Table S1. Definitions of features used for model development.

Feature		**Definition and clinical significance**	**Engineering formula or logic**
Original clinical variables	age_y	Patient age (years) at TDM sampling; affects distribution volume and clearance.	-
	admin_id	Sequential administration number; indicates therapy progression and steady-state likelihood.	-
Derived clinical variables	daily_dose_mg_kg	Daily dose normalized by body weight; primary dosing parameter.	(dose × 24 / interval) / weight
	egfr	Estimated glomerular filtration rate via revised Schwartz equation; major determinant of elimination.	0.413 × (height / creatinine)
	hours_since_prev_dose	Time between consecutive doses; defines dosing schedule.	Δtime (hours)
	hours_from_admin_to_sample	Time from latest dose to sampling; determines trough validity.	Δtime (hours)
Engineered features	dose_per_weight_hour	Hourly dose normalized by weight; reflects dosing intensity.	daily_dose_mg_kg / 24
	clearance_estimate	Estimated clearance based on dose and trough concentration.	daily_dose_mg_kg / (conc + $1e^{-6}$)
	clearance_age_adjusted	Clearance adjusted for age-related decline.	clearance_estimate × (1 − (age_y − 40)/100)
	age_egfr_ratio	Age scaled by renal function; indicator of physiological burden.	age_y / (egfr + 1)
	dose_egfr_interaction	Interaction between dose intensity and renal reserve.	daily_dose_mg_kg × egfr / 100
	is_mid_dose / is_late_dose	Categorical phase indicators: mid-phase (4th–8th dose), late-phase (≥9th dose).	if 4≤admin_id≤8 or ≥9
	time_dose_interaction	Product of daily dose intensity and sampling delay; represents cumulative exposure.	hours_from_admin_to_sample × daily_dose_mg_kg

