

Reproducibility, and sensitivity to motor unit loss in ALS, of a novel MUNE method: MScanFit MUNE

Jacobsen AB ^a, Bostock H ^b, Fuglsang-Frederiksen A ^a, Duez L ^a, Beniczky S ^a, Møller AT ^c, Blicher JU ^c, Tankisi H ^a

^a *Department of Clinical Neurophysiology, Aarhus University Hospital, Aarhus, Denmark*

^b *Institute of Neurology, Queen Square House, London, United Kingdom*

^c *Department of Neurology, Aarhus University Hospital, Aarhus, Denmark*

Corresponding author:

Hatice Tankisi

Department of Neurophysiology

Aarhus University Hospital

Nørrebrogade 44, 8000 Aarhus C, Denmark

Tel.: +45 78 46 24 31

Fax: +45 78 46 31 40

E-mail address: hatitank@rm.dk

Acknowledgements

Conflict of interest: HB receives royalties from UCL for sales of his Qtrac software used in this study. The other authors have no potential conflicts of interest. All authors have approved the final article.

Funding: This study was financially supported by the Lundbeck Foundation, Knud og Edith Eriksens Mindefond, Søster og Verner Lipperts Fond, Fonden til Lægevidenskabens Fremme and Aage og Johanne Louis Hansens Fond.

We would like to thank the patients and healthy controls for their participation in the study.

Highlights:

- MScan is a novel promising MUNE method with good reproducibility and clinical potential.
- MScan and MPS are better at discriminating between ALS patients and healthy controls than MUNIX.
- Only MScan manages to distinguish patients with less certainty of ALS from healthy controls.

Abstract

Objective: To examine inter- and intra-rater reproducibility and sensitivity to motor unit loss of a novel motor unit number estimation (MUNE) method, MScanFit MUNE (MScan), compared to two traditional MUNE methods; Multiple point stimulation MUNE (MPS) and Motor Unit Number Index (MUNIX).

Methods: Twenty-two ALS patients and 20 sex- and age-matched healthy controls were included. MPS, MUNIX, and MScan were performed twice each by two blinded physicians. Reproducibility of MUNE values was assessed by coefficient of variation (CV) and intra class correlation coefficient (ICC). Ability to detect motor unit loss was assessed by ROC curves and area under the curve (AUC). The times taken for each of the methods were recorded.

Results: MScan was more reproducible than MPS and MUNIX both between and within operators. The mean CV for MScan (12.3%) was significantly lower than for MPS (24.7%) or MUNIX (21.5%). All methods had $ICC > 0.94$. MScan and MUNIX were significantly quicker to perform than MPS (6.3m vs. 13.2m). MScan (AUC= 0.930) and MPS (AUC=0.899) were significantly better at discriminating between patients and healthy controls than MUNIX (AUC=0.831).

Conclusions: MScan was more consistent than MPS or MUNIX and better at distinguishing ALS patients from healthy subjects.

Significance: MScan may improve detection and assessment of motor unit loss.

Keywords: Motor unit number estimation; MScanFit; Multiple point stimulation; MUNIX; ALS; reproducibility

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a fatal neuromuscular disease characterized by a decline in upper and lower motor neurons and a corresponding denervation of muscle fibers. This leads to progressive paralysis with a fatal outcome within an average of 3 years after diagnosis of the disease. (Couratier, et al. 2016)

Overall, there is no method that can measure the exact number of motor units. Nerve conduction studies and electromyography cannot provide accurate information about the number of motor unit loss. Because of collateral sprouting, where healthy axons take over the muscle territory of axons that have been lost, compound muscle action potential (CMAP) amplitude does not decrease until 50% or more of motor units are lost and does not correlate well with the degree of denervation (Daube. 2006, McComas, et al. 1971). Similarly, in muscles with denervation, measurement of muscle strength does not accurately reflect the number of remaining motor units. Thus a reliable method to monitor loss of motor units in ALS is needed to improve diagnostics and as an outcome measure in drug trials seeking a treatment of the disease.

For this reason, several motor unit number estimation (MUNE) methods have been developed as an attempt to quantify the number of motor units. The first MUNE method was described in 1971 (McComas, et al. 1971), and since then MUNE has been of interest to measure lower motor neuron loss. In studies already published, MUNE has been found to be useful in diagnosis, follow-up and prognosis of many clinical disorders including ALS and also in the evaluation of drug trials (Bromberg and Brownell. 2008, Daube, et al. 2000, Gooch, et al. 2014, Doherty and Brown. 1993).

Most MUNE methods have been based on estimating the size of an average surface-recorded motor unit potential (sMUP) and dividing that value into the maximal CMAP. These include incremental stimulation (McComas, et al. 1971), multiple point stimulation (Bromberg. 1993, Daube. 1995, Porter, et al. 2008) and spike-triggered averaging (Bromberg. 1993). Other methods have used statistical techniques based on the response of the motor unit to a stimulus either with Poisson statistics (Lomen-Hoerth and Slawnych. 2003, Henderson, et al. 2007) or Bayesian statistics (Ridall, et al. 2006, Nandedkar, et al. 2004). Motor Unit Number Index (MUNIX) is a more recently introduced method which uses the surface interference patterns recorded during voluntary contractions to infer the average size of surface-recorded motor unit potentials (Nandedkar, et al. 2010, de Carvalho, et al. 2008, Doherty and Brown. 1993, Uncini and Kuwabara. 2012). However,

the mentioned MUNE methods suffer from several limitations including the presence of subjectivity in the estimation process, the failure to obtain a representative sample of units or the fact that it is taking too long to perform or analyse.

