Study Title: OBAEG Study –Acceleromyography versus Electromyography in Obesity

Official Title: Comparison between acceleromyography (AMG) and electromyography (EMG) in patients with obesity

Document Type: Study Protocol and Statistical Analysis Plan

Version: v1.0 – August 2025

Document Date: October 2025

NCT Number: Pending assignment

Study Protocol and Statistical Analysis Plan (Combined)

Official Title: Comparison of Acceleromyography and Electromyography in Obese Patients Undergoing General Anesthesia With Rocuronium: A Prospective Observational Study Organization's Unique

Acronym: OBAEG

Protocol ID: AOP3692_CET

Secondary ID: 6296/EST/25 (Territorial Ethics Committee – Central-East Veneto Area)

This is the English translated version, faithfully aligned with the version registered and approved by the Ethics Committee. The original Italian version can be provided upon request.

Table of Contents

Part A – Study Protocol Synopsis

Introduction

Obesity is defined using the Body Mass Index (BMI), calculated as weight in kilograms divided by height in meters squared (kg/m²). Severity is classified as follows: Normal weight (BMI 20–24), Overweight (25–29), Obesity Class I (30–34), Class II (35–39), and Class III (>40). Obesity is a global public health problem due to its association with increased morbidity and mortality. Metabolic bariatric surgery is the most effective long-term treatment, requiring optimal perioperative management and tailored anesthetic approaches, especially in obese patients at higher risk of postoperative residual curarization (PORC). Quantitative neuromuscular monitoring is essential to prevent PORC. Acceleromyography (AMG) measures muscle movement but may have precision limitations, while electromyography (EMG) measures electrical muscle activity and is more reliable.

Study Objective

To compare the accuracy and precision of AMG and EMG in measuring the Train-of-Four (TOF) ratio during general anesthesia in patients with obesity undergoing bariatric surgery, and to assess perioperative usability, recovery times, and postoperative respiratory complications.

Study Design

Prospective observational cohort study. Thirty adult patients with pathological obesity (Class II with comorbidities or Class III) scheduled for bariatric surgery under general anesthesia will be monitored using both AMG and EMG (one sensor applied to each hand) throughout anesthesia and recovery.

Population

Inclusion criteria: Adults aged 18–65 years, pathological obesity (Class II with comorbidities or Class III), scheduled for bariatric surgery under general anesthesia, and able to provide informed consent. *Exclusion criteria:* Severe acute or chronic respiratory or cardiac disease, end-stage hepatic or renal disease, intolerance/allergy to study drugs, or absence of informed consent.

Sample Size Justification

Based on Wedemeyer Z. et al., who reported a mean difference of 0.12 (SD 0.13) between AMG and EMG TOF ratios in a non-paired comparison, the required sample size for the present paired study was estimated using a two-sample t-test for the comparison of means, with a significance level of $\alpha = 0.05$

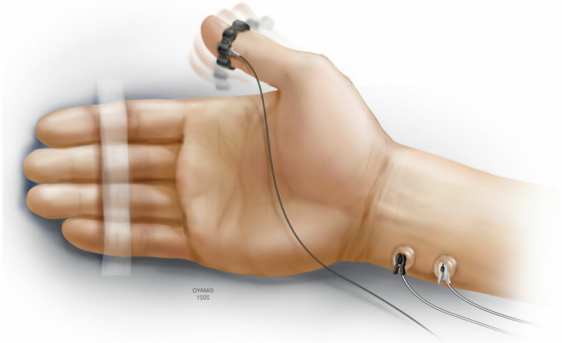
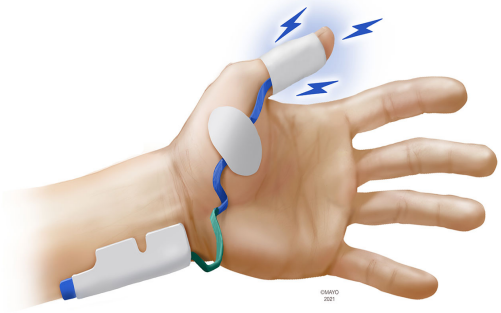
and a statistical power of 90%. The required number of patients is 25. Allowing for an anticipated dropout rate of 16%, a total of 30 patients will be enrolled. This approach is deliberately conservative: it provides protection against higher variability in the obese population, ensures robustness for the primary analysis of paired mean differences, and offers sufficient power for exploratory secondary endpoints. The calculation was performed in R. Unlike Wedemeyer et al., this study evaluates patients with obesity under general anesthesia with paired AMG–EMG measurements in the same subject; repeated measures will be collected but are analyzed as secondary outcomes and were not included in the sample size calculation. In the analysis, the distribution of paired differences will be assessed (Shapiro–Wilk, QQ-plots), and if normality is not satisfied, non-parametric methods (Wilcoxon signed-rank test) will be applied.

