Article

Not peer-reviewed version

Leveraging Machine Learning to Enhance Preventive Care Continuity: A Predictive Analytics Case Study from Qassim in Alignment with Saudi Vision 2030

Mohammed Awad Alanazi *

Posted Date: 6 March 2026

doi: 10.20944/preprints202602.1900.v1

Keywords: preventive screening; no-show; follow-up; machine learning; predictive analytics; Saudi Arabia; Vision 2030; Qassim

doi:10.20944/preprints202602.1900.v1

Article

Leveraging Machine Learning to Enhance Preventive Care Continuity: A Predictive Analytics Case Study from Qassim in Alignment with Saudi Vision 2030

Mohammed Awad Alanazi

Saudi Model of Care – Qassim Health Cluster, Buraidah, Saudi Arabia; malsumairy@gmail.com

Abstract

Background: Preventive health screening is a cornerstone of population health, but many patients fail to return for follow-up care, undermining early disease detection. This issue is highly pertinent in Saudi Arabia's Qassim region, aligning with the national Vision 2030 healthcare transformation. Machine learning (ML) offers a promising predictive approach to identify patients at risk of "non-return" to enable targeted interventions. **Objective:** To develop and evaluate an ML-based model for predicting patient non-return after preventive screening in the Qassim region, identify associated risk factors, and align the findings with the Saudi Model of Care reforms. **Methods:** A retrospective observational analysis of electronic health records from Qassim's primary care screening program (2019–2024) was conducted. The primary outcome was "non-return" within 6 months of an indicated follow-up. Multiple ML algorithms were evaluated using 10-fold cross-validation. **Results:** Among 18,752 screened patients, 5,230 (27.9%) did not return for follow-up. Ensemble tree-based methods performed best. The random forest classifier achieved the highest predictive performance (AUROC 0.812, accuracy 78.5%). Key predictors of non-return included extended lead time until the scheduled follow-up, prior appointment no-shows, and a lack of critical clinical findings during the initial screening. **Conclusion:** The developed ML model successfully predicts patient loss to follow-up with high accuracy. Integrating such predictive analytics into routine primary care enables early, personalized interventions, directly supporting Saudi Arabia's healthcare efficiency and preventive care goals.

Keywords: preventive screening; no-show; follow-up; machine learning; predictive analytics; Saudi Arabia; Vision 2030; Qassim

---

Introduction

Patient failure to attend recommended follow-up appointments after initial preventive health screenings is a persistent challenge in healthcare delivery. Such non-adherence limits the effectiveness of early disease detection, leads to delayed interventions, and negatively impacts resource allocation. Addressing appointment non-adherence is therefore a primary focus for improving overall healthcare quality [1].

Within Saudi Arabia, the Vision 2030 Health Sector Transformation Program emphasizes preventive care as a foundational element for improving population health [2]. The recently

 doi:10.20944/preprints202602.1900.v1

implemented Saudi Model of Care explicitly integrates proactive screening initiatives, such as the "Keep Well" program [3]. However, the clinical utility of these initiatives relies heavily on care continuity. In the Qassim region, which serves approximately 1.3 million residents through a network of Primary Healthcare Centers (PHCs), internal audits indicate a substantial rate of patient dropout following initial screenings. Traditional methods to mitigate this attrition, such as standard SMS reminders, often fall short, highlighting a need for targeted, data-driven operational strategies.

Machine learning (ML) models are increasingly utilized to predict outpatient appointment adherence. Recent literature indicates that ensemble algorithms, such as Random Forest and Gradient Boosting, can effectively forecast no-shows using routine operational data. These models typically achieve Area Under the Receiver Operating Characteristic Curve (AUROC) values between 0.75 and 0.85 [1, 4]. Studies utilizing large-scale primary care and hospital outpatient datasets frequently identify scheduling lead time, prior attendance history, and specific demographic variables as primary predictors of non-adherence [5, 6].

