

**Prevention and rehabilitation of musculoskeletal pain among abdominal and pelvic surgeons with Intelligent Physical Exercise Training (IPET) and Ergonomic Recommendations (ERGO) versus ERGO only: statistical analysis plan**

This statistical analysis plan (SAP) is based on the protocol registered in ClinicalTrials.gov (NCT06112106) and the protocol paper: Christiansen et al. 2025 (1).

The structure and content of the SAP is adopted from the Guidelines for the Content of Statistical Analyses Plans in Clinical Trials (2).

SAP version 1.4, December 15, 2025

# Table of contents

## **1. Introduction 1**

- 1.1. *Background 1*
- 1.2. *Objectives 1*

## **2. Study Methods 1**

- 2.1. *Trial design 1*
- 2.2. *Randomization 1*
- 2.3. *Sample size 2*
- 2.4. *Timing of outcome assessments and final analyses 2*

## **3. Statistical Principles 2**

- 3.1. *Confidence intervals and P-values 2*
- 3.2. *Adherence and analysis population 2*

## **4. Trial Population 3**

- 4.1. *Eligibility 3*
- 4.2. *Recruitment 3*
- 4.3. *Withdrawal and Follow-up 3*
- 4.4. *Baseline participant characteristics 3*

## **5. Analysis 4**

- 5.1. *Outcome definitions 4*
- 5.2. *The primary outcome 4*
- 5.3. *Secondary outcomes 5*
- 5.4. *Analysis methods 7*
- 5.5. *Blind interpretation procedure 7*
- 5.6. *Additional analyses 7*
  - Adjusted analysis 7
  - Per protocol analysis 7
  - Sub-group analysis 8
  - Sensitivity analysis 8
- 5.7. *Quality Control and Reproducibility 9*
  - Data management and verification 9
- 5.8. *Statistical software 9*

*Analyses will be implemented using validated statistical software, specifically Stata (StataCorp LLC, College Station, TX) and/or R (R Foundation for Statistical Computing, Vienna, Austria). The specific software version used will be reported in the final study report 9*

## **References 10**

## 1. Introduction

### 1.1. Background

Surgeons have a high prevalence of musculoskeletal pain which is a significant threat to their career longevity and overall wellbeing. While improving intraoperative ergonomics and raising awareness are important preventive strategies, physical exercise training remains a cornerstone in both primary and secondary prevention of musculoskeletal pain (MSP).

Intelligent Physical Exercise Training (IPET) offers a tailored, evidence-based approach by designing targeted physical exercise training programs based on the location of MSP, job-specific physical demands, individual physical capacity, and overall health. Incorporating IPET into the daily routines of surgeons has the potential to reduce pain, enhance physical resilience, and extend professional longevity.

### 1.2. Objectives

The objective of this trial is to investigate the added effectiveness of Intelligent Physical Exercise Training (IPET) in addition to ergonomic recommendations (ERGO) versus ERGO only, on MSP among abdominal and pelvic surgeons.

The hypothesis is that ERGO alone may prevent or lead to a slight, but not clinically relevant, reduction in MSP, and that the combination of IPET+ERGO will have a synergistic effect resulting in clinically relevant and statistically significant reduction in MSP compared to ERGO alone.

## 2. Study Methods

### 2.1. Trial design

This trial is a 20-week multi-center, open-label randomized, superiority, parallel-group trial with a treatment allocation ratio of 1:1, assessing the added effectiveness of IPET (delivered via the Intelligent Motion app) in addition to ERGO (intervention group), compared to ERGO only (control group). Information about the Intelligent Motion app and ERGO can be found in the protocol article (accepted, <https://doi.org/10.1186/s13063-025-09178-x>).

### 2.2. Randomization

Participants will be randomly allocated to either a control or intervention group with a 1:1 allocation ratio. The allocation follows a computer-generated sequence with permuted and random block sizes of 4, 6 and 8, stratified by pain case status. A pain intensity  $\geq 3$  NRS points (0-10) is considered the threshold for clinically important pain (3), and defines a pain case in this trial.

The allocation sequence list will be prepared prior to study start by a central REDCap data manager not involved in the trial. A member of the research group, blinded to the allocation sequence, will press the “randomize” button in the digital REDCap system. The randomization and allocation process will then be carried out automatically by the REDCap system.

Details about the randomization can be obtained from the REDCap data manager at Dept. of Clinical Medicine, Aarhus University.

### 2.3. Sample size

To detect a 1.5 NRS point difference in worst MSP intensity between the intervention and control groups, a two-sided independent sample t-test was used for the sample size calculation. With a standard deviation of 2.0 NRS points, a power of 80%, and a significance level (alpha) of 0.05, 58 participants in total are required. To account for a drop-out rate of up to 30%, the total sample size is 83 participants. The sample size calculation was performed using Stata (version 17).