Anesthesiologic Approach

Pre-anesthesia: Standard cardio-respiratory monitoring and peripheral venous access placement in the preoperative area. Intraoperative: Completion of monitoring with anesthesia depth sensor, TOF monitoring (AMG on one hand, EMG on the other), and radial arterial catheter placement.

General anesthesia will follow the unit's standardized bariatric protocol. Neuromuscular parameters will be recorded in the anesthesia record and on a dedicated case report form by an independent investigator.

Post-anesthesia: After TOF ratio > 0.9 is reached, the patient will be extubated, monitored in PACU, and discharged to the ward once criteria are met.

AMG	EMG
	

Outcome Measures

Primary: The primary endpoint will be the mean paired difference in TOF ratio values between AMG and EMG, evaluated on standardized paired measurements at baseline (before induction) and recovery (after reversal). The corresponding 95% confidence interval (CI) of the mean paired difference will be reported.

Supportive analyses Accuracy and precision of AMG vs EMG will be further assessed using Bland–Altman analysis (bias and limits of agreement, with SD of the bias as measure of precision) and Lin's concordance correlation coefficient (CCC). Repeated intraoperative measurements (induction, maintenance, recovery progression) will be analyzed separately as exploratory endpoints to evaluate phase-related variations in AMG–EMG comparison.

Secondary: Time to optimal intubation conditions (TOF count = 0 with PTC confirming deep block), duration and stability of deep block (PTC < 5), time to recovery (TOF ≥ 0.9) after sugammadex, usability/quality ratings of AMG and EMG (1–5 scale, including setup time, signal stability, and measurement quality), incidence of postoperative respiratory complications (from extubation until ward discharge), and perioperative changes in HR, MAP, SpO_2, and perfusion index.

Statistical Analysis

Continuous variables will be summarized as mean ± standard deviation (SD) or median and interquartile range (IQR), with normality assessed using the Shapiro–Wilk test. Paired comparisons between AMG and EMG will be performed using the paired t-test or the Wilcoxon signed-rank test, as appropriate. Categorical variables will be expressed as counts and percentages. TOF ratios will be repeatedly measured at predefined perioperative timepoints (baseline, during block, and recovery). These repeated measures will be analyzed using repeated-measures ANOVA if assumptions hold, or the Friedman test otherwise. If the global test is significant, post-hoc pairwise comparisons between timepoints (e.g., vs baseline) will be performed with appropriate correction for multiple testing. Multivariate logistic regression will be used to explore predictors of dichotomous outcomes (e.g., postoperative respiratory complications). Stepwise selection with the Akaike Information Criterion (AIC) will guide model building. Multicollinearity among predictors will be assessed using the Variance Inflation Factor (VIF). Results will be reported as odds ratios (OR) with 95% confidence intervals (CIs). Statistical significance will be set at $p < 0.05$. Analyses will be performed in R. Missing data will not be imputed unless more than 10% of recovery values are missing and the absence is plausibly random, in which case a Last Observation Carried Forward (LOCF) approach will be considered. Accuracy will be evaluated relative to the clinical reference TOF ratio of 1.0 in predefined windows where 1.0 is expected (baseline pre-block and recovery prior to PACU discharge). Results will be reported both as raw TOF values and as baseline-normalized TOF (TOF/current ÷ TOF/baseline). Accuracy summaries will include bias and precision, the proportion of measurements within ±0.05 and ±0.10 of 1.0, and an equivalence assessment to 1.0 using two one-sided tests (TOST) with margins of ±0.10. Agreement between methods (AMG vs EMG) will be reported as supportive analyses using Bland–Altman plots (bias and 95% limits of agreement) and Lin's concordance correlation coefficient (CCC). Primary comparisons will use paired statistical methods. Accuracy versus 1.0 will be summarized separately in baseline and recovery windows (bias, SD/RMSE, proportions within ±0.05 and ±0.10, TOST ±0.10). The detailed Statistical Analysis Plan (SAP) is included as a technical appendix to this protocol.