Beyond general outpatient settings, predictive analytics have been successfully applied to anticipate loss to follow-up in longitudinal chronic disease care, such as HIV and tuberculosis management [7, 8], as well as in specific preventive procedures like mammography screening [9]. Despite these methodological advancements, the majority of existing ML models for appointment adherence are trained on data from Western or East Asian healthcare systems. Localized models are necessary to account for regional and cultural variations, as demonstrated by a recent study highlighting the significant impact of local weather and scheduling infrastructure on primary care attendance in southern Saudi Arabia [4]. Furthermore, predictive modeling specifically targeting post-screening follow-up behaviors within the Saudi healthcare context remains largely unexamined.

To address this gap, this quality improvement (QI) study aims to develop and validate an ML model to predict patient non-return following preventive screenings in Qassim's PHCs. Specifically, this study outlines four primary objectives:

1. To develop an ML model capable of estimating the probability of a patient failing to attend a recommended follow-up appointment within the designated timeframe.
2. To identify key demographic, clinical, and logistical factors associated with follow-up non-adherence in the local screened population.
3. To evaluate model performance using standard metrics (e.g., AUROC, precision, recall) and simulate its practical application for risk-stratified patient outreach.
4. To discuss the alignment of these predictive analytics with the Saudi Vision 2030 Model of Care, and to detail the ethical considerations of utilizing routinely collected operational health data for QI initiatives without formal human-subjects research oversight [10].

Methods

Study Design and Setting

A retrospective, observational case study was conducted using electronic health records (EHRs) from the preventive health services of the Qassim Health Cluster. Healthcare delivery in the Qassim region is overseen by the Ministry of Health (MoH) and includes approximately 90 Primary

Healthcare Centers (PHCs). The study focused on adult patients attending routine preventive screenings for chronic conditions and age-appropriate cancer surveillance.

This project was executed as a formal Quality Improvement (QI) initiative. The primary objective was to analyze existing operational data to identify patterns in care continuity without introducing new clinical interventions. The Qassim regional health administration granted approval to access de-identified operational data for service enhancement purposes. Consequently, formal Institutional Review Board (IRB) approval was not required, consistent with established ethical guidelines distinguishing QI initiatives from human-subjects research [10].

Data Sources and Study Population

Data were extracted from the Qassim regional EHR and the centralized MoH primary care appointment system (Mawid) for the period between January 2019 and December 2024. The study cohort included 18,752 unique patients who completed an initial preventive screening and were explicitly advised by a provider to return for a follow-up appointment. Patients advised that "no follow-up is needed" were excluded from the analysis.

The primary outcome variable was binary: “non-return” (coded as 1) versus “Return” (coded as 0). A patient was classified as a "Return" if a completed follow-up visit was documented within 1.5 times the provider-recommended interval (allowing a clinical grace period). Patients lacking any recorded follow-up visit within this timeframe were classified as "non-return."

The cohort’s mean age was 47.3 years (SD ± 11.4), comprising 54% females and 46% males. Demographically, 88% were Saudi nationals, reflecting the local population structure. Baseline clinical metrics revealed a high burden of non-communicable disease risk factors: 35% presented with obesity (BMI ≥ 30), 30% with hypertension, 23% with diabetes or pre-diabetes, and 12% exhibited abnormal cancer screening findings. Of the 18,752 patients, 5,230 (27.9%) failed to attend their recommended follow-up. This non-return rate of approximately 28% is consistent with loss-to-follow-up metrics reported in similar outpatient settings globally [6].

Feature Engineering and Predictor Variables

A comprehensive set of 58 predictive features was engineered based on clinical relevance and literature precedents. The variables were categorized as follows:

- **Demographic Factors:** Age, sex, nationality, marital status, and a neighborhood-level socioeconomic deprivation index based on national census data.
- **Medical History:** Presence of chronic comorbidities (e.g., diabetes, cardiovascular disease), smoking status, polypharmacy indicators, and historical appointment adherence (i.e., prior no-shows within the past two years), which is a well-documented predictor of future attendance [5].
- **Screening Findings:** The severity of the initial screening results (e.g., normal, minor abnormalities, critical findings requiring urgent intervention).
- **Appointment Logistics:** Scheduling location (PHC versus hospital referral), lead time (days between scheduling and the target appointment date), and whether the follow-up appointment was booked immediately at the time of screening.
- **Contextual and Temporal Factors:** Seasonality (e.g., summer months, Ramadan) and regional meteorological data (average temperature and sandstorm occurrence), reflecting

evidence that extreme weather influences adherence in the Saudi context [4]. Distance to the clinic (in kilometers) was also calculated to estimate travel burden.

