

Additional file 1: Supplementary methods, tables and figures

**Multilevel Predictors and Cross-Level Mechanisms of Functional Recovery After Hip Fracture
in Older Adults: A Prospective Cohort Study**

Lujin Lian, PhD, et al.

BMC Geriatrics

Contents:

Supplementary Text S1. Detailed three-stage feature selection and Bayesian hyperparameter optimization procedures.

Supplementary Table S1. XGBoost model performance and optimized hyperparameters.

Supplementary Table S2. Five-level contribution analysis of final selected predictors.

Supplementary Figure S1. Bootstrap-LASSO stability selection frequency plot.

Supplementary Figure S2. XGBoost feature importance (measured by gain).

Supplementary Figure S3. SHAP mean absolute bar plot.

Supplementary Figure S4. Bootstrap out-of-bag AUC distribution.

Supplementary Figure S5. Calibration plot (predicted vs observed).

Supplementary Figure S6. Five-level contribution analysis of ecological levels.

Supplementary Text S1. Detailed three-stage feature selection and Bayesian hyperparameter optimization procedures.

Stage 1: Filtering. The 24 candidate predictors were screened for coding and range consistency and near-zero variance; no variable was excluded on the basis of univariable outcome association at this stage. Missing data were handled by complete-case analysis (overall missingness <5%).

Stage 2: Bootstrap-LASSO stability selection. Least absolute shrinkage and selection operator (LASSO) logistic regression was fitted within each of 1,000 bootstrap resamples of the analytical sample, and the selection frequency of each variable (the proportion of resamples in which it received a non-zero coefficient) was computed. Variables with a selection frequency of 0.50 or higher were retained (Supplementary Figure S1).

Stage 3: Clinical constraint. To preserve the theoretical completeness of the Health Ecological Model, retention was constrained so that every ecological level (L1–L5) was represented by at least one variable; where a level would otherwise have been unrepresented, the variable with the highest selection frequency within that level was retained. This procedure yielded the final 10 predictors used in model development.

Bayesian hyperparameter optimization. XGBoost hyperparameters (max_depth, learning_rate, n_estimators, subsample, colsample_bytree, min_child_weight, and gamma) were tuned by Bayesian optimization with 5-fold cross-validation on the full sample, optimizing the cross-validated area under the receiver operating characteristic curve. The optimized values are reported in Supplementary Table S1. All preprocessing and tuning parameters were fixed before bootstrap out-of-bag validation to prevent information leakage.

Supplementary Table S1. XGBoost model performance and optimized hyperparameters.

Metric	Value
Algorithm	XGBoost
AUC (full sample)	0.98
Bootstrap OOB AUC (B = 1,000)	0.92 (95% CI: 0.89-0.95)
Calibration (Hosmer-Lemeshow test)	P < .001
Optimal max_depth	4
Optimal learning_rate	0.15
Optimal n_estimators	34
Optimal subsample	0.8
Optimal colsample_bytree	1
Optimal min_child_weight	3
Optimal gamma	0.1

Note: The Hosmer-Lemeshow test is sensitive to large sample sizes; the calibration curve (Supplementary Figure S5) showed acceptable agreement between predicted and observed probabilities.

