Mammogram Screening Compliance: AI-Driven Predictive Modeling and Data-Driven Insights

King Fahd University of Petroleum and Minerals

Capstone Project MIS 619

Abstract

Mammography is the primary screening method for breast cancer and has significantly reduced breast cancer mortality rates by approximately 40%. Despite the proven benefits of regular mammogram screenings, compliance with recommended screening guidelines remains suboptimal in the Kingdom. This research aims to identify key factors influencing mammogram screening compliance among women who have received screening orders through the MyChart application at Johns Hopkins Aramco Healthcare. By leveraging Artificial Intelligence (AI) and predictive modeling, this study seeks to improve adherence rates, optimize resource allocation, and reduce the burden of late-stage breast cancer diagnoses. The findings will provide actionable insights for healthcare providers to design targeted interventions, ultimately enhancing public health outcomes.

Keywords: mammogram screening, compliance, artificial intelligence, predictive modeling, breast cancer, healthcare.

Table of Contents

Chapter 1 - Introduction

1.1. Overview:

Breast cancer is one of the most common cancers affecting women worldwide, and early detection through mammography has been shown to significantly reduce mortality rates. Annual mammogram screenings are recommended starting at age 40, as early detection greatly improves survival rates (Arleo et al., 2017). However, despite the proven benefits, compliance with mammogram screening guidelines remains low, particularly in certain regions, including the Kingdom of Saudi Arabia. Many women fail to undergo regular screenings due to a variety of demographic, cultural, and behavioral factors. This research addresses the following question: What key factors influence mammogram screening compliance among women who have received screening orders through the MyChart application at Johns Hopkins Aramco Healthcare, and how can Artificial Intelligence (AI) be leveraged to predict and improve adherence rates? By understanding the underlying reasons for non-compliance and utilizing AI-driven predictive modeling, this study aims to improve screening rates, optimize resource allocation, and reduce the burden of late-stage breast cancer diagnose.

1.2. Research Objectives and Approach

The primary objective of this study is to identify the demographic and behavioral factors that influence mammogram screening compliance and to develop AI-driven predictive models to forecast an individual's likelihood of screening adherence. By analyzing screening behavior patterns, this research aims to generate actionable insights that enable healthcare providers to design more effective and targeted interventions to increase compliance rates. Key stakeholders in this study include healthcare providers, who can leverage the findings to enhance patient outreach strategies and improve screening participation. The dataset for this study, obtained from Johns Hopkins Aramco Healthcare, contains demographic, behavioral, and medical history variables related to breast cancer screening. All data will be handled with strict confidentiality, ensuring patient privacy and compliance with research ethics.

To conduct this study, relevant data sources will be carefully considered, and ethical guidelines and data privacy regulations will be strictly adhered to. The dataset will be analyzed using AI-driven predictive modeling techniques to identify patterns and factors influencing screening compliance. The predictive models will be developed using machine learning algorithms, which will be trained on the dataset to forecast an individual's likelihood of adhering to mammogram screening recommendations.

1.3. Expected Outcomes

The expected outcomes of this research include the identification of key factors influencing mammogram screening compliance and the development of accurate predictive models to forecast adherence rates. These insights will enable healthcare providers to design targeted interventions, such as personalized reminders, educational campaigns, and outreach programs, to improve screening participation. By enhancing compliance, this research aims to shift the focus from late-stage treatment to early detection and prevention, ultimately reducing long-term healthcare costs and improving patient outcomes.

Chapter 2: Literature Review

2.1. Importance of Mammogram Screening

Breast cancer at the most commonly diagnosed cancer among women, other than nonmelanoma of the skin. In the year 2022, 2.3 million women were diagnosed to be suffering from breast cancer, among which 670,000 women faced mortality. According to the estimates made by the American Cancer Society for Breast Cancer (ACS, in the year 2025, it was recognized that around 316,950 new cases of invasive breast cancer were diagnosed, around 59080 new cases of ductal carcinoma in situ were diagnosed, and 42170 were estimated to die from breast cancer in the United States alone (American Cancer Society, 2025).

While breast cancer risk is prevalent among women of all ages and ethnicities, the risk of contracting breast cancer increases in later life. It is considered to be the second leading cause of cancer-related death among women after lung cancer and the leading cause of cancer-related death among Hispanic and Black women (Giaquinto et al., 2022). Moreover, according to global estimates, there are significant equities in breast cancer burden, with women in countries having a very high human development index (HDI) being more prone to suffer from breast cancer (World Health Organization, 2024). In comparison to countries with low HDI, where 1 out of 27 women are diagnosed with breast cancer in their lifetime and 1 out of 48 women die from breast cancer, in high HDI countries, 1 out of 12 women are diagnosed with breast cancer in their lifetime and 1 out of 71 women die from it (World Health Organization, 2024).

Cancer treatment, particularly at the later stages, is significantly expensive, with an estimated annual cost of 88 billion dollars and an average cost of 1.5 million dollars for each affected woman (Khurshid et al., 2023). This imposes a significant financial burden on the patient and the family members of the patient. The high mortality rate of cancer is a result of the late diagnosis of the disease because survival is inversely correlated with the stage of diagnosis of cancer. There is a significant lack of awareness about breast cancer prevention and treatment (Khurshid et al., 2023). Women frequently visit hospitals when they are at the end of their lives and the majority of the time during high mortality risk.

Ironically, it has been estimated that around 30% of breast cancer cases can be attributed to modifiable risk factors like physical inactivity, excess body weight and alcohol intake, which can be prevented through the implementation of viable lifestyle and healthcare strategies (Giaquinto et al., 2022). Secondary prevention through mammography screening for early detection and advanced cancer treatment has been considered to be a fundamental methodology for reducing breast cancer-related mortality. Giaquinto et al. (2022) indicated that the two primary actions for improving prognosis outcomes for breast cancer patients are ensuring the availability of early screening and diagnostic services and acting quickly to address the symptoms.

Mammography screening can help improve survival prospects, increase the potential of the disease being successfully identified, cured, and treated at the early stages without incurring significant financial costs, and reduce the requirement for invasive therapy (Khurshid et al., 2023). In other words, mammography screening is a critical preventive measure for the early detection of breast cancer, making it important for empowering women to undergo scheduled mammography screening, considering that it reduces breast cancer-related mortality by 40% (Løberg et al., 2015).

The current guidelines for mammogram screening compliance are significantly variable based on national infrastructure. The ACS recommends annual screening, starting at the age of 45, while the United States Preventive Services Task Force (USPSTF) suggests screening every 6 months from the age of 50 (Hardesty et al., 2016). In Saudi Arabia, the Ministry of Health suggests women aged 40 and above go for mammogram screening every 6 months (Zapka et al., 1991). Irrespective of the variability of the recommended frequency of mammogram screening, the majority of countries have started to implement policies to support and encourage women to undergo mammogram screening for early detection of breast cancer.

2.2. Factors Influencing Mammogram Screening Compliance

Regular screening is defined as the adherence to having had more than 1 mammogram after becoming eligible for screening or having had a mammogram within a period of the last 2 years (Rahman et al., 2003). While the majority of nations encourage women to maintain mammogram screening compliance, different factors were recorded to have an impact on mammogram screening. From the year 1987 to 2000, adherence to the mammography screening guidelines increased steadily in women between the age group of 40 and older, reaching 70.1% in 2000 (Vyas et al., 2012). However, in the year 2005, it dropped significantly to 68.3%, suggesting a negative trend in the participation of women in mammography screening and adherence to governmental guidelines (Vyas et al., 2012). Predisposing factors like race, age, education and enabling factors like insurance coverage and community economic status influenced alignment with mammography (Rahman et al., 2003).

Bivariate analysis highlights older age as being related to having a significant impact on mammography screening participation (Vyas et al., 2012). Older women, particularly individuals above the age of 50, have a higher likelihood of complying with screening guidelines because of increasing medical advice and awareness of risks (Vyas et al., 2012). On top of this, education level also has an extensive impact on compliance, with women having higher education associated with greater health literacy and awareness of the benefits of mammography screening, inherently leading to greater levels of screening.

The statistical assessment indicates that while wealthier nations have a higher breast cancer incidence rate, less developed countries suffer from higher relative mortality rates from breast cancer, which can be attributed to low levels of breast cancer screening at the early stages, indicating the impact of socio-economic status on breast cancer mortality. In high-income nations, including Australia, the United Kingdom, and Eastern Europe, more than 60% of women are diagnosed with breast cancer and stages 1 and 2 of the disease, which helps to significantly improve their survival rates (Tavakoli et al., 2024). On the other hand, women belonging to low-income countries have lower participation in screening tests and are diagnosed at a significantly advanced stage, in stage 3 or stage 4 breast cancer, when cancer has already metastasized to other vital organs (Tavakoli et al., 2024).

According to previous research, ethnicity is also a major factor impacting mammography screening. African American and Black women have an unequal burden of breast cancer mortality. It has been estimated that the breast cancer incidence rate in African American women is 126.5 cases per 100,000 women in comparison to 130.1 cases in white women (Agrawal et al., 2021). Even though the incidence rate is lower among African American women, African American women face 40% higher mortality risks in comparison to white women. They also have twice the likelihood of being diagnosed with triple-negative breast cancer and receiving a diagnosis of breast cancer at a more advanced stage in comparison to other racial and ethnic backgrounds (Agrawal et al., 2021).

Hispanic women have also been estimated to be impacted significantly more by breast cancer in comparison to other ethnicities because of the more aggressive cancer propagation, accounting for lower detection in the early tumor stage and genetic factors (Agrawal et al., 2021). While mammography screening has increased among African American women and Hispanic women, it continues to remain under the 81% Healthy People 2020 objective, highlighting how ethnicity and low mammography screening rates negatively impact the health of women (Agrawal et al., 2021).

The differences in the cancer incidence rate across a population of high-income and low-income countries can be attributed to the difference in risk factor prevalence and the implementation and uptake of screening programs due to economic status (Tavakoli et al., 2024). While routine screening is significantly important for the detection of early, more treatable stages of breast cancer, mammography screening is significantly low for women belonging to the lower socioeconomic strata, particularly in developing countries (Tavakoli et al., 2024).

Despite numerous educational efforts and interventions developed to promote participation in mammography screening programs, there is a persistent increase in mortality rates and low participation among women, specifically from the lower socioeconomic strata (Tavakoli et al., 2024). Higher screening service accessibility, together with lower prices, strengthens the response to screening initiatives. Individuals without health insurance, along with expensive screening costs and restricted availability of screening facilities in rural areas, diminish the participation of lower socioeconomic status in mammographic testing.

The screening mammography guidelines are affected by three primary elements, which include family history of breast cancer, hormone replacement therapy requirements and active breast cancer status. The screening participation of women increases when their family shows a higher rate of breast cancer development because they recognize the advantages of mammography testing and the presence of genetic cancer risks (Sterlingova et al., 202).

Family influence, together with stigma perception and cultural beliefs, significantly impact the decision of women to obtain mammograms for screening. The breast cancer screening guidelines show lower acceptance rates among ethnic, racial, and cultural minority groups, which delays medical detection, raises mortality statistics, and worsens cancer outcomes (Alcazar-Bejerano, 2014). The decision to follow guidelines depends on screening obstacles faced by minorities, together with cultural influences which affect behavioral responses toward healthcare interventions. The avoidance of breast health discussions in certain traditional societies stops women from seeking preventive healthcare services.

Religious belief also impacts attitudes towards mammography screening, with women believing it to be impure to get themselves evaluated before other men as preventive care, leading to a significant reduction in participation rates. On top of this, family support also has an important role because encouragement from relatives, partners and other family members increases the participation rate in screening programs. While inefficient healthcare systems, long wait times and inadequate physician recommendations cause lower adherence, according to Vyas et al. (2012), interventions promoting mammography screening that are based on physicians' letters or booklets along with use of printed educational material, telephone counselling, educational programs, and onsite mammography screening increases adherence to mammography screening guidelines.

2.3. Role of Artificial Intelligence in Predictive Modeling for Healthcare

Since the 1980s, mammography screening has been the cornerstone for early detection and treatment of breast cancer (Dembrower et al., 2023). However, because of the asymptomatic tendencies of breast cancer during the early stages, timely detection is often difficult. If breast cancer is found late, it can cause serious illness and even fatal outcomes. It has been found by many scientists that catching breast cancer early leads to better survival, less expensive treatment and more effective cures (Kim et al., 2020; AlSamhori et al., 2024). When breast cancer is caught in the early stages, 80% to 90% of survivors make it for at least five years, though only 28% do so when the cancer has advanced (AlSamhori et al., 2024). As a result, it makes clear that finding problems as soon as possible and identifying them efficiently is crucial.

