Counterfactual Prescriptions via Hierarchical ML for Missed Chemotherapy Appointment Prevention

Mohammadreza Rajabzadeh

A Thesis submitted to the Faculty of Graduate Studies
in partial fulfillment of the requirements for the degree of

Master of Arts

in Information Systems & Technology
York University, Toronto, Ontario

October 2025

Abstract

Missed medical appointments, including cancellations and no-shows, disrupt clinical workflows, reduce efficiency, and compromise patient care. Using 1.8 million chemotherapy appointments from the Dana-Farber Cancer Institute, this study develops a hierarchical machine learning framework that first predicts cancellations and then no-shows among remaining cases. Combining operational and temporal features, the model achieves F1-scores of 0.76 and 0.82, outperforming a multinomial baseline by 7–10 points on minority classes. To extend interpretability, we infer missed-visit reasons through semi-supervised learning on short-notice cancellations, training supervised predictors with weighted F1-scores of 0.57 and 0.54. Counterfactual analysis identifies modifiable scheduling factors—such as appointment timing and provider consistency—while showing that standard interventions like reminders are less effective than previously reported. These findings underscore the value of predictive models with prescriptive intent, highlighting their role in designing tailored, context-specific interventions to improve clinical efficiency and patient care.

Keywords: Hierarchical Machine Learning; Counterfactual Analysis; Predictive and Prescriptive Analytics; Patient No-Shows; Appointment Cancellations

Acknowledgements

I wish to express my deepest gratitude to Professor Arik Senderovich and Professor Adam Diamant for their invaluable guidance, insightful feedback, and sustained encouragement throughout this research. Their expertise and support were instrumental to the completion of this thesis. Finally, I dedicate this work to my family, whose love and encouragement sustained me throughout my studies.

Contents

List of Tables

List of Figures

List of Abbreviations

ML	Machine Learning
IS	Information Systems
XGBoost	Extreme Gradient Boosting
DT	Decision Tree
F1-score	Harmonic mean of precision and recall

Chapter 1

Introduction

Medical appointments in outpatient services that are scheduled but not attended, whether due to no-shows (without prior notice) or cancellations (with advance notice), create substantial inefficiencies, financial losses, and reduced access to care within healthcare systems [1]–[3]. These disruptions strain clinic operations and may delay critical treatments in high-risk care settings such as oncology, where treatment delays are associated with increased mortality [4]. In the UK, primary care facilities report over 12 million missed appointments annually (approximately 6.5% of all bookings), with the NHS estimating the annual cost at more than £600 million [5]. In the U.S., no-shows are reported to cost the healthcare sector around $150 billion annually, with each unused slot averaging $200 in lost physician revenue [6]. Missed appointments are common (typically 15–30% of scheduled visits and averaging ∼23%) and can disrupt workflow, prolong patient wait times, and weaken continuity of care [3], [5], [7].

Although their definitions can vary, a no-show is typically defined as a missed appointment or one canceled within 24 hours of the scheduled time, whereas a cancellation refers to an appointment that is canceled more than 24 hours in advance [3]. The distinction lies in (i) when the clinic receives notice; and (ii) how it can respond operationally, that is, whether there is sufficient time to reallocate the newly vacant slot to another patient. While the consequences of missed appointments are significant across all areas of medicine, they are particularly critical in the delivery of chemotherapy. Treatments are scheduled at precise intervals, and disruptions from cancellations or no-shows can compromise therapeutic efficacy, increase toxicity risks, and lead to poorer survival outcomes [4]. Furthermore, from a management perspective, chemotherapy appointments require specialized hospital resources such as infusion chairs and nursing staff, making no-shows and cancellations especially disruptive to effective resource utilization and system-wide efficiency [8], [9].

To combat this issue, researchers have developed predictive models to identify patients at risk of missing an appointment. Early approaches use logistic regression combined with

Bayesian inference to estimate no-show probabilities in near-real time [10], while later work extends this methodology using multinomial logistic regression to jointly predict no-shows and cancellations [11]. More recently, modern machine learning (ML) methods such as XGBoost have demonstrated strong predictive capacity – for example, achieving an AUC of 0.90 in predicting attendance for pediatric ophthalmology follow-up visits [12]. A systematic review confirms that a wide range of ML models, including logistic regression, decision trees, and ensemble methods, now appear across many clinical contexts with varying degrees of effectiveness [13]. Despite these advances, several gaps remain.

Most prior work either conflates cancellations and no-shows or focuses on predicting only one outcome. Operationally, however, they differ: advance cancellations can often be rebooked, while no-shows leave the clinic little opportunity to recover capacity [14]. Collapsing them into a single "non-attendance" category or ignoring one outcome is practically limiting as both are operationally and clinically consequential. Furthermore, developing useful prediction models with the limited patient information available to administrative staff is challenging due to severe class imbalance: most appointments are attended, leading models to underpredict missed visits. While imbalance-correction techniques – such as oversampling, undersampling, and cost-sensitive learning – can improve detection sensitivity, studies show that these methods often fail to improve discrimination and may impair model calibration and reliability, reducing clinical usefulness [15], [16].

Another challenge lies in the selection of predictors. Static features, such as demographic variables or distance to the clinic, capture who is generally at risk, whereas temporal features – such as the time until the next appointment, time-of-day, and the patient's cumulative history of no-shows and cancellations – capture evolving risk over time and refine personalized risk scores as new data accumulate from repeated interactions with the clinic. Prior studies often emphasize one set of variables over the other, focusing either on static models [17], [18] or emphasizing temporal variables [11], [19]. Understanding how these feature sets interact can inform targeted interventions, such as personalized reminders for time-of-day risks or adjustments to the lead time (i.e., the time interval between successive appointments) for patients with specific risk profiles [20], [21]. However, this comparative perspective remains under-explored.

Equally important, existing models rarely consider the underlying reasons for no-shows and cancellations. That is, most predictive frameworks stop at forecasting whether an appointment will be missed, offering little insight as to *why*. Prior studies have identified potential factors through descriptive analyses [21], while patient surveys highlight general causes such as forgetfulness or competing demands, situational barriers like a lack of transportation or childcare, and structural issues including long lead times or inadequate reminders [22],

[23]. Other literature discusses recurrent determinants, including prior attendance history, sociodemographic characteristics, and clinic-level processes [3]. Although these findings provide valuable contextual information, existing machine learning models do not explain the reasoning behind their predictions [3], [24]. This creates a gap between a general but descriptive knowledge of causes and the ability to design actionable, personalized interventions. Embedding a reason-prediction model into the ML pipeline would help overcome this limitation by not only flagging patients at risk of missing their scheduled appointment, but would also predict the likely underlying cause, enabling targeted interventions.

Finally, prior work typically stops short of translating predictions into actionable strategies for process improvement. While interpretability methods – like feature importance and SHapley Additive exPlanations (SHAP) – can identify dominant factors contributing to missed visits [25], counterfactual analyses can extend this by demonstrating how outcomes might change under new operational protocols, such as shifting appointments to earlier hours [26]. Despite this potential, few studies use these methods to evaluate interventions before their implementation, even though effectiveness can vary widely across clinical environments, patient populations, and resource constraints [3]. In fact, without prospective evaluation, common interventions (such as appointment reminders) may prove ineffective for a given clinical context, and eliciting improvements in some variables may be too costly to yield meaningful gains. Therefore, moving from prediction to prescription is essential to connect ML modeling forecasts with initiatives that improve operational performance and patient outcomes.

To address these limitations, we develop an ML pipeline that jointly predicts appointment cancellations and no-shows – along with their underlying reasons – using a hierarchical approach. In the first stage, the model identifies whether a patient is at risk of canceling their appointment; in the second stage, it predicts the likelihood of a no-show. If either stage predicts a missed appointment, the framework also provides a reason as we train a multi-class classifier across five managerially distinct categories using semi-supervised learning with last-minute cancellations to infer reasons for no-shows (which are absent). Utilizing a dataset of 1,825,975 scheduled chemotherapy appointments from 95,282 patients from the Dana-Farber Cancer Institute, a major U.S. cancer center in Boston, MA, the pipeline incorporates both static and temporal features, enabling it to capture nuanced patterns in patient behavior and appointment logistics over time. Despite severe class imbalances and being limited to fields available in the scheduling system (i.e., information accessible to schedulers and care coordinators at booking or rescheduling), the ML pipeline achieves F1-scores of 0.76 for cancellations and 0.82 for no-shows—representing a 7–10 percentage point improvement on minority classes as compared to a single-stage multinomial baseline. Further, incorporating both static and temporal variables improves minority-class performance by up to 12% over models using either feature type alone. Finally, our approach achieves weighted F1-scores of 0.57 and 0.54 when predicting reasons for cancellations and no-shows, respectively.

To support clinical utility, we conduct post-hoc analyses using a feature importance assessment, SHAP, and counterfactual simulations. Feature importance analysis and SHAP consistently identify encounter type, visit type, lead time, provider changes, and prior visit intervals as dominant predictors. These findings highlight both structural factors (e.g., type of clinical service, treatment-related visits) and scheduling dynamics (e.g., time since last visit, lead time) as key determinants of attendance. Building on these insights, we utilize our predictive model to estimate the effect of targeted process improvement initiatives on the rate of no-shows and cancellations. That is, by adjusting feature values to reflect potential operational or behavioral interventions that can be implemented by the clinic, we can evaluate the resulting changes in cancellation and no-show rates. We evaluate both literature-informed strategies (e.g., SMS reminders, reduced scheduling lead times) as well as novel interventions (e.g., provider consistency, a novel patient confirmation strategy) enabled by the structure and output of our ML pipeline.

