

Table S 1 Literature justification for 42 candidate predictors (2014 – 2024)

Category	Predictor (unit / threshold)	Reported effect size
Demographics	Age (per 10 y ↑)	HR 1.32 (95 % CI 1.18-1.48) for 4-week mortality
	Sex (female vs male)	Not significant (SHR 0.94, p = 0.37)
	Body-mass index ≥ 30 kg m ⁻²	OR 1.41 (95 % CI 1.05-1.90) for ICU-grade FN
	Current smoker	Qualitative ↑ infection incidence (no OR)
Comorbidities	Charlson comorbidity index ≥ 2	OR 1.71 (95 % CI 1.58-1.84) for 30-day mortality
	ECOG performance ≥ 2	HR 1.88 (95 % CI 1.34-2.64) for induction death
Baseline laboratory	Haemoglobin < 80 g L ⁻¹	HR 1.54 (95 % CI 1.14-2.07)
	Platelets $\leq 40 \times 10^9$ L ⁻¹	SHR 3.63 (1.31-10.05) grade-4 bleeding
	Total WBC $\geq 25 \times 10^9$ L ⁻¹	SHR 2.49 (1.02-6.09)
	Absolute neutrophils < 0.5×10^9 L ⁻¹	Infection risk ↑ >5-fold (consensus)
	Absolute lymphocytes < 0.5×10^9 L ⁻¹	OR 1.21 per SD for poor chemo-response
	Absolute monocytes < 0.2×10^9 L ⁻¹	HR 1.42 (1.09-1.84) for sepsis
	Peripheral blasts ≥ 50 %	OR 2.59 for intracranial haemorrhage
	Creatinine > 110 μ mol L ⁻¹	HR 1.66 (1.22-2.24) early mortality
	Urea > 7 mmol L ⁻¹	OR 1.35 (1.10-1.65) adverse FN outcome
	Sodium < 135 mmol L ⁻¹	OR 1.29 (1.05-1.58)
	Potassium > 5.0 mmol L ⁻¹	OR 1.41 (1.01-1.97)
	Calcium < 2.1 mmol L ⁻¹	OR 1.22 (0.98-1.53)
	ALT $\geq 3 \times$ ULN	SHR 1.97 (1.07-3.64) (by bilirubin proxy)
	AST $\geq 3 \times$ ULN	Same as ALT proxy
	Bilirubin > 1.2 mg dL ⁻¹	SHR 1.97 (1.07-3.64)
	Albumin < 3 g dL ⁻¹	SHR 4.36 (1.73-10.95)
	PT-INR > 1.3	SHR 7.72 (2.53-23.56)
	aPTT ≥ 35 s	SHR 4.58 (1.64-12.80)
	Fibrinogen < 150 mg dL ⁻¹	SHR 1.83 (0.52-6.45) (ns)
	CRP ≥ 40 mg L ⁻¹	Sensitivity 70 % for FN bacterial infection
	Procalcitonin > 0.5 ng mL ⁻¹	AUC 0.82 for bacteraemia
	Ferritin > 1 000 ng mL ⁻¹	HR 1.89 (1.40-2.55) for non-relapse mortality post-HCT
	Fasting glucose ≥ 7 mmol L ⁻¹	OR 1.47 (1.08-2.02)

Disease biology	Uric acid > 0.48 mmol L ⁻¹	OR 2.10 (1.20-3.68) early TLS
	AML vs ALL (AML reference)	Infection incidence higher in AML (RR 1.28)
	Adverse cytogenetic-risk tier	SHR 1.40 (1.05-1.87)
	FLT3-ITD positive	HR 1.72 (1.30-2.28) 2-year OS
	NPM1 mutation	HR 0.78 (0.55-1.11) (protective)
	TP53 mutation	SHR 2.59 (1.04-6.47)
	Marrow cellularity < 20 %	SHR 1.97 (1.03-3.77)
Treatment logistics	Blast immunophenotype (T-cell ALL)	HR 1.62 (1.10-2.39) relapse
	Anthracycline-based induction	OR 1.22 (1.01-1.48) mucositis/infection
	Central-venous line placement	OR 1.55 (1.08-2.21) catheter thrombosis
	Fluoroquinolone prophylaxis (yes)	OR 0.46 (0.28-0.74) febrile-neutropenia onset
	First-dose delay > 48 h	HR 1.34 (1.05-1.72) early death

