Occupational Health Toolkit – Statistical Report for Supervisor

February 1, 2026

Purpose

This report documents the statistical workflow, model choices, transformations, assumption checks, multiple-comparison corrections, and results for hypotheses H1–H6. It is a rigorous methods + results + interpretation briefing (not a paper) intended to justify why each model and adjustment was chosen and to summarize outputs with diagnostics.

Study Design and Data Structure

Dataset Summary

- Population: 38 office workers (Front-Office vs Back-Office)
- Repeated measures: daily observations per subject (`subject_id`)
- Key outcomes: EMG trapezius activity, perceived workload, sitting behavior, postural sway, and self-report sitting (OSPAQ)
- Modeling strategy: Linear Mixed Models (LMMs) for repeated measures; OLS for subject-level validation (H4)
- Repeated-measures implication: within-subject correlation requires random effects to avoid inflated type-I error

Hypotheses and Models (Confirmatory vs Exploratory)

All confirmatory models use ML estimation for valid likelihood ratio tests (LRT). Day effects are categorical (`C(day_index)`), avoiding linear-trend assumptions.

Hypothesis	Outcome	Model	Formula (fixed effects)	Notes
H1 (Confirmatory)	EMG p90 (%MVC)	LMM	log(EMG p90) ~ work_type + C(day_index)	Log transform for skew/heteroscedasticity; EMG p90 not bounded in [0,1]
H2 (Confirmatory)	Workload mean	LMM	workload_mean ~ work_type + C(day_index)	No transform
H3 (Confirmatory)	Sitting proportion	LMM	logit(har_sentado_prop) ~ workload_mean + work_type + C(day_index)	Proportion -> logit
H4 (Confirmatory)	OSPAQ validation	OLS	logit(har_sentado_prop) ~ ospaq_sitting_frac + work_type	Subject-level aggregation
H5 (Exploratory)	EMG p90 (%MVC)	LMM	EMG p90 ~ hr_ratio_mean_within + hr_ratio_mean_between + noise_mean_within + noise_mean_between + posture_95_confidence_ellipse_area_within + posture_95_confidence_ellipse_area_between + work_type + C(day_index)	Within-between decomposition
H6 (Confirmatory)	Posture area (cm^2)	LMM	posture_95_confidence_ellipse_area ~ work_type + C(day_index)	No transform

Transformations and Units

Transform Summary

- EMG p90: %MVC; can exceed 100% -> not a strict proportion; log transform used in H1
- Workload mean: questionnaire score; no transform
- Sitting proportion: proportion in (0,1); logit transform used in H3 and H4
- OSPAQ sitting: proportion in (0,1); predictor, no transform
- Posture ellipse area: cm^2; no transform
- HAR durations: seconds; used for duration-weighted sitting proportion in H4

Assumption Checks and Corrections

Diagnostics

- Normality: Q–Q plot + Shapiro-Wilk/Jarque-Bera summary
- Homoscedasticity: residuals vs fitted + Breusch–Pagan proxy
- Outliers: standardized residuals > 3 flagged
- Auto-correction: if violations and outcome ≥ 0, apply log transform and refit
- Bootstrap: cluster bootstrap p-values if violations persist and configured

Note: H5 produced convergence warnings (boundary of parameter space), retained to avoid masking instability in the exploratory model.

Multiple Comparisons

Confirmatory family: H1, H2, H3, H4, H6

Correction Strategy

Holm step-down procedure (FWER control) applied to the confirmatory family. H5 is exploratory and excluded from correction. Primary p-values are LRT (full vs reduced model); Wald p-values retained for sensitivity.

