BJA Open

BJA Open, 11 (C): 100293 (2024)

doi: 10.1016/j.bjao.2024.100293
Original Research Article

ORIGINAL RESEARCH ARTICLE

Area under the curve and amplitude of the compound motor action potential are clinically interchangeable quantitative measures of neuromuscular block: a method comparison study

Richard H. Epstein[1,*], Réka Nemes[2], Johnathan R. Renew[3] and Sorin J. Brull[3]

[1]Department of Anesthesiology, Perioperative Medicine & Pain Management, University of Miami, Miller School of Medicine, Miami, FL, USA, [2]Department of Anesthesiology and Intensive Care, University of Debrecen, Debrecen, Hungary and [3]Department of Anesthesiology and Perioperative Medicine, Mayo Clinic College of Medicine and Science, Jacksonville, FL, USA

*Corresponding author. E-mail: repstein@med.miami.edu

Abstract

Background: Current guidelines recommend quantitative neuromuscular block monitoring during neuromuscular blocking agent administration. Monitors using surface electromyography (EMG) determine compound motor action potential (cMAP) amplitude or area under the curve (AUC). Rigorous evaluation of the interchangeability of these methods is lacking but necessary for clinical and research assurance that EMG interpretations of the depth of neuromuscular block are not affected by the methodology.

Methods: Digitised EMG waveforms were studied from 48 patients given rocuronium during two published studies. The EMG amplitudes and AUCs were calculated pairwise from all cMAPs classified as valid by visual inspection. Ratios of the first twitch (T_1) to the control T_1 before administration of rocuronium (T_1c) and train-of-four ratios (TOFRs) were compared using repeated measures Bland–Altman analysis.

Results: Among the 2419 paired T_1/T_1c differences where the average T_1/T_1c was $\leq$0.2, eight (0.33%) were outside prespecified clinical limits of agreement (−0.148 to 0.164). Among the 1781 paired TOFR differences where the average TOFR was $\geq$0.8, 70 (3.93%) were outside the prespecified clinical limits of agreement ((−0.109 to 0.134). Among all 7286 T_1/T_1c paired differences, the mean bias was 0.32 (95% confidence interval 0.202–0.043), and among all 5559 paired TOFR differences, the mean bias was 0.011 (95% confidence interval 0.0050–0.017). Among paired T_1/T_1c and TOFR differences, Lin's concordance correlation coefficients were 0.98 and 0.995, respectively. Repeatability coefficients for T_1/T_1c and TOFR were <0.08, with no differences between methods.

Conclusions: Quantitative assessment neuromuscular block depth is clinically interchangeable when calculated using cMAP amplitude or the AUC.

Keywords: action potentials; electromyography; monitoring; neuromuscular block; signal processing

The American Society of Anesthesiologists and the European Society of Anaesthesiology and Intensive Care guidelines strongly recommend quantitative neuromuscular monitoring when administering neuromuscular blocking agents.[1,2] There are subtle but important differences in the interpretation of physiological neuromuscular variables obtained when using different monitoring modalities, such as electromyography (EMG) or acceleromyography (AMG), the two technologies used most frequently. For example, the count of twitches reported by EMG monitors after train-of-four (TOF) stimulation is affected by the control first twitch amplitude (T_1c) in the TOF sequence before administering any neuromuscular blocking drugs.[3] In contrast to EMG, AMG signals often exhibit reverse fade, a phenomenon in which the control fourth response (T_4c)

Received: 24 April 2024; **Accepted:** 22 May 2024

after TOF stimulation is larger than the control first response, T_1c (i.e. TOFR >100%), a physiologically unlikely phenomenon.[4] When reverse fade is present, a proportional increase (normalisation) is required to the 90% threshold TOF ratio (TOFR) to demonstrate adequate minimal recovery from neuromuscular block (e.g. if the control TOFR is 110%, the recovery threshold is a TOFR ≥99%, not 90% because 90% of 110% is 99%, not 90%).[5] Thus, anaesthesia practitioners should not accept insentiently that all quantitative neuromuscular block monitors provide clinically interchangeable information but rather need to understand the underlying fundamentals to appropriately and effectively use the information provided by the various devices.[6]

In this study, using the digitised EMG waveforms obtained from a single manufacturer's device, we explored two current signal processing methodologies used by various other EMG devices to determine the quantitative extent of neuromuscular block: the amplitude and the area under the curve (AUC) of the compound motor action potential (cMAP; Fig 1).[7] Although studies from >25 yr ago in cats[7] and a small number of patients suggested that the two methods are clinically interchangeable,[8,9] an evaluation using rigorous statistical methods has not been performed to determine if they provide comparable information. This issue is both scientifically and clinically relevant because there are EMG neuromuscular monitors that use one or the other method (Table 1). Multiple clinical research studies have been conducted with devices using either method,[10–18] but a significant clinical care limitation may exist if the results based on one method were not generalisable to devices using the other method. Furthermore, it is unknown if any adjustment in clinical interpretation is needed based on the different EMG measurement methodologies.

We hypothesised that the pairwise assessments of the T_1/T_1c and the TOFR would fall within the prespecified, clinically relevant limits of agreement based on a Bland–Altman paired comparison of measurements between two EMG methodologies with multiple measurements per subject.[19–21]

Methods

The institutional review board of the University of Miami determined on 16 March 2023 that this analysis of deidentified data from two prior randomised clinical trials where informed consent was obtained does not meet the regulatory definition of human subjects research. Those two studies had been approved by the Mayo Clinic Institutional Review Board (#20–00629) and the University of Debrecen Ethical Board (No. OGYÉI2690/2018). We followed the checklist recommended for studies comparing measurement methods (Supplementary Table S1).[22]

Data sources

Digitised EMG data were provided by two co-authors (JRR, RN) from previous studies of the TetraGraph™ EMG neuromuscular block monitor (Senzime AB, Uppsala, Sweden). In all studies, the negative stimulating electrode was distal, the skin was prepped with alcohol to remove skin oil and wiped with