See protocol for further details.

### 2.4. Timing of outcome assessments and final analyses

The primary and secondary outcome variables will be assessed at baseline and at 12 and 20 weeks.

All outcomes will be analyzed collectively once the sample size has been reached, or the deadline for data collection has been reached, which is 15 March 2026.

## 3. Statistical Principles

### 3.1. Confidence intervals and *P*-values

The effect estimate will be described as a point estimate with accompanying confidence limits of 95% confidence intervals or an interquartile range. *P*-values will be reported as their actual value. All applicable statistical tests will be 2-sided and will be performed using a 5% significance level.

### 3.2. Adherence and analysis population

Definition of adherence to the intervention: Adherence to the intervention is considered as completing (self-reported indication of completion) an average of minimum 70% of the weekly 50 min. ( $\geq 35$  min.) of IPET (4). No protocol deviations are considered.

Assessment of adherence: Adherence to the intervention is assessed at the 20-week follow-up by asking participants to report the average weekly minutes of IPET performed during the 20-week intervention period. Participants who complete at least 70% of the prescribed 50 minutes per week are considered adherent. In the intervention group, adherence will be reported as the mean weekly minutes completed.

#### Analysis populations

The intention-to-treat population will include all randomized participants with completed baseline questionnaires, regardless of their degree of adherence to the intervention and whether they complete the follow-up questionnaires. Participants who withdraw consent will be excluded and data up to withdrawal will be deleted.

The per protocol analyses will be conducted with the control group and participants from the intervention group who are defined as adhering to the intervention.

Complete case population is defined as the participants with complete primary outcome data at baseline, 12 and 20 weeks follow-up.

## 4. Trial Population

### 4.1. Eligibility

Eligible participants are primary and assisting surgeons who regularly perform abdominal or pelvic surgeries, either via open, conventional laparoscopic or robot-assisted laparoscopic approaches, with an average surgical workload of at least four hours per week over the past six months. Surgeons will be excluded from the trial if their general practitioner advises against participating in physical exercise training.

### 4.2. Recruitment

A CONSORT diagram will outline the recruitment and retention process, detailing the number of surgeons invited to participate, the number who consented, those who met eligibility criteria, and those who were randomized. The number of participants contributing data at each follow-up time point will be shown in the CONSORT diagram, along with the number of participants who withdrew or were lost to follow-up between each time point.

### 4.3. Withdrawal and Follow-up

Levels of withdrawal:

- Participants may discontinue the intervention (i.e., cease IPET, not follow ERGO) but continue with follow-up assessments. These cases will not be reported separately.
- Participants may withdraw from follow-up (i.e., no longer complete questionnaires) but allow previously collected data to be used. These are defined as Loss to follow-up.
- Participants who withdraw from both the intervention and follow-up and additionally withdraw consent for the use of all previously collected data (app and questionnaire) will be reported separately.

### 4.4. Baseline participant characteristics

Baseline characteristics will be described for each randomized group. Variables will include demographics (age, sex), body mass index (BMI), occupational details (surgical specialty, weekly working hours), and health-related variables (use of pain medication, physical activity level, physical resources, workability, general health, burnout, and musculoskeletal pain (MSP) intensity). In addition, the proportion of participants classified as "pain cases" (defined as  $MSP \geq 3$  at baseline) will be reported, as randomization was stratified based on this criterion.

Baseline characteristics will be descriptively summarized depending on the nature of the variables:

- Interval variables will be presented as mean/median values with standard deviation (SD)/interquartile range (IRQ).
- Categorical variables will be presented as numbers and percentages.

No statistical comparisons of baseline characteristics between groups will be conducted.

## 5. Analysis

### 5.1. Outcome definitions

All outcome variables described below are assessed at baseline, 12 weeks, and 20 weeks. The primary follow-up time point is 20 weeks for all the outcome variables.

### 5.2. The primary outcome

The primary effect of interest is the between-group difference in worst MSP intensity at 20 weeks, measured on an 11-point numerical rating scale (NRS 0-10). The worst MSP is defined as the highest reported MSP intensity score at baseline across the ten assessed body parts (neck, upper back, lower back, shoulders, elbows, wrists/hands, hips, knees, ankles/feet).

MSP intensity is collected using a modified version of The Nordic Musculoskeletal Questionnaire (NMQ) (5). For each body part, participants are first asked whether they have had trouble (pain, ache, discomfort) in the past three months, and if so, they are then asked to rate the intensity of that discomfort on the NRS (0 = no discomfort, 10 = maximum discomfort).

