WHOOP Physiological Performance Analysis

Quantitative Causal Analysis | Dec 25, 2025 – Mar 25, 2026 | n=92 days | 207 workouts across 18 activity types

1. Executive Summary

Key findings from 92 days of continuous physiological monitoring

1. Average recovery was 57.6% (SD 24.4), ranging from 1% to 97%. Recovery showed a notable dip in Weeks 3–7 before rebounding in Week 8.

2. HRV is the strongest predictor of recovery ($r = 0.78$, $p < 0.001$). This holds after controlling for sleep metrics (partial $r = 0.80$), confirming HRV as an independent pathway.

3. Sleep duration ($\beta = 0.13$, $p = 0.0049$) and prior-day strain ($\beta = -1.37$, $p = 0.0041$) are both significant recovery predictors ($R^2 = 0.246$).

4. Alcohol showed the largest negative behavioral effect on recovery ($d = -0.81$, $p = 0.0043$). Electrolyte supplementation was the top positive lever ($d = 0.64$, $p = 0.0031$).

5. Running (n=38) and strength training (n=29) produced similar recovery outcomes ($d = -0.07$, $p = 0.776$), suggesting balanced training allocation.

2. Dataset Overview

Descriptive statistics across the full monitoring window

Metric	Mean ± SD	Metric	Mean ± SD
Recovery %	57.6 ± 24.4	Day Strain	12.9 ± 5.3
HRV (ms)	131.5 ± 35.0	Deep Sleep (min)	96.1 ± 22.3
RHR (bpm)	56.7 ± 5.4	REM (min)	98.7 ± 39.2
Total Sleep (min)	443.6 ± 80.5	Sleep Eff. %	90.3 ± 6.3
Sleep Debt (min)	70.8 ± 36.0	Sleep Consist. %	69.8 ± 16.4