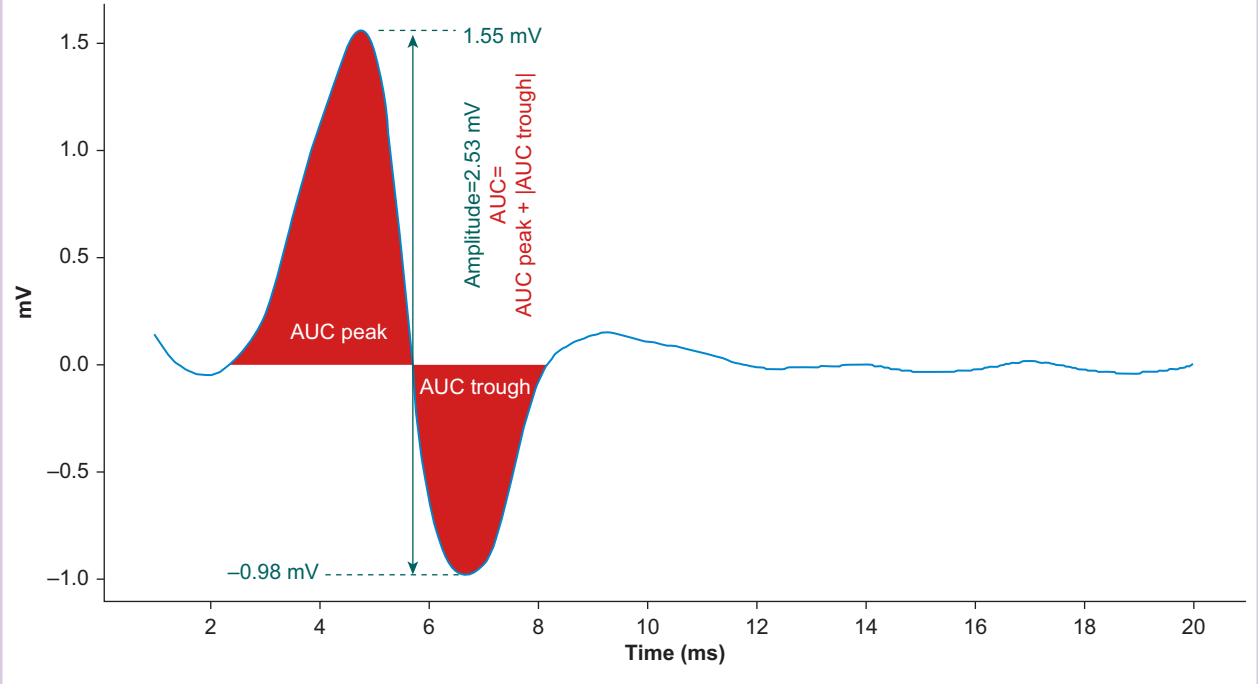

Fig 1. Graphical representation of the algorithm used to determine the amplitude and the adaptive area under the curve (AUC) for a representative electromyographic (EMG) compound motor action potential (cMAP) from the adductor pollicis muscle after ulnar nerve stimulation at time=0 ms. The amplitude (2.53 mV) is the difference between the peak voltage (1.55 mV) and the trough voltage (−0.98 mV). The AUC is the sum of the areas of the red-shaded portions of the curve after rectification of the voltages (i.e. the absolute value of each voltage is taken, and then the curve is numerically integrated). The integration interval started where the cMAP curve first crossed the 0.0 mV baseline before the peak and ended where the curve first crossed the baseline after the trough.

Table 1 Method of calculation of twitch strength by clinically available EMG devices. AUC, area under the curve; cMAP, compound motor action potential; ms, millisecond; mV, millivolt.

Device	Manufacturer	Method	Reference electrode	Measurement details
E-NMT	GE Healthcare Chicago, IL (USA)	AUC	Yes	AUC integration between 3 and 15 ms after stimulus
NMT Pod	Nihon Kohden Tokyo (Japan)	Amplitude	Yes	Amplitude threshold =0.7 mV for a valid cMAP
StimPod™	Xavant Technology Silverton (South Africa)	AUC	No	Limits of integration determined adaptively from the waveform
TetraGraph™	Senzime AB Uppsala (Sweden)	Amplitude	No	Amplitude threshold =0.4 mV for a valid cMAP
TwitchView™	Blink Device Company Seattle, WA (USA)	AUC	Yes	The company declined to provide any details

gauze to remove dead skin and lower the electrode contact impedance, adequate time was allowed for the electrodes to cure, and supramaximal currents were applied, as recommended for pharmacodynamic studies of neuromuscular blocking agents.[23] The first study included 29 adult patients undergoing robotic and laparoscopic procedures under deep neuromuscular block with rocuronium, antagonism of neuromuscular block with sugammadex, and, in most cases, volatile anaesthetics for maintenance of anaesthesia.[24] The second study comprised 19 adult patients undergoing elective surgery where anaesthesia was maintained with a target-controlled propofol infusion.[15] Rocuronium was administered under a protocol designed to allow spontaneous recovery to a TOFR >0.9, with neostigmine administered if this endpoint was not reached. In both studies,[15,24] active forced-air warming devices were used to maintain core body temperature >36°C; however, maintenance of hand temperatures was not targeted.

We emphasise that the TetraGraph was only used to provide the cMAP waveforms for the comparative analysis, and all assessments of cMAP amplitudes and AUCs were computed from those waveforms. We did not use that device's smoothing or interpolation algorithms and performed these independently using standard signal processing methods (e.g. locally estimated scatterplot smoothing [LOESS] peak detection). Consideration of the threshold amplitude and other signal processing analyses applied by the TetraGraph to ascertain that a valid cMAP is present is irrelevant to our comparison of amplitude and AUC analytical methods for several reasons. First, in our previous study, we determined by visual inspection whether a candidate cMAP was valid or not and in the current study only made comparisons between the amplitude and AUC methods for valid waveforms. There were many instances where the amplitude was <0.4 mV (the TetraGraph threshold for reporting the presence of a twitch) and the waveform was clearly valid.[3] Clinical observation indicates that the count of twitches reported by the TetraGraph often undercounts the count that is obtained from visual inspection when the baseline amplitude is small.[3] The algorithms applied by the TetraGraph, and other EMG devices, are proprietary, and we can only draw conclusions based on the waveforms themselves. Thus, the study represents a basic science comparison between two different methods of quantitative assessment of neuromuscular block based on the EMG waveform, not a study of how the TetraGraph computes amplitude or a comparison of different devices.

The data from each study subject were downloaded from the Senzime server using the company's TetraConnect data management portal. These files contained the digitised EMG waveforms with 0.01 mV resolution and sampled at 1000 Hz from the adductor pollicis muscle, from 1 to 100 ms after the end of supramaximal stimulation of the ulnar nerve at the wrist.[25] The monitor processed the signals with low-pass filters tuned to the sampling rate to remove high-frequency noise, a standard technique when acquiring electrical signals from patients (personal communication, Fredrik Norrby, Senzime AB, 1 September 2023). The data included the timestamp of each stimulus, the stimulation pattern (e.g. single twitch, TOF, or tetanus), the sequence in the TOF (i.e. first to fourth twitch), if applicable, and the stimulus current and duration. Data were uploaded into an Excel workbook (Microsoft, Redmond, WA, USA) for initial processing and visual examination. Only the first 20 ms after the stimulus were included, as there was no relevant information related to the cMAP after this interval. For each patient, the cMAP from the first twitch of the initial TOF after the determination of the supramaximal stimulation current was used as the control (baseline) T_1 response (T_1c). Each waveform had previously been evaluated by two of the co-authors (RHE, SJB) and tagged as representing a valid cMAP or not based on a visual comparison of the waveform and alignment of the peak and trough location with other cMAPs from the same patients' baseline amplitudes. The dataset has previously been used for the development of a neural network to identify valid cMAPs, and all waveforms classified as representing a valid cMAP were included without restriction as to whether the amplitude was ≥0.4 mV, the threshold used by the TetraGraph.[3] To reiterate, we did not study the algorithms used by the TetraGraph to measure the responses to neurostimulation; rather, we used the digitised waveforms recorded routinely by the TetraGraph and performed all processing *de novo*. Sets of the four responses after TOF stimulation in which the T_1 was considered a valid cMAP were then selected and exported to a comma-separated value (CSV) file, with the cMAP designation (i.e. T_1–T_4) noted for each of the four responses in the TOF sequence.