The results indicate that the literature-informed interventions provide little to no benefit: appointment reminders and lead-time adjustments have minimal impact on reducing cancellations and may even increase the number of no-shows. In contrast, model-guided strategies deliver measurable improvements, particularly in lowering cancellations. For instance, enhancing provider consistency reduces cancellations by more than 2% (although it slightly increases no-shows) while our patient confirmation strategy decreases both cancellations and no-shows. These findings highlight the importance of context-specific evaluation and the value of predictive models in guiding prescriptive strategies.

1.1 Contributions of this Thesis

The main contributions of this research are summarized as follows:

- Develop a data-aware hierarchical framework to jointly predict chemotherapy appointment cancellations and no-shows, achieving F1-scores of 0.76 (cancellations) and 0.82 (no-shows), a 7–10 point improvement over a multinomial baseline.
- Engineer both temporal and static features to improve minority-class F1-scores by up to 12% as compared to using either feature-type alone.
- Apply semi-supervised learning to infer missing no-show reason labels from late cancellations, enabling supervised prediction of reasons for both cancellations and no-shows,

with weighted F1-scores of 0.57 and 0.54, respectively.

- Conduct feature importance analyses and counterfactual simulations to identify actionable scheduling interventions, including those emphasized in the literature and those unique to our modeling approach.
- Present a data-driven, prescriptive framework for assessing the effectiveness of strategies to improve patient attendance and resource optimization in outpatient health programs.

The remainder of this thesis is organized as follows: Chapter 2 reviews related work; Chapter 3 presents the dataset, feature construction, and methodology; Chapter 4 presents results of the prediction modeling; Chapter 5 presents the counterfactual analysis; Chapter 6 discusses implications and limitations.

Chapter 2

Literature Review

2.1 Factors Influencing Patient No-Shows and Cancellations

Understanding patient no-shows and cancellations requires a structured analysis of contributing factors, as prior studies have used diverse sets of predictors. To provide a comprehensive foundation to our study, we conducted a systematic review of 116 peer-reviewed articles indexed in Scopus, the largest repository of academic literature. The search spanned publications from January 1980 to December 2024 and used keywords such as "*no-shows*," "*missed appointments*," "*non-attendance*," and "*cancellation*," applied to titles, abstracts, and keywords, with spelling variations and wildcard operators included for coverage. Relevant articles were screened, and thematic analysis grouped predictors into categories (see Table 2.1).

The review identified four categories of predictors: *Patient Demographics* (e.g., age, insurance type), *Medical History* (e.g., comorbidities, diagnostic information), *Appointment Details* (e.g., visit type, lead time), and *Patient Behavior* (e.g., prior no-show history, cancellation frequency). These categories are generally consistent with prior reviews [3], [13] and offer a clinically interpretable framework for understanding attendance risk. A detailed breakdown of these predictors and associated references is provided in Table 2.1.

The most frequently reported predictors of missed appointments include static features such as age, insurance, or payment method (*Patient Demographics*), diagnosis severity (*Medical History*), and temporal predictors such as a history of no-shows and cancellation frequency (*Patient Behavior*). Appointment-related details like lead time and the type of visit also represent important inputs. Previous studies typically incorporate both static and dynamic features, though the balance between them varies. For example, Daggy et al. [17] and Huang and Hanauer [18] use logistic regression and tree-based models primarily based on demographic (static) and appointment-related (temporal) variables, while Goffman et al. [27] emphasize features based on demographics and a patients medical and behavioral data (temporal). Other studies explicitly engineer temporal information from patients' appointment sequences: Huang et al. [28] demonstrate that prior attendance history improves model sensitivity, while Lenzi et al. [19] and Li et al. [29] architect temporal features to enhance no-show detection. Collectively, this body of work underscores the important roles that static and temporal features have in improving predictive performance. However, few have directly evaluated the extent to which each category contributes to predictive accuracy or how their interaction affects performance. Without this insight, models may miss key risk interactions that evolve over time. Furthermore, cancellations are often excluded from these models, despite substantial overlap in risk factors and their unique operational consequences.

Table 2.1: Grouped factors affecting no-shows and cancellations

Category	**Factors**	**References**
Patient Demographic	Age	[17], [30]–[38]
	Gender	[32]–[34], [39]–[42]
	Language	[18], [28], [33], [36], [42]–[44]
	Race/Ethnicity	[33], [34], [40], [45]–[48]
	Employment	[44], [49]–[52]
	Marital Status	[17], [32], [33], [48], [51], [53]–[56]
	Economic Status	[36], [38], [40], [45], [57]–[59]
	Education Level	[35], [38], [52], [55], [59]–[62]
	Insurance/Payment	[18], [31], [32], [34], [41], [46], [58], [61], [63], [64]
	ZIP Code	[11], [65]–[67]
	Distance/Transportation	[17], [18], [46], [53], [68]–[70]
	Religion	[28], [51], [56]
	Access to Phone	[71]–[75]
	Citizenship	[44], [63], [76]–[78]
	Telephone number Recorded	[38], [72], [79], [80]
	Residence (Urban/Rural)	[30], [48], [61], [81], [82]
	Season of Birth	[76], [83]
	Speciality	[65], [66], [84]–[86]
Medical History	Previous Visit	[65], [84], [86]–[89]
	Provider	[19], [29], [33], [46], [90]–[94]
	Referral Source	[44], [81], [95]–[100]
	Diagnosis	[22], [30], [69], [90], [98], [99], [101]–[105]
	Case Duration	[65], [67], [74], [86], [106]
	First/Follow-Up	[66], [74], [75], [84], [85], [107]

Continued on next page

Table 2.1 – *Continued from previous page*

Category	**Factors**	**References**
	Substance Abuse	[17], [35], [36], [44], [46], [52], [59], [103], [108]
	Symptoms	[22], [30], [62], [63], [81], [90], [109], [110]
	Use of Medication	[35], [54], [58], [61], [62]
Appointment Detail	Month	[18], [28], [39], [79], [92], [111]–[114]
	Weekday	[19], [75], [86], [115], [116]
	Visit Time	[31], [65], [73], [75], [90], [117]
	Holiday Indicator	[10], [11], [67], [87]
	Same Day Visit	[19], [27], [67]
	Weather	[29], [32], [64], [111], [118], [119]
	Season	[17], [33], [34], [48], [91], [120]
	Visit Interval	[27], [65], [75], [86], [121]
	Lead Time	[17], [18], [31]–[33], [69], [93], [95], [97], [120], [122]
	Waiting Time	[19], [67], [123]
	Scheduling Mode	[22], [70], [71], [88], [90], [106], [121], [124]
	Hospital Admissions	[17], [22], [72], [90], [102]
	Clinic Characteristics	[10], [22], [30], [66], [67], [86], [90], [125]
	Type of Procedure	[30], [36], [90], [114], [126]–[128]
Patient Behaviour	Previous No-Show	[22], [29], [31], [60], [75], [90], [107], [121], [129], [130]
	Previous Cancel	[22], [27], [29], [30], [65], [90]
	Last Visit Status	[67], [70], [86], [125]
	Visit Late	[49], [67]
	Satisfaction	[49], [106]
	Number of Previous Scheduling Visit	[17], [36], [94], [131]
	Family Support	[61], [132]
	Days Since Last Visit Appointment	[17], [131]

2.2 Predicting Missing Appointments in Oncology

While many studies have examined various predictors of patient attendance in outpatient healthcare settings, relatively few studies focus on oncology. Chemotherapy appointments pose unique operational challenges because they have long and variable durations, require patients to be scheduled at regular intervals over extended periods of time (e.g., weeks), depend on specialized infusion chairs, and require highly trained nursing staff [8], [9]. Missed appointments or last-minute cancellations in this context are especially disruptive, leading to wasted resources and potential delays in life-saving treatment. Evidence from cancer care research shows that even modest delays in therapy can adversely affect survival outcomes [4].

Most existing studies in this domain have concentrated on optimizing scheduling templates and resource allocation under uncertainty [8], [9], rather than predicting the patient-level risk of no-shows or cancellations. Furthermore, our work is unique in that it combines predictive modeling with counterfactual analysis to determine the effect that targeted interventions have on reducing missed appointments for chemotherapy treatment.

2.3 Methodological Challenges in Training Prediction Models

Early efforts to predict patient attendance primarily used statistical techniques such as logistic regression and survival analysis, relying on demographic characteristics and variables on the appointment-type [17], [27]. With the growing availability of electronic health records, ML methods – particularly tree-based approaches such as random forests and gradient boosting – have demonstrated improved predictive performance [18], [33], [121]. Recently, several studies have explored the integration of temporal variables [19], [29] and unstructured text data from clinical notes [133], though such approaches remain less common in the prediction of missed appointments.

Despite these advances, methodological challenges persist. First, most studies focus exclusively on no-shows, with cancellations either excluded or grouped into a general "missed appointment" category [134], [135]. Second, appointment can be highly imbalanced: attended visits dominate, with no-shows typically accounting for $<$20% of data points, leading standard algorithms to favor the majority class and under-predict missed visits [85], [133]. Third, reliable labels can be scarce in electronic health record systems; in our setting, reasons for no-shows are rarely documented due to a lack of follow-up by administrative staff. Hierarchical classification offers one such solution by decomposing a multi-class prediction task into sequential sub-problems, reducing imbalance at each decision point and improving minority-class detection. Zheng and Zhao [136] introduce a cost-sensitive hierarchical method that incorporates an imbalance-aware loss function, significantly improving performance on skewed datasets. Pereira et al. [137] further show that resampling strategies tailored to hierarchical structures can outperform single-stage models. We develop a hierarchical classification pipeline to separately predict cancellation and no-show risks. To address the missing labels for no-show reasons, we employ semi-supervised learning (SSL), which leverages large volumes of unlabeled data alongside limited labeled examples. SSL has shown strong results in domains such as cancer diagnostics [138], and we extend its use to infer reasons for missed appointments by exploiting a proxy signal in the reasons behind short-notice ($\leq$24 hrs) cancellations.