To overcome these limitations, a new MUNE method was recently developed, 'MScanFit MUNE', (MScan) (Bostock. 2016). This method meets the above-mentioned criticisms by fitting a model to a detailed stimulus-response curve or 'CMAP scan' (Maathuis, et al. 2012, Visser and Blok. 2009), taking into account the threshold variability of all the units, and avoiding subjectivity. The method has been found to be accurate (mean absolute error 7%) for simulated data (Bostock. 2016). In a previous study, MScan was tested on patients with neurofibromatosis type 2, where it revealed denervation and reinnervation in peripheral nerves suggesting that it may be used to quantify and monitor disease progression (Farschtschi, et al. 2016).

The overall objective of this project was to examine the inter-and intra-rater reproducibility and sensitivity to motor unit loss of MScan in the assessment of ALS patients and healthy controls, and to compare it with two other more traditional MUNE methods, i.e. Multiple point stimulation (MPS) MUNE and MUNIX.

2. Material and Methods

2.1 Subjects

Twenty-two patients with ALS or progressive muscular atrophy (PMA) (7 females and 15 males – age range 44-83 years, mean 65.3) and 20 healthy sex- and age-matched healthy subjects (7 females and 13 males – age range 44-76 years, mean 65.6) were included prospectively in the study between December 2015 and June 2016. Fourteen of the patients had just received the diagnosis before participating in the project and eight were diagnosed earlier. Mean duration of symptoms was 34 months (range 3-120 months). Patients were categorized according to the revised El Escorial criteria (Brooks, et al. 2000) which were used as a clinical reference standard (Bossuyt, et al. 2015 STARD 2015). Five of the patients were categorized as definite, five as probable, seven as laboratory supported ALS, one as clinically possible and four with PMA with only lower motor neuron signs. Patient demographics are summarized in Table 1. Patients were recruited from the departments of Neurology and Neurophysiology at Aarhus University Hospital. All subjects gave their written consent prior to examinations and the protocol was approved by the Regional Scientific Ethical Committee and the Danish Data Protection Agency.

Subjects with a history of carpal tunnel syndrome, polyneuropathy or diseases that could induce polyneuropathy were excluded. Moreover, patients with too severe atrophy of the abductor pollicis brevis muscle to allow for reliable recordings were also excluded. Two patients were excluded because of this.

Patients (n=22)	
Age (years)	65.3±9.98 (44-83)
Sex	Females (n=7), Males (n=15)
Duration of symptoms (months)	34.0±33.93 (3-120)
ALSFRS-R	35.3±10.98 (9-48)
Region of onset	Bulbar (n=3) Upper limb (n=11) Lower limb (n=8)

Table 1. Patient demographics. Values expressed as mean ± standard deviation. Values in parenthesis indicate range. n=number of patients.

2.2 Clinical Examination

Detailed neurological examination including evaluation of force by manual muscle testing and grading according to the MRC scale and deep tendon reflexes were done on all subjects. ALS patients were scored by use of revised ALS Functional Rating scale (ALSFRS-R) (Cedarbaum, et al. 1999) (Table 1).

2.3 Electrophysiological Evaluation

All examinations were carried out at Department of Neurophysiology in Aarhus University. In all subjects, an electrophysiological evaluation was performed, which included a motor and sensory nerve conduction study and three MUNE methods; MPS, MUNIX and MScan. All three MUNE methods were repeated twice each by two blinded physicians (ABJ and HT) in order to evaluate inter- and intra-rater reproducibility of the methods. One physician had two years of experience with all MUNE methods while the other physician had 2 months of training. Examinations were done either on the same day or within a week. For the examinations done on the same day, all electrodes were removed and new electrodes were applied for the retest and the physicians took turns. The order of the MUNE tests remained the same for each participant: 1) MPS, 2) MUNIX and 3) MScan. The final calculation of MUNE values for each of the methods were done after all of the examinations of the subject were finished, and cursors for supramaximal CMAPs and sMUPs

were placed so that the physicians were blinded both to their own and the other physician's MUNE values. The final calculations of MUNE values were made by one of the physicians (ABJ). The exact times taken for each of the three MUNE examinations were recorded.

The subject's hand was cleansed with skin prepping gel and alcohol. The skin temperature was held between 32 and 36°C by a warming lamp. Recordings were made on the thenar muscles on the dominant hand in healthy controls and on the least affected side on patients. The active electrode was placed over the abductor pollicis brevis muscle (APB) and the reference electrode was placed on the interphalangeal joint of the first digit. A ground electrode was placed on the dorsum of the hand. Apart from MScan, all electrophysiological examinations were done using an EMG-equipment Keypoint version 2.11 (Dantec, Skovlunde, Denmark).

2.4 Nerve conduction studies (NCS)

Routine motor and sensory nerve conduction studies were performed on the median nerve to exclude a carpal tunnel syndrome. Motor NCS was performed by delivering supramaximal stimulation at the wrist and the elbow and recording at APB, and sensory NCS was performed by delivering supramaximal stimulation at the wrist and recording at the second and third digits.

