

American Society Clinical Neurophysiology

Guideline 9A: Guidelines on Evoked Potentials¹

Introduction and Standards

INTRODUCTION TO 2006 GUIDELINES

Since the last Guidelines were published, there has been a considerable change in the way in which evoked potentials are used clinically. The primary clinical application of evoked potentials has shifted from the diagnosis of neurological disorders to monitoring during neurological, orthopedic and other types of surgeries to prevent neurological injury. Nonetheless, the use of diagnostic evoked potentials continues and the need for guidelines is ever present.

Many parts of the previous Guidelines on Evoked Potentials have been preserved with this latest update. Readers will notice significant changes in “Recommended Standards for the Clinical Practice of Evoked Potentials.” The last decade has seen an evolution in the training and certification examinations for interpreters. Due to changing requirements of various certification examinations, references have been made to websites of the respective certification boards for more information.

It is hoped that these Guidelines continue to provide useful information to newcomers to the field, established neurophysiologists, technologists and others interested in evoked potentials. Another set of Guidelines will address intraoperative use of evoked potentials.

INTRODUCTION TO 1991 GUIDELINES

It was predicted in the “Introduction to the First Guidelines” that “because of the dynamic nature of this field, it is likely that revisions of the recommendations will be required in the future.” This second version of the Guidelines gives truth to that prediction. In the interval between these guidelines, many thousands of research and clinical diagnostic studies have been performed, adding greatly to our knowledge and necessitating the many changes in the current Guidelines. There is little doubt, however, that there will be further revisions in the future.

In contrast to the First Guidelines, these current Guidelines, where appropriate, address some of the better-documented sources of generation of various evoked potential components.

The sections on visual and somatosensory evoked potentials have undergone the greatest modification with relatively fewer changes in the section on auditory evoked potentials. Several components of the introductory section were left largely unchanged, although some additions were made.

These are guidelines and as such are not intended to stifle thought, investigation, or inventiveness. They are the product of negotiation between many committee members. In some areas, there was less than total agreement. Although one should feel free to deviate from these Guidelines when appropriate, it should be emphasized that workers in the field should know the content of these guidelines and understand the logic behind them. Further, the effects of deviations from these guidelines should be well understood.

¹ This topic was previously published as Guideline 9.

In the initial meeting of the Guidelines committee for the 1991 revision, there was discussion regarding the desirability of adding sections on evoked potential studies in neonates and patients requiring intensive care. We also considered making recommendations for cognitive evoked potentials, somatosensory studies isolating the functioning of single dorsal roots (dermatomal, segmental spinal evoked potentials), and motor evoked potentials. Although data from such studies may in theory be of clinical value, it was felt that there was an insufficient number of replicated reports to warrant their inclusion at this time. Since the 1984 Guidelines, there has been a separate report on intraoperative monitoring.

There have been slight modifications in nomenclature of electrode location to be consistent with the recently distributed Guidelines on this topic.

Not only newcomers to the field but also established workers have used the initial Guidelines to help in acquiring the knowledge necessary for performing and interpreting clinical evoked potential studies. Physicians preparing for the examinations of the American Board of Clinical Neurophysiology and technologists preparing for examinations of the American Board of Registration of Electroencephalographic and Evoked Potential Technologists have found the Guidelines a useful study guide. The committee hopes the current version will be equally helpful.

INTRODUCTION TO 1984 GUIDELINES

Electrical responses of the nervous system that are time-locked to a sensory stimulus, an electrical excitation, a movement, or other identifiable events are referred to as “event-related” potentials. In the context of the present recommendations, the term “evoked potential” is used more specifically to designate the responses of the sensory pathways to sensory or electrical stimuli. Because of the small amplitude of evoked potentials recorded by noninvasive methods in humans, computer summation or averaging generally is necessary to resolve them from background “noise.”

Optimal recording conditions, criteria for abnormality, and the clinical usefulness of some evoked potentials have been sufficiently well-established to warrant the formulation of standards for their performance and interpretation. By contrast, our knowledge of other evoked potentials is undergoing such rapid evolution that it would be premature to suggest any precise standards at this time. Thus, the standards recommended by the present Guidelines are limited to the following areas: (1) clinical practice of evoked potentials; (2) normative studies of evoked potentials, statistical analysis of results, and criteria for clinically significant abnormality; (3) visual system evoked potentials; (4) short-latency auditory evoked potentials; and (5) short-latency somatosensory evoked potentials.

