic factors so that an individual comparison standard is generated. Assuming that the test scores are normally distributed, we can describe a patient’s performances in terms of how far they deviate from demographic expectations (Busch, Chelune, & Suchy, 2006). For example, an 81-year-old woman with 14 years of education who obtains a PSI of 91 on the WAIS-III falls in the “average” range for the general population in her age group, but is 1.5 standard deviations (T-score of 35) below expectations for her

demographic cohort. The clinical significance of this test result will depend on the patient's history and the referral question at hand. If the patient has been bedridden and in the hospital for 3 weeks for a hip fracture and is now being considered for discharge home in a wheelchair, the finding may not be considered abnormal or a contradiction to discharge. However, if the patient is referred to a dementia clinic for work-up and has a history of diabetes and hypertension, this finding in the context of other results might help distinguish vascular cognitive impairment versus Alzheimer's disease (Levy & Chelune, 2007).

Using the discrepancy between various factor scores derived from the Third Edition of the Wechsler Adult Intelligence and Memory Scales and the direction of these discrepancies, Hawkins and Tulsy (2003) describe how the base rates of discrepancy scores can inform the clinical decision-making process and be helpful in classifying patients with lateralized brain lesions, traumatic brain injuries and neurodegenerative conditions. However, the test user should be mindful and note whether the base rates for the discrepancy scores are one-tailed or two-tailed (Tulsy, Rolhus, & Zhu, 2000); that is, whether the base rates represent the cumulative percentage of the absolute values of the discrepancy regardless of the direction or whether the base rates are for directional values. The base rates of test-retest reliable change scores are also increasingly used to monitor meaningful change in cognitive status as a function of disease progression, treatment response, surgical or pharmacological intervention, and recovery of function (Busch et al., 2006; Chelune, 2002a; Collie, Darby, Falletti, Silbert, & Maruff, 2002; Strauss et al., 2006). Whether the clinician is looking at a single score or the discrepancy between scores, at a basic level simply reporting the base-rate information for the test findings that we interpret as meaningful for our conclusions allows others to follow our logic and objectively evaluate our work. This is the beginning of EBCNP practice—making our inferences and opinions publicly verifiable.

(3) *Incorporate and use outcomes research to guide our assessments.* It is incumbent on the clinician in EBCNP to use the best available information to guide and inform the evaluation of the patient. This begins with selecting published test instruments and normative data most relevant to the clinical questions being posed. Test measures with poor reliability are likely to yield spurious findings, and those with poor sensitivity to the COI will yield overly conservative results (Strauss et al., 2006). If base-rate information is to be derived from normative samples, the clinician must know the characteristics of the reference sample to which the patient is to be compared, how it was obtained, and what inclusion and exclusion criteria were used in composing the sample (Busch et al., 2006; Ritchie, Frerichs, & Tuokko, 2007; Strauss et al., 2006).

Once the referral question has been converted into an answerable question and we have selected the most appropriate tests and norms to evaluate our patient, we can begin to seek relevant outcomes studies in the literature to further guide our evaluation and interpretation of the patient's test data. Textbooks would seem to be an obvious source of relevant information, but often the information contained is dated and presented in broad strokes, making it difficult to use at the level of the individual patient. Journals, especially electronic journals specifically devoted to evidence-based practice such as the *Cochrane Report*, *Journal of Evidence-Based*

*Medicine*, and *Evidence-Based Mental Health*, are the best sources of current outcomes data.

Unfortunately there is no specific repository for neuropsychological outcomes research as yet, and the EBCNP clinician will most likely need to conduct a search of the research literature. Fortunately, the Internet offers ready access to a wide array of journals and databases through public gateways such as *PubMed* ([www.ncbi.nlm.nih.gov/PubMed](http://www.ncbi.nlm.nih.gov/PubMed)) or more specialized databases such as *PsycArticles* or *PsycINFO* that may be available through university libraries, and finding patient-relevant studies is relatively easy if the clinician has clearly defined questions.

Sorting through the “hits” from electronic searches for the best empirical evidence may be challenging. For single studies, the clinician should look for well-designed studies that follow the recommendations of the American Psychological Association’s Task Force on Statistical Inference (Wilkinson & Task Force on Statistical Inference, 1999); namely, studies that have clear designs, a specified population that the study is designed to reflect, and sampling procedures that characterize the inclusion and exclusion criteria for selecting participants, well-defined variables that are explicitly assessed, sample sizes that are appropriately powered to test the hypotheses, appropriate analyses, reporting of effect sizes for the primary outcomes measures, and interpretations that are credible, robust, and generalizable to other patients that mirror the characteristics of the sample studied. Well-designed studies that report base-rate analyses such as ORs, pre- and post-test odds, and LR<sub>s</sub> deserve special attention. However, it is rare that a single study will be definitive. Therefore, clinicians should also look for systematic reviews that summarize the research literature on the topic of interest, specify the methods used to generate the systematic search, and critically appraise and synthesize the results of the studies reviewed. Systematic reviews based on meta-analyses of the literature are especially useful as they are based on an empirical method for statistically integrating the results of many studies involving large numbers of patients by emphasizing the effect sizes of the group differences and the factors that moderate these differences. By carefully specifying the criteria for their searches, researchers make explicit any biases that may be inherent in their reviews. For example, Belanger and colleagues (Belanger, Curtiss, Demery, Lebowitz, & Vanderploeg, 2005) performed a meta-analysis of the literature on the neuropsychological effects of mild traumatic brain injury based on 39 studies involving 1463 patients and 1191 controls. They report an overall moderate effect size ( $d = .54$ ), with the findings moderated by such factors as time since injury, cognitive domain, sampling methods, and patient characteristics. In their review they report that the acute effects of injury were most apparent for delayed memory and fluency ( $d = 1.03$  and  $.89$ , respectively). These findings are derived across a large number of patients and different sampling approaches, and are likely to be quite robust and generalizable, making the information useful in EBCNP practice.