References

1. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in underweight and obesity from 1990 to 2022: a pooled analysis of 3663 population-representative studies with 222 million children, adolescents, and adults. Lancet. 2024 Mar 16;403(10431):1027-1050
2. Murphy GS, Szokol JW, Marymont JH, Greenberg SB, Avram MJ, Vender JS. Residual neuromuscular blockade and critical respiratory events in the postanesthesia care unit. Anesth Analg. 2008 Jul;107(1):130-7.
3. Carron M, Tamburini E, Iepparielloo G, Linassi F. Br J Anaesth. 2023 Jun;130(6):e461-e463.
4. Fuchs-Buder T, Romero CS, Lewald H, Lamperti M, Afshari A, Hristovska AM, Schmartz D, Hinkelbein J, Longrois D, Popp M, de Boer HD, Sorbello M, Jankovic R, Kranke P. Peri-operative management of neuromuscular blockade: A guideline from the European Society of Anaesthesiology and Intensive Care. Eur J Anaesthesiol. 2023 Feb 1;40(2):82-94.
5. Naguib M, Brull SJ, Johnson KB. Conceptual and technical insights into the basis of neuromuscular monitoring. Anaesthesia. 2017 Jan;72 Suppl 1:16-37
6. Wedemeyer Z, Michaelsen KE, Jelacic S, Silliman W, Lopez A, Togashi K, Bowdle A. Accuracy and Precision of Three Acceleromyographs, Three Electromyographs, and a Mechanomyograph Measuring the Train-of-four Ratio in the Absence of Neuromuscular Blocking Drugs. Anesthesiology. 2024 Aug 1;141(2):262-271.
7. Carron M, Safaee Fakhr B, Iepparielloo G, Foletto M. Perioperative care of the obese patient. Br J Surg. 2020 Jan;107(2):e39-e55.
8. Kotake Y, Ochiai R, Suzuki T, Ogawa S, Takagi S, Ozaki M, Nakatsuka I, Takeda J. Reversal with sugammadex in the absence of monitoring did not preclude residual neuromuscular block. Anesth Analg. 2013 Aug;117(2):345-51.
9. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014 Dec 19;14:135.

Part B – Statistical Analysis Plan (SAP)

1. Introduction and Purpose

This Statistical Analysis Plan (SAP) details the pre-specified analyses for the OBAEG study and complements the Study Protocol.

2. Study Overview

Design: Prospective observational cohort.

Population: Adults 18–65 years with pathological obesity (Class II with comorbidities or Class III) undergoing bariatric surgery under general anesthesia.

Sample size: 30 patients (25 required by power calculation + 16% anticipated dropout).

3. Objectives and Endpoints

Primary Objective: To evaluate the difference in Train-of-Four (TOF) ratio measurements between AMG and EMG during general anesthesia.

Primary Endpoint: The mean paired difference in TOF ratio values between AMG and EMG, evaluated on standardized paired measurements at baseline (before induction) and recovery (after reversal). The corresponding 95% confidence interval will be reported.