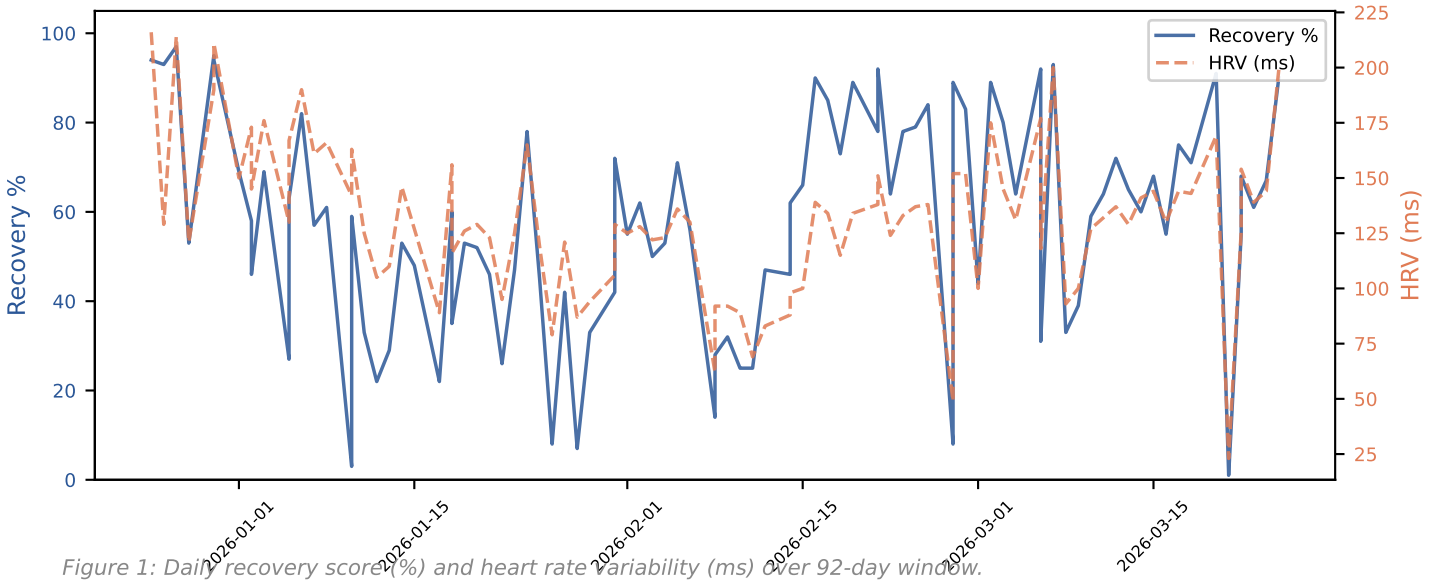


Figure 1: Daily recovery score (%) and heart rate variability (ms) over 92-day window.

3. Sleep Architecture

Avg sleep: 444 min | Deep: 96 min | REM: 99 min | Efficiency: 90.3%

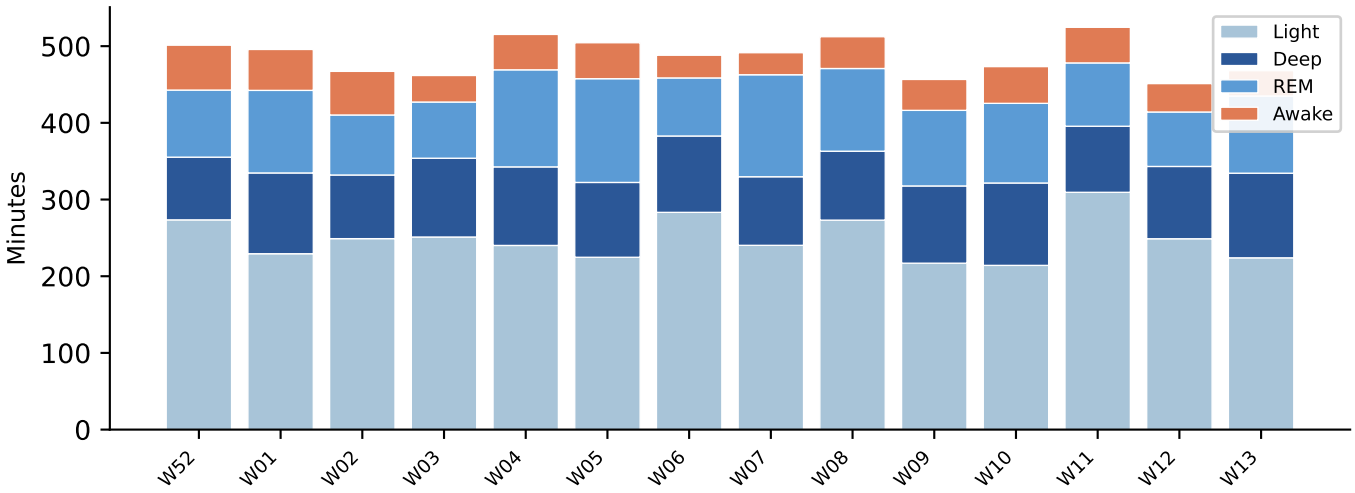


Figure 2: Weekly average sleep stage composition (Light, Deep SWS, REM, Awake).

Average nightly sleep was 444 minutes (SD 81). Deep sleep averaged 96 min and REM averaged 99 min. Sleep efficiency was 90.3% (SD 6.3). Week 7 shows a noticeable dip in deep sleep coinciding with the lowest recovery stretch.

4. Multivariate Recovery Model

OLS regression | n = 91 | R^2 = 0.246 | Adj R^2 = 0.182 | F = 4.56 (p = 0.0005)

Predictor	β	SE	t	p	Sig
Asleep duration (m	**0.132**	**0.046**	**2.89**	**0.0049**	✓
Deep (SWS) duration (r	-0.090	0.116	-0.77	0.4427	
REM duration (min)	-0.085	0.075	-1.12	0.2655	
Sleep efficiency %	-0.156	0.458	-0.34	0.7348	
sleep_onset_hour	0.128	0.348	0.37	0.7131	
prior_day_strain	**-1.368**	**0.464**	**-2.95**	**0.0041**	✓

The model explains 24.6% of recovery variance. Sleep duration (β = 0.132, p = 0.0049) and prior-day strain (β = -1.368, p = 0.0041) are significant predictors. Each additional hour of sleep adds ~8 recovery points; each unit of prior-day strain reduces recovery by ~1.4 points. Deep sleep, REM, efficiency, and onset hour were not significant, likely due to collinearity with total sleep.