Neuropsychological research is published in an ever-growing list of neuropsychology and medical journals. At some point there may be a journal or service that culls this literature for outcomes-related findings and makes it available to clinical neuropsychologists who have specific patient-related questions.

Until then, clinicians should learn to develop search strategies for identifying the “best available research” to assist them in evaluating individual patients, and apply this information in a routine manner. If EBCNP truly becomes the standard of care, practitioners themselves will eventually need to become outcomes researchers, monitoring and tracking the outcomes of their own practices. This will require clinicians to work prospectively, set up automated and continuous data collection systems that provide access to archival data in a comfortable and easy to use manner, ask answerable questions of their patient data, and be willing to make their findings available to others (Dodrill, 1999). Such practice-based databases of individual clinicians will, of course, be subject to the same types of sampling and selection biases that researchers face, limiting generalizability but providing publicly verifiable information about the clinician’s practice that can be shared with consumers, patients and payers. These individual practice-based monitoring systems also offer the opportunity for clinicians to pool their data sets to answer broader questions, provide within study replication, and identify between-site variables that affect outcomes (Costa, 1983).

### **ADVOCATING FOR EBCNP: THE MARRIAGE OF SCIENCE AND PRACTICE**

Advocacy for neuropsychology is not necessarily limited to efforts directed at changing policies or laws. As noted by Howe and colleagues (Howe, Sweet, & Bauer, in press), advocacy can also involve efforts to change the views and practices of individuals and organizations. The proposition advanced here is that engaging in EBCNP practice and research is to “advocate” for neuropsychology. If our goal is to advance the science and practice of neuropsychology in a manner that is publicly verifiable by our consumers, our patients, and payers, then our research and our clinical work products should be the “best” arguments for why neuropsychological services demonstrably provide “added value” for individual patient care. Within the proposed EBCNP framework, we are all potentially advocates by virtue of our day-to-day efforts.

Outcomes-based clinical neuropsychological research and practice are only beginning to emerge, and a definition for EBCNP has yet to be clearly formulated. What I have offered here are only suggestions for what I see as potentially relevant steps that can simultaneously advance evidence-based practice and research, while providing a basis for advocating the value of the profession. I believe we need to think of outcomes in terms of discrete events and findings, and we should be prepared to make explicit the criteria for our decisions regarding these outcomes. Making EBCNP the standard of care in our field and a financially viable service will require the combined “advocacy” efforts of both outcomes researchers and clinical practitioners. If researchers would supplement their data analyses with simple contingency table analyses that provide additional information on the validity (sensitivity and specificity) and potential impact and value of cognitive assessment (odds ratios, likelihood ratios, and changes in pre- to post-test odds), it will clearly advance both the science and potentially the practice of neuropsychology. However, this information will only have an impact on EBCNP if it used and applied by practitioners in a way that payers can determine the value of neuropsychological services.