2.5 Multiple point stimulation MUNE (MPS)

The recordings were made with a stimulus duration of 0.1 ms and a gain of 0.1 mV/division when determining SMUP's and 2 mV/division when determining CMAP. The median nerve CMAP was recorded by delivering supramaximal percutaneous stimulation to the wrist by a handheld bipolar stimulator. The recording electrode on the APB muscle was moved several times before accepting the CMAP with the highest amplitude. The CMAP amplitude was measured from baseline to negative peak. The SMUPs were recorded by gradually increasing and decreasing the stimulus intensity just enough to produce an all-or-none response. After recording one SMUP with a minimum amplitude of $\geq 25 \mu\text{V}$ at one position, the stimulation site was moved along the median nerve in 1 cm steps and the process repeated 10 times, until 10 different SMUPs had been recorded. In patients with less than 10 SMUPs as many as possible were recorded. To calculate the MUNE, the supramaximal CMAP amplitude was divided by the average amplitude of the ~10 SMUPs.

2.6 Motor Unit Number Index (MUNIX)

The recordings were performed with a 300 ms window and with a stimulus frequency of 1 Hz and gain of 0.5 mV/division. A new supramaximal percutaneous stimulation was delivered to the wrist by a bipolar stimulator. The CMAP amplitude was measured from baseline to negative peak by

keeping the recording electrodes at the same position. Next, the electromyographic surface interference pattern (SIP) was recorded during ten voluntary contractions at increasing isometric force levels during abduction of digit 1. The examiner provided manual resistance to abduction of digit 1 to maintain near isometric conditions as the participant applied an increasing force. During the last recordings, there was a pause of 30 seconds between the contractions in order to avoid fatigue (Fuglsang-Frederiksen. 2006). Only SIPs with an area above 20 mV/ms were accepted. The recordings were exported to an excel file designed to process the raw data to obtain MUNIX values (Neuwirth, et al. 2015).

2.7 MScanFit MUNE (MScan)

The recordings were made using the TRONDNF recording protocol, which is part of the Qtrac software. (Institute of Neurology, University College, London, distributed by Digitimer Ltd at www.digitimer.com.) Other equipment used with the set up was a DS5 bipolar stimulator (Digitimer Ltd), a HumBug 50 Hz noise eliminator and a D440 amplifier (Digitimer Ltd). To eliminate noise and artefacts due to voluntary movement, the fingers were taped together and the most relaxing position of the arm was found. A muting box was developed that provided the patients with audio feed-back if movement occurred during the recordings. The subjects were asked to remain silent and avoid movement during the examination. The stimulus was gradually increased manually until supramaximal stimulation was reached. After the responses to 20 supramaximal stimuli had been recorded, the stimulus intensity was automatically decreased with a frequency of 2 Hz in 0.2% steps, from supramaximal stimulation until there was no longer a motor response, when a further set of 20 CMAPs was recorded. The 20 recordings before and 20 recordings after the main scan were used to assess the variability of the responses and the baseline noise. The data were plotted as a `CMAP scan` or detailed stimulus response curve with the amplitude of the motor response on the y-axis and stimulus intensity on the x-axis. In healthy subjects this created a smooth S-shaped curve. In ALS patients with a loss of motor units, the curve was shaped as a stair with steps as a sign of collateral reinnervation or loss of motor units (Figure 1). The number of motor units was estimated offline, using the MScanFit component of the QtracP analysis program. The program generated a preliminary model based on the variance and slope of scan points, which was used to simulate the recorded CMAP scan. The model was then refined by making serial adjustments to minimize the difference between the simulated and recorded CMAP scans (Bostock. 2016). The off-line analysis is fast and is done automatically including the modelling without requirements of making any parameter adjustments. The analysis took less than five min in total.

2.8 Statistical analysis

The inter- and intra-rater reproducibility of MPS, MUNIX and MScan was tested by intraclass correlation coefficient (ICC) calculated for absolute agreement and coefficient of variation (CV) as defined by the ratio of the standard deviation to the mean of the MUNE values. The Wilcoxon signed rank test was used to test if the three methods differed from each other. To evaluate the relation between the three MUNE methods and CMAP amplitude, regression analysis was used. A Mann-Whitney test was used to compare the absolute MUNE values of each of the MUNE methods. Moreover, the MUNE methods and CMAP amplitude were evaluated with a receiver operating characteristic (ROC) analysis, by determining the area under the ROC curve (AUC). MedCalc provides two methods for estimating P values, both of which make use of the fact that the same subjects were used for each test. Pa refers to the method of DeLong et al (DeLong, et al. 1988) and Pb to the method of Hanley and MacNeil (Hanley and McNeil. 1983). These methods were also used to evaluate the differences between the three MUNE methods. Comparisons between the methods were done by using a 2-tailed test and comparisons between each of the MUNE methods and the corresponding CMAP amplitude by using a 1-tailed test. When making single comparisons, P values < 0.05 were considered significant. In tables where either single or multiple comparisons can be made, and a Bonferroni correction may be appropriate, we provide exact P values. QtracP was used for the statistical analyses and for generating the figures.

3. Results

3.1 Test of MPS, MUNIX and MScan between groups

A total of 168 recordings were performed for MPS and MScan each corresponding to two recordings by two different observers for each subject. For MUNIX, only 164 recordings were performed, because in two patients the examinations could not be performed due to too severe muscle weakness. The median MUNE values of MPS, MScan and MUNIX in patients and controls are summarized in table 2.

