INTERNATIONAL JOURNAL OF STEM RESEARCH
Science • Technology • Engineering • Mathematics

Manuscript Type: Original Research
Published: *[27-06-2026]*
DOI: *https://doi.org/10.19895/ijstemr.2026.1.7*

SLECare: Explainable Machine Learning for Kidney Damage and Remission Prediction in Malaysian Lupus Nephritis

Muhammad Izzul Islam Faisal[1*], Muhammad Ezairiel Azieq Shahromnizal[1], Shahnorbanun Sahran[1], Syahrul Sazliyana Shaharir[2]

[1] *Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
[2] *Faculty of Medicine, Universiti Kebangsaan Malaysia / Hospital Canselor Tuanku Muhriz, 56000 Cheras, Kuala Lumpur, Malaysia*
[*]**Corresponding Author:** A200363@siswa.ukm.edu.my

ABSTRACT

Background and Objectives — Lupus nephritis is a severe renal manifestation of systemic lupus erythematosus that may progress to chronic kidney disease and delayed remission, yet patient-level risk prediction remains difficult because clinical patterns are heterogeneous and differ across ethnic groups. This study developed and evaluated SLECare, an explainable machine learning clinical decision-support framework for predicting kidney damage and remission outcomes among Malaysian lupus nephritis patients. Methods — Retrospective clinical data from Hospital Canselor Tuanku Muhriz covering demographic, clinical, laboratory, biopsy, treatment, and outcome variables were processed through clinically guided cleaning, leakage-aware feature removal, stratified splitting, imputation, feature engineering, and recursive feature elimination. CatBoost, random forest, multilayer perceptron, support vector machine, weighted soft voting, and stacked ensemble models were trained across baseline, six-month, and twelve-month scenarios and evaluated using precision, recall, and F1-score. Shapley additive explanations provided interpretability. Results — CatBoost delivered consistently strong and clinically interpretable performance, the stacked ensemble achieved the best six-month kidney damage prediction (F1 = 0.80), and the twelve-month CatBoost model produced the strongest remission prediction (F1 = 0.77). Influential drivers included baseline creatinine, ethnicity, lupus anticoagulant, chronicity index, treatment timing, and global sclerosis. Conclusions — SLECare shows potential as an explainable, dashboard-ready tool supporting earlier and locally relevant risk stratification and clinician decision-making in Malaysian lupus nephritis care.

Keywords: *Systemic lupus erythematosus; Explainable artificial intelligence; CatBoost; SHapley additive explanations; Clinical decision support; Renal outcome prediction*

1. INTRODUCTION

Systemic lupus erythematosus is a chronic autoimmune disease with highly heterogeneous clinical manifestations, treatment responses, and long-term complications, and it may affect the skin, joints, blood, nervous system, cardiovascular system, and kidneys. Among its severe manifestations, lupus nephritis is one of the most clinically important because renal involvement can progress to chronic kidney disease, end-stage renal disease, and irreversible organ damage. In Malaysia, systemic lupus erythematosus has been reported at approximately 40–60 cases per 100,000 population, and around 40–60% of these patients may develop lupus nephritis. The risk of end-stage renal disease remains substantial, reaching roughly 3–11% within five years and 19–25% within fifteen years after diagnosis. Earlier identification of patients at risk of kidney damage and delayed remission is therefore important to support timely treatment adjustment, closer monitoring, and more proactive clinical decision-making.

Lupus nephritis is difficult to predict because disease progression differs between patients: some achieve remission after induction therapy, while others experience delayed remission despite treatment. This challenge is more complex in Malaysia because the patient population is multi-ethnic and clinical patterns may differ across ethnic groups. A Malaysian study on proliferative lupus nephritis showed that renal response and remission outcomes were influenced by ethnicity,

hypertension, recurrent nephritis, and treatment delay, with Malay patients carrying a higher risk of delayed remission than non-Malay patients (Syahrul Sazliyana et al., 2021). These findings indicate that risk prediction for Malaysian lupus nephritis should not rely solely on general international evidence but should also be developed from local clinical data that reflect Malaysian patients.

Conventional statistical methods have contributed important knowledge by identifying associations between clinical factors and renal outcomes, but statistical association alone does not directly provide patient-level risk prediction. This limits usefulness when clinicians need to estimate whether an individual patient is more likely to develop kidney damage or experience delayed remission. Lupus nephritis datasets also contain mixed demographic, clinical, laboratory, biopsy, treatment, and outcome variables that may interact in non-linear ways, making the task well suited to machine learning. Recent studies have increasingly applied machine learning to systemic lupus erythematosus and autoimmune disease research because such models can learn complex patterns from high-dimensional clinical data (Adamichou et al., 2021; Chen et al., 2021; Gu et al., 2021).

This study introduces SLECare, an explainable machine learning clinical decision-support innovation for Malaysian lupus nephritis. The specific objectives are to: (1) develop predictive models for two clinically important outcomes, namely kidney damage represented by chronic kidney disease and delayed remission; (2) compare CatBoost with benchmark and ensemble models across baseline, six-month, and twelve-month scenarios; (3) apply Shapley additive explanations to identify clinically meaningful risk drivers; and (4) translate the validated outputs into a clinician-facing dashboard. The remainder of the paper reviews related work, describes the methodology, presents the results, discusses their clinical implications and limitations, and concludes.

2. LITERATURE REVIEW

Machine learning has been applied to systemic lupus erythematosus because it can learn complex, high-dimensional relationships that are difficult to capture with classical statistics. Adamichou et al. (2021) developed the SLE risk probability index, a clinician-friendly model to assist diagnosis, while Chen et al. (2021) and Gu et al. (2021) applied machine learning to diagnosis and disease-activity prediction using clinical and laboratory features. Across this body of work, the dominant emphasis has been on diagnosis or disease classification rather than on predicting long-term clinical outcomes such as renal damage or remission.

Explainability has become a central requirement for clinical artificial intelligence, because a model used to support patient monitoring or treatment should not only output a class label but also justify why a prediction was made. AlShareedah et al. (2023) demonstrated this in an Oman-based systemic lupus erythematosus cohort, in which CatBoost achieved strong predictive performance and Shapley additive explanations were used to explain the factors contributing to individual predictions, supporting early clinical suspicion. However, that study addressed systemic lupus erythematosus classification in an Omani cohort rather than Malaysian lupus nephritis outcomes such as kidney damage and remission.

Within the Malaysian context, Syahrul Sazliyana et al. (2021) identified important ethnic disparities and other determinants of renal response, but these findings have not been translated into an explainable, patient-level prediction system. More broadly, Ceccarelli et al. (2023) reviewed machine learning in systemic lupus erythematosus and emphasized that clinical value depends on careful methodology, valid data processing, and interpretable outputs that clinicians can understand.

