

Motor Unit Action Potential Duration, I: Variability of Manual and Automatic Measurements

I. Rodríguez,* L. Gila,† A. Malanda,* I. G. Gurtubay,† F. Mallor,‡ S. Gómez,‡
J. Rodríguez,* and J. Navallas*

Summary: To analyze the variability in manual measurements of motor unit action potential (MUAP) duration and to evaluate the effectiveness of well-known algorithms for automatic measurement. Two electromyographers carried out three independent duration measurements of a set of 240 MUAPs. The intraexaminer and interexaminer variabilities were analyzed by means of the Gage Reproducibility and Repeatability method. The mean of the three closest manually marked positions was considered the gold standard of the duration markers positions (GSP). The results of four well-known automatic methods for estimating MUAP duration were compared with the GSP. Manual measurements of duration showed a lot of variability, with the combined intraoperator and interoperator variability greater than 30%. The greatest difference between manual positions was 11.2 ms. The mean differences between the GSP and those obtained with the four automatic methods ranged between 0.6 and 8.5 ms. Both manual and automatic measurements of MUAP duration show a high degree of variability. More precise methods are needed to improve the accuracy and reliability of the estimates of this parameter.

Key Words: Motor unit action potential, Duration, Quantitative electromyography, Reproducibility

(*J Clin Neurophysiol* 2006;23: 000–000)

Analysis of the motor unit action potential (MUAP) is an essential aspect of needle electromyography (EMG) studies. The MUAP waveform is quantitatively characterized by several parameters, of which duration is one of the most important. MUAP duration is related to the number of muscle fibers in the motor unit (MU) and to the temporal dispersion of the activation times of the fibers and their conduction velocities (Stalberg et al., 1996). On the other hand, the

duration markers define the boundaries of the MUAP waveform and thereby separate what parts of the recorded signal will be analyzed from what can be considered as baseline (BL) or background activity. Thus, duration is the MUAP parameter that must be determined first.

A number of algorithms for automatic measure of the MUAP duration are available (Stalberg et al., 1986), but they show a low stability, making visual inspection always necessary and manual cursor adjustments is frequently required (Bischoff et al., 1994; Stalberg et al., 1995; Takehara et al., 2004a). However, even the manual placement of duration markers does not guarantee an accurate duration measurement and low degrees of reliability of manual duration readings have been reported (Chu et al., 2003; Nandedkar et al., 1988; Stalberg et al., 1986; Takehara et al., 2004b).

In this study, we present a quantitative estimation of the intraexaminer and interexaminer variability in MUAP duration measurements. A gold standard of the duration markers positions (GSP), obtained from the actual manual measurements, is proposed. Finally, the accuracy of four well-known automatic methods for duration measurement is analyzed by comparing the results of these methods with the GSP.

A summary of the most salient results of the present study was presented at the 42nd Annual Meeting of the Spanish Society of Clinical Neurophysiology (Gila et al., 2004).

METHODS

Subjects, Signal Recording, and MUAP Extraction

The EMG signals were recorded from the tibialis anterior (TA) and first dorsal interosseous (FDI) muscles of eight healthy voluntary subjects, five men and three women, mean age of 44 years with standard deviation of 11.2 and range of 27 to 61. Informed consent was obtained. The recordings were continuous, of 2-second duration, and were taken during slight muscle contraction.

Signals were acquired with a Counterpoint electromyograph (Dantec, Skovlunde, Denmark), using concentric needle electrodes (type DCN37; diameter = 0.46 mm, recording area = 0.07 mm²; Dantec). The filter setting was 3 Hz to 10 kHz, with a sampling rate of 25.6 kHz and 16-bit analog-to-digital conversion. The digitized signals were stored on the hard disk of a PC and analyzed off-line, using a software tool developed by us and implemented with the Signal Processing

*Departamento de Ingeniería Eléctrica y Electrónica, Universidad Pública de Navarra; †Servicio de Neurofisiología Clínica, Hospital Virgen del Camino; and ‡Departamento de Estadística, Universidad Pública de Navarra, Pamplona, Spain.

