



# Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience

## Use of Databases in QEEG Evaluation

Jack Johnstone PhD <sup>a</sup> & Jay Gunkelman QEEG-D <sup>a</sup>

<sup>a</sup> Q-Metrx, Inc. , Burbank, CA

Published online: 07 Sep 2008.

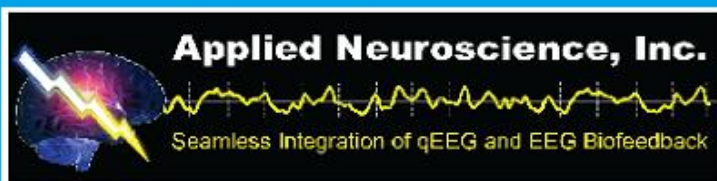
**To cite this article:** Jack Johnstone PhD & Jay Gunkelman QEEG-D (2003) Use of Databases in QEEG Evaluation, Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience, 7:3-4, 31-52, DOI: [10.1300/J184v07n03\\_02](https://doi.org/10.1300/J184v07n03_02)

**To link to this article:** [http://dx.doi.org/10.1300/J184v07n03\\_02](http://dx.doi.org/10.1300/J184v07n03_02)

PLEASE SCROLL DOWN FOR ARTICLE

© International Society for Neurofeedback and Research (ISNR), all rights reserved. This article (the "Article") may be accessed online from ISNR at no charge. The Article may be viewed online, stored in electronic or physical form, or archived for research, teaching, and private study purposes. The Article may be archived in public libraries or university libraries at the direction of said public library or university library. Any other reproduction of the Article for redistribution, sale, resale, loan, sublicensing, systematic supply, or other distribution, including both physical and electronic reproduction for such purposes, is expressly forbidden. Preparing or reproducing derivative works of this article is expressly forbidden. ISNR makes no representation or warranty as to the accuracy or completeness of any content in the Article. From 1995 to 2013 the *Journal of Neurotherapy* was the official publication of ISNR ([www.isnr.org](http://www.isnr.org)); on April 27, 2016 ISNR acquired the journal from Taylor & Francis Group, LLC. In 2014, ISNR established its official open-access journal *NeuroRegulation* (ISSN: 2373-0587; [www.neuroregulation.org](http://www.neuroregulation.org)).

THIS OPEN-ACCESS CONTENT MADE POSSIBLE BY THESE GENEROUS SPONSORS



# Use of Databases in QEEG Evaluation

Jack Johnstone, PhD  
Jay Gunkelman, QEEG-D

**SUMMARY.** *Background.* Quantitative electroencephalography (qEEG) analysis incorporating the use of normative or reference database comparison has developed from being primarily a research tool into an increasingly widely used method for clinical neurophysiological evaluation.

*Method.* A survey of several of the most widely used qEEG databases as well as issues surrounding the construction and use of these databases is presented, comparing and contrasting the various features of these databases, followed by a discussion of critical issues in this developing technology.

*Results.* This review considers the concept of normalcy, norming of qEEG features, and validation of clinical findings. Technical issues such as methods for recording and analysis, filter use, broad bands versus single Hz finer frequency resolution, the number of variables relative to the number of cases, and the problem of multiple statistical testing are addressed. The importance of the recording electrode and montage reformatting for normative EEG data is emphasized. The use of multiple references is suggested.

*Discussion.* A brief review of the characteristics of several major da-

---

Jack Johnstone is President and CEO, and Jay Gunkelman is Executive Vice President, Q-Metrx, Inc., Burbank, CA.

Address correspondence to: Jack Johnstone, President, Q-Metrx, Inc., 2701 West Alameda Avenue, Suite 304, Burbank, CA 91505.

[Co-indexing entry note]: "Use of Databases in QEEG Evaluation." Johnstone, Jack, and Jay Gunkelman. Co-published simultaneously in *Journal of Neurotherapy* Vol. 7, No. 3/4, 2003, pp. 31-52; and: *Quantitative Electroencephalographic Analysis (QEEG) Databases for Neurotherapy: Description, Validation, and Application* (ed: Joel F. Lubar)

tabases is presented. Each has advantages and disadvantages, and newer databases will exploit new technological developments and increasing sophistication in statistical analysis of EEG data. Implementing new measures such as variability over time and extraction of features such as event-related desynchronization (see Pfurtscheller, Maresch, & Schuy, 1985) and gamma synchrony (Rennie, Wright, & Robinson, 2000) are likely to have important clinical impact. Caution is urged in the use of automated classification by discriminant analysis.

**KEYWORDS.** qEEG, databases, normalcy, montage reformatting, Laplacian, discriminant analysis

## ***INTRODUCTION***

Quantitative electroencephalography (qEEG) analysis refers to signal processing and extraction of features from the EEG signal. In typical practice, multichannel EEG is digitized, edited or adjusted to remove extra cerebral artifact, and subjected to spectral analysis using the fast Fourier transform (FFT). Extraction of features such as amount of power at each electrode for each frequency band or coherence among channels as a function of frequency is carried out for an individual and then compared to a “normal” group or another clinically defined group.

The use of a database has become an integral part of qEEG reporting, which usually also includes topographic color-graduated representation of EEG features (Duffy, Burchfiel, & Lombroso, 1979). Another important component of the qEEG study is visual inspection of the raw waveforms by a clinically experienced electroencephalographer (Duffy, Hughes, Miranda, Bernad, & Cook, 1994). Visual inspection of the EEG data is required to identify the possible presence of significant transient events as well as to evaluate transitions evolving over time, and assess the influence of extracerebral artifacts on the record.