Missing demographic data (accounting for <5% of the dataset) were imputed using median values. Continuous variables exhibiting non-linear relationships were binned into categorical tiers.

Machine Learning Development and Evaluation

The prediction of non-return behavior was formulated as a binary classification task. Analysis was conducted using Python (scikit-learn and XGBoost libraries). To establish baseline performance, standard Logistic Regression (LR) with L2 regularization was utilized. More complex, non-linear algorithms evaluated included Decision Trees, Random Forests (RF), Extreme Gradient Boosting (XGBoost), a Multi-Layer Perceptron (MLP) Neural Network, and Gaussian Naïve Bayes.

The dataset was partitioned into a 70% training set and a 30% independent testing set using stratified sampling to preserve the class distribution (27.9% minority class). Model hyperparameters (e.g., tree depth, learning rate, number of estimators) were optimized using 5-fold cross-validation grid search on the training data. The moderate class imbalance was addressed by applying class-weight adjustments rather than synthetic resampling, optimizing the loss function for the minority class. Feature selection was refined using Recursive Feature Elimination (RFECV), identifying approximately 15 to 20 top features that captured the majority of the predictive signal.

Model performance was rigorously assessed on the hold-out test set using the Area Under the Receiver Operating Characteristic Curve (AUROC) as the primary discrimination metric, targeting benchmark values ≥ 0.75 [1]. Secondary evaluation metrics included overall accuracy, precision (Positive Predictive Value), recall (Sensitivity), and the F1-score. To ensure the reliability of the probabilistic outputs for future clinical deployment, probability calibration was performed using isotonic regression. Finally, a temporal validation check was conducted by training the optimal model on data from 2019–2022 and testing it on 2023–2024 records to confirm longitudinal stability.

Results

Descriptive Findings

Prior to predictive modeling, an exploratory analysis of the cohort ($N = 18{,}752$) was conducted to compare the characteristics of patients who returned for follow-up (72.1%) versus those who did not (27.9%).

Demographically, non-returners were significantly younger than returners (mean age 45.0 vs. 48.3 years, $p < 0.001$). The highest non-return rate was observed in the 20–39 age bracket (32%), while patients aged 60 and above demonstrated the highest adherence (80% return rate). Gender differences were marginal and not statistically significant (28.5% non-return for males vs. 27.4% for females, $p = 0.08$).

Socioeconomic and logistic barriers strongly influenced adherence. Patients residing in the lowest-income quartile neighborhoods exhibited a 33% non-return rate compared to 22% in the highest quartile. Furthermore, patients with an estimated clinic distance exceeding 30 km without documented personal transportation defaulted at a higher rate (35%).

Clinically, a paradoxical trend emerged regarding screening severity. Patients with critical abnormal findings had an 85% return rate, whereas 31% of patients with completely normal screenings failed to return for routine follow-up. Historical attendance was also a strong indicator; patients with at least one prior missed appointment in the preceding year had a 45% non-return rate. Logistically, referrals requiring cross-facility visits (e.g., PHC to a hospital specialist) increased the default rate to 38%. Additionally, leaving the clinic without a booked follow-up appointment resulted in a 41% non-return rate.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202602.1900.v1

Temporally, non-return rates peaked during the extreme summer months (June–August; ~32%) and dropped to their lowest in December (~24%). Elevated non-return rates strongly correlated with months characterized by temperatures exceeding 42°C or frequent sandstorms.