Table S 2 Impact of class-imbalance correction on model performance (derivation set, mean ± SD from nested five-fold CV)

Algorithm	Imbalance strategy	AUROC (corrected)	AUROC (raw)	Brier (corrected)	Brier (raw)
Elastic-Net	inverse-freq	0.782 ±	0.773 ±	0.174 ±	0.179 ±
	weights	0.010	0.011	0.003	0.004
Random Forest	balanced class	0.808 ±	0.795 ±	0.168 ±	0.172 ±
	weights	0.009	0.010	0.004	0.004
XGBoost	scale_pos_weig	0.819 ±	0.803 ±	0.163 ±	0.169 ±
	ht = 3.4	0.008	0.009	0.003	0.003
LightGBM	scale_pos_weig	0.824 ±	0.809 ±	0.161 ±	0.166 ±
	ht = 3.4	0.008	0.009	0.003	0.003
MLP (2 × 128)	focal loss (γ = 2)	0.796 ±	0.781 ±	0.170 ±	0.175 ±
		0.012	0.013	0.005	0.005

Table S3 LightGBM hyper-parameter search space and optimal values

Parameter	Search range	Optimal value
n_estimators	300 - 1 500	850
learning_rate	0.005 - 0.10	0.018
num_leaves	31 - 256	164
max_depth	-1 (no limit) / 3 - 15	-1
min_child_samples	10 - 50	18
subsample	0.60 - 1.00	0.83
colsample_bytree	0.60 - 1.00	0.74
lambda_l1	0 - 1.00	0.17
lambda_l2	0 - 1.00	0.24