Supportive Endpoints:

- Accuracy and precision of AMG vs. EMG will be further characterized through Bland–Altman analysis (bias and 95% limits of agreement, with the SD of the bias as measure of precision).
- Concordance between AMG and EMG will be quantified using Lin's concordance correlation coefficient (CCC) with 95% confidence intervals.
- Repeated intraoperative AMG–EMG measurements (induction, maintenance, recovery progression) will be analyzed separately as exploratory endpoints to assess phase-related variations in AMG–EMG comparison.

Secondary Endpoints:

- Time to optimal intubation conditions (TOF count = 0 with PTC confirming deep block).
- Duration and stability of deep neuromuscular block (PTC < 5).
- Time to recovery to TOF ratio ≥ 0.9 after sugammadex.
- Usability/quality ratings (1–5 scale) for AMG and EMG (setup time, signal stability, measurement

quality).

- Incidence of postoperative respiratory complications after tracheal extubation until ward discharge.
- Perioperative changes in HR, MAP, SpO_2, and perfusion index.

4. Analysis Populations

Full Analysis Set (FAS): All enrolled participants with at least one valid paired AMG–EMG measurement.

Per-Protocol Set (PPS): Subset of FAS without major protocol deviations that could affect neuromuscular monitoring.

Safety Set: All participants receiving general anesthesia and neuromuscular monitoring.

5. General Statistical Principles

Two-sided tests with $\alpha = 0.05$. Continuous variables summarized as mean ± SD or median (IQR) per distribution (Shapiro–Wilk). Categorical variables summarized as counts and percentages with 95% CIs. Given the within-subject design, primary comparisons use paired methods. No formal multiplicity adjustment is planned; secondary analyses are supportive and considered exploratory. Therefore, p-values for secondary and exploratory endpoints will be reported descriptively rather than confirmatory.

6. Handling of Data and Missing Values

The primary endpoint requires paired AMG–EMG values. If either value in a pair is missing at a time point, that pair is excluded from paired analyses but retained for descriptive summaries. No routine imputation is planned. If >10% of recovery-time data are missing and missingness is plausibly at random, a sensitivity analysis using last-observation-carried-forward (LOCF) may be conducted.

7. Statistical Methods

7.1 Primary Endpoint: Mean Paired Difference Between AMG and EMG TOF Ratio

The primary analysis will evaluate the mean paired difference in TOF ratio values between AMG and EMG, based on standardized paired measurements at baseline (before induction) and recovery (after reversal). The mean difference and corresponding 95% confidence interval will be estimated and reported. The hypothesis test will be performed using a paired t-test, or a Wilcoxon signed-rank test if normality assumptions are not met.

Supportive Analyses

Agreement between AMG and EMG will be further characterized using Bland-Altman analysis (mean bias and 95% limits of agreement) and Lin's concordance correlation coefficient (CCC) with 95% confidence intervals. Accuracy relative to the "ideal" TOF value of 1.0 will be evaluated as a supportive analysis in two specific time windows:

- **Baseline** - median of at least 3 stable TOF measurements before rocuronium administration.
- **Recovery -** median of at least 3 TOF measurements after the TOF ratio has reached ≥0.9.

In addition to raw TOF values, a normalized TOF will be calculated as (current TOF / baseline TOF) to account for the known tendency of AMG to yield baseline values greater than 1.0.

Repeated intraoperative paired measurements (post-induction, during surgery, and pre-extubation) will be analyzed as exploratory endpoints to assess phase-related variations.

7.2 Secondary Endpoints

Time-to-event analyses

- **Induction (TOF = 0)**

  Paired times to reach the clinical threshold will be analyzed using the Wilcoxon signed-rank test. Kaplan–Meier curves (AMG vs EMG) will be presented to visualize time distributions and the proportion of patients not yet at the threshold over time. If censoring is non-negligible, a stratified log-rank test (stratified by patient) will be performed as a supportive analysis.