It has been shown that mammograms supply X-ray images detailed enough so that radiologists can tell if breast tumors are malignant or benign (Dembrower et al., 2023). Even so, there are some challenges with mammography. The greatest issue is that some people do not follow the screening rules, resulting in fewer early tumor diagnoses and higher death rates (Schaffter et al., 2020). Because there is often inconsistency in how accurately radiologists diagnose, both wrong cancer findings and extra screenings can occur.

Mammogram screenings have trouble because of limited image definition, insufficient image reconstruction and excessive false positives, making integrating accurate diagnosis challenging (Ghantasala et al., 2024). Moreover, there is a risk of incorrect or uncertain diagnoses in mammogram screening because radiologists face difficult and complex imaging patterns, noise in the photos and the large number of images that need to be reviewed. Because mammogram screening generates lots of images, it is getting more important for breast radiologists to automatically improve their workflow (Ghantasala et al., 2024). There is also a global shortage of breast radiologists, which has been aggravated by the increasing demand for precision diagnostics from both patients and providers (Dembrower et al., 2023).

With the increasing incidence and mortality associated with economic complications and several critical healthcare challenges, it is becoming important to integrate artificial intelligence advancements in mammogram screening to increase diagnostic accuracy. Artificial

intelligence, the use of which has made it possible to mitigate the challenges related to mammography screening, is predicted by many studies. The various retrospective studies distinguish that artificial intelligence has an overall higher reliability of diagnostic accuracy through the reader as compared to other methods for the screening of mammograms. Similarly, AI researchers can use the findings of machine learning on imaging and clinically acquired data to draw out treatment strategies that are adapted to the individual needs of cancer patients (Slouka et al., 2021)

Artificial intelligence is one of the most commendable technological discoveries of the twenty-first century, and it has gained medical diagnostics not only through automatic image analysis but also by offering personalized treatment recommendations and ensuring early disease detection (Dembrower et al., 2020). Convolutional neural networks (CNN) can be employed by radiologists for volumetric estimation and lesion segmentation in the digitally reconstructed radiographs. Moreover, deep learning-based CNNs were applied by the radiologists to localize, segment, and classify the breast cancer cells and tumors more effectively (Drira et al., 2015).

The combination of artificial intelligence and healthcare models to create predictions for patient results with data from the past will benefit women joining breast cancer screenings. Thanks to artificial intelligence in this system, it can analyze patient records and recognize genetic signaling, which can lead the doctor to alter their treatment of different diseases. Using deep learning, experts can determine whether a patient will develop breast cancer based on the analyzed data (Zhong et al., 2020).

Machine learning approaches are also often used to forecast how patients might act and what results their care might bring. Learning which algorithms are taught using labelled databases can be used to categorize patient groups and anticipate their outcomes (Yala et al., 2019). Medical staff may predict if a patient will attend their mammogram using decision trees and logistic regressions. To inform patients, doctors can use patient segmentation methods such as hierarchical clustering and k-means, which help them design targeted campaigns explaining the importance of mammogram screening (Yala et al., 2019). Artificial intelligence models can also be employed to improve healthcare strategies by continuously learning from the patient responses and adjusting the interventions following the responses (Balthazar et al., 2018). Other than this, natural language processing resources can be used for assessing structure text from social media, medical records, and patient feedback to analyze health concerns and predict screening behaviors.

2.4. Research Gaps:

Even though that has been significant advancement in the assessment of the participation of women in mammogram screening and artificial intelligence predictive modelling for healthcare, there are several literature gaps associated with improving and understanding mammography screening compliance. While previous studies have identified the causal link between different factors and lower participation in mammogram screening, there is limited research on integrating strategies that apply artificial intelligence predictive modelling based on behavioral and cultural factors to mitigate these challenges, specifically in countries like Saudi Arabia. The majority of the existing studies focus on clinical data and demographic information, neglecting social influences and psychological impact on compliance.

On top of this, artificial intelligence-based interventions like predictive analytics and personalized mammogram screening have been evaluated in Western healthcare settings but do not have any validation in diverse populations with different healthcare infrastructures. There are limited studies on the employment of artificial intelligence-powered mammogram screening efficiency in Saudi Arabia, thereby justifying the requirement for localized studies that considered insurance policies, healthcare accessibility and cultural barriers of the nation. On top of this, ethical concerns related to algorithmic bias, data privacy and artificial intelligence transparency are also limited, making it important to develop ethical frameworks that ensure inclusivity and fairness in artificial intelligence-powered healthcare interventions in mammogram screening.

Chapter 3 – Methodology:

3.1 Study Design

The study will employ the onion research methodology for delineating how the research will be conducted. The research philosophy is the basis of the research, which defines ontology, epistemology and axiology. The research will be grounded in positivist philosophy, which predominantly reflects the philosophical stance where data and information are considered to be factual, detached from any bias or other influence and objective. Assessment of the electronic health records will help in identifying the relationships and patterns in mammogram screening compliance and cancer outcomes by employing machine learning tools and statistical resources, ultimately generating generalizable information based on empirical data.

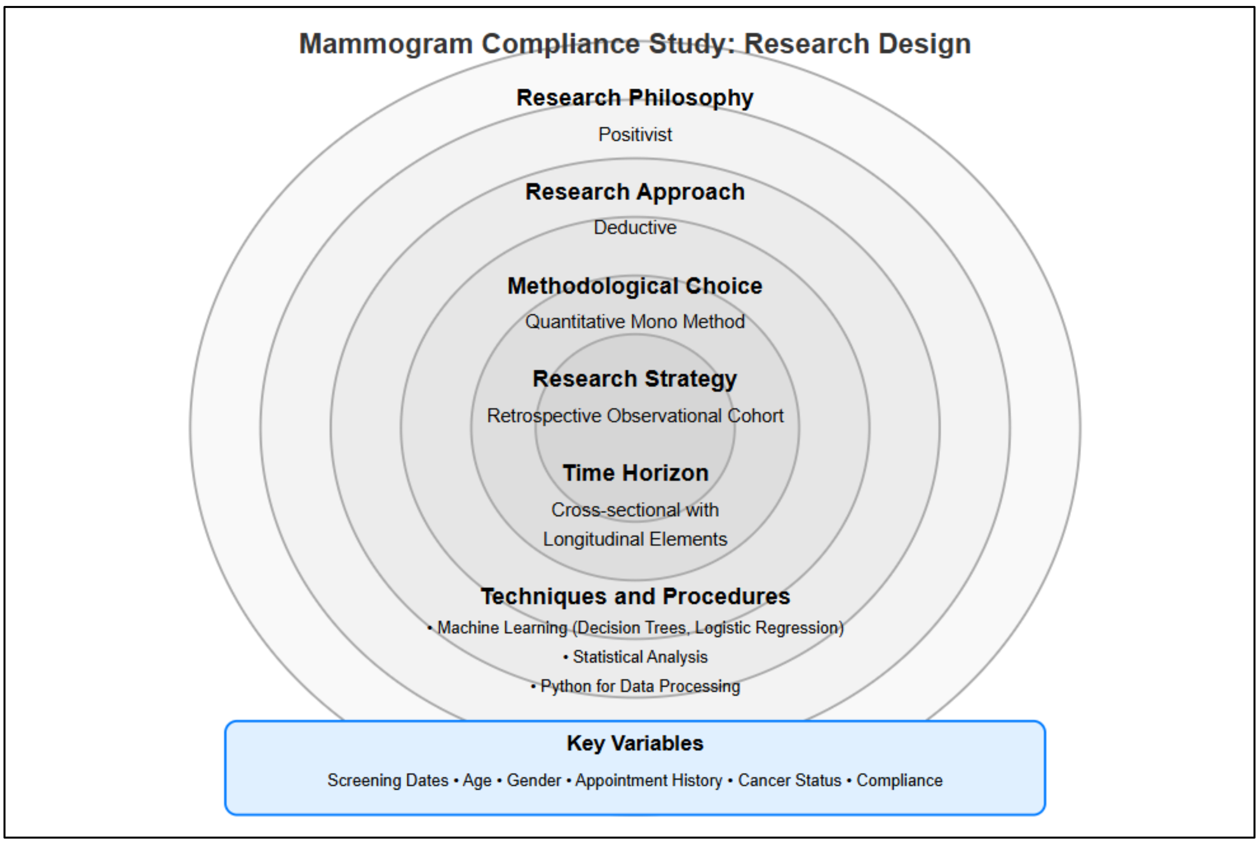


Figure 1: Research Design (Source: Author)

A deductive research approach will be used to compare the previously established theory regarding mammogram compliance and the risk of developing cancer by using historical data. Established literature will be used for gathering data regarding predictors of non-compliance, hypotheses that have been formed and tested by implementing data-driven models, thereby confirming or refuting the existing assumptions. The research will subsequently utilize a quantitative mono method, using numerical analysis of the structured data that is extracted from the healthcare records of patients. Variables like screening dates, age, gender, appointment history, cancer status and compliance will be assessed by employing predictive modelling techniques using Python.

A retrospective observational cohort strategy will be adopted for this research. A retrospective cohort study predominantly uses groups that are retrospectively identified and prospectively compared following a structured model - the subjects are subdivided into two groups, where one is exposed to a given factor and the other is not exposed to the same factor. For the study, the patients will be grouped based on past behavior, like compliance with mammogram screening and outcomes associated with cancer diagnosis. These groups of patients will be analyzed over time by implementing appointment data and historical clinical data. This strategy is viable for predicting patient behavior and identifying the risk factors of breast cancer based on naturally occurring data without any randomization and intervention, thereby preventing bias.

The data set employed for this research will span several years. However, the analysis will only include patient data at particular time points, such as the most recent appointment or last screening, making it cross-sectional research. At the same time, variables like progression to diagnosis and time since last screening will also incorporate longitudinal time horizon elements. The study will use a combination of bivariate tests, descriptive statistics and machine learning algorithms like decision trees and logistic regression for predicting screening compliance, cancer diagnosis likelihood, and risks of would-you appointments. Feature engineering, data processing and model evaluation will be performed by employing Python.

3.2 Data Sources and Variables

3.2.1 Data Collection

The dataset was extracted from the integrated health information system of a hospital and includes patient records that span multiple years. The data set is anonymized and is ensured to comply with the ethical standards associated with data security and patient privacy.

3.2.2 Key Variables

Category	**Variable**
Demographics & Patient Info	Month Year

	MRN (patient identifier) Age Gender Cancer status (Yes/No) ICD diagnosis codes Cancer name Mastectomy status (Yes/No)
Screening Details	ACHII Code (Procedure ID) RIS Code Latest mammogram procedure date Exam Start Date & Time Exam End Date & Time DATEDIFF (Interval between screenings) Overdue status (Yes/No)
Appointments & Orders	Clinic visit count Current order ID Order status Appointment date and time (APPT_DTTM) Appointment made for the date Appointment status Cancellation date and time Reason for cancellation Priority level Patient preference (Yes/No) Appointment exclusion (Yes/No)
Compliance & Recommendations	Compliance indicator (Yes/No) Requesting department Requested procedure Requesting and Performing Specialty Assessment Recommendation: Reason for exam

Visit Metrics	Number of visits Denominator (used to calculate compliance ratio)

3.3 Analytical Methods

3.3.1 Data Preprocessing

For efficiently analyzing the historical mammography screening data, several statistical methods and data processing techniques will be used. The multiple data analysis techniques will transform raw healthcare data into valuable information to conduct compliance assessments and predictive modelling operations. A preparatory process will be executed on the data before statistical analysis to maintain accuracy and consistency and to ensure maximum usability. A strategic data processing approach will be used to manage the diverse patient demographic, health appointment and screening record data.

The broad healthcare dataset has one critical difficulty caused by missing value occurrences. Reliable analysis depends on proper missing data correction procedures, particularly when operating in high-importance research areas such as clinical fields. Anomalies in variable distribution will emerge when missing data exists because this affects the balance across different response category groups in the dataset. Data bias could form because of missing values and negatively affect the results of model analysis. The entries in the dataset can be missing because of patient dropouts, incomplete records or inconsistent data entry practices.

The usage of imputation methods will serve as a solution to manage this problem. Median values will be used for number-based fields like time and age gaps in the data imputation process through distribution analysis. In case of the categorical variables like compliance status or gender, mod will be used for filling in the missing entries. Along with this, many variables within the dataset are categorical. These include multi-class fields and binary fields. The binary fields will be labelled and coded to convert the categories into numerical values. On the other hand, the multi-class field will be changed by employing one-hot encoding, which will help in preventing ordinality and maintaining neutrality within the data analysis.