Model Performance

Feature importance analysis based on Mean Decrease in Impurity (Gini importance) identified the primary drivers of the RF model's predictions:

1. **Follow-up Lead Time:** The number of days until the target appointment was the dominant predictor. Partial dependence plots indicated a sharp exponential increase in non-return risk when the lead time exceeded 30 days.
2. **Prior No-Show History:** A documented history of missed appointments nearly doubled the likelihood of default.
3. **Screening Result Severity:** "Normal" findings significantly increased the predicted non-return probability.
4. **Age:** Younger adults independently elevated the risk score.
5. **Booking Status:** The absence of a scheduled appointment generated at checkout was a strong indicator of subsequent non-return.
6. **Distance to Facility and Referral Type:** Cross-facility referrals coupled with distances exceeding 20 km compounded the non-return probability.
7. **Socioeconomic Index:** Lower neighborhood-level socioeconomic status incrementally raised the risk probability.

Meteorological variables (e.g., month of appointment, summer temperature indicators) contributed a modest but measurable improvement to the model's overall discriminative ability.

Temporal Generalization and Predictive Risk Stratification

To confirm longitudinal stability, a temporal validation test was conducted by training the RF model on 2019–2022 records and evaluating it in 2023–2024 data. The model maintained an AUROC of 0.805, demonstrating robustness against systemic temporal shifts.

To evaluate practical utility for resource allocation, the testing cohort was stratified into risk quintiles based on the RF model's calculated probabilities. The highest-risk quintile (top 20%) captured patients with an observed non-return rate of 67%, compared to a mere 5% in the lowest-risk quintile. For resource-constrained settings requiring high precision, adjusting the threshold to flag only the top 15% of patients (predicted risk > 0.60) yielded a Positive Predictive Value of ~0.80.

Discussion

Principal Findings and Alignment with Literature

This study successfully developed and validated a machine learning model to predict patients non-return following preventive health screenings in the Qassim region. The Random Forest classifier demonstrated strong discriminative ability (AUROC = 0.812), outperforming traditional regression methods. This performance aligns with recent comprehensive reviews of outpatient no-

 doi:10.20944/preprints202602.1900.v1

show prediction models, which typically report AUROCs between 0.75 and 0.85 [1], as well as specific hospital-based studies in similar demographic settings [6].

Our findings confirm that follow-up adherence is driven by a complex interplay of logistical, clinical, and contextual factors. The dominance of "scheduling lead time" as the primary predictor mirrors findings by Tuan et al. in U.S. primary care settings, were extended intervals significantly degraded attendance [5]. Similarly, the identification of geographic distance and lower socioeconomic status as substantial barriers to follow-up is consistent with recent predictive models evaluating mammography screening completion [9].

Interestingly, our model identified that patients with entirely "normal" initial screening results had a higher risk of non-return compared to those with critical abnormalities. This suggests that perceived clinical urgency strongly dictates patient behavior. While prior ML applications have largely focused on generic outpatient appointments, our model's ability to accurately predict preventive care attrition demonstrates predictive success comparable to models used for loss-to-follow-up in longitudinal chronic disease management, such as HIV and tuberculosis care [7, 8]. Furthermore, the measurable impact of extreme summer temperatures on non-return rates corroborates recent findings in southern Saudi Arabia, underscoring the necessity of incorporating local meteorological data into regional ML frameworks [4].

Strategic Alignment with Saudi Vision 2030

The practical application of this predictive model directly supports the Saudi Vision 2030 Health Sector Transformation Program [2]. A core tenet of the newly implemented Saudi Model of Care is a shift toward proactive, patient-centered preventive services, exemplified by the "Keep Well" initiative [3].

Identifying high-risk patients before they leave the clinic enables healthcare administrators to transition from reactive outreach (e.g., automated SMS blasts) to proactive, targeted interventions. For instance, prioritizing the top 15% highest-risk patients—who our model identified with an 80% positive predictive value—allows for the efficient allocation of intensive care coordination resources, such as personal patient navigators or immediate appointment scheduling. This data-driven resource allocation exemplifies the value-based healthcare delivery model envisioned by national reforms.

Ethical Considerations in Predictive Quality Improvement

The deployment of ML models in healthcare necessitates rigorous ethical oversight. This study was conducted as a formalized Quality Improvement (QI) initiative utilizing retrospective, de-identified operational data. Consequently, formal Institutional Review Board (IRB) approval was not required. This approach is consistent with established ethical frameworks which differentiate QI activities aimed at immediate service enhancement from human-subjects research designed to generate generalizable knowledge [10].