We emphasise that the signal processing algorithms described below are not identical to those implemented in any of the devices listed in Table 1. Rather, we used the TetraGraph as the source of the digitised EMG waveforms, analogous to other studies published using waveform data obtained from the Relaxograph (Datex, Helsinki, Finland).[12,26]

Waveform smoothing and processing

The CSV data file was imported into the RStudio Integrated Development Environment version 2023.03.01 (RStudio) running R version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria) for processing. To remove noise and improve the accuracy of digital integration of the area under the EMG curve, each waveform was smoothed using the local polynomial regression fitting procedure LOESS with span=0.25 and a resolution=0.1 ms. To adjust for baseline drift in the potential difference between the recording electrode pair, each interpolated value was modified by subtracting the waveform's mean voltage between 1 and 20 ms (i.e. demeaned). Adjusting by a constant does not affect amplitude measurements, only the waveform offset relative to the 0.0 mV baseline.

Calculation of the EMG amplitude

Among waveforms identified as representing a valid cMAP response, the amplitude of processed EMG waveform was calculated as the difference between the peak amplitude measured between 3.7 and 8.7 ms after the stimulus artifact and the trough amplitude measured between 7.2 and 14.2 ms after the stimulus artifact. Those limits were established from previous analyses of the locations of the peaks and troughs of the waveforms.[3]

Calculations of the cMAP AUC

Examination of the EMG waveforms revealed considerable variability in the interval between the stimulus and the peak and trough voltage. Thus, we followed an adaptive approach for calculating the AUC rather than applying fixed limits of integration, as described in van Steen's Master's thesis in electrical engineering, for which raw surface EMG waveforms from the Relaxograph NMT-100 were used.[27] The Relaxograph was the precursor for the GE E-NMT device (GE Healthcare, Chicago, IL, USA), which continues to use fixed limits of integration (Table 1). Our approach more closely resembles that described in the historical study by Engbæk[7] of EMG responses in cats, in which integration of the AUC was performed between the isoelectric baseline before the start and after the end of the cMAP.

Our algorithm to identify the start and end of the cMAP waveform for purposes of numerical integration was as follows, with the location of the peak and trough voltage used as reference points. The lower limit of integration was the latest time before the peak amplitude, where the waveform crossed the 0.00 mV baseline or there was an inflexion in the waveform (i.e. the first derivative changed from positive to negative). If there was no inflexion point and no zero crossing, we used 1.0 ms as the lower limit of integration. The upper limit of integration was the first time after the trough where the 0.00 mV baseline was crossed. If there was no zero crossing, we used the minimum of the time of the trough plus 7.8 ms (the 75th percentile for the interval from the trough to the zero crossing among all cMAPs with a zero crossing) or 20 ms (i.e. the last data point). Limits were calculated to the nearest 0.1 ms. A graphical representation of the adaptive AUC method is shown in Fig. 1 for an archetypal EMG waveform, and examples of varying cMAP waveforms and their integration limits are shown in Fig. 2. The waveform was rectified (i.e. voltages were converted to their absolute values) before numerical integration using the trapezoidal rule.[28]

After this process, the medians of the lower and upper integration limits were 2.9 ms (inter-quartile range 2.0–3.4 ms) and 17.6 ms (inter-quartile range 16.8–18.3 ms) after the stimulus interval, respectively, similar to the fixed limits of integration from the van Steen Master's thesis (i.e. 3–17 ms).[27]

Calculation of depression of the T_1 compared with baseline and the TOFR

The baseline T_1 and AUC were determined for each patient and used to transform the subsequent T_1 determinations into a fraction. That is, the strength of each T_1 was calculated by dividing the T_1 by the corresponding baseline T_1 amplitude (T_{1c}) or AUC (T_{1c} AUC). For each TOF where the T_4 represented a valid cMAP, the TOFR was computed as the amplitude of the T_4 divided by the amplitude of the corresponding T_1 (TOFR amplitude) and the AUC of the T_4 divided by the AUC of the T_1 (TOFR AUC). Such transformation was necessary to allow comparison because the units of amplitude are mV, and those of AUC are mV ms^{-1}.

Statistical analysis

All analyses were performed within the RStudio 2022.12.0 integrated development environment (RStudio) running R version 4.2.1 (R Foundation). To compare the two methods (amplitude and AUC), we used a nested repeated measures Bland–Altman method[21] for the paired measures of the T_1/T_1c and the TOFR using the SimplyAgree v0.1.2 package.[29] This enhancement of the Bland–Altman method is a much more rigorous approach than simple Pearson correlation for establishing the agreement between two different measurement methods, the latter having been applied to previous studies comparing amplitude and AUC methods.[7–9] Furthermore, we used the nested repeated measures version because the originally described Bland–Altman approach is not appropriate when there are multiple measurements for each patient.[21] The difference between the paired amplitude and AUC methods was plotted against the average of the two methods. The statistical package we used adjusts for repeated measurements within subjects and measurements across the continuum of neuromuscular block (i.e. 0–100%).

To determine whether the two methods provided clinically interchangeable information for the T_1/T_1c, we analysed the paired differences when the average T_1/T_1c was $\leq$20%. We required that 95% of those differences would be between 0.9×95% lower confidence limit of the T_1/T_1c bias and 1.1×95% upper confidence limit of the bias. The T_1/T_1c range $\leq$20% was selected because the primary clinical use of the T_1/T_1c is for assessing and managing moderate to deep levels of block (e.g. >95% suppression of the T_1, corresponding to a T_1/T_1c of 0.05).[30] That is, this range represents the level of neuromuscular block commonly utilised by anaesthesia practitioners to ensure that surgical levels of muscle paralysis are maintained.[31] As a secondary analysis, we calculated over the entire range of non-zero average T1/T1c values.

