MSR244

Methodological Characteristics of Pharmacogenomics-Integrated AI Clinical Decision Support Systems (AI-CDSS): A Systematic Review

Shinyoung Park[1,2], Hae Sun Suh[1,2,3*]

1 Department of Regulatory Science, Graduate School, Kyung Hee University, Seoul, Republic of Korea
2 Institute of Regulatory Innovation through Science (IRIS), Kyung Hee University, Seoul, Republic of Korea
3 College of Pharmacy, Kyung Hee University, Seoul, Republic of Korea

*Corresponding author

PEBD
Pharmaceutical Economics
Big Data Analysis and Policy Lab

Background

- **Pharmacogenomics (PGx)** has the potential to support patient-centered precision medicine. However, Evidence on **artificial intelligence-based clinical decision support systems (AI-CDSS)** integrating PGx remains fragmented.
- This systematic review examined the methodological characteristics of PGx-integrated AI-CDSS modeling and validation studies.

Methods

Study Design

- Systemic Review (following PRISMA 2020 guidelines)

Databases & Search Period

- Sources: PubMed, EMBASE, Cochrane Library, and ACM Digital Library
- Search Period: Through December 31, 2025

Keywords

No.	Category	Query
#1	AI/ML	"Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Algorithms"[Mesh] OR "Deep Learning"[Mesh] OR "Neural Networks, Computer"[Mesh]
#2	CDSS	"Decision Support Systems, Clinical"[Mesh] OR "Clinical Decision Support" OR "CDSS"
#3	Pharmacogenetics	"Pharmacogenetics"[Mesh] OR "Pharmacogenomics" OR "Genotype-guided"
#4	Model/Validation	"Models, Theoretical"[Mesh] OR "Modeling" OR "Validation" OR "Performance" OR "Accuracy"
Final Combined		#1 AND #2 AND #3 AND #4

Inclusion & Exclusion Criteria

- Inclusion criteria
  - Studies utilizing pharmacogenomic data as a primary input.
  - Studies focusing on pharmacological interventions or outcomes.
- Exclusion criteria:
  - Studies lacking pharmacogenomic data analysis.
  - Studies restricted to disease screening or prognostic diagnosis excluding treatment-related outcomes.

Analysis

- Quality Assessment: Risk of bias and applicability were assessed using the PROBAST (Prediction model Risk Of Bias Assessment Tool) framework.
- Data Extraction: Key parameters included clinical domain, PGx inputs, machine learning algorithms, validation strategies, explainability techniques, and performance metrics.

Acknowledgment

- This research was supported by grants from the Ministry of Food and Drug Safety (RS-2024-00331719) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00345981) in 2026.

References

1. Doig KD et al. Using artificial intelligence (AI) to model clinical variant reporting for next generation sequencing (NGS) oncology assays. BioData Min. 2025 Oct 29;18(1):74.
2. Gómez-Luque MÁ et al. Towards Precision Oncology: A New Predictive Machine Learning Model for Early Progression to Castration Resistant Prostate Cancer. Prostate. 2026 Feb;86(3):320-328.
3. Liu Y et al. Comparative performance of warfarin pharmacogenetic algorithms in Chinese patients. Thromb Res. 2012 Sep;130(3):435-40.
4. Mansoor M et al. Predictive Modeling of Novel Somatic Mutation Impacts on Cancer Prognosis: A Machine Learning Approach Using the COSMIC Database. medRxiv (2024).
5. Abd Al Rahman E et al. A multioutput classifier model for breast cancer treatment prediction. Intell Based Med. 2024;10:100158.
6. Zhang C et al. Comparative Study of AI Algorithms in Personalized Ovarian Stimulation Protocol Optimization: Predictive Performance Analysis Based on Patient Baseline Characteristics. In: Proceedings of the 4th International Conference on Artificial Intelligence and Intelligent Information Processing (AIIIP '25); 2026. p. 654-662.