	MPS	MSCAN	MUNIX
	(Median ± range from min to max)		
Patients	30.1 (1-143.3)	32.4(1-123.5)	103.1(2-259.8)
Controls	132.9 (79.8-251.8)	111.1(71.75-166.8)	207.5(98-423.3)

P-values	1.9×10^{-6}	1.2×10^{-6}	0.00026
----------	----------------------	----------------------	---------

Table 2. Comparison of MUNE-values in MPS, MScan and MUNIX in patients and healthy subjects. MUNE values are expressed as medians across 4 recordings on each subject (two per observer).

The MUNE values were significantly lower in patients compared with controls in both MPS, MScan and MUNIX. There was also a significant difference in MUNE values between all three methods both for patients and for healthy subjects. ($p < 0.05$).

Figure 2 shows the distribution of MUNE values for each of the methods in healthy subjects and patients divided into grade 1 (clinically definite ALS), grade 2-3 (clinically probable and lab-supported) and grade 4-5 (clinically possible or PMA) based on the revised El Escorial criteria. In MScan, there was a significant difference between all grades of ALS compared to healthy controls. In MPS and MUNIX, there was a significant difference between healthy controls and grade 1 and 2-3 respectively. However, no significant difference was found between grade 4-5 and healthy subjects. The ALSFRS-R scores were significantly different for the El Escorial subgroups. Grade 1 had significantly lower ALSFRS-R scores compared to grade 2-3 ($p = 0.024$) and grade 4-5 ($p = 0.0037$). Similarly, grade 2-3 had significantly lower scores than grade 4-5 ($p = 0.0041$).

3.2 Reproducibility

Table 3 summarizes the coefficients of variation (CV) of MUNE values of MPS, MScan and MUNIX expressed as means and their corresponding p-values when comparing the methods. The CV values are based on two recordings for each observer for each subject.

MUNE Coefficients of variation (%)						
	(Mean \pm SE)			(P for Wilcoxon signed rank test)		
	MPS	MSCAN	MUNIX	MPS v MSCAN	MSCAN v MUNIX	MPS v MUNIX
Intra-rater	14.4 \pm 1.5	8.6 \pm 1.6	14.7 \pm 1.3	0.00042	0.00056	0.35
Inter-rater	20.3 \pm 1.5	9.7 \pm 1.3	16.7 \pm 2.3	2.9×10^{-6}	0.00067	0.0068
Patients	28.5 \pm 2.7	14.7 \pm 2.7	25.0 \pm 4.6	0.00088	0.0030	0.081
Controls	20.5 \pm 1.6	9.6 \pm 1.3	17.9 \pm 2.2	0.00071	0.0024	0.14
Combined	24.7 \pm 1.7	12.3 \pm 1.6	21.5 \pm 2.6	3.4×10^{-6}	3.8×10^{-5}	0.031

Table 3. Comparison between coefficients of variation across two repeats by two different observers. *P* values <0.05 are shown in bold. Combined refers to inter- and intra-rater measurements on all subjects.

MScan had significantly lower CV values than MPS and MUNIX both with regards to the inter- and intra-rater recordings and in patients and healthy controls separately and combined. MUNIX had a significantly lower CV than MPS in inter-rater recordings and inter- and intra-rater recordings combined. However, there was no significant difference between CV values of MPS and MUNIX in the intra-rater recordings and in patients and healthy controls separately.

Figure 3 A illustrates the CV values of MPS, MScan and MUNIX in patients and controls separately and combined (patients and controls). Figure 3B illustrates the CV values of MPS, MScan and MUNIX in inter-and intra-rater recordings separately and combined based on both patients and controls.

Table 4 summarizes the intraclass correlation coefficients (ICC) and the corresponding confidence intervals of MPS, MScan and MUNIX for inter- and intra-rater measurements. The ICC values were higher for MScan than MPS and MUNIX both between and within operators.

	Intra class correlation coefficient (ICC)		
	MPS	MScan	MUNIX
Inter-rater	0.970(0.906-0.987)	0.993(0.987-0.996)	0.944(0.898-0.970)
Intra-rater (observer 1)	0.979 (0.961-0.989)	0.983 (0.968-0.991)	0.970 (0.945-0.984)
Intra-rater (observer 2)	0.980 (0.963-0.989)	0.985(0.972-0.992)	0.974 (0.952-0.986)

Table 4. Summary of intra class correlation coefficients (ICC) across two repeats by two different observers. Confidence intervals are shown in parentheses.

3.3 Correlations between MPS, MUNIX, MScan and CMAP

In patients and healthy subjects combined, MScan and MPS values were similar and a strong linear correlation was found between them ($R^2=0.962$). MUNIX values were higher and less correlated with either MScan ($R^2=0.832$) or MPS ($R^2=0.891$). (figure 4). All correlations were statistically significant. In ALS patients only, there was also a strong and significant linear correlation between

all methods: MPS and MUNIX ($R^2=0.87$), MPS and MScan ($R^2=0.97$) and MScan and MUNIX ($R^2=0.79$). In healthy subjects, weaker correlations were found: MPS and MUNIX ($R^2=0.62$), MScan and MUNIX ($R^2=0.43$) and MPS and MScan ($R^2=0.12$). The correlations were all significant except for MPS and MScan ($p=0.14$).