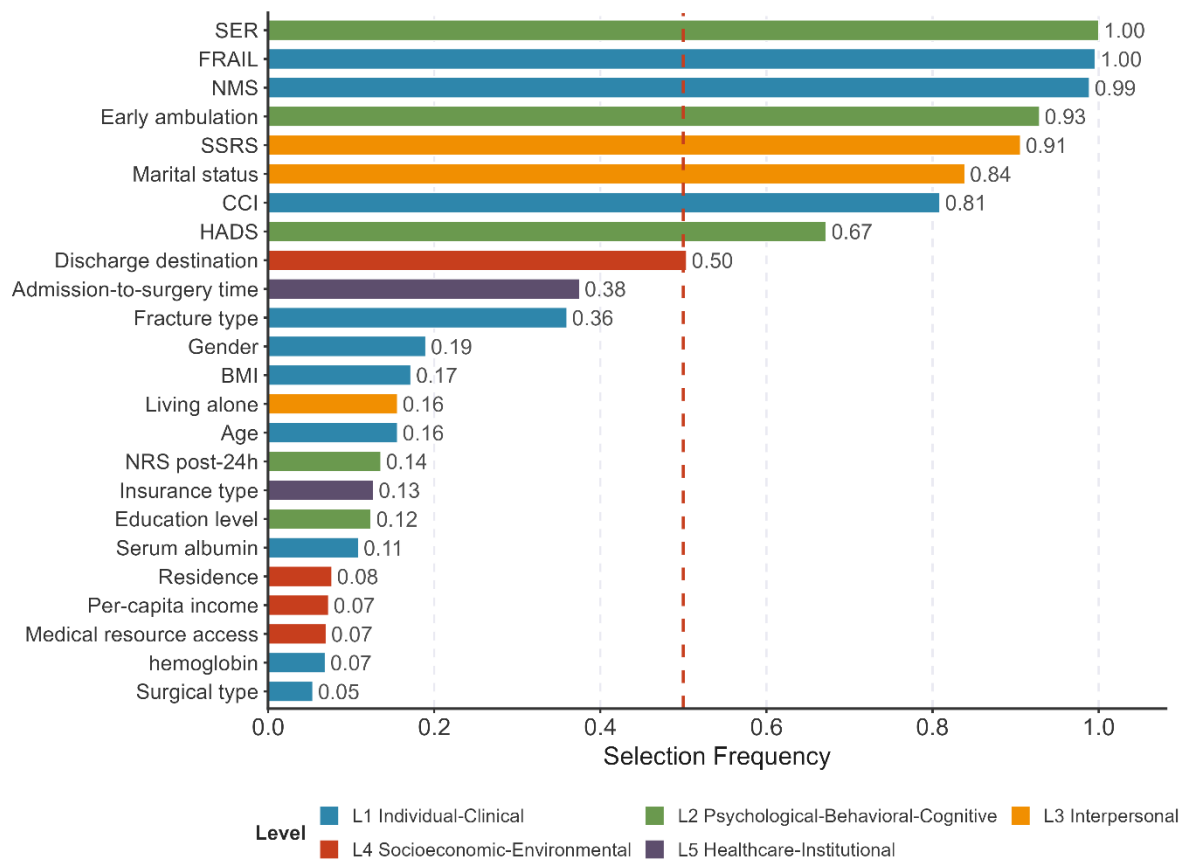
OOB = out-of-bag; AUC = area under the receiver operating characteristic curve; CI = confidence interval.

Supplementary Table S2. Five-level contribution analysis of final selected predictors.

Ecological Level	Contributors	Mean SHAP	Contribution (%)
L1-Individual-clinical	FRAIL, NMS, CCI	0.45	32.4%
L2-Psychological-behavioral-cognitive	SER, HADS, early ambulation	0.48	34.6%
L3-Interpersonal	SSRS, marital status	0.55	26.3%
L4-Socioeconomic-environmental	Discharge destination	0.14	3.3%
L5-Healthcare-institutional	Admission-to-surgery time	0.14	3.4%
Combined (contextual)	L2-L5 -	-	67.6%