These standards reflect the “state of the art” in clinical evoked potential studies. Because of the dynamic nature of this field, it is likely that revisions of the recommendations will be required at some future time.

No attempt was made to discuss in any detail the generator sources of the individual response components nor to delineate specific areas of clinical application for each type of evoked potential. Moreover, no recommendations were formulated on evoked potential recording in newborns and infants as well as in the intensive care unit and the operating room.

RECOMMENDED STANDARDS FOR THE CLINICAL PRACTICE OF

EVOKED POTENTIALS

I. Introduction

Sensory evoked potentials are recorded in different clinical contexts. They may be used to assess peripheral sensory function, to evaluate the functional integrity of sensory projection pathways in the central nervous system, or both. Collaboration among the different disciplines that utilize evoked potential measures should be fostered as much as possible. The following Guidelines recommend standards for recording and interpreting evoked potentials primarily aimed at evaluating the function of sensory pathways in the central nervous system with the intent of providing clinically relevant information.

II. Qualifications for Practice

Recommended Qualifications for Interpreters of Clinical Evoked Potential Studies

The training of a qualified interpreter of clinical evoked potential studies is accomplished in a post-residency fellowship in clinical neurophysiology. Such fellowships are usually a minimum of 12 months' duration, and general skills in clinical neurophysiology, in addition to those required for clinical evoked potential interpretation, should be provided. A listing of such fellowships is currently available from the American Academy of Neurology (www.aan.com). Many of these fellowships may be accredited by the Accreditation Council for Graduate Medical Education (ACGME).

Fellowship training in evoked potentials should be designed to provide thorough understanding and direct familiarity with all aspects of evoked potential data acquisition, processing, and interpretation, including the influences on evoked potentials of stimulus and recording parameters; knowledge of anatomical structures, neurophysiological events and other factors underlying the generation of evoked potentials; the clinical significance and pathological correlates of dysfunction of neural pathways demonstrated by evoked potential alterations; and an understanding of the relevant statistics.

Becoming credentialed by a national examining organization, assessing adequacy of knowledge of evoked potentials, is the only objective method of demonstrating competency in interpretation of clinical evoked potential studies. The American Board of Clinical Neurophysiology (ABCN) offers a two-part examination (a written followed by an oral examination), a part of which is devoted to clinical evoked potentials. Details regarding eligibility for acceptance to this examination can be found at www.abcn.org.

The American Board of Psychiatry and Neurology (ABPN) has created a written examination in the "Subspecialty of Clinical Neurophysiology," first offered to eligible candidates in March 1992. This written examination focuses on EEG and EMG topics with lesser coverage of evoked potentials, long-term monitoring, sleep, and others. Eligibility information for this examination can be found at www.abpn.com.

Recommended Qualifications for Evoked Potential Technologists

The qualifications for an evoked potential technologist derive directly from the tasks requisite to conducting standard brainstem auditory evoked potential (BAEP), pattern visual evoked

potential (PVEP), and somatosensory evoked potential (SSEP) studies. This includes the ability to explain the nature and purpose of the study to the patient, to allay fears related to the study, to obtain medical history necessary to establish appropriate clinical correlations, to obtain high-quality data, to reduce the data, to annotate data sheets with all pertinent and appropriate information, and to troubleshoot and maintain equipment. The background of such an individual will in most instances require the completion of a formal educational program in clinical neurophysiology, as well as direct training and experience totaling not less than 12 months and usually more. Such skills constitute the qualifications of an entry-level evoked potential technologist.

More advanced qualifications are needed to perform complex studies such as those obtained in neonatal and intensive care units and during surgery. Such studies demand skill, patience, and ingenuity. They require an evoked potential technologist with advanced training and experience, initiative, and capacity for independent judgment. Such a technologist would also provide training to junior staff, technologist trainees, and other medical personnel.

Details regarding training of technologists can be obtained from the Association of Electroneurodiagnostic Technologists (ASET), www.aset.org. Information about credentialing of technologists is available from the American Board of Electroencephalographic and Evoked Potential Technologists (ABRET), www.abret.org.