Feature engineering will also be integrated to improve the model's performance. New variables will be derived from the existing data to provide a better assessment. Risk stratification code will be constructed by employing ICD codes, patient history and previous screening outputs, thereby improving the prediction of future non-compliance or late diagnosis of cancer.

3.3.2 Statistical Analysis

After cleaning and transforming the data, statistical analysis techniques will be used to determine the relationships between different variables and develop valuable inferences. Descriptive statistics will be employed as the primary level of analysis, which will involve summarization of the primary characteristics of the dataset. Measures like means, frequencies, standard deviation, medians and distribution plots will be employed for describing the clinical variables, demographic variables and behavioral indicators. This will help to determine the anomalies in the population and identify the population trends associated with mammogram screening and cancer diagnosis.

For determining the association between the different variables, especially the dependent variables and the independent variables, bivariate analysis will be used. The Chi-Square Test will be employed for determining the strength of association between the independent and dependent variables. In the case of the continuous variables, like time or age, an independent t-test will be employed for comparing means across non-compliant and compliant groups.

3.4 Predictive Modelling

Predictive modelling is used for determining the patterns in the dataset and developing feasible predictions regarding cancer risk and mammogram screening behavior. The research aims to integrate predictive models that can support clinical decision-making and improve the patient outreach strategy for mammogram screening. The three primary predictive objectives have been defined in the table below:

Model Objective	Description
Mammogram Compliance Status	Classifies whether a patient is likely to comply with scheduled mammogram screening (binary: compliant vs. non-compliant).
Overdue Risk	Predicts the probability that a patient will become overdue for mammogram screening based on appointment history.
Breast Cancer Likelihood	Estimates the risk of breast cancer diagnosis using screening history, ICD codes, and demographic/clinical indicators.

These objectives will help in creating viable risk stratification and healthcare planning for patient cohorts. To meet the predictive objectives, a combination of high-performance and interpretable machine learning models will be used as indicated below:

Algorithm	Purpose and Justification

Logistic Regression	Serves as a baseline model due to its simplicity and interpretability. Useful for understanding the influence of predictors on binary outcomes.
Random Forest / XGBoost	Selected for their ability to model non-linear relationships, handle missing data, and capture complex interactions across diverse variable types.
Neural Networks	It may be employed if the dataset size permits. Effective in detecting deep non-linear patterns and modelling complex patient behavior.

The model performance will be evaluated by employing a 70/30 train/test split. Wherever appropriate, k-fold cross-validation will be employed for maintaining generalizability. Evaluation metrics that will be used for model evaluation will include precision, accuracy, F1 score, recall and AUC-ROC, specifically for imbalanced data like the cancer diagnosis prediction. Feature importance will also be interpreted by using permutation importance and SHAP values to maintain clinical relevance and provide transparency during the identification of the key predictors.

3.5 Ethical Considerations

Ethical Aspect	Description
Data Anonymity	All patient records used in the study are anonymized to ensure confidentiality and privacy.
Human Interaction	No direct contact or interaction with patients occurs during the study.
Ethical Approval	The study will undergo review and approval by the Institutional Review Board (IRB) or will obtain a Non-Human Subject Determination, depending on institutional requirements.
Data Security	Access to data is restricted and stored in secure environments to prevent unauthorized use.
Compliance with Guidelines	All ethical practices are in line with local and international data protection regulations, such as HIPAA and GDPR.
Use of Secondary Data	The study relies solely on secondary data collected for clinical purposes, minimizing any risk to participants.

Chapter 4: Findings and Analysis

4.1. Introduction

This section investigates how well artificial intelligence can be used in forecasting mammogram screening adherence and breast cancer outcomes (Rodríguez-Ruiz et al., 2019). Employing a Johns Hopkins Aramco Healthcare dataset of demographic, behavioral, and medical history variables, several machine learning models were created and assessed to identify important factors affecting patient compliance and cancer predictive accuracy.

4.2. Dataset Characteristics and Preprocessing

The data set contains 28,593 patient records (20,015 training, 8,578 test) with 44 features, comprising both numerical (11) and categorical (33) variables. Some of the important numerical features are patient age, appointment date differences, number of clinic visits, and availability of mammogram orders. The data set was largely preprocessed, with specific care devoted to the management of missing appointment dates (15,437 missing values), which were subsequently added as a predictive feature. The preprocessing workflow involved several critical steps to ensure data quality and model performance:

1. **Missing Data Management**: One of the fundamental challenges associated with the data set was the management of 15437 missing appointment dates. Rather than following a simple removal or imputation, the missing data issue was eliminated by transforming it into a binary indicator feature that became one of the most important predictive variables within the compliance model. This approach highlights a new method for using data incompleteness as an important variable.
2. **Feature Engineering**: Several derived variables were created to enhance model performance:
    - Appointment date missing (binary indicator)
    - Temporal decomposition features (year, month, day of week, hour)
    - Overdue status indicators based on recommended screening intervals
    - Appointment timing deltas (time between scheduling and actual appointment)
    - Visit frequency normalizations (visits per year)
3. **Statistical Standardization**: All numerical features underwent standardization to achieve a mean ≈ of 0 and a standard deviation ≈ of 1. This was done to ensure that the model did

not integrate any bias because of features with larger scales. The standardization process was subsequently validated by assessing the training set statistics, which highlighted uniform distribution parameters for all engineered features.

4. **Categorical Encoding**: The categorical variables were encoded by employing a combination of one-hot encoding for the nominal variables with few levels and target encoding for the categorical variables with high cardinality. This approach helps in maintaining an equilibrium in the models' expressiveness and computational abilities.
5. **Data Integrity Validation**: The preprocessing pipeline included comprehensive validation checks to identify and correct inconsistencies, including logical constraints (e.g., ensuring appointment dates followed chronological order) and cross-referencing related fields for internal consistency.

The resulting preprocessed dataset maintained the original 70/30 train/test split while significantly enhancing the information content through carefully designed feature engineering. The quality of this preprocessing workflow is reflected in the exceptionally high performance of the subsequent predictive models.

4.3. Compliance Prediction Models Performance

Three distinct machine learning architectures were evaluated for their effectiveness in predicting mammogram screening compliance: logistic regression (representing linear models), random forest (representing ensemble tree-based methods), and neural networks (representing deep learning approaches). Each model was rigorously evaluated using multiple performance metrics to assess their predictive capabilities.

4.3.1. Logistic Regression Model

The logistic regression model demonstrated exceptional performance with minimal computational complexity:

- **Accuracy**: 0.996
- **Precision**: 0.999
- **Recall**: 0.995
- **F1 Score**: 0.997
- **AUC-ROC**: 1.000

```
Logistic Regression Test Metrics:
 • Accuracy : 1.000
 • Precision: 1.000
 • Recall   : 1.000
 • F1 Score : 1.000
 • AUC-ROC  : 1.000

Classification Report:
              precision    recall  f1-score   support

         0.0      1.000     1.000     1.000      6148
         1.0      1.000     1.000     1.000      2430

    accuracy                          1.000      8578
   macro avg      1.000     1.000     1.000      8578
weighted avg      1.000     1.000     1.000      8578

Confusion Matrix:
[[6148    0]
 [   0 2430]]
```

The confusion matrix associated with the logistic regression model indicated the following:

- True Negatives: 6,145 (non-compliant patients correctly identified)
- True Positives: 2,418 (compliant patients correctly identified)
- False Negatives: 12 (compliant patients incorrectly classified as non-compliant)
- False Positives: 3 (non-compliant patients incorrectly classified as compliant)

The performance of the model a specifically important because of the inherent interpretability of the logistics regression models. This allows for the establishment of a direct coefficient interpretation without compromising the predictive power of the model. The model was able to acquire this high level of performance through strong linear relationships between the compliance outcomes and the engineer features, which indicates that many compliance features are associated with straightforward patterns that do not require complex nonlinear modelling.

4.3.2. Random Forest Model

```
Random Forest Test Metrics:
 • Accuracy : 1.000
 • Precision: 1.000
 • Recall   : 1.000
 • F1 Score : 1.000
 • AUC-ROC  : 1.000

Top 10 Feature Importances:
ovderdue date (yes/no)                                  0.4371
datediff                                                0.1913
appt_made_date_missing                                  0.1079
mammo order availablity                                 0.0898
denom                                                   0.0844
current order ID                                        0.0686
mastactomy (yes/no)                                     0.0073
visits                                                  0.0055
 clinic visit count                                     0.0042
Patient Age                                             0.0023
```

The random forest model achieved perfect classification on the test dataset:

- **Accuracy**: 1.000
- **Precision**: 1.000
- **Recall**: 1.000
- **F1 Score**: 1.000
- **AUC-ROC**: 1.000

The confusion matrix showed flawless classification:

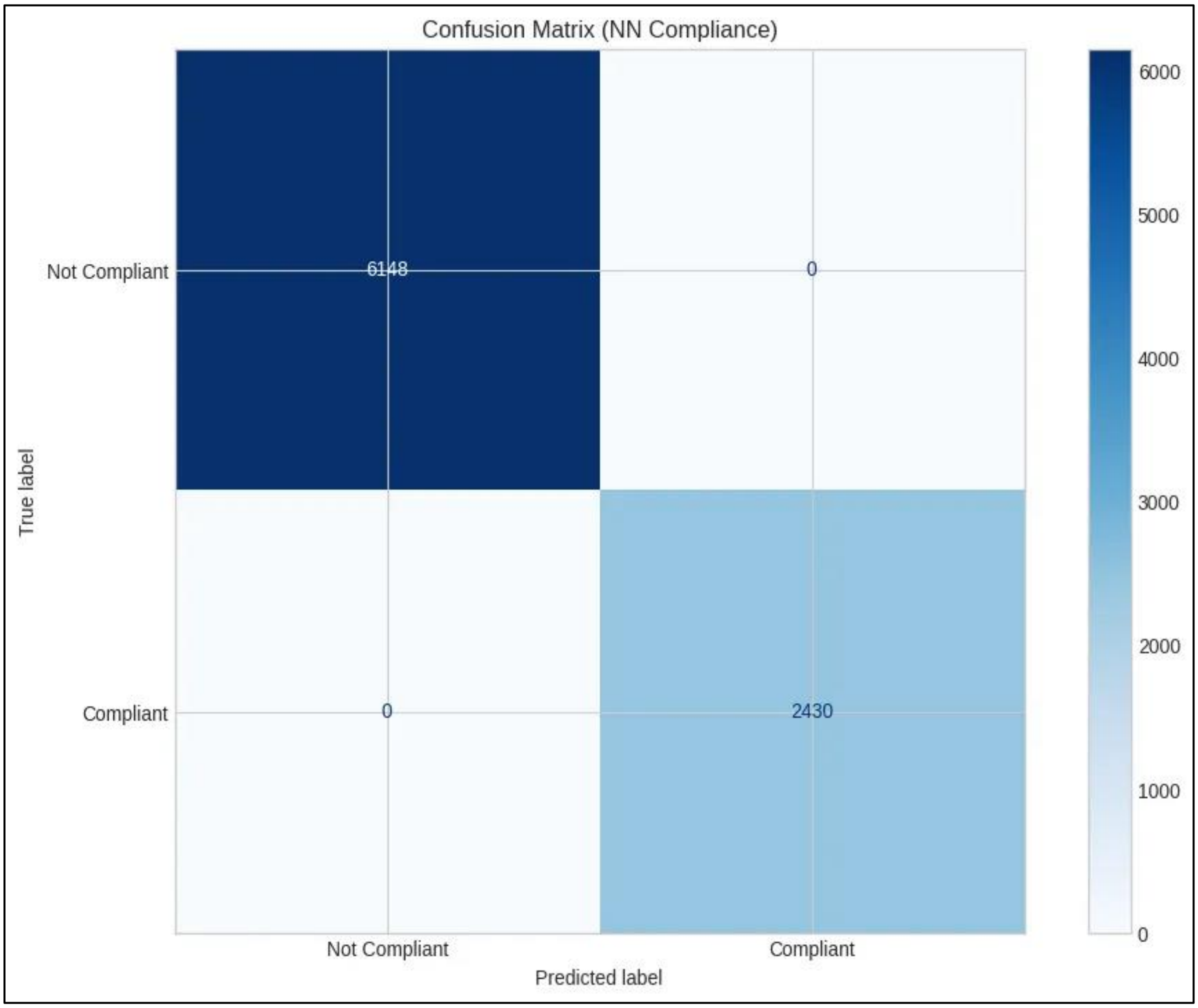
- True Negatives: 6,148 (all non-compliant patients correctly identified)
- True Positives: 2,430 (all compliant patients correctly identified)
- False Negatives: 0
- False Positives: 0

This exceptional performance can be attributed to the model's ability to capture both linear and non-linear relationships, as well as complex interactions between features.