- **Maintenance – duration of deep neuromuscular block (PTC < 5)**

Paired times to block resolution (PTC ≥ 5) will be analyzed using the Wilcoxon signed-rank test. Kaplan–Meier curves will be used for visualization and to estimate the proportion of patients remaining in deep block over time. If censoring is non-negligible, a stratified log-rank test will be applied as supportive.

- **Reversal (TOF ≥ 0.9 after sugammadex)**

  Paired times to reach the clinical threshold will be analyzed using the Wilcoxon signed-rank test. Kaplan–Meier curves will be plotted to visualize time distributions and the proportion of patients still below threshold over time. If censoring is non-negligible, a stratified log-rank test will be used as supportive analysis.

- **Usability and signal quality ratings**: Analyzed as paired ordinal data using the Wilcoxon signed-rank test; effect sizes will also be reported.
- **Postoperative respiratory complications**: Will be presented as absolute counts and percentages**.**
- **Physiological variables (heart rate, mean arterial pressure, SpO_2, perfusion index)**

  1. **Longitudinal description across predefined time points.**

     Results will be reported as mean ± SD or median (IQR) with 95% CIs alongside p-values**,** and within-patient changes will be tested using repeated-measures ANOVA (if approximately normal) or Friedman's test (if non-normal). Post-hoc pairwise comparisons will use Bonferroni/Holm adjustments.

  2. **Event-aligned comparisons at clinical thresholds**

     For each patient, physiological parameters recorded at the AMG trigger will be paired with those recorded at the EMG trigger for the same clinical threshold

     - Induction: values at TOF = 0 (AMG) vs TOF = 0 (EMG)
     - Reversal**:** values at TOF ≥ 0.9 (AMG) vs TOF ≥ 0.9 (EMG)

     This approach allows exploration of whether specific physiological variables (e.g., perfusion index, respiratory and hemodynamic parameters) show systematic differences depending on which monitoring modality reaches the threshold first.

  3. Paired comparisons will use the paired t-test or Wilcoxon signed-rank test, as appropriate.

Operational rule: the measurement at the trigger is the closest available within ±10 s (or within the same minute); otherwise set to missing. Results will be reported as mean (or median) paired differences with 95% confidence intervals, in addition to p-values, to allow assessment of both statistical and clinical relevance.

7.3 Exploratory/Sensitivity Analyses

Mixed-effects models (patient-level random effects) to account for repeated paired measurements with fixed effect for method and time window.

8. Sample Size Justification

Based on Wedemeyer Z. et al., who reported a mean difference of 0.12 (SD 0.13) between AMG and EMG TOF ratios in a non-paired comparison, the required sample size for the present paired study was estimated using a two-sample t-test for the comparison of means, with a significance level of $\alpha = 0.05$ and a statistical power of 90%. The required number of patients is 25. Allowing for an anticipated dropout rate of 16%, a total of 30 patients will be enrolled. This approach is deliberately conservative: it provides protection against higher variability in the obese population, ensures robustness for the primary analysis of paired mean differences, and offers sufficient power for exploratory secondary endpoints. The calculation was performed in R. Unlike Wedemeyer et al., this study evaluates patients with obesity under general anesthesia with paired AMG–EMG measurements in the same subject; repeated measures will be collected but are analyzed as secondary outcomes and were not included in the sample size calculation. In the analysis, the distribution of paired differences will be assessed (Shapiro–Wilk, QQ-plots), and if normality is not satisfied, non-parametric methods (Wilcoxon signed-rank test) will be applied.

9. Software

Analyses will be conducted in R.

10. Tables, Listings, and Figures (TLFs)

TLFs will include: demographics/baseline characteristics; descriptive summaries of TOF measurements by method and time; Bland–Altman plots; concordance coefficients; Kaplan–Meier curves for time to event (induction and reversal); distributions of detection and recovery times; complications summary; and longitudinal plots of physiological parameters.

11. Deviations from SAP

Any deviation from the pre-specified analyses will be documented in the final report with justification and impact assessment.