III. Standards for Clinical Evoked Potential Equipment

The standards given in this section represent the minimum necessary to obtain good-quality clinical evoked potentials. In many respects, analog and digital technology have advanced well beyond the minimum, and these advances are often incorporated in modern commercial equipment. This trend should be welcomed but does not imply obsolescence of older or noncommercial systems provided they comply with these Guidelines.

Amplifier

Input signals with peak-to-peak amplitudes of 5 μV to 50 μV should be amplified equal to the full range of the analog-to-digital (A-D) converter. Gain should be adjustable in steps of not more than 2.5 to 1. The differential input impedance of the amplifier must be at least 100 megOhm. The common mode rejection should be at least 80 dB (10,000:1) at the highest sensitivity of the amplifier when the common mode signal is applied between both inputs and neutral. With filters wide open, the amplifier bandpass measured at the -3 dB points must be at least 0.1-5,000 Hz. The rolloff slopes of the filters must be specified.

The noise level of the amplifier must not exceed 2 μV rms with the inputs connected to neutral and with a bandpass of 0.1-5,000 Hz. The amplifier must meet all specifications in the presence of a sustained 300-mV offset applied differentially between the input terminals 1 and 2 or commonly to inputs 1 and 2 (with respect to the patient “ground” or neutral lead).

Averager

Time (horizontal) resolution of the system should be 20 μs /data point or less. An amplitude resolution of eight bits at the A—D converter is adequate for many applications, but 12-bit A—D converters are preferred because they are usable across a broader dynamic range. At least 500

addresses of memory should be available for each channel. The system should allow averaging at least 4,000 trials. The onset of the averaging sweep should be easily and accurately synchronized to stimulus production. A mechanism whereby artifact-contaminated trials can be simply and quickly excluded from the averaging process is essential. This is most commonly achieved by rejecting those trials that exceed the limits of the A-D converter or some adjustable percentage thereof. At least four channels are the minimum necessary for PREP and SSEP studies. Two channels are the minimum for BAEP studies.

Display and Writeout

A cathode-ray-tube (CRT) or equivalent display must be available to show the average waveforms and the ongoing unaveraged EEG. Such a display must have easily understandable voltage and time scales. A permanent hard copy of the evoked potentials must be available. Postacquisition data manipulations potentially affecting the reliability of the study (smoothing, additive transfers, etc.) should be routinely displayed on the hard copy as part of the data report.

Optional Features

Additional features may be useful in certain situations but are not considered essential. These include additional channels (as will be noted in the individual sections, eight channels may often be used to advantage), lower amplifier noise, DC input capability, electronic data transfer and storage, phase-shift-free digital filtering, data cursors, and continuous trend analysis. Electronic storage should always include the original data, and any electronic data manipulation performed after acquisition should automatically be detailed in the writeout. Data should be electronically stored in their raw form prior to postacquisition alterations. Incorporation of such advanced features should not be at the cost of interpretable data outputs or user accessibility.

IV. Standards for Clinical Evoked Potential Recording

Electrical Safety

In recording evoked potentials, measures must be taken to assure the patient's safety. The grounding and the chassis leakage current of all instruments connected to the patient or located in the same room as the patient must be periodically tested. Chassis leakage should be less than 300 μ A rms with ground open. Special caution must be exercised when recording evoked potentials with portable equipment at the patient's bedside, in the intensive care unit, and in the operating room. A direct connection from patient "ground" to chassis or power line ground should not be present. An isolation amplifier or a solid-state current limiter must be incorporated to actively prevent such connections.

Equipment should be designed to prevent inadvertent shock during power-on, power-off, and failures. For the operating room or intensive care unit, current limiting must be present in every patient lead. In these areas the maximum leakage current through patient leads should be 10 μ A

rms to ground with 120 VAC applied. The maximum leakage current should be 50 uA rms at the patient end of the cable when the power line ground is disconnected.

Electrical stimulators for SSEP and TES should have all outputs isolated from ground, and should meet the same standards as the recording connections when they are not energized. The selected output current should not be exceeded for any reason including system failure.