Address correspondence and reprint requests to Dr. Armando Malanda Trigueros, Universidad Pública de Navarra D.I.E.E., Campus de Arrosadía, 31.006 Pamplona, Spain; e-mail: malanda@unavarra.es

Presented in part at the XVth Congress of the International Society of Electrophysiology and Kinesiology (ISEK 2004), celebrated in Boston in June 2004, and in the 42nd Annual Meeting of the Spanish Society of Clinical Neurophysiology celebrated in Zaragoza in September 2004.

Copyright © 2006 by the American Clinical Neurophysiology Society
ISSN: 0736-0258/06/2306-0001

Toolbox 3.0b of Matlab 4.2c.1 (The Math Works, Natick, Massachusetts).

Each record was revised for manual extraction of MUAPs. Epochs of 50 ms containing discharges of the same MUAP were isolated by using a software specially devised to assess the similarity of the waveforms of different discharges. Well-defined waveforms (avoiding superimpositions, gross BL fluctuations, and distortions of other sources) of at least three discharges were selected for each studied MUAP.

The selected epochs were first aligned regarding to the maximal negative peak of the discharge, which was located 15 ms after the start of the 50-ms analysis window. Then, a correlation maximization algorithm (Campos et al., 2000; Proakis and Manolakis, 1996) was applied with visual supervision of the waveform alignment. The MUAP waveform was finally obtained as the average waveform from the aligned discharges. All of the MUAP waveforms that we included in subsequent studies were well-defined over BL activity and had a “rise-time” of <1 ms (most of them <500 μ s). A total of 240 MUAPs, 152 from 8 TA muscles and 88 from 8 FID muscles, were accepted for analysis.

Manual Measurements of MUAP Duration

The durations of the MUAPs were measured independently by two senior electromyographers. Measurements were made three times, that is, on three occasions, each separated by at least 2 weeks. To achieve their task, the electromyographers were provided with the averaged MUAP and the set of the extracted discharges in raster and superimposed modes (Fig. 1). The time base and sensitivity could be changed by the operator, but the sensitivity for placing duration markers was fixed at 100 μ V/cm.

Statistical Analysis of the Manual Measurements

For TA and FDI MUAPs considered separately, a one-factor analysis of variance (ANOVA) test was performed to assess the differences between the six measurements of duration.

To estimate the intraexaminer and interexaminer variability in the duration measurements, we used the Gage Reproducibility and Repeatability (Gage R & R) method (Montgomery and Runger, 1993a, 1993b). This method,

based on the ANOVA of repeated measurements of a given feature, is currently applied in industrial quality-control studies. It was designed for assessing both the variability in product magnitudes caused by the production process itself (part-to-part variability) and that attributable to the measurement system (gauge). The latter component of variability includes a part attributable to the measurement device, assessed by repeated measurements by the same operator (repeatability or intraoperator variability) and other part attributable to the operator intervention, assessed by comparison of the measurements made by different operators (reproducibility or interoperator variability). In the context of our study, the part-to-part variability is related to the intrinsic variability of MUAP duration present in each set of MUAP extracted from a given muscle (TA or FID). This intrinsic variability of MUAP duration (caused by the variability of the system being measured, not the process of measurement) is due to the differences of size, structure, and position relative to the electrode between the MUs within the muscle. We applied the Gage R & R method for each muscle independently because the durations of the MUAPs of the two studied muscles have different ranges of variation.

Assessment of the Accuracy of the Automatic Measurements

For analyzing the accuracy of any automatic method for MUAP duration measurement, it was necessary to provide a GSP. Given the variability in the manual placement of duration markers and the impossibility of determining which of the manual positions is the best, a method was designed to determine the “most likely” start and end points and to consider it as the GSP. The procedure (Fig. 2) sets the GSP as the mean position of the three (of six) manual marker positions that are closest together.