To determine whether the two methods provided clinically interchangeable information for the TOFR, we analysed the paired differences when the average TOFR was $\geq$80%. *A priori*, we required that 95% of those differences would be between

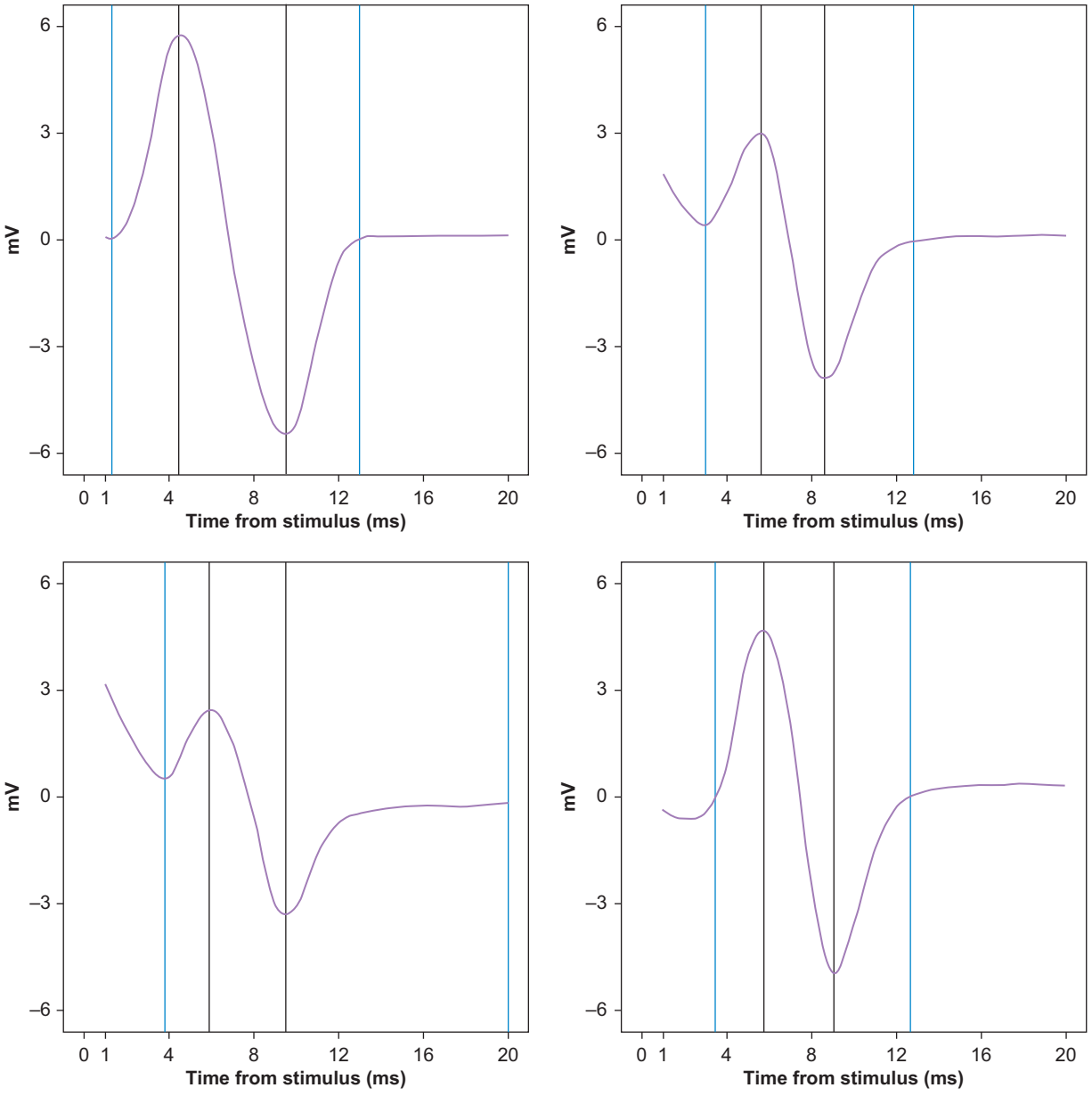

Fig 2. Examples of determination of the amplitude and adaptive area under the curve from a compound motor action potential (cMAP). The purple vertical lines represent the limits of integration of the waveform, and the black vertical lines represent the peak and trough locations. The amplitude is the difference between the peak and the trough. A typical waveform in the top left panel is displayed starting and ending on the 0.0 mV baseline. In the top right panel, the signal is above baseline for the entire interval before the peak; the change in slope of the first derivative of the line (the inflexion point) defines the lower limit of the numerical integration. In the bottom left panel, the signal never returns to baseline, so the upper limit of integration is 20 ms. In the bottom right panel, there is a zero crossing before the peak and after the trough; those form the limits of integration. The signal is rectified (i.e. the voltages are converted to their absolute values) before integration between the intervals marked by the purple lines.

0.95×95% lower confidence limit of the TOFR bias and 1.05×95% upper confidence limit of the bias. The TOFR range ≥80% was selected because the primary clinical use of the TOFR is to assess that there has been sufficient recovery from neuromuscular block (i.e. TOFR ≥0.9) to permit safe tracheal extubation. A narrower range (i.e. 5% *vs* 10%) was selected because of the potentially more serious consequences of underestimating the extent of recovery from neuromuscular block. There is no established use of the TOFR to guide the redosing of neuromuscular blocking agents to maintain a given level of neuromuscular block; rather, that is typically assessed by the TOF count (TOFC). As a secondary endpoint, we analysed the entire range of non-zero average TOFR values (i.e. when one of the two methods provided a ratio >0%).

We calculated Lin's concordance correlation coefficient (LCCC) for repeated measures within subjects[32] from the SimplyAgree v0.1.2 package to assess the reliability of the two measurements.[29]

We also calculated the repeatability coefficient, defined as the upper limit of the specified prediction interval (e.g. 95%) for the absolute value of the difference between replicate measurements using a given method performed under identical conditions.[33] Lower values of the repeatability coefficient correspond to better repeatability. We had a calculation challenge because the extent of neuromuscular block often changes substantively between successive measurements, particularly during onset and after reversal, violating the requirement for identical conditions. It is not possible to measure responses to TOF stimulation in rapid succession because doing so results in facilitation of the subsequent response.[23] Furthermore, the patients in the first study, who underwent robotic or laparoscopic surgery, were maintained at deep levels of block for most of the surgical procedure, resulting in frequent absence of a fourth twitch in the TOF sequence and often with a TOFC=0, precluding assessment of either the TOFR or the T_1/T_1c.[24] Thus, we could only evaluate the repeatability coefficients of the two methods using data from the second study, where patients were given a smaller initial muscle relaxant dose and allowed to recover spontaneously.[15] We considered replicates as the non-overlapping sequential q-15 second TOF stimulations occurring within 60 s after the first stimulation in that window. For example, if TOF stimulation sequences were recorded at 09:00:05, 09:00:20, 09:00:50, and 10:00:05, the first three would be included as the 09:00:05 replicates, and the next 60-s window started at 10:00:05. If there were an interruption of >20 s (i.e. longer than the nominal 15-s TOF stimulation interval) between the second and third TOF stimulation in a window, the first two stimulations were used as the replicate values. If there was an interruption of >20 s between the first and second TOF stimulation, that window was dropped, and the next 60-s window started when TOF stimulation resumed. Our grouping of TOF stimulations within 60 s as replicates introduced a degree of bias toward lower repeatability for both methods because of the continuous waning of neuromuscular block during spontaneous recovery. However, limiting the window to 60 s mitigated this effect. Multilevel mixed effects linear regression in Stata (StataCorp, College Station, TX, USA) was used to calculate the within-subject standard deviations (SD) of the differences between the replicate measurements for each method and was calculated separately for each of the 19 patients. The mean and standard errors of the mean repeatability indices were then calculated using the individual patient results (N=19). For the repeatability calculations, data were analysed after the $T_1/_{T1c}$ returned to at least 0.1 (after the nadir of neuromuscular block was reached) until the end of EMG recording.