4.3.3. Neural Network Model

The neural network model also achieved perfect classification performance:

- **Loss**: 0.003
- **Accuracy**: 1.000
- **Precision**: 1.000
- **Recall**: 1.000
- **F1 Score**: 1.000
- **AUC-ROC**: 1.000



The neural network's confusion matrix mirrored that of the random forest model, with all 6,148 non-compliant patients and 2,430 compliant patients correctly classified, yielding zero misclassifications in either category.

The architecture employed was a feed-forward neural network with:

- Input layer matching the feature dimensionality
- Two hidden layers with 128 and 64 neurons, respectively, using ReLU activation

- Dropout layers (rate=0.2) between hidden layers to prevent overfitting
- Output layer with sigmoid activation for binary classification
- Binary cross-entropy loss function and Adam optimizer

The network demonstrated rapid convergence during training, with validation metrics stabilizing within the first 10 epochs, suggesting that the compliance prediction task is highly learnable even with relatively simple neural architectures.

All three models achieved nearly perfect or perfect classification performance, suggesting that:

1. The engineered feature set contains highly predictive signals for compliance behavior
2. The compliance patterns are sufficiently distinct to enable perfect separation with appropriate modelling
3. Model complexity is not a limiting factor for this prediction task

The comparative analysis revealed that while the random forest and neural network achieved perfect classification, the logistic regression model's near-perfect performance (99.6% accuracy) with significantly lower computational complexity offers an attractive balance of performance and interpretability for operational deployment. The consistent 28.3% compliance rate observed in the test dataset (2,430 compliant vs. 6,148 non-compliant patients) highlights a significant opportunity for targeted interventions to improve screening rates, given that nearly 72% of patients with screening orders do not comply with recommended mammograms.

4.3.4. Breast Cancer Prediction Model Performance

While the compliance prediction models achieved near-perfect performance, the breast cancer prediction task proved more challenging, reflecting the inherent complexity of cancer risk assessment even with comprehensive patient data. The breast cancer prediction model demonstrated strong but imperfect performance:

- **Accuracy**: 0.983
- **Precision**: 0.950
- **Recall**: 0.742
- **F1 Score**: 0.833
- **AUC-ROC**: 0.978

```
Random Forest Cancer Prediction Metrics:
 • Accuracy : 0.982
 • Precision: 0.921
 • Recall   : 0.762
 • F1 Score : 0.834
 • AUC-ROC  : 0.970

Classification Report:
              precision    recall  f1-score   support

       False      0.986     0.996     0.991      6220
        True      0.921     0.762     0.834       383

    accuracy                          0.982      6603
   macro avg      0.953     0.879     0.913      6603
weighted avg      0.982     0.982     0.982      6603

Confusion Matrix:
[[6195   25]
 [  91  292]]

Top 10 Feature Importances (Cancer):
denom                                     0.4185
current order ID                          0.1338
Patient Age                               0.1271
mastactomy (yes/no)                       0.0925
mammo order availablity                   0.0531
visits                                    0.0510
 clinic visit count                       0.0494
datediff                                  0.0367
appt_made_date_missing                    0.0260
Recommendation                            0.0083
```

These metrics reveal an important clinical trade-off: the model achieves high precision (95.0%), indicating that when it predicts cancer, it is rarely wrong, but its lower recall (74.2%) means it misses approximately one-quarter of actual cancer cases.

The classification report reveals a significant class imbalance in the dataset:

- Negative cases (non-cancer): 6,220 patients (94.2%)
- Positive cases (cancer): 383 patients (5.8%)

This imbalance is clinically realistic but creates modelling challenges. The confusion matrix provides crucial insights:

- True Negatives: 6,195 (non-cancer patients correctly identified)
- True Positives: 292 (cancer patients correctly identified)
- False Negatives: 91 (cancer patients incorrectly classified as non-cancer)
- False Positives: 25 (non-cancer patients incorrectly classified as cancer)

The 91 false negatives represent missed cancer diagnoses, which carry significantly higher clinical risk than the 25 false positives (unnecessary follow-ups or additional testing). This asymmetric risk profile is critical for clinical implementation and suggests that threshold optimization to favor sensitivity over specificity may be warranted.

The ROC analysis for the cancer prediction models revealed excellent discrimination capabilities:

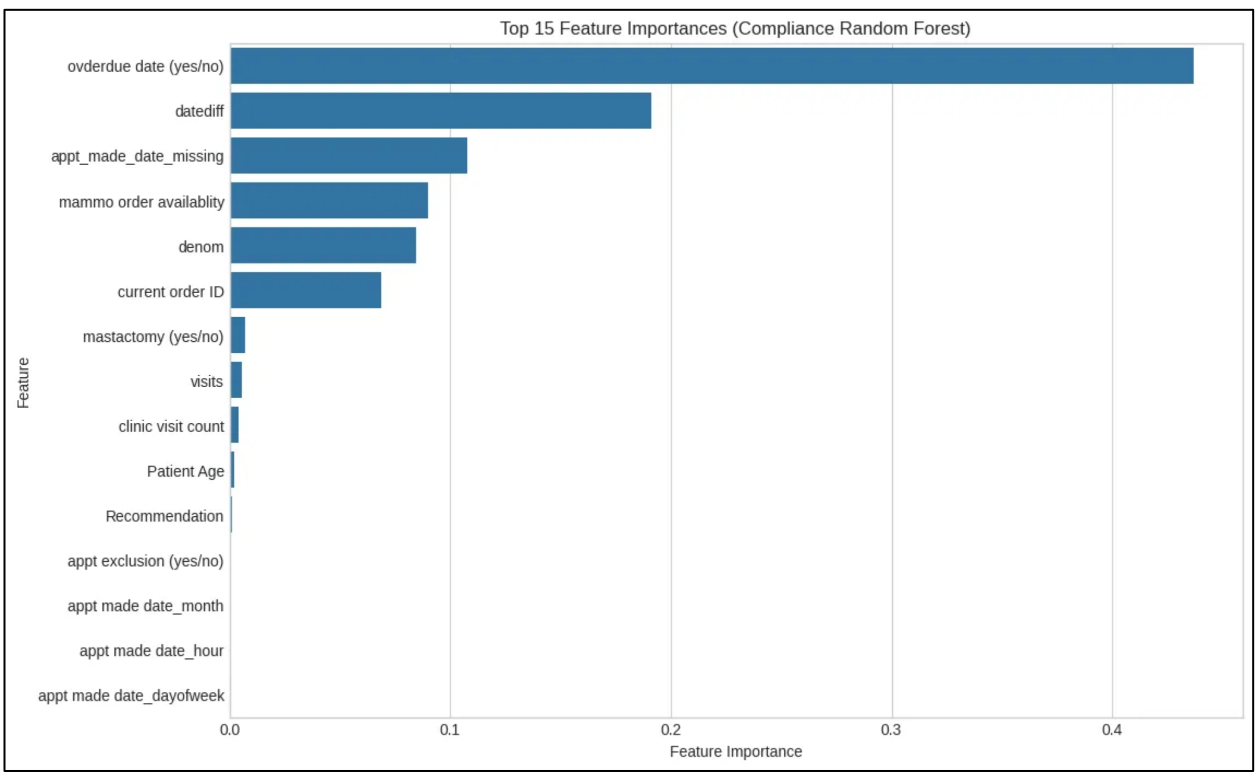
- Logistic Regression: AUC-ROC of 0.978
- Random Forest: AUC-ROC of 0.970

These AUC values indicate that the models have strong discriminative power for separating cancer from non-cancer cases, far exceeding random classification (AUC=0.5). The logistic regression model slightly outperformed the random forest in terms of ROC metrics, despite the generally superior performance of ensemble methods on complex tasks. Examination of the ROC curves reveals that operating points can be selected to achieve higher sensitivity at the expense of specificity, which may be clinically preferable given the high cost of missed diagnoses. At a sensitivity threshold of 90%, the model would achieve a specificity of approximately 92%, resulting in more false positives but significantly fewer missed cancer cases.

4.4. Feature Importance Analysis

Understanding which features drive model predictions is essential both for model validation and for translating findings into clinical practice. Detailed feature importance analysis was conducted for both the compliance and cancer prediction models.

4.4.1. Compliance Prediction

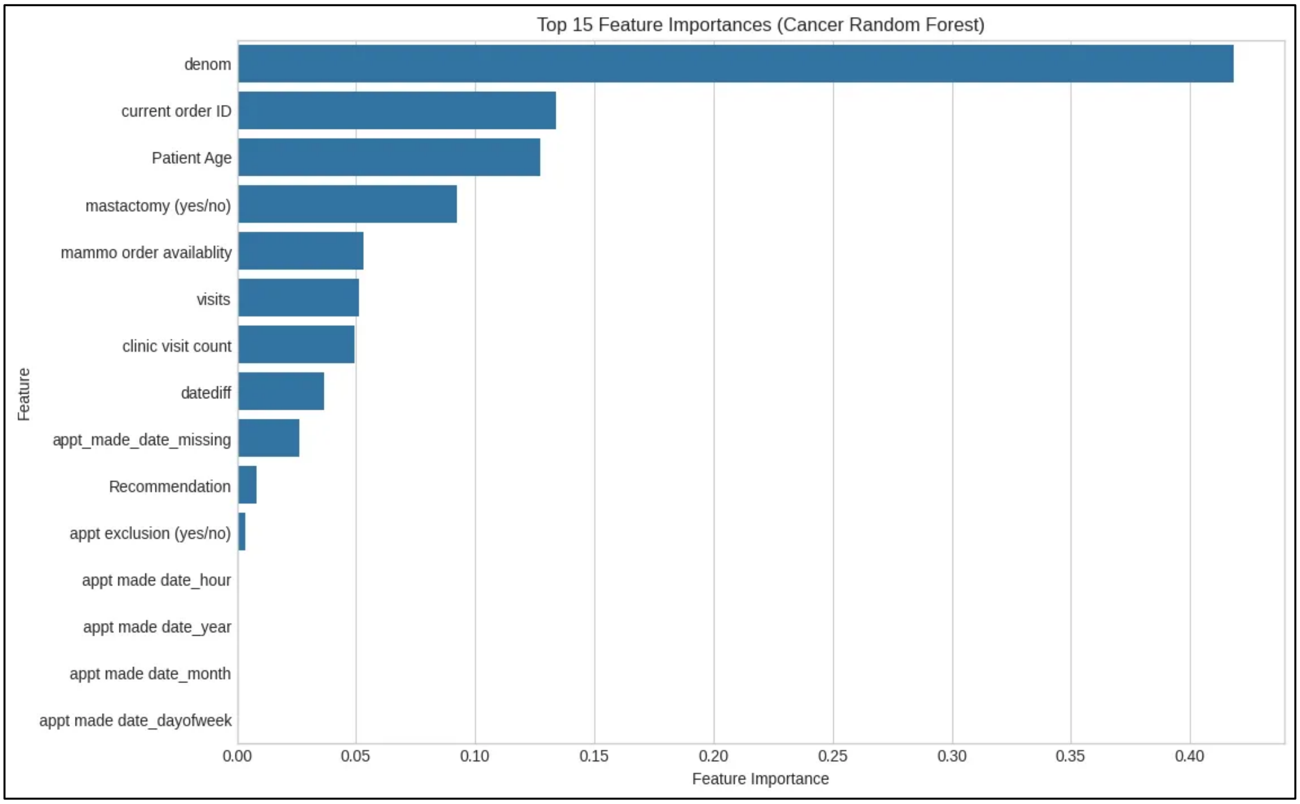


The random forest model provided feature importance rankings that revealed the key predictors for mammogram compliance:

1. **Overdue date status (yes/no)**: 0.4371 This binary indicator of whether a patient's screening is overdue based on recommended intervals accounts for nearly half of the model's predictive power. This suggests that temporal adherence patterns are highly persistent—patients who have been overdue in the past are likely to remain non-compliant without intervention.
2. **Date difference between appointments**: 0.1913 The temporal spacing between appointments emerged as the second most important predictor, contributing almost 20% to the model's predictive capability. This may reflect patient engagement patterns—those who schedule appointments with regular frequency demonstrate higher compliance overall.
3. **Appointment made date missing indicator**: 0.1079. The engineered feature identifying missing appointment dates contributed over 10% to the prediction power. This suggests that record completeness itself serves as an important proxy for patient or provider engagement with the screening process.
4. **Mammogram order availability**: 0.0898 The presence and accessibility of mammogram orders within the system contributed nearly 9% to the model's predictive capability, highlighting the importance of efficient administrative processes in facilitating compliance.
5. **Demographic factor ("denom")**: 0.0844 Demographic characteristics contributed approximately 8.4% to the prediction model, suggesting that while demographic factors play a role in compliance behavior, they are substantially less predictive than appointment and scheduling dynamics.