2.4 Interventions to Reduce No-Shows and Cancellations

Building on predictive modeling research, prior work has explored operational and behavioral interventions to reduce missed appointments. These fall into three main categories: *reminders*, *scheduling strategies*, and *incentives*. Reminders – via SMS, phone, or multi-channel approaches – are the most common and show consistent effectiveness, particularly when utilized with risk prediction models [139], [140]. Scheduling interventions, such as open-access booking, overbooking, and real-time rescheduling, aim to optimize capacity and proactively respond to cancellations and no-shows [141], [142]. Incentives, though less studied, address structural barriers (e.g., transport vouchers) or apply penalties to deter non-attendance [143], [144]. Despite these advances, most interventions emphasize patient-focused actions, with limited attention to provider engagement or continuity of care – factors shown to influence attendance in clinics with multiple appointments [145]. Moreover, evidence suggests that some of the most common interventions, i.e., automated reminders, achieve only modest reductions in no-show rates and fail to address underlying systemic issues [142].

While interventions are often evaluated retrospectively or tested in single-institution studies, few attempts have been made to assess their expected impact prior to deployment. That is, they are typically assumed to be broadly effective even though their efficacy can vary significantly with the operational context and patient population [3]. To our knowledge, no prior work has systematically evaluated the anticipated effects of a wide range of clinic interventions on the rate of no-shows and cancellations. We address this gap through a counterfactual analysis based on our hierarchical ML model, which estimates the likely effect of implementing literature-informed interventions alongside several new operational approaches derived from our modeling framework.

Chapter 3

Data and Methodology

3.1 Operational Setting & Data

Our analysis focuses on outpatient chemotherapy appointments at the Dana-Farber Cancer Institute, a comprehensive cancer treatment and research center affiliated with the Harvard Medical School, where patients receive treatments at precisely scheduled intervals. Each visit requires advance coordination of infusion chairs, oncology-trained nursing staff, and pharmacy preparation, making reliable attendance essential for both patient care and operational planning [8], [9], [146]. Unanticipated no-shows or late cancellations disrupt these tightly managed workflows, limit opportunities to reallocate capacity, and reduce overall clinic efficiency [147]. Furthermore, missing or rescheduling a chemotherapy sessions is not merely inconvenient: even modest delays in the prescribed treatment protocol have been associated with an increased risk of mortality [4].

At the time of scheduling, there are three possible appointment outcomes: completed visits (patients attend their scheduled appointment), no-shows (including late cancellations within 24 hours and those who left without being seen), and cancellations made at least 24 hours in advance of the next appointment. Distinguishing between these outcomes is important, as each has different clinical and operational consequences.

3.1.1 Dataset

The dataset comprises daily records of chemotherapy appointments from June 12, 2021, to December 14, 2023. While the initial data cut contained 4,560,061 records for all patients and visit types, only appointments related to chemotherapy treatment were retained; patients with only one recorded appointment, incomplete records, and duplicate entries were excluded. The final cohort includes 1,825,948 appointments from 95,282 unique patients who were either

Table 3.1: Dataset features and descriptions

Feature	Description	Data Type
Encounter Date	Date of the patient encounter	Date
Patient ID	Patient identifier	Labeling
Encounter Datetime	Date and time of the patient encounter	Datetime
Encounter ID	Unique encounter identifier	Labeling
Arrived Datetime	Date and time of patient arrival	Datetime
Visit Type	Type of visit (e.g., Blood Draw, Treatment Chemo)	Factor
Encounter Type	Blood Draw, Infusion, Appointment, etc.	Factor
Encounter Provider ID	Provider identifier for the encounter	Labeling
Attending Provider ID	Attending provider identifier	Labeling
Appointment provider	Appointment provider identifier	Labeling
Department	Name of the clinical department	Factor
Floor Location	Name of the floor location	Factor
StD Duration	Standard duration of the appointment (minutes)	Numeric
Status	Completion, no-show, cancellation indicator	Factor
Scheduled Datetime	Date and time of appointment scheduling	Datetime
Canceled Datetime	Date of appointment cancellation	Datetime
Rescheduled Datetime	Date and time of appointment rescheduling	Datetime
Reason Description	Description of the reason for cancellation	Factor

undergoing chemotherapy, receiving cancer care, or attending monitoring visits. All data fields originate from the operational scheduling system, reflecting information available to appointment schedulers and clinical care coordinators at the time of booking or rescheduling. Practitioner notes and other unstructured data sources are not accessible in this database, which is consistent with data available in real-world scheduling workflows.

Table 3.1 summarizes the available features that were collected. The dataset includes structured variables such as appointment details, visit type, and provider information, but lacks demographic and socioeconomic indicators (e.g., age, gender, language, race/ethnicity) previously associated with attendance patterns [17], [32]–[38]. Although the absence of demographic and socioeconomic features constitutes a limitation, it also underscores the strength of our design: if accurate predictions can be achieved using only operational scheduling variables, augmenting the model with additional features derived from demographic data or unstructured clinical text should be expected to further enhance performance. Reason labels were available only for cancellations (not no-shows) due to the lack of follow-up by scheduling staff. These reasons were consolidated into clinically meaningful categories informed by prior literature on missed appointments and consultation with clinical collaborators. This process yielded five categories: *patient choice*, *clinic management*, *patient health*, *socioeconomic barriers*, and *provider-related factors* (see 3.2). This structured categorization provided the framework for our analysis of both outcome prediction and reason inference.

The following summary statistics describe the dataset we collected. No-shows comprise

Table 3.2: Structured reason categories for missed appointments

Category	Explanation	Example Reasons	Citations
Patient Choice	Patient voluntarily cancels due to personal preferences or conflicting priorities.	"Personal Reasons", "Sought Care Elsewhere"	[31], [148]
Clinic Management	Appointment canceled due to administrative rescheduling or internal errors.	"Schedule Order Error", "Changed by Radiology"	[31]
Patient Health	Medical or clinical conditions prevent the patient from attending.	"Hospitalized", "Clinically Caused"	[59], [100]
Socioeconomic Factors	Attendance prevented by external social or economic barriers.	"Lack of Transportation", "Financial Concerns"	[149]
Provider Choice	Appointment missed due to provider-related scheduling issues.	"Provider Departure", "MD Appointment"	[31], [148]

5.7% of scheduled appointments, while cancellations (≥24 hours in advance) account for 29.8%. Appointment histories are highly skewed: most patients have few visits, while a minority accumulate many (Figure 3.1). Patients average 19.16 visits (range 2–520). Excluding those with only one visit yields a final cohort of 1,825,948 appointments. On average, patients record 1.08 no-shows (affecting 35% of patients) and 5.71 cancellations (affecting 81%).

Several derived features were computed from the raw data to create temporal features that capture both cumulative and recent attendance behavior (Table 3.3). Cumulative variables represent long-term patterns (e.g., historical no-show and cancellation rates), while recent variables capture short-term trends (e.g., missed last appointment). Categorical variables were one-hot encoded, and numerical variables were standardized. Missing values in numeric variables were imputed using the median, and missing categorical values were assigned an "Unknown" category.

3.1.2 Data Partitioning

The ML models were trained and evaluated using an event-based design: at the time an appointment is scheduled or rescheduled, the model predicts whether it will be completed, canceled, or result in a no-show. If the outcome is a no-show or cancellation, the framework also predicts the most likely reason. To assess robustness, we implemented three complementary data-splitting strategies. As a baseline, we used a row-wise stratified random splitting approach, as it is the simplest procedure to implement. However, this strategy can overestimate performance by allowing the same patient to appear in both training and test sets [150]. Thus, we also implement a patient-level split which ensures all appointments for a given patient appear in one partition, enforcing independence across individuals. Finally, we analyze an appointment–time-based temporal split where models are trained on earlier appointments