A power analysis was not performed to determine the sample size (i.e. number of patients) because no prior data (i.e. means and SD of differences between AUC- and amplitude-based EMG determinations for either the T_1/T_1c or the TOFR) were available to inform such a decision. Rather, all available patient data (n=46) were used, making this a convenience sample.

Results

Among the 46 patients studied, 7286 previously validated T_1 cMAPs[3] were evaluated, of which 2419 corresponded to a deeper level of block (i.e. where the average $T_1/T_1c \leq 0.2$, Table 2). Overall T_1/T_1c bias between the AUC and amplitude methodologies was 0.032 (95% confidence interval [CI] 0.022–0.043) but only 0.0077 (95% CI 0.0017–0.014) in the range of primary clinical interest (i.e. deeper block, Table 2, Fig 3). Only 0.33% of the paired measurements were outside the clinical level of agreement in the range of clinical interest, and 4.71% were outliers overall (Table 2, Fig 3). Overall, Lin's concordance correlation coefficient=0.98 (95% CI 0.98–0.98), demonstrated high reliability (Table 2, Fig 3). Thus, the methods can be considered clinically interchangeable for determination of the T_1/T_1c, There were 5541 evaluable TOFRs (i.e. with both a valid T_1 and T_4; Table 2), of which 1781 corresponded to the interval during which patients are in the process of recovering from their neuromuscular block (i.e. TOFR $\geq$0.8) (Table 2). Overall TOFR bias between the AUC- and amplitude-based methodologies was 0.011 (95% CI 0.0050–0.017) and similar for the recovery interval (mean 0.012, 95% CI 0.0061–0.019, Table 2, Fig 4). TOFR differences between the methodologies in the range of clinical interest (recovery) outside the clinical limits of agreement occurred in 3.93% of the pairs and, overall, in 5.45% (Table 2, Fig 4). Overall, Lin's concordance correlation coefficient=0.995 (95% CI 0.995–0.996), demonstrating high reliability (Table 2, Fig 4). Thus, the methods can be considered clinically interchangeable for determination of the TOFR.

A summary of the numbers of paired T_1/T_1c and TOFR comparisons among the patients is presented in Appendix 1.

Repeatability was assessed using 1162 replicates among the N=19 patients from the second study.[15] The mean (SD, 95% CI) of the 95% repeatability coefficients for T_1/T_1c were 0.068 (0.035, 0.054–0.083) for the amplitude method and 0.067 (0.036, 0.052–0.081) for the AUC method. Similarly, the 95% repeatability coefficients for the TOFR were 0.077 (0.022, 0.068–0.085) and 0.076 (0.022, 0.068–0.085), respectively. Thus, using either method, 95% of the absolute values of the differences between replicate values would be within 7% for the T_1/T_1c and within 8% for the TOFR. Because the 95% CIs for the amplitude and AUC methods overlap for the T_1/T_1c and the TOFR, there is no statistically significant difference between the two methods.

Discussion

The results of this study confirm the previous historical assertion that amplitude and AUC methods of analysis of the EMG waveform (T_1/T_1c and TOFR) after nerve stimulation are clinically interchangeable, despite those previous analyses being based on statistical methodology that subsequently was determined to be inappropriate for such assessment.[7–9] Agreement was good across the entire range of T_1/T_1c and TOFR (Figs 3 and 4). There was no indication that either method had better repeatability. Thus, clinical decisions about the level of neuromuscular block are agnostic as to whether the neuromuscular monitor used a method based on calculating the cMAP amplitude or the cMAP AUC. Furthermore, research studies using one EMG method *vs* the other can validly be compared. However, because the identification of valid twitches in current devices is dependent, in part, on arbitrary amplitude thresholds (Table 1), one cannot equate either the TOFC or the post-tetanic count (PTC) as reported by different devices based solely on cMAP amplitude. We recently published a methodology that used a fully connected neural network to identify valid cMAPs with an accuracy exceeding 99.5%; this method would resolve threshold amplitude-related discrepancies among devices.[3] We have no

Table 2 Comparison of the amplitude and adaptive AUC methods for determination of the T_1/T_1c and the TOFR. AUC, area under the curve; BA, Bland–Altman; CI, confidence interval; LCCC, Lin's concordance correlation coefficient; LoA, limits of agreement; T_1, first twitch in the train-of-four; T_1c, control (baseline) first twitch in the train-of-four; TOFR, train-of-four ratio. *The T_1/T_1c was calculated as the ratio of the amplitude or AUC of the first twitch after TOF stimulation divided by the baseline T_1 amplitude or AUC, respectively, measured before administration of the neuromuscular blocking drug. The TOFR was calculated as the ratio of the amplitude or AUC of the fourth twitch after TOF stimulation divided by the corresponding T_1 amplitude or AUC, respectively. †The range of the clinically acceptable limits of agreement for the T_1/T_1c was prespecified as 10% (0.1) below the 95% lower confidence limit of the bias to 10% (0.1) above the 95% upper confidence limit of the bias. The range of the clinically acceptable limits of agreement for the TOFR was prespecified as 5% (0.05) below the 95% lower confidence limit of the bias to 5% (0.05) above the 95% upper confidence limit of the bias. *‡The number of paired differences that were outside the range of the clinical LoA (previous column in the table).

Measure*	Range	Bias (95% CI)	Bland–Altman lower LoA (95% CI)	Bland Altman upper LoA (95% CI)	LCCC (95% CI)	Number of pairs compared	Range of the clinical LoA†	Number outside the clinical LoA‡ *n* (%)
T_1/T_1c	All	0.032 (0.022–0.043)	−0.078 (−0.092 to −0.067)	0.14 (0.13 to 0.16)	0.98 (0.98 to 0.98)	7286	−0.12 to 0.14	343 (4.71)
T_1/T_1c	≤20%	0.0077 (0.0017–0.014)	−0.039 (−0.048 to −0.032)	0.055 (0.047 to 0.064)	0.91 (0.903 to 0.91)	2419	−0.10 to 0.11	8 (0.33)
TOFR	All	0.011 (0.0050–0.017)	−0.058 (−0.066 to −0.052)	0.080 (0.074 to 0.088)	0.995 (0.995 to 0.996)	5559	−0.055 to 0.067	303 (5.45)
TOFR	≥80%	0.012 (0.0061–0.019)	−0.051 (−0.059 to −0.044)	0.076 (0.069 to 0.084)	0.90 (0.89 to 0.91)	1781	−0.056 to 0.069	70 (3.93)

information from manufacturers whose devices use AUC methods as to how valid cMAPs are identified. Thus, we lacked a basis on which to compare the TOFC or PTC using the two methodologies and could only analyse the unit-independent ratios (i.e. TOFR and T_1/T_1c).