Four automatic methods for MUAP duration measurement were analyzed: Turku Method 1, Turku Method 2, Uppsala Method 2, and Aalborg (Stalberg et al., 1986). The results were compared with the GSP, and the relative mean differences for each method were compared by using a one-factor ANOVA test. For each method, we calculated the mean of the differences between automatic marker position and the GSP (i.e., the bias of each method) and the standard

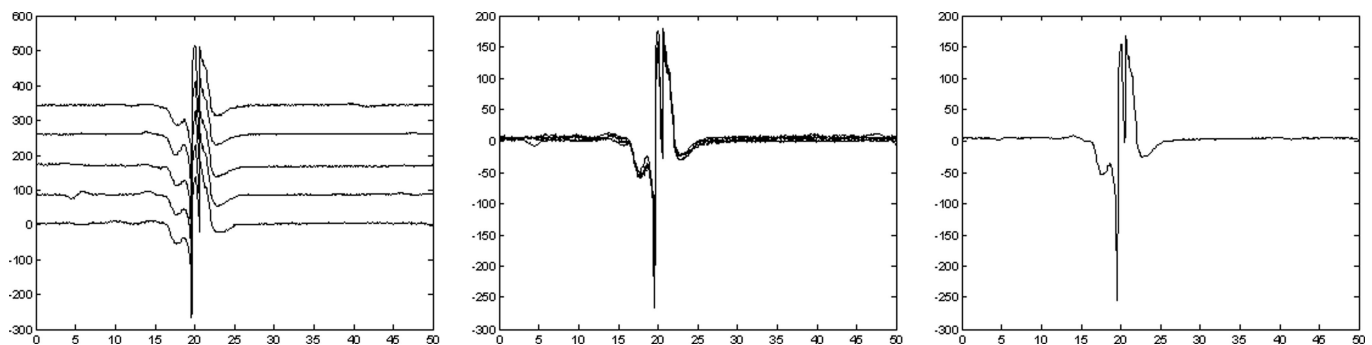


FIGURE 1. Durations of MUAPs, measured independently by two senior electromyographers. Measurements were made three times, on three occasions, each separated by at least 2 weeks. Electromyographers were provided with the averaged MUAP and the set of the extracted discharges in raster and superimposed modes.

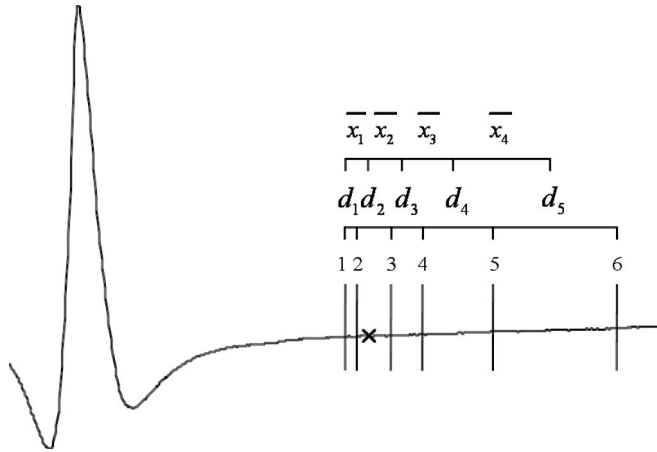


FIGURE 2. Gold standard of the duration marker positions.

deviation of such differences (the precision). We used the estimated mean square error (EMSE) of the differences as criterion for choosing the most efficient method. We used Friedman and Wilcoxon tests to evaluate the significance level of EMSE differences between the methods.

The number of cases in which the absolute difference between the GSP and the automatic marker position was greater than 5 ms was counted for each method. The limit of 5 ms is arbitrary, but such cases can be considered as gross errors. The proportions of gross errors corresponding to each method were compared by means of the χ^2 test.

RESULTS

Manual Measurements

The mean values of the duration, calculated separately for TA and FDI MUAPs, for each of the six manual measurements (three for each electromyographer) are given in Table 1.