Existing studies therefore leave three gaps. First, most machine learning work on systemic lupus erythematosus targets diagnosis or classification rather than long-term outcome prediction. Second, Malaysian lupus nephritis research has identified relevant renal-response factors but has not produced an explainable patient-level prediction tool. Third, many healthcare artificial intelligence models remain difficult to interpret, reducing clinical usability. SLECare addresses these gaps by combining local Malaysian clinical data, dual-outcome predictive modelling, individual and ensemble model comparison, and Shapley-based clinical interpretation within a dashboard-ready framework.

3. RESEARCH METHODOLOGY

3.1 Research Design

This study adopted a machine learning-based clinical prediction design to develop and evaluate SLECare. The work focused on two prediction outcomes: kidney damage risk, represented by chronic kidney disease, and remission outcome, represented by delayed remission status. The design was intended not only to compare predictive performance but also to

evaluate whether model outputs could be explained in a clinically meaningful way and translated into a practical clinician-facing dashboard. Three time-based scenarios — baseline, six-month, and twelve-month — were defined to test whether follow-up clinical information improves prediction.

3.2 Participants / Samples

Retrospective clinical data were obtained from lupus nephritis patients treated at Hospital Canselor Tuanku Muhriz, Universiti Kebangsaan Malaysia, between 2000 and 2020. The dataset contained demographic, clinical manifestation, laboratory, biopsy, treatment, and outcome variables. Records were included if they were suitable for machine learning analysis after data quality checking and clinical review. Before feature removal and preprocessing, the dataset comprised more than 200 treatment records and approximately 74 clinical variables. The data were divided into training and testing sets using stratified train-test splitting to preserve the distribution of each target class in both sets.

3.3 Instruments and Materials

CatBoost served as the main model because it is well suited to structured clinical data and handles categorical variables effectively. Random forest, multilayer perceptron, and support vector machine models were developed as benchmarks, and two ensemble strategies — weighted soft voting and a stacked ensemble — were evaluated as enhancements. Shapley additive explanations were used for interpretation because they can be linked directly to the original clinical variables. All modelling and analysis were implemented in Python using established open-source libraries, including scikit-learn, CatBoost, and the SHAP package.

3.4 Data Collection Procedures

Data were extracted retrospectively from de-identified hospital records. Guided by a data quality plan and clinical expert input, irrelevant identifiers, constant variables, and post-outcome leakage-prone variables were removed so that models learned only from clinically available predictors rather than from information that would not exist at the intended prediction time. Examples of removed variables included patient identifiers and fields that directly encoded future outcome information. Ethical approval and the use of de-identified clinical data are described in the Ethical Approval and Consent to Participate statement.

3.5 Data Analysis

Preprocessing was performed using training-derived information to reduce leakage. Missing numerical values were imputed with training-set medians, while selected categorical missing values were imputed using clinically guided codes representing negative, not-tested, or mode-based values according to feature type. Skewed numerical variables were log-transformed where appropriate, and clinically meaningful extreme values were retained. Feature engineering derived the interval between original diagnosis and active lupus nephritis onset to represent the timing of renal involvement. Recursive feature elimination was then applied to retain the most relevant predictors for each target. Feature sets were separated by scenario: the baseline scenario used pre-treatment variables only; the six-month scenario added early follow-up variables; and the twelve-month scenario added three-month and six-month variables. Class imbalance in the chronic kidney disease target was handled using class weighting for CatBoost and training-only synthetic minority oversampling for selected non-CatBoost models, whereas the more balanced delayed-remission target did not require this. Models were tuned using grid search with cross-validation on the training data and evaluated on the held-out test set using precision, recall, F1-score, and confusion matrices. Global and local Shapley additive explanations were computed primarily for the CatBoost model.

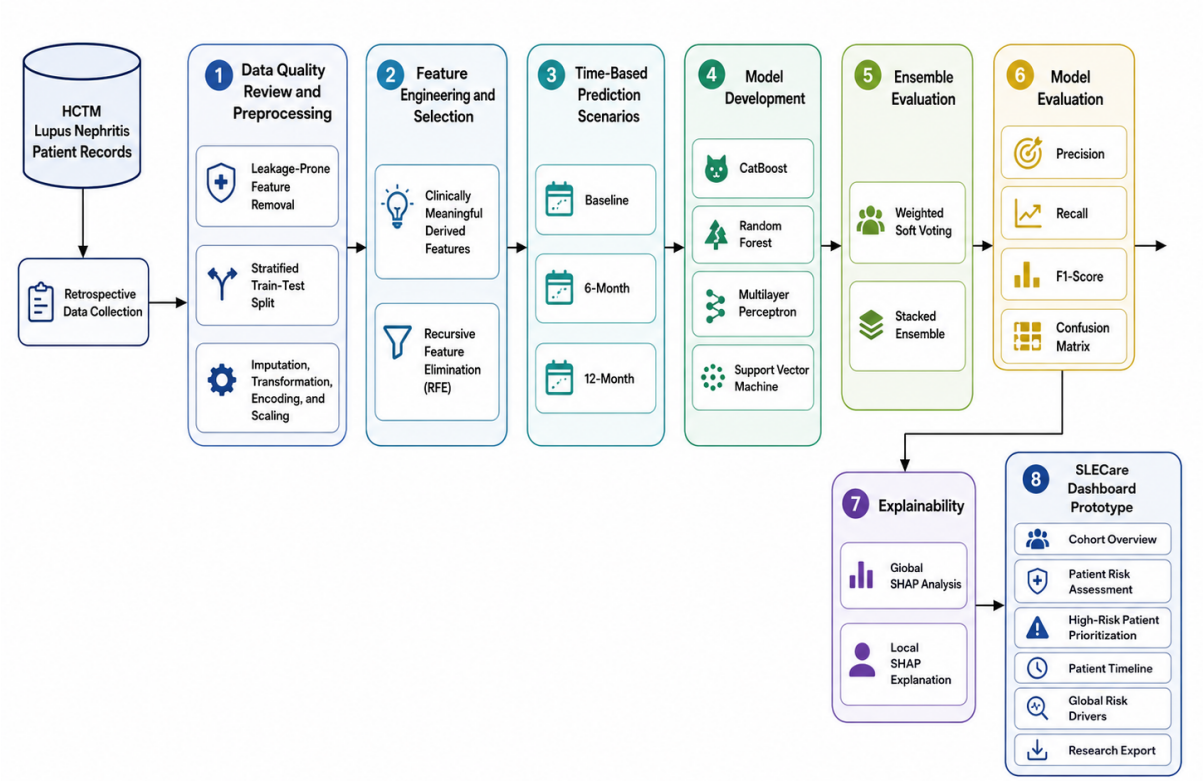

Figure 1. *SLECare methodological workflow.*

3.6 System Architecture and Explainable Inference

The validated models were deployed in a three-tier decision-support application (Figure 2). A React and TypeScript front end presents the dashboard and communicates through a compact JSON REST interface with a Flask application tier, which delegates all domain logic to a single service layer. The service layer loads the serialised CatBoost models together with their recursive-feature-elimination feature lists, performs cohort-based imputation and categorical encoding, computes predictions and explanations, and applies process-level memoisation and a bounded least-recently-used prediction cache to keep interactive response times low. The anonymised cohort, serialised models, and precomputed global SHAP plots are served as static assets.