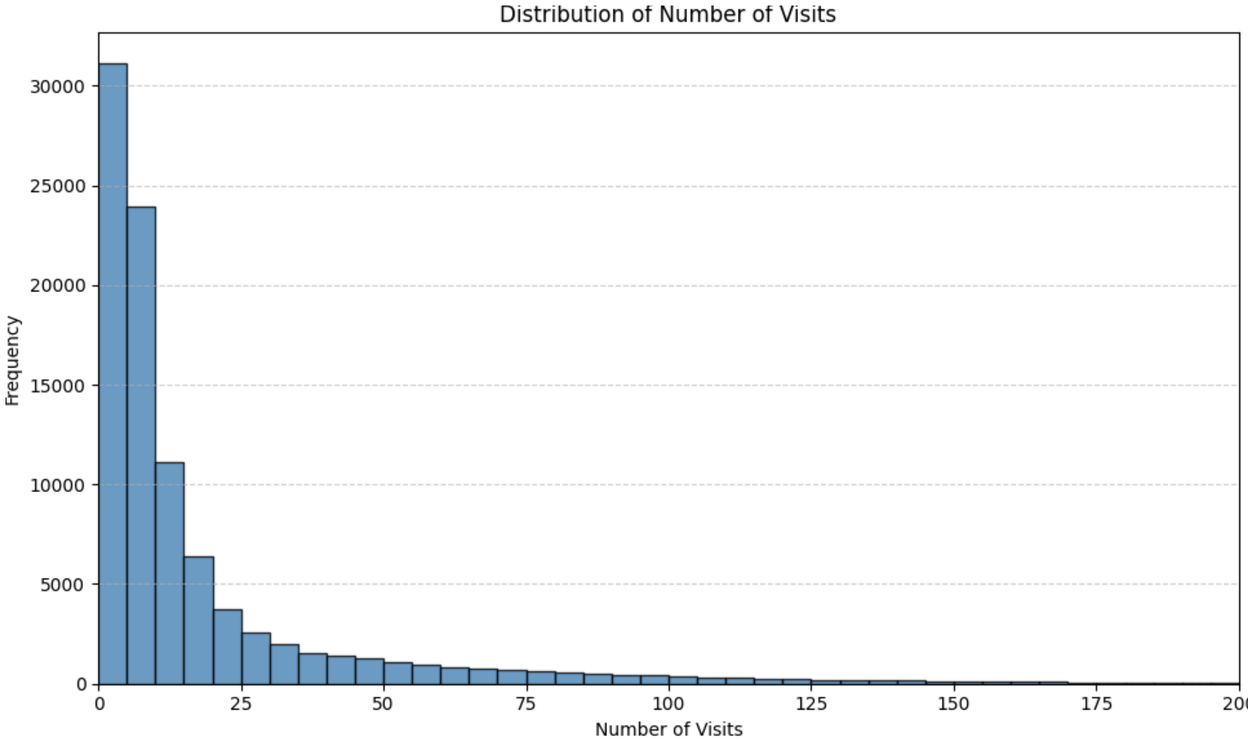

Figure 3.1: Distribution of the number of visits per patient

and evaluated on later ones; this reflects the real-world setting where predictions generalize to future events [151].

3.2 Methodology

At time t, only historical information up to t is available as inputs to the prediction model. Let $X_{0:t}$ denote the sequence of observed feature vectors from time 0 to t, and let $o_{t+\tau} \in \{\text{show}, \text{no-show}, \text{cancellation}\}$ be the scalar outcome of the next scheduled appointment at time $t+\tau$ for $\tau > 0$. The objective is to estimate

$$P(o_{t+\tau} \mid X_{0:t}) \approx f(X_{0:t}; \theta),$$

where $f(\cdot;\theta)$ is a supervised learning model that outputs a distribution over the outcomes, while θ represents a vector of trainable parameters. Predictions are made for an upcoming appointment using only information that would have been available at the time it was scheduled.

Table 3.3: Derived variables generated from raw scheduling data.

Derived Variable	**Description**
Cancellation Count	Total cancellations up to current appointment
No-Show Count	Total no-shows or “left without seen” events up to current appointment
Provider Change Count	Number of provider changes up to current appointment
Department Change Count	Number of department changes up to current appointment
Last Visit Duration	Duration of the last visit
Average Duration	Average standard duration of past appointments
Reschedule Count	Number of rescheduled appointments up to current appointment
Average Encounter	Average hours between consecutive appointments
Average Scheduled	Average hours from scheduling to next encounter
Average Arrival Lag	Average hours between scheduled time and arrival
Time Since Last Visit	Hours since the last patient visit
Cancellation Reason	Patient Choice, Clinic Mgmt, Patient Health, Socioeconomic, Provider Choice
Provider Changed Last Visit	Whether provider changed since last visit
Department Changed Last Visit	Whether department changed since last visit
Visit Type	Established Patient Follow-Up, New Chemotherapy, Treatment Chemotherapy
Encounter Type	Type of encounter (e.g., Appointment, Telemedicine, Infusion)
First or Follow-Up	Whether the visit is first or follow-up
Encounter Month	Month of appointment
Encounter Weekday	Day of the week
Last Appointment Status	Status of the last appointment
Encounter Hour Category	Time of day (Morning, Afternoon, Night)
Weekend Indicator	Whether appointment falls on weekend

3.2.1 Hierarchical Framework

Multiclass models typically optimize a single decision across all classes, but this can be problematic under severe class imbalance and overlapping decision boundaries. Hierarchical ML models decompose the task into sequential binary decisions, simplifying decision boundaries and improving minority-class detection. In this setting, cancellations and no-shows resemble each other more closely than they do shows as is illustrated in the t-SNE projection provided by Figure 3.2.

Prior ML studies have demonstrated that hierarchical approaches enhance both predictive performance and interpretability in structured healthcare settings [152]. Following this precedent, our design first isolates cancellations from the union of no-shows and shows (Stage 1), then distinguishes shows from no-shows among the residual cases (Stage 2). This ordering reflects operational realities: cancellations are explicitly recorded and often accompanied by reason codes, making them easier to separate reliably, whereas no-shows are rarer and less well-documented.

Formally, let $\mathcal{D}_1 = \{(x_i, y_i^{(1)})\}_{i=1}^{N}$ denote the training set for Stage 1, where $x_i \in \mathbb{R}^d$ is a feature vector of dimension d and $y_i^{(1)} \in \{0, 1\}$ is a binary label indicating membership in the

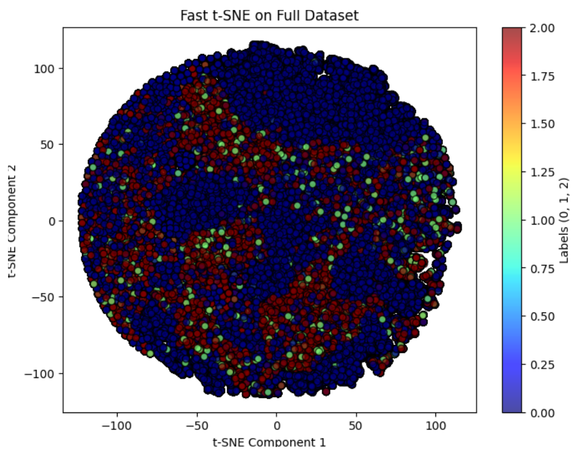

Figure 3.2: Visualization of t-SNE plot. Appointment outcomes: 0 = show, 1 = no-show, 2 = cancellation.

selected class (versus all others). Here, N is the number samples used for training. A binary classifier is trained on $\mathcal{D}_1$ to separate the target class (cancellations) from the rest (shows + no-shows). Stage 2 operates on the residual subset, partitioning between the two remaining outcomes. Let $\mathcal{D}_2 = \{(x_j, y_j^{(2)})\}_{j=1}^{M}$ denote this dataset, where $x_j \in \mathbb{R}^d$ is a feature vector, $y_j^{(2)} \in \{0, 1\}$ is a binary label distinguishing between the remaining outcomes (shows versus no-shows), and $M < N$ is the number of samples in the residual subset.

Each binary stage is trained using the standard binary cross-entropy loss function

$$\begin{aligned}\mathcal{L}(f^{(k)}(x), y^{(k)}) = &-y^{(k)} \log p_1(x) \\ &- (1 - y^{(k)}) \log p_0(x).\end{aligned}$$

where $f^{(k)}$ and $y^{(k)}$ denote the stage-$k \in \{1, 2\}$ classifier or outcome, respectively. Default optimizer settings from the respective libraries are used and no additional class weighting are applied. Instead, operating thresholds are tuned post-hoc using precision–recall analysis (Section 3.2.4) to emphasize minority-class performance.

3.2.2 Model Architectures

For the hierarchical approach, we utilize three model families – Decision Trees (DT), XGBoost, and a Multilayer Perceptron (MLP) – to compare traditional, ensemble, and neural methods in a structured manner. We also evaluate each of these models as a single-stage, multi-class classifier.

Decision Trees (DT): They are simple, fast to train, and highly interpretable, which makes them attractive for clinical use. However, they are prone to overfitting, unstable under small perturbations of the data, and generally underperform ensembles on large tabular

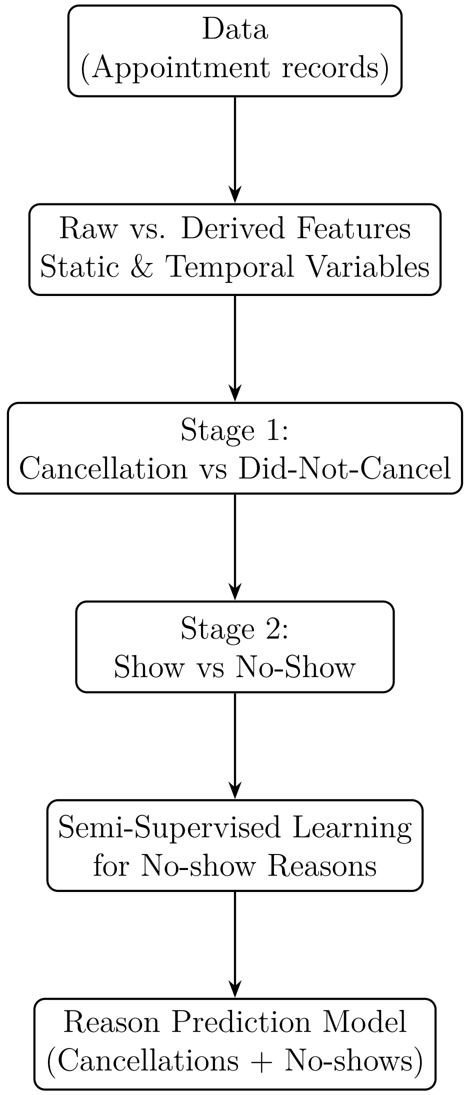

Figure 3.3: Hierarchical ML pipeline

datasets. Nevertheless, they serve as an interpretable baseline model [153].

XGBoost (Gradient-Boosted Decision Trees): Captures complex nonlinear feature interactions while controlling for overfitting through shrinkage, regularization, and subsampling. It has become a de facto state-of-the-art method for structured tabular data due to its scalability, robustness, and strong empirical performance [154].