There are many issues to be considered in construction and use of a comparison database in the clinical assessment of individuals. This review will be concerned with problems such as the definition of normalcy, how individuals are recruited and screened for inclusion in the database, types of EEG features that are normed, and the use of statistical analysis of EEG data. It is important to consider the specific reason

for use of database comparisons. Frequency tuning in neurofeedback applications may require analysis and display of single Hz information, while other applications (medical, legal, pharmacologic) may have different requirements. A comprehensive database should allow for a number of applications. Several databases currently in use will be examined with respect to these issues.

### *NORMALCY*

The definition of a database as a representation of the range of “normal” within a population raises the issue of what is meant by normal. A database could be comprised of many individuals in a population who are not rigorously screened for neuropsychiatric disorders, space occupying lesions, or aberrant neurophysiological functioning. If the population is very large, a simple statistical definition could be applied with those individuals falling close to the mean of a particular variable considered as normal and those with deviant scores considered abnormal. That is, the normal group falls within the bell shape of the normal curve and the abnormal group at the tails. Of course the problem here is that an individual may fall close to the mean for one variable but not other variables. A “pure” normal would be close to the mean for *all* variables. It is important to keep in mind that deviations from a database represent differences from average, not from optimal.

This is different from a clinically normal database where individuals are carefully screened for relevant abnormalities using other clinical tools such as psychometric assessment or MRI. In common clinical practice clinical and statistical deviation are often combined with limited screening (e.g., questionnaires) and statistical significance assessed conservatively (e.g., greater than three standard deviations from the mean). Further, there are technical concerns about calling a database representative of the normal population. To truly represent a particular population, stratified sampling must be employed. The database should represent the mix of ages, gender, ethnicity, socioeconomic status and other demographic factors present in the overall population. Most databases in current use in qEEG do not meet criteria for this level of norming, and are more appropriately considered reference rather than normative databases. Using a rigorously screened population, deviations represent the difference from well functioning as compared to average.

The method of recruitment of putatively normal individuals also should be considered. Often laboratory or office personnel are desig-

nated as normal because of good work skills and are used to populate the database. We have seen repeated occurrences of abnormal test results from office personnel who with sufficient questioning admit to occasional migraine headache or use of over the counter medications, etc. This points to the danger of poor screening procedures resulting in admitting individuals into a normal group who have significant clinical problems. When advertisements are used and paid volunteers are recruited, care must be exercised not to entice individuals to participate and falsify information because of the financial reward. We have advertised for boys to participate in a brain wave study and had parents bring in children that they suspected as having neurological or psychiatric problems for the hidden purpose of obtaining a free evaluation.

With respect to the extent of screening required to construct a proper database of normal individuals, certain practical constraints are relevant. It would be desirable to have full MRI, PET, fMRI, complete neuropsychological evaluation, genetic analysis, blood and urine testing, etc., but this may be prohibitively time consuming and costly. We argue that some form of screening is useful beyond simple self-report measures, which are well known to be unreliable. In addition to questionnaire information, objective assessment of general health, social and intellectual functioning is critical.

Special concerns apply to pediatric databases where dramatic developmental changes occur over relatively short time intervals. It is inappropriate to compare a four-year-old patient to norms derived from six-year-olds. The same two-year difference would be trivial in adults. It is possible to compute developmental equations over a relatively wide age range in childhood that reflects normal changes in development (Ahn et al., 1980). Departure from the normal trajectory of development of particular qEEG features could be considered a potentially clinically relevant finding. The size of the population required for computation of a stable set of developmental equations depends on the number of qEEG variables being studied. The larger the *cases to variables ratio*, the more stable and reliable the assessment. The number of subjects needed in a given database is larger as more measures are normed in order to account for use of multiple statistical tests.

### ***NORMING EEG FEATURES***

Evaluation of the pattern of deviations compared to a reference database is typically an integral part of qEEG evaluation. Most often a set of

parametric univariate Z-scores is computed as a way of detecting and characterizing potential abnormalities. The use of parametric statistics assumes a normal (gaussian) distribution of the variable(s) in question. Data transformation (log, square root, etc.) may be useful in meeting assumptions of parametric statistics. In addition, further exploration of other statistical methods (e.g., nonparametric) may be useful for application to the special requirements of EEG data.

It is not strictly the number of deviations, rather the *pattern* of deviations that is most relevant. Most database analyses do not allow for quantitative multivariate assessment of such patterns, and the overall pattern of significance must be reviewed by an individual with relevant experience in EEG, and both clinical and statistical evaluation. John et al. (1983) describe the use of the Mahalanobis distance statistic to capture patterns of regional deviation, for example, deviations involving the entire the left lateral or anterior brain regions.

Spectral power is an often used measure in qEEG studies. The amount of power for each frequency or frequency band for a given electrode is computed and compared to the database mean value. The results are usually represented as Z-scores, which is the difference between the mean score of a population and the patient's individual score divided by the standard deviation of the population.

There is a clear trend in the field toward more recording electrodes. Many recording systems now offer 32 to 40 channels. Further, using faster sampling rates, wider band passes (e.g., increasing the frequency setting of the high frequency filter), and higher resolution analog to digital conversion (A/D) allows for a more complete evaluation of the EEG signal over a wider range of frequencies. It is expected that new databases will provide norms past the 40 Hz range. The number of Z-scores increases as the number of electrodes, frequency bands, and recording conditions increases. The use of a large number of variables without corresponding increases in the sample size of the normal population increases the likelihood of deviation occurring by chance, unrelated to true neurophysiological abnormality. The number of false positive findings can be limited by requiring replication of patterns of deviation on independent samples of individual patient data.