Limitations and strengths

Our study was performed using the digitised EMG output from a single manufacturer's monitor (TetraGraph), and waveforms may be slightly different from those obtained from other devices as a result of varying implementations of electronic and digital filtering processes. We did not have access to the digitised waveforms from other devices, so we could not perform comparative assessments. However, we think it unlikely that subtle changes in EMG waveform morphology would make a substantive difference in calculating the amplitude or the AUC. Nonetheless, such determinations represent an area of research that we are pursuing currently. The information provided by Senzime AB (which computes the amplitude of the cMAP) indicated that their process of determining the peak and trough voltage also involves smoothing and interpolation. However, our approach was not designed to replicate the proprietary algorithms in that monitor. The NMT Pod (Nihon Kohden, Tokyo, Japan) also identifies the peak and trough voltage to compute the amplitude of the cMAP. We think substantive differences would unlikely accrue from this standard signal processing methodology to identify such inflexion points in the waveform. A strength of the study is that the data used for analysis included the entire range of possible values for both the T_1/T_1c and TOFR.

Regarding the calculation of the AUC, methods are not as well defined as those related to peak finding (e.g. limits of integration, numerical integration methods, baseline voltage offset adjustment, etc.). Our method of calculation of the AUC undoubtedly differs somewhat in precise details from those implemented in clinical monitors, but is qualitatively similar to that described by Xavant Technology (Pretoria, South Africa), which also adaptively integrates the cMAP. We cannot comment on the AUC algorithm implemented in the TwitchView (Blink Device Company, Seattle, WA, USA) because the company declined to provide any details on its process despite several requests. Thus, while our study of digitised EMG waveforms provides fundamental assurance that quantitative information on the degree of neuromuscular block is clinically interchangeable whether calculated using amplitude or AUC, this might not be the case if a manufacturer's algorithm to remove noise, smooth the signal, and calculate the AUC were poorly implemented. From inspection of the digitised EMG data, waveforms were often encountered where there was considerable noise outside the portion of the cMAP being integrated, and occasionally, some baseline shift of the signal was noted. In addition, there were often cMAPs where the peak and trough were clearly identifiable despite the presence of noise within the limits of integration. Thus, fixed limits of integration (as in the E-NMT device) may result in a lower degree of agreement than an amplitude-based approach.

Another limitation is that hand temperatures were not actively maintained during the studies; rather, core temperatures were maintained above 36°C. There is an inverse relationship between temperature and twitch amplitude or AUC, but these have different slopes when measured in the cat. For example, with each 1°C increase in muscle temperature, twitch amplitude decreased by 2%, whereas the AUC

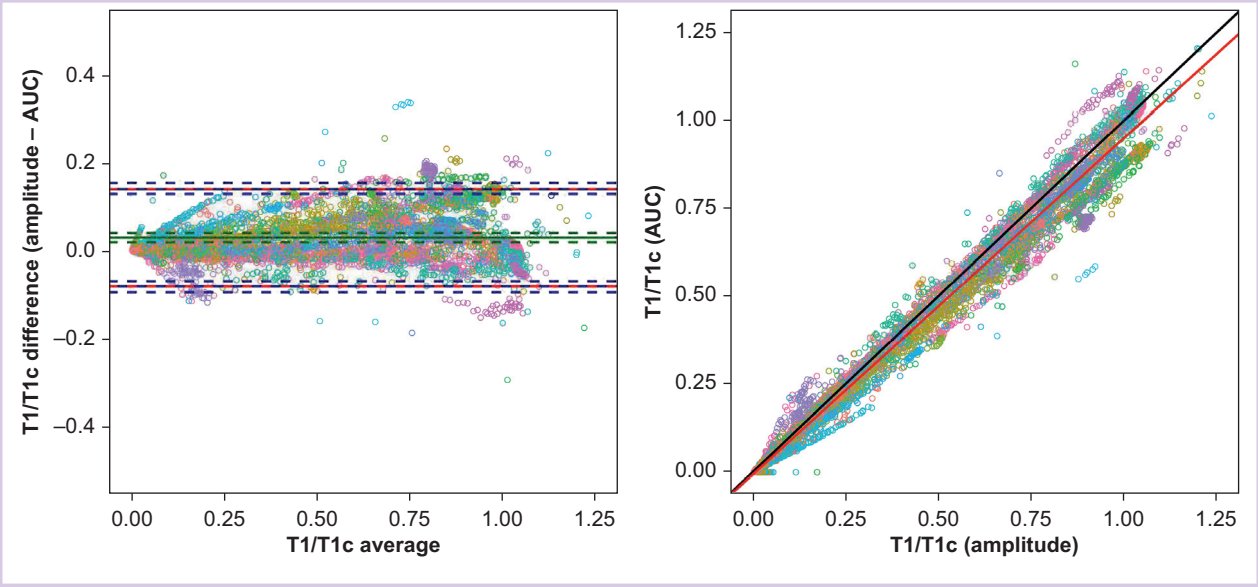

Fig 3. Bland–Altman (left panel) and correlation (right panel) plots for the ratio of the first twitch (T_1) divided by the control (baseline) first twitch (T_1c) after train-of-four stimulation (T_1/T_1c) between the adaptive area under the curve (AUC) and amplitude methods. On the left, the bias and its 95% confidence interval are shown (green lines), and the upper and lower limits of agreement with their corresponding 95% confidence intervals are displayed (blue lines). The clinical limits of agreement are displayed as the red line. On the right, Lin's concordance correlation coefficient=0.98, comparing the two methods with the identity line (black) and the best fit linear line (red), diverging slightly from the common intercept at the origin. The prespecified clinical limits of agreement of plus or minus 5% from the upper and lower confidence limits of bias are displayed (red dashed line). The coloured circles represent the individual paired measurements, with a different colour corresponding to the values contributed by each patient. See Table 2 for the values of the various reference lines and the percentages of measurements falling outside those limits.