Multilayer Perceptron (MLP): A fully connected feedforward neural network that transforms input feature vectors through multiple layers of linear transformations followed by nonlinear activation functions (e.g., ReLU). Although MLPs often lag behind tree ensembles on tabular data, recent benchmarks demonstrate that they can achieve competitive performance [155].

3.2.3 Reason Prediction Model

In addition to appointment outcomes, we train a model to estimate the underlying reason for a missed appointment, i.e., patient choice, clinic management, patient health, socioeconomic, provider choice. Let $\mathcal{D}_3 = \{(x_j, r_j)\}_j^L$ denote the subset of appointments with observed or inferred reason labels, where $x_k \in \mathbb{R}^d$ is the feature vector, $r_k \in \{1, \dots, R\}$ indexes one of R clinically defined categories (Appendix 3.2), and L is the number of samples labeled as a no-show or cancellation. The reason classifier is again trained using the standard multiclass cross-entropy loss,

$$\mathcal{L}_{\text{reason}}(f^{(3)}(x), r) = -\sum_{c=1}^{R} \mathbf{1}[r = c] \log p_c(x),$$

where $f^{(3)}$ denotes the reason model and $p_c(x)$ is the predicted probability of category c.

During training, the reason for a cancellation or no-show is learned from available labels, while at inference the reason model is applied only when the hierarchical classifier predicts a no-show or cancellation. In our dataset, however, reasons are documented only for cancellations, with no ground-truth labels available for no-shows. To address this, we employ confidence-based self-training (i.e., semi-supervised learning) to propagate labels from the 77,702 cancellations within 24 hours to the 25,377 no-shows.

The training approach is as follows: we first train a classification model using the dataset of appointments canceled within 24 hours, where reason labels are readily available. We then predict labels for the no-show data and augment the training set with the most confident predictions. This process repeats until no additional data can be classified with confidence. More formally, at iteration m, the calibrated model generates class posteriors $\{p_c(x)\}_{c=1}^{R}$ for the unlabeled data pool of no-show appointments; samples with $\max_c p_c(x) \geq \tau$ receive pseudo-labels $\hat{r} = \arg\max_c p_c(x)$ and are added to the labeled training set. The model is then retrained on the augmented dataset, and this process is repeated until no further samples exceed the threshold (typically within 20 iterations). The final model is evaluated on a holdout set.

3.2.4 Evaluation Metrics

Model performance is assessed using class-wise F1-scores for the three outcomes (show, no-show, cancellation). For class c, we define

$$\text{Precision}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c}, \quad \text{Recall}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c},$$

Table 3.4: Per-class F1-score of single-stage vs two-stage models.

Class	DT		XGBoost		MLP	
	Single	Two-stage	Single	Two-stage	Single	Two-stage
Show (0)	0.94	0.94	0.95	0.95	0.94	0.94
No-Show (1)	0.23	0.30	0.13	0.18	0.04	0.12
Cancellation (2)	0.78	0.87	0.85	0.92	0.83	0.92

and the F1-score is

$$\mathrm{F1}_c = \frac{2\,\mathrm{Precision}_c\,\mathrm{Recall}_c}{\mathrm{Precision}_c + \mathrm{Recall}_c}.$$

We use class-wise F1 as the primary evaluation metric because it balances precision and recall, which provides meaningful insight under severe class imbalance. As a complementary measure of overall discrimination, we also compute the area under the ROC curve (AUROC). This choice is consistent with recommendations emphasizing F1's reliability over ROC-based metrics when minority-class performance is critical [156].

3.2.5 Feature Set Analysis

To assess the relative importance of different feature groups, we conduct two sets of ablation experiments. These experiments isolate subsets of variables to evaluate how each type contributes to predictive accuracy, with full results reported in Section 4.4. We also examine feature importance plots and Shapley values – both the marginal contributions of variables and their interactions – to identify those most relevant for prediction.

Static vs. Temporal Features: Static features represent time-invariant appointment properties, such as visit type, department, and patient demographics. Temporal features, in contrast, encode event histories and sequential patterns, including timestamps and summaries of prior behavior (e.g., cumulative cancellation count, time since last visit) [157], [158].

Raw vs. Derived Features: Raw features are extracted directly from the scheduling system, such as appointment timestamps, encounter type, or visit type. Derived features are engineered from patient histories, such as average lead time, provider change count, or cumulative rescheduled appointments.

Chapter 4

Results

This chapter presents the empirical findings of the study. The analysis yields two primary outcomes. First, the proposed hierarchical machine learning framework substantially improves minority-class performance, particularly for no-shows, without compromising overall accuracy, and outperforms its single-stage counterparts. These improvements are consistent across train–test splits and random seeds. Second, integrating static and temporal features provides complementary gains in F1-scores and accuracy over using either set alone, with feature importance and SHAP analyses confirming the value of both variable types. Strong interaction effects between static and temporal variables—as well as raw and derived data—further enhance predictive accuracy. The following sections present these results in detail.

Training and evaluation were conducted on Google Colab (v2-8 TPU system) with 334 GB RAM and 225 GB disk space, running Python 3. Full details of software libraries and code are provided in the project repository to ensure reproducibility. Random seeds were fixed across all experiments. Each reported result corresponds to a single training run, as performance variance across repeated runs using different seeds was negligible. Unless otherwise noted, model hyperparameters were set to library defaults, consistent with evidence that default settings often yield performance comparable to extensive hyperparameter searches [159].

4.1 Predictive Performance

The hierarchical classification framework was compared against a single-stage multiclass baseline while holding model architecture fixed. For no-shows, F1-scores increased from 0.23 to 0.30 with Decision Trees (DTs), 0.13 to 0.18 with XGBoost, and 0.04 to 0.12 with MLPs. For cancellations, F1-scores increased from 0.78 to 0.87 (DT), 0.85 to 0.92 (XGBoost), and 0.83 to 0.92 (MLP). F1-scores for patients who attended their appointment remained high and stable across all architectures (0.94–0.95). These findings indicate that the hierarchical modeling framework yields consistent gains for the minority classes without degrading performance for the majority class.

To validate the framing of outcome subsets at each stage, Table 4.1 reports results for three distinct two-stage permutations using DTs for interpretability and consistency: (i) Show → No-Show/Cancel, (ii) Cancel → Show/No-Show, and (iii) No-Show → Cancel/Show. Among these, the **Cancel → Show/No-Show** permutation achieved the best overall performance, balancing accuracy across stages while maintaining strong discriminative power for no-shows.

Table 4.1: F1-scores for Shows, No-Shows, and Cancellations across three key two-stage permutations. Per-stage macro F1 scores (F1) and accuracy (Acc) are also reported. Bold values represent the best score per column.

	F1-Scores			Stage 1		Stage 2	
Permutation	**Show**	**No-Show**	**Cancel**	**F1**	**Acc**	**F1**	**Acc**
Show → No-Show/Cancel	0.94	0.30	**0.87**	**0.92**	**0.92**	0.58	0.78
Cancel → Show/No-Show	**0.99**	0.60	0.78	0.84	0.87	0.79	**0.97**
No-Show → Cancel/Show	0.96	**0.68**	0.78	0.84	0.87	**0.82**	0.93

Next, we fixed the hierarchical structure to *Stage 1: Cancellation vs Not-Canceled* and *Stage 2: Show vs No-Show*, and compared architectures. Table 4.3 reports F1 results for all combinations. Two pipelines (DT → DT and DT → XGBoost) achieved the best balance, especially for no-shows. To refine trade-offs, decision thresholds were tuned via precision–recall sweeps. Table 4.2 presents the selected thresholds for each architecture, rounded to two decimals. As shown, the DT → XGBoost pipeline achieves the best balance. The final optimized results for this pipeline are then reported in Table 4.4, where no-show F1 increased from 0.67 to 0.75 and cancellation F1 improved similarly.

Table 4.2: Two-stage permutations with optimal thresholds (rounded to two decimals). Thresholds were selected by maximizing F1-score using the precision–recall curve.

Architecture	**Cancellations**	**Threshold 1**	**Shows**	**No-Shows**	**Threshold 2**
Decision Tree → Decision Tree	0.78	0.33	0.96	0.69	0.33
Decision Tree → XGBoost	0.78	0.43	0.97	0.75	0.83
XGBoost → Decision Tree	0.85	0.43	0.97	0.00	0.00
XGBoost → XGBoost	0.85	0.43	0.98	0.27	0.62

These results demonstrate that the hierarchical approach yields substantial improvements in minority-class detection compared with prior baselines such as heuristics or logistic regression [17], [31], [35], which typically achieve modest discrimination (AUC 0.55–0.65) and limited precision.

Table 4.3: F1-scores for all combinations of two-stage model architectures (Stage 1 = Cancellation vs Not-Canceled; Stage 2 = Show vs No-Show). Bold values represent the best score per column.

Permutation	Cancellations	Shows	No-Shows
DT → DT	0.78	0.96	**0.67**
DT → XGBoost	0.78	0.96	**0.67**
DT → MLP	0.78	0.96	0.64
XGBoost → DT	**0.85**	**0.97**	0.35
XGBoost → XGBoost	**0.85**	**0.97**	0.27
XGBoost → MLP	**0.85**	**0.97**	0.28
MLP → DT	0.83	0.97	0.32
MLP → XGBoost	0.83	0.97	0.22
MLP → MLP	0.83	0.97	0.18

Table 4.4: Detailed results for the best-performing two-stage pipeline (Decision Tree → XGBoost). Results are reported after threshold optimization (Stage 1 threshold = 0.43, Stage 2 threshold = 0.83).