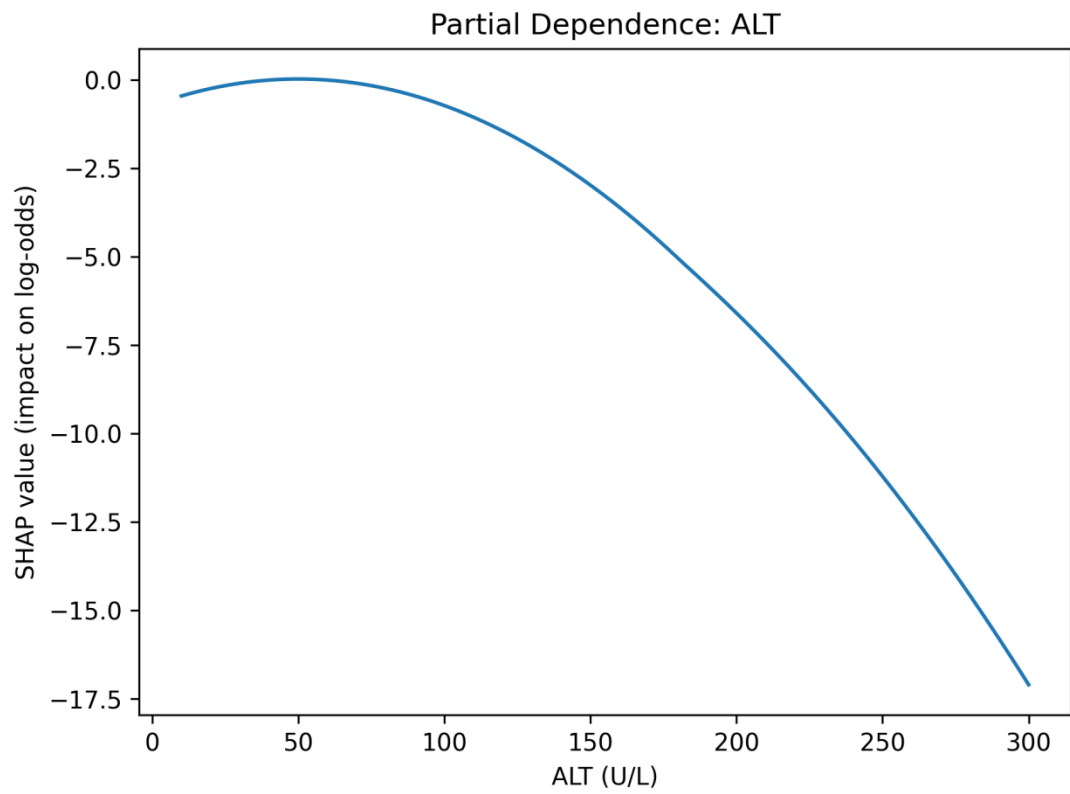
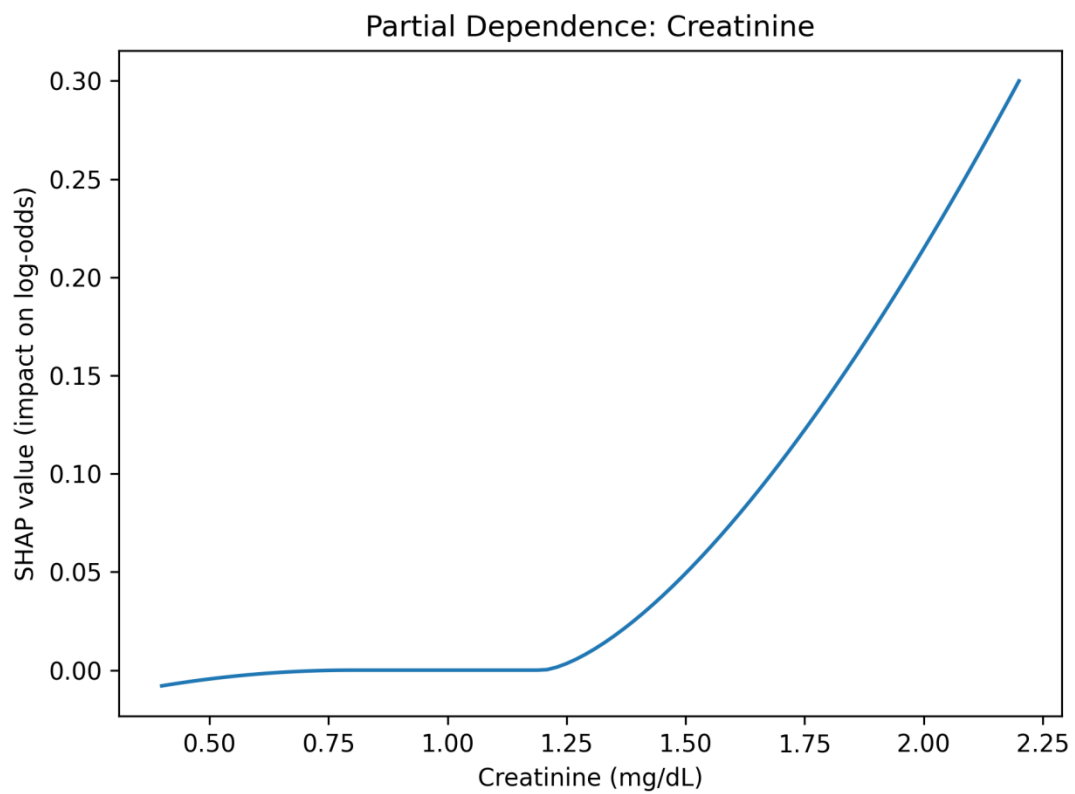
**Figure S 1 ALT Dependence**