Most databases available for clinical use contain the mean values of particular EEG features and the standard deviation of the feature across the normal population (see John, Pritchep, & Easton, 1987). Certain databases provide not only spectral power (or magnitude, the square root of power) but other derived measures such as relative power. Relative power represents the percentage of power in any bands compared with



the total power in the patient's EEG (e.g., relative theta is the percentage of theta of the combined sum of delta, theta, alpha, and beta). Other derived measures that have been normed include hemispheric asymmetry for power: comparing homologous electrode sites over the two hemispheres, as well as anterior/posterior power gradients. It should be recognized that these measures are not statistically independent, and significant deviations on more than one feature may be representing the same neurophysiological process.

Derived EEG features also include correlation or similarity measures such as coherence or the cross-spectrum, sometimes referred to as the comodulation (Serman & Kaiser, 2001). These measures index the similarity of activity between two recordings. When two electrodes are placed closely together on the scalp they pick up a large amount of common signal and recordings are highly correlated. Database comparisons are useful in showing when the signals are too correlated or not correlated enough as a function of the distance between the recording electrodes.

Phase measures the time delay between activities at two sites. The phase measures have also been normed. It is clear that phase is an important measure in understanding propagation of neuronal activity. Progressive phase delays are measurable as a volley travels from a source to a destination. In addition, 180 degree differences in phase denote polarity inversion and suggest the location of underlying generators. Using polarity inversion to model the source of brain macropotentials is useful in localizing activity and is commonly used in electroencephalography (Niedermeyer & Lopes da Silva, 1999). However, the meaning and utility of measures of average phase over time is not as clear.

Nearly all databases utilize features extracted in the frequency domain. Phase is a measure of timing derived from frequency domain analysis. It is also possible to norm measures directly in the time domain. The sequence and timing of neuronal activities following sensory stimulation can be measured very precisely and the timing and sequence of these events can be normed. Time domain analysis usually is carried out on the average response to the presentation of many sensory stimuli, the so-called averaged evoked potential (EP) or event-related potential (ERP; see Misulis & Fakhoury, 2001). Time domain techniques are very powerful in minimizing extra cerebral artifact not specifically linked to the presentation of the sensory stimuli. The ERP method is therefore a good candidate for recording under conditions of increased artifact, such as performance of complex psychomotor tasks.

The P300 is a well-characterized component of the ERP (for example, see Donchin, 1987). The usual procedure involves presentation of many standard stimuli, intermixed with occasional target stimuli. The stimuli are most often auditory tones or clicks but the procedure works generally independent of the sensory modality. The recognition of targets embedded in a series of standard stimuli is accompanied by a positive-wave recorded from the scalp at about 300 milliseconds following the stimulus, called the P300. The size and timing of the P300 component of the ERP is sensitive to the detectability of the stimulus, the speed of presentation, and a host of patient alerting, attentional and memory processes. We advocate the use of the P300 as a sort of treadmill test. The ability of a patient to detect and respond to increasingly difficult stimuli will be reflected in the timing (latency) and size (amplitude) of the P300. It is possible to determine at what point the P300 changes in character for a given individual responding to the increasing challenge of the P300 task. Norming this type of feature should provide a more robust measure of brain activity under stress because of the suppression of artifact by signal averaging. This procedure appears more amenable to routine clinical evaluation than presentation of complex tasks such as reading or math where frequency domain analysis is often so severely contaminated by muscle, eye motion, and other movement artifact that data are unable to be interpreted.

A problem with most of the currently available databases is the sole reliance on the linked ear reference. Often this reference is not only active but asymmetric. Problems using a single linked ear reference point for all analyses can be reduced by use of multiple references and by montage reformatting, as described below.

### ***MONTAGE REFORMATTING***

To record an EEG a multichannel recording amplifier is used. Each channel has a differential preamplifier, which has three electrical contacts: a ground contact, typically via a system or chassis ground which has a ground electrode contact on the patient; the two other contacts go directly into the differential preamplifier (Tyner & Knott, 1983). The preamplifier amplifies the voltage (E) difference between two electrodes placed in the two inputs, which are designated as grid one (G1) and grid two (G2) in electronic terms. This may be expressed as the following equation:



EEG equals grid 1 voltage minus grid 2 voltage

$$\text{EEG} = \text{G1(E)} - \text{G2(E)}$$

These two inputs give the EEG preamplifier the differential voltage, which fluctuates, or oscillates, over time creating the EEG waveform. This is simply showing the first grid's input electrode activity with respect to the grid two electrode activity. These combinations of inputs, summed to show the whole set of electrodes being monitored, is called the montage (French for mountings). A montage is selected to most clearly demonstrate the EEG pattern being monitored. One example is the controversial "14 and 6 positive spikes," which are *visualized* in ear reference montages more clearly than with sequential montages, though focal spike discharges are more easily *localized* with the sequential montages or with Laplacian/Hjorth techniques (Scherg, Ille, Bornfleth, & Berg, 2002).

Many will refer to the active electrode (grid 1) and the reference electrode (grid 2). Commonly used references include the ear references (linked ear, ipsilateral and contralateral ear), the Cz or vertex reference, and the sequential references (commonly termed bipolar). A more modern reference is based on Laplacian mathematics; it is variously called the Hjorth reference, local average, reference free, and virtual reference. Other computerized references include the common average or global average and the weighted average reference (Scherg et al., 2002).