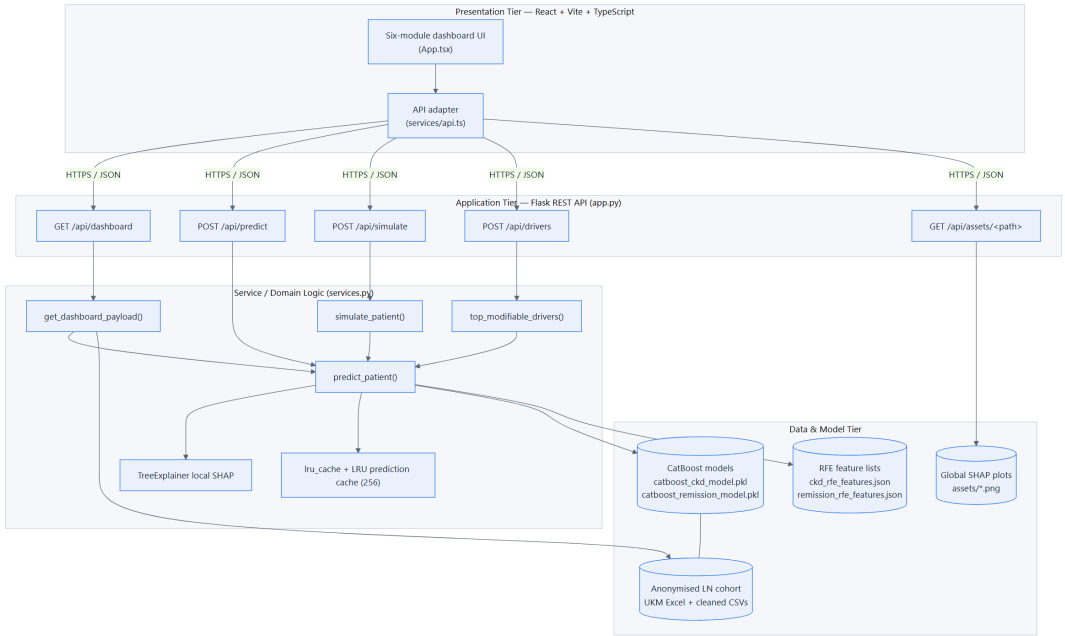

Figure 2. *SLECare three-tier system architecture: a React/TypeScript front end, a Flask REST application tier, and a single service layer performing CatBoost inference, SHAP explanation, and two-layer caching over the anonymised cohort, serialised models, and SHAP assets.*

All predictions follow a single explainable inference pipeline (Figure 3) that is reused by every dashboard feature. Patient inputs are completed with cohort-derived values and cast to the categorical types expected by CatBoost; each model then outputs a probability that is mapped to a Low, Moderate, or High risk band, while a tree-based SHAP explainer ranks the local feature contributions into the strongest risk-increasing and protective factors. A constrained presentation layer converts this structured output into clinician-readable text and never itself generates predictions, ensuring that every explanation is consistent with the model that produced it.

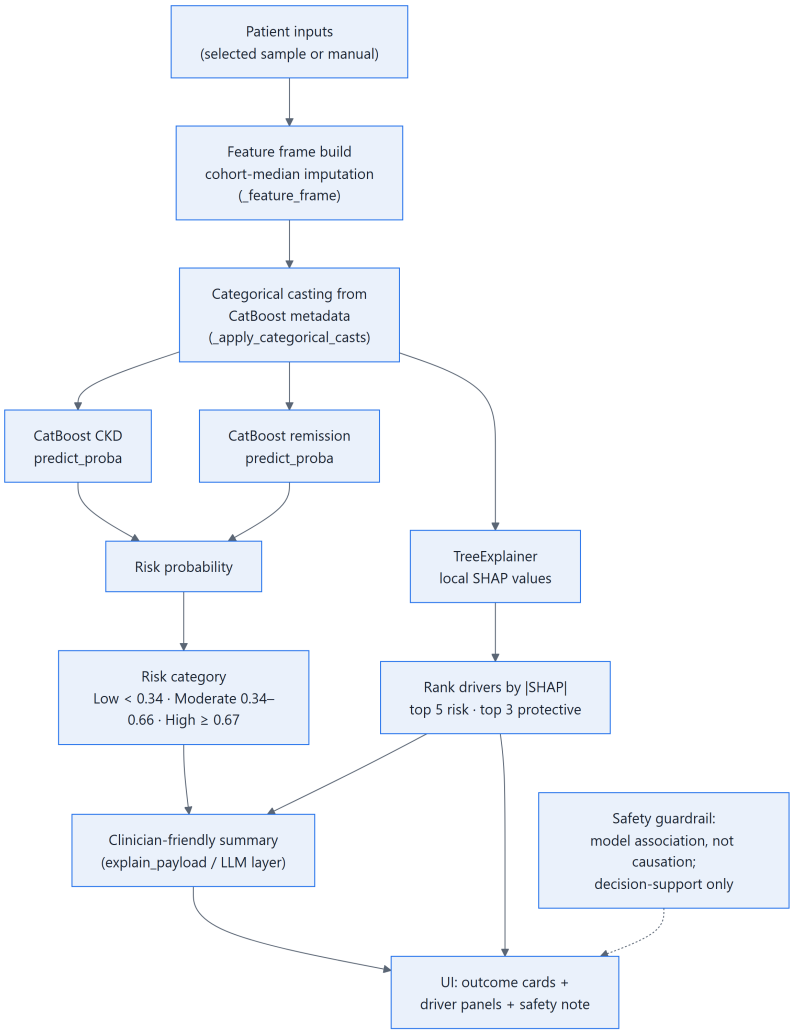

Figure 3. *Explainable patient-level inference pipeline shared by all dashboard features: input completion and categorical casting feed both CatBoost models, probabilities are mapped to risk bands, and local SHAP values are ranked into risk-increasing and protective drivers before a constrained layer renders clinician-readable text.*

4. RESULTS

Evaluation focused on balanced clinical performance using precision, recall, and F1-score rather than accuracy alone, because the clinical aim was to detect high-risk patients while limiting false alarms. Table 1 summarizes the best-performing model for each outcome and time-based scenario.