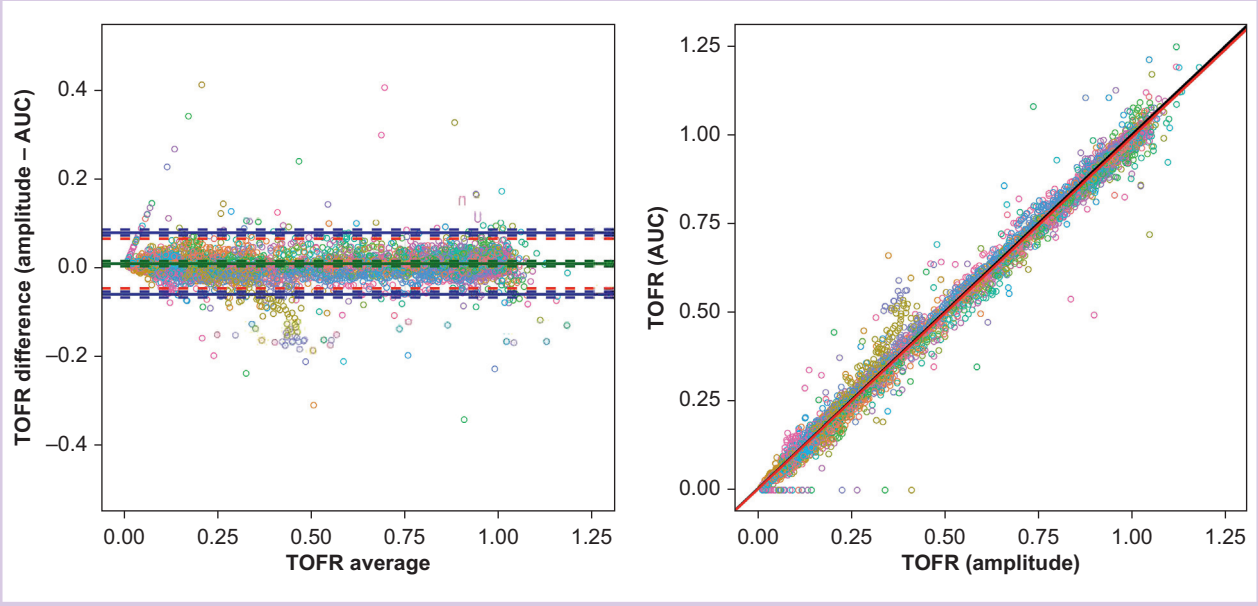

Fig 4. Bland–Altman (left panel) and correlation (right panel) plots for the train-of-four ratio (TOFR) between the adaptive area under the curve (AUC) and amplitude methods. On the left, the bias and its 95% confidence interval are shown (green lines), and the upper and lower limits of agreement with their corresponding 95% confidence intervals are displayed (blue lines). The clinical limits of agreement are displayed as the red line. On the right, Lin's concordance correlation coefficient is 0.99, comparing the two methods with the identity line (black) and the best fit linear line (red) nearly overlapping. The prespecified clinical limits of agreement of plus or minus 5% from the upper and lower confidence limits of bias are displayed (red dashed line). The coloured circles represent the individual paired measurements, with a different colour corresponding to the values contributed by each patient. See Table 2 for the values of the various reference lines and the percentages of measurements falling outside those limits.

decreased by 6%.[34] The consequence in the cat model would be that if the peripheral temperature increases, the T_1/T_1c measured using the AUC would be systematically lower than when measured using amplitude. However, there would be no effect on the TOFR because the fractional increase would occur in both the numerator (T_4) and the denominator (T_1), cancelling the effect. In 24 patients undergoing ophthalmic surgery under general anaesthesia, Santanen and Paloheimo[35] studied the effect of deliberate cooling and re-warming of the hand on the AUC T_1 response from the first dorsal interosseous muscle after every 1-min TOF stimulation of the ulnar nerve at the wrist. Similar to the findings in the cat,[34] each 1°C change in muscle temperature (measured with a myocardial temperature sensor inserted into the belly of the muscle) resulted in an 8.0% (SD 0.8%) difference in the T_1 response compared with its stable baseline value. The effect was 4.1% (SD 1.6) per °C change in the skin temperature probe over the muscle.

A final limitation is that the study included only adult patients without diseases potentially affecting neuromuscular transmission. Thus, the generalisability of our results to children and patients with neuromuscular disorders remains to be established.

Conclusions

Using digitised EMG waveforms exported from a clinically available neuromuscular block monitor, computing the amplitude of the cMAP or the AUC provided clinically interchangeable information. There was no evidence suggesting a preference for one method over the other. The implication is that clinicians can continue to follow the historical assertion that the two methods provide interchangeable information. Furthermore, clinical research studies using different methods of assessing the degree of block based on EMG can likely be regarded as interchangeable.[4]

Authors' contributions

Conceptualisation, validation: RHE
Resources: RHE, JRR, RN
Data curation, formal analysis: RHE
Visualisation, writing original draft: RHE, SJB
Writing review and editing: all authors

Declarations of interest

SJB has intellectual property assigned to Mayo Clinic (Rochester, MN), is a consultant for Merck (Kenilworth, NJ), is a principal, shareholder, and Chief Medical Officer in Senzime AB (Uppsala, Sweden), and is an unpaid member of the Scientific/Clinical Advisory Boards for Coala Life (Irvine, CA), NMD Pharma (Aarhus, Denmark), and Takeda Pharmaceuticals (Cambridge, MA). JRR has ongoing industry-sponsored research from Merck (with funds to his employer) and has served on a Scientific Advisory Board for Senzime AB (Uppsala, Sweden). The other authors declare that they have no conflicts of interest.

Funding

Departmental resources only.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bjao.2024.100293.