Within the six evaluations of start marker position for a given MUAP, the biggest range for a TA MUAP was 6.6 ms and for an FID MUAP was 3.2 ms. Broader ranges were observed for end marker positions: the biggest range for a TA MUAP being 11.2 ms and for an FID MUAP being 10.5 ms. The biggest such ranges were observed in end marker positions for MUAPs with a long, slight slope at the terminal portion of their waveform (Fig. 2). Examples of other diffi-

culties encountered in manual placement of duration markers are given in Fig. 3.

The results of the reproducibility and repeatability analyses of the manual duration measurements, obtained by the Gage R & R method, are given in Table 2. For each muscle, the total variability of the duration measurements is decomposed into that intrinsic to the sampled MUAPs (variability of the proper measured parameter or part-to-part variability) and the variability attributable to the electromyographers. The latter component represented greater than 30%

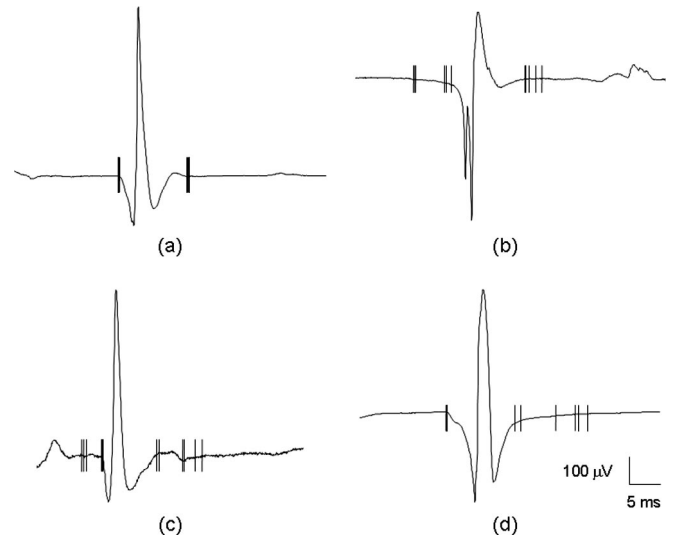


FIGURE 3. Examples of other difficulties encountered in manual placement of duration markers.

TABLE 2. Gage R & R Analysis of Reproducibility and Repeatability of Manual Measurements of MUAP Duration

Source of Variability	Rate of Variability Contribution (%)	
	TA (n = 152)	FID (n = 88)
Duration variability (part-to-part variability)	69.3	67.9
Total Gage R & R	30.7	32.1
Repeatability (intraexaminer variability)	21.4	32.0
Reproducibility (interexaminer variability)	9.3	0.1

TABLE 1. MUAP Duration Values Obtained in Six Manual Measurements Performed by Two Electromyographers

Muscle	Electromyographer 1 Measurements: Mean/SD (ms)			Electromyographer 2 Measurements: Mean/SD (ms)		
	1	2	3	1	2	3
TA n = 152	13.5/3.2	13.4/3.1	14.0/3.2	14.4/3.8	14.6/3.4	14.5/3.4
FDI n = 88	9.3/1.9	9.5/2.7	10.1/3.0	10.0/2.4	10.0/2.9	9.4/1.8

All differences of mean duration between TA and FDI MUAPs were significant. Significant differences were not found among the six measurements performed on the FDI MUAPs.

* $P < 0.001$ (one-factor ANOVA).