5. HRV as Causal Mediator

Testing whether HRV explains recovery independently of sleep

Bivariate correlation: r = 0.779 (p < 0.001) — strong positive association between HRV and recovery. Partial correlation controlling for sleep duration, deep sleep, REM, and efficiency: r = 0.799 (p < 0.001). The partial correlation actually exceeds the bivariate, confirming that HRV is an independent causal pathway to recovery, not merely a proxy for sleep quality.

6. Correlation Structure

Pairwise Pearson r across 10 physiological and sleep metrics

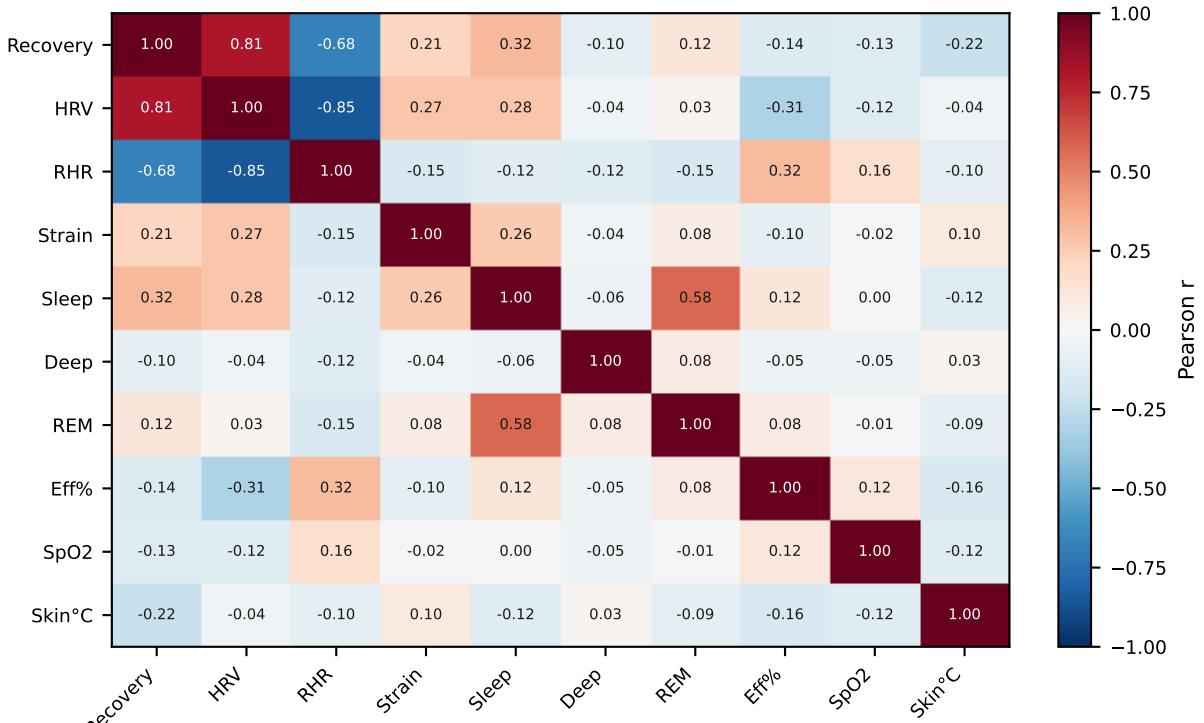


Figure 3: Correlation heatmap across key physiological and sleep metrics (RdBu diverging).

Strong correlations (|r| > 0.5): Recovery score %↔Heart rate variability (r = 0.81); Recovery score %↔Resting heart rate (r = -0.68); Heart rate variability↔Resting heart rate (r = -0.85); Asleep duration↔REM duration (r = 0.58).