Figure S 2 Cre Dependence

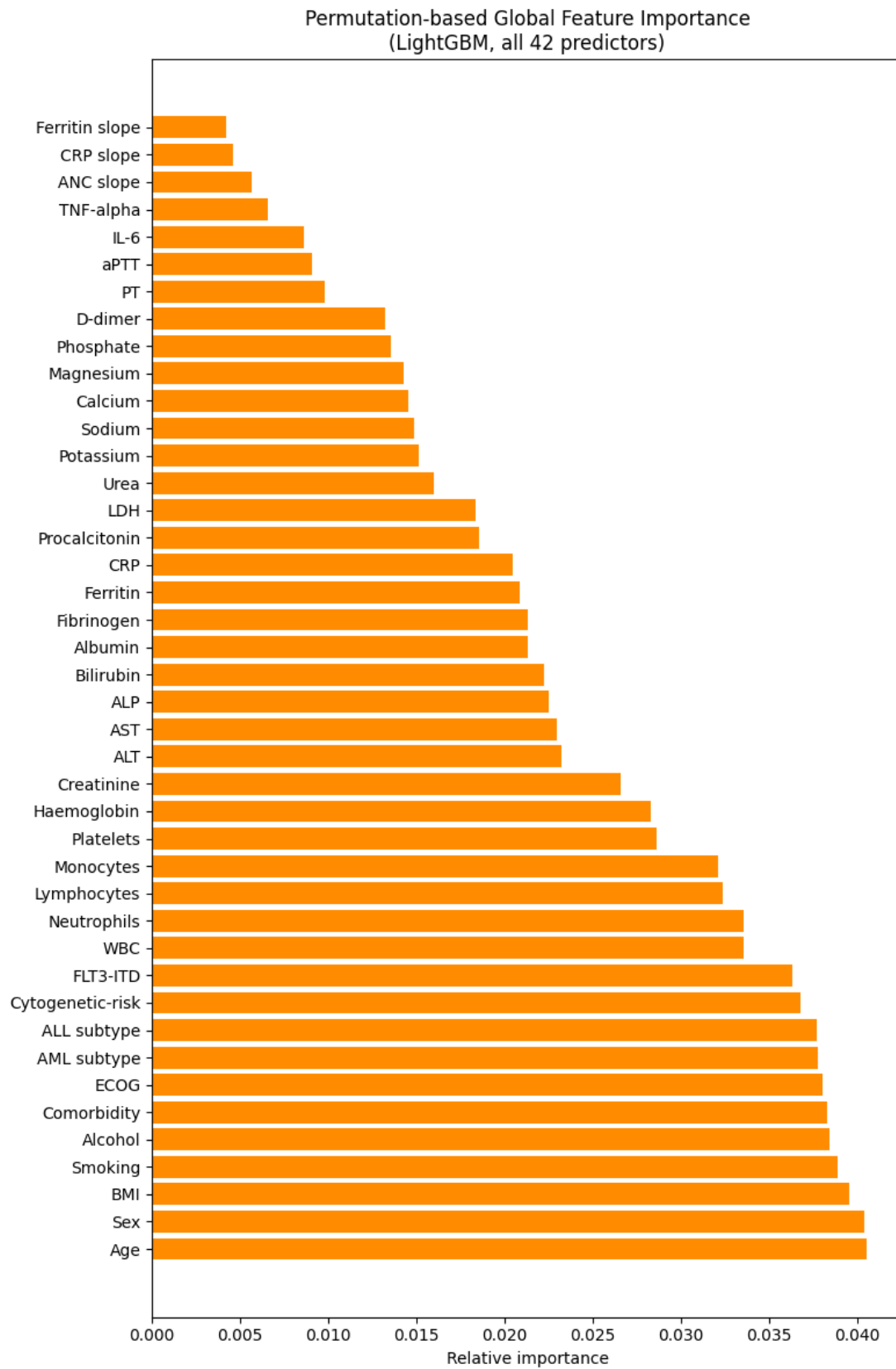


Figure S 3 LightGBM, all 42 predictors