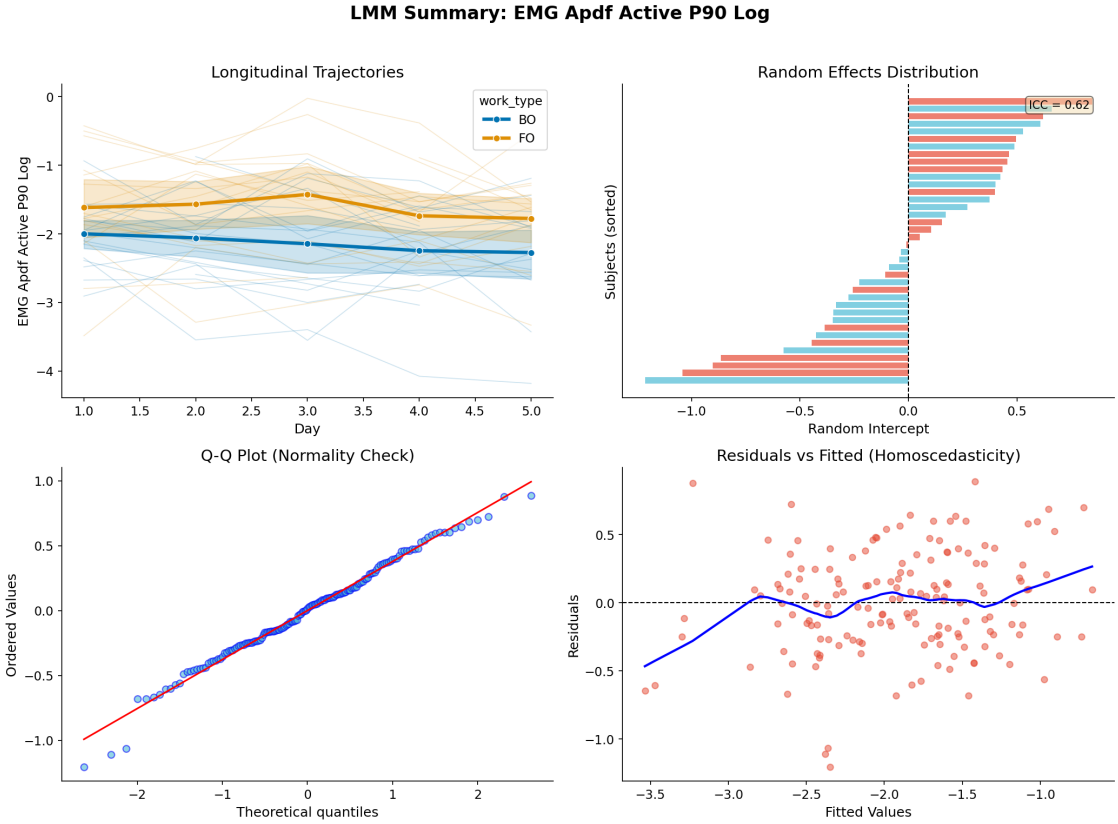
Results: Estimates + Interpretation

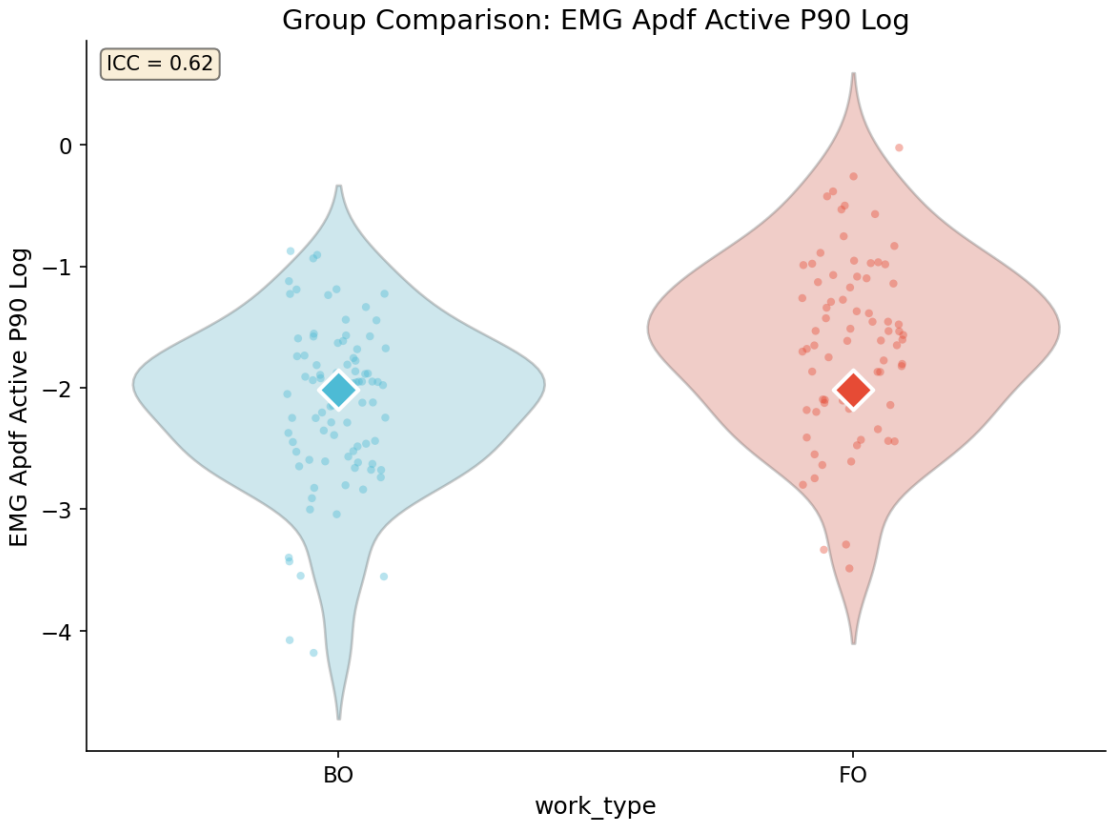
Hypothesis	N_obs / N_subjects	Primary term	Estimate (95% CI)	Wald p	LRT p	Holm p	Interpretation
H1	161 / 38	work_type	0.5027 [0.1335, 0.8719]	0.0076	0.0107	0.0537	FO higher than BO on log scale; narrowly misses Holm threshold
H2	176 / 38	work_type	-0.0176 [-0.3930, 0.3578]	0.9268	0.9268	1.0000	No evidence of FO/BO workload difference
H3	168 / 38	workload_mean	-0.0505 [-0.1416, 0.0405]	0.2768	0.0431	0.1294	LRT suggests model-level improvement but coefficient not significant; tentative evidence
H4	38 / 38	ospaq_sitting_fra	0.1709 [-0.4316, 0.7734]	0.5684	—	1.0000	No evidence of strong self-report vs objective association
H5 (Expl.)	160 / 38	posture_within	0.0041 [-0.0834, 0.0916]	0.9266	0.9266	—	No within-day posture–EMG association; exploratory with convergence warnings
H6	180 / 38	work_type	-0.1482 [-0.2660, -0.0305]	0.0136	0.0174	0.0698	Suggestive FO/BO difference; not confirmatory under Holm

