Letters to the Editor

Reference values in concentric needle electrode studies

The paper by Kokubun et al. describes results from a multicenter study to establish reference values for jitter measurements performed with concentric needle electrodes (CNEs) (Kokubun et al., 2012). Because the variation in timing of composite signals recorded with CNEs is affected not only by the neuromuscular jitter, but also by indeterminate interactions among the composing action potentials (APs), Stålberg and Sanders (2009), it is imperative to obtain reference values for jitter measured with these electrodes, and the authors are to be commended for undertaking to do this. However, we have theoretical and operational concerns about their conclusions.

The authors’ statement that “there is no theoretical reason that jitter values should be different between SF and CN electrodes” is at variance with published observations (Stålberg and Sanders, 2009). Signals recorded with larger recording surfaces, even if visibly indistinguishable from single fiber APs, may actually be composed of synchronized and near-simultaneous APs from more than one muscle fiber, which will inevitably alter jitter measurements.

Jitter measured between two APs during voluntary activation represents the neuromuscular transmission variability at two end-plates, and is thus inherently larger than the jitter measured in single APs during axonal stimulation. Reference values for stimulation jitter studies should reflect this difference, unlike those reported by the authors, which are the same for voluntary and stimulation jitter studies. The authors’ reference values also differ from those that we have previously reported from CNE studies in which jitter was measured using the peak-detection algorithm (Kouyoumdjian and Stålberg, 2008, 2012). Their mean overall and individual MCD values were only slightly greater, but their SD values were much larger, which gives much higher normal cut-off limits. Several factors may have contributed to this:

(1) The same type of CNE was not used by all centers, inevitably adding variability to the resulting measurements.

(2) The jitter measurement technique (“voltage threshold” or “peak-detection”) was not standardized among laboratories. These techniques may give different results, particularly for riding signals, common in CNE, in which case the peak-detection method is more accurate. It is surprising that only signals with interpotential intervals (IPIs) <2 ms were accepted; for SFEMG this value is 4 msec, allowing well separated signals to be included.

(3) Another possible explanation for the somewhat higher jitter values and particularly the outlier values is the way data are treated before final analysis. In the original descriptions of SFEMG, IPI values more than 4 standard deviations from the mean were discarded; this excludes most spurious signals that do not belong to the motor unit (MU) under study. Obvious interfering signals with intervals within the 4 SD range should also be excluded. The authors should indicate how such signals were identified and handled, particularly since outliers seem to be more frequent than in other studies (Gilchrist et al., 1992).

(4) To obtain the most reliable results from CNE jitter measurements, it is essential to measure only from signals with no visible evidence that they are composite, i.e., with “shoulders”, notches or rounded peaks or with shape variation from discharge to discharge. Superimposition of several sweeps with a rapid sweep is often the only way to assure stable and clean signal shapes with parallel rising phases, a technique we do not see having been used.

(5) The possibility of errors is even greater in stimulation jitter studies, particularly regarding the issue of subliminal stimulation. Each spike to be analyzed must be checked individually by increasing the stimulus intensity to assure supraliminal stimulation. It is not clear if this was done for entire recordings, or if it was done for each measured spike. Because the cut-off limits in this study do not reflect the expected difference between jitter with stimulation and volitional activation, it must be concluded that one or more of these factors has added variability (also commented by the authors).

(6) Various statistical methods were presented, but all of them showed higher values than ours. Thus the difference between the results cannot be explained only by the choice of statistical method.

A recommendation of considerable concern is that different cut-off limits be used “depending on the clinical purpose”. The electromyographer must report the results objectively and distinguish clearly between the results of electrophysiologic testing and conclusions based on other factors. There should only be one cutoff value, but the ultimate interpretation should take into account clinical information.

A strength of this study is the diversity of equipment and experience of the participating laboratories, which may actually represent the way jitter is typically being measured with CN electrodes. This was not the case in the referred studies by Kouyoumdjian and Stålberg, where most of the protocol was highly standardized, only two collaborators were involved and studies were performed in only one laboratory.

Some of these disturbing factors could be considerably reduced by strictly defining the electrode used, filter settings, signal quality criteria and analysis method. Use of the reference values proposed by Kokubun et al. would result in reduced diagnostic sensitivity compared to studies performed with these strict criteria, and we cannot recommend accepting their reference values as final. We suggest the formation of a multicenter study with well-defined criteria and methods before data collection is started.