Outcome	Precision	Recall	F1-Score	Support
Stage 1: Decision Tree (Cancellation vs Not-Canceled)				
Not-Canceled	0.91	0.90	0.91	256,327
Cancellation	0.78	0.78	0.78	108,863
Accuracy			0.87	365,190
Macro Avg.	0.84	0.84	0.84	365,190
Weighted Avg.	0.87	0.87	0.87	365,190
Optimal Threshold			0.43	—
Stage 2: XGBoost (Show vs No-Show)				
No-Show	0.87	0.65	0.75	31,855
Show	0.95	0.99	0.97	223,765
Accuracy			0.94	255,620
Macro Avg.	0.91	0.82	0.86	255,620
Weighted Avg.	0.94	0.94	0.94	255,620
Optimal Threshold			0.83	—

4.2 Semi-Supervised Learning

To structure the prediction of reasons for missed appointments, we define five broad categories (Table 3.2) based on prior literature. These categories capture the dominant drivers of cancellations and no-shows.

To optimize pseudo-labeling for no-show reason prediction, thresholds between 0.60 and 0.90 were tested. As shown in Figure 4.1, a threshold of 0.60 incorporated over 90% of unlabeled data, whereas a threshold of 0.85 only included 7,901 of 25,377 samples. Predictive performance was higher at 0.60 due to the increased training volume, despite noisier labels. Thus, 0.60 was selected as the final threshold.

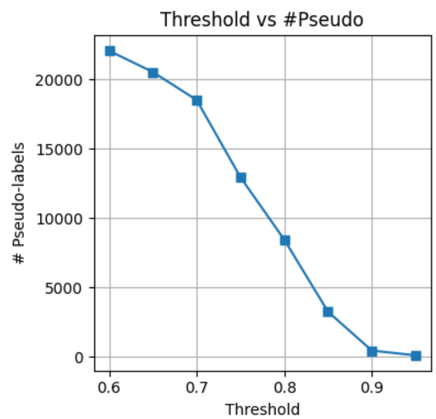

Figure 4.1: Effect of the pseudo-labeling confidence threshold on the number of pseudo-labeled examples used in training. As the threshold increases (from 0.60 to 0.95), fewer examples meet the confidence criteria and are included.

Table 4.5 compares results across models with and without pseudo-labeling. XGBoost consistently achieved the best performance, with accuracy 0.71 and weighted F1-score 0.70. Pseudo-label distributions across categories (e.g. 11,495 *Patient Choice*, 50,872 *Clinic Management*) mirrored observed cancellation distributions, suggesting plausible inference.

Table 4.5: Model performance on predicting reasons for no-shows using semi-supervised learning with pseudo-labeling thresholds of 0.60 and 0.85. F1-scores are weighted averages. The best performing model and threshold are bolded.

	No Pseudo-Labels		**Pseudo-Labels (0.60)**		**Pseudo-Labels (0.85)**	
Model	**Accuracy**	**F1-Score**	**Accuracy**	**F1-Score**	**Accuracy**	**F1-Score**
DecisionTree	0.59	0.56	0.59	0.55	0.56	0.53
XGBoost	0.67	0.67	**0.71**	**0.70**	0.68	0.70
MLP	0.52	0.46	0.61	0.61	0.62	0.62

The final results in Table 4.6 confirm that XGBoost outperforms DT and MLP for both no-show and cancellation reasons.

Table 4.6: Final model performance on predicting reasons for no-shows and cancellations using pseudo-labeling at threshold 0.60. F1-scores are weighted averages. Bold values represent the best score per row.

	No-Shows		**Cancellations**	
Model	**Accuracy**	**F1-Score**	**Accuracy**	**F1-Score**
DecisionTree	0.59	0.55	0.49	0.47
XGBoost	**0.71**	**0.70**	**0.56**	**0.55**
MLP	0.61	0.61	0.49	0.44

4.3 Generalization Across Splits

Robustness of the hierarchical pipeline (DT → XGBoost) was evaluated under three partitioning strategies: row-wise random, patient-level, and appointment-time splits. Table 4.7 shows consistent performance across cancellations and shows, while no-shows achieved the highest F1-score (0.82) under appointment-time splitting. This effect is attributed to recurrence patterns captured by temporal splitting [160]. For comparability, row-wise splitting was used for all main experiments.

Table 4.7: F1-scores across data splitting strategies using the final hierarchical pipeline (i.e., DT → XGBoost). Bold values represent the best score per column.

Splitting Method	**Cancellations**	**No-Shows**	**Shows**
Row-Wise (Random)	**0.78**	0.75	**0.97**
Patient-Level	0.75	0.74	**0.97**
Appointment-Time	0.76	**0.82**	**0.97**

4.4 Feature Importance and Ablation Analysis

To enhance interpretability, feature importance was assessed using SHAP, Permutation Importance, Mutual Information (MI), and Mean Decrease in Impurity (MDI). Results indicate that both static and temporal features contribute meaningfully to prediction. Static appointment characteristics (e.g. encounter and visit type) and temporal features (e.g. lead time, time since last visit) consistently ranked among the strongest predictors. Derived features such as provider change count also proved influential. Table 4.8 lists the raw and derived temporal variables alongside static features used in the modeling framework.

Table 4.8: Unprocessed variable names for static and temporal features

Feature	**Description**	**Data Type**
Static features		
Visit_Type	Type of visit (e.g., Blood Draw, Chemo)	Categorical
Encounter_Type	Encounter category (e.g., Infusion, Appointment)	Categorical
Clin_Dept_Nm	Clinical department name	Categorical
Floor_Location_Nm	Floor location of the appointment	Categorical
Temporal features		
Raw temporal variables		
Encounter_Dttm	Datetime of the patient encounter	Datetime
Schd_Dttm	Datetime of appointment scheduling	Datetime
Derived temporal variables		
Cancellation_Count	Total past cancellations before this appointment	Numeric
No_Show_Left_Count	Past no-shows or left-without-being-seen visits	Numeric
Provider_Change_Count	Number of times provider changed	Numeric
Department_Change_Count	Number of department changes in history	Numeric
Last_Visit_Duration	Duration of the last visit	Numeric
Avg_Std_Duration	Average scheduled duration of past appointments	Numeric
Rescheduled_Appointments_Count	Past rescheduled appointments	Numeric
Avg_Encounter_Frequency_Hours	Average time between encounters	Numeric
Avg_Schd_Encounter_Diff_Hours	Average time from scheduling to encounter	Numeric
Average_Arrival_Lag_Hours	Average delay from scheduled to arrival time	Numeric
Time_Since_Last_Visit	Time since previous visit	Numeric
Patient_Choice	Cumulative cancellations due to patient-driven reasons	Numeric
Clinic_Management	Cumulative cancellations due to clinic management	Numeric
Patient_Health	Cumulative cancellations due to patient health	Numeric
Socioeconomic	Cumulative cancellations due to socioeconomic factors	Numeric
Provider_Choice	Cumulative cancellations due to provider-driven choice	Numeric
Provider_Changed_Last_Visit	Whether provider changed since last visit	Boolean
Department_Changed_Last_Visit	Whether department changed since last visit	Boolean
First_Follow_Up_First	First visit in a sequence	Boolean

Continued on next page

Table 4.8 – *Continued from previous page*

Feature	Description	Data Type
First_Follow_Up_Follow_Up	Follow-up visit	Boolean
Last_Appointment_Status_{...}	Status of the last appointment	Categorical
Encounter_Hour_{...}	Hour-of-day bucket	Categorical
Weekend_Indicator_{...}	Weekend indicator	Boolean
Encounter_Dttm_{...}	Components of encounter datetime	Datetime
Schd_Dttm_{...}	Components of scheduling datetime	Datetime

Ablation results in Tables 4.9 and 4.10 show that combining static and temporal features, as well as raw and derived variables, yields the strongest results. Interaction analyses using SHAPIQ further revealed that static–temporal interactions accounted for 24.6% of predictive contribution, exceeding static–static or temporal–temporal interactions (Table 4.11). This highlights the synergistic role of combining diverse feature types.

Table 4.9: Class-wise F1-scores by feature family. Stage 1 evaluates cancellations; Stage 2 evaluates no-shows and shows. Bold values represent the best score per column.

Feature Family	Cancellations (Stage 1)	No-Shows (Stage 2)	Shows (Stage 2)
Static only	**0.83**	0.12	**0.98**
Temporal only	0.44	0.03	0.82
Static + Temporal	0.76	**0.69**	0.96

Table 4.10: Class-wise F1-scores by variable type. Stage 1 evaluates cancellations; Stage 2 evaluates no-shows and shows. Bold values represent the best score per column.

Feature Type	Cancellations (Stage 1)	No-Shows (Stage 2)	Shows (Stage 2)
Raw Only	0.76	0.69	0.96
Derived Only	0.58	0.11	0.87
Raw + Derived	**0.78**	**0.74**	**0.97**

Table 4.12 summarizes the mapping between raw and derived features. Ablation results in Table 4.10 show how these groups affect predictive performance.

Table 4.11: Relative contribution of feature categories based on SHAPIQ values (main effects and pairwise interactions). Results are from 100 subsamples, background = full test set, max order = 2.