Some montages need to have special electrodes applied, though these montages are not the subject of this paper. These include some obscure placements such as the tip of nose, the mastoid process, as well as the more obscure sternum-spinal reference (which cancels the EKG), or the more invasive references used in epilepsy research (e.g., the naso-pharyngeal or sphenoid leads) as well as direct cortical measurement (see Niedermeyer & Lopes da Silva, 1999).

Though it might seem comforting to be told which montage is the "right one" or the "best" montage, this is a more complex issue. The montages all have significant strengths and weaknesses, with the benefit being the ability to customize the montage to fit the finding that needs to be displayed. The weakness of this flexibility is in missing or inadvertently trivializing a phenomenon, or even creating a false image in the EEG if the reference selected is not a relatively neutral area electrically. An example is when strong temporal alpha contaminates the mastoid due to the lateral spread of the EEG through the skull. Thus, the ear references when contaminated create false alpha, displaced to areas

without alpha by the differential amplification system, which is blind as to the source of an oscillation, whether grid 1 or the reference at grid 2, the EEG output is an oscillation (see Gunkelman, 2000).

The selection of the montage needs to be based on the EEG but using a *variety* of montages is a part of the minimum guidelines for EEG, a guideline to which insurance companies can audit you for compliance. It is a commonly adopted guideline developed by the American Society of Electroneurodiagnostic Technologists. Their guidelines may be found on their web site ([www.aset.org](http://www.aset.org)). This document also specifies 20 to 30 minutes of total recording time to meet these guidelines.

This practice of switching perspectives avoids the reader being fooled by a false localization when a reference is contaminated with voltages as described above. The montage is like a perspective, it is used to present information from a particular point of view.

Each montage has strengths as well as weaknesses, including the problem of the neutrality of the reference. A single ear reference will avoid any problem associated with the other ear's contamination though it does not cancel the EKG as well as the linked ear reference, as well as creating an apparent asymmetry due to the systematically different inter-electrode distances between the two hemispheres.

The Cz reference will give good resolution for the temporal areas but not the central area. The Cz montage is also a poor choice where there is drowsiness, with the associated vertex sharp waves and spindles which are seen maximally along the midline anteriorly and at the vertex which will contaminate the vertex electrode.

The sequential placements, whether anterior-posterior or transverse, all have good localization of cortical events through phase reversal, but the raw wave morphologies are distorted by the phase cancellations as well.

The linked ear montage is subject to temporal lobe activity being seen on the reference, likely via volumetric conduction from the temporal activity and the influences of lateral diffusion of activity seen with the skull. The activity seen at T5 and T6 are most likely to be seen in the electrodes on the ears but activity at T3 and T4 may be influential as well.

The Laplacian technique or local average (or the Hjorth) is quite good at showing localized findings. The Laplacian montages are effective in elimination of the cardioballistic artifacts. The Hjorth derivation will also make the focal nature of the electrode artifacts seen easily. This is often not used for display due to the unforgiving nature of the display. There is a minor distortion at the edge electrodes, commonly a

small percentage error, though not a real problem for clinical utility. More problematic is the poor display of generalized or regional EEG findings, with false localization to the perimeter or edge of the finding (Scherg et al., 2002).

The global average is also a Laplacian technique, with the average of all electrodes used as the reference. This technique also has distortions when displaying generalized changes, especially generalized paroxysmal activity. This global reference may be done with a spatial weighting factor, which tunes the filter spatially to be more or less sensitive to focal findings. This weighted average reference is very popular when using dense arrays in topographic mapping.

One of the major factors in selecting montages is that the database selected has to have the same montage for the norms and to have any relevance to the data. Comparisons must be carried out with the same montage. This does not preclude reviewing the raw EEG data with these various montages, as well as creating topographic maps of these various montages to display the EEG, though it is mandatory that the proper montage be compared to the database.

The proper selection of montages will allow the reader to get a good view of the EEG phenomenon and its distribution across the cortex. Understanding of the topographic distribution of brain function is required to understand what the neuropsychological impacts of the EEG changes might be.

## **VALIDATION**

In order to yield valid representations of neurophysiological abnormality by statistical deviation, the influence of artifact must be taken into account. Artifact can be generated by a variety of extra cerebral sources commonly including muscle, cardio-ballistic propagation, eye motion, sweat (GSR), and movement (see Hammond & Gunkelman, 2001). Since these artifacts occur to some extent in virtually all EEG recordings it is useful to specifically record and characterize the extent of artifacts. Concurrent recording of EMG, eye movement, and EKG are commonly used in EEG evaluation. High correlation between scalp recorded data and data from artifact channels increases the likelihood that the EEG is contaminated by artifact. The effects of correlated artifactual data can be reduced by statistically removing the signal using partial correlation. Although this is not commonly done, there is no substitute for making every attempt to minimize artifact at the time of the record-

ing. It should be emphasized that at this time there are no validated procedures for automatic rejection of EEG artifact.

Another important influence on the EEG involves the effects of psychoactive medication. Normative databases do not include individuals taking medication whereas many if not most patients are taking medication, and often multiple medications, and in fact may not even report the use of over the counter medication or street drugs. Effects of medications are known to substantially alter the EEG frequency content, often causing large increases in slow or fast activity. The only methods useful in limiting these effects are (a) have a referring physician withdraw the medications prior to the test, which is generally not practical; or (b) take known effects of medications into account in the interpretation. It is clear that it is desirable to also verify medication status with drug screening for individuals included in a normal database.

A related issue involves non-medication supplements, or agents such as used in hormone replacement therapy. Since these agents replace normally occurring hormones it may not be necessary to consider them in the same way as pharmaceuticals in general. However, many replacements are synthetic and may not cause the same effects as naturally occurring hormones. Many vitamins, particularly b-vitamins, and nutraceuticals also have direct effects on the EEG.