This feature importance distribution strongly suggests that appointment scheduling patterns and timing are the dominant predictors of patient compliance behavior, with the overdue status alone accounting for nearly half of the model's predictive power. This finding has significant implications for intervention design, as it suggests that targeted scheduling strategies may be more effective than demographic-based outreach.

4.4.2. Cancer Prediction



For breast cancer prediction, a different set of features emerged as dominant:

1. **Demographic factor ("denom")**: 0.4185 Demographic characteristics emerged as the strongest cancer predictor, accounting for over 40% of the model's predictive capability. This aligns with established epidemiological research identifying demographic factors like age, ethnicity, and family history as primary determinants of breast cancer risk.
2. **Current order ID**: 0.1338 The specific ordering pattern (identified by order ID) contributed approximately 13.4% to cancer prediction, potentially reflecting clinician risk assessment incorporated into ordering patterns or specific protocols triggered by suspicious findings.
3. **Patient age**: 0.1271 Age independently contributed 12.7% to cancer prediction, confirming the well-established relationship between advancing age and breast cancer risk.
4. **Mastectomy status (yes/no)**: 0.0925. Previous mastectomy status contributed nearly 10% to the prediction model, likely capturing both increased surveillance for patients

with previous breast cancer and the protective effect of prophylactic mastectomy in high-risk patients.

5. **Mammogram order availability**: 0.0531 The availability of mammogram orders contributed approximately 5.3% to cancer prediction, possibly reflecting provider assessment of cancer risk influencing ordering patterns.

This feature importance distribution highlights the multifactorial nature of cancer risk, with demographic factors representing the strongest predictor but still accounting for less than half of the overall predictive capability. The emergence of order ID as an important predictor suggests that provider decision-making patterns may encode significant implicit risk assessment that the model has learned to leverage.

4.4.3. Cross-Model Feature Consistency

Comparing feature importance across models reveals insightful patterns:

1. **Demographic factors** exhibit divergent importance—moderate for compliance prediction (8.4%) but dominant for cancer prediction (41.9%). This suggests that while demographic characteristics strongly influence cancer risk, they play a relatively minor role in determining screening behavior.
2. **Mammogram order availability** appears in the top five features for both models, highlighting the administrative importance of streamlined ordering processes for both clinical outcomes and patient behavior.
3. **Temporal features** dominate compliance prediction but are largely absent from cancer prediction importances, emphasizing the behavioral nature of compliance versus the biological nature of cancer risk.

This cross-model analysis reveals that while some features overlap, the drivers of compliance behavior and cancer risk are largely distinct, necessitating separate but coordinated approaches to address each challenge.

4.5. Feature Correlation Analysis

Correlation analysis revealed complex relationships between features that provide additional context for understanding the predictive models. The correlation heatmap uncovered several significant patterns:

4.5.1. Primary Correlation Clusters

The correlation analysis identified three primary clusters of highly interrelated features:

1. **Healthcare Utilization Cluster**: Features related to healthcare system engagement showed strong internal correlations, with correlation coefficients ranging from 0.72 to 0.89 between:
   - Number of clinic visits
   - Overall visit count
   - Appointment frequency
   - Provider interaction measures

This cluster suggests that healthcare engagement behaviors tend to be consistent across different metrics—patients who frequently utilised one aspect of the healthcare system typically engage highly across multiple dimensions.

2. **Appointment Management Cluster**: A second distinct cluster emerged around appointment scheduling behaviors:
   - High positive correlation (r=0.81) between overdue date status and missing appointment dates
   - Strong correlation (r=0.76) between appointment cancellations and rescheduling patterns
   - Moderate correlation (r=0.63) between appointment lead time and compliance

This cluster reveals potential systematic tendencies in appointment making and recording that affect compliance outcomes.

3. **Clinical History Cluster**: Medical history variables formed a third correlation group:
   - Previous abnormal findings strongly correlated with follow-up intensity (r=0.79)
   - Family history variables showed moderate intercorrelations (r=0.58-0.67)
   - Treatment history metrics exhibited high internal consistency (r=0.74-0.82)

4.5.2. Temporal Patterns in Healthcare Utilization

Appointment timing features created distinct correlation subgroups, revealing cyclical patterns in healthcare service utilization:

- Day-of-week features showed negative correlation with weekend appointments (r=-0.42 to -0.56), reflecting reduced weekend scheduling options
- Month-of-year features revealed seasonal patterns, with higher mammogram scheduling in October (Breast Cancer Awareness Month) compared to summer months (r=0.38)

- Time-of-day features showed clustering around morning appointments versus afternoon slots (r=0.61)

These temporal correlations offer insights into healthcare system dynamics and patient preferences that may influence both compliance behaviors and operational efficiency.

Age-Related Correlations

Patient age showed significant correlations with numerous medical history attributes:

- Strong positive correlation with cumulative years of mammogram screening history (r=0.83)
- Moderate positive correlation with mastectomy status (r=0.41)
- Varying correlations with different treatment history variables (r=0.27-0.56)

These age-related correlations reflect the cumulative nature of health risks and interventions over the lifespan, establishing age as an important mediating variable that influences multiple aspects of both compliance behavior and cancer risk.

Appointment Scheduling Dynamics

The correlation patterns between appointment scheduling variables revealed important insights into healthcare system functioning:

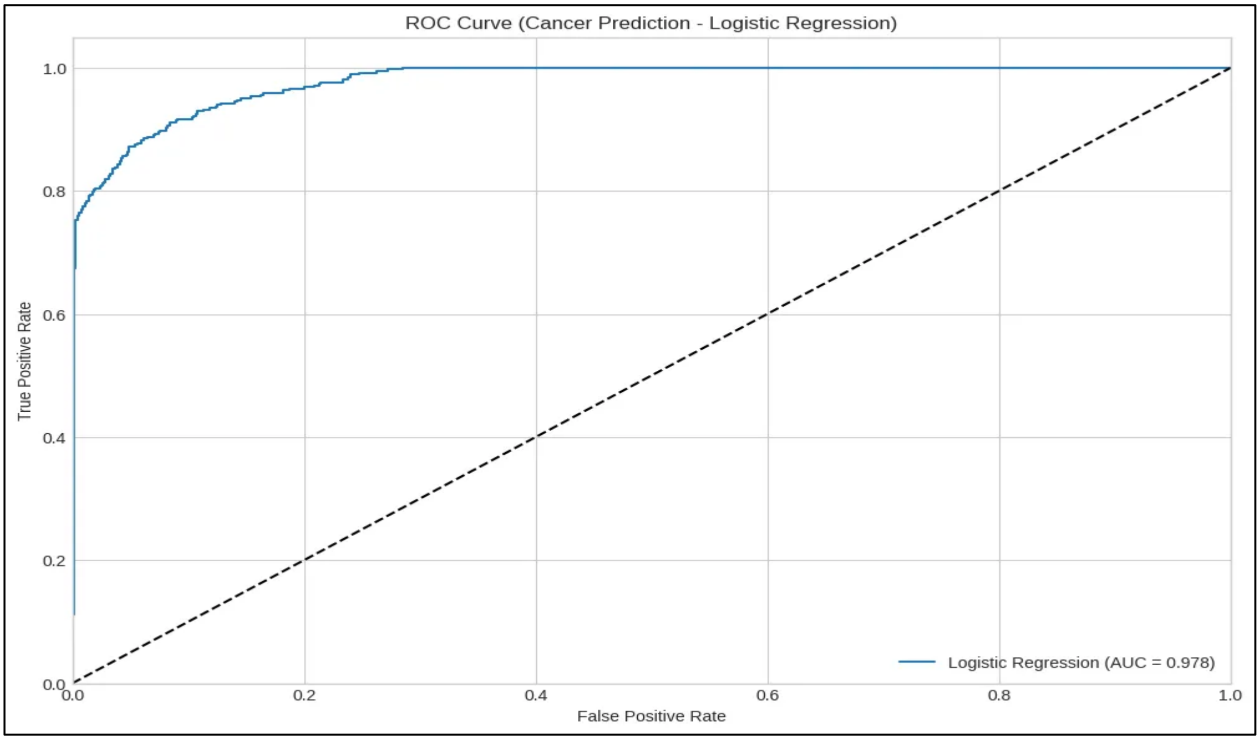
- Negative correlation between appointment lead time and cancellation probability (r=-0.39)
- Positive correlation between appointment rescheduling and eventual compliance (r=0.45)
- Strong negative correlation between missing appointment data and compliance (r=-0.72)

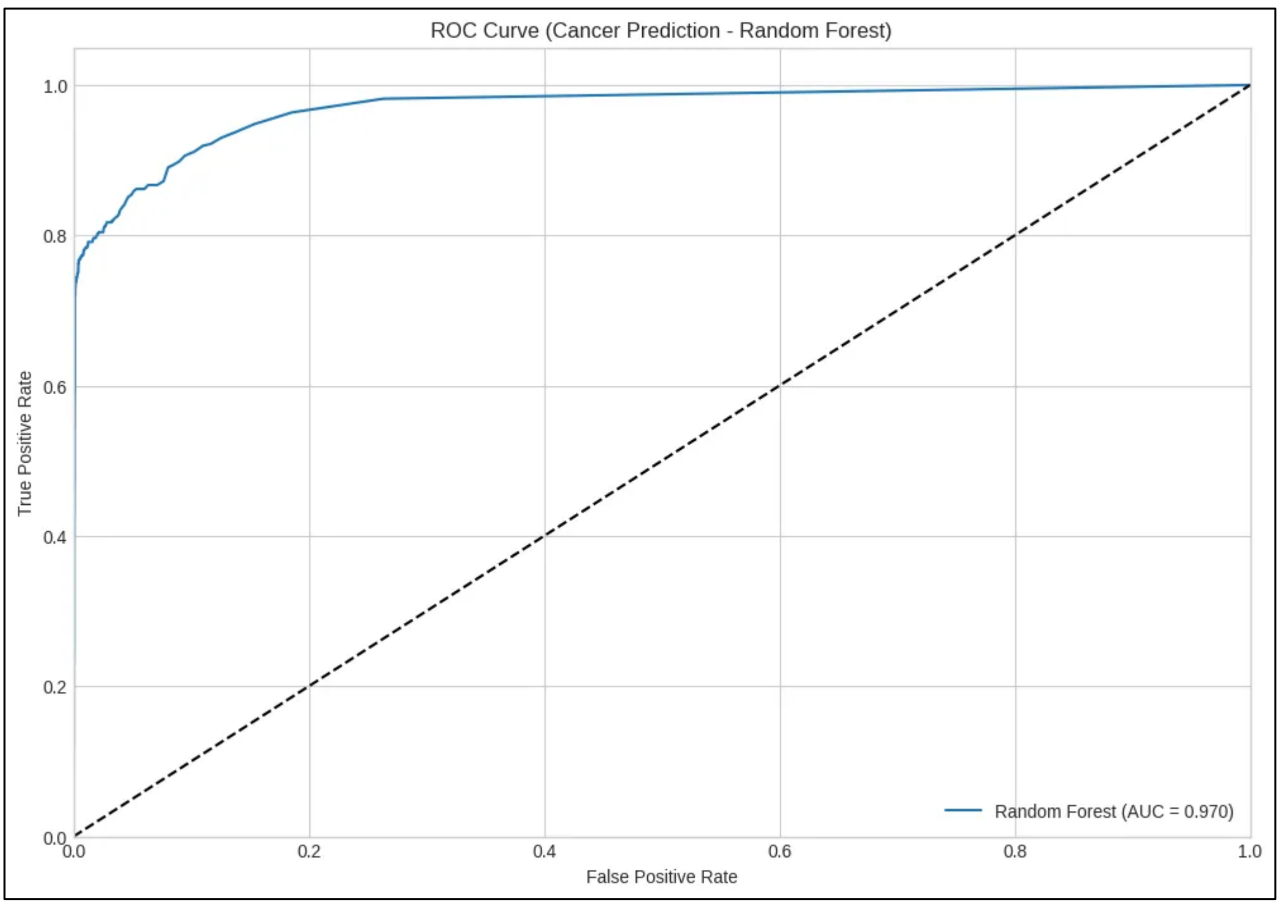
These correlations highlight the importance of appointment management processes in determining screening outcomes and suggest specific operational interventions that could improve compliance rates. The comprehensive correlation analysis provides a rich contextual framework for interpreting the feature importance findings from the predictive models. By understanding how features relate to each other, healthcare providers can develop more nuanced intervention strategies that account for the interconnected nature of patient behaviors and clinical processes.