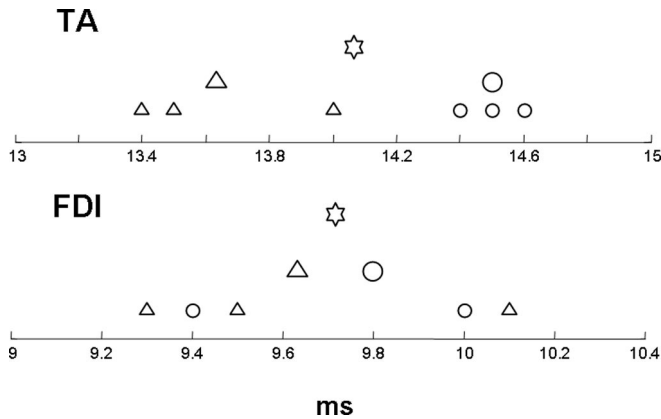


FIGURE 4. Schematic for FID muscle.

of total variability and was mainly due to variability in repeated measurements by the same examiner (repeatability or intraoperator variability). Intraoperator variability was higher in FID muscle, where the estimated interoperator variability (reproducibility) results were much lower.

The Gage R & R results are in accordance with the mean values of Table 1. As can be seen schematically in Fig. 4, for the FID muscle, the distance between the mean of the six manual measurements (star) and the mean of the three manual measurements for each electromyographer (big triangle and big circle) is much smaller than the variability in the manual measurements within the two examiners (small triangles and circles, respectively). Thus, the two examiners obtained similar measurements, so the interoperator variability was very low. For the TA muscle, the means of each examiner are further apart than in the FDI muscle. Thus, interoperator variability was greater and represents approximately one third of the total gauge variability, whereas the other two thirds is attributable to intraoperator variability.

Automatic Measurements

The GSP were compared with the marker positions obtained with the four automatic methods. Differences were calculated for both start and end marker positions (Table 3). The highest mean difference was -8.5 ms for Turku Method 1 end markers in FDI MUAPs. The lowest mean difference was

0.6 ms for the Uppsala Method 2 end marker in TA MUAPs. Standard deviations, between 5.5 and 11.9 ms, were high in all the cases. An absolute difference of more than 5 ms between the GSP and an automatic marker position were found in many cases (Table 4): between 15.0% (for Aalborg end markers) and 49.6% (for Uppsala 2 end markers). To compare the accuracy of each method, we calculated the EMSE of the respective differences between the GSP and the automatic positions (Fig. 5). The Aalborg method had the lowest EMSE and was thus the best method in this respect.

TABLE 4. Rate of Automatic Marker Placements With Differences to the GSP Greater Than 5 Milliseconds

Method	Turku 1	Turku 2	Aalborg	Uppsala 2
Start marker	37.5%	28.3%	17.9%	31.3%
End marker	46.7%	46.3%	15.0%	49.6%

n = 240. *P < 0.01 (χ^2 test).

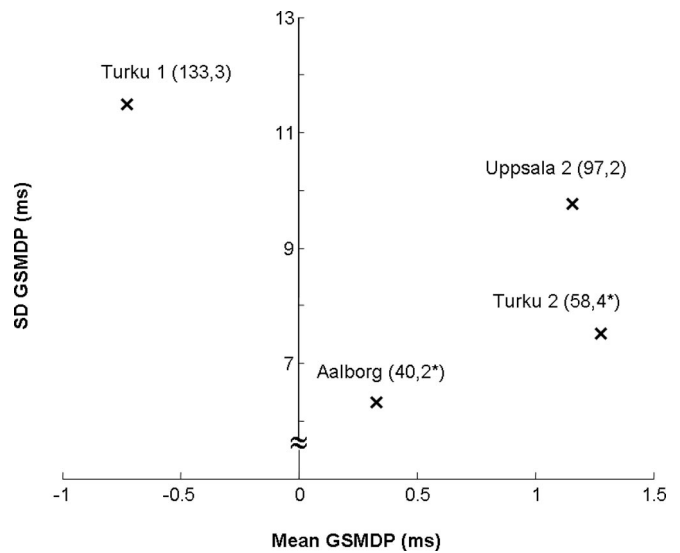


FIGURE 5. Calculation of EMSE of the respective differences between the GSP and the automatic positions.