Plots and Diagnostics

Each LMM includes a 4-panel summary: trajectories, random intercepts, Q–Q plot, residuals vs fitted.

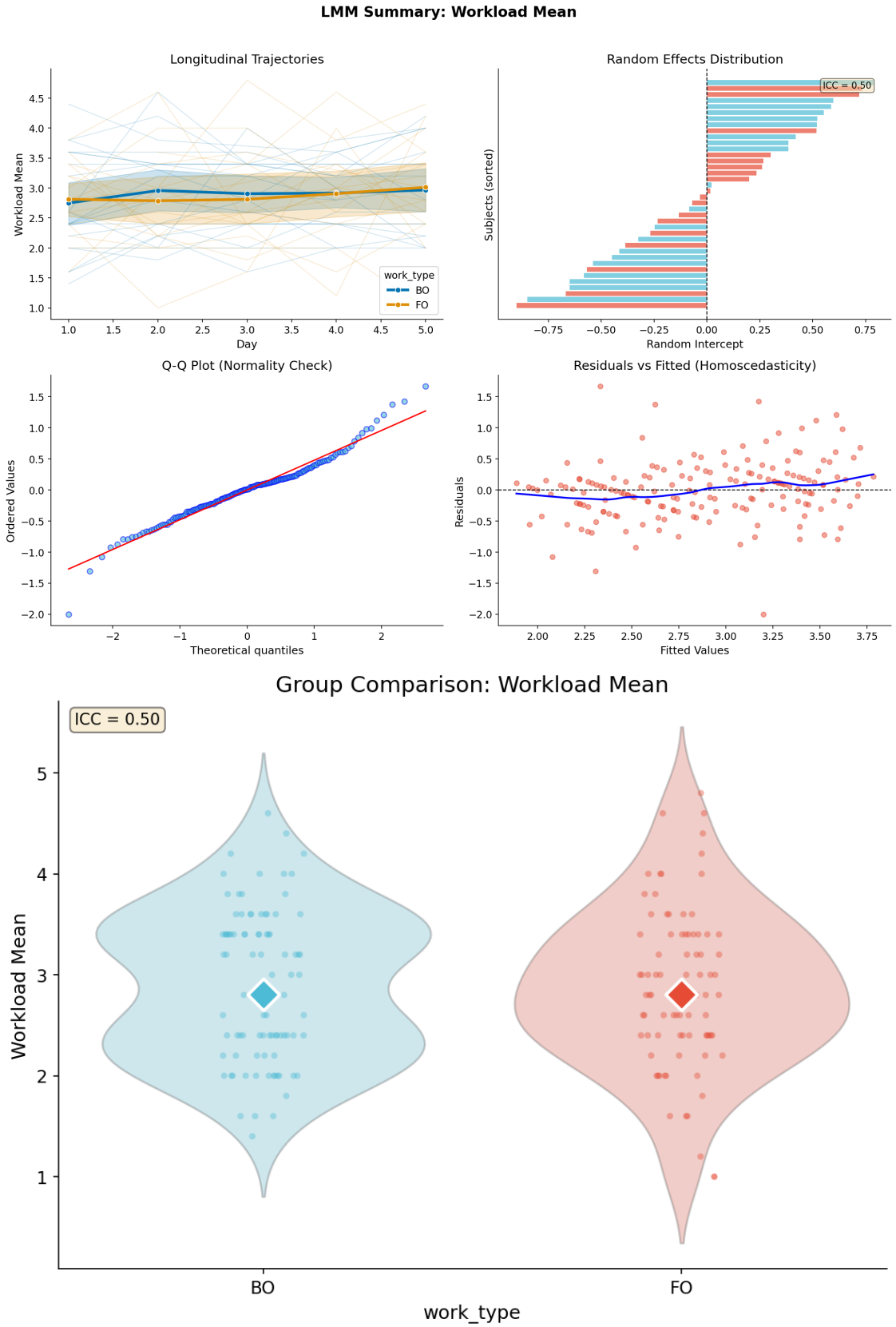
H1 – EMG p90





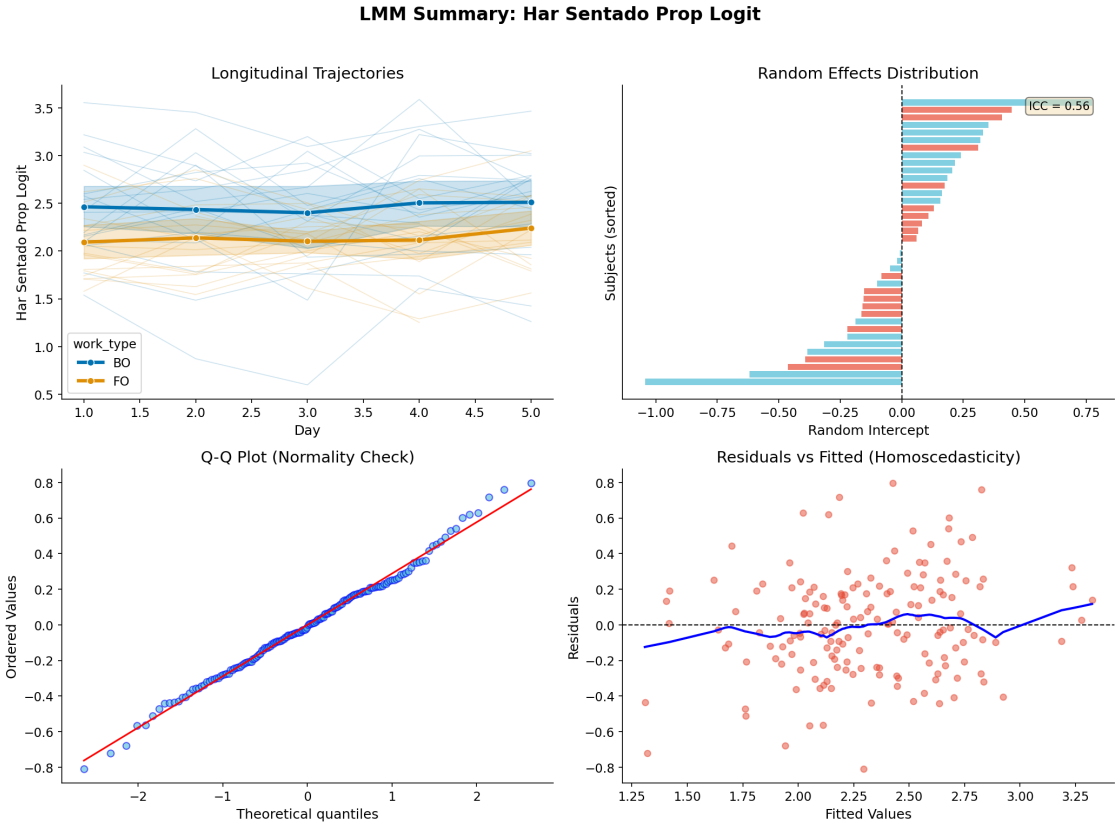
Diagnostics: improved normality after log transform; residual variance stabilized.