TDM = therapeutic drug monitoring

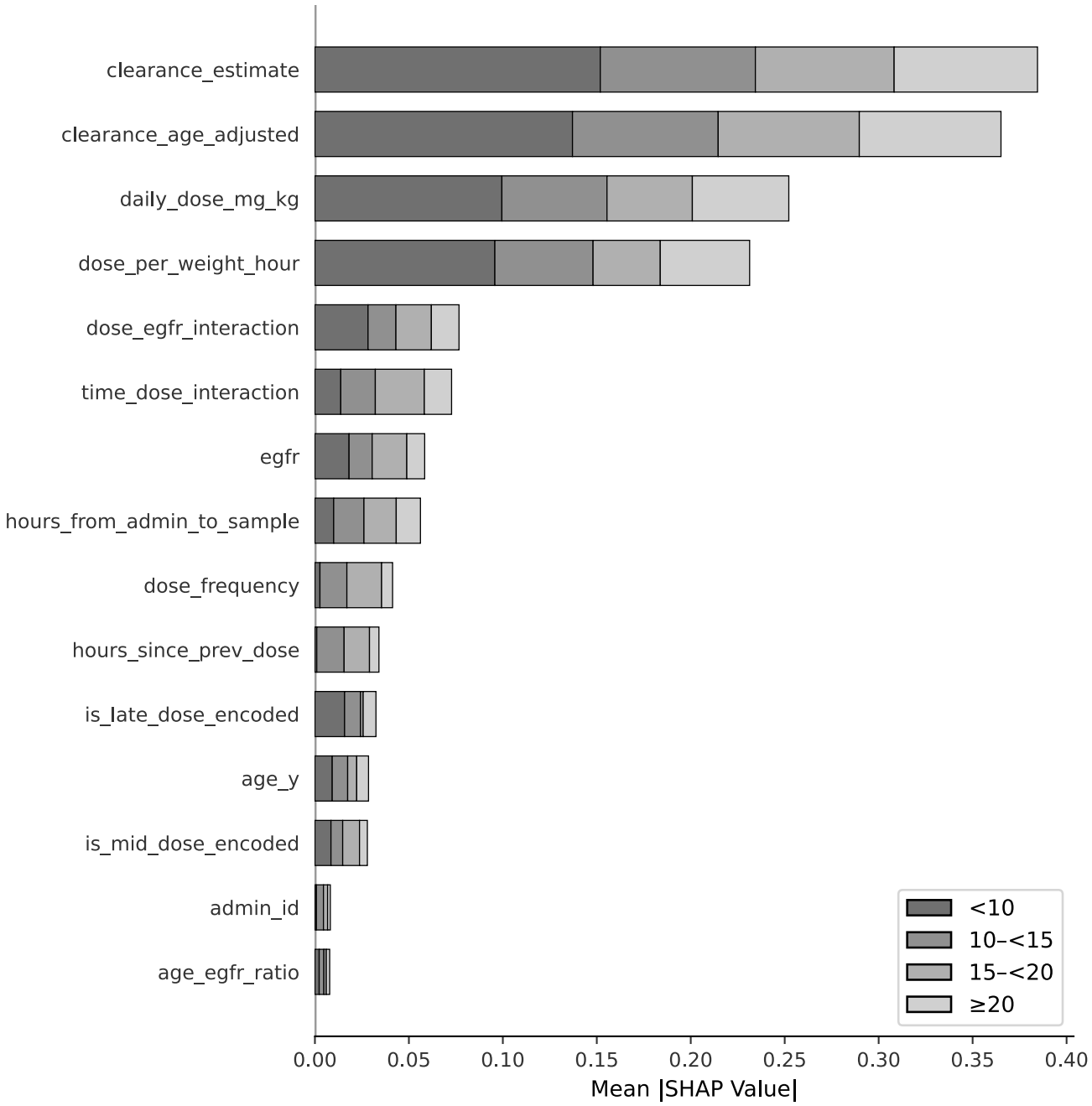

Figure S1. SHAP feature importance for predicting vancomycin trough concentrations.
Bar plot of mean absolute SHAP values showing the relative importance of input variables in predicting vancomycin trough serum concentration categories (<10, 10–<15, 15–<20, and ≥20 μg/mL). Each bar represents the average magnitude of a feature's contribution to model output across all samples; higher SHAP values indicate greater influence on the model's prediction. Variables are ranked in descending order of importance, highlighting the clinical and pharmacokinetic determinants most strongly associated with vancomycin trough level classification. Feature descriptions are provided in Supplementary Table S1.
SHAP = SHapley Additive exPlanations, TDM = therapeutic drug monitoring, eGFR = estimated glomerular filtration rate.