Filtering

Appropriate analog filtering is useful in rejecting frequencies of minimal or no clinical interest. As a result, evoked potentials are improved with less artifact and the desired level of signal-to-noise ratio improvement can be obtained with fewer stimuli. Excessive analog filtering will alter the latency, amplitude, and morphology of evoked potential components. Excessive low-frequency filtering will cause a peak or valley to appear earlier, leading to possible falsely normal results. Excessive high-frequency filtering will delay (prolong) a peak or valley leading to a possible falsely abnormal study. A P100 peak of PREP origin will be delayed by 8-10 ms when recorded at a high-frequency filter setting of 30 Hz as compared to a filter setting of 250 Hz. The exact amount of time shift cannot be predicted because the time shift is a function of both the filter and the frequency of the evoked potential component. A broad slower frequency wave will be shifted more than a sharp higher frequency potential. In general, the ratio of the high-to-low frequency filter settings should be at least 100:1 to be confident that filter-induced phase shifts are minimized.

Equipment made by different manufacturers may perform differently with identical filter settings. This is because the approximate break points of the filter (in Hz) do not describe the rolloff slopes (dB/ octave). Differences in rolloff slopes can cause differences in amplitude, morphology, and most importantly latency.

Digital filtering and smoothing algorithms will affect amplitude and morphology but usually do not affect latency. Note that the term “digital filtering” is a general term and does not actually describe the construction of the filter. A digital filter could, for instance, emulate an analog RC filter.

Use of the 60- (or 50-) Hz notch filter is discouraged because it may ring when activated by a sudden transient such as an auditory or somatosensory stimulus artifact. As a result, a burst of 60-Hz activity of decreasing amplitude (interpeak interval of 16.66 ms or 20.0 ms for a 50-Hz filter) will contaminate the response. Under some circumstances this may falsely appear as a neural response when none is present or shift the latency of a true response.

The filter settings suggested in various portions of these guidelines were arrived at empirically and are known to have minimal clinically significant effect on obligate components (Desmedt,, 1977).

Polarity Convention

The majority of commercially available evoked potential systems now employ an amplifier polarity convention with input terminals designated by “+” and “-“. A positive event occurring in the lead connected to the + terminal (or a negative event occurring in the lead connected to the - terminal) will result in an upward deflection. Downward deflections occur when a positive or negative event is applied to a terminal of opposite polarity. In a majority of current literature, SSEP and PREP negative events are displayed upward. The opposite is true of BAEP

components; positive events are usually displayed upward. The choice of display is ultimately at the discretion of the individual; however, it is imperative that the polarity convention be understood and clearly labeled.

Calibration

The recording system must have the capacity to be calibrated periodically as needed to insure the integrity of analog and digital components. Generally, this is achieved by injecting into the input jacks of each channel rectangular pulses of appropriate amplitude, usually 0.5-100 μV , time-locked to the onset of the sweep. Calibration pulses should not be injected after a stage of amplification. The calibration pulses must be amplified and averaged, and their amplitude measured in conditions identical to those to be employed for the recording of the evoked potential under study.

Replication

To replicate is to obtain two or more temporally independent averages. Replication of the response is imperative to demonstrate that clinical evoked responses are consistently repeatable and therefore are of neural and not artifactual origin. Nonsuppressed artifact can mimic biologic responses in a single average but usually not in subsequent replications. This can produce false components leading to incorrect assumptions regarding the normality or abnormality of a study.

Replication is demonstrated by the consistency of latency and amplitude measures of evoked potential components recorded in successive averages. Latency replication within 1.0% of the total sweep time and amplitude replication within 15% of the peak-to-peak amplitude can usually be achieved. Poor replication may be caused by (1) unusually low (but not necessarily abnormal) amplitude responses, (2) excessive artifact (frequently myogenic), and (3) insufficient number of responses in the average. Decreasing electrode impedance will often reduce noise by improving common mode rejection. Increasing the number of stimuli (responses) will reduce all sources of noise. Relaxation and sleep will reduce various biologic artifacts. With low-amplitude responses, increasing the number of stimuli will improve the signal-to-noise ratio, but there are practical limits. In BAEP and SSEP studies, increasing the intensity of the stimulus may increase the amplitude of the evoked potential components. Some clinical studies will have poor reproducibility not amenable to any clinically practical technique to correct the problem. Poor reproducibility per se does not imply abnormality.

Poorly replicating data should not be used in calculating normal values for evoked potential data. In clinical studies, more than one replication may be required to obtain consistent data. When poor replication in clinical studies cannot be controlled, it should be taken into account by extending the usual upper limits of normal. This is particularly true in the evaluation of side-to-side asymmetries of latency or amplitude.