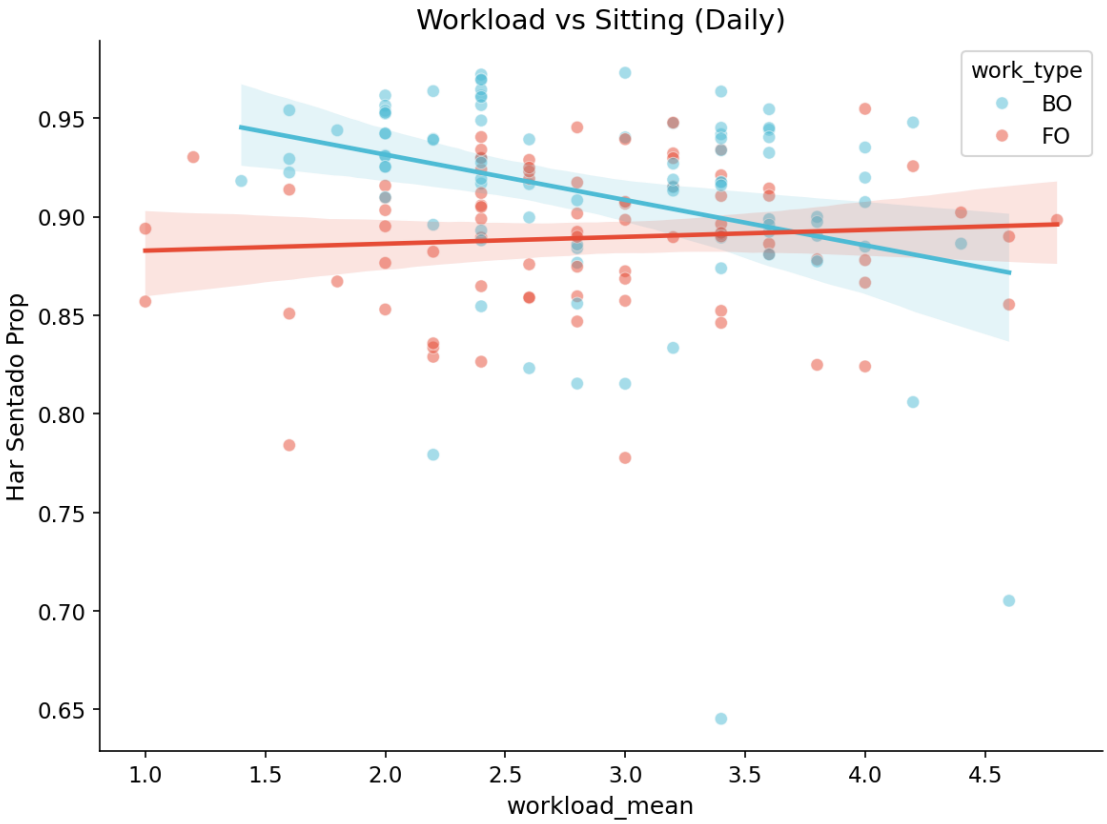
H2 – Workload



Diagnostics: approximately symmetric residuals; no strong heteroscedasticity.