Table 1. *Summary of best model performance across prediction outcomes and time-based scenarios.*

Outcome	Best Model	Scenario	Precision	Recall	F1-score
Kidney damage prediction	CatBoost	Baseline	0.6700	0.8300	0.7400
Kidney damage prediction	Stacked ensemble	6-month	0.7692	0.8333	0.8000
Kidney damage prediction	CatBoost	12-month	0.7143	0.8333	0.7692
Remission prediction	CatBoost	Baseline	0.6522	0.6522	0.6522
Remission prediction	CatBoost	6-month	0.7391	0.7391	0.7391
Remission prediction	CatBoost	12-month	0.8095	0.7391	0.7727

Note: *Values are reported on the held-out test set. The best-performing configuration is shown for each outcome and time-based scenario; bold rows in the original analysis correspond to the strongest F1-score per outcome.*

Figure 4 visualises the results in Table 1, showing that recall is generally maintained at or above precision — appropriate for a screening setting that prioritises detection of high-risk patients — and that performance tends to improve as follow-up information accumulates from the baseline to the twelve-month scenario.

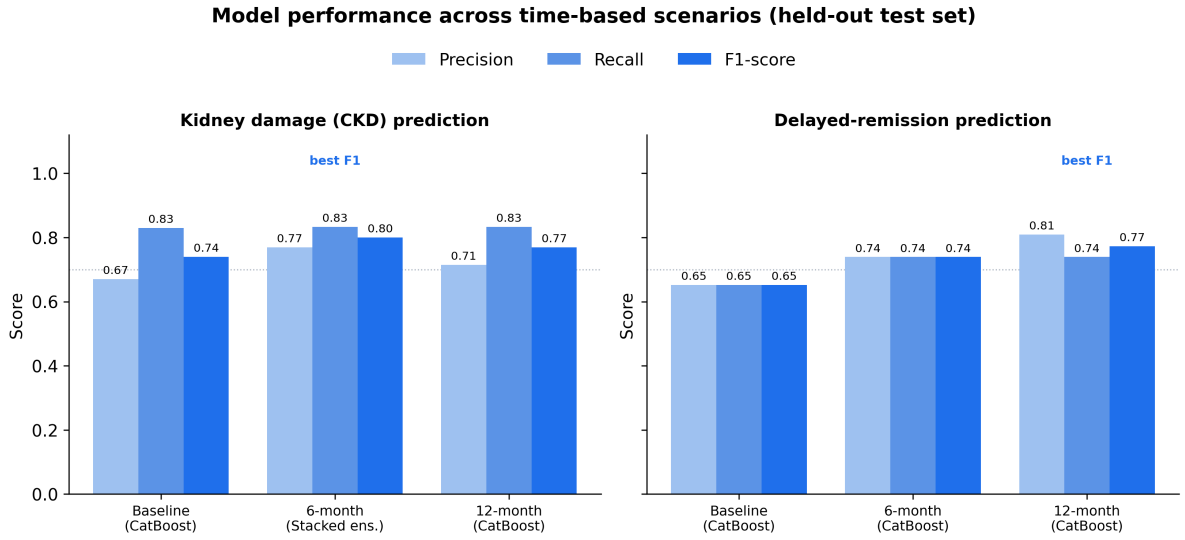

Figure 4. *Precision, recall, and F1-score of the best model in each outcome and time-based scenario on the held-out test set (values from Table 1).*

The baseline clinical characteristics of the cohort are summarised in Table 2.

Table 2. *Baseline clinical characteristics of the lupus nephritis cohort (median, mean, and standard deviation of key renal, immunological, and biopsy markers).*

Variable	n	Median	Mean	SD
Baseline creatinine	213	75.0	95.01	65.22
Creatinine 24 mth	214	68.3	91.0	90.08
Baseline albumin	212	30.0	29.25	8.61
Pretreatment C3	209	48.0	54.54	29.6
Pretreatment C4	208	9.73	11.17	8.34
Pretreatment UPCI	211	0.35	0.63	1.55
Chronicity index	179	3.0	3.26	2.51
Activity index	178	7.0	7.47	4.77

For kidney damage prediction, CatBoost provided balanced performance at baseline and at twelve months, while the stacked ensemble achieved the strongest six-month result (F1 = 0.80). For remission prediction, CatBoost was the strongest model across all scenarios, with the twelve-month model producing the best overall result (F1 = 0.77). Figure 5 presents the global Shapley feature importance for kidney damage prediction.

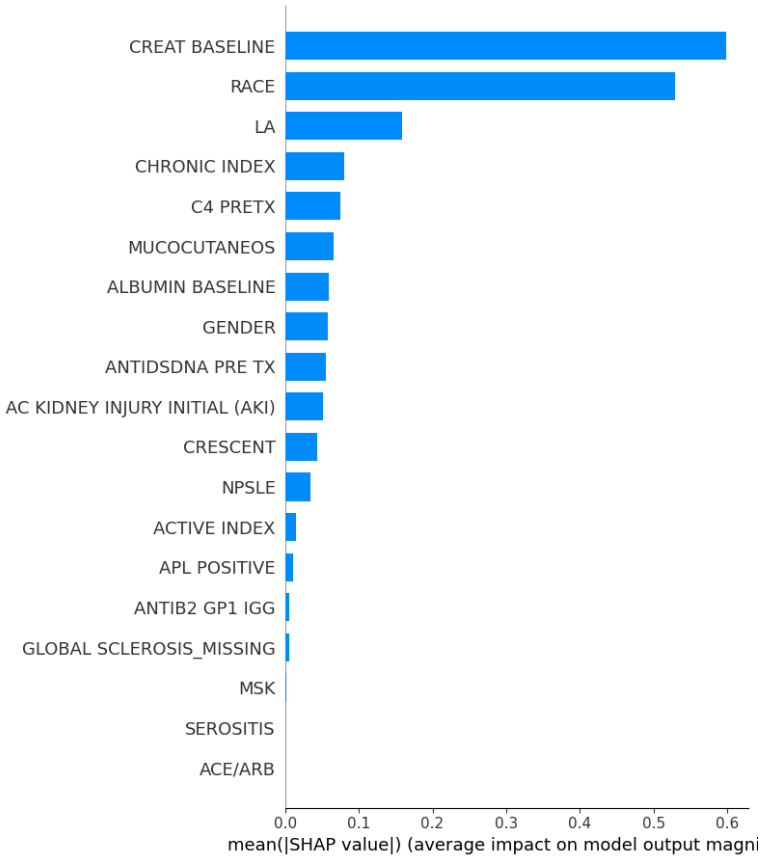

Figure 5. *SHAP global feature importance for kidney damage prediction.*

The global Shapley plot for kidney damage prediction indicated that baseline creatinine had the highest average impact on model output, followed by race, lupus anticoagulant, chronicity index, C4 level, mucocutaneous involvement, baseline albumin, gender, anti-dsDNA status, acute kidney injury, crescent formation, and neuropsychiatric lupus. Figure 6 shows the corresponding dependence plots.