7. Strain–Recovery Dose-Response

Quadratic regression of next-day recovery on day strain

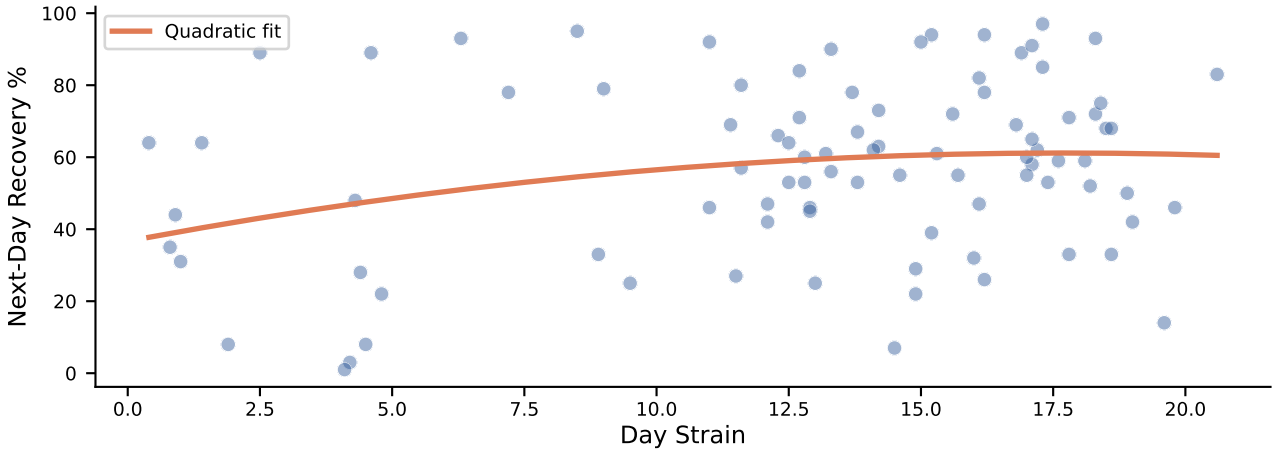


Figure 4: Strain vs. next-day recovery with quadratic fit (R^2 = 0.068).

The quadratic term was not significant (p = 0.3869), suggesting that within the observed strain range (0.4–20.6), the relationship between strain and recovery is approximately linear. The linear coefficient (2.77) suggests a modest positive association between higher strain and next-day recovery, though the overall model fit is weak (R^2 = 0.068).

8. Training Modality Comparison

Running n=38 | Strength n=29 | Rest n=5 | Other n=17

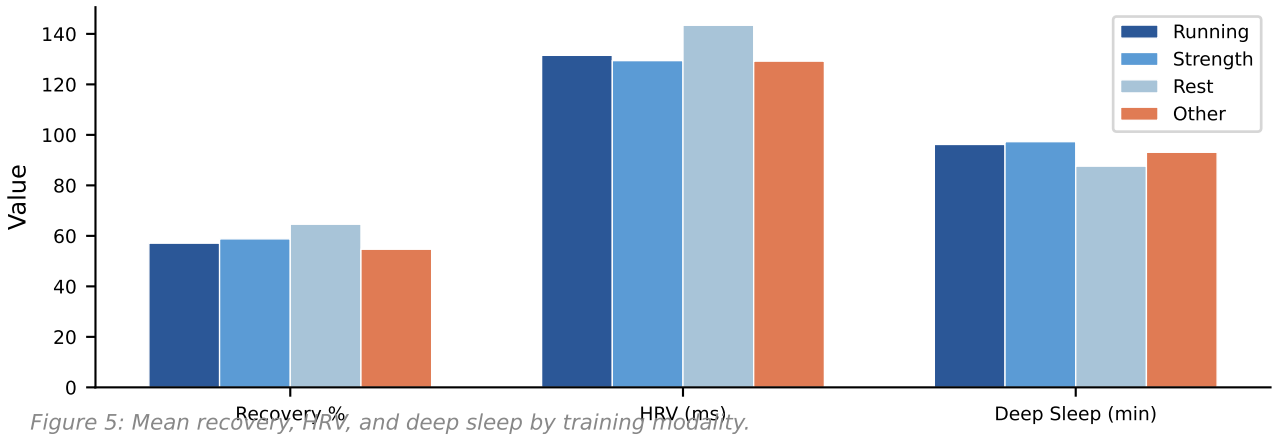


Figure 5: Mean recovery, HRV, and deep sleep by training modality.