Category	**Share (%)**
Static Main	27.1
Temporal Main	10.4
Static–Static Interaction	22.2
Temporal–Temporal Interaction	15.8
Static–Temporal Interaction	24.6

Table 4.12: Unprocessed variable names for raw and derived features

Raw Features	**Derived Features**
Visit_Type	Cancellation_Count
Encounter_Type	No_Show_Left_Count
Clin_Dept_Nm	Provider_Change_Count
Floor_Location_Nm	Department_Change_Count
Schd_Dttm	Avg_Std_Duration
Encounter_Dttm	Last_Visit_Duration
	Rescheduled_Appointments_Count
	Avg_Encounter_Frequency_Hours
	Avg_Schd_Encounter_Diff_Hours
	Average_Arrival_Lag_Hours
	Time_Since_Last_Visit
	Patient_Choice
	Clinic_Management
	Patient_Health
	Socioeconomic
	Provider_Choice
	Provider_Changed_Last_Visit
	Department_Changed_Last_Visit
	First_Follow_Up_First
	First_Follow_Up_Follow_Up
	Last_Appointment_Status_Arrived
	Last_Appointment_Status_Canceled
	Last_Appointment_Status_Completed

Continued on next page

Table 4.12 – *Continued from previous page*

Raw Features	**Derived Features**
	Last_Appointment_Status_Left_without_-seen
	Last_Appointment_Status_No_Show
	Encounter_Hour_After_18PM
	Encounter_Hour_Before_12PM
	Encounter_Hour_Between_12_18PM
	Weekend_Indicator_0
	Weekend_Indicator_1

4.5 Summary

In summary, the hierarchical framework consistently improved predictive performance, particularly for minority classes, across multiple evaluation strategies. Semi-supervised learning with pseudo-labeling further enhanced classification of no-show reasons. Finally, feature importance and ablation analyses confirmed that static and temporal variables interact synergistically to improve predictive accuracy.

Chapter 5

Counterfactual Analysis

While predictive models can identify patients at risk of missed appointments, they do not by themselves provide actionable guidance for administrators or schedulers. From a managerial perspective, the important question is not only *who* is at risk of being a no-show or canceling their appointment, but also *which interventions are likely to reduce this risk and by how much*. A counterfactual analysis addresses this concern by evaluating the impact of alternative scheduling or engagement strategies on predicted outcomes. However, many prior studies have emphasized predictive accuracy without assessing how proposed quality improvement initiatives might affect clinic operations [161]. A few recent studies—a trial of predictive model–driven appointment reminders [139] and simulations of rescheduling effects [162]—have begun to explore this, but efforts remain limited. This lack of evaluation is problematic, as strategies may be less effective when applied in different contexts and may even reinforce existing inequities without careful policy design [140].

To this end, we utilize our modeling framework to evaluate several interventions with the goal of identifying operationally feasible changes that could reduce the predicted risk of cancellations and no-shows. The evaluated interventions are structured along two complementary axes:

- Interventions informed by previous research.
- Novel operational strategies motivated by the design of the hierarchical ML pipeline.

Tables 5.1 and 5.2 provide an overview of intervention categories drawn from prior work, distinguishing between predictive, real-time, and hybrid strategies. Our analysis builds upon this taxonomy by quantifying their potential impact in the context of oncology treatment scheduling.

Table 5.1: Intervention strategies for reducing missed appointments, categorized by delivery mechanism. Interventions are grouped by predictive modeling, real-time operations, or hybrid approaches.

Category	Predictive Modeling	Real-Time / Operational	Hybrid (Predictive + Real-Time)
Reminders	Predictive SMS/text reminders [139] Predictive phone-call reminders [139] Patient-preferred reminders [163] Staff vs automated reminders [164] Dual reminders [165] Behavioral nudges [166] Targeted reminder calls [167] Postal + digital reminders [168] Predictive navigator outreach [139] Predictive ML + targeted phone outreach [140]	–	–
Scheduling	AI-based predictive scheduling [169] Overbooking + ML risk alerts [139] Telehealth / remote consults [170]	Open-access scheduling [142] Automated rescheduling ("Fast Pass") [171] Real-time AI-driven overbooking [172] Intraday dynamic rescheduling [173] Dynamic slot filling [172] Portal message after missed appointment [174]	Two-stage stochastic scheduling [175] Predictive no-show model + dynamic overbooking [141] Multi-method intervention [176]
Incentives	Transportation vouchers [143] No-show fine ($30) [144]	–	–

5.1 Literature-Informed Interventions

Studies have assessed strategies such as appointment reminders, reduced lead time, and confirmation requirements. These interventions generally show reductions in no-shows and cancellations [140], [164], [167], [168], [177]. Within our modeling framework, we simulate their likely impact by perturbing associated input variables, as mapped in Table 5.2.

Table 5.2: Interventions and corresponding model variables used in the counterfactual analysis. Literature-informed interventions (first two groups) and novel interventions (last three groups) are shown.

Intervention	**Variable**	**Studies**
SMS/AutoCall Reminders	Cancellation Count	SMS reduces cancellations [177]
	No Show Left Count	No-shows ↓ 1.9–7% abs. [164], [168], [178]
	Rescheduled Appointments	Reminders improve rescheduling [171]
	Patient Choice	Tailored reminders effective [178]
Scheduling Timing	Avg Schd Encounter Diff	Shorter lead time ↓ no-shows [179]
	Cancellation Count	Shorter intervals ↓ cancellations [179]
	Patient Choice Cancels	Preference-based scheduling ↓ cancellations [179]
Overnight Patient Confirmation Calls	Cancellation Count	–
Overnight Provider Confirmation	Provider Choice Cancels	–
Provider/Department Consistency	Provider Change Count	–

5.2 Novel Operational Strategies

Analysis of reason codes shows cancellations and no-shows are primarily driven by operational or preference-based factors. Two novel interventions are therefore modeled: *Provider Consistency* and *Commitment-Based Scheduling*. Table 5.3 summarizes their simulated effects.

Table 5.3: Counterfactual analysis: absolute change in cancellation and no-show rates for each intervention. Positive values = improvement; negative values = worsening. Mean ± SD across 10 models.

Category	**Intervention**	**Δ** Cancel (%)	**Δ** No-show (%)
Literature-Informed	SMS Reminders	-0.16 ± 0.04	-0.43 ± 0.03
	Patient Confirmation	-0.09 ± 0.05	-0.40 ± 0.03
	Reduced Lead Time	-0.14 ± 0.62	-1.80 ± 0.31
Novel Strategies	Provider Consistency	2.19 ± 0.04	-1.92 ± 0.02
	Commitment-Based Scheduling	0.13 ± 0.03	0.11 ± 0.01

5.3 Simulation Setup

Counterfactual simulations were implemented by modifying test set variables according to intervention mappings. Adjustments included proportional changes, categorical flips, and caps

(Table 5.4). Outcomes were measured as absolute percentage point changes in cancellation and no-show rates.

Table 5.4: Variable-level changes applied in counterfactual simulations. Each intervention modifies inputs for both cancellation (Stage 1) and no-show (Stage 2) predictions.

Intervention	Variable	Change	Rationale
SMS Reminder	Cancellation_Count	+12%	More clarity reduces cancellations
	No_Show_Left_Count	-9%	Improves accountability
	Rescheduled_Appointments	+15%	Better follow-through
Patient Confirmation	Cancellation_Count	+24%	Improved tracking
	No_Show_Left_Count	-18%	Higher compliance
Reduced Lead Time	Avg_Schd_Encounter_Diff	Cap: 168 hrs	Max 7 days lead time
	Cancellation_Count	-50%	Shorter planning reduces cancels
Provider Consistency	Provider Change Count	Cap: 6	Improved continuity
	Department Change Count	Cap: 6	Consistent care
Commitment-Based Scheduling	No_Show_Left_Count	-20%	Limits stacking of missed appts
	Socioeconomic	-20%	More support provided

5.4 Extreme-Value Simulations

Extreme-value tests were run by setting individual variables to zero for all patients. Table 5.5 reports changes in predicted outcomes. These represent best-case but less feasible scenarios.

Table 5.5: Absolute reduction in predicted cancellation and no-show rates when each variable is individually set to zero. Mean ± SD across 10 models.

Variable	Max	Δ Cancel (%)	Δ No-show (%)
Provider Changed Last Visit	0/1	2.21 ± 0.04	-0.36 ± 0.02
No Show Left Count	59	-0.19 ± 0.01	0.46 ± 0.01
Department Change Count	10	0.48 ± 0.07	-0.66 ± 0.08

5.5 In-Situ Results

Table 5.3 summarizes the absolute change in predicted cancellation and no-show rates for each intervention. Among the literature-informed policies, *SMS Reminders* worsened both outcomes (cancellations: –0.16%; no-shows: –0.43%), *Patient Confirmation* also worsened both outcomes (cancellations: –0.09%; no-shows: –0.40%), while *Reduced Lead Time* produced the largest degradation (cancellations: –0.14%; no-shows: –1.80%). These results contrast with prior studies that found these policies to be effective in other clinical settings.

The novel strategies yield more promising results. *Provider Consistency* achieved the largest overall reduction in cancellations (+2.19%) but increased no-shows (–1.92%), suggesting that continuity-of-care primarily reduces patient-driven cancellations while failing to address structural barriers to attendance. In contrast, *Commitment-Based Scheduling* generated modest improvements in both outcomes (cancellations: +0.13%; no-shows: +0.11%), making it the only intervention to improve both simultaneously. Although smaller in magnitude, these effects are managerially relevant as they demonstrate potential to concurrently reduce cancellations and no-shows. Furthermore, while the policy is informed by prior literature, it is adapted to address the unique behavior of the patient population.

The results of the extreme-value ablation study (Table 5.5) reinforce these insights. The largest reduction in cancellations was observed when `Provider Changed Last Visit` was set to zero (+2.21%), which underscores the central role of care continuity in preventing cancellations. The largest improvement in no-shows occurred when `No Show Count` was eliminated (+0.46%), which indicates that a small subset of patients with repeated prior no-shows accounts for the majority of missed appointments. Smaller but consistent gains were also observed from reducing provider changes (+0.44%) and department changes (+0.48%), although the extent to which these aspects of care can be modified may be constrained by medical considerations.