In patients and healthy subjects combined, there was a strong quadratic correlation between CMAP amplitude and MUNIX ($R^2=0.944$) (figure 5), indicating that 94% of the variation in MUNIX values was accounted for by the variation in CMAP amplitude. A weaker correlation was found for CMAP amplitude and MPS ($R^2=0.801$) and CMAP amplitude and MScan ($R^2=0.724$). All correlations were significant, also for patients and controls separately.

3.4 Sensitivity and specificity of MPS, MUNIX, MScan and CMAP

In figure 6, the potential of MPS, MScan, MUNIX and CMAP amplitude to distinguish ALS patients from healthy subjects is illustrated in a ROC curve. (MedCalc Statistical Software version 16.8 (MedCalc Software bvba, Ostend, Belgium; <https://www.medcalc.org>; 2016). The best cut-off MUNE value for distinguishing between healthy controls and patients was 96.5 for MPS, 75.5 for MScan and 146.5 for MUNIX. This yielded the following sensitivities and specificities respectively for each of the methods: MPS: 90% and 80.7%, MScan: 92.5% and 80.7% and MUNIX: 82.5% and 69.5%. MUNE methods were compared with their corresponding CMAP amplitude and evaluated by the area under the curve (AUC) as a measure of how well they were able to distinguish between patients and healthy subjects. MUNIX and CMAP had the same AUC (0.831). Both MPS and MScan had an AUC (0.899 and 0.903 respectively) that were significantly higher than their CMAP amplitude (0.841 and 0.845 respectively). When comparing MPS and MScan there was no significant difference ($P_a=0.11$ and $P_b=0.14$). There were significant differences between MPS and MUNIX ($P_a=0.029$ and $P_b=0.0084$) and between MScan and MUNIX ($P_a=0.0005$ and $P_b=0.0001$).

3.5 Difference in time to perform MPS, MUNIX and MScan

Table 5 and figure 7 show the difference in recording times between MPS, MScan and MUNIX both in patients, in controls, and combined. In all groups the recording time of MPS was significantly longer than of MScan and MUNIX. There was no significant difference between MScan and MUNIX. When comparing the recording times between the two observers, the experienced observer did MPS recordings significantly faster ($p<0.001$), but there was no

significant difference between the two observers' recording times for MScan ($p=0.82$) and MUNIX ($p=0.26$).

MUNE Recording time (min)						
	(Mean \pm SE)			(P for Wilcoxon signed rank test)		
	MPS	MSCAN	MUNIX	MPS v MSCAN	MSCAN v MUNIX	MPS v MUNIX
Patients	12.73 \pm 0.44	6.08 \pm 0.28	5.92 \pm 0.23	2.1$\times 10^{-5}$	0.40	4.8$\times 10^{-5}$
Controls	13.79 \pm 0.66	6.48 \pm 0.29	6.76 \pm 0.18	6.5$\times 10^{-5}$	0.14	7.1$\times 10^{-5}$
Combined	13.24 \pm 0.39	6.27 \pm 0.20	6.34 \pm 0.16	1.1$\times 10^{-8}$	0.27	2.7$\times 10^{-8}$

Table 5. Comparison of recording times (min) between MPS, MScan and MUNIX in patients and healthy subjects expressed as means across 4 recordings on each subject (two per observer). P values <0.05 are shown in bold.

4. Discussion

This study was the first one to test the reproducibility and clinical utility of MScan in patients with ALS. We have shown that the inter-and intra-rater reproducibility of MScan is significantly better than that of MPS and MUNIX. MScan was capable of differentiating ALS patients from healthy controls in an earlier stage of the disease than MPS and MUNIX. There is no gold standard in MUNE methods. However, the very close agreement between MScan and MPS, which are two quite different methods, gives confidence in the accuracy of their estimates.

4.1 Reproducibility

All methods had excellent reproducibility with ICC values of >0.94 . However, the ICC values are all high in part because of the wide variation in motor unit loss between the ALS patients. A more relevant measure of reproducibility for drug studies, for example, is CV, since a halving of CV implies that only one quarter of the number of patients would be required to detect a difference in treatments. The CVs were significantly lower in MScan both for inter-and intra-rater recordings respectively (9.7% and 8.6%) compared with MPS (20.3% and 14.4%) and MUNIX (16.7% and 14.7%). This study was the first one to test MScan on ALS patients. MScan is a novel MUNE method, and reproducibility of this method has never been tested before. Several previously published papers have evaluated reproducibility of MUNIX. A recent study measured the intra-rater

reproducibility in APB in ALS patients and found a CV of 19.4% and an ICC of 0.74. (Boekestein, et al. 2012). Another study measured the inter- and intra-rater reproducibility in the ADM muscle of ALS patients and healthy controls and found inter-rater CV values of 23.7% in ALS patients and 17.5% in healthy subjects and intra-rater CV values of 24.0% in patients and 15.3% in healthy subjects. (Ahn, et al. 2010) Thus, similar to our study, there was less variability in healthy controls than in ALS patients. Another research group found similar inter-rater reproducibility in ALS patients (ICC>0.89) and discovered that the variability was slightly larger in APB (CV=20.5%) than in ADM (CV=18.4%) (Nandedkar, et al. 2011). This was also seen in a multicentre MUNIX study comparing reproducibility of ADM and APB in healthy subjects. This research group found a higher inter-rater reproducibility in ADM (ICC=0.68) than in APB (ICC=0.43) but a lower intra-rater reproducibility in ADM (ICC=0.67) than in APB (ICC=0.75) (Neuwirth, et al. 2011). In our study, the participants were examined in the APB muscle where we found better reproducibility than some of the studies mentioned above. This may be because we have paid big attention for placing the recording electrode on APB muscle and moved the electrode several times to find the highest supramaximal CMAP amplitude which is a major determinant of MUNIX values. Reproducibility of MPS has also been assessed in previous studies in ALS patients and healthy controls (Doherty and Brown. 1993, Felice. 1995). One study measured intra-rater reproducibility in ALS patients and healthy subjects combined and found similar results to ours (CV=12.0%). (Lomen-Hoerth and Olney. 2000). This was performed in ADM. Another study found a lower intra-rater CV of 10.4% in healthy subjects using the adapted MPS methods in thenar muscles (Wang and Delwaide. 1995). This, however, was only carried out in 10 subjects.