Another difficult issue is the influence of patient drowsiness. Decreases in patient arousal and increased drowsiness can be expected routinely in EEG recording. These effects may be very subtle, and in fact may be the essence of the patient presentation and complaint. The individual recording the EEG should be aware of the common effects of mild drowsiness, namely a decrease in posterior alpha activity and increased slow activity, usually over the frontal midline. Attempts should be made to monitor for the effects of drowsiness and alert the patient as necessary at the time of the recording. Sophisticated and cost-effective monitors of the effects of drowsiness and loss of conscious awareness are now available and may be used concurrently to assure that the patient is alert, or if not, quantify the level of patient awareness (see Sigl & Chamoun, 1994; Johnstone, 2002).

A number of methods for deletion of artifact are available but automated methods are not generally used in clinical practice at this time. Typically, segments of EEG containing significant artifact are simply deleted from analysis based on visual inspection by a trained technician or clinical specialist.

In addition to validating the quality of EEG data, another type of clinical validation involves assuring that appropriate deviations occur with

cases of known pathology, such as amplitude suppression of focal slowing post-stroke or diffuse excessive slow activity in advanced dementia.

### ***DATABASES IN PRACTICE***

Several databases are commercially available for clinical use. Several widely used databases have been reviewed and compared by Lorensen and Dickson (2001). Following is a brief review of several currently available databases and a description of a new multifactorial, comprehensive database currently under construction.

#### ***Neurometrics***

The first database developed for the purpose of general neurophysiological evaluation was constructed by John, Prichep and Easton (1987). The term “neurometrics” was first used by this group to describe an analogy to psychometric assessment, commonly used in clinical psychology (John et al., 1977). Neurometrics refers to the comparison of individual EEG features with a reference database and is used in much the same way as IQ testing. A standardized test is constructed using a large population of individuals and the relative standing of the test results for a given individual within that overall population is assessed. John et al. (1987) stressed the need for standardization of recruitment, recording, and analysis procedures.

The Neurometric database is based on a specific set of EEG features: absolute power, relative power, coherence, mean frequency within band, and symmetry (left-right and front-back) extracted from approximately two minutes of data selected for being minimally contaminated by artifact. Only recordings made with eyes closed at rest were analyzed and normed. The EEG frequency range analyzed extends from .5 to 25 Hz. Extracted features were transformed to assure a gaussian (normal) distribution. Two thousand and eighty-four (2,084) variables are computed for each member of the database. The correlation of EEG features with age was noted and best fit age regression equations were developed to account for age effects. Univariate and multivariate Z-scores were computed for the purpose of characterizing an individual's deviations from the mean of the population. This database includes measures from some 782 normal individuals. Of this total 356 cases were between the ages of 6 to 16 and 426 cases were between the ages 16 to 90. Over

4,000 clinical cases were used in the discriminant section of the software.

Individuals selected for inclusion in the Neurometric database were screened by questionnaire to exclude head injury, neurological or psychiatric disease, any history of psychological problems, alcohol or drug abuse, any use of psychotropic medication, and academic or social problems.

One important feature of the Neurometric database is the availability of normed features for specific sequential (bipolar) electrode pairs. This feature allows for at least some assessment of the effects of activity recorded with the linked ear reference.

We have had extensive experience with the Neurometric database and have found it to be useful in both characterizing abnormalities detected by visual inspection, as well as in identifying patterns of deviation which appear to be clinically significant but are not easily detected by inspection of the raw EEG signals. We appreciate that the Neurometric database has received a 510(k) clearance by the FDA (July, 1998, #K974748), indicating that construction of the database has been scrutinized for good manufacturing practices (GMPs). The 510(k) also signifies the legitimacy of marketing claims made concerning the database.

The most significant problem with the Neurometric database is exclusive reliance on banded EEG. Only information about delta, theta, alpha, and low frequency beta bands are available. Findings restricted to narrow frequencies are often seen when data are displayed in single Hz increments but are obscured with the use of the relatively wide bands as normed in the Neurometric database.

#### ***Thatcher Lifespan Normative EEG Database (LSNDB/NeuroGuide)***

The database developed by Robert W. Thatcher has been described in detail (Thatcher, 1998). Subsequently, during 1999-2000 new analyses presented under the commercial name "NeuroGuide" were completed (see [www.appliedneuroscience.com](http://www.appliedneuroscience.com)). The lifespan database was reconstructed starting with the same raw digital EEG values from the same normal subjects. This database now contains information from 625 individuals, covering the age range two months to 82.6 years. More advanced methods were used to compute the revised database, including more extensive cross-validation and tests of gaussian distributions for average reference, linked ears, Laplacian, eyes open and eyes closed. The NeuroGuide database has been tested and re-tested and the sensi-

tivity of the statistical distributions has been calculated for each montage and condition.

Normalcy was determined by response to a neurological history questionnaire which was given to the child's parents and/or filled out by each subject. IQ and other age appropriate psychometric testing, academic achievement, as well as class room performance as determined by school grades and teacher reports also were used in determining normalcy.

Nine hundred and forty-three (943) variables were computed for each subject including measures of absolute and relative power, coherence, phase, asymmetry, and power ratios. Z-score transforms are available in single Hz bins. Sliding averages were used to compute age-appropriate norms. Results were inspected for gaussian distribution. Recording with task challenges was not performed (for further details see Thatcher, Walker, Biver, North, & Curtin, Sensitivity and cross-validation of a lifespan normative EEG database at: <http://www.appliedneuroscience.com/LSNDBWEB.htm>).

