STATISTICAL ANALYSIS PLAN

Version 1.0 — 18 March 2025

A preparatory lifestyle-based physical activity intervention before intensive rehabilitation for chronic low back pain: a controlled clinical study

Protocol identifier	PFRP-FRP-CHB-2023
ClinicalTrials.gov	NCT[to be inserted after PRS QC review]
Ethics approval	CERNI 2025-01-07 / 2025-51
	Université de Picardie Jules Verne
Sponsor	Université d'Artois
Principal investigator	Steven Crombecque, PhD
Affiliated laboratory	SHERPAS, URePSSS (ULR 7369)
Study site	Centre Hélène Borel, Raimbeaucourt, France
Document version	1.0
Document date	18 March 2025

Preamble

This Statistical Analysis Plan (SAP) documents the analytical strategy applied to the controlled clinical study comparing a 12-week preparatory lifestyle-based physical activity program (PFRP) followed by an intensive Functional Restoration Program (FRP) against an intensive FRP alone, in patients with chronic low back pain. The study protocol was approved by the Comité d'Éthique pour les Recherches Non-Interventionnelles (CERNI) of Université de Picardie Jules Verne under reference 2025-01-07/2025-51, as part of a broader doctoral research programme covering three related studies on the optimisation of pre- and post-rehabilitation care pathways for chronic low back pain patients. The present SAP concerns Study 2 of that protocol.

This SAP documents the analytical strategy planned a priori for Study 2, in accordance with the CERNI-approved protocol. The outcomes, sample size rationale, and analyses described herein were defined prior to data analysis and were subsequently applied as documented in the resulting doctoral thesis (defended December 2025). No outcome was selected on the basis of observed results.

1. Study objectives and hypotheses

1.1 Primary objective

To evaluate whether a 12-week autonomy-focused lifestyle intervention (PFRP) delivered before intensive multidisciplinary rehabilitation (FRP) prevents behavioural deterioration and improves functional outcomes during the waiting period before FRP entry, and to compare trajectories between groups across the full care pathway.

1.2 Hypotheses

The primary hypothesis is that a 12-week preparatory phase grounded in autonomy-based adapted physical activity, delivered before intensive rehabilitation, will (i) prevent the deterioration of sedentary behaviour and functional capacity observed during unstructured waiting periods, and (ii) support better medium-term maintenance of behavioural and functional gains after intensive rehabilitation completion.

2. Study design

Prospective, controlled, non-randomised, single-centre clinical study conducted at Centre Hélène Borel (Raimbeaucourt, France). Two parallel groups of patients with chronic low back pain were assigned to either PFRP followed by FRP (experimental group) or FRP alone after an equivalent waiting period (active comparator group). Group assignment was determined by admission scheduling and the residential programme's organisational constraints (groups of 2 to 6 patients per admission cycle), which precluded randomisation.

Identical inclusion and exclusion criteria were applied to both groups to limit selection bias. Baseline comparability across all 28 measured variables was prospectively verified before any between-group inferential analysis. Recruitment took place between March 2023 and February 2024.

3. Population

3.1 Inclusion criteria

- Chronic low back pain (duration > 3 months)
- Previous treatment failure (e.g., ineffective physical therapy)
- For participants over 40 years of age: prior cardiovascular stress test to rule out cardiovascular risk
- Aged 18 to 65 years
- Written informed consent obtained prior to inclusion

3.2 Exclusion criteria

- Spinal or related surgery within the six months prior to programme entry
- Under 18 or over 65 years of age
- Previous participation in a similar multidisciplinary rehabilitation programme

4. Sample size

Sample size was estimated a priori using G*Power software (version 3.1.9.7) for a repeated-measures ANOVA within-between interaction with the following assumptions: 2 groups, 4 measurement timepoints, medium effect size ($f = 0.25$), $\alpha = 0.05$, statistical power = 0.70, correlation among repeated measures = 0.50, and nonsphericity correction $\varepsilon = 0.90$. The minimum required sample size was 22 participants. Anticipating attrition, the planned target was 24 participants, distributed across PFRP and FRP groups according to admission scheduling.

Given the resulting small sample size and the likely non-normal distribution of repeated measures in this clinical population, non-parametric tests were planned for the inferential analyses regardless of the initial parametric framework used for sample size estimation.

5. Outcomes

5.1 Co-primary outcomes

Two co-primary outcomes were defined a priori based on the theoretical framework of the PFRP intervention, which emphasises autonomy development and behavioural change:

(1) **Sedentary behaviour (ONAPS-SED)** — measured by the Sedentary subscale of the French ONAPS Physical Activity Questionnaire (ONAPS-PAQ), expressed in minutes per week. This subscale was selected as a direct measure of the behavioural change targeted by the lifestyle-based intervention.

(2) **Functional impact on Daily Activities (DALLAS-DA)** — measured by the Daily Activities subscale of the DALLAS Pain Questionnaire, expressed as a percentage (0% to 100%). This subscale reflects functional autonomy in everyday life.

The ONAPS questionnaire was not administered at T2 because intensive hospitalisation is not representative of habitual physical activity behaviour. ONAPS-SED analyses therefore focused on the

T0–T1 (pre-programme / waiting phase) and T0–T3 (overall) periods. DALLAS-DA was administered at all four timepoints.

5.2 Measurement timepoints

- T0 — baseline, at the start of the PFRP (experimental group) or of the equivalent waiting period (comparator group), approximately three months before FRP entry.
- T1 — at FRP entry / end of PFRP phase (week 12).
- T2 — at FRP exit (week 16).
- T3 — four-month follow-up after FRP completion (approximately week 32).

5.3 Secondary outcomes

Secondary outcomes were organised a priori by their hypothesised role in the intervention's mechanism of action.