Results

shinyoung.park@khu.ac.kr

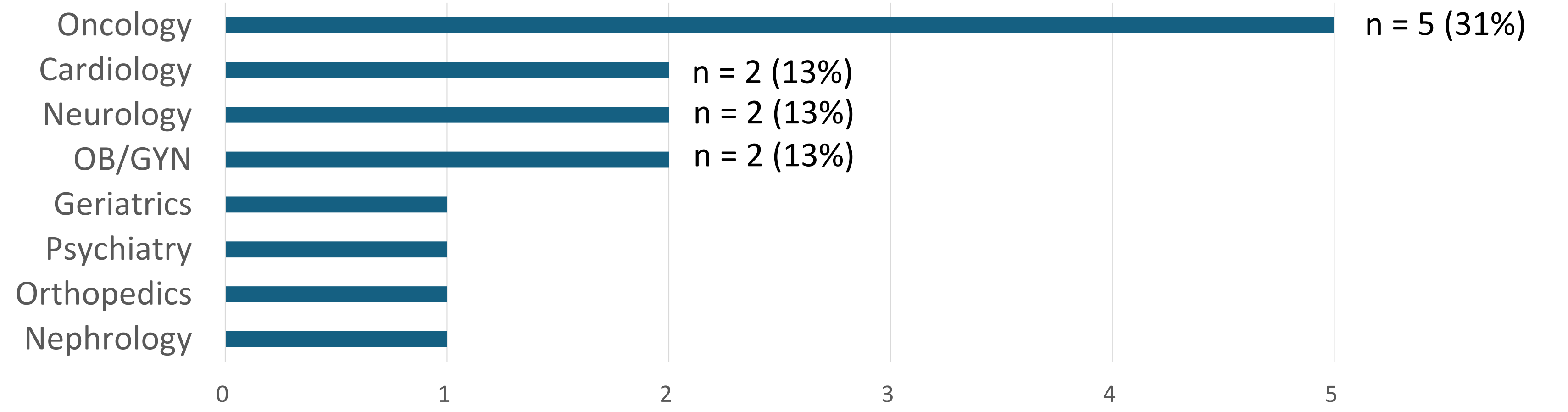
Study Selection

- Search results: A total of 147 articles were identified.
- Final inclusion: 16 articles met the eligibility criteria for final analysis.

Distribution of Clinical Domains

- Oncology (31%) was the most prevalent clinical domain, followed by Cardiology (13%), Neurology (13%), and Obstetrics and Gynecology (13%).
- Major applications included individualized treatment strategies.

Fig 1. Distribution of Clinical Domains (n=16)



Methodological and Performance Summary of Low Risk of Bias Studies (n = 6)

- **PGx inputs**
  - Genomic inputs: somatic mutations, targeted sequencing data, HPV genotyping, MSI-related markers, and mutation counts
  - Molecular and biomarker data: hormonal markers (AMH, FSH, LH, estradiol, inhibin B) and variant effect predictors
  - Clinical integration: clinicopathological characteristics, treatment history, and patient baseline features for personalized therapeutic decision-making
- **Explainability**: Feature importance was evaluated using SHAP values. Notably, an integrated SHAP-hypetune pipeline was employed in cutting-edge analysis to simultaneously optimize feature selection and hyperparameter tuning.
- **Validation strategies**: k-fold cross-validation, repeated cross-validation, and external validation
- **Performance**: Most studies evaluated model performance using discrimination and classification metrics, including the area under the curve (AUC), accuracy, sensitivity, specificity, precision, recall, and F1-score.

Table 1. Distribution of AI Algorithms Used in Included in Studies With Low Risk of Bias

Percentages were calculated based on total algorithm occurrences as multiple AI algorithms could be used within a single study.