Table 1 Primary outcome, assessment method and types of variables

Variable	Assessment	Type of variable
<b><i>Days with MSP past 3 months (secondary outcome)</i></b>	will be assessed by asking: "How many days in total have you had trouble (pain, ache, discomfort) in [body part] in the past three months?", with five response options ranging from "zero days" to "daily".  The question is repeated for each body part, i.e. neck, upper back, lower back, shoulders, elbows, wrists/hands, hips, knees, ankles/feet.  Any number of days with trouble (MSP) leads to the following question.	Categorical (ordinal)
<b><i>MSP intensity past three months</i></b>	Will be assessed by asking: "Indicate the level of discomfort in your [body part] over the past	Continuous

<b>(primary outcome)</b>	<p>three months. A score of 0= no discomfort, 10= maximum discomfort”.</p> <p>The question is repeated for each body part, i.e. neck, upper back, lower back, shoulders, elbows, wrists/hands, hips, knees, ankles/feet.</p>
<b>Worst MSP</b>	<p>Selection of the “worst MSP”:</p> <p>For each participant, the “worst MSP” body part is determined at baseline using the following rules:</p> <ol style="list-style-type: none"> <li>1. The body part with the highest NRS intensity score is selected.</li> <li>2. If two or more body parts have the same intensity, the body part with the longest reported duration of pain is selected.</li> </ol> <p>If both intensity and duration are identical across tied body parts, the body part is selected according to the most prevalent painful body region among Danish surgeons (6).</p>

### 5.3. Secondary outcomes

Below table provides an overview of the secondary outcome variables, how they are being assessed and the type of variable.

*Table 2 Secondary outcomes, assessment method and type of variable*

<b>Variable</b>	<b>Assessment</b>	<b>Type of variable</b>
<b><i>Pain case status</i></b>	Proportions in each group considered a pain case at 20-week follow-up	Count (%)
<b><i>Days with MSP past 3 months (used as a trigger in the primary outcome)</i></b>	<p>will be assessed by asking: ”How many days in total have you had trouble (pain, ache, discomfort) in [body part] in the past three months?”, with five response options ranging from “zero days” to “daily”.</p> <p>The question is repeated for each body part, i.e. neck, upper back, lower back, shoulders, elbows, wrists/hands, hips, knees, ankles/feet.</p> <p>Any number of days with trouble (MSP) leads to the following question.</p>	Categorical (ordinal)

<b><i>MSP intensity past 7 days</i></b>	will be assessed by asking: “Indicate the level of discomfort in your [body part] over the past 7 days. A score of 0= no discomfort, 10= maximum discomfort”	Continuous
	The question is repeated for each body part, i.e. neck, upper back, lower back, shoulders, elbows, wrists/hands, hips, knees, ankles/feet.	
<b><i>Number of days prevented from work (domestic and professional) because of MSP</i></b>	will be assessed by asking: “How many days has trouble prevented you from doing work (at home or away from home) during the past three months?” with five response options ranging from “zero days” to “daily”.	Categorical (ordinal)
	The question is repeated for each body part, i.e. neck, upper back, lower back, shoulders, elbows, wrists/hands, hips, knees, ankles/feet.	
<b><i>Use of pain medication</i></b>	is assessed by asking: “How many days during the past three months have you taken pain medication for MSP?” with four response options ranging from “rarely or never” to “daily”.	Categorical (ordinal)

#### 5.4. Analysis methods

The primary outcome variables are worst MSP at 12 and 20 weeks. The primary analysis will be conducted according to the intention-to-treat principle using a linear mixed model, with MSP at weeks 12 and 20 forming the dependent variable. The model will include time (12 weeks, 20 weeks), group allocation (intervention vs. control), and a time×group allocation interaction as fixed effects with baseline MSP intensity value as covariate. Random intercepts for participants and randomization blocks will be included to account for within-person correlation and clustering induced by block randomization. The primary effect of interest is the treatment-group difference at 20 weeks. Results will be presented with corresponding 95% confidence intervals. A two-sided significance level of 0.05 will be used for all statistical tests. Missing data will be handled under the missing-at-random assumption, inherent to the mixed-model approach, and the extent of missingness will be described by group and key baseline characteristics.

For categorical outcomes, differences between groups will be assessed using either the chi-squared test or the Mann–Whitney U test, depending on the nature and distribution of the data.

Model assumptions will be assessed by examining residual plots for linearity, homoscedasticity, and influential observations. Normality of residuals will be evaluated using Q-Q plots. If applicable, residual independence will be assessed.

If violations are detected, data transformations or alternative modeling approaches will be considered.