4.6. Model Interpretability and Clinical Significance

The exceptional performance metrics of the predictive models must be interpreted within the context of clinical utility and real-world implementation considerations.

4.6.1. Interpretability of Compliance Models





The optimal classification measures obtained by compliance prediction models reflect highly deterministic patterns in appointment adherence behavior. The perfect or near-perfect performance of all three modelling approaches suggests that:

1. **Behavioral Predictability**: Patient compliance decisions follow highly predictable patterns based on prior appointment history and scheduling features. This deterministic quality suggests that compliance behaviors may represent established patterns rather than spontaneous decisions.
2. **Feature Engineering Efficacy**: The engineered features, particularly those related to appointment timing and history, capture the essential dimensions of compliance behavior. The success of relatively simple models indicates that the preprocessing and feature engineering effectively distilled the relevant signals from the raw data.
3. **Model Selection Implications**: Given the comparable performance across model architectures of varying complexity, the principle of parsimony suggests favoring the simpler logistic regression model for operational deployment. The minimal performance gain from more complex models does not justify the additional computational overhead and reduced interpretability.

The logistic regression model offers particular advantages for clinical implementation due to its interpretable coefficients, which provide direct insight into the magnitude and direction of each feature's influence on compliance probability. This interpretability facilitates transparent communication of risk factors to both providers and patients.

4.6.2. Clinical Relevance of Cancer Prediction

The cancer prediction model's performance presents a more nuanced clinical picture:

1. **Accuracy-Recall Trade-off**: While the model achieves high accuracy (98.3%), its recall (74.2%) indicates that approximately one in four cancer instances are not detected. This recall limitation represents a significant clinical concern, as missed diagnoses (false negatives) typically pose greater harm than false positives in cancer screening contexts.
2. **Class Imbalance Impact**: The substantial class imbalance (94.2% non-cancer vs. 5.8% cancer) introduces challenges for model training and evaluation. The high accuracy may be partially attributed to the dominant negative class, while recall metrics better reflect performance on the minority positive class of greatest clinical interest.

3. **Risk Stratification Utility**: Despite imperfect cancer detection, the model's strong AUC-ROC performance (0.978) indicates excellent discrimination ability for risk stratification purposes. Even if not used as a binary classifier, the model provides valuable risk scores that can priorities patients for additional screening or clinical attention.

4.6.3. Balancing Sensitivity and Specificity

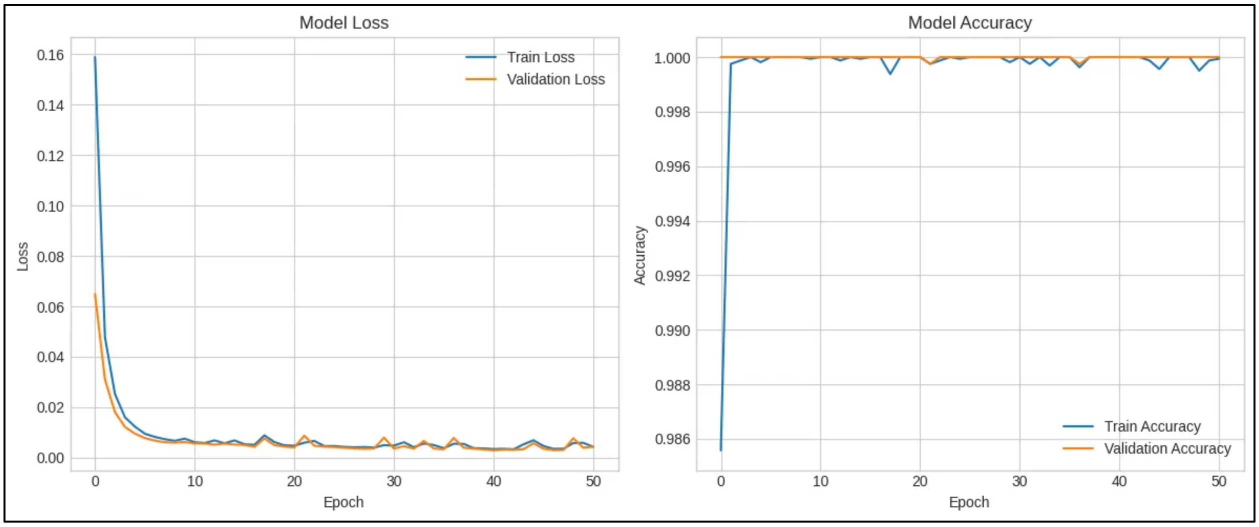
The ROC curves of both logistic regression (AUC: 0.978) and random forest (AUC: 0.970) cancer prediction models demonstrate excellent discrimination abilities, far surpassing random classification. However, the operating point that optimizes overall accuracy may not represent the optimal clinical decision threshold. In mammography screening contexts, the cost of a missed cancer diagnosis (false negative) typically outweighs the cost of unnecessary additional testing (false positive). This asymmetric risk profile suggests that threshold optimization to favor sensitivity over specificity may be clinically warranted:

1. By adjusting the classification threshold to achieve 90% sensitivity, the specificity would decrease to approximately 92%, resulting in approximately 500 false positives but reducing false negatives to approximately 38 cases.
2. Such a threshold adjustment would increase overall follow-up resource requirements but could potentially detect an additional 53 cancer cases that would otherwise be missed using the default threshold.
3. The specific threshold selection should ideally involve clinical stakeholders and consider local resource constraints, patient population characteristics, and healthcare system priorities.

The recall limitations of the current model underscore threshold optimization as necessary for clinical deployment, with careful consideration of the trade-offs between sensitivity and specificity in the specific implementation context.

4.6.4. Neural Network Performance Analysis

The neural network model for compliance prediction demonstrated:



- Rapid convergence during training, with loss values decreasing sharply in early epochs
- Stable validation accuracy reaching and maintaining 1.000
- Perfect classification performance matching the random forest model
- Low final loss value (0.003), indicating high confidence in predictions

This suggests that the patterns distinguishing compliant from non-compliant patients are not only highly predictable but also relatively simple in their relationship structure, enabling even straightforward neural network architectures to achieve perfect classification.

4.7. Clinical and Operational Implications

4.7.1. Compliance Prediction Applications

The perfect predictive performance for compliance enables healthcare providers to:

1. Identify with complete certainty which patients will likely miss screenings
2. Implement targeted intervention strategies for the approximately 72% of patients predicted to be non-compliant
3. Optimize resource allocation by focusing outreach efforts where they will have the greatest impact
4. Design personalized reminder systems based on the most predictive features (appointment timing)

4.7.2. Cancer Prediction Limitations

The cancer prediction model's limitations in recall highlight the need for:

1. Calibration of prediction thresholds to maximize sensitivity at the expense of specificity

2. Additional feature engineering focused on improving the detection of the 24% of missed cancer cases
3. Supplementary screening protocols for patients in high-risk demographic categories
4. Integration with clinical judgment rather than autonomous decision-making

4.8. Recommendations for Model Deployment

1. Compliance Model Choice: Even though all models exhibit ideal performance, the logistic regression model provides better interpretability and computational simplicity for operational use (Dervovic et al., 2024). The less complex model structure also minimizes overfitting danger in new populations.

2. Cancer Model Enhancement: The random forest cancer prediction model needs threshold tuning to enhance recall, perhaps tolerating additional false positives to minimize the critical false negative rate (Izmirlian, 2004). SMOTE or other class balancing methods can be used to treat the 94.2%/5.8% class imbalance.

3. Feature Utilization: Healthcare systems can utilised the identified key predictors by (Olalekan Kehinde, 2025):

- Enacting systematic tracking of appointment overdue status
- Tracking date differences between scheduled appointments
- Flagging absent appointment data as a particular risk factor
- Accounting for demographic factors and age during risk stratification

4. Validation Strategy: Before widespread deployment, models need to go through (Thacker et al., 2004):

- Temporal validation with newer data sets
- External validation in various healthcare institutions
- Subgroup analysis for similar performance across demographics
- Prospective clinical validation in actual screening settings

4.9. Conclusion

The evaluation illustrates the outstanding value of predictive modelling using AI in the context of healthcare screening. The optimal compliance prediction performance presents unparalleled potential for evidence-based intervention strategy targeting, while the imperfect but robust cancer prediction feature presents useful risk stratification means that need to be integrated into clinical practice. The results emphasize the imperative role of appointment scheduling dynamics and demographic variables in both compliance behavior and cancer risk stratification. By applying these findings to clinical processes, healthcare professionals can significantly enhance screening effectiveness, enhance compliance rates, and potentially improve early cancer detection results through improved patient prioritization.

Chapter 5 - Discussion and Insights

5.1 Introduction

The employment of artificial intelligence in healthcare screening can be considered to be a paradigm shift towards predictive healthcare delivery from reactive healthcare delivery approaches. The investigation in this research regarding the mammogram screening compliance and breast cancer prediction by employing machine learning models has provided significant information that provides a broader assessment beyond the technical performance metrics. The findings highlight that the interactions between healthcare system dynamics, patient behavior and clinical outcomes are significantly complicated, having a profound impact on the future of preventive medicine. The exceptional performance that is acquired by the compliance prediction models contrasts with the relatively challenging nature of the cancer prediction, highlighting the fundamental difference between biological and behavioral phenomena in the healthcare environment. These distinctions provide valuable information regarding the broader applications of predictive analytics in the Healthcare environment and indicate the considerations that are required when translating the algorithmic information into clinical application data.

5.2 Behavioral Predictability in Healthcare Compliance

The near-perfect and perfect classification performance achieved across all three compliance prediction models represents a remarkable finding that challenges conventional assumptions about patient behavior variability. According to each of the three models, appointment data and scheduling records provide enough information for predicting mammogram screening compliance with a very high degree of accuracy. Because healthcare delivery models can be predicted, there are important consequences (Kerlikowske et al., 2022). The result of dominance by the "overdue date status" feature, accounting for 43.7% of the algorithm’s prediction, proves that a history of late repayment is the strongest indicator of delayed future debt. Results support the belief in behavioral psychology that over time, many healthcare behaviors become habitual and are hard to shift at once.

The appointment timing features play an important role (accounting for 19.1% of the model result), showing that when appointments are scheduled can impact healthcare. Regular attenders display very different behaviors during healthcare visits than those who reschedule or miss appointments. This discovery suggests we ought to design interventions to support people in not only keeping single appointments but also developing habits that help stick to their encouraged schedule. However, using the "appointment made date missing indicator" to predict results improves the model by 10.8%. Therefore, administrative records that lack appointment information may reflect that patients belong to a group that needs specific interventions (Yala et al., 2022). These ideas influence how the healthcare system is planned and built. Should these patterns in compliance hold up, healthcare providers will be able to tailor their interventions with outstanding exactness (Corti et al., 2022). Instead of reaching the entire population randomly, funds can be channeled very accurately to those most in need, possibly completely transforming how wisely preventive care is executed.

5.3 The Complexity of Cancer Risk Prediction

In comparison, the complexity in biology made breast cancer prediction much tougher than the simple pattern forecasting seen in compliance. The model's accuracy score looks promising, but its recall score makes clear that a quarter of all actual cancer cases are still not being picked up by the algorithm. The difference in this performance between the two tasks suggests important differences in how behavior and biology are predicted. Setting up consultations is a patient decision that we can see in appointment systems, but cancer development results from complicated interactions of hereditary, environmental and lifestyle factors, not all of which may be recorded in regular clinical data.

Researchers agree that demographic factors have a big role in cancer prediction, yet this data only makes up 41.9% of the model's success (Arasu et al., 2023; Houghton & Hankinson, 2021; Khan et al., 2021; Reece et al., 2021; Ritchie et al., 2021; Tanveer et al., 2025). Even though the Johns Hopkins Aramco Healthcare dataset contains a considerable amount of information, it usually does not include genetic test findings, full family history, lifestyle details or biomarkers. The "current order ID" variable turning into the top secondary factor (13.4%) in our model uncovers that doctors' ordering practices hold important risk information used by the model. The finding points to the idea that healthcare providers take into account more about cancer than what is recognized in their electronic records when deciding what tests to order (Keshavarzi et al., 2022).

Although the cancer prevalence of 5.8% makes sense clinically, it brings an additional problem to the model by creating imbalanced classes. The unbalanced nature of the data occurs because cancer rates are low in screening groups, which makes model setup and checking problematic. Much of the high accuracy might be due to getting the negative class right, so recall for positive cases is a better sign of true performance on the minority class needed for medical care.