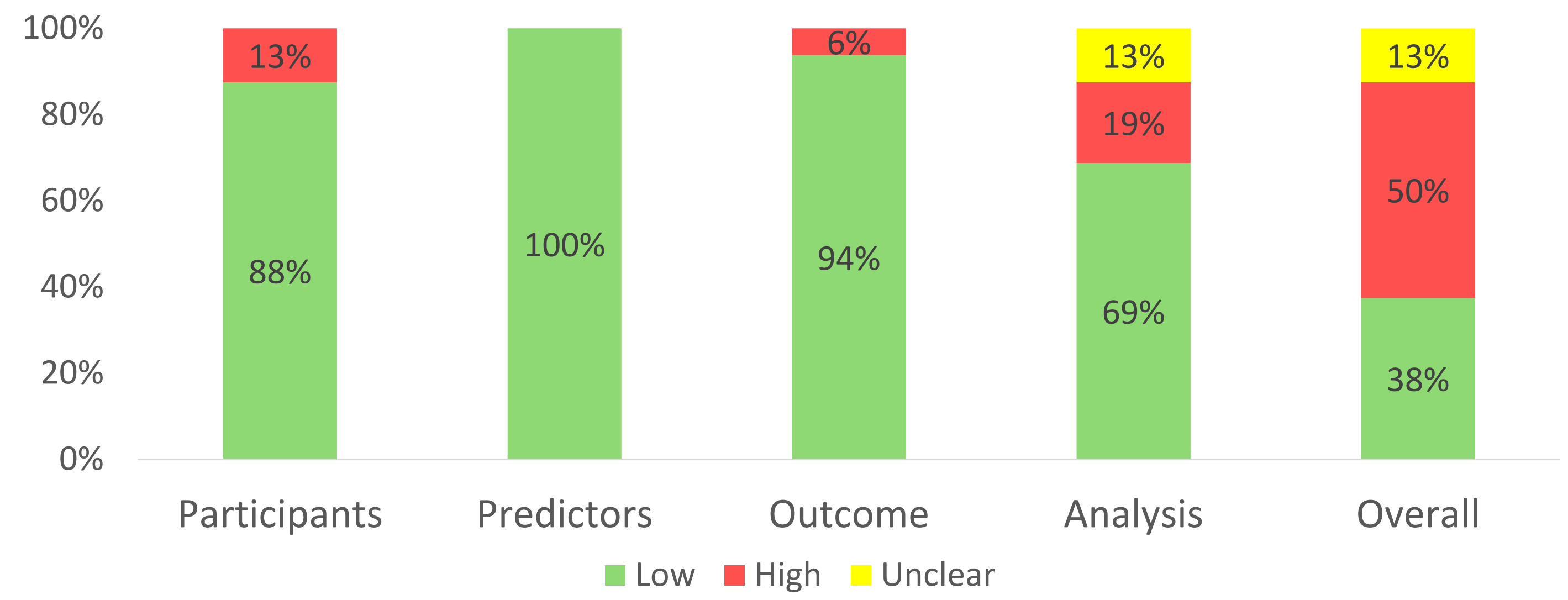
AI Algorithm category	Included Algorithms	Frequency, n (%)
Tree-based	Random Forest, XGBoost, CatBoost, AdaBoost, GentleBoost, LogitBoost, RusBoost, Tree20	13 (46.4%)
Neural Network (NN)	NN, DNN, NN Binary	4 (14.3%)
Support Vector Machine (SVM)	SVMLI, SVMGAM, SVMCU, SVM	4 (14.3%)
Regression	Logistic Regression, Linear Regression	2 (7.1%)
K-nearest neighbor	KNN10, SKNN	2 (7.1%)
Custom / Proprietary	RINH, RLI	2 (7.1%)
Discriminant Analysis Methods	SDiscriminant	1 (3.6%)

Abbreviations: DNN, deep neural network; KNN10, k-nearest-neighbors with k=10; NN, neural network; RINH, rivality index based algorithm; RLI, classification discriminant; SDiscriminant, subspace discriminant; SKNN, subspace k-nearest neighbors; SVMCU, supportvector machine with cubic kernel; SVMGAM, support vector machine with Gaussian kernel; SVMLI, support vector machine with linear kernel; Tree20, decision tree.

Risk of bias assessment according to PROBAST

- 6 studies were identified as having low risk of bias and adequate applicability.
- 8 studies were assessed as having high risk of bias, primarily due to:
  - Small sample sizes (< 100) along with model overfitting concerns
  - Inconsistent outcome definitions across study participants
  - Inappropriate integration of heterogenous data sources
- 2 studies were assessed as uncertain, primarily due to:
  - Insufficient reporting on missing data handling
  - Unclear validation procedures addressing model overfitting and optimism in performance measures

Fig 2. Risk of bias assessment using PROBAST (n=16)



Discussion

Summary

- Methodological Standards: Rigorous validation (e.g., k-fold cross-validation and external validation) is fundamental to achieving low-bias predictions.
- The Role of SHAP: SHAP has emerged as a leading approach for providing feature-level transparency in complex ML models.
- Future Direction: Integrated pipelines like Shap-hypetune represent the next step in streamlining model development while maintaining high interpretability.

Limitations

- A potential limitation is the heterogeneity of PGx data formats across the reviewed studies, which may affect the direct comparability of feature importance across different clinical domains.

Implications

- These findings emphasize that integrating explainable AI (XAI) with high-quality genomic inputs is essential for building clinical trust and enabling personalized therapeutic decisions.

Conclusions

- PGx-integrated AI-CDSS studies demonstrated methodological innovation and increasing adoption of explainability techniques; however, limitations including small sample sizes, overfitting risks, and insufficient validation highlight the need for larger, rigorously validated studies to support trustworthy clinical implementation in precision medicine.