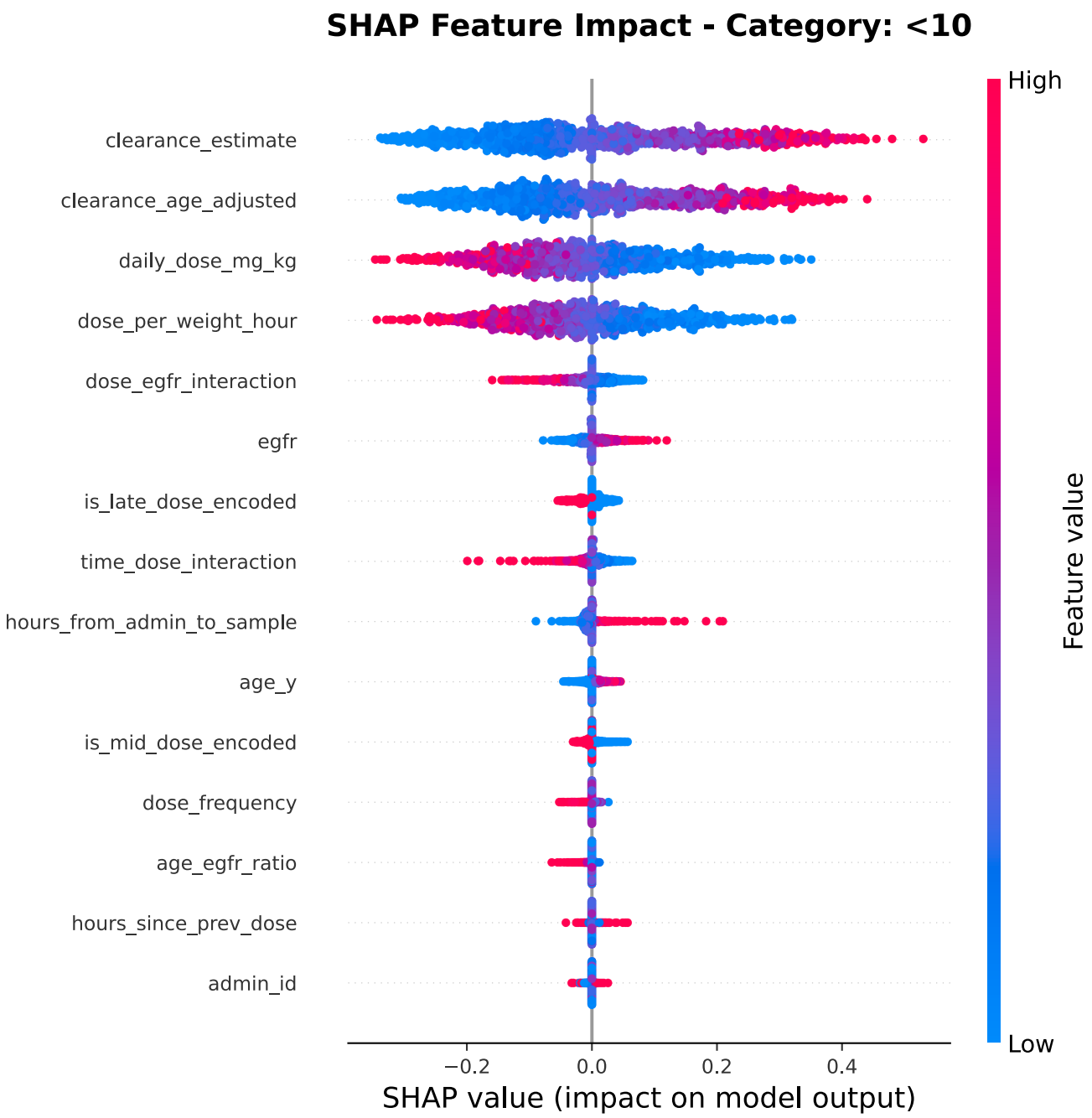

Figure S2. SHAP summary for the <10 µg/mL vancomycin trough concentration category.
Beeswarm plot of SHAP values showing the contribution of each input feature to the model's prediction that a TDM episode belongs to the <10 µg/mL vancomycin trough serum concentration category. Each point represents one episode; its horizontal position is the SHAP value (impact on model output), where positive values increase the predicted probability of the <10 µg/mL class and negative values decrease it. Features are ordered from top to bottom by mean absolute SHAP value, so higher positions indicate greater overall importance. Point color encodes the underlying feature value (blue = low, red = high), enabling visual assessment of how low or high values of a feature drive predictions toward or away from this category. See Supplementary Table S1 for full feature definitions and clinical interpretations.
SHAP = SHapley Additive exPlanations, TDM = therapeutic drug monitoring, eGFR = estimated glomerular filtration rate.