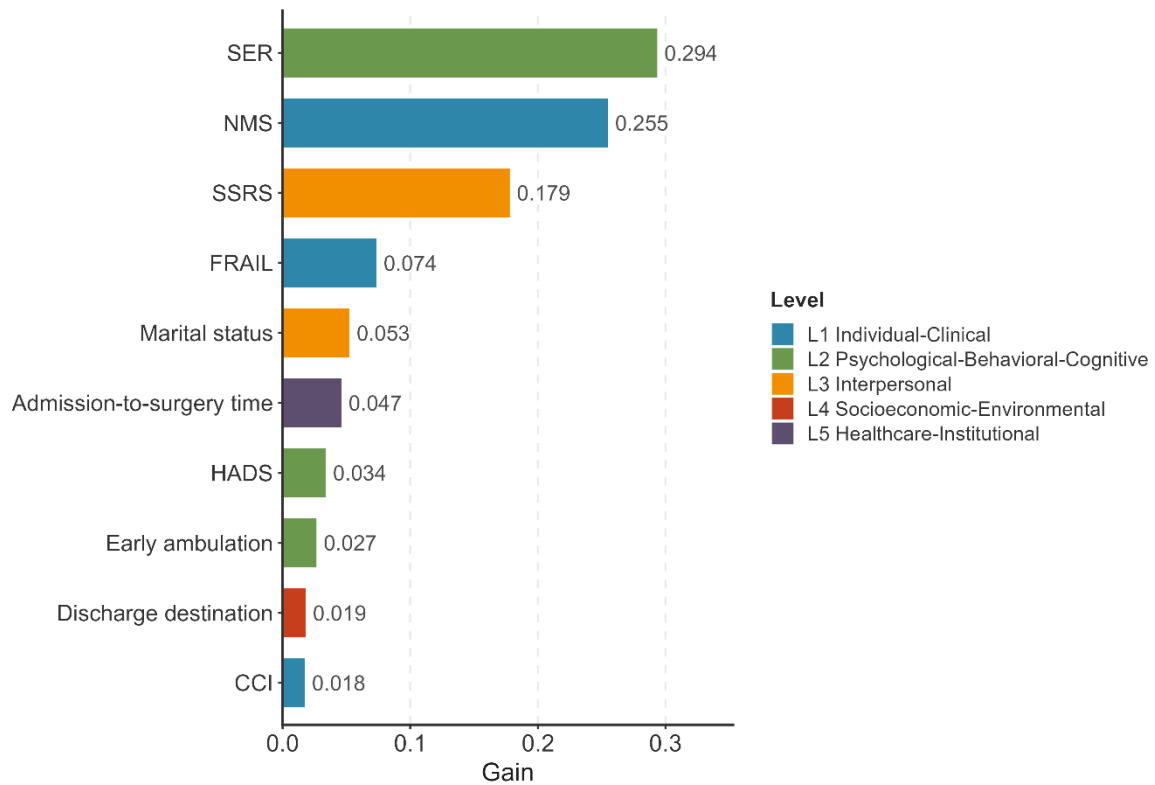
Note: Mean |SHAP| values are averaged across all selected variables within each level. Only the top contributors are listed for brevity. FRAIL = Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight scale; NMS = New Mobility Score; CCI = Charlson Comorbidity Index; SER = Self-Efficacy for Rehabilitation; HADS = Hospital Anxiety and Depression Scale; NRS = Numeric Rating Scale; SSRS =

Social Support Rating Scale.



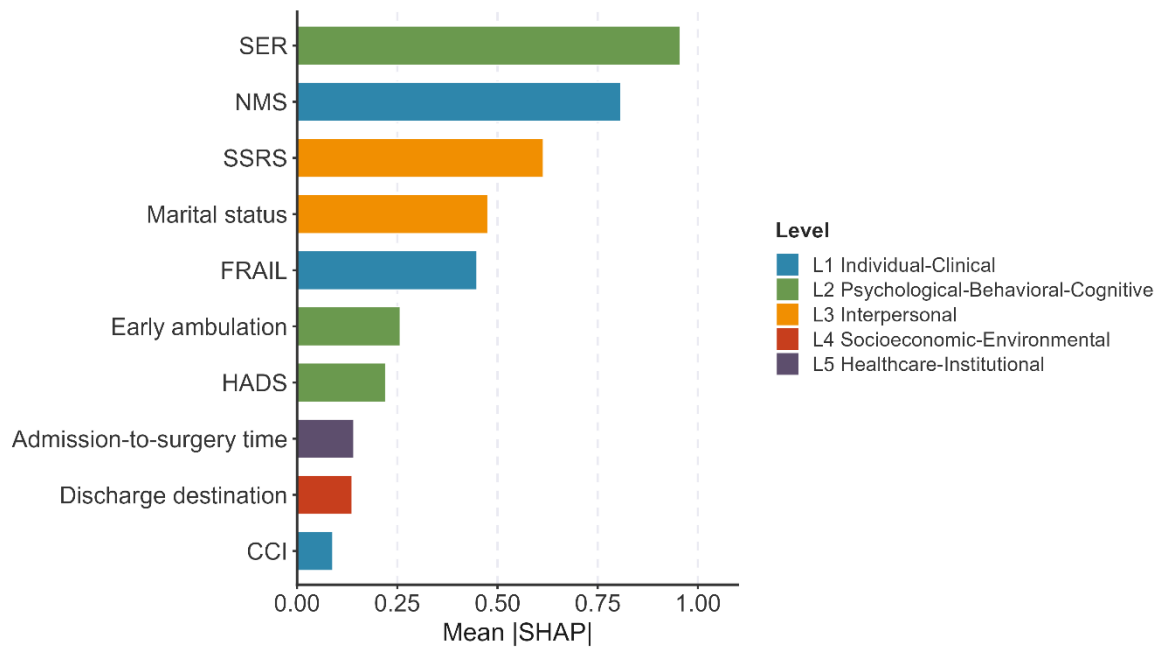
Supplementary Figure S1. Bootstrap-LASSO stability selection frequency plot.