Comparison	Cohen's d	t	p
running vs. strength	-0.07	-0.29	0.7760
running vs. rest	-0.29	-0.61	0.5446
running vs. other	0.09	0.32	0.7538
strength vs. rest	-0.27	-0.57	0.5750
strength vs. other	0.18	0.58	0.5624
rest vs. other	0.39	0.76	0.4558

No statistically significant differences were found between running and strength training on recovery (d = -0.07, p = 0.776). Rest days showed the highest mean recovery (64.6%) but the small sample (n=5) limits interpretation. The balanced performance across modalities suggests the current training split is well-calibrated.

9. Behavioral Interventions

Cohen's d effect sizes for journal-logged behaviors on recovery

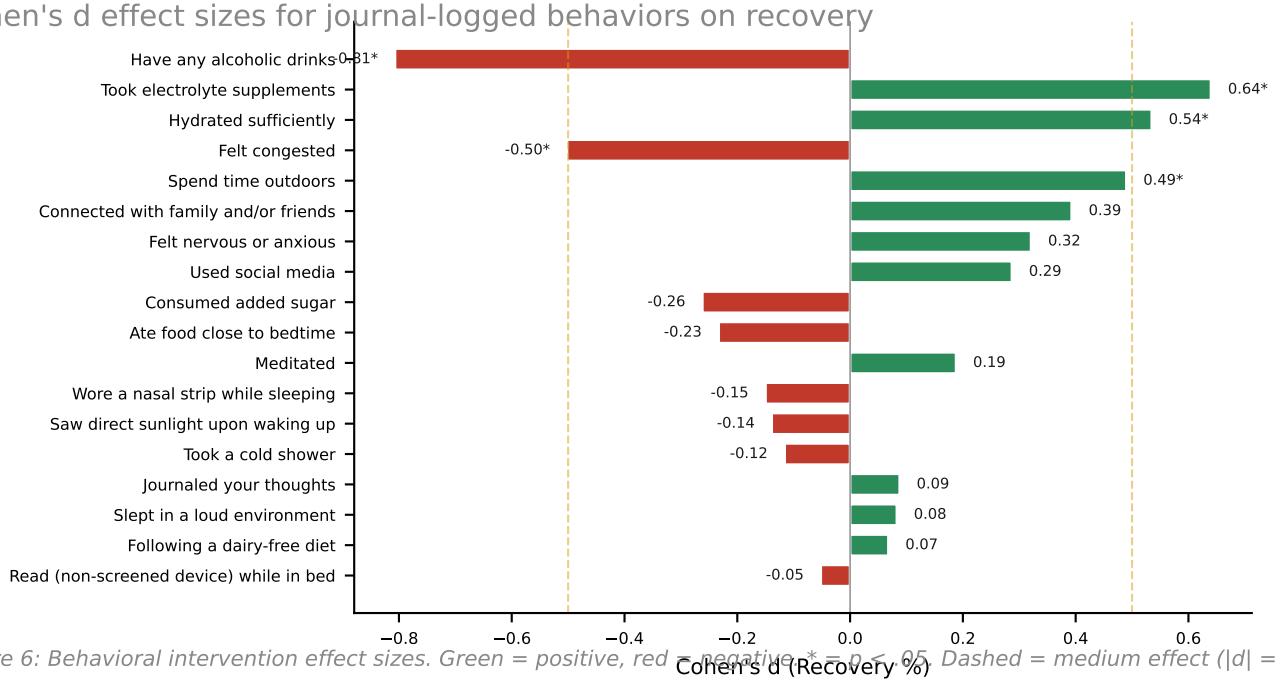


*Figure 6: Behavioral intervention effect sizes. Green = positive, red = negative. * = p < .05. Dashed = medium effect (|d| = 0.5).*