NeuroGuide was considered to not require FDA 510(k) clearance, based on both the non-medical nature of the intended use and the fact that databases are considered tables of numbers involving library functions. Overall, the construction and composition of this database are relatively well documented.

### ***Sterman-Kaiser (SKIL) Database***

The SKIL database currently includes 135 adults ranging from 18 to 55 years old (see Sterman & Kaiser, 2001, Appendix: Adult Database Description, available at: <http://www.skiltopo.com/manual.htm>). No normative information is currently available for children or young adults, although data is being collected to cover the younger ages. SKIL does not consider age as a factor in computing Z-score deviations. The reference population is comprised of students and laboratory personnel (50%), volunteers recruited from the community (25%), and U.S. Air Force personnel (25%). Screening was based on questionnaires regarding medical history, drug use, and recent life events.

The SKIL database incorporates recordings at rest (eyes closed and open) and during task challenges involving audio-visual information processing and visual-spatial tracking. A correction for the time of day of recording is available which is based on combined cross-sectional and longitudinal data rather than the preferred method of tracking within-subject changes over time.



The SKIL database covers a restricted range of frequencies, from 2 to 25 Hz. This method deletes significant slow and fast frequency data which may be of clinical importance. However, the database does have the advantage of providing norms for each single Hz increment over this frequency range. The SKIL database relies exclusively on the linked ear reference. The SKIL analysis has not received FDA 510(k) clearance and is labeled as not intended for medical use.

The SKIL database does not have a measure of EEG coherence but rather includes a similarity measure termed comodulation, which is quite like coherence but does not yield a measure of phase. Comodulation is essentially the correlation of the spectrum for two recording electrodes over time using a sliding one second data window moved in 250 millisecond increments. Although this effectively deals with windowing issues, it also is clear that the degree of correlation between electrode sites is not computed on independent, but rather overlapping, spectral analyses.

Since the SKIL database relies exclusively on the linked ear reference, the comodulation similarity measure is strongly influenced by the fact that both sites are connected to a common source (for a review of problems with a common reference when using similarity measures see Fein, Raz, Brown, & Merrin, 1988). Nevertheless, the comodulation metric is currently being explored for possible clinical utility.

As discussed previously, there must be a balance between the number of individuals in a database and the number of variables used in assessment to account for multiple statistical tests. The SKIL database has the advantage of a large number of features (e.g., multiple conditions, single Hz bins) but has the disadvantage of a relatively small number of individuals represented in the database. This tends to increase the number of false positive findings. A solution to the problem of false positive results is to replicate findings on independent samples of data from the individual patient.

### ***The International Brain Database***

One of the most exciting developments involving qEEG database construction is the development of the first standardized International Brain Database. It overcomes the ubiquitous problems about databases that Chicurel (2000) summarized—namely that “. . . technical problems are huge, and reaching a consensus on what to archive won’t be easy.” A consortium of leading neuroscientists were consulted to resolve an

optimal choice of tests that tap the brain's major networks and processes in the shortest amount of time. Six sites have been set up with identical equipment and software (New York, Rhode Island, London, Holland, Adelaide, and Sydney) under the auspices of a publicly listed company (The Brain Resource Company—[www.brainresource.com](http://www.brainresource.com)), with new sites to be added progressively.

Hundreds of normative subjects have been acquired and the assessment of clinical patient groups has also recently begun. One thousand (1,000) normal controls and 1,000 patients (across the age spectrum) will be collected in the first phase.

A key dimension of this initiative (in addition to the database) is new and sophisticated analyses of EEG, ERP and autonomic activity (heart rate and electrodermal activity which are collected at the same time as the EEG/ERP). This allows not only the evaluation of state (arousal) versus trait effects, but in addition, a numerical simulation of the brain allows interpretation of EEGs according to fundamental whole brain physiological principles (in addition to simply quantifying frequency power).

Another of the most interesting new analysis methods is of 40 Hz activity (Gamma synchrony). Gamma synchrony related to cognitive processing has been observed even up to the whole brain level, and with widely separated EEG electrodes (e.g., between hemispheres). It seems, therefore, that synchrony may be an important coding mechanism across multiple scales of brain organization.

This International Brain Database involves data collection not only of EEG/ERP/autonomic in a battery of psychophysiological activation tasks, but also a comprehensive psychological test battery undertaken using a touch-screen monitor. The individual tests are listed below:

#### *Psychological Test Battery*

- Choice reaction time (speed of motor performance)
- Timing test (capacity to assess time)
- Digit span (short term memory).
- Memory Recall Test (12 words repeated 5 times with a matched distracter list after trial 4)
- Spot The Word Test (word:non-word index of IQ)
- Span of Visual Memory Test (4-second delay test of spatial short term spatial memory)
- Word Generation Test (verbal fluency test)

- Malingering Test (number recognition malingering test)
- Verbal Interference Test (test of inhibitory function)
- Switching of Attention (alternation between numbers and letters)

*Psychophysiology Paradigms (NeuroScan Nuamps 40 channel/Grass electrodermal)*

- Startle paradigm (fight and flight reflex)
- Go-NoGO (inhibition)
- Resting EEG (cortical stability)
- Visual tracking task (automatic tracking)
- Habituation paradigm (novelty learning)
- Auditory oddball (efficiency of target processing)
- Visual oddball (visual novelty target processing)
- Conscious and subconscious processing of facial emotions
- Visual working memory task (memory and sustained attention)
- Executive maze task (planning and error correction)

Specific event related potential measures, including P300 will be obtained. Structural and functional MRI will also be obtained for many selected individuals. Further, genetic information will be systematically collected for comparison to neuroanatomical, neurophysiological, and psychometric measures. The goal is to construct a database which can be used to integrate information directly across a variety of indices of brain structure and function.

