Journal of Neuroscience Methods 192 (2010) 138–145

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

High inter-rater reliability in analyzing results of decomposition-based quantitative electromyography in subjects with or without neuromuscular disorder

Shaun G. Boe[a,b,*], Nathan M. Antonowicz[b,1], Vanessa W. Leung[b,2], Susan M. Shea[b,1], Toby C. Zimmerman[b,3], Timothy J. Doherty[c]

[a] School of Physiotherapy, Dalhousie University, Rm. 418, 4th Floor Forrest Building, 5869 University Avenue, Halifax, Nova Scotia, Canada B3H 3J5
[b] School of Physical Therapy, The University of Western Ontario, London, Ontario, Canada
[c] Department of Clinical Neurological Sciences, Schulich School of Medicine and Dentistry, The University of Western Ontario, London, Ontario, Canada

ARTICLE INFO

Article history:
Received 2 February 2010
Received in revised form 16 July 2010
Accepted 21 July 2010

Keywords:
Electrodiagnostic
Motor unit
Multi-center
Neuromuscular disease
Neuropathy

ABSTRACT

Decomposition-based quantitative electromyography (DQEMG) comprises a group of methods used to obtain information related to the health of the neuromuscular system. Although primarily objective, aspects of the data analysis protocol include operator decisions that may impact its reliability and reduce the applicability of the technique among multiple users. Thus, the objective of this study was to establish the inter-rater reliability of the protocol used for DQEMG analysis among five raters. Seventy data files previously obtained using DQEMG from healthy control subjects and patients with disorders of the neuromuscular system were analyzed by four novice and one experienced rater. Values obtained from this analysis were then evaluated for reliability within the novice raters and in contrast to the results of the experienced rater to examine the influence of the level of rater experience on the results obtained. The majority of the parameters associated with the number of motor unit potentials and their physiological characteristics were found to be reliable among all raters, with moderate-high ICC values observed for both the biceps brachii and first dorsal interosseous muscles. The data suggest that the level of rater experience does not greatly influence the results obtained and that the analysis can be reliably performed by a rater who is given suitable instruction. These findings are important particularly given the potential use of DQEMG as an outcome measure in multi-center studies.

1. Introduction

Decomposition-based quantitative electromyography (DQEMG) has been designed to provide information pertaining to the physiological characteristics and numbers of motor units (MUs) within a given muscle or muscle group (Doherty and Stashuk, 2003). Clinically this information is valuable in that it provides insight into the changes occurring at the level of the MU in response to disorders of the neuromuscular system. Evidence to support this has come from several studies, including the observation of decreased MU number estimates and increased MU size and complexity in patients with amyotrophic lateral sclerosis (ALS) (Boe et al., 2007) and Charcot-Marie Tooth (CMT) disease (Shy et al., 2007), in addition to demonstrating age-related MU remodeling and an associated decrease in the estimated numbers of MUs in old and very old men (McNeil et al., 2005).

To provide value as a clinical tool, it is important to ensure the data obtained using DQEMG are reliable from one test to the next both within and across raters. When studies are performed by experienced raters, DQEMG has been found to be reliable, with high

Abbreviations: AAR, area-to-amplitude ratio; ALS, amyotrophic lateral sclerosis; BB, biceps brachii; CMT-X, Charcot-Marie Tooth disease, Type X; COV, coefficient of variation; DQEMG, decomposition-based quantitative electromyography; EMG, electromyographic; FDI, first dorsal interosseous; FR, firing rate; ICC, intra-class correlation coefficient; IDI, interdischarge interval; MU, motor unit; MUP, motor unit potential; NpAmp, negative-peak amplitude; S-MUP, surface-detected motor unit potential.

* Corresponding author at: School of Physiotherapy, Dalhousie University, Rm. 418, 4th Floor Forrest Building, 5869 University Avenue, Halifax, Nova Scotia, Canada B3H 3J5. Tel.: +1 902 494 6360; fax: +1 902 494 1941.
E-mail address: s.boe@dal.ca (S.G. Boe).
[1] Address: School of Physical Therapy, *The* University of Western Ontario, Elborn College, Rm 1588, London, Ontario, Canada N6G 1H1.
[2] Address: E.W. Bickle Centre for Complex Continuing Care, Toronto Rehabilitation Institute, 130 Dunn Ave., Toronto, Ontario, Canada M6K 2R7.
[3] Address: VHA Rehab Solutions, 633 Colborne St., 2nd Floor, London, Ontario, Canada N6B 2V3.