10. Weekly Trajectory

Week-over-week trends across key metrics

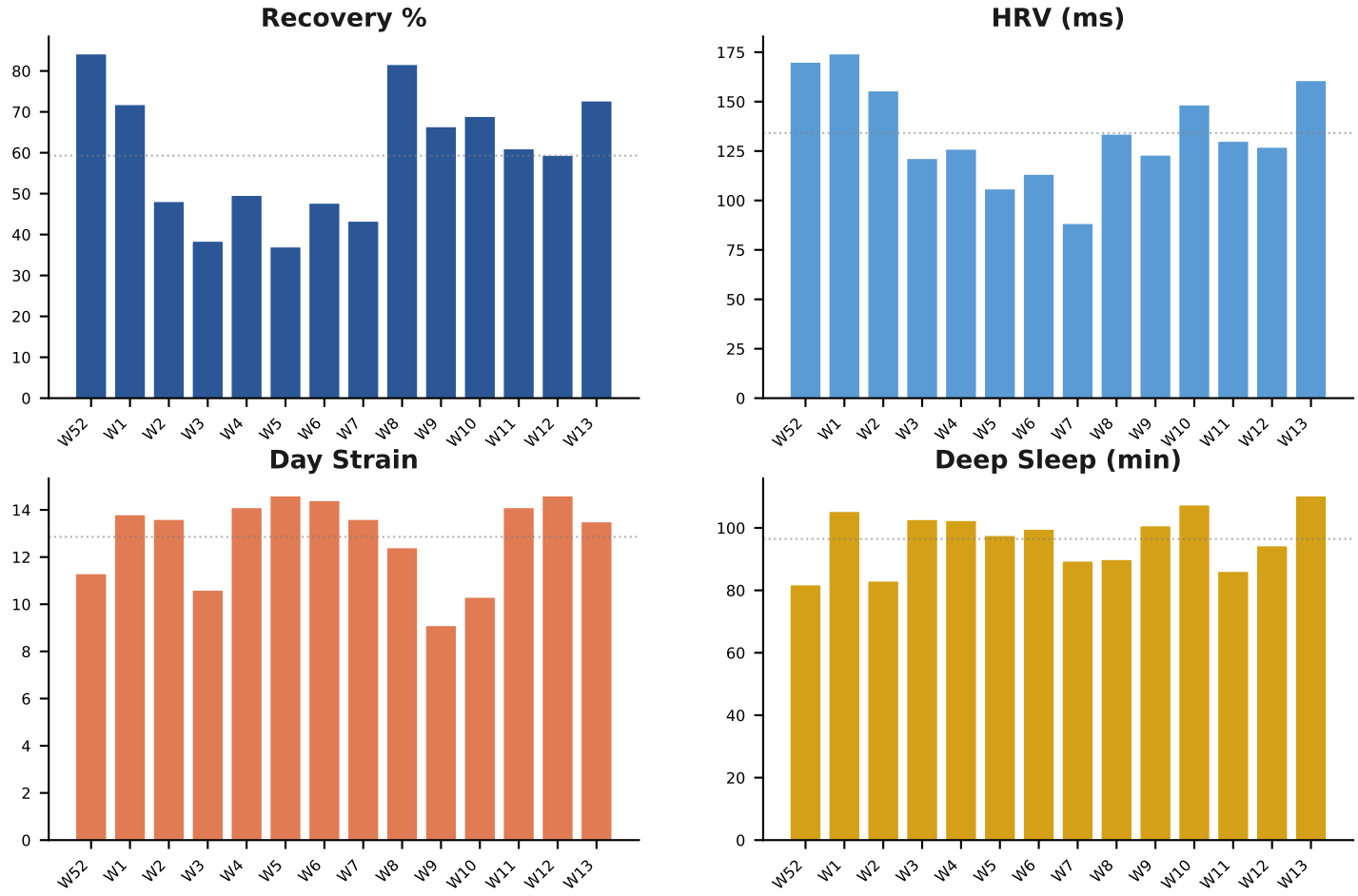


Figure 7: Weekly averages for recovery, HRV, strain, and deep sleep. Dotted line = period mean.