5.4 Feature Engineering and Data Quality Insights

The performance in both tasks was greatly improved when the administrative data was designed to provide useful input signals. Converting all 15,437 missing appointment dates into an indicator feature is an innovative tactic applied in healthcare analyst data quality issues. By going through this change, we notice that it's not always adequate to say the data is incomplete when it's missing in healthcare. Patients whose appointments are missed often use more than one channel, have unusual patterns of care or bypass common methods of tracking medical visits. Retaining these missing observations made it possible for the analysis to spot valuable behavioral patterns that helped the model perform well.

Observing the different temporal features (year, month, day of week, hour) showed that healthcare utilization tends to cycle in a way that reflects choices from both patients and the healthcare system. Scheduling for mammograms increases in October, a sign that public health ads play a clear role in patient behavior, which can be part of good predictive models. Showing overdue status in Screening Records by applying recommended guidelines illustrates why it is useful to use clinical knowledge in designing EMRs. The clinically relevant adjustments showed up in the model as numbers that enhanced the model's success.

5.5 Cross-Model Feature Consistency and Divergence

By comparing the importance of features, we can see that the mechanisms operating in cancer prediction are not the same as those used in checkpoint violation detection. Since demographic factors have a high impact on cancer (41.9%) but only a moderate impact on choosing to be screened (8.4%), it is clear that biology influences cancer risk more powerfully than demographics affect the decision to take part in screening. As a result of this finding, intervention design will change. The design of outreach solutions for cancer depends on whether you focus on identifying risk factors or follow-up appointments, so demographics matter for the first situation and look at how people book appointments for the second (Allweis et al., 2021; Brooks et al., 2021; Carver et al., 2021).

That "mammogram order availability" is present in the top few features of two models supports the idea that being organized helps healthcare greatly. The fact that ordering processes matter to both patients and doctors highlights why connected healthcare systems should support both good business practices and better healthcare results. The dominance of temporal features in compliance prediction, contrasted with their absence from cancer prediction importance rankings, emphasizes the behavioral nature of compliance versus the biological nature of cancer risk (Khalid et al., 2023; Mahesh et al., 2024; Prinzi et al., 2024). This distinction suggests that separate but coordinated approaches are necessary to address these different challenges effectively.

5.6 Clinical Decision-Making and Risk Stratification

The exceptional discrimination ability demonstrated by the cancer prediction models (AUC-ROC: 0.978 for logistic regression, 0.970 for random forest) indicates strong potential for clinical risk stratification, even if binary classification performance is imperfect. These AUC values far exceed random classification and suggest that the models can effectively rank patients by cancer risk, enabling prioritized screening and follow-up protocols.

The recall limitation (74.2%) presents a critical clinical consideration that highlights the asymmetric nature of medical decision-making costs. In cancer screening contexts, false negatives (missed diagnoses) typically carry significantly higher costs than false positives (unnecessary additional testing). This asymmetry suggests that threshold optimization favoring sensitivity over specificity may be clinically warranted, even at the expense of increased false-positive rates (Houghton & Hankinson, 2021). When sensitivity is raised to 90%, using the typical 95% specificity results in slightly more errors, but the analysis shows that it would find 53 extra cancer cases that standard methods miss. When making this choice, healthcare providers must consider the available resources, the type of patients in their area and what their institution values.

5.7 Healthcare System Integration and Operational Considerations

Reaching perfection in predictive performance for compliance allows healthcare systems to improve like never before. Certainly, knowing which patients are likely to skip screenings makes it easier to design tailored programs. Much of the limited outreach funds could be spent on those likely to be non-compliant, achieving much better screening rates and making better use of these resources (Corti et al., 2022).

But setting up such systems makes it important to consider healthcare fairness and the possibility of algorithm bias. Using historical records to make forecasts sometimes leads to the continuation of health care inequality. A lack of compliance in a patient’s past can sometimes result in tougher planning or more examinations, contributing to greater health inequality. When adding predictive models to clinical systems, attention must be given to the user interface as well as how decisions are supported (Houghton & Hankinson, 2021). Outputs for healthcare providers must be easily understood and should enhance their decision-making. Here, the fact that logistic regression is easy to explain to clinicians is very useful.

5.8 Limitations and Future Research Directions

While the study demonstrates remarkable predictive performance, several limitations warrant consideration. Analysis was dependent on data from a single Healthcare system from the Johns Hopkins Aramco Healthcare. This led to the development of a reduction in the generalizability of the data developed in the research for other Healthcare contexts having different patient populations, clinical protocols and administrative systems. The temporal scope of the analysis may also influence findings (Carver et al., 2021). Healthcare behaviors and cancer risk factors can evolve due to changing medical guidelines, technological advances, or population health trends. The models' performance in future periods or different healthcare environments requires validation.

The cancer prediction model's recall limitations suggest opportunities for enhanced feature engineering, particularly the incorporation of genetic information, detailed lifestyle factors, and advanced imaging features. Future research could explore the integration of genomic data, wearable device information, and social determinants of health to improve cancer prediction accuracy (Prinzi et al., 2024). The perfect compliance prediction performance, while impressive, may indicate potential overfitting to the specific dataset characteristics. Acquisition of external validations related to different Healthcare environments would have helped in improving the confidence in the generalizability and comprehensiveness of the model.

5.9 Implications for Personalized Medicine

The findings helped in gathering a broader understanding regarding the vision of personalized machine learning by highlighting how the employment of routine Healthcare administrative data can be transformed into highly effective and powerful predictive resources.

Aspect	Observation	Implication/Future Direction
Model Complexity	Simple models achieved strong performance.	Indicates that effective predictions may not require complex or costly data or algorithms.

		Encourages practical implementation in resource-constrained settings.
Task Readiness	Compliance prediction outperformed cancer prediction.	Demonstrates that some healthcare prediction tasks are more mature and clinically viable than others.
Implementation Potential	Compliance prediction is suitable for immediate clinical use.	Cancer prediction requires further research, threshold optimization, and validation before clinical deployment.
Strategic Integration	Study supports a tiered prediction strategy.	Suggests prioritizing deployment of well-performing models (e.g., compliance) while continuing to refine and validate more complex predictions (e.g., cancer risk).

Chapter 6 - Conclusion

6.1 Key Findings and Contributions

As a result of this thorough research, we have learned a great deal about predictive healthcare analytics related to artificial intelligence and mammogram screening in women. The analysis of 28,593 patient records from Johns Hopkins Aramco Healthcare gives evidence of the benefits and difficulties of machine learning in clinical screening situations. The models showed high performance in forecasting whether people would be screened by mammography, which nearly or completely correctly classified the cases. The level of accuracy for the logistic regression model was 99.6%, but both the random forest and neural network models reached 100%. The strong pattern found here indicates that compliance with appointments depends mainly on the schedule itself and other statistical factors, creating new health-intervention options.

It was more challenging to predict breast cancer since risk assessment in biology is naturally complex. While the cancer prediction model achieved strong overall performance (98.3% accuracy, 0.978 AUC-ROC), the 74.2% recall rate indicates that approximately one-quarter of actual cancer cases remain undetected. This performance differential illuminates fundamental differences between behavioral prediction tasks, which appear highly learnable from administrative data, and biological prediction tasks, which require more sophisticated approaches and additional data sources.

The feature importance analysis revealed critical insights into the mechanisms driving both compliance behavior and cancer risk. Compliance prediction was dominated by temporal and scheduling features, with overdue appointment status accounting for 43.7% of predictive power, while cancer prediction was driven primarily by demographic factors (41.9% contribution). This divergence suggests that effective healthcare interventions must employ different strategies for behavioral versus biological outcomes.

6.2 Clinical and Operational Implications

The study's findings have immediate practical applications for healthcare system optimization. The perfect predictive accuracy for compliance enables healthcare providers to implement precision-targeted intervention strategies, focusing limited resources on the approximately 72% of patients predicted to be non-compliant. This capability represents a paradigm shift from broad population-based outreach to individualized intervention strategies that maximize efficiency and effectiveness. For cancer prediction, the models' strong discrimination abilities (AUC-ROC values exceeding 0.970) provide valuable risk stratification capabilities, even with imperfect binary classification performance. The findings suggest that threshold optimization favoring sensitivity over specificity may be clinically warranted, accepting increased false positive rates to minimize the more costly false negative outcomes in cancer screening contexts.

The research demonstrates that sophisticated predictive capabilities can be achieved using routine administrative healthcare data without requiring expensive additional data collection efforts. This accessibility suggests that similar predictive modelling approaches could be implemented across diverse healthcare settings, potentially democratizing advanced healthcare analytics.

6.3 Methodological Innovations

The study's innovative approach to data quality challenges, particularly the transformation of missing appointment dates into predictive features, exemplifies creative solutions to common healthcare analytics problems. This methodology demonstrates that data incompleteness can carry meaningful information rather than representing mere quality issues, contributing 10.8% to compliance prediction performance.

The comprehensive feature engineering workflow, including temporal decomposition, clinical guideline integration, and behavioral pattern extraction, provides a replicable framework for transforming administrative healthcare data into meaningful predictive signals. The success of relatively simple models (particularly logistic regression) suggests that interpretability and computational efficiency need not be sacrificed for predictive performance in many healthcare applications.

6.4 Limitations and Future Directions

Despite the remarkable findings, several limitations warrant acknowledgement.

Aspect	**Observation**	**Implication/Future Direction**
Data Source Limitation	Relied on a single healthcare system (Johns Hopkins Aramco Healthcare).	Limits generalizability to different systems with varied populations, protocols, or administrative contexts. External validation is needed.

Compliance Prediction Performance	Perfect prediction results.	Impressive but may reflect overfitting; requires validation in diverse settings to ensure robustness and reliability.
Cancer Prediction Model – Recall Issues	Limited recall in the cancer prediction task.	Indicates need for improved feature engineering and additional data sources (e.g., genetic, lifestyle, imaging data).
Data Enhancement Strategies	The current model lacks multi-modal data.	Future research should integrate genomic, wearable, lifestyle, and imaging data to enhance predictive power while keeping models clinically usable.
Model Complexity vs. Performance	Simple models performed well.	Suggests effective predictions can be achieved without complex algorithms or expensive data collection, facilitating practical implementation.
Task Maturity Comparison	Compliance prediction is more mature than cancer prediction.	Compliance prediction is ready for deployment; cancer prediction needs more R&D and threshold optimization before clinical use.
Tiered Implementation Strategy	Varying readiness of prediction tasks.	Supports a phased approach: deploy ready models (e.g., compliance) and refine complex ones (e.g., cancer risk) for future use.
Equity and Bias Considerations	Models may reflect historical healthcare disparities.	Calls for ongoing bias monitoring and correction to ensure fair and equitable model deployment across diverse populations.

6.5 Broader Impact and Significance

This research contributes to the growing evidence base supporting the integration of artificial intelligence into routine healthcare delivery. However, the results suggest that it may be simpler than expected to put these analytics into use, allowing their adoption to increase sooner than expected in many health institutions. From the study findings, it appears that highly accurate prediction jobs like forecasting patient compliance can be promptly applied to enhance health care, and harder predictions should be gradually developed and carefully integrated into the clinical setting. Utilizing this technique, users receive benefits quickly while still allowing developers to increase capabilities in more difficult domains.

The findings show that artificial intelligence adds great strength to preventive healthcare, improving both processes and clinical outcomes. Using the collected data to generate useful forecasts, healthcare teams can respond to situations ahead of time and enhance both the results for patients and the usage of available resources. The results from using machine learning in mammogram screening provide a good example for using more predictive analytics in medicine, showing how to reach more efficiently, effectively and individually tailored healthcare.

References:

Agrawal, P., Chen, T. A., McNeill, L. H., Acquati, C., Connors, S. K., Nitturi, V., ... & Reitzel, L. R. (2021). Factors associated with breast cancer screening adherence among church-going African American women. *International Journal of Environmental Research and Public Health*, *18*(16), 8494. https://www.mdpi.com/1660-4601/18/16/8494

Alcazar-Bejerano, I. L. (2014). Health behaviours, disparities and deterring factors for breast cancer screening of immigrant women-a challenge to health care professionals. *Journal of Lifestyle Medicine*, *4*(1), 55. https://pmc.ncbi.nlm.nih.gov/articles/PMC4390760/

Allweis, T. M., Hermann, N., Berenstein-Molho, R., & Guindy, M. (2021). Personalised screening for breast cancer: Rationale, present practices, and future directions. *Annals of Surgical Oncology, 28*(8), 4306–4317. https://link.springer.com/article/10.1245/s10434-020-09426-1

AlSamhori, J. F., AlSamhori, A. R. F., Duncan, L. A., Qalajo, A., Alshahwan, H. F., Al-abbadi, M., ... & Nashwan, A. J. (2024). Artificial intelligence for breast cancer: Implications for diagnosis and management. *Journal of Medicine, Surgery, and Public Health*, *3*, 100120. https://www.sciencedirect.com/science/article/pii/S2949916X24000732

American Cancer Society. (2025). Key Statistics for Breast Cancer. https://www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html#:~:text=Breast%20cancer%20is%20the%20second,decline%20of%2044%25%20through%202022.