0165-0270/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2010.07.024

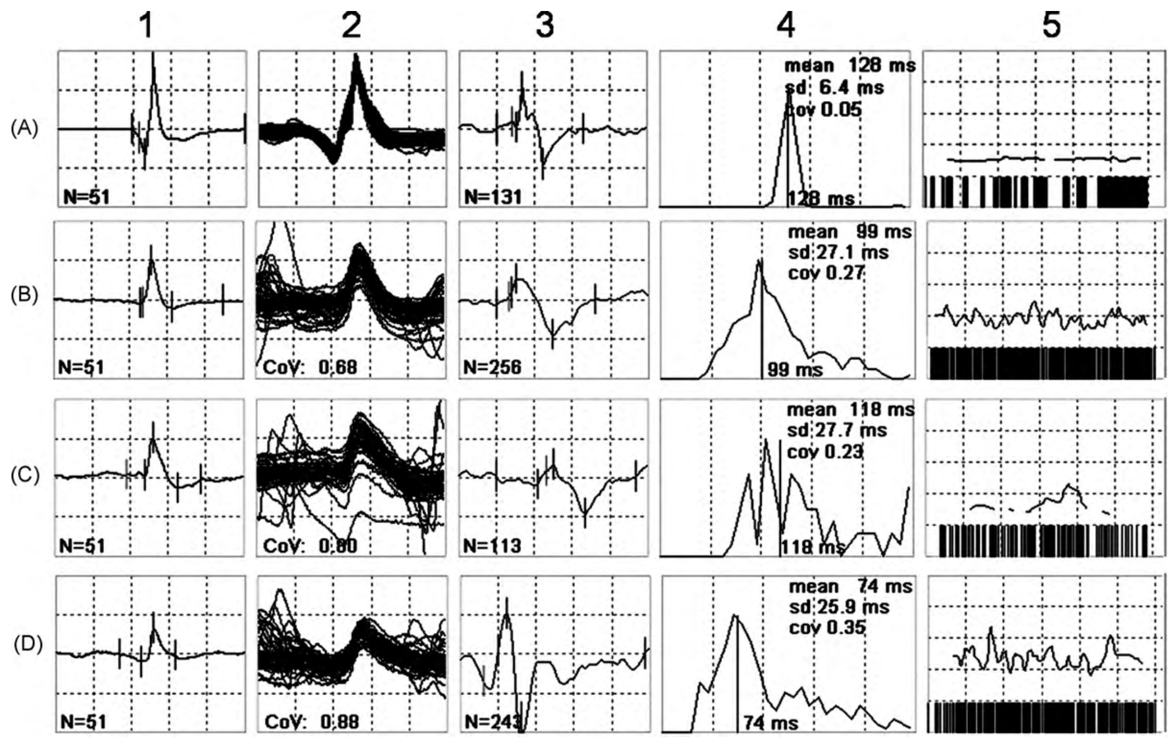

Fig. 1. Decomposition summary. Sample of four MUP trains (vertical, labeled A–D) that represent the decomposition of a typical needle-detected EMG signal and subsequent analysis of the needle- and surface-detected signal. Columns from left to right represent (1) prototypical needle-detected MUP; (2) individual MUPs of each MUP train superimposed in a shimmer plot; (3) surface-detected MUP and number of contributing discharges; (4) IDI histogram; and (5) firing rate vs. time plot representing MUP train discharge times and instantaneous firing rate plots. Raters are required to review each MUP train based on objective and subjective criteria to determine whether to include or exclude a particular MUP train. Briefly, this process includes: (i) does the MUP train include a minimum of 51 contributing MUPs (column 1), (ii) is the COV value associated with the IDI < 0.30 (column 4), (iii) does the firing rate vs. time plot display a consistent firing rate pattern (column 5), and (iv) does the IDI histogram have a normally distributed main peak (column 4).

intra- and inter-rater reliability observed for data collection and analysis (Boe et al., 2006, 2009; Calder et al., 2008).

Although DQEMG has been designed to utilize objective criteria during data collection and analysis, some aspects of the analysis are not amenable to quantification or classification and thus are at the discretion of the operator. These include the decisions to include or exclude needle-detected MU potential (MUP) trains and surface-detected MUPs (S-MUPs) that do not meet the objective inclusion criteria of the DQEMG software. Additionally, raters may be required to re-position inaccurately placed markers used to calculate MUP duration and amplitude measures. Although these subjective components do not impact reliability in studies performed by experienced raters (Boe et al., 2006, 2009; Calder et al., 2008), this may not be the case when analysis is performed by less experienced raters, which may occur in larger, multi-center studies. Thus, the purpose of this study was to examine the potential impact of rater experience on the intra-rater reliability of the procedures used for the analysis of DQEMG data obtained from subjects with and without neuromuscular disorder.

2. Methods

2.1. Subjects

DQEMG data from 28 healthy control subjects (27 ± 5 years), nine patients with ALS (52 ± 12 years) and three patients with CMT, Type X (CMT-X, 46 ± 9 years) who had previously undergone examination using DQEMG were analyzed in the current study. Biceps brachii (BB) data were obtained from 28 healthy control subjects and seven patients with ALS. First dorsal interosseous (FDI) data were obtained from 24 healthy control subjects, eight patients with ALS and three patients with CMT-X. All subjects had previously provided informed consent and the University of Western Ontario ethics review board approved the study.

2.2. DQEMG data collection

The DQEMG method, associated algorithms and data collection protocols for the FDI and BB muscles have been described in detail elsewhere (Boe et al., 2004, 2007; Doherty and Stashuk, 2003; Stashuk, 1999). Utilizing a series of pattern recognition algorithms in addition to spike-triggered averaging, DQEMG is able to breakdown both a needle and surface-detected EMG signal, acquired simultaneously during a voluntary muscle contraction, into their individual needle and surface-detected MUPs (Stashuk, 1999, 2001). Briefly, DQEMG decomposes the composite needle-detected EMG signal into its constituent MUP trains using shape and temporal information related to the individual MUP discharges in addition to MU firing time statistics. Using these needle-detected MUPs as triggers for spike-triggered averaging, a sample of MUPs, detected via surface electrodes (S-MUPs), are obtained. These S-MUPs are representative of the sizes of the MUs in the underlying muscle of interest (Stalberg, 1980; Stalberg and Fawcett, 1982). Parameters associated with the needle and surface-detected MUPs provide information regarding MU size, complexity and discharge rate. Although not presented here, if a representative sample of these S-MUPs is attained (≥20), a mean S-MUP size can be determined and divided into a corresponding size-related parameter of a maximal M wave to produce a MU number estimate (Boe et al., 2004).

2.3. Raters

The novice raters were four second-year students enrolled in a Masters level clinical program at the University of Western Ontario.