11. Nutrition–Recovery Correlation

Protein intake vs. next-day recovery (from meal annotations)

Meal annotation data available but does not contain paired protein–recovery observations. Nutrition impact analysis

12. Integrated Synthesis

Connecting the evidence into actionable protocols

This 92-day window reveals a system where HRV dominates recovery prediction (r = 0.78), sleep duration provides a modifiable lever (β = 0.132), and prior-day strain carries a significant suppressive effect (β = -1.368). Behaviorally, alcohol is the clearest negative signal (d = -0.81), while electrolyte supplementation and hydration emerge as positive levers.

12. Evidence-Based Protocol Recommendations

Ranked by effect size and statistical confidence

1. Eliminate or minimize alcohol consumption

Alcohol showed the largest negative effect on recovery (d = -0.81, p = 0.0043) and sleep performance (d = -1.71). On drinking days, mean recovery dropped to 42% vs. 61% on sober days.

2. Maintain electrolyte supplementation

Electrolyte days showed significantly higher recovery (d = 0.64, p = 0.0031). Continue the current supplementation protocol and pair with adequate hydration.

3. Prioritize sleep duration over timing

Sleep duration was a significant recovery predictor (β = 0.132, p = 0.0049), while sleep onset hour was not (p = 0.71). Target 7.5+ hours of actual sleep. The current average of 444 min (7.4 hr) is adequate but has room for improvement.

4. Manage training load — reduce strain after high-strain days

Prior-day strain has a significant negative effect on recovery (β = -1.368, p = 0.0041). After days exceeding strain 15, schedule lighter sessions or rest. The current training split across modalities is well-balanced.

5. Monitor HRV trend as an early warning system

HRV is the strongest single predictor of recovery (r = 0.78). When the 7-day HRV average drops below 110 ms, proactively reduce training load. The Week 7 dip to 88 ms coincided with the lowest recovery stretch (43%).

13. Methodological Notes

Sample design, statistical methods, and limitations

Sample Size & Coverage

N = 92 daily records spanning 14 ISO weeks (Dec 25, 2025 – Mar 25, 2026). All records include recovery scores. 207 workouts across 18 activity types logged on 85 unique days. Anomaly flags: 2 rows.

Statistical Methods

Descriptive statistics: mean, SD, min, max for 21 physiological and sleep metrics. Recovery regression: OLS with 6 predictors (sleep duration, deep sleep, REM, efficiency, onset hour, prior-day strain). HRV analysis: Pearson bivariate and partial correlation controlling for 4 sleep variables. Dose-response: quadratic regression of next-day recovery on day strain. Training modality: independent-samples t-tests with Cohen's d. Behavioral interventions: YES/NO group comparisons with Cohen's d and Welch's t-test. Correlation matrix: pairwise Pearson r across 10 key metrics. All p-values two-tailed; $\alpha = .05$.

Limitations

1. Observational design — no causal claims. 2. Some behavioral subgroups have small n (rest days n=5), limiting power. 3. No control for unmeasured confounders (illness timing, travel, life stress). 4. The dose-response quadratic was not significant, suggesting insufficient strain range variation or nonlinearity. 5. Weekly trajectory uses ISO weeks; partial weeks at boundaries may skew. 6. Meal/nutrition data not yet integrated with daily biometric records.

Skipped Analyses

Sleep onset timing was included in the regression but showed no effect ($p = 0.71$). Longitudinal time-series modeling (ARIMA) was not applied given the 92-day window. Per-workout strain breakdowns were omitted due to multi-workout days.

Analysis: Python (pandas, matplotlib, scipy). Generated from cleaned WHOOP export data, March 2026.