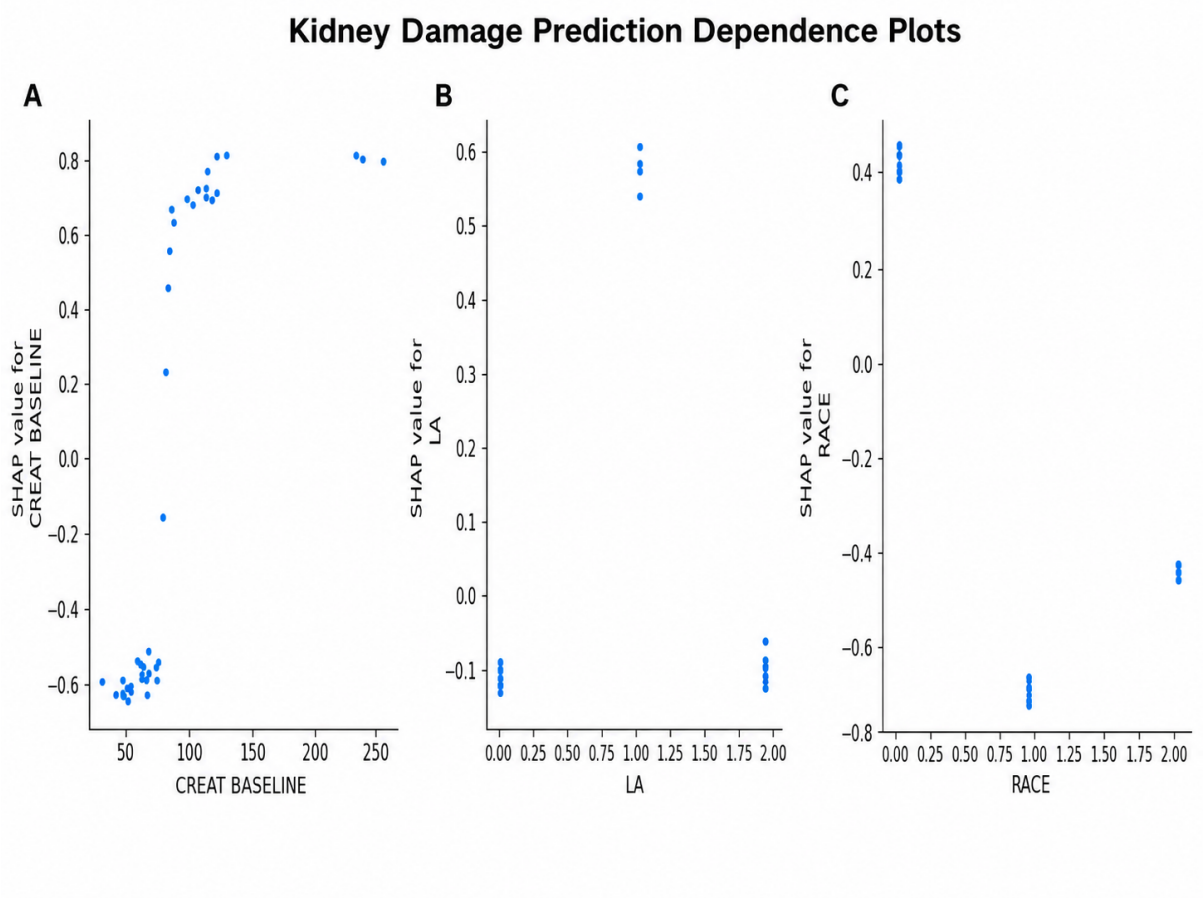

Figure 6. *SHAP dependence plots for kidney damage prediction.*

In the dependence plots, baseline creatinine produced mostly negative Shapley values below approximately 70, increased sharply between about 80 and 90, and reached strongly positive contributions above 100. Lupus anticoagulant category 1 produced positive contributions of approximately 0.55 to 0.60, while categories 0 and 2 contributed near −0.10. Race code 0 contributed positively, whereas race codes 1 and 2 contributed negatively. Figure 7 presents the global Shapley feature importance for remission prediction.

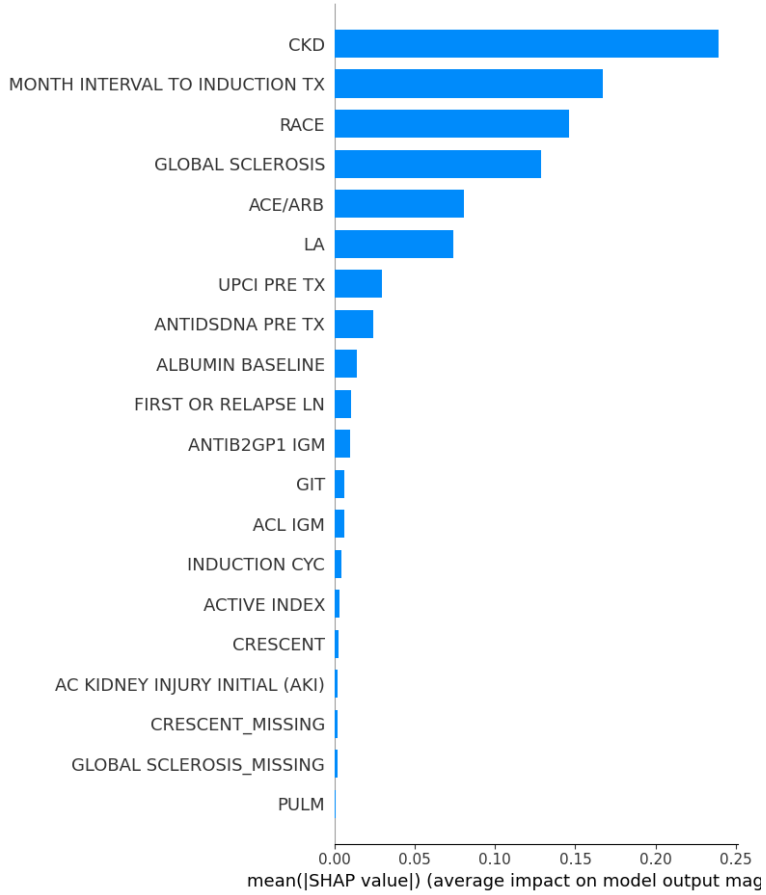

Figure 7. *SHAP global feature importance for remission prediction.*

For remission prediction, chronic kidney disease status had the highest average impact, followed by month interval to induction treatment, race, global sclerosis, ACE/ARB use, lupus anticoagulant, urine protein-creatinine index before treatment, anti-dsDNA status, baseline albumin, and first-or-relapse lupus nephritis status. Figure 8 shows the dependence plots.