References

1. Thilen SR, Weigel WA, Todd MM, et al. 2023 American Society of Anesthesiologists practice guidelines for monitoring and antagonism of neuromuscular blockade: a report by the American Society of Anesthesiologists task force on neuromuscular blockade. *Anesthesiology* 2023; **138**: 13–41
2. Fuchs-Buder T, Romero CS, Lewald H, et al. Peri-operative management of neuromuscular blockade: a guideline from the European Society of Anaesthesiology and Intensive care. *Eur J Anaesthesiol* 2023; **40**: 82–94
3. Epstein RH, Perez OF, Hofer IS, Renew JR, Nemes R, Brull SJ. Validation of a convolutional neural network that reliably identifies electromyographic compound motor action potentials following train-of-four stimulation: an algorithm development experimental study. *BJA Open* 2023; **8**, 100236
4. Viby-Mogensen J, Jensen E, Werner M, Kirkegaard Neilsen H. Measurement of acceleration: a new method of monitoring neuromuscular function. *Acta Anaesthesiol Scand* 1988; **32**: 45–8
5. Murphy GS. Neuromuscular monitoring in the perioperative period. *Anesth Analg* 2018; **126**: 464–8
6. Carvalho H, Verdonck M, Brull SJ, et al. A survey on the availability, usage and perception of neuromuscular monitors in Europe. *J Clin Monit Comput* 2023; **37**: 549–58
7. Engbæk J. Monitoring of neuromuscular transmission by electromyography during anaesthesia. A comparison with mechanomyography in cat and man. *Dan Med Bull* 1996; **43**: 301–16
8. Pugh ND, Kay B, Healy TE. Electromyography in anaesthesia. A comparison between two methods. *Anaesthesia* 1984; **39**: 574–7
9. Engbæk J, Roed J, Hangaard N, Viby-Mogensen J. The agreement between adductor pollicis mechanomyogram and first dorsal interosseous electromyogram. A pharmacodynamic study of rocuronium and vecuronium. *Acta Anaesthesiol Scand* 1994; **38**: 869–78
10. Iwasaki H, Takahiko K, Matsui T, et al. The impact of electrosurgical devices on electromyography-based neuromuscular monitoring during abdominal laparotomy: a case series. *J Clin Monit Comput* 2023; **37**: 1111–4
11. Chaves-Cardona HE, Fouda EA, Hernandez-Torres V, et al. Comparison of onset of neuromuscular blockade with electromyographic and acceleromyographic monitoring: a prospective clinical trial. *Braz J Anesthesiol* 2023; **73**: 393–400
12. Sato H, Iwasaki H, Doshu-Kajiura A, et al. Comparison of two electromyography-based neuromuscular monitors, AF-201P and TetraGraph, in rocuronium-induced neuromuscular block: a prospective comparative study. *Anaesth Crit Care Pain Med* 2022; **41**, 101145
13. Iwasaki H, Yamamoto M, Sato H, et al. A comparison between the adductor pollicis muscle using TOF-Watch SX and the abductor digiti minimi muscle using TetraGraph

in rocuronium-induced neuromuscular block: a prospective observational study. *Anesth Analg* 2022; **135**: 370–5
14. Giudici G, Piccioni F, Proto P, Valenza F. A comparison of accelerometric monitoring by TOF Watch® SX and electromyographic monitoring by Tetragraph® for recovery from neuromuscular blockade. *J Clin Anesth* 2021; **75**, 110481
15. Nemes R, Lengyel S, Nagy G, et al. Ipsilateral and simultaneous comparison of responses from acceleromyography- and electromyography-based neuromuscular monitors. *Anesthesiology* 2021; **135**: 597–611
16. Renew JR, Hernandez-Torres V, Logvinov I, et al. Comparison of the TetraGraph and TOFscan for monitoring recovery from neuromuscular blockade in the post anesthesia care unit. *J Clin Anesth* 2021; **71**, 110234
17. Bussey L, Jelacic S, Togashi K, Hulvershorn J, Bowdle A. Train-of-four monitoring with the TwitchView monitor electromyograph compared to the GE NMT electromyograph and manual palpation. *J Clin Monit Comput* 2021; **35**: 1477–83
18. Bowdle A, Bussey L, Michaelsen K, et al. Counting train-of-four twitch response: comparison of palpation to mechanomyography, acceleromyography, and electromyography. *Br J Anaesth* 2020; **124**: 712–7
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**: 307–10
20. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; **8**: 135–60
21. Zou GY. Confidence interval estimation for the Bland-Altman limits of agreement with multiple observations per individual. *Stat Methods Med Res* 2011; **22**: 630–42
22. Abu-Arafeh A, Jordan H, Drummond G. Reporting of method comparison studies: a review of advice, an assessment of current practice, and specific suggestions for future reports. *Br J Anaesth* 2016; **117**: 569–75
23. Fuchs-Buder T, Brull SJ, Fagerlund MJ, et al. Good clinical research practice (GCRP) in pharmacodynamic studies of neuromuscular blocking agents III: the 2023 Geneva revision. *Acta Anaesthesiol Scand* 2023; **67**: 994–1017
24. Hernandez V, Chaves-Cardona H, Renew JR, Brull SJ. Electromyographic and acceleromyographic monitoring in restricted arm movement surgical setting: a prospective, randomized trial 2021 Annual Meeting of the American Society of Anesthesiologists, San Diego, CA. (abstract). *Anesthesiology* 2021: A2069
25. Senzime. *TetraGraph model SEN 2001 neuromuscular monitor operating instructions*. Available from: https://www.manualslib.com/manual/1848525/Senzime-Tetragraph-Sen-2001.html. [Accessed 6 December 2023]
26. Kalli I. Effect of surface electrode positioning on the compound action potential evoked by ulnar nerve stimulation in anaesthetized infants and children. *Br J Anaesth* 1989; **62**: 188–93
27. van Steen MHA. *Validation of muscle relaxation measurements 1998*. Available from: https://research.tue.nl/en/studentTheses/validation-of-muscle-relaxation-measurements. [Accessed 6 December 2023]
28. Wikipedia. *Trapezoidal rule*. Available from: https://en.wikipedia.org/wiki/Trapezoidal_rule. [Accessed 30 December 2023]
29. Caldwell AR. SimplyAgree: an R package and jamovi module for simplifying agreement and reliability analyses. *J Open Source Softw* 2022; **7**: 4148
30. Brull SJ, Kopman AF. Current status of neuromuscular reversal and monitoring: challenges and opportunities. *Anesthesiology* 2017; **126**: 173–90
31. Martini CH, Boon M, Bevers RF, Aarts LP, Dahan A. Evaluation of surgical conditions during laparoscopic surgery in patients with moderate vs deep neuromuscular block. *Br J Anaesth* 2014; **112**: 498–505
32. Lin Li-K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989; **45**: 255–68
33. Carstensen B, Simpson J, Gurrin LC. Statistical models for assessing agreement in method comparison studies with replicate measurements. *Int J Biostat* 2008; **4**. Article 16
34. Engbæk J, Skovgaard Briis B, Kann T, Viby-Mogensen J. Monitoring of the neuromuscular transmission by electromyography (I). Stability and temperature dependence of evoked EMG response compared to mechanical twitch recording in the cat. *Acta Anaesthesiol Scand* 1992; **36**: 495–504
35. Santanen OA, Paloheimo MPJ. Effects of small peripheral temperature changes on the evoked baseline electromyographic response. *Acta Anaesthesiol Scand* 1999; **43**: 338–42

Handling editor: Phil Hopkins

Appendix 1.

Summary of the numbers and magnitudes of the paired comparisons between the amplitude and the area under the curve methods for assessment of the electromyographic compound motor action potentials among the n=26 patients. IQR, inter-quartile range; SD, standard deviation; T_1, first twitch of train-of-four; T_1c, control (baseline) first twitch of train-of-four; TOFR, train-of-four ratio.

Measure	Range	Number of comparisons, median (IQR)	Mean difference (SD)
T_1/T_1c	All	138 (52.5–240)	0.024 (0.030)
T_1/T_1c	≤20%	67 (24–67)	0.0030 (0.032)
TOFR	All	65 (24–226)	−0.0022 (0.044)
TOFR	≥80%	42 (11–89)	0.0026 (0.047)