TABLE 3. Differences Between the GSP and the Marker Positions Obtained by Four Automatic Methods of Duration MUAP Measurement

Method	Turku 1	Turku 2	Aalborg	Uppsala 2
TA (n = 152)				
Start marker	4.2/7.2/3.1–5,4	0.9/5.5/0.04–1,8	2.4/5.6/1.5–3.3	2.9/6.7/1.8–4,0
End marker	-5.6/11.5/-7.5–-3,8	2.4/9.0/0.9–3,8	-1.7/5.9/-2.6–-0,7	0.6/10.2/-1.0–-2,2
FDI (n = 88)				
Start marker	6.9/8.2/5.2–8,7	3.0/6.6/1.7–4,5	2.6/6.7/1.2–4,0	5.3/7.8/3.6–6,9
End marker	-8.5/11.7/-10.9–-6,0	-1.8/7.7/-3.4–-0,1	-1.9/5.8/-3.1–-0,6	-5.0/11.9/-7.6–-2,5

GSP, Gold standard of the duration marker positions.
 Values are mean/SD/95% confidence interval (ms).
 *P < 0.001; one-factor ANOVA.

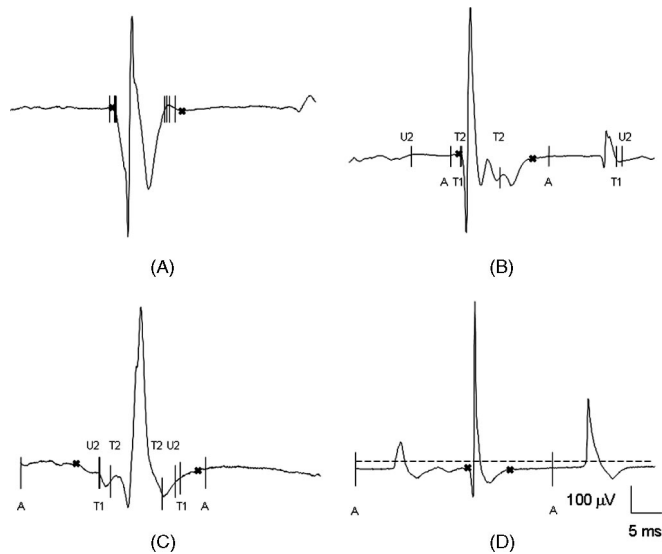


FIGURE 6. Discharges of secondary MUAPs (i.e., MUAPs other than the measured one) in the analysis window.

The main source of error was the presence of discharges of secondary MUAPs (i.e., MUAPs other than the measured one) in the analysis window, which highlight over the BL despite epoch averaging (Fig. 6, B and D). Eventual superimposition of secondary MUAPs on initial and final portions of the MUAP waveform can also interfere with the automatic methods (start markers in Fig. 6C). The presence of a gradual terminal slope was associated with high dispersion in automatic marker placements (Fig. 6C), as was also observed in the manual measurements (Fig. 2).

DISCUSSION

The duration of the MUAP can be defined as the time between the beginning and the end of the MU muscle fiber action potentials as detected by the recording electrode. It has been demonstrated in real records and in simulation studies that the extinction of the action potentials continues for more than 20 ms after the main spike of the MUAP (Dumitru and King, 1999; Dumitru et al., 1999; Lateva and McGill, 1998). In real recordings, a very stable BL and a large number of averaged discharges are needed to observe the very slow return to the BL in the final portion of the MUAP (Dumitru and King, 1999). However, routine recordings almost invariably have slow BL fluctuations and other noise such that it is very difficult to distinguish the full extension of the final portion of the MUAP. Thus, as pointed out by Dumitru et al., there are two meanings of “duration”: the “physiologic MUAP duration” (as it has been defined above) and the “clinical MUAP duration” (Dumitru and King, 1999; Dumitru et al., 1999).