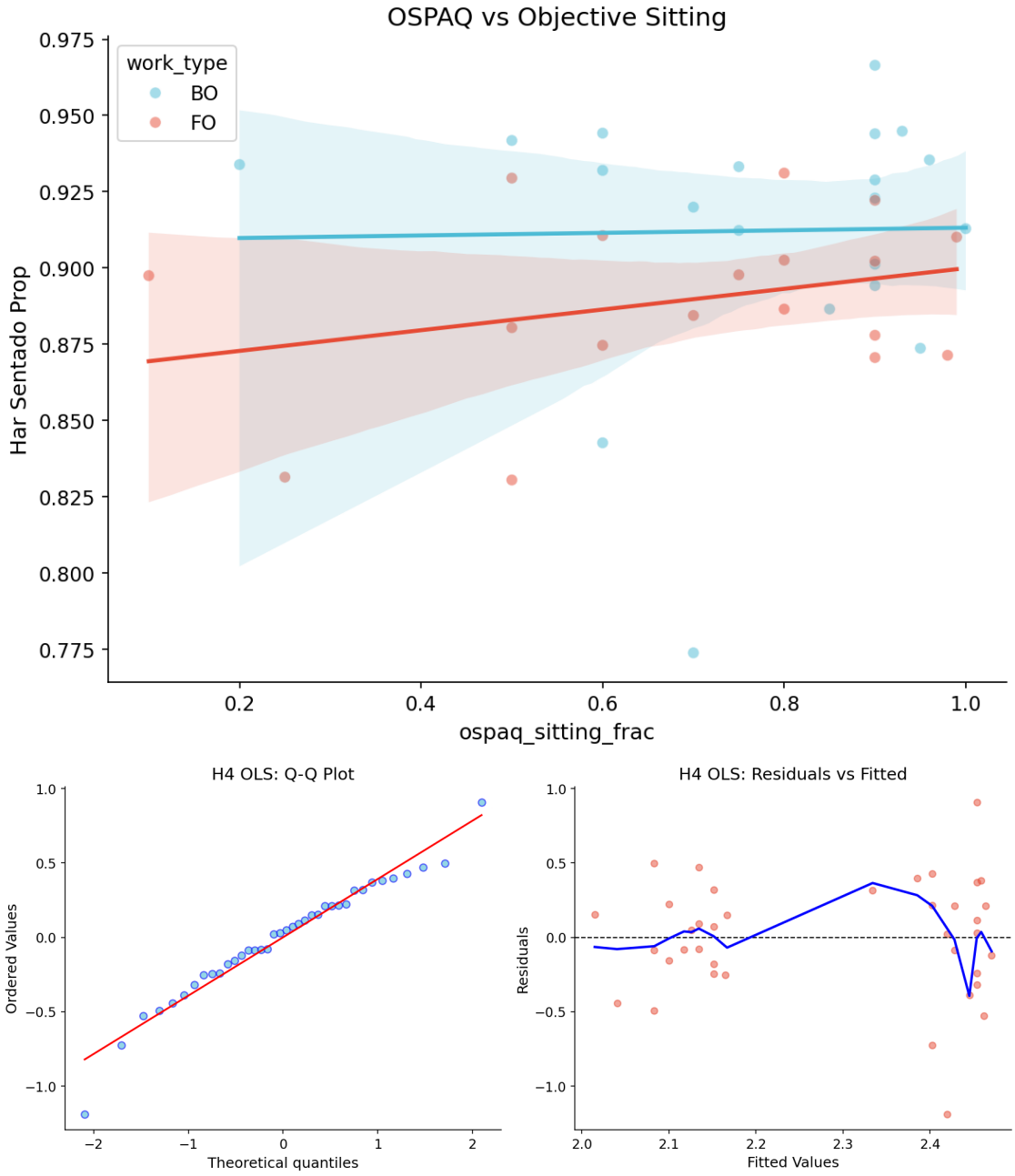
H3 – Workload → Sitting





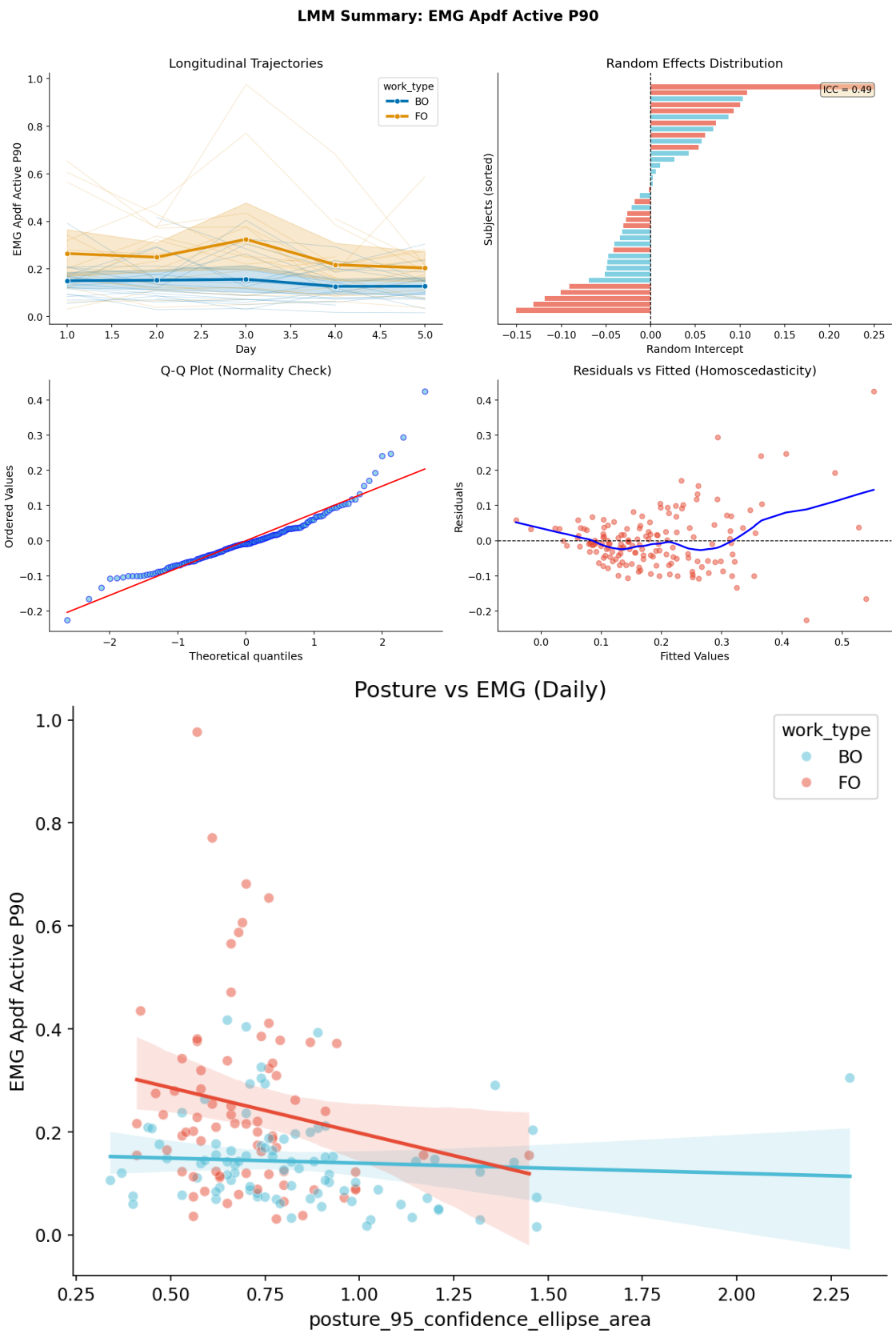
Diagnostics: logit scale yields acceptable residual structure; mild tail deviation.