## REFERENCES

- APA Presidential Task Force on Evidence-Based Practice (2006). Evidence-based practice in psychology. *American Psychologist*, 61(4), 271–285.
- Belanger, H. G., Curtiss, G., Demery, J. A., Lebowitz, B. K., & Vanderploeg, R. D. (2005). Factors moderating neuropsychological outcomes following mild traumatic brain injury: A meta-analysis. *Journal of the International Neuropsychological Society*, 11, 215–227.
- Bieliauskas, L. A., Fastenau, P. S., Lacy, M. A., & Roper, B. L. (1997). Use of the odds ratio to translate neuropsychological test scores into real-world outcomes: From statistical significance to clinical significance. *Journal of Clinical and Experimental Neuropsychology*, 19(6), 889–896.
- Busch, R. M., Chelune, G. J., & Suchy, Y. (2006). Using norms in neuropsychological assessment of the elderly. In D. K. Attix, & K. A. Welsh-Bohmer (Eds.), *Geriatric neuropsychology: Assessment and intervention* (pp. 133–157). New York: Guilford Press.
- Chelune, G. J. (2002a). Assessing reliable neuropsychological change. In R. Franklin (Ed.), *Prediction in forensic and neuropsychology: New approaches to psychometrically sound assessment* (pp. 123–148). Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Chelune, G. J., (2002b). Making neuropsychological outcomes research consumer friendly: A commentary on Keith et al., (2002). *Neuropsychology*, 16(3), 422–425.
- Chelune, G. J., & Najm, I. (2001). Risk factors associated with postsurgical decrements in memory. In H. O. Lüders, & Y. Comair (Eds.), *Epilepsy surgery* 2nd ed., pp. 497–504. New York, Lippincott: Williams & Wilkins.
- Chelune, G. J., Naugle, R. I., Luders, H., & Awad, I. A. (1991). Prediction of cognitive change as a function of preoperative ability status among temporal lobectomy patients seen at 6-month follow-up. *Neurology*, 41(3), 399–404.
- Chelune, G. J., & Stone, L. (2005). Risk of processing speed deficits among patients with relapsing and remitting and secondary progressive multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, 11, 52.
- Chelune, G. J., Stott, H., & Pinkston, J. (2008). Multiple sclerosis. In J. E. Morgan, & J. H. Ricker (Eds.), *Textbook of clinical neuropsychology* (pp. 599–615). New York: Taylor & Francis.
- Collie, A., Darby, D. G., Falletti, M. G., Silbert, B. S., & Maruff, P. (2002). Determining the extent of cognitive change after coronary surgery: A review of statistical procedures. *Annals of Thoracic Surgery*, 73(6), 2005–2011.
- Costa, L. (1983). Clinical neuropsychology: A discipline in evolution. *Journal of Clinical Neuropsychology*, 5(1), 1–11.
- Dodrill, C. B. (1999). Myths of neuropsychology: Further considerations. *The Clinical Neuropsychologist*, 13(4), 562–572.
- Evidence-Based Medicine Working Group (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, 268(17), 2420–2425.
- Hawkins, K. A., & Tulsy, D. S. (2003). WAIS-III WMS-III discrepancy analysis: Six-factor model index discrepancy base rates, implications, and a preliminary consideration of utility. In D. S. Tulsy, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 211–272). New York: Academic Press.
- Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Demographic effects and use of demographically corrected norms with the WAIS-III and WMS-III. In D. S. Tulsy, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III*. New York: Academic Press.

- Horwitz, R. I. (1996). The dark side of evidence-based medicine. *Cleveland Clinic Journal of Medicine*, 63(6), 320–323.
- Howe, L., Sweet, J., & Bauer, R. (in press). Advocacy 101: A step beyond just complaining: How the individual practitioner can become involved and make a difference. *The Clinical Neuropsychologist*.
- Hunsley, J. (2007). Addressing key challenges in evidence-based practice in psychology. *Professional Psychology: Research and Practice*, 38, 113–121.
- Ivnik, R. J., Smith, G. E., Cerhan, J. H., Boeve, B. F., Tangalos, E. G., & Petersen, R. C. (2001). Understanding the diagnostic capabilities of cognitive tests. *The Clinical Neuropsychologist*, 15(1), 114–124.
- Johnson, L. A. (1997). Outcomes management a decade out: An interview with Paul Ellwood. *Group Practice Journal*, 46, 12–15.
- Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, 63(3), 146–159.
- Levy, J. A., & Chelune, G. J. (2007). Cognitive-behavioral profiles of neurodegenerative dementias: Beyond Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*, 20(4), 227–238.
- Prigatano, G. P., & Morrone-Stupinsky, J. (in press). Advancing the profession of clinical neuropsychology with appropriate outcome studies and demonstrated clinical skills. *The Clinical Neuropsychologist*.
- Rao, S. M., Leo, G. J., Bernardin, L., & Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis: Frequency, patterns, and prediction. *Neurology*, 41, 685–691.
- Rascovsky, K., Salmon, D. P., Hansen, L. A., Thal, L. J., & Galasko, D. (2007). Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology*, 21(1), 20–30.
- Ritchie, L. J., Frerichs, R. J., & Tuokko, H. (2007). Effective normative samples for the detection of cognitive impairment in older adults. *The Clinical Neuropsychologist*, 21, 863–874.
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-based medicine: How to practice and teach EBM* 2nd ed.,. New York: Churchill Livingstone.
- Smith, G. E. (2002). What is the outcome we seek? A commentary on Keith et al. (2002). *Neuropsychology*, 16(3), 432–433.
- Strauss, E., Sherman, M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests* 3rd ed.,. New York: Oxford University Press.
- The Psychological Corporation. (2002). *WAIS-III-WMS-III-WIAT-II scoring assistant (rev.)*. San Antonio, TX: The Psychological Corporation.
- Tulsky, D. S., Chiaravalloti, N. D., Palmer, B. W., & Chelune, G. J. (2003). The Wechsler Memory Scale: A new perspective. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* 3rd ed., (pp. 93–137). New York: Academic Press.
- Tulsky, D. S., Rolfhus, E. L., & Zhu, J. (2000). Two-tailed versus one-tailed base rates of discrepancy scores in the WAIS-III. *The Clinical Neuropsychologist*, 14, 451–460.
- Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals. *American Psychologist*, 54, 594–604.
- Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: Formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology*, 16(7), 653–667.