4.2 Correlations

We found that overall, MScan provided similar MUNE values to MPS. In patients, we found that MPS and MScan were strongly correlated and both were less correlated with MUNIX. In healthy subjects, we found weaker correlation between the methods and no correlation between MPS and MScan. Numerous previous studies have evaluated correlation of MUNE methods with varying results. A recent study compared MPS and MUNIX on patients with inflammatory demyelinating polyneuropathy and found a linear correlation in patients but no correlations in healthy subjects (Paramanathan, et al. 2016). This was also the case in another study comparing high density MUNE with MUNIX in ALS patients and healthy subjects (Boekestein, et al. 2012). Another study compared incremental stimulation MUNE with MUNIX in healthy subjects and found no correlations (Furtula, et al. 2013). MPS has also been compared to the statistical method in which

one study found a correlation in ALS patients and healthy controls (Lomen-Hoerth and Olney. 2000), and another study did not show a correlation in healthy controls (Blok, et al. 2010). The different results across the studies could partly be explained by the fact that the MUNE methods are based on different technical and physiological aspects, and the spread of the values vary between the methods.

We found that CMAP amplitude and MUNIX were strongly correlated and the correlations between CMAP amplitude and MPS and MScan respectively were weaker. The strong correlation between CMAP amplitude and MUNIX has been reported in numerous previous studies, both in healthy subjects and in ALS patients (Nandedkar, et al. 2010, Neuwirth, et al. 2011, Furtula, et al. 2013, Ahn, et al. 2010). Thus our study confirms that MUNIX values depend too much on CMAP amplitude to provide useful estimates of motor unit numbers.

4.3 Sensitivity and specificity as a test for motor unit loss

All three methods had the ability to distinguish patients from healthy subjects. Only MScan managed to distinguish patients with possible ALS and PMA from healthy subjects suggesting that the discriminative potential of MScan in ALS patients with less disease certainty is better than the two other methods. We found a higher sensitivity for MScan (92.5%) compared to MPS (90%) and MUNIX (82.5%) and a higher specificity for MScan and MPS (80.7%) compared to MUNIX (69.5%). Another important issue is whether MUNE methods are better than CMAP amplitude in discriminating patients from healthy subjects. To our knowledge, the sensitivity of MUNE methods has not been compared to CMAP amplitude in previous studies. In our study, ROC curves showed that MPS (AUC=0.899) and MScan (AUC=0.930) were significantly better at discriminating between patients and healthy subjects than MUNIX and CMAP amplitude. (AUC=0.831). However, MUNIX did not differ from CMAP amplitude in distinguishing patients from healthy subjects. MScan had the largest AUC suggesting the highest discriminating value, although there was no significant difference between MPS and MScan. ROC curves of different MUNE methods have been assessed in previous studies. In one study, MUNIX was performed on ALS patients and healthy controls and a slightly higher AUC of 0.897 was found compared with the AUC we found. (Furtula, et al. 2013). This was measured in ADM and only 13 patients were included. An explanation for the higher AUC could be that the recordings were done on the most affected side. Moreover, it is not specified how many of the patients had onset in their upper limb. Another study comparing MPS and MUNIX on patients with inflammatory demyelinating polyneuropathy did not find difference in AUC between the two methods (Paramanathan, et al. 2016).

4.4 Recording times

The recordings times were significantly longer for MPS than MUNIX and MScan, and for MPS it was dependent on the experience of the observer. The experienced observer was significantly faster than the less experienced observer. This was not the case for MScan and MUNIX. The recording times have not been evaluated in previous studies.

4.5 Limitations

There were limitations in this study which are important to address. The sample size was small and the recruitment of patients was limited to patients who had sufficient physical and mental strength to participate. However, this was not different from other similar studies with ALS patients. We did the recordings in APB muscle and due to the split hand phenomenon (Kuwabara, et al. 1999, Wilbourn. 2000), this muscle can be too atrophic in ALS to allow for MUNE recordings. Because of this, we had to exclude two patients from the study. However, MPS and MScan turned out to be more applicable to patients with severe muscle weakness and atrophy. In two of the patients included in this study MUNIX, which requires voluntary contraction, could not be performed due to severe muscle weakness whereas MPS and MScan could be applied. The limitation of severe weakness and atrophy could be avoided by choosing ADM. On the other hand, APB is preferable because of higher sensitivity than ADM due to the split hand phenomenon. We examined only APB muscle on the least affected side and we also included patients with bulbar or lower limb onset who had APB MUNE values comparable to healthy subjects. If we had chosen to examine more muscles on the most affected side, the sensitivity could probably have been improved. This, however, would have been too time consuming in a reproducibility study such as this one, where the examinations for each participant lasted for approximately four hours in total, because two observers did the recordings twice each. If we had done the recordings on the most affected side, it would have been necessary to exclude several of the patients since their most affected side was too atrophic and weak to allow for MUNE examinations. Another limitation is that the ALS patients were compared to healthy subjects, and we did not address the question of differentiating ALS from ALS-mimicking diseases.

