Chapter 5. Future Research

In this section, we discuss particular shortcomings of study design and research in the available literature. By inference, these shortcomings point the way towards future research. We then discuss the optimal designs of trials that could answer many outstanding questions. While these are optimal design characteristics, they may not always be practical. It is impossible, for example, to blind patients to the fact that they have received surgery. However, to the extent that it is possible to adopt optimal procedures, they should be adopted.

Gaps in Current Research

Lack of adequate statistical power

A consistent theme observed throughout the literature on WRUEDs is a lack of statistical power. Studies that do not contain adequate numbers of patients cannot detect clinically meaningful differences in outcomes between-treatment groups. When designing clinical trials, *a priori* power analysis calculations can be used as a guide to ensure that sufficient numbers of patients are enrolled so that the proposed trial can answer the questions it is investigating.

Inclusion of hands, rather than patients

It is tempting, in bilateral cases of WRUEDs, to count the number of arms/hands treated rather than the number of treated patients. However, when one does so, the data are not independent. Therefore, statistical procedures that take this lack of independence into account must be used for data analysis.

Outcomes

The primary outcome measures in trials of WRUEDs are often physiological measurements such as nerve conduction velocity and grip strength. Although such outcomes are of interest, the correlation between the effect of WRUEDs on physiology and their effect on patients' lives is not well established. Outcomes of greater applicability include measurements of the effect of the disorder on the patient's quality of life and on the patient's ability to work and perform common activities of daily living. An additional shortcoming of the available literature is the incomplete reporting of harms, morbidities, and complications of treatment.

Sufficient length of time of followup

WRUEDs are often chronic conditions that affect patients for many years. Studies that evaluate the effect of a treatment for only a few weeks are unlikely to have followed patients for a long enough period of time to allow for definitive conclusions about the effectiveness of a treatment.

Intent-to-treat statistical analysis

Intent-to-treat statistical analysis is the accepted method of handling attrition from clinical trials. Trials that do not use intent-to-treat statistical analysis may come to incorrect conclusions.

Diagnostics

It is difficult to evaluate the usefulness of a diagnostic test without first establishing a "gold standard" diagnostic method. This difficulty appears likely to remain because there currently appears to be no test that is widely accepted as a gold standard. Nevertheless, improvements in studies of diagnostic tests are possible. Thus, although it is appropriate to perform pilot studies of diagnostic methods on groups pre-selected to contain only definite "normals" and "diseased", the specificities derived from such studies will be inaccurate. Therefore, these specificities may not reflect those one will obtain in actual clinical practice. The relevance of diagnostic studies to this practice can be increased by evaluating a diagnostic test in a population like the one in which it will be used in clinical practice. The accepted method of analyzing diagnostic data, ROC analysis, has been rarely used in this literature. Most of the published articles on diagnostic tests for WRUEDs reported results at only one diagnostic threshold, and usually selected thresholds based on arbitrary criteria rather than on an objective analysis of the consequences of false positive and false negative results.

Optimal Study Designs

Prospective, randomized double-blinded controlled trials are widely considered to provide the highest quality of evidence for treatment effectiveness. Non-randomized trials may have differences in outcomes between patient groups because of differences in the characteristics of the patient groups, rather than the treatment applied. Trials without a control group are unable to examine the potential for recovery in the absence of treatment, and they do not allow one to accurately gauge the magnitude of any change that occurs after treatment. Blinding of patients and evaluators to treatments avoids the potential for placebo effects and previously held beliefs about the effectiveness of treatments to impact on the results of trials.

Studies of diagnostic tests need not be randomized or contain concurrent control groups. In the absence of a "gold standard" test, longitudinal studies that employ clinical outcomes as the gold standard are desirable for assessing diagnostic tests for WRUEDs. In these studies, patients are first given the diagnostic test, and then they are followed for a period of time to see if they develop symptoms of a WRUED. Repeating the tests at regular intervals during the trial could yield insights into the etiology of the conditions as well as measure test-retest variability. Controlled studies designed to gather epidemiological data and identify risk factors are often not possible. Thus, well-designed observational cohort studies are accepted as the optimal design to gather this sort of information. In order to generate generalizable data, it is important that cohort studies enroll sufficient numbers of patients and follow the patients for sufficient periods of time.