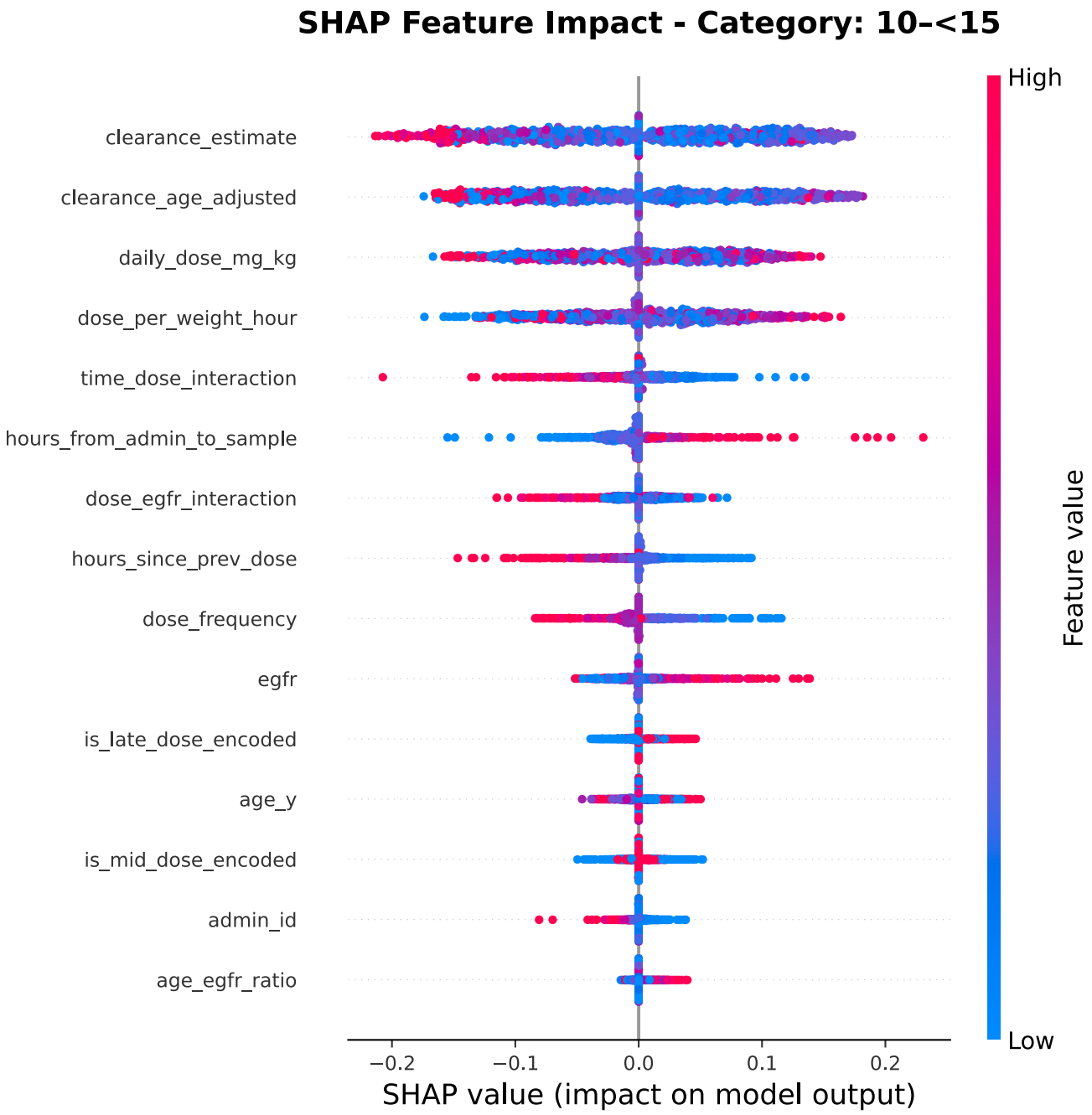

Figure S3. SHAP summary plot for the 10–<15 μg/mL vancomycin trough concentration category.
The figure displays the SHAP summary plot illustrating the relative contribution of each clinical and pharmacokinetic feature to the model's prediction of vancomycin trough serum concentrations within the 10–<15 μg/mL range. Each point represents one TDM episode. The horizontal position corresponds to the SHAP value (impact on model output), where positive values indicate stronger influence toward predicting this concentration category. Features are ranked by mean absolute SHAP value, with those at the top exerting the greatest effect on model prediction. Point colors represent actual feature values (red = high, blue = low), revealing how specific feature magnitudes affect prediction directionality. Detailed definitions and clinical interpretations of all features are provided in Supplementary Table S1.
SHAP = SHapley Additive exPlanations, TDM = therapeutic drug monitoring, eGFR = estimated glomerular filtration rate.