Clinical MUAP duration is defined as the time between the start and end points of the MUAP, when observed at a sensitivity of 100 $\mu\text{V}/\text{cm}$ (Dumitru and King, 1999; Stalberg et al., 1986; Stalberg et al., 1996). At higher gains, duration results longer because more of the slight initial or terminal slopes are visible before merging with the random noise of

background activity. The gain of 100 $\mu\text{V}/\text{cm}$ was arbitrarily chosen to standardize the visual resolution at which duration markers should be manually placed. In this way, duration can be conceived as a morphologic feature, although defined in accord to a given magnitude of visual resolution of the recorded signal.

Similar quantitative definitions are applied to the analysis of morphologic features of the MUAP, such as the automatic counting of turns and phases (Bromberg et al., 1999; Pfeiffer and Kunze, 1992). Algorithms for automatic measurement of duration are based on quantitative criteria of amplitude and slope to look for the limit points between the MUAP waveform and the BL (Stalberg et al., 1986; Stalberg et al., 1996). One might expect such algorithms to be more reliable than manual measurement, but in fact they have several limitations when dealing with real signals.

As can be seen in Fig. 6, a noisy BL, in particular the presence of discharges of MUAPs other than the one under consideration, can induce to gross errors in automatic duration measurement. These errors are frequent with Turku Method 1 and Uppsala Method 2, which work by scanning the signal from the extremes of the analysis window toward the center while looking for samples with specific amplitude or amplitude and slope criteria. The Aalborg method and the Turku Method 2 can overcome these errors because they signal scanning starts in the sample corresponding to the location of the trigger point and run toward the extremes of the analysis window. This is probably the main reason why these methods give better results in terms of bias, precision, and EMSE.

Another problem encountered by automatic algorithms is that the amplitude criteria they need is dependent on the BL definition. The four algorithms used here define the BL as a straight line at the level of the mean voltage of the first and last segments of the analysis window (Rodríguez et al., 2001; Stalberg et al., 1986). Such a BL definition is unlikely to adequately cancel the slow BL fluctuations contained in real recordings. Misplacements of duration markers caused by this effect can also be seen, as in Fig. 6D.

A low degree of agreement between automatic duration estimates has been reported (Bromberg et al., 1999), and so these methods require continuous supervision of the placement of the duration markers. According to various studies, manual correction of automatic placements was required in 20% to 50% of MUAPs (Chu et al., 2003; Bischoff et al., 1994; Stalberg et al., 1995; Takehara et al., 2004a). Such manual adjustments are time-consuming and cause operator fatigue. Moreover, the manual corrections do not guarantee reliable estimates of MUAP duration. As evinced by our work, even with a selection of relatively undistorted MUAP discharges and with optimal display and interface operator facilities, manual measurements of duration showed a lot of variability. The maximum range of manual marker placements for a MUAP was 11.2 ms, and the combined intraoperator (repeatability) and interoperator (reproducibility) variability was greater than 30%. In industrial contexts, where the Gage R & R method applied here is frequently used, degrees

of operator variability greater than 10% are considered as poor and greater than 30% as unacceptable (AIAG, 2002).

Despite the appointed limitations, the placement of duration markers is essential to quantitative EMG, and it cannot be overcome. The duration of the MUAP has itself intrinsic physiologic information about the MU because it is related to the number of muscle fibers in it and the temporal dispersion of their firing. Moreover, the delimitation of the MUAP waveform span is the first task that must be accomplished, since all the MUAP parameters are measured within such limits.

Since establishing MUAP duration is unavoidable, there is great need of a method that provides “acceptable” estimations with a lower degree of variability. By the term “acceptable,” we recognize there is not a unique true value of clinical duration. In the present work, a GSP of the duration markers has been obtained in a probabilistic manner from six manual measurements, being impossible to decide which of them is best. A good automatic method should provide placements close to such standards in the vast majority of MUAPs, working in real time and with relatively noisy signals found in everyday clinical practice.

In conclusion, we observed a high degree of variability in the manual readings of MUAP duration. Currently available automatic methods show low degrees of precision and high variability. Improved automatic methods are needed to provide not only more comfortable work sessions reducing the requirement of manual interventions but also more consistent duration estimates and to overcome the variability and subjectivity of the manual procedure.

