# STATISTICAL ANALYSIS PLAN

**Analysis populations.**

The primary analysis population includes all eligible athletes who completed baseline and at least one follow-up physiological assessment ($VO_2$max, lactate threshold, Wingate, CMJ, HRR, or HRV). A secondary analysis population includes all athletes who provided daily perceptual load data (sRPE, readiness) for at least 80% of sessions. A safety population includes all participants exposed to the prescribed training program.

**Primary endpoint analysis.**

The primary endpoints are changes in $VO_2$max, lactate threshold, Wingate peak power, CMJ height, HRR-1, and HRV RMSSD from baseline (week 0) to post-phase (week 24). Because the study is observational and within-subject, analyses will use repeated-measures linear mixed-effects models with fixed effects for time (baseline, week 12, week 24) and random intercepts for each athlete. Robust standard errors will be clustered at the participant level. Primary inference will be based on time contrasts (baseline→week 12, baseline→week 24).

**Secondary endpoints.**

- Continuous secondary endpoints (sRPE, readiness, PLR, PRI, biomechanical asymmetry index) will be analysed using mixed-effects linear models with time as a categorical predictor.
- HRV suppression and lag coefficients will be modelled using generalized estimating equations (GEE) with an exchangeable correlation structure.
- Compliance and monotony indices will be summarised descriptively and tested using repeated-measures ANOVA or equivalent mixed models.
- Associations between training load (external and internal) and physiological outcomes will be assessed with multivariable linear regression models including covariates (age, sex, baseline $VO_2$max).

**Missing data.**

For physiological outcomes missing due to drop-out or incomplete testing, multiple imputation by chained equations (MICE) will be used under missing-at-random assumptions, generating 20 imputed datasets combined with Rubin's rules. Sensitivity analyses will include last-observation-carried-forward and complete-case analysis. For daily perceptual data, missingness <10% will be tolerated; otherwise, time-series imputation with Kalman smoothing will be applied.

**Subgroups and interactions.**

Exploratory subgroup analyses will compare male vs female athletes and high vs low baseline $VO_2$max (median split). Interaction terms (time×sex, time×baseline $VO_2$max) will be included in mixed-effects models but interpreted cautiously as exploratory.

**Sensitivity analyses.**

Analyses will be repeated excluding athletes with <85% training compliance. PRI outcomes will be re-analysed excluding taper-week data to test robustness. To address potential non-normality, outcomes will also be modelled with rank-based methods.

**Multiplicity and inference.**

The primary endpoints will be tested with a two-sided alpha of 0.05, applying Bonferroni-Holm correction across the six physiological outcomes. Secondary endpoints will be considered exploratory with false discovery rate (FDR) control where appropriate. Effect estimates will be reported with 95% confidence intervals and standardized effect sizes (Cohen's d).

**Software and reporting.**

Analyses will be performed in **R (version 4.3+)** and **Python (version 3.11+)** using validated packages (lme4, nlme, mice, statsmodels). A reproducible analysis log and code repository will be archived with the final manuscript.