H4 – OSPAQ Validation



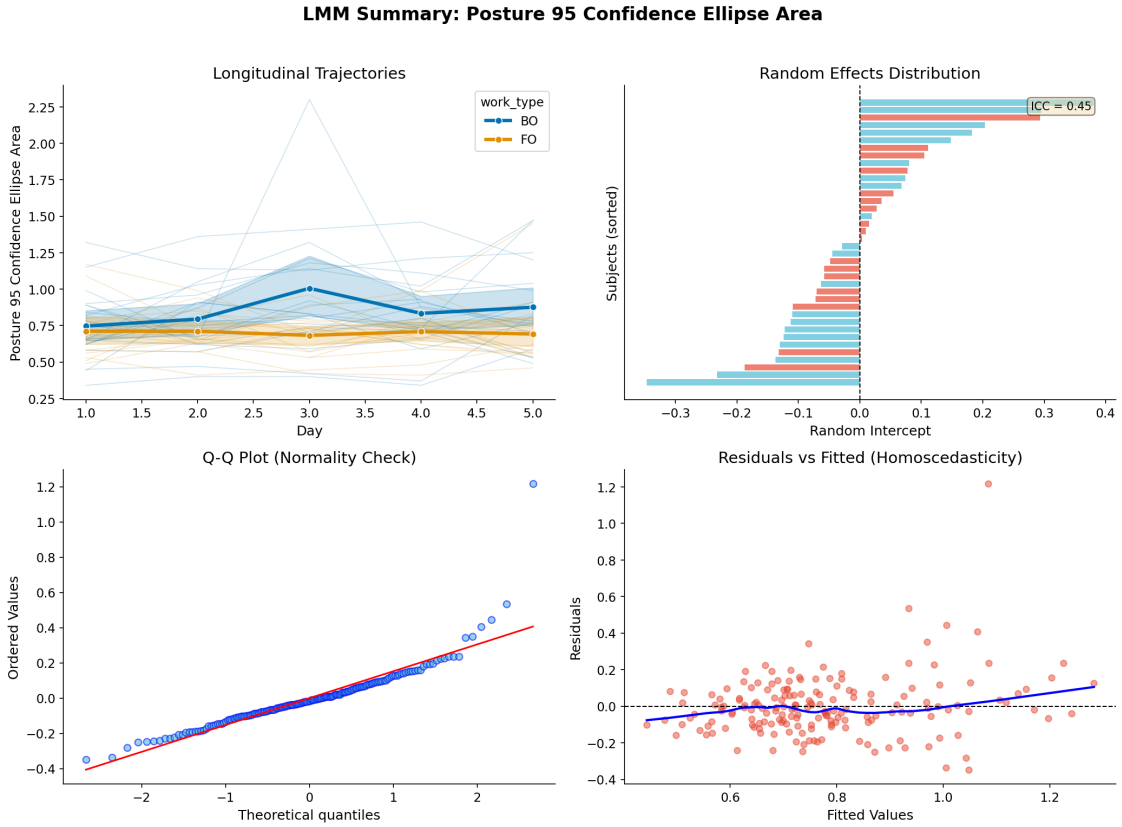
Diagnostics: OLS residuals show no strong pattern; normality acceptable for N=38.