V. Standards for Documentation and Interpretation of Results

Documentations

All evoked potential record should bear the following information:

1. The patient's name, identifying number, age, and gender.
2. The date of the examination and the procedure number.
3. The technologist's name or initials.
4. The derivation recorded in each channel in the form of abbreviated accepted designations of the electrode locations connected to the input terminals 1 and 2 of the amplifier, in that order.
5. The type, polarity, field size, check size, full or partial field, intensity, and rate of presentation of the stimulus, and the side of stimulation (when relevant).
6. Other information relevant to test results such as masking of the nonstimulated ear, state of retinal adaptations, etc.
7. The number of individual trials averaged.
8. The time calibration in the form of a horizontal line, corresponding to the epoch averaged, with subdivisions appropriate to the temporal dimensions of the evoked potential recorded. Whenever a prestimulus baseline or poststimulus delay is used, it should be clearly displayed.
9. The voltage calibration in the form of a vertical line indicating the amplitude of deflection produced in terms of sensitivity (voltage/linear distance).
10. Indication of the polarity convention followed by the user. This may be in the form of a plus or minus sign near the upper and lower ends of the voltage calibration line, respectively. These signs should indicate the polarity of the electrode connected to the input terminal 1 of the amplifier, relative to that connected to the input terminal 2, during an upward and downward deflection to the record, respectively. Alternatively, the polarity convention may be indicated above the derivation(s). An arrow above or below the first named electrode locations, pointing to a plus or minus sign indicates the direction of deflection produced by a given polarity event.
11. Modern EP equipment will leave marks indicating the points at which "measurements" were taken. These marks can be used to calculate latencies and amplitudes of various peaks and valleys of interest.

Interpretation

See individual sections for each modality, the section on "Recommended Standards for Normative Studies of Evoked Potentials, Statistical Analysis or Results, and Criteria for Clinically Significant Abnormality." Also see the separate "Guidelines for the Writing of Evoked Potential Interpretations."

REFERENCES

1. American National Standard Safe Current Limits of Electromedical Apparatus. ES-1, 1993. Arlington, VA: Association for the Advancement of Medical Instrumentation, NFPA-99, 2002.
2. Desmedt JE. Some observations on the methodology of cerebral evoked potentials in man. In: Desmedt JE, ed. *Attention, voluntary contraction and event-related cerebral potentials*. Basel: Karger. 1977:12—29. (Progress in clinical neurophysiology~ vol. 1.)
3. Donchin E, Callaway E, Cooper R, Desmedt JE, Goff WR, Hillyard SA, Sutton S. Publication criteria for studies of evoked potentials (EP) in man. Report of a committee. In:

Attention, voluntary contraction and event-related cerebral potentials. Basel: Karger. 1977:12—
29. (Progress in clinical neurophysiology: vol. 1.)

RECOMMENDED STANDARDS FOR NORMATIVE STUDIES OF EVOKED POTENTIALS, STATISTICAL ANALYSIS OF RESULTS, AND CRITERIA FOR CLINICALLY SIGNIFICANT ABNORMALITIES

I. Introduction

The successful clinical application of evoked potentials depends, in large measure, on the availability of carefully collected and skillfully analyzed normative data. When differences exist in normative data values between different laboratories, there is a limited number of causes for such differences. These are (1) subject characteristic (age, gender, non-random sampling), (2) stimulation parameters, (3) recording parameters, and (4) data reduction algorithms (“peak picking” rules). In organizing new laboratories, it is acceptable to utilize as a reference the normative data published by another center provided the following requirements are satisfied:

1. Stimulus, recording, and other conditions are used that are *identical* to those of the reference laboratory as determined by appropriate calibration equipment and methods. Further, there must be detailed familiarity with the peak identification rules used by the reference laboratory.

2. At least 20 normal subjects are studied, spanning the age range of the patients to be examined in the particular laboratory, and it is determined that a specified proportion (such as 95% or 99%) of this subset of normal values falls within the limits derived from the subset studied in the reference laboratory. This exercise is imperative to verify the accuracy with which the conditions described in the above paragraphs are met.