***Others***

Other databases are also under development, including one using advanced EEG tomographic analysis called Low Resolution Electromagnetic Tomography (LORETA). The NovaTech EEG database currently has 84 cases and is actively adding cases. This EEG imaging technology allows for a tomographic representation of EEG sources in 3-dimensional space (see [www.NovaTechEEG.com](http://www.NovaTechEEG.com)). This database will be useful not only in identifying deviations but approximating the location of the brain regions involved.

Hudspeth offers the NeuroRep AQR (Adult QEEG Reference Database; see [www.neurorep.com](http://www.neurorep.com)). One of the most useful features of Hudspeth's work is the emphasis on *reliability* of measures obtained from individual patients, and the importance of EEG variability over time as a clinical index. EEG data are available for both eyes open and

closed conditions. The database includes measures of absolute and relative power for 19 scalp electrodes, and all combinations ( $N = 171$ ) of pair wise electrode comparisons for coherence, phase, asymmetry and correlation indices. High quality graphic representations of raw data and database comparisons also are included.

The total number of individuals in the AQR is now rather small ( $< 50$ ) but additional data is being collected. Largely because of Hudspeth's extensive experience with Neurometric analysis and EEG data handling, this method can be expected to develop into an increasingly useful clinical product over time. Hudspeth (personal communication, September 14, 2002) also has presented new materials on individual patient assessment without using a database.

Frank H. Duffy, inventor and promoter of brain electrical activity mapping (BEAM; see Duffy et al., 1979), also has constructed an EEG database. This database includes both eyes-closed and eyes-open conditions, and spans a wide age range, including both children and adults. EEG data are available for 19 electrodes, and auditory and visual evoked potential measures also are included. This database was previously used in several commercial neurodiagnostic instruments (Nicolet, QSI) but to our knowledge, is not currently commercially available.

### *Comparison of Databases*

Since QEEG databases are based on different inclusion criteria, recording methods and statistical analysis techniques, it is apparent that comparing a patient to multiple databases will likely yield different patterns of deviations. Comparing an individual to different databases confounds patient characterization with differences among the databases themselves. Therefore, we do not recommend using multiple databases in the characterization of individual patients, but rather the selection of a database that best suits the individual case.

## **DISCRIMINANT ANALYSIS**

Both John's Neurometric analyses and Thatcher's LSNDB/Neuro-Guide include the use of discriminant analysis and clinically defined databases. The discriminant analysis is a statistical technique based on the general linear model which is used to select and weight specific variables recorded from an individual. The linear combination of these variables may be optimized so that individual patient scores are generated which correctly categorize the patient as a member of their known

clinical group. When a rule is developed that successfully categorizes patients into their correct group with a high degree of accuracy, the rule is then applied to unknown individuals to see if they fit in a specific group. Neurometric analysis includes discriminant analyses for overall abnormality, learning disability, attentional disorders, alcohol abuse/addiction, schizophrenia, depression, dementia, and head injury. However, as John has pointed out, the discriminants should only be used with individuals who have known histories of the disorders being classified. It is not legitimate to run a patient through all the discriminants to see which one “hits” and then use the hit as a diagnosis. The Thatcher LSNDDB has discriminant analyses available for head injury evaluation and learning disability/hyperactivity.

There are several important issues in the use of discriminant analysis (Duffy et al., 1994). First, the discriminant attempts to classify behavioral disorders based on EEG features. It is well established that a variety of underlying patterns may be seen in seemingly homogeneous behavior disorders. Suffin and Emery (1995) have clearly demonstrated subgroups of neurometric patterns in two groups of patients with either affective or attentional disorders. Seeking invariant neurophysiological markers that are pathognomonic of complex behavioral disorders is difficult at best. If one is attempting to diagnose a learning disorder such as dyslexia, it would seem important to administer reading tests, not perform EEG discriminant analysis.

Secondly, since discriminant analyses rely on optimizing techniques, it is important to assure that a rule developed on a training set of data applies to a similarly defined test set. Without independent cross validation it is not possible to use a rule on a given individual. Even though replication by jackknifing (leave-one-out and repeat the analysis iteratively) procedures appear successful at identifying internal validity and consistency, discriminants may not yield information valuable to the treating clinician, or hold any external validity. Reports of success of discriminants often exceed 90% classification accuracy, but when used in patients with real-life comorbidities and polypharmacy, or excessive drowsiness, the techniques cannot be legitimately applied.

Further, most discriminants have been most extensively implemented as two-way classifiers. For example, a depression discriminant can only yield information on whether a patient has depression. It does not suggest that the patient is alcoholic or head injured. For these reasons, despite the apparent accuracy and reliability of certain discriminants, the discriminant analysis method has not been widely adopted in clinical practice.

## CONCLUSIONS

This brief review has considered some of the important issues in construction and use of normative EEG databases, including operationalizing the concept of normalcy. Methods for recording and analysis constrain the conclusions that can be made using database comparisons. If filters are set to pass only a relatively narrow range of EEG frequencies, no statements can be made about the frequencies filtered out and lost to subsequent analyses. If broad bands are used, important deviations appearing at single frequencies or a restricted range of frequencies will be missed.

A critical concern is the number of variables in a database relative to the number of cases along with the problem of multiple statistical testing. Most databases contain an unfavorable variables/cases ratio and replication of findings should be strongly encouraged in order to limit false positive results. As databases grow to incorporate more recording electrodes, frequency bands, and task conditions the problem of false positive findings is more relevant.