Horizontal bar chart showing the selection frequency of each candidate predictor across 1,000 Bootstrap-LASSO iterations. Variables are colored by ecological level (L1–L5). The red dashed line indicates the 50% selection frequency threshold; variables at or above this threshold were retained for subsequent clinical constraint validation. LASSO = least absolute shrinkage and selection operator.



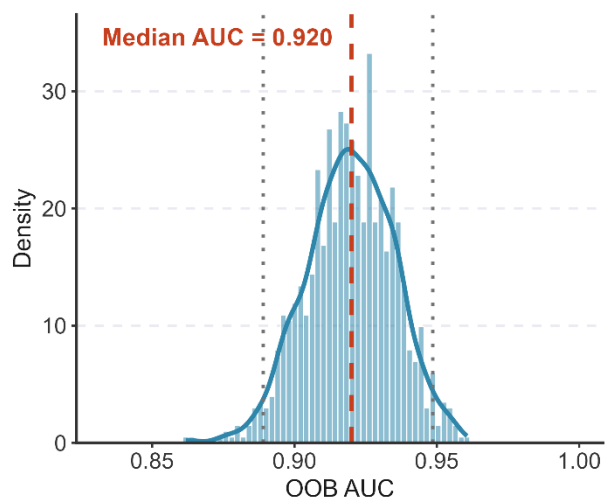
Supplementary Figure S2. XGBoost feature importance (measured by gain).

Feature importance calculated using the “Gain” metric, which reflects the average improvement in prediction accuracy contributed by each feature when used as a split node in the XGBoost model. Colors correspond to the five ecological levels (L1–L5).



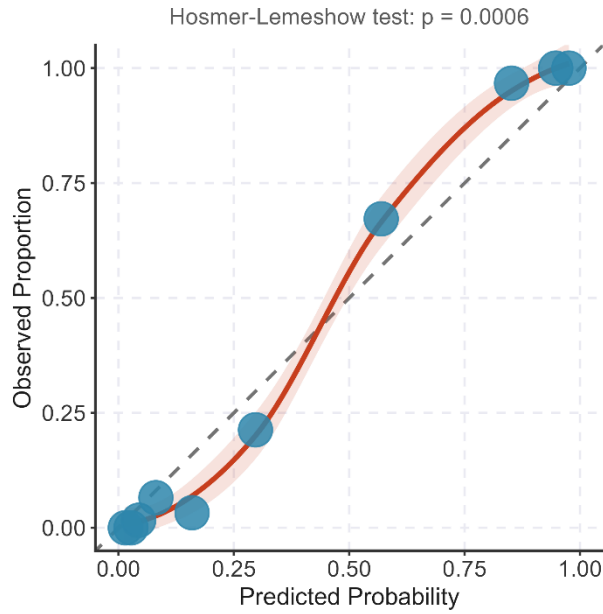
Supplementary Figure S3. SHAP mean absolute bar plot.