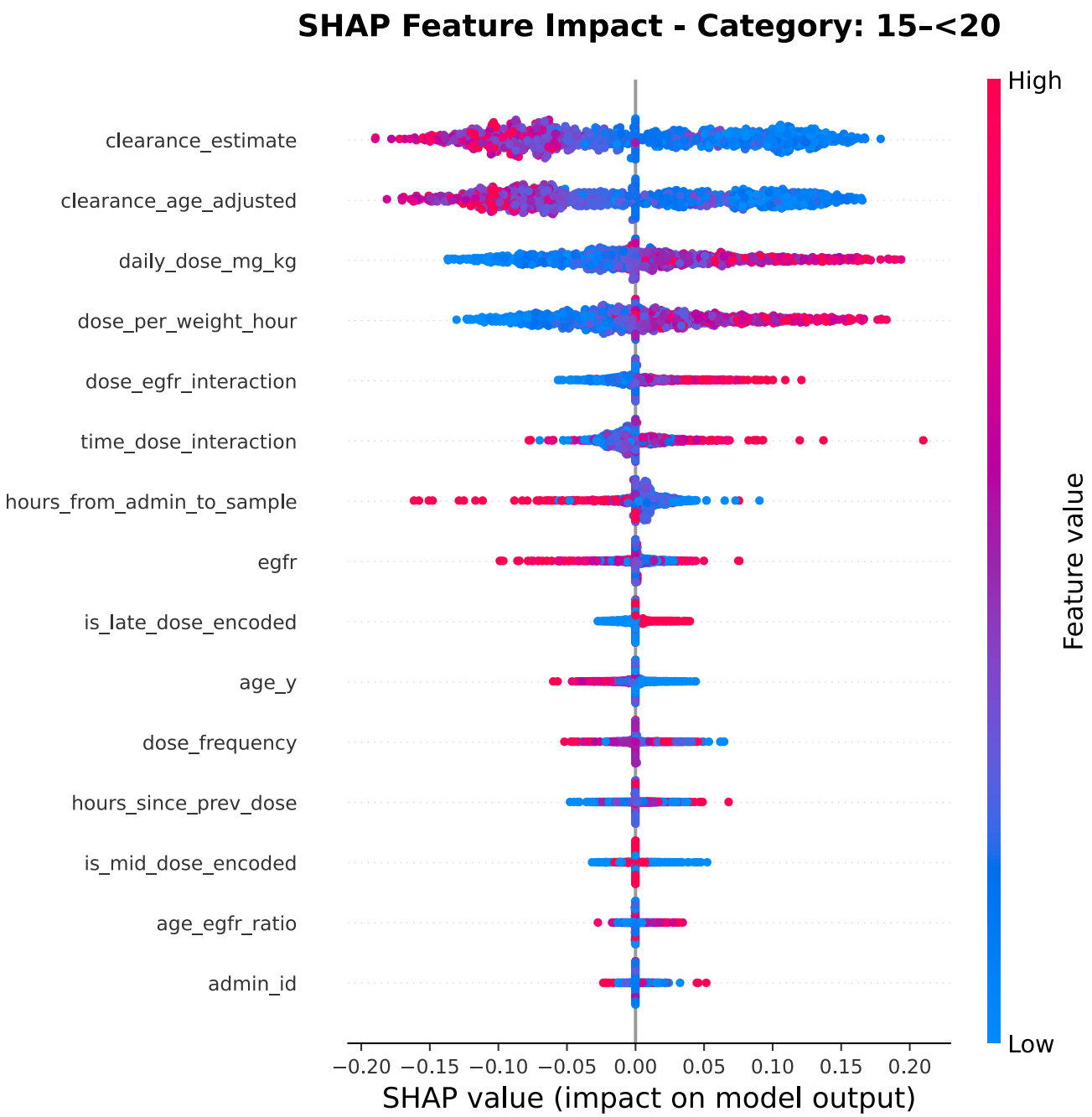

Figure S4. SHAP summary plot for the 15–<20 μg/mL vancomycin trough concentration category.
This figure presents the SHAP summary plot depicting the relative importance of clinical and pharmacokinetic features in predicting vancomycin trough serum concentrations within the 15–<20 μg/mL range.
Each dot represents a single TDM episode, with its horizontal position corresponding to the SHAP value (impact on model output) and its color indicating the feature value (red = high, blue = low). Features are sorted by mean absolute SHAP value, such that higher-ranked variables exert greater influence on prediction outcomes. Detailed variable definitions and clinical interpretations are provided in Supplementary Table S1.
SHAP = SHapley Additive exPlanations, TDM = therapeutic drug monitoring, eGFR = estimated glomerular filtration rate.

