Choosing Adequate Average Number in Electrophysiological Studies using Correlation Coefficient

F.-Z. Shaw, R.-F. Chen*, H.-W. Tsao and C.-T. Yen*

Department of Electrical Engineering and *Department of Zoology,
National Taiwan University, Taipei, Taiwan 10764, Republic of China

Abstract— A theoretical relationship was derived among the number of trials averaged (N), initial signal-to-noise ratio (SNR) and correlation coefficient (r). The validity of this relationship was tested with a simulate signal in different amount of noise. By selecting 0.95 as an acceptable r value, the adequate average number could be determined for different SNR's conditions.

INTRODUCTION

Ensemble average is commonly used to enhance the extraction of signal waveform in many applications, for example, brain evoked potential, ECG,....,etc. If the signal is consistent and uncorrelated to background noise, better waveform can be obtained by increasing the number of trials averaged. However, more trials averaged means more computing time, more man power, and more suffering of the patients (either human or animal). Furthermore, less averaging efforts seems to be needed in situations with high signal-to-noise ratio (SNR). To find the adequate number of averaging, we chose the correlation coefficient (r), a general similarity index between two signals, as the judging criterion [1]. The theoretical relationship among the trial number averaged, the SNR and the correlation coefficient was derived first. Simulate signal with different amounts of white noise were averaged and r values obtained between the averaged waveform and the true waveform. The validity of the theoretical relationship was examined with the simulation data.

METHOD

The Ensemble average is based on the assumption that the evoked response by each stimulus is the same from trial to trial, and that the noise is random, uncorrelated with the response and has a mean value of zero. According to the definition of the correlation coefficient (covariance of the two signals normalized by the square root of the variances of the two signals) and the assumptions described above, a concise relationship of the average number (N), initial SNR and correlation coefficient r can be obtained:

$$r = \cfrac{1}{\sqrt{1 + \cfrac{1}{N}\cfrac{\text{var}(n)}{\text{var}(s)}}} = \cfrac{1}{\sqrt{1 + \cfrac{1}{N}\cfrac{1}{SNR}}} \qquad (1)$$

$$N = \frac{r^2}{1 - r^2} \times \frac{1}{SNR} \qquad (2)$$

where $SNR = \dfrac{\text{var}(s)}{\text{var}(n)}$. Var(s) and var(n) represent the variances of the evoked signal and the noise, respectively. If r is fixed, N is inversely related to SNR as in equation 2.

The recorded signals contained two components: the evoked response by the stimulus and the background noise. We used a one cycle sinusoidal wave as the evoked response. Gaussian white noise was selected as the background noise. We obtained different SNR by changing the variance of the white noise. The whole simulation was developed using a commercial package (Labview, National Instruments, Austin).

RESULTS

The theoretical relationship between r and N under three different SNRs is shown in Fig. 1. The performance of the average process was similar to a half sigmoidal curve. The capacity to enhance evoked response decreased as the N value grew larger. If we selected r=0.95 as the acceptable level, the adequate average number (Na) related inversly to the SNR (Fig. 2). The deviation of the adequate average number between of the theoretical prediction and of the simulation results was very small as illustrated in Fig. 2.

DISCUSSION

The limited usefulness of over-averaging could be clearly seen in Fig. 1. The capacity of enhancing signal SNR descreased as the number of trials averaged increased to near saturation level. This implied that excessive average was not necessary. In other words, too much average could consume more computing time, waste more man power and cause more suffering of the patients. By using a commonly used similarity index, the correlation coefficient, and using r=0.95 as the acceptable threshold, adequate average number, Na, was estimated. This gives the information about suitable average number *in posterior.*

Modification of the same principle could be used to estimate adequate trial number before experiment. This will be very helpful in situations where stimuli are painful or uncomfortable (e.g. somatosensory evoked potential) or the same situation cannot be repeated many times (e.g. most event-related potentials). A way of modification may be when all leads have been connected, a few test trials may be given such that the SNR condition could be estimated. Fig. 2 could then be consulted about the adequate number of trials under the circumstances.

REFERENCE

[1] R.E. Challis and R.I. Kitney " Biomedical signal processing. Part 1 time-domain methods," Med. & Biol. Eng. & Comp., vol. 28, pp. 509-524, 1990.
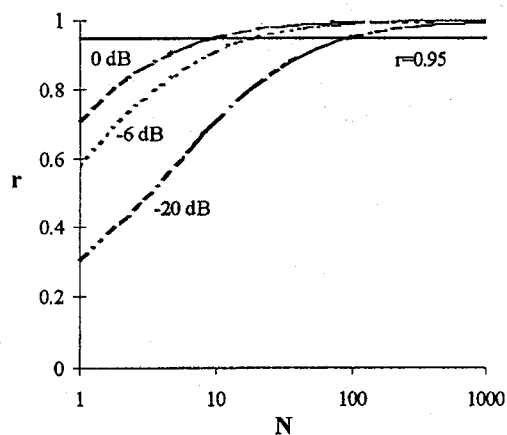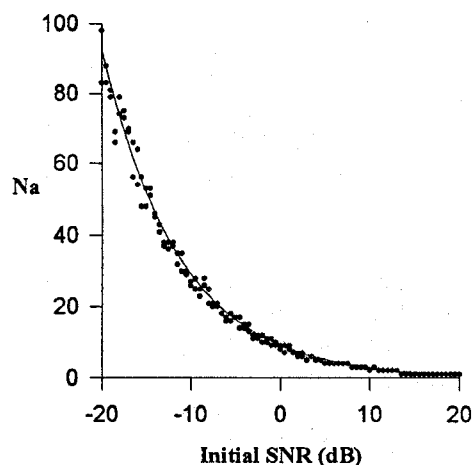
Fig. 2 The adequate average number (Na) plot against initial SNR conditions. The deviation between prediction data (solid line) and simulation data (solid dots) is very small.

Fig.1 The semilogrithic plots of the theoretical relationship of r, N and initial SNR.