4.6 Conclusion

In this study, we found that MScan is a method that takes into account most of the limitations that have been found in other MUNE methods in previous published studies. We found excellent reproducibility both with regards to inter- and intra-rater measurements, and it was quick and easy

to perform and did not depend on the experience of the observer. It had the ability to differentiate ALS patients from healthy controls in an earlier stage of the disease than MPS and MUNIX. Our study suggests that MScan is a method that has the potential to be implemented in the clinic, where the ability to detect motor unit loss may be helpful in ALS and other neuromuscular disorders. The good reproducibility may enable more efficient assessment of treatments designed to slow or prevent motor neurone degeneration. More and larger studies are needed in order to evaluate whether our findings can be reproduced and to explore the potential of MScan as a follow-up tool.

Reference list

Ahn SW, Kim SH, Kim JE, Kim SM, Kim SH, Park KS, et al. Reproducibility of the motor unit number index (MUNIX) in normal controls and amyotrophic lateral sclerosis patients. *Muscle Nerve* 2010;42:808-813.

Blok JH, van Dijk JP, Drenthen J, Maathuis EM, Stegeman DF. Size does matter: the influence of motor unit potential size on statistical motor unit number estimates in healthy subjects. *Clin.Neurophysiol.* 2010;121:1772-1780.

Boekestein WA, Schelhaas HJ, van Putten MJ, Stegeman DF, Zwarts MJ, van Dijk JP. Motor unit number index (MUNIX) versus motor unit number estimation (MUNE): a direct comparison in a longitudinal study of ALS patients. *Clin.Neurophysiol.* 2012;123:1644-1649.

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Radiology* 2015;277:826-832.

Bostock H. Estimating motor unit numbers from a CMAP scan. *Muscle Nerve* 2016;53:889-896.

Bromberg MB. Motor unit estimation: Reproducibility of the spike-triggered averaging technique in normal and ALS subjects. *Muscle Nerve* 1993;16:466-471.

Bromberg MB, Brownell AA. Motor Unit Number Estimation in the Assessment of Performance and Function in Motor Neuron Disease. *Phys.Med.Rehabil.Clin.North Am.* 2008;19:509-532.

Brooks BR, Miller RG, Swash M, Munsat TL, World Federation of Neurology Research Group on Motor Neuron Diseases. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler.Other Motor Neuron.Disord.* 2000;1:293-299.

Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, et al. The ALSFRS-R: A revised ALS functional rating scale that incorporates assessments of respiratory function. *J.Neurol.Sci.* 1999;169:13-21.

Couratier P, Corcia P, Lautrette G, Nicol M, Preux PM, Marin B. Epidemiology of amyotrophic lateral sclerosis: A review of literature. *Rev.Neurol.(Paris)* 2016;172:37-45.

Daube JR. Motor unit number estimates - From A to Z. *J.Neurol.Sci.* 2006;242:23-35.

Daube JR. Estimating the number of motor units in a muscle. *J.Clin.Neurophysiol.* 1995;12:585-594.

Daube JR, Gooch C, Shefner J, Olney R, Felice K, Bromberg M. Motor unit number estimation (MUNE) with nerve conduction studies. *Suppl.Clin.Neurophysiol.* 2000;53:112-115.

de Carvalho M, Dengler R, Eisen A, England JD, Kaji R, Kimura J, et al. Electrodiagnostic criteria for diagnosis of ALS. *Clin.Neurophysiol.* 2008;119:497-503.

DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-845.

Doherty TJ, Brown WF. The estimated numbers and relative sizes of thenar motor units as selected by multiple point stimulation in young and older adults. *Muscle Nerve* 1993;16:355-366.

Farschtschi S, Gelderblom M, Buschbaum S, Bostock H, Grafe P, Mautner VF. Muscle action potential scans and ultrasound imaging in neurofibromatosis type 2. *Muscle Nerve* 2016. doi: 10.1002/mus.25256.

Felice KJ. Thenar motor unit number estimates using the multiple point stimulation technique: reproducibility studies in ALS patients and normal subjects. *Muscle Nerve* 1995;18:1412-1416.

Fuglsang-Frederiksen A. The role of different EMG methods in evaluating myopathy. *Clin.Neurophysiol.* 2006;117:1173-1189.

Furtula J, Johnsen B, Christensen PB, Pugdahl K, Bisgaard C, Christensen M-, et al. MUNIX and incremental stimulation MUNE in ALS patients and control subjects. *Clin.Neurophysiol.* 2013;124:610-618.

Gooch CL, Doherty TJ, Chan KM, Bromberg MB, Lewis RA, Stashuk DW, et al. Motor unit number estimation: A technology and literature review. *Muscle Nerve* 2014;50:884-893.

Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839-843.

Henderson RD, Ridall PG, Hutchinson NM, Pettitt AN, McCombe PA. Bayesian statistical MUNE method. *Muscle Nerve* 2007;36:206-213.