II. Selection of Subjects

Appropriate selection of subjects for normative studies of evoked potentials is of critical importance (Blumhardt et al., 1982). All such subjects should be neurologically normal and have no family history of inherited neurologic disorders. For *visual evoked potentials*, (VEPs), subjects should have no history of a disorder of vision for their age (other than refractive error) or ophthalmic migraine. Ophthalmologic examination should include testing of visual acuity, refractive error (not to exceed -5 diopters for myopes), visual fields, and color sensitivity as well as fundoscopic examination. Subjects selected for normative studies of *auditory evoked potentials* (AEPs) should have no personal history or history of familial disease of the ear and the nervous system and should be otologically, audiological, and neurologically normal for age. Audiometric examinations should include standard pure tone audiometry with determination of air-conduction and bone-conduction hearing threshold levels and tests of speech discrimination, acoustic impedance, and crossed acoustic reflexes. Persons contributing norms for *SSEPs* should have no history of familial neurologic disease or personal history of neurologic disease and must be neurologically normal for age. Any personal history of trauma, bone fractures, and alterations of sensation must be carefully evaluated. It is recommended that for studies dealing with cross-hemispheric comparisons, handedness, eye dominance, or ear dominance be specified, at least as perceived by the subject. Thorough inquiry should be made into the use of drugs by prospective normal subjects, including narcotics, stimulants, and neurotropic drugs. Individuals taking such medications should be excluded from normative studies.

III. Number, Age, and Sex of Subjects

Each control group should contain an equal number of age-matched individuals of the two sexes. As a general rule, age-specific norms should be obtained by week in the perinatal period, by month in infants, and by decade in children and adults. It is desirable that each subgroup consist of a minimum of 20 subjects. *Regression analysis* (Edwards, 1976; Celesia and Daly, 1977) of data collected on individuals evenly spread over a given age range permits more parsimonious use of subjects.

IV. Paired Observations

Measures of responses of right and left eyes, ears, or peripheral nerves of the same individuals should not be treated as independent observations, i.e., lumped together (Ederer, 1973). In general, a high positive correlation exists between such paired evoked potential observations in normal subjects.

V. Description of Results and Criteria for Clinically Significant Abnormality

The first step to be taken in the statistical analysis of evoked potential measures obtained in a normative study is to examine the *shape of the distribution of the observations* in the particular sample examined. Should this distribution be or approximate a normal bell-shaped (gaussian) curve, it is appropriate to describe the characteristics of the sample by computing standard measures of central tendency and dispersion, such as *the mean and the standard deviation*. It should be emphasized that these statistics assume normal distribution of values and have little validity unless this assumption is met. Unfortunately, the distribution of evoked potential measures obtained from the small samples generally studied, frequently exhibits deviations from normality including significant skewness (deviation of the curve from symmetry), kurtosis (relative peakedness or flatness of the curve), or both. Ratios (i.e., amplitude ratios) even when the numerator and denominator are both normally distributed, are usually markedly skewed even when large samples are utilized. In these instances, it is recommended that the *observed data be transformed* (Cohen and Cohen, 1973; Kirk, 1982; Oken, 1990) with the intent of obtaining a normal distribution or a distribution more closely approximating normalcy, before computing mean and standard deviation statistics. Taking the logarithm, the square root, reciprocal, or various other transformations of the values not conforming to normal distribution may be attempted. Some data sets cannot be transformed to a normal distribution. The original data and various transforms should be assessed for deviations of the distributions from normal. Various statistics may be used, but for small sample sizes, the Shapiro-Wilk's goodness of fit (W) is sensitive to a wide range of deviations from normality (Shapiro et al., 1968). Once data are normalized, the mean and upper (or lower) limits are calculated from the normalized data and then transformed back to the original units (Judd and McClelland, 1989).

Clinical diagnosis frequently requires that measures obtained in individual patients be compared to population norms with the intent of determining whether they are "normal" or "abnormal." Because a small sample from the normal population represents a very limited part of