Other dimensions of the primary constructs.

- Remaining DALLAS subscales: Work/Leisure Activities (WLA), Anxiety/Depression (A/D), Social interest (SOC).
- Other ONAPS subscales: total physical activity (TOT), Moderate-to-Vigorous Physical Activity Intensity (MVPAI).
- EIFEL questionnaire (French validated version of the Roland-Morris Disability Questionnaire), for functional disability comparison with the literature.

Potential cognitive mediators.

- Fear-Avoidance Beliefs Questionnaire (FABQ): subscales FABQ-PA (Physical Activity) and FABQ-W (Work).

Physical fitness measures, considered prerequisites for behavioural change.

- Trunk muscular endurance: Sorensen test (extensors), Ito-Shirado test (flexors), bilateral lateral plank, Killy test.
- Mobility: finger-to-floor distance, bilateral lateral flexion, bilateral trunk rotations, bilateral hamstring flexibility.
- Isokinetic trunk flexor/extensor strength at 30°/s and 90°/s on a Contrex® dynamometer; the flexor/extensor (F/E) ratio averaged across both speeds was used for analyses.

Distal outcomes.

- Pain intensity: Visual Analogue Scale (VAS), 0–10 cm.
- Psychological well-being: General Health Questionnaire 12-item (GHQ-12).

All secondary outcomes were administered at all four timepoints (T0, T1, T2, T3), with the exception of ONAPS subscales as specified above.

6. Statistical methods

6.1 Descriptive statistics

Continuous variables are summarised as mean ± standard deviation when the distribution was approximately symmetrical, and as median with interquartile range when skewed. Categorical variables are summarised as frequencies and percentages.

6.2 Baseline comparability

Between-group baseline comparability across all measured variables was assessed using Mann-Whitney U tests for continuous variables and Fisher's exact test for categorical variables. Comparability was considered acceptable when all between-group baseline comparisons yielded $p > 0.05$. This assessment was performed before any inferential analysis of treatment effect.

6.3 Within-group changes over time

Within-group changes between consecutive timepoints (T0–T1, T1–T2, T2–T3) and across the full study period (T0–T3) were assessed using Wilcoxon signed-rank tests applied to paired observations.

6.4 Between-group comparisons (Group × Time interactions)

Group × Time interactions were assessed by computing change scores (Δ) for each participant between the relevant timepoints, then comparing the distributions of these change scores between groups using Mann-Whitney U tests. This approach is the non-parametric analogue of the interaction term in a two-way repeated-measures ANOVA and was preferred over rank-based ANOVA-type statistics given the small sample size and the focus on phase-specific effects. Between-group comparisons of absolute values at each timepoint were performed using Mann-Whitney U tests.

6.5 Effect sizes

For all between-group comparisons, effect sizes were reported using rank-biserial correlation (RBC), computed from the Mann-Whitney U statistic. Interpretation followed conventional thresholds: $|RBC| \approx 0.10$ small effect, 0.30 medium effect, 0.50 large effect.

Given the exploratory nature of the study and the small sample size limiting statistical power, effect sizes are emphasised in the interpretation of results alongside p-values, to allow assessment of the magnitude of effects independently of statistical significance.

6.6 Multiple comparisons correction

The Benjamini-Hochberg procedure was applied to control the false discovery rate (FDR) at $\alpha = 0.05$, separately within outcome families (co-primary outcomes, DALLAS subscales, ONAPS subscales, FABQ subscales, physical fitness measures, distal outcomes). Both uncorrected and FDR-corrected p-values are reported to support transparent assessment of robustness.

7. Handling of missing data

Missing data were handled following intention-to-treat principles using Last Observation Carried Forward (LOCF). The expected proportion of missing values across the dataset was anticipated to be

low (< 5%) given the structured collection schedule and the supervised setting. Sensitivity analyses on complete cases were considered if missingness exceeded the anticipated threshold.

8. Significance level

Statistical significance was set at $\alpha = 0.05$, two-sided, for all primary and secondary inferential analyses. P-values are reported to three decimal places (or as "< 0.001" when smaller). Both uncorrected and FDR-corrected p-values are presented for each outcome family.

9. Software

All analyses were performed using Python 3.12.2, with the following libraries: pandas (data management), scipy (non-parametric tests), pingouin (effect sizes and additional statistical functions), statsmodels (where supplementary regression-based sensitivity analyses were considered), and matplotlib / seaborn (graphical reporting). Sample size calculation was performed in G*Power 3.1.9.7.

10. Reporting

Reporting followed the TREND statement (Transparent Reporting of Evaluations with Nonrandomized Designs), given the controlled non-randomised design of the study.

11. Additional considerations

11.1 Relationship to other studies under the same CERNI protocol

The CERNI approval (2025-01-07/2025-51) covers a broader doctoral research programme encompassing three related studies on chronic low back pain rehabilitation. A subset of participants in the present Study 2 also contributed to Study 1 of that protocol, which addressed a distinct research question on modalities of supervision within multidisciplinary rehabilitation and was published as a research letter in Annals of Physical and Rehabilitation Medicine (Crombecque et al., 2026; PMID 42035666). No outcome reported in the present analysis is duplicated from that prior publication; outcomes, analytical timepoints, and research questions differ between the two studies.

11.2 Data protection

The study was conducted in compliance with the French Loi Informatique et Libertés and the European General Data Protection Regulation (GDPR), under engagement of compliance with CNIL reference methodology MR-004 (research not involving the human person, studies and evaluations in the health domain). Data were pseudonymised during collection and anonymised prior to analysis. Anonymised data are stored at the SHERPAS laboratory for a maximum duration of 10 years following CERNI provisions.

— End of Statistical Analysis Plan —