References

We thank Stålberg et al. for the comments on our study (Kokubun et al., 2012) to establish reference values for SFEMG using concentric needle electrodes (CN-SFEMG) (Stålberg et al., 2013). We also acknowledge the great efforts of Kouyoumdjian and Stålberg (2008a,b, 2011) who completed a series of studies that also investigated reference values of CN-SFEMG. The difference between the two studies is that ours is a multicenter study whereas theirs were performed at one laboratory, and the actual examination was done by one investigator (the first author of their studies). Both methods of investigation have advantages and disadvantages. The merit of the study from a single laboratory is that the equipment is standardized and the technique is uniform. However, the results from such a study may not be directly applicable to examinations in other laboratories. Lack of uniformity is the disadvantage of a multicenter study. However, such differences may be “diluted” and the results may be more rationally applicable to results at other laboratories (Stålberg and Trontelj, 1994); especially when inter-institutional differences are negligible, an issue which is discussed later in this reply.

In order to reduce the differences between institutions and investigators as much as possible, before starting this study, we discussed the many technical aspects of SFEMG examinations extensively and established a rigid protocol, especially on the management of possible pitfalls. Special care was taken to avoid composite potentials, dull potentials that may interfere with the jitter measurements, velocity recovery function (VRF) jitter in the voluntary SFEMG (V-SFEMG), and subthreshold stimulation in the stimulated SFEMG (S-SFEMG) (Kokubun et al., 2012).

We would like to explain our measures against individual pitfalls mentioned by Stålberg et al. (2013). The regulation that the stimulus intensity should be set at least 20% over the threshold for each accepted individual potential was strictly adhered. Due to this restriction, we even felt that S-SFEMG may take a longer time than V-SFEMG because when the stimulus intensity was increased from the threshold intensity for a potential to 20% over that, the potential was frequently lost by the slight displacement of the needle due to a stronger contraction, or was interfered by the newly-recruited potentials, which were not accepted either.

Composite potentials were identified not only by their complicated shape but also by the variation of the shape after minute displacements of the needle or during consecutive discharges. Such potentials were carefully excluded from the analysis when it was judged that the jitter measurement became inaccurate due to the contamination of the disturbing potentials. Furthermore, erroneous inclusion of composite potentials should decrease the jitter value, and not increase it, since the jitters of constituent potentials would be averaged and partially cancelled out.

Although we did not automatically exclude potentials with interpotential interval (IPI) values outside ±4 SD, the contamination of SF potentials from non-target motor unit is a fundamental pitfall in SFEMG examinations, and this was well appreciated by all participants, and we believe that exclusion of such potentials were strictly performed.

The criterion that IPI should be less than 2 ms in V-SFEMG might have been too strict considering the description that the contamination of the VRF jitter may make the jitter value inaccurate when IPI is more than 3–4 ms (Stålberg and Trontelj, 1994). However, a more strict criterion in this regard would just decrease the jitter value if anything.

There are a few factors that cannot be unified in the multicenter study. The selection of needle is one such factor because each institution had favorite needles that have been long employed at that laboratory. The EMG equipment is also a typical one, and the triggering method, peak detection or voltage threshold, is often specific to the equipment and is unchangeable. However, we think that the difference of the triggering methods would not have caused much distortion in the obtained results. Riding signals are a situation for which the voltage threshold method may give incorrect results, but we excluded such potentials when we judged that the distortion was significant: a policy belonging to the exclusion of composite potentials. We think that both methods have advantages and disadvantages. Even the peak detection method may have limitations for potentials that have a steep rising phase but a very gradual falling phase, as was often observed.

It is evident that the increased jitter and blocking in the example presented in our manuscript (Fig. 3 in Kokubun et al., 2012) is not due to above known pitfalls such as composite, riding or dull potentials, or contamination of potentials from other motor units. Plural cut-off values were presented by the following reasons. When the data have a single-layer structure, one can easily judge the “degree” of abnormality using the given mean and SD values. A result with +3 SD would be definitely abnormal. However, the interpretation of an observation with +1.8 SD (> the 95th percentile) might depend on the clinical situation (pre-test probability). The problem is that such a judgment is not simply achieved when the data have double-layer structure, such as in the outlier analysis, since the selection of the cut-off level is done at the first layer. Because +2 SD is widely employed for the judgment of abnormality (Kouyoumdjian and Stålberg, 2008a,b, 2011), we have even newly added the cut-off values corresponding to the +2 SD (Table 1). For EDC-V, for instance, if more than 10% of collected potential pairs have MCD values >50.2 μs, we can infer that this is an abnormality of the degree that would allow 5% false-positive rate (+1.68 SD...