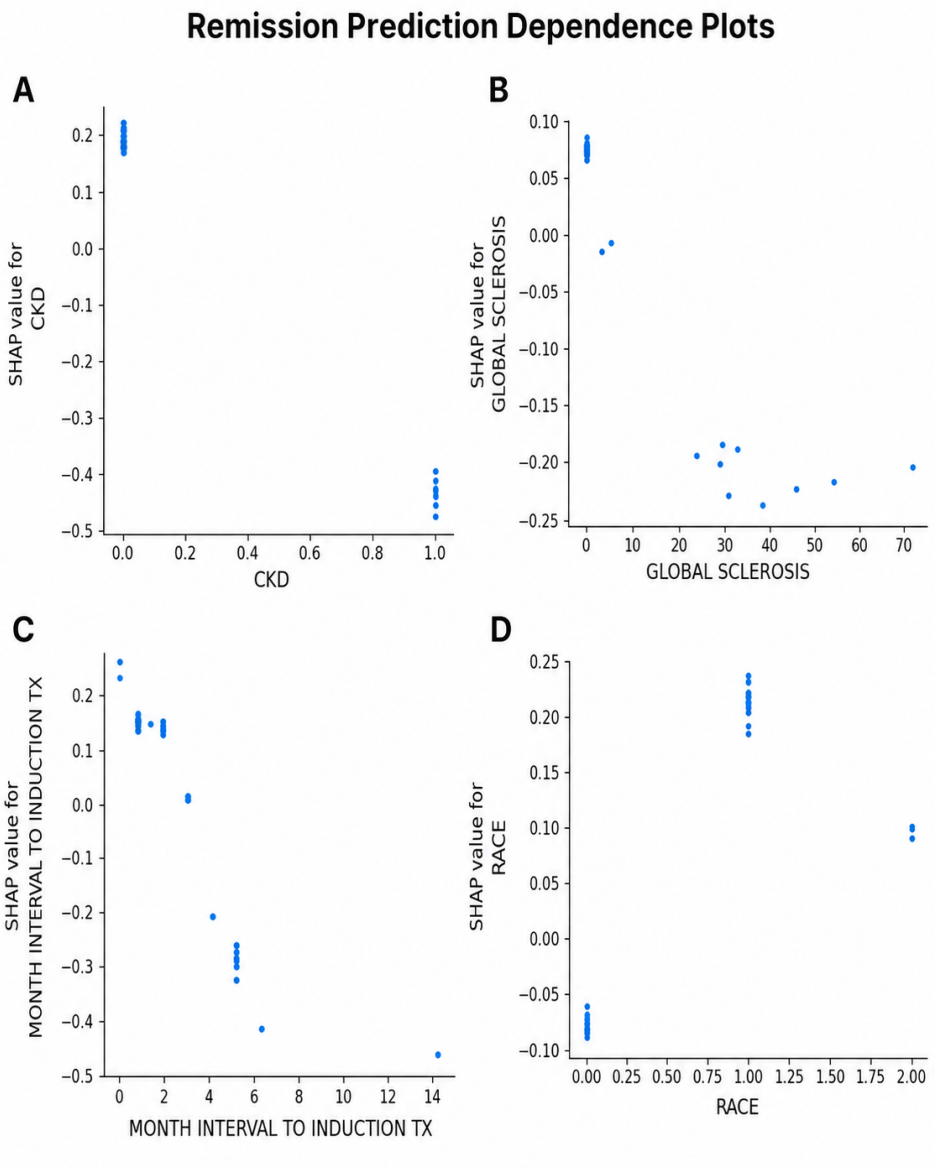

Figure 8. *SHAP dependence plots for remission prediction.*

In the remission dependence plots, chronic kidney disease status of 0 contributed positively (approximately 0.17 to 0.22) while a status of 1 contributed negatively (approximately −0.35 to −0.48). Higher global sclerosis values and longer intervals to induction treatment, particularly beyond about four to six months, produced increasingly negative contributions, and race codes again showed distinct effects. Figure 9 presents the SLECare clinician-facing dashboard.

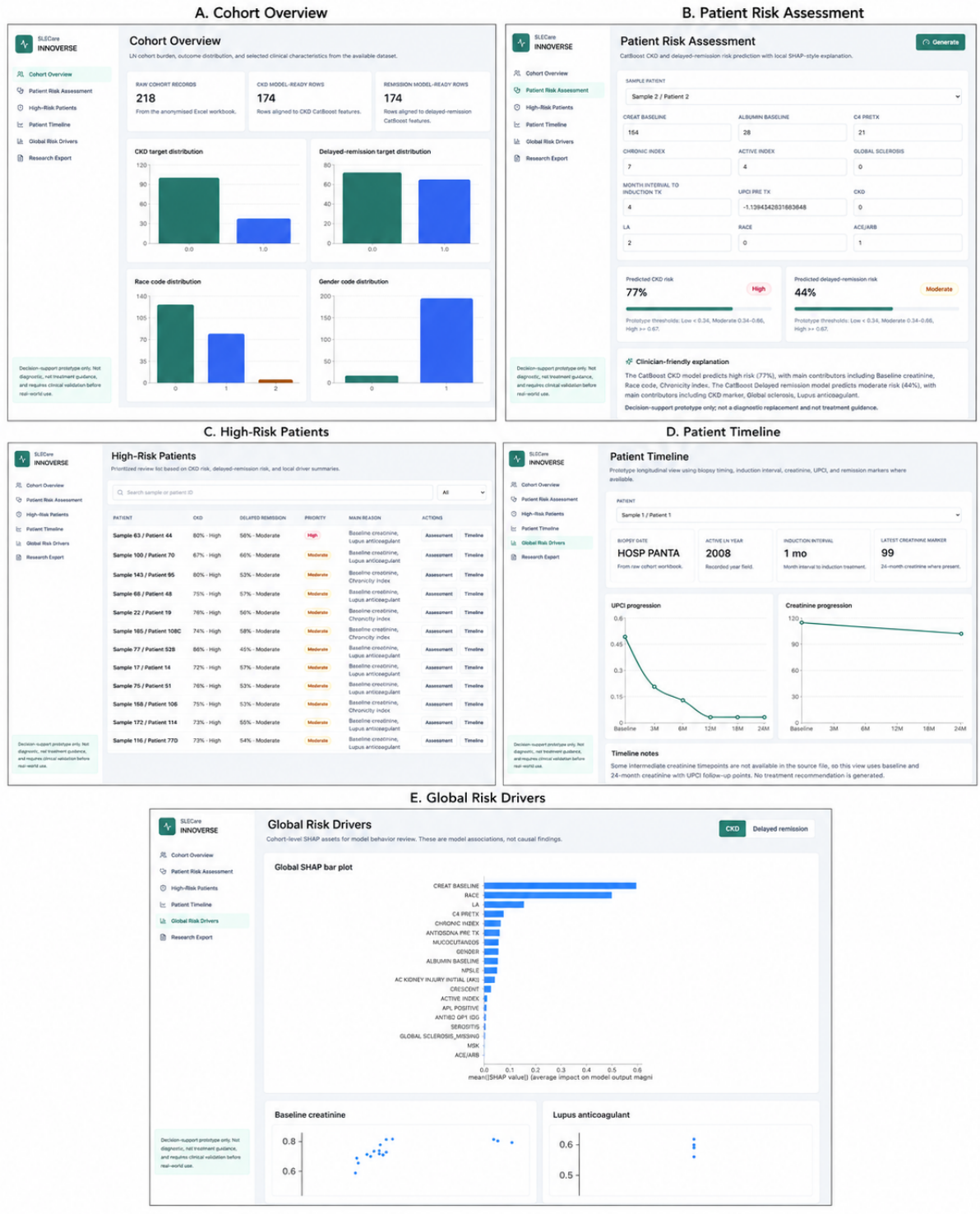

Figure 9. *SLECare clinician-facing dashboard prototype.*

The dashboard comprises a cohort overview, an individual patient risk-assessment view with clinician-friendly explanations, a high-risk patient prioritization list, a patient timeline for longitudinal monitoring of biopsy timing, induction interval, creatinine, and urine protein-creatinine index trends, and a cohort-level global risk-driver view based on Shapley explanations.