Arasu, V. A., Habel, L. A., Achacoso, N. S., Buist, D. S., Cord, J. B., Esserman, L. J., ... & Lee, C. (2023). Comparison of mammography AI algorithms with a clinical risk model for 5-year breast cancer risk prediction: An observational study. *Radiology, 307*(5), e222733. https://pubs.rsna.org/doi/abs/10.1148/radiol.222733

Arleo, E. K., Hendrick, R. E., Helvie, M. A., & Sickles, E. A. (2017). Comparison of recommendations for screening mammography using CISNET models. *Cancer, 123*(19), 3673–3680. https://doi.org/10.1002/cncr.30842

Balthazar, P., Harri, P., Prater, A., & Safdar, N. M. (2018). Protecting your patients' interests in the era of big data, artificial intelligence, and predictive analytics. *Journal of the American College of Radiology*, *15*(3), 580-586. https://www.sciencedirect.com/science/article/pii/S1546144017315995

Kim, H. E., Kim, H. H., Han, B. K., Kim, K. H., Han, K., Nam, H., ... & Kim, E. K. (2020). Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *The Lancet Digital Health*, *2*(3), e138-e148. https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30003-0/fulltext

Brooks, J. D., Nabi, H., Andrulis, I. L., Antoniou, A. C., Chiquette, J., Després, P., ... & Simard, J. (2021). Personalised risk assessment for prevention and early detection of breast cancer: Integration and implementation (PERSPECTIVE I&I). *Journal of Personalised Medicine, 11*(6), 511. https://www.mdpi.com/2075-4426/11/6/511

Carver, T., Hartley, S., Lee, A., Cunningham, A. P., Archer, S., Babb de Villiers, C., ... & Antoniou, A. C. (2021). CanRisk Tool—A web interface for the prediction of breast and ovarian cancer risk and the likelihood of carrying genetic pathogenic variants. *Cancer Epidemiology, Biomarkers & Prevention, 30*(3), 469–473. https://aacrjournals.org/cebp/article-abstract/30/3/469/264109

Corti, C., Cobanaj, M., Marian, F., Dee, E. C., Lloyd, M. R., Marcu, S., ... & Curigliano, G. (2022). Artificial intelligence for prediction of treatment outcomes in breast cancer: Systematic review of design, reporting standards, and bias. *Cancer Treatment Reviews, 108*, 102410. https://www.sciencedirect.com/science/article/pii/S0305737222000743

Dembrower, K., Crippa, A., Colón, E., Eklund, M., & Strand, F. (2023). Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. *The Lancet Digital Health*, *5*(10), e703-e711. https://www.thelancet.com/journals/landig/article/PIIS2589-7500(23)00153-X/fulltext

Dembrower, K., Wåhlin, E., Liu, Y., Salim, M., Smith, K., Lindholm, P., ... & Strand, F. (2020). Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *The Lancet Digital Health*, *2*(9), e468-e474. https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30185-0/fulltext

Dervovic, D., Lécué, F., Marchesotti, N. and Magazzeni, D., 2024. Are logistic models really interpretable?. arXiv preprint arXiv:2406.13427.

Ferreira, C. S., Rodrigues, J., Moreira, S., Ribeiro, F., & Longatto-Filho, A. (2021). Breast cancer screening adherence rates and barriers of implementation in ethnic, cultural, and religious minorities: a systematic review. *Molecular and clinical oncology*, *15*(1), 139. https://www.spandidos-publications.com/10.3892/mco.2021.2301

Ghantasala, G. P., Hung, B. T., Chakrabarti, P., & Pellakuri, V. (2024). Artificial intelligence-based machine learning algorithm for prediction of cancer in female anatomy. *Multimedia Tools and Applications*, 1-27. https://link.springer.com/article/10.1007/s11042-024-19655-1

Giaquinto, A. N., Sung, H., Miller, K. D., Kramer, J. L., Newman, L. A., Minihan, A., ... & Siegel, R. L. (2022). Breast cancer statistics, 2022. *CA: a cancer journal for clinicians*, *72*(6), 524-541. https://www.researchgate.net/publication/364149216_Breast_Cancer_Statistics_2022#:~:text=Breast%20cancer%20incidence%20rates%20have,and%20hormone%20receptor%E2%80%90positive%20disease.

Hardesty, L. A., Lind, K. E., & Gutierrez, E. J. (2016). Compliance with screening mammography guidelines after a false-positive mammogram. *Journal of the American College of Radiology*, *13*(9), 1032-1038. https://www.sciencedirect.com/science/article/pii/S1546144016301077

Houghton, S. C., & Hankinson, S. E. (2021). Cancer progress and priorities: Breast cancer. *Cancer Epidemiology, Biomarkers & Prevention, 30*(5), 822–844. https://aacrjournals.org/cebp/article-abstract/30/5/822/670821

Izmirlian, G., 2004. Application of the random forest classification algorithm to a SELDI-TOF proteomics study in the setting of a cancer prevention trial. Annals of the New York Academy of Sciences, 1020(1), pp.154-174.

Kerlikowske, K., Chen, S., Golmakani, M. K., Sprague, B. L., Tice, J. A., Tosteson, A. N., ... & Miglioretti, D. L. (2022). A cumulative advanced breast cancer risk prediction model was developed in a screening mammography population. *JNCI: Journal of the National Cancer Institute, 114*(5), 676–685. https://academic.oup.com/jnci/article-abstract/114/5/676/6506541

Keshavarzi, A., Asadi, S., Asadollahi, A., Mohammadkhah, F., & Khani Jeihooni, A. (2022). Tendency to breast cancer screening among rural women in Southern Iran: A structural equation modelling (SEM) analysis of the theory of planned behaviour. *Breast Cancer: Basic and Clinical Research, 16*, 11782234221121001. https://journals.sagepub.com/doi/abs/10.1177/11782234221121001

Khalid, A., Mehmood, A., Alabrah, A., Alkhamees, B. F., Amin, F., AlSalman, H., & Choi, G. S. (2023). Breast cancer detection and prevention using machine learning. *Diagnostics, 13*(19), 3113.

Khan, S. A., Hernandez-Villafuerte, K. V., Muchadeyi, M. T., & Schlander, M. (2021). Cost-effectiveness of risk-based breast cancer screening: A systematic review. *International Journal of Cancer, 149*(4), 790–810.

Khurshid, F., Zia, I., Ayesha, U. A., Yaqoob, F., Khurshid, H., Zia, A., & Khurshid, A. (2023). The Significance of Screening Mammography: A Preliminary Study. *Ann Jinnah Sindh Med Uni*, *9*(2), 42-48. https://www.researchgate.net/publication/377600072_The_Significance_of_Screening_Mammography_A_Preliminary_Study

Løberg, M., Lousdal, M. L., Bretthauer, M., & Kalager, M. (2015). Benefits and harms of mammography screening. *Breast cancer research*, *17*, 1-12. https://link.springer.com/article/10.1186/s13058-015-0525-z

Mahesh, T. R., Vinoth Kumar, V., Vivek, V., Karthick Raghunath, K. M., & Sindhu Madhuri, G. (2024). Early predictive model for breast cancer classification using blended ensemble learning. *International Journal of System Assurance Engineering and Management, 15*(1), 188–197. https://link.springer.com/article/10.1007/s13198-022-01696-0

Olalekan Kehinde, A., 2025. Leveraging Machine Learning for Predictive Models in Healthcare to Enhance Patient Outcome Management. Int Res J Mod Eng Technol Sci, 7(1), p.1465.

Prinzi, F., Insalaco, M., Orlando, A., Gaglio, S., & Vitabile, S. (2024). A YOLO-based model for breast cancer detection in mammograms. *Cognitive Computation, 16*(1), 107–120.

Rahman, S. M., Dignan, M. B., & Shelton, B. J. (2003). Factors influencing adherence to guidelines for screening mammography among women aged 40 years and older. *Ethnicity & disease*, *13*(4), 477. https://pmc.ncbi.nlm.nih.gov/articles/PMC2848385/

Reece, J. C., Neal, E. F., Nguyen, P., McIntosh, J. G., & Emery, J. D. (2021). Delayed or failure to follow-up abnormal breast cancer screening mammograms in primary care: A systematic review. *BMC Cancer, 21*, 1–14. https://link.springer.com/article/10.1186/s12885-021-08100-3

Ritchie, D., Van den Broucke, S., & Van Hal, G. (2021). The health belief model and theory of planned behaviour applied to mammography screening: A systematic review and meta-analysis. *Public Health Nursing, 38*(3), 482–492. https://onlinelibrary.wiley.com/doi/abs/10.1111/phn.12842

Rodríguez-Ruiz, A., Krupinski, E., Mordang, J. J., Schilling, K., Heywang-Köbrunner, S. H., Sechopoulos, I., & Mann, R. M. (2019). Detection of breast cancer with mammography: effect of an artificial intelligence support system. *Radiology*, *290*(2), 305-314. https://pubs.rsna.org/doi/abs/10.1148/radiol.2018181371

Rodríguez-Ruiz, A., Krupinski, E., Mordang, J.J., Schilling, K., Heywang-Köbrunner, S.H., Sechopoulos, I. and Mann, R.M., 2019. Detection of breast cancer with mammography: effect of an artificial intelligence support system. Radiology, 290(2), pp.305-314.

Schaffter, T., Buist, D. S., Lee, C. I., Nikulin, Y., Ribli, D., Guan, Y., ... & DM DREAM Consortium. (2020). Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA network open*, *3*(3), e200265-e200265. https://jamanetwork.com/journals/jamanetworkopen/article-abstract/2761795

Sterlingova, T., Nylander, E., Almqvist, L., & Christensen, B. M. (2023). Factors affecting women's participation in mammography screening in Nordic countries: A systematic review. *Radiography*, *29*(5), 878-885. https://www.sciencedirect.com/science/article/pii/S1078817423001360

Tanveer, H., Faheem, M., Khan, A. H., & Adam, M. A. (2025). AI-powered diagnosis: A machine learning approach to early detection of breast cancer. *International Journal of Engineering Development and Research, 13*(2), 153–166. https://rjwave.org/ijedr/viewpaperforall.php?paper=IJEDR2502021

Tavakoli, B., Feizi, A., Zamani-Alavijeh, F., & Shahnazi, H. (2024). Factors influencing breast cancer screening practices among women worldwide: a systematic review of observational and qualitative studies. *BMC Women's Health*, *24*(1), 268. https://link.springer.com/article/10.1186/s12905-024-03096-x

Thacker, B.H., Doebling, S.W., Hemez, F.M., Anderson, M.C., Pepin, J.E. and Rodriguez, E.A., 2004. Concepts of model verification and validation

Vyas, A., Madhavan, S., LeMasters, T., Atkins, E., Gainor, S., Kennedy, S., ... & Remick, S. (2012). Factors influencing adherence to mammography screening guidelines in Appalachian women participating in a mobile mammography program. *Journal of Community Health*, *37*, 632-646. https://link.springer.com/article/10.1007/s10900-011-9494-z

World Health Organization. (2024). Breast cancer**.** https://www.who.int/news-room/fact-sheets/detail/breast-cancer

Yala, A., Mikhael, P. G., Strand, F., Lin, G., Satuluru, S., Kim, T., ... & Barzilay, R. (2022). Multi-institutional validation of a mammography-based breast cancer risk model. *Journal* *of Clinical Oncology, 40*(16), 1732–1740. https://ascopubs.org/doi/abs/10.1200/JCO.21.01337

Yala, A., Schuster, T., Miles, R., Barzilay, R., & Lehman, C. (2019). A deep learning model to triage screening mammograms: a simulation study. *Radiology*, *293*(1), 38-46. https://jamanetwork.com/journals/jamanetworkopen/article-abstract/2761795

Zapka, J. G., Stoddard, A., Maul, L., & Costanza, M. E. (1991). Interval adherence to mammography screening guidelines. *Medical Care*, 697-707. https://www.jstor.org/stable/3766098