Bar plot displaying the top 10 predictors ranked by mean absolute SHAP value. Each bar represents the average magnitude of a feature's contribution to the model output across all observations, with colors corresponding to the five ecological levels (L1–L5). SHAP = SHapley Additive exPlanations.



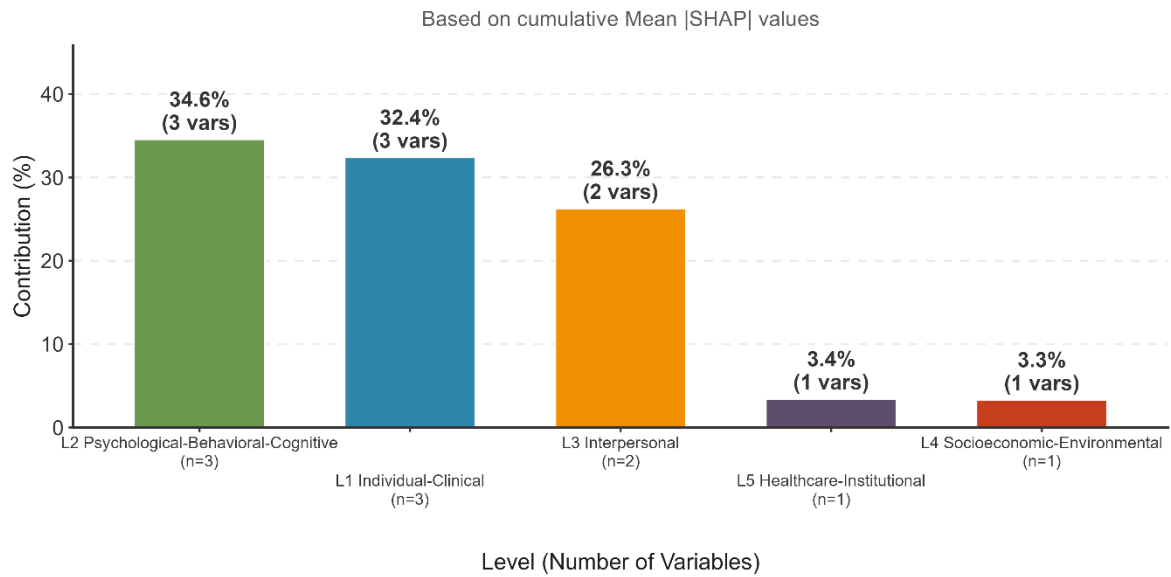
Supplementary Figure S4. Bootstrap out-of-bag AUC distribution.

Histogram of the bootstrap out-of-bag (OOB) AUC values across 1,000 resampling iterations. The red dashed vertical line indicates the median OOB AUC (0.920); the gray dotted vertical lines indicate the 95% confidence interval (0.889–0.949). AUC = area under the receiver operating characteristic curve; OOB = out-of-bag.



Supplementary Figure S5. Calibration plot (predicted vs observed).

Calibration plot showing predicted versus observed probabilities of ADL dependency at 6 months post-surgery. The dashed diagonal line represents perfect calibration. Points represent mean predicted and observed probabilities within each decile ($n = 10$ groups). The red line represents locally estimated scatterplot smoothing (LOESS). ADL = activities of daily living.



Supplementary Figure S6. Five-level contribution analysis of ecological levels.

Vertical bar chart showing the summed mean absolute SHAP contribution of each ecological level (L1 individual-clinical, 32.4%; L2 psychological-behavioral-cognitive, 34.6%; L3 interpersonal, 26.3%; L4 socioeconomic-environmental, 3.3%; L5 healthcare-institutional, 3.4%). Bar-top labels show the contribution percentage and the number of selected variables (n) within each level. The combined contextual contribution (L2–L5, 67.6%) substantially exceeds the prespecified 30% relevance threshold. SHAP = SHapley Additive exPlanations.