Ethically, the predictive risk scores generated by this model must be utilized to mitigate health disparities rather than exacerbate them. Identifying patients from lower socioeconomic backgrounds or remote areas as "high-risk" should exclusively trigger supportive interventions (e.g., transportation assistance, telemedicine options) to ensure equitable access to care continuity.

Limitations and Future Directions

Despite robust predictive performance, this study presents certain limitations. The analysis was restricted to a single region's public healthcare network, potentially limiting generalizability to highly transient or private-sector populations. Additionally, the dataset lacked granular psychosocial variables (e.g., patient health literacy, family support), which inherently influence adherence behavior. Future prospective studies are necessary to evaluate the real-world clinical impact of implementing these ML predictions into the electronic health record (EHR) workflow and to measure the actual improvement in follow-up rates following targeted interventions.

 doi:10.20944/preprints202602.1900.v1

Conclusion

This case study demonstrates the feasibility and clinical utility of employing machine learning to predict loss to follow-up in preventive screening programs within Saudi Arabia. The developed Random Forest model achieved high predictive accuracy by leveraging routine clinical, logistical, and sociodemographic data. By integrating such localized predictive analytics into primary care workflows, health systems can proactively identify at-risk patients and deploy tailored interventions. This proactive, data-driven methodology not only enhances patient care continuity and early disease management but also directly advances the overarching efficiency and population health objectives of the Saudi Vision 2030 Model of Care.

References

1. Toffaha K., Simsekler M.C.E., Omar M.A., ElKebbi I. (2025). Predicting patient no-shows using machine learning: A comprehensive review and future research agenda. Intelligence-Based Medicine, 11: 100229. DOI: 10.1016/j.ibmed.2025.100229
2. Vision 2030 Health Sector Transformation Program. (2022). Health Sector Transformation: A Journey Towards a Healthy and Vibrant Community. Vision2030.gov.sa.
3. Health Holding Company. (2025). The Saudi Model of Care – Early Screening for Chronic Diseases Initiative. HHC Saudi – Model of Care documentation.
4. Alshehri A., Saeed A., AlShafea A., et al. (2024). Machine Learning Approaches to Predict No-Shows in Saudi Arabian Primary and General Healthcare Settings. Saudi J Health Syst Res, 5(1): 24–40. DOI: 10.1159/000542701
5. Tuan W.J., Yan Y., Abou Al Ardat B., Felix T., Chen Q. (2025). Predicting Missed Appointments in Primary Care: A Personalized Machine Learning Approach. Annals of Family Medicine, 23(4): 294–301. DOI: 10.1370/afm.240316
6. Hamdan A.F.A., Abu Bakar A., et al. (2023). Machine Learning Predictions on Outpatient No-Show Appointments in a Malaysia Major Tertiary Hospital. Malays J Med Sci, 30(5): 169–180. DOI: 10.21315/mjms2023.30.5.14
7. Taye G., Deressa W. (2025). Development of a machine learning prediction model for loss to follow-up in HIV care using routine electronic medical records in a low-resource setting. BMC Med Inform Decis Mak, 25(1): 30. DOI: 10.1186/s12911-025-03030-7
8. Chen J., Jiang Y., Li Z., et al. (2024). Predictive machine learning models for anticipating loss to follow-up in tuberculosis patients throughout anti-TB treatment journey. Scientific Reports, 14: 24685. DOI: 10.1038/s41598-024-74942-z
9. Davis M., Simpson K., Lenert L., Diaz V., Alekseyenko A. (2023). Predicting Mammogram Screening Follow Through with Electronic Health Record and Geographically Linked Data. Cancer Research Communications, 3(10): 2126–2132. DOI: 10.1158/2767-9764.CRC-23-0263
10. Hall S., Lee V., Haase K. (2020). Exploring the challenges of ethical conduct in quality improvement projects. Can Oncol Nurs J, 30(1): 36–45. PMID: 33118977