4.1 Interactive and Explainable Decision Support

In addition to static risk scores, SLECare provides two interactive features built on the same inference pipeline. The What-If Explorer (Figure 10) allows a clinician to modify one or more clinical variables for a selected patient; the edited record is re-evaluated and the baseline and simulated predictions are compared to display the signed change in predicted probability, its direction, and any shift in risk band. The Top Modifiable Driver Engine (Figure 11) automates this

analysis: for each clinically adjustable variable it shifts the value toward the healthier quartile observed in the cohort, keeping the input within the observed data range, re-runs the prediction, and ranks the variables by the resulting reduction in predicted risk together with a deliberately conservative confidence label. Both features are presented as exploratory analyses of model associations rather than treatment recommendations.

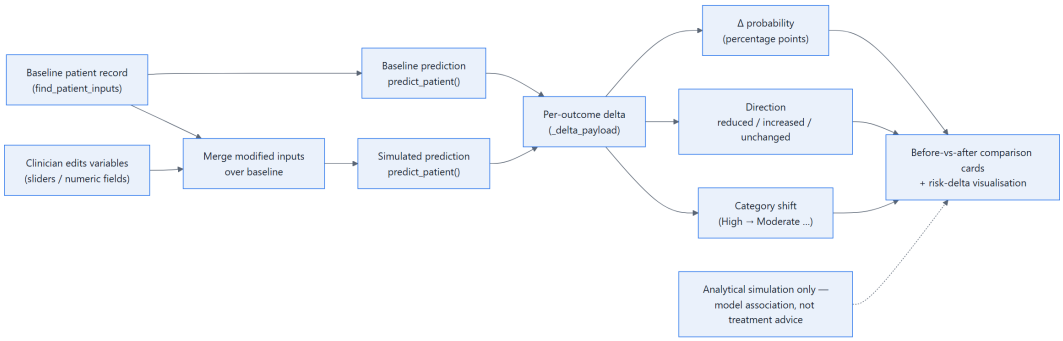

Figure 10. *Clinical What-If Explorer: clinician edits are merged over a baseline record and re-evaluated, and baseline-versus-simulated predictions are compared as signed probability changes, directions, and risk-band shifts.*

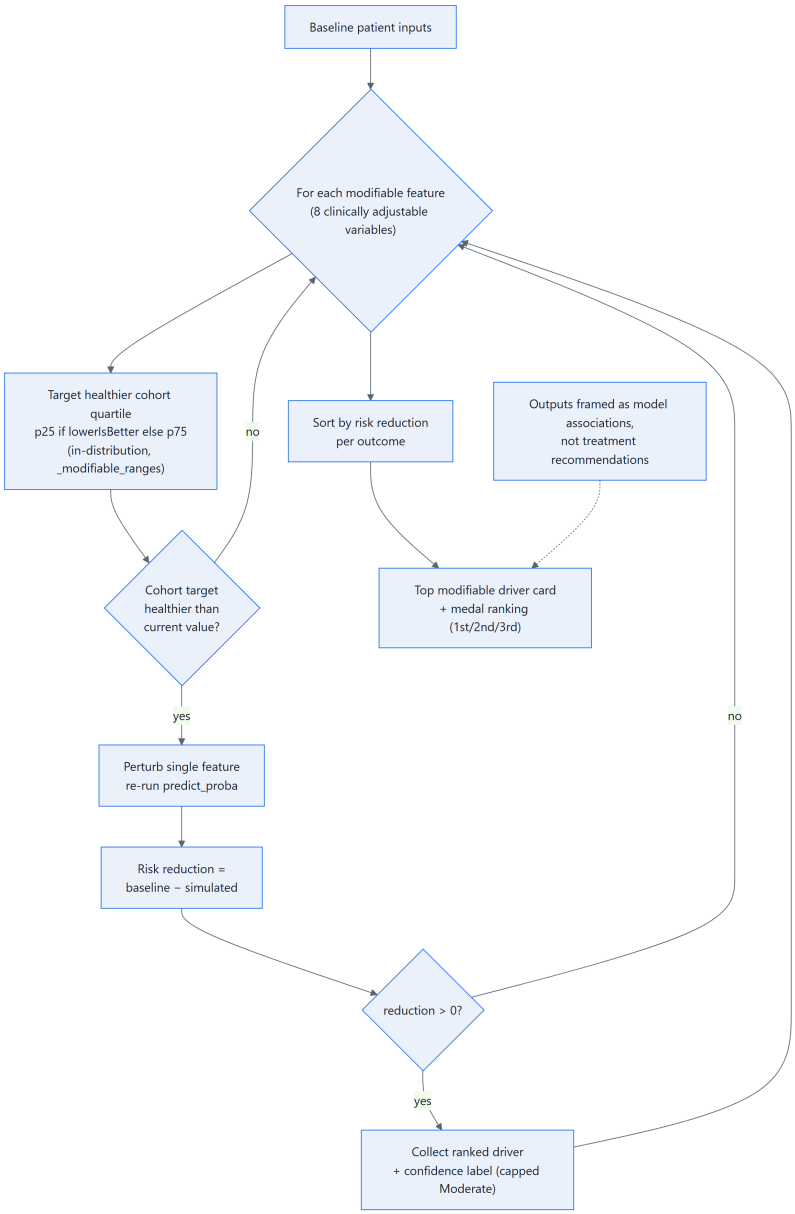

Figure 11. *Top Modifiable Driver Engine: each adjustable variable is shifted toward the healthier in-range cohort quartile and the prediction re-run; variables are ranked by the resulting reduction in predicted risk with a conservative confidence label and explicit association-not-causation framing.*

5. DISCUSSION

The comparison between individual and ensemble models revealed an important performance–interpretability trade-off. CatBoost remained the most suitable model for clinical-facing deployment because it achieved balanced performance while allowing Shapley explanations to be linked directly to the original clinical variables, which clinicians need in order to understand the factors behind each prediction. Although the stacked ensemble achieved the strongest six-month kidney damage result, it is less transparent because the meta-model learns from base-model outputs rather than from patient features directly. Weighted soft voting was useful as a research benchmark by combining probability outputs from different model families, but it did not consistently justify replacing CatBoost as the deployment model. The enhanced ensembles are therefore best treated as research-performance benchmarks, while baseline CatBoost remains the most practical model for early dashboard deployment.