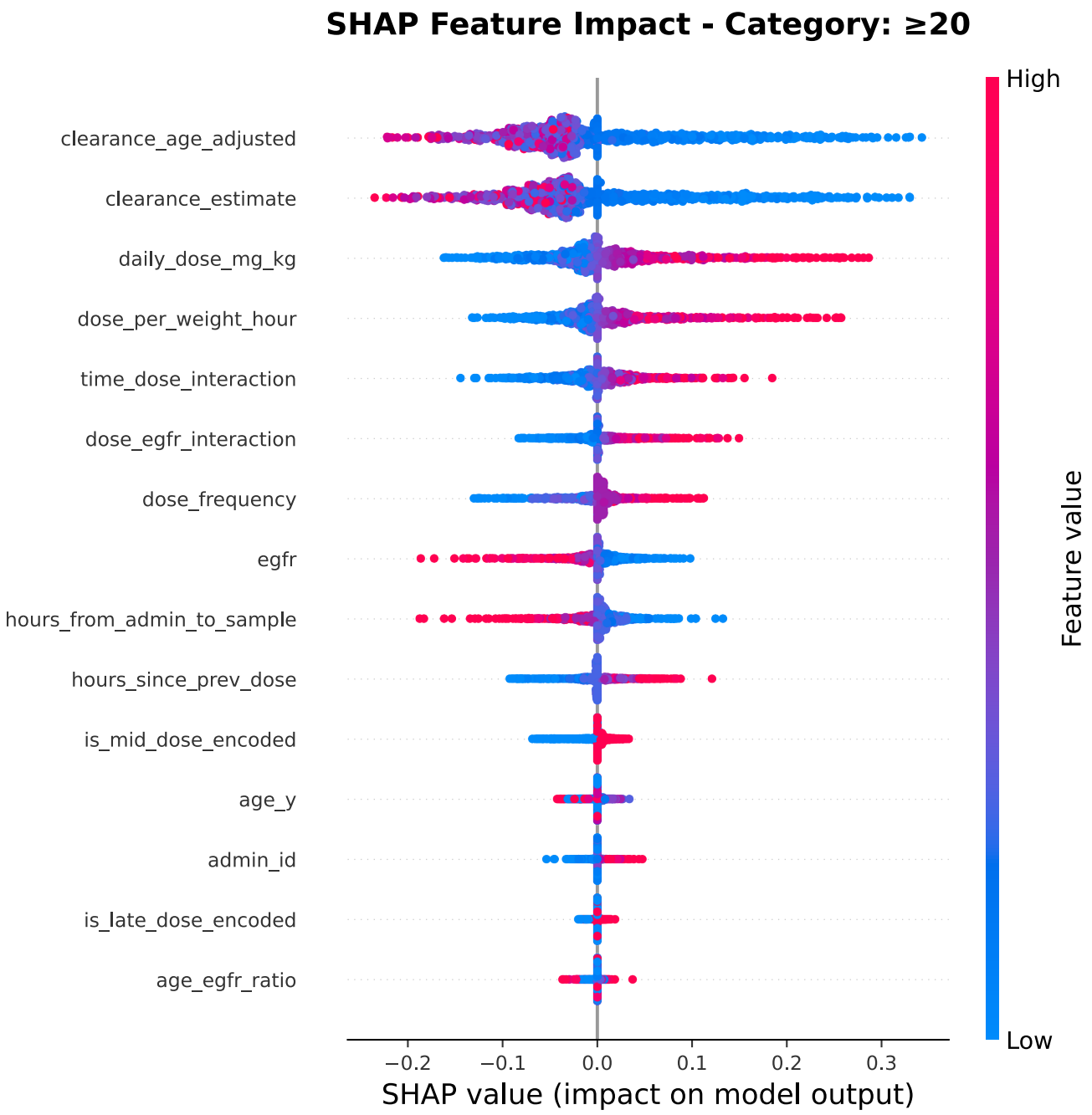

Figure S5. SHAP summary plot for the ≥20 µg/mL vancomycin trough concentration category.
This figure illustrates the SHAP summary plot showing the relative contribution of clinical and pharmacokinetic features to the model's prediction of vancomycin trough serum concentrations ≥20 µg/mL. Each dot represents an individual TDM episode. The horizontal axis corresponds to the SHAP value, reflecting the magnitude and direction of each feature's influence on the prediction. Red points denote higher feature values, and blue points represent lower values. Features are ordered by their mean absolute SHAP value, such that the most influential variables appear at the top. Detailed feature definitions and clinical interpretations are available in Supplementary Table S1.
SHAP = SHapley Additive exPlanations, TDM = therapeutic drug monitoring, eGFR = estimated glomerular filtration rate.