ACKNOWLEDGMENT

This work was funded by the Instituto de Salud Carlos III (Project PIO21510) and by the Ministerio de Ciencia y Tecnología (Project BFM 2001-1667-03-01).

REFERENCES

AQ: 8 Automotive Industry Action Group (AIAG). Measurement systems analysis. 3rd edition. 2002.
Bischoff C, Stalberg E, Falck B, Eeg-Olofsson KE. Reference values of motor unit action potentials obtained with multi-MUAP analysis. *Muscle Nerve*. 1994;17:842–851.

Bromberg MB, Smith AG, Bauerle J. A comparison of two commercial quantitative electromyographic algorithms with manual analysis. *Muscle Nerve*. 1999;22:1244–1248.
Campos C, Malanda A, Gila L, et al. Quantification of jiggle in real electromyographic signals. *Muscle Nerve*. 2000;23:1022–1034.
Chu J, Takehara I, Li TC, Schwartz I. Skill and selection bias has least influence on motor unit action potential firing rate/frequency. *Electromyogr Clin Neurophysiol*. 2003;43:387–392.
Dumitru D, King JC. Motor unit action potential duration and muscle length. *Muscle Nerve*. 1999;22:1188–1195.
Dumitru D, King JC, Zwartz MJ. Determinants of motor unit action potential duration. *Clin Neurophysiol*. 1999;110:1876–1882.
Gila L, Gurtubay IG, Rodríguez I, et al. Variabilidad en las medidas manuales y automáticas de la duración del potencial de acción de unidad motora. *Rev Neurol*. 2004;39:1083–1084.
Lateva Z, McGill K. The physiological origin of the slow afterwave in muscle action potentials. *Electroencephalogr Clin Neurophysiol*. 1998;109:462–469.
Montgomery DC, Runger GC. Gage capability and designed experiments, I: basic methods. *Qual Eng*. 1993a;6:115–135.
Montgomery DC, Runger GC. Gage capability and designed experiments, II: experimental design models and variance component estimation. *Qual Eng*. 1993b;6:289–305.
Nandedkar S, Barkhaus P, Sanders D, Stalberg E. Analysis of the amplitude and area of the concentric needle EMG motor unit action potentials. *Electroencephalogr Clin Neurophysiol*. 1988;69:561–567.
Pfeiffer G, Kunze K. Turn and phase counts of individual motor unit potentials: correlation and reliability. *Electroencephalogr Clin Neurophysiol*. 1992;85:161–165.
Proakis JG, Manolakis DG. Digital signal processing: principles, algorithms and applications. 3rd ed. London: Prentice-Hall International, 1996: pp.118–133.
Rodríguez I, Gila L, Malanda A, et al. Baseline removal from EMG recordings. Proceedings of the 23rd Annual International Conference of the IEEE in Engineering in Medicine and Biology Society. *Istanbul*. 2001;Vol2:pp.1034–1037.
Stalberg E, Andreassen S, Falck B, et al. Quantitative analysis of individual motor unit potentials: a proposition for standardized terminology and criteria for measurement. *J Clin Neurophysiol*. 1986;3:313–348.
Stalberg E, Falck B, Sonoo M, Astrom M. Multi-MUP EMG analysis: a two year experience with a quantitative method in daily routine. *Electroencephalogr Clin Neurophysiol*. 1995;97:145–154.
Stalberg E, Nandedkar S, Sanders DB, Falck B. Quantitative motor unit potential analysis. *J Clin Neurophysiol*. 1996;13:401–422.
Takehara I, Chu J, Schwartz I, Aye HH. Motor unit action potential (MUAP) parameters affected by editing duration cursors. *Electromyogr Clin Neurophysiol*. 2004a;44:265–269.
Takehara I, Chu J, Li TC, Schwartz I. Reliability of quantitative motor unit action potential parameters. *Muscle Nerve*. 2004b;30:111–113.