Kuwabara S, Mizobuchi K, Ogawara K, Hattori T. Dissociated small hand muscle involvement in amyotrophic lateral sclerosis detected by motor unit number estimates. *Muscle Nerve* 1999;22:870-873.

Lomen-Hoerth C, Olney RK. Comparison of multiple point and statistical motor unit number estimation. *Muscle Nerve* 2000;23:1525-1533.

Lomen-Hoerth C, Slawnych MP. Statistical motor unit number estimation: From theory to practice. *Muscle Nerve* 2003;28:263-272.

Maathuis EM, Henderson RD, Drenthen J, Hutchinson NM, Daube JR, Blok JH, et al. Optimal stimulation settings for CMAP scan registrations. *J.Brachial Plex Peripher.Nerve Inj.* 2012;7:4.

McComas AJ, Fawcett PR, Campbell MJ, Sica RE. Electrophysiological estimation of the number of motor units within a human muscle. *J Neurol Neurosurg Psychiatry* 1971;34:121-131.

Nandedkar SD, Barkhaus PE, Stalberg EV. Reproducibility of MUNIX in patients with amyotrophic lateral sclerosis. *Muscle Nerve* 2011;44:919-922.

Nandedkar SD, Barkhaus PE, Stålberg EV. Motor unit number index (MUNIX): Principle, method, and findings in healthy subjects and in patients with motor neuron disease. *Muscle Nerve* 2010;42:798-807.

Nandedkar SD, Nandedkar DS, Barkhaus PE, Stalberg EV. Motor unit number index (MUNIX). *IEEE Trans.Biomed.Eng.* 2004;51:2209-2211.

Neuwirth C, Barkhaus PE, Burkhardt C, Castro J, Czell D, de Carvalho M, et al. Tracking motor neuron loss in a set of six muscles in amyotrophic lateral sclerosis using the Motor Unit Number Index (MUNIX): a 15-month longitudinal multicentre trial. *J.Neurol.Neurosurg.Psychiatry.* 2015;86:1172-1179.

Neuwirth C, Nandedkar S, Stålberg E, Barkhaus PE, Carvalho MD, Furtula J, et al. Motor Unit Number Index (MUNIX): A novel neurophysiological marker for neuromuscular disorders; test-retest reliability in healthy volunteers. *Clin.Neurophysiol.* 2011;122:1867-1872.

Paramanathan S, Tankisi H, Andersen H, Fuglsang-Frederiksen A. Axonal loss in patients with inflammatory demyelinating polyneuropathy as determined by motor unit number estimation and MUNIX. *Clin.Neurophysiol.* 2016;127:898-904.

Porter CL, Alvarez A, Jones KE, Ming Chan K. Test--retest reliability of a modified multiple point stimulation technique for motor unit number estimation. *Clin.Neurophysiol.* 2008;119:2287-2290.

Ridall PG, Pettitt AN, Henderson RD, McCombe PA. Motor unit number estimation--a Bayesian approach. *Biometrics* 2006;62:1235-1250.

Uncini A, Kuwabara S. Electrodiagnostic criteria for Guillain-Barrè syndrome: A critical revision and the need for an update. *Clin.Neurophysiol.* 2012;123:1487-1495.

Visser GH, Blok JH. The CMAP scan. *Suppl.Clin.Neurophysiol.* 2009;60:65-77.

Wang FC, Delwaide PJ. Number and relative size of thenar motor units estimated by an adapted multiple point stimulation method. *Muscle Nerve* 1995;18:969-979.

Wilbourn AJ. The "split hand syndrome". *Muscle Nerve* 2000;23:138.

Figure 1. Examples of MScan recordings from A: a patient with ALS and B: a healthy subject

Figure 2. Boxplot showing the distribution of MUNE values of MPS, MScan and MUNIX across four repetitions in normal controls (NC) and in patients divided into grade 1, grade 2-3 and grade 4-5. Horizontal line represents the medians and the ends of the boxes the first and third interquartile. The whiskers indicate the 10% and 90% interquartile. Comparisons between controls and grade of ALS are made. $P < 0.05$ are expressed with *. NS= Not significant. PMA= progressive muscular atrophy. Controls are shown in white and patients in striped.

Figure 3. Box plot showing the distribution of coefficients of variations (CV) in A) patients, controls and combined (patients and controls) across two repeats by two different observers and B) inter-rater, intra-rater and combined (inter-and intra-rater) recordings based on all patients and controls. MPS is shown in white, MScan in black and MUNIX in striped.

Figure 4. Linear regression between the three methods (black line). The dashed line indicates the line of equality. A. MPS vs MScan. B MUNIX vs MScan. C. MPS vs MUNIX. Patients are shown with closed circles and controls with open circles. R^2 is based on patients and healthy subjects combined.

Figure 5. Quadratic regression (black line) between compound muscle action potential (CMAP) amplitude and: A. MPS. B. MScan and C. MUNIX. Patients are shown with closed circles and controls with open circles. R^2 is based on patients and healthy subjects combined.

Figure 6. ROC curves comparing area under the curve (AUC) of CMAP amplitude with A. MPS, B. MScan and C. MUNIX. P_a is obtained from the method of DeLong et al (1988) and P_b is obtained from the method of Hanley and MacNeil (1983).

Figure 7. Box plot showing the distribution of recording times (min) in MPS, MScan and MUNIX. MPS is shown in white, MScan in black and MUNIX in striped.