The Shapley analyses identified clinically coherent risk drivers. For kidney damage, baseline creatinine dominated, consistent with the established role of pre-treatment renal function in chronic kidney disease risk, and immunological and biopsy markers such as lupus anticoagulant, chronicity index, and crescent formation contributed in clinically plausible directions. For remission, existing chronic kidney disease, structural renal damage reflected by global sclerosis, and treatment timing were the most influential, aligning with the clinical understanding that established damage and delayed induction reduce the likelihood of remission.

These findings are consistent with and extend prior work. They reinforce the value of combining CatBoost with Shapley explanations demonstrated by AlShareedah et al. (2023) in an Omani systemic lupus erythematosus cohort, but apply it to dual renal outcomes in a Malaysian lupus nephritis population. The observed influence of ethnicity is consistent with Syahrul Sazliyana et al. (2021), who reported ethnic disparities in renal response, and the importance of treatment timing echoes evidence that treatment delay worsens outcomes. By translating these explanations into a clinician-facing dashboard, SLECare moves beyond statistical association toward patient-level risk stratification, high-risk prioritization, and longitudinal monitoring.

The interactive What-If and modifiable-driver analyses further distinguish SLECare from a passive predictor by letting clinicians explore, within the observed data range, how changes in modifiable factors relate to predicted risk; because these analyses reflect model associations rather than causal effects, they are reported with conservative confidence labelling and explicit non-causal framing.

Several limitations should be acknowledged. The dataset was modest in size and drawn from a single centre, and the retrospective design constrains causal interpretation; some scenarios produced only moderate performance, so the models should be regarded as decision-support aids rather than standalone tools. The ethnicity, or race, variable is encoded categorically and should be interpreted with particular caution: its model effects reflect statistical patterns within this cohort and may capture socioeconomic, access-to-care, or unmeasured clinical factors rather than any biological cause, so it must not be read as a deterministic or causal driver. External validation on larger, multi-centre, and prospective cohorts is needed, together with model calibration and fairness auditing of ethnicity-related effects, before clinical deployment.

6. CONCLUSION

SLECare was developed as an explainable machine learning framework for predicting kidney damage and remission outcomes among Malaysian lupus nephritis patients. CatBoost provided strong and clinically interpretable performance, while ensemble modelling offered additional research value for performance comparison, with the stacked ensemble giving the best six-month kidney damage result and the twelve-month CatBoost model the best remission result. Shapley analysis identified meaningful clinical drivers spanning renal, immunological, demographic, biopsy, and treatment-related factors, and the prototype dashboard translated these outputs into a practical decision-support tool for risk assessment, high-risk prioritization, longitudinal monitoring, and explainable review. By integrating local clinical data, dual-outcome prediction, and interpretable outputs, SLECare addresses the need for earlier, explainable, and locally relevant risk stratification in Malaysian lupus nephritis care. Future work should pursue external and prospective validation on larger multi-centre cohorts, formal calibration, and fairness auditing of ethnicity-related effects to support safe clinical adoption.

ACKNOWLEDGEMENTS

The authors thank Universiti Kebangsaan Malaysia and Hospital Canselor Tuanku Muhriz for supporting the development of SLECare and for facilitating access to the clinical data used in this study.

AUTHOR CONTRIBUTIONS

Author	Contribution (CRediT Taxonomy)
Muhammad Izzul Islam Bin Faisal	Conceptualization, Methodology, Software, Formal Analysis, Writing – Original Draft
Muhammad Ezairiel Azieq Bin Shahromnizal	Software, Data Curation, Visualization, Validation
Shahnorbanun Binti Sahran	Supervision, Methodology, Writing – Review & Editing
Syahrul Sazliyana Shaharir	Clinical Resources, Validation, Writing – Review & Editing, Funding Acquisition

Note: *Contributions follow the CRediT taxonomy. All listed authors meet ICMJE authorship criteria.*

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The clinical dataset analyzed in this study is not publicly available because it contains sensitive patient information. De-identified data may be made available from the corresponding author on reasonable request and subject to the necessary institutional and ethical approvals.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This study used retrospective, de-identified clinical data and was conducted with the approval of the relevant Research Ethics Committee of Universiti Kebangsaan Malaysia / Hospital Canselor Tuanku Muhriz (Approval No.: [to be inserted]) in accordance with the Declaration of Helsinki. The requirement for individual informed consent was handled in line with the approving committee's conditions for the use of de-identified retrospective records.

FUNDING

Participation in INNOVERSE 2026 was supported by HEP-UKM through Tabung Program Mobiliti Tanpa Kredit — TAP317 (Luar Negara) and by the Faculty of Information Science and Technology (FTSM), Universiti Kebangsaan Malaysia. This research was further supported by research grants TT-2024-002 and Tubitak -MIGHT Collaboration 2.0.

REFERENCES

Adamichou, C., Genitsaridi, I., Nikolopoulos, D., Nikoloudaki, M., Repa, A., Bortoluzzi, A., … Tzanakaki, G. (2021). Lupus or not? SLE risk probability index (SLERPI): A simple, clinician-friendly machine learning-based model to assist the diagnosis of systemic lupus erythematosus. *Annals of the Rheumatic Diseases, 80*(6), 758–766. https://doi.org/10.1136/annrheumdis-2020-219069

AlShareedah, A., Zidoum, H., Al-Sawafi, S., Al-Lawati, B., & Al-Ansari, A. (2023). Machine learning approach for predicting systemic lupus erythematosus in an Oman-based cohort. *Sultan Qaboos University Medical Journal, 23*(3), 328–335. https://doi.org/10.18295/squmj.12.2022.069

Ceccarelli, F., Natalucci, F., Picciariello, L., & Tuscano, D. (2023). Application of machine learning models in systemic lupus erythematosus. *Journal of Clinical Medicine, 12*(11), 3761.

Chen, Y., Liu, J., Zhang, H., Wang, Y., Li, X., & Zhang, F. (2021). Machine learning-based prediction and classification of systemic lupus erythematosus using clinical and laboratory features. *Frontiers in Immunology, 12*, 756752.

Gu, Y., Huang, Y., Chen, X., Wang, Y., Chen, J., & Liu, Z. (2021). Machine learning for systemic lupus erythematosus diagnosis and disease activity prediction: A clinical data-based study. *Frontiers in Medicine, 8*, 746810.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

Syahrul Sazliyana, S., Mohd, R., Kamaruzaman, L., Wan Daud, W. R., Abu Shamsi, Y., Lim, K. Y., Mustafar, R., & Abdul Gafor, A. H. (2021). Ethnic disparities in renal response among proliferative lupus nephritis patients: A single-centre study from Malaysia. *Lupus, 30*(5), 767–776.

For submission inquiries, contact the Editorial Office:
ocso.info@gmail.com | www.ijstemr.org
International Journal of STEM Research — Indexed in Scopus, Web of Science, DOAJ