The importance of the recording electrode and the need for remontaging norms has been emphasized. The linked ear reference is a particularly poor choice of references since it can be both active and asymmetric, yet historically it has been the default reference for most databases. Use of multiple references is suggested to minimize the bias inherent in use of a single reference point.

A brief review of the characteristics of several major databases is presented. Each has advantages and disadvantages but newer databases will exploit new technological developments and increasing sophistication in statistical analysis of EEG data. Implementing new measures such as variability over time and extraction of features such as event-related desynchronization (see Pfurtscheller, Maresch, & Schuy, 1985) and gamma synchrony (Rennie, Wright, & Robinson, 2000) are likely to have important clinical impact. Caution is urged in the use of automated classification by discriminant analysis.

## REFERENCES

- Ahn, H., Prichep, L. S., John, E. R., Baird, H., Trepetin, M., & Kaye, H. (1980). Developmental equations reflect brain dysfunctions. *Science*, *210*, 1259-1261.
- Chicurel, M. (2000). Databasing the brain. *Nature*, *406*, 822-825.
- Donchin, E. (1987). The P300 as a metric for mental workload. *Electroencephalography and Clinical Neurophysiology*, *39* (Suppl.), 338-343.

- Duffy, F. H., Burchfiel, J. L., & Lombroso, C. T. (1979). Brain electrical activity mapping (BEAM): A method for extending the clinical utility of EEG and evoked potential data. *Annals of Neurology*, 5 (4), 309-321.
- Duffy, F. H., Hughes, J. R., Miranda, F., Bernad, P., & Cook, P. (1994). Status of quantitative EEG (QEEG) in clinical practice. *Clinical Electroencephalography*, 25 (4), VI-XXII.
- Fein, G., Raz, J., Brown F. F., & Merrin, E. L. (1988). Common reference coherence data are confounded by power and phase effects. *Electroencephalography and Clinical Neurophysiology*, 69 (6), 581-584.
- Gunkelman, J. (2000). Hjorth referencing in QEEG. *Journal of Neurotherapy*, 4 (1), 57-62.
- Hammond, D. C., & Gunkelman, J. (2001). *The art of artifacting*. Salt Lake City, Utah: Society for Neuronal Regulation (see [www.snr-jnt.org](http://www.snr-jnt.org)).
- John, E. R., Karmel, B. Z., Corning, W. C., Easton, P., Brown, D., Ahn, H. et al. 1977. Neurometrics: Numerical taxonomy identifies different profiles of brain functions within groups of behaviorally similar people. *Science*, 196, 1393-1410.
- John, E. R., Prichep, L., Ahn, H., Easton, P., Fridman, J., & Kaye, H. (1983). Neurometric evaluation of cognitive dysfunctions and neurological disorders in children. *Progress in Neurobiology*, 21 (4), 239-290.
- John, E. R., Prichep, L. S., & Easton, P. (1987). Normative data banks and neurometrics. Basic concepts, methods, results of norm construction. In A. S. Gevins & A. Remond (Eds.), *Methods of analysis of brain electrical and magnetic signals* (pp. 449-495). Amsterdam: Elsevier.
- Johnstone, J. (2002). Bispectral analysis of the EEG: A brief technical note. *Journal of Neurotherapy*, 6 (3), 77-81.
- Lorensen, T., & Dickson, P. (2002). Quantitative EEG normative databases: A comparative investigation [Abstract]. *Journal of Neurotherapy*, 6 (1), 89-90.
- Misulis, K. E., & Fakhoury, T. (2001). *Spehlmann's evoked potential primer* (3rd ed.). Philadelphia: Elsevier (Butterworth-Heinemann).
- Niedermeyer, E., & Lopes da Silva, F. (Eds.). (1999). *Electroencephalography: Basic principles, clinical applications, and related fields* (4th ed.). Baltimore: Lippincott, Williams & Wilkins.
- Pfurtscheller, G., Maresch, H., & Schuy, S. (1985). ERD (event-related desynchronization) mapping: A new procedure for functional brain diagnosis. *Biomedical Technology (Berlin)*, 3 (1-2), 2-6.
- Rennie C. J., Wright, J. J., & Robinson P. A. (2000). Mechanisms of cortical electrical activation and emergence of gamma rhythm. *Journal of Theoretical Biology*, 205 (1), 17-35.
- Scherg, M., Ille, N., Bornfleth, H., & Berg, P. (2002). Advanced tools for digital EEG review: Virtual source montages, whole-head mapping, correlation, and phase analysis. *Journal of Clinical Neurophysiology*, 19 (2), 91-112.
- Sig J. C., & Chamoun N. C. (1994). An introduction to bispectral analysis for the EEG. *Journal of Clinical Monitoring*, 10, 392-404.
- Sterman, M. B., & Kaiser, D. (2001). Comodulation: A new qEEG analysis metric for assessment of structural and functional disorders of the central nervous system. *Journal of Neurotherapy*, 4 (3), 73-83.



- Suffin, S. C., & Emory, W. H. (1995). Neurometric subgroups in attentional and affective disorders and their association with pharmacotherapeutic outcome. *Clinical Electroencephalography*, 26, 76-83.
- Thatcher, R. W. (1998). EEG normative databases and EEG biofeedback. *Journal of Neurotherapy*, 2 (4), 8-39.
- Tyner, F., & Knott, J. (1983). *Fundamentals of EEG technology Vol. 1*. New York: Raven Press.