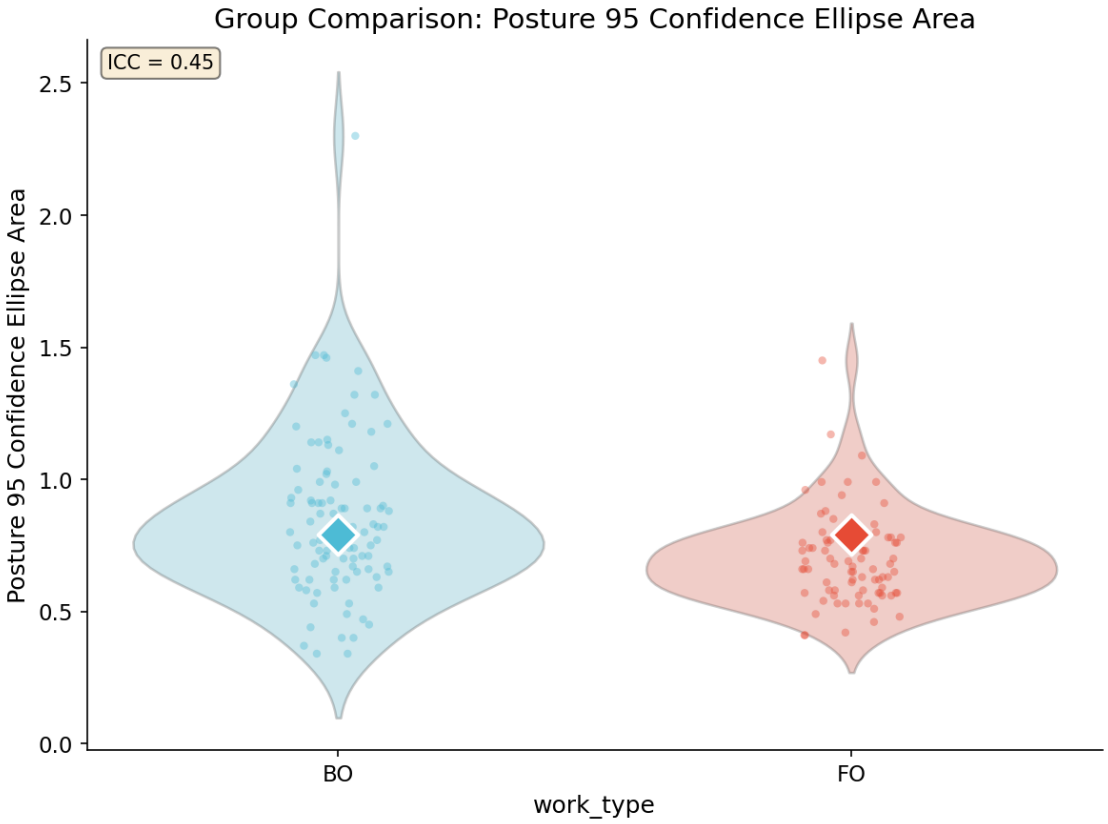
H5 – Physiological → EMG (Exploratory)



Diagnostics: convergence warnings; interpret cautiously.

H6 – Posture


LMM Summary: Posture 95 Confidence Ellipse Area
Longitudinal Trajectories
work_type
BO
FO
Posture 95 Confidence Ellipse Area
Day
Random Effects Distribution
ICC = 0.45
Subjects (sorted)
Random Intercept
Q-Q Plot (Normality Check)
Ordered Values
Theoretical quantiles
Residuals vs Fitted (Homoscedasticity)
Residuals
Fitted Values



Diagnostics: residuals broadly acceptable.

Limitations and Practical Considerations

- Sample size (38 subjects) limits power under strict FWER control
- LMMs assume Gaussian residuals on the transformed scale; diagnostics support adequacy but not certainty
- EMG p90 can exceed 100% MVC; logit is invalid for this outcome
- H5 is exploratory and relatively complex for the available sample size
- H4 is cross-sectional; Bland–Altman could be added if agreement (not only correlation) is desired

Recommendations for Next Steps

- Provide an FDR sensitivity analysis if a less conservative correction is desired
- Consider power analysis for future data collections
- Add Bland–Altman analysis for OSPAQ validation if agreement metrics are needed