#### 5.5. Blind interpretation procedure

To reduce the risk of biased interpretation of results, the following procedure will be used: Two interpretations will be drafted based on the results from the main analyses, with treatment groups arbitrarily labelled as A and B. One interpretation assumes that A is the intervention group, and B is the control group, and the other interpretation assumes that A is the control group and B is the intervention group. After agreeing on both interpretations, the randomization code is broken, and the correct interpretation will be used.

#### 5.6. Additional analyses

The following additional analyses are planned for:

##### Adjusted analysis

The primary analysis will be repeated with additional adjustment for baseline levels of potentially important prognostic factors, including age, sex, BMI, weekly working hours, primary surgical approach, pain duration, pain intensity in other body parts, multi-site pain, fitness level, burnout, and the extent to which participants applied the ERGO principles.

##### Per protocol analysis

The primary analysis will be repeated in the per-protocol population, including the control group and participants from the intervention group who achieved  $\geq 70\%$  adherence to the

prescribed IPET sessions (corresponding to an average of  $\geq 35$  minutes of training per week, or  $\geq 70\%$  of the total prescribed 50 minutes per week).

This analysis aims to estimate the effect of the intervention under conditions of adequate adherence.

#### Sub-group analysis

Subgroup and interaction analyses will explore whether the effect of the intervention (IPET + ERGO) compared to control (ERGO only) on MSP intensity differs across predefined participant characteristics. These analyses are exploratory and hypothesis-generating.

The following subgroups will be examined for potential effect modification on the primary outcome (worst MSP intensity at 20 weeks):

Sub-group	Variable
Baseline pain case status	Generated: Pain cases = NRS $\geq 3$ , Non-pain case $< 3$
Seniority	Age as a proxy
Country	Denmark vs. United States
Sex	Female vs. Male
Surgical specialty	Gynecology / colorectal / urology
Baseline physical activity level	Low / moderate / high (Saltin–Grimby Physical Activity Scale) Weekly minutes of moderate and vigorous physical activity
Physical resources compared to peers	Generated: Above average compared to peers = $\geq 5$ , or below average compared to peers $< 5$ on a 11-point ordinal NRS scale from 0-10 (0= weak/poor, 10= strong/good, Strøyer self-reported physical resources scale)

#### Sensitivity analysis

To assess the robustness of the primary outcome to the model-imputed missing data, a sensitivity analysis will be performed comparing results from the intention-to-treat analysis to the complete cases only.

No interim analysis or a priori defined stopping rules are defined or implemented for this trial.

## 5.7. Quality Control and Reproducibility

### Data management and verification

All datasets will be exported from REDCap using standardized export scripts to ensure consistent variable naming and data structure. Before analysis, the primary investigator will perform a structured data verification procedure, including:

- Checking for duplicates, range errors, and internal inconsistencies (e.g., negative time stamps, implausible values).
- Independent verification of key derived variables (e.g., index (worst MSP) body part, pain case status, BMI).
- Any detected discrepancies will be resolved through a double-checking process documented in a data query log (date, issue, correction, responsible person).

## 5.8. Statistical software

Analyses will be implemented using validated statistical software, specifically Stata (StataCorp LLC, College Station, TX) and/or R (R Foundation for Statistical Computing, Vienna, Austria). The specific software version used will be reported in the final study report

## References

1. Christiansen HJ, Sandal LF, Mogensen O, Norasi H, Chrouser K, Hallbeck MS, et al. Prevention and rehabilitation of musculoskeletal pain among abdominal and pelvic surgeons with intelligent physical exercise training (IPET) and intraoperative ergonomic recommendations (ERGO): study protocol for a multicenter open-label randomized controlled trial in Denmark and North America (USA). *Trials*. 2025;26(1):565.
2. Gamble C, Krishan A, Stocken D, Lewis S, Juszcak E, Dore C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017;318(23):2337-43.
3. Kaergaard A, Andersen JH, Rasmussen K, Mikkelsen S. Identification of neck-shoulder disorders in a 1 year follow-up study. Validation of a questionnaire-based method. *Pain (Amsterdam)*. 2000;86(3):305-10.
4. Sjøgaard G, Sjøgaard K, Hansen AF, Østergaard AS, Teljigovic S, Dalager T. Exercise Prescription for the Work-Life Population and Beyond. *J Funct Morphol Kinesiol*. 2023;8(2).
5. Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sørensen F, Andersson G, et al. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. *Applied ergonomics*. 1987;18(3):233-7.
6. Dalager T, Sogaard K, Boyle E, Jensen PT, Mogensen O. Surgery Is Physically Demanding and Associated With Multisite Musculoskeletal Pain: A Cross-Sectional Study. *J Surg Res*. 2019;240:30-9.