the entire set of relevant observations, it cannot be identified with the population. Thus, statements that clinically observed values, such as the latency or amplitude of a given wave exceeding 2, 2.5, or 3 standard deviations of the mean of a normal control group, are “*abnormal*” are acceptable provided the following requirements are satisfied: (1) it is clearly specified that the values in question are regarded as abnormal compared to “*a control sample from the normal population.*” and (2) no precise probability is implied in predicting where these values are located relative to the normal population. Qualifying these same clinical observations as abnormal compared to “a normal control population,” or to “the normal population” is statistically erroneous. Predictive statements giving the interpreting clinical neurophysiologist as well as the referring physician a quantitative appreciation of the statistical significance of the abnormality of a given clinical observation are made possible by the use of “*tolerance limits*” (TL) (Stockard et al., 1978). For a normally distributed control sample of a given size and of known mean and standard deviation tolerance limits that include a *given proportion* of the normal population with a *given level of confidence* can be computed with the aid of appropriate tables of tolerance factors. For practical clinical purposes, generally, latencies and interpeak intervals are regarded as abnormal when they are excessively long, and amplitudes are viewed as aberrant when they are excessively small.’ Thus, the use of *one-tailed* tolerance limits is recommended for most evoked potential studies. Many laboratories have adopted tolerance limits of 99% and 95%, meaning that at least 99% (a given proportion), of the normal population is less than the upper tolerance limit with a confidence of 95% (a given level of confidence). This is the equivalent of the mean plus 3.064 standard deviations if the sample size is 30. Given the same percentage of the population and confidence level, when smaller sample sizes are used, the equivalent number of standard deviations increases, i.e., 3.295 with a sample size of 20 and 3.981 with a sample size of 10 (Owen, 1970).

For any given limit (upper or lower) of normality, there is a certain probability of falsely interpreting normal values as abnormal, i.e., of false-positive results and of conversely qualifying abnormal values as normal, i.e., of “false-negative” findings. Adopting more stringent normal limits (lowering the upper limits of absolute latency, interpeak latency (IPL), side-to-side comparison, etc.) has the advantage of decreasing the number of false-negative results (test more sensitive) but carries the penalty of increasing the proportion of false-positive (test less specific) decisions. The opposite is true when more liberal limits of normality are adopted. Setting normal limits is a decision to be made by each individual laboratory with full understanding of its statistical implications.

Further confounding these issues is the usual practice of applying multiple criteria of abnormality. It is common, particularly in BAEP and SSEP studies, to consider several IPLs and side-to-side asymmetries of these IPLs in determining the normalcy of a given clinical study. If only a single criterion is being considered, a 95% or 98% TL maybe appropriate, but when multiple criteria are being applied, a 99% TL helps reduce the errors inherent in the overapplication of univariate statistics to multi-variate problems (Oken, 1990).

Some of the statistical analyses alluded to, including the transformation of values not conforming to a normal distribution and the use of techniques such as regression analysis, among others, require extensive computational capabilities and advanced statistical skills. It is suggested that clinical laboratories undertaking the collection of normative data seek experienced advice on the design of their studies and support in the analysis of their results.

Ultimately, the adequacy of any given normal limit in discriminating between normal and diseased individuals must be supported by appropriate clinical and/or clinico-pathologic

correlations.

'In electroretinography "supranormal" responses may be clinically relevant. The same may be true of flash VEPs in some neurological conditions.

VI. References

- Celesia GG, Daly RF. Effects of aging on visual evoked responses. *Arch Neurol* 1977;34:403—7.
- Cohen J, Cohen P. *Applied multiple regression/correlation analysis for the behavioral sciences*. New York: Wiley, 1973.
- Donchin E, Callaway E, Cooper R, et al. Publication criteria for studies of evoked potentials (EP) in man. Report of a Committee. In: Desmedt JE, ed. *Attention, voluntary contraction and event-related potentials*. Basel: Karger, 1977:1-11. (Progress in clinical neurophysiology: vol 1.)
- Ederer F. Shall we count the number of eyes or numbers of subjects? *Arch Ophthalmol* 1973;89:1-2.
- Edwards AL. *An introduction to linear regression and correlation*. San Francisco: Freeman, 1976.
- Judd CM, McClelland GH. *Data analysis: a model-comparison approach*. Orlando, FL: Harcourt, Brace, Jovanovich, 1989.
- Kirk RE. *Experimental design: procedures of the behavioral sciences*. New York: Brooks/Cole, 1982.
- Oken BS. Statistics for evoked potentials. In: Chiappa KH, ed. *Evoked potentials in clinical medicine*, 2nd ed. New York: Raven Press, 1990:593-608.
- Owen DB. *Handbook of statistical tables*. Reading, MA: Addison-Wesley, 1970.
- Shapiro SS, Wilk MB, Chen HJ. A comparative study of various tests for normality. *J Am Stat Assoc* 1968;63:1343-72.
- Stockard JJ, Stockard JE, Scarborough FW. Nonpathologic factors influencing brainstem auditory evoked potentials. *Am JEEG Technol* 1978;18:177—209.