5.6 Summary

This chapter presented counterfactual analyses simulating both literature-informed and novel interventions. Literature-based policies such as reminders and confirmations showed mixed or negative effects in this setting, while provider consistency and commitment-based scheduling offered more promising improvements. Extreme-value simulations highlighted provider and patient history variables as critical leverage points for reducing missed appointments.

Chapter 6

Discussion

This study demonstrates that decomposing appointment outcome prediction into a hierarchical pipeline improves minority-class detection as compared to a single-stage multiclass baseline. The hierarchy first isolates cancellations and then separates show from no-shows, consistent with the observed proximity of cancellation and no-show classes in the t-SNE projection. In doing so, we observe F1-score improvements of 7–10 points in comparison to their single-stage counterparts. Ablation studies further reveal that combining static and temporal variables outperforms using either alone, and that enriching raw features with derived variables further enhances predictive performance. In addition, for cancellations and no-shows, the pipeline predicts the most likely underlying reason for the missed appointment. This enables a counterfactual analysis to evaluate the effectiveness of operational interventions that clinics can implement to reduce the numbers of missed appointments. We find that common strategies (e.g., appointment reminders) may be much less effective than often claimed, and develop alternative strategies that, using our modeling approach, have greater impact. Thus, our study contributes to the literature by introducing a novel ML model that achieves superior results—despite data constraints imposed by real-world operational workflows—and showcases its unique prescriptive value when applied to a specialized clinical setting, such as chemotherapy treatment.

The results support two practical levers. First, integrating static attributes with derived temporal features substantially improves the identification of patients that may be at risk of missing an appointment, which is critical for targeted outreach. Second, the counterfactual analysis suggests multiple strategies—patient confirmation programs and better provider/department care continuity—can effectively reduce the number of no-shows and/or cancellations. This suggests that multiple interventions can reduce their incidence, providing providers with flexibility to balance improved care quality with cost-effectiveness.

Importantly, this framework can be readily implemented in clinical practice through integration with existing scheduling tools. Model outputs can be directly embedded into dashboards as color-coded alerts or flags, which would allow care teams to immediately provide outreach or proactively adjust appointments. Moreover, rather than merely flagging at-risk patients, the framework also attempts to explain *why* they might miss their next appointment. Thus, our approach supports prescriptive decision-making with data-driven, scenario-specific estimates of impact. This is particularly important in the context of outpatient chemotherapy treatment, where the magnitude of previously published interventions does not necessarily generalize directly to our setting.

This study is limited by the type of data that is collected. Due to the sensitivity of patient-level information and the challenges of mining electronic health records, our analysis relies exclusively on operational-level data. Thus, future work should incorporate additional modalities such as demographic information, socioeconomic data, as well as unstructured electronic health records (e.g., clinical notes, pathology reports, communication logs). These sources may capture patient- or context-specific factors influencing attendance that are not reflected in structured scheduling data. In addition, our analysis reports offline predictive performance and simulated counterfactual effects, and we do not claim causal interpretation of the results. Future work should validate the accuracy and robustness of these findings in experimental settings. Nevertheless, although our findings are grounded in oncology scheduling, the hierarchical ML and prescriptive framework is broadly applicable to other domains—such as primary care, mental health, and chronic disease management—where appointment adherence is critical. Core elements such as the hierarchical approach, a reason prediction model, and counterfactual analysis may translate well. Thus, future work should look to validate the general framework in other outpatient settings.

Reproducibility and Code Availability

All tables and figures can be reproduced from the defined features, data partitions, and thresholds. Hyperparameters, search ranges, computing environment, and full training/evaluation scripts are maintained in the following GitHub repository:

`https://github.com/mrrajabzadeh/Hierarchical-ML-for-Missed-Chemotherapy-Appointment-Prevention`

Bibliography

[1] L. H. A. Salazar, W. D. Parreira, A. M. da Rocha Fernandes, and V. R. Q. Leithardt, "No-Show in Medical Appointments with Machine Learning Techniques: A Systematic Literature Review," *Information*, vol. 13, no. 11, p. 507, 2022.

[2] M. Bech, "The Economics of Non-Attendance and the Expected Effect of Charging a Fine on Non-Attendees," *Health Policy*, vol. 74, pp. 181–191, 2005.

[3] L. F. Dantas, J. L. Fleck, F. L. C. Oliveira, and S. Hamacher, "No-shows in Appointment Scheduling: A Systematic Literature Review," *Health Policy*, vol. 122, no. 4, pp. 412–421, 2018.

[4] T. P. Hanna, W. D. King, S. Thibodeau, *et al.*, "Mortality Due to Cancer Treatment Delay: Systematic Review and Meta-Analysis," *BMJ*, vol. 371, p. m4087, 2020.

[5] S. M. McLean, A. Booth, M. Gee, *et al.*, "Appointment Reminder Systems Are Effective but Not Optimal: Results of a Systematic Review and Evidence Synthesis Employing Realist Principles," *Patient Preference and Adherence*, vol. 10, pp. 479–499, 2016.

[6] K. M. Toffaha, M. C. E. Simsekler, M. A. Omar, and I. ElKebbi, "Predicting Patient No-Shows Using Machine Learning: A Comprehensive Review and Future Research Agenda," *Intelligence-Based Medicine*, vol. 11, p. 100 229, 2025.

[7] C. G. Moore, P. Wilson-Witherspoon, and J. C. Probst, "Time and money: Effects of No-shows at a Family Practice Residency clinic," *Family Medicine*, vol. 33, no. 7, pp. 522–527, 2001.

[8] A. F. Hesaraki, N. P. Dellaert, and T. de Kok, "Generating Outpatient Chemotherapy Appointment Templates with Balanced Flowtime and Makespan," *European Journal of Operational Research*, vol. 275, no. 1, pp. 304–318, 2019.

[9] S. Karakaya, S. Gül, and M. Çelik, "Stochastic Scheduling of Chemotherapy Appointments Considering Patient Acuity Levels," *European Journal of Operational Research*, vol. 305, no. 2, pp. 902–916, 2023.

[10] A. Alaeddini, K. Yang, C. Reddy, and S. Yu, "A Probabilistic Model for Predicting the Probability of No-Show in Hospital Appointments," *Health Care Management Science*, vol. 14, no. 2, pp. 146–157, 2011.

[11] A. Alaeddini, K. Yang, P. Reeves, and C. K. Reddy, "A Hybrid Prediction Model for No-Shows and Cancellations of Outpatient Appointments," *IIE Transactions on Healthcare Systems Engineering*, vol. 5, no. 1, pp. 14–32, 2015.

[12] J. Chen, I. H. Goldstein, W.-C. Lin, M. F. Chiang, and M. R. Hribar, "Application of Machine Learning to Predict Patient No-Shows in an Academic Pediatric Ophthalmology Clinic," in *AMIA Annual Symposium Proceedings*, vol. 2020, 2021, pp. 293–302.

[13] D. Carreras-García, D. Delgado-Gómez, F. Llorente-Fernández, and A. Arribas-Gil, "Patient No-show Prediction: A Systematic Literature Review," *Entropy*, vol. 22, no. 6, p. 675, 2020.

[14] S. L. Harris, Q. Feng, and Y. Liu, "Predictive Analytics Model for Healthcare Planning and Scheduling," *European Journal of Operational Research*, vol. 253, no. 1, pp. 121–131, 2016.

[15] R. van den Goorbergh, E. Steyerberg, B. van Calster, and L. Wynants, "Calibration of Clinical Prediction Models: The Forgotten Aspect of Model Performance," *JAMA Internal Medicine*, vol. 182, no. 9, pp. 1049–1056, 2022.

[16] C. Yang, E. A. Fridgeirsson, J. A. Kors, J. M. Reps, and P. R. Rijnbeek, "Impact of Random Oversampling and Random Undersampling on the Performance of Prediction Models Developed Using Observational Health Data," *Journal of Big Data*, vol. 11, no. 1, p. 7, 2024.

[17] J. Daggy, M. Lawley, D. Willis, *et al.*, "Using No-show Modeling to Improve Clinic Performance," *Health Informatics Journal*, vol. 16, no. 4, pp. 246–259, 2010.

[18] Y. Huang and D. A. Hanauer, "Patient No-show Predictive MOdel Development Using Multiple Data Sources for an Effective Overbooking Approach," *Applied Clinical Informatics*, vol. 5, no. 3, pp. 836–860, 2014.

[19] H. Lenzi, Â. J. Ben, and A. T. Stein, "Development and Validation of a Patient No-show Predictive Model at a Primary Care Setting in Southern Brazil," *PLOS ONE*, vol. 14, no. 4, e0214869, 2019.

[20] L. Liu, Y. Zhang, J. Gong, *et al.*, "Predicting Appointment No-Shows in Real Time Using Interpretable Machine Learning," *npj Digital Medicine*, vol. 5, no. 1, p. 69, 2022.

[21] M. Samorani, S. L. Harris, L. G. Blount, H. Lu, and M. A. Santoro, "Overbooked and Overlooked: Machine Learning and Racial Bias in Medical Appointment Scheduling," *Manufacturing & Service Operations Management*, vol. 24, no. 6, 2021.

[22] M. Koushan, L. C. Wood, and R. Greatbanks, "Evaluating Factors Associated with the Cancellation and Delay of Elective Surgical Procedures: A Systematic Review," *International Journal for Quality in Health Care*, vol. 33, no. 2, mzab092, 2021.

[23] M. Alkomos, O. Alkhoshaiban, S. Almazyad, *et al.*, "Understanding Patient Reasons for Missed Appointments in a Saudi Arabian Hospital: A Cross-Sectional Survey," *Patient Preference and Adherence*, vol. 14, pp. 2293–2303, 2020.

[24] W.-J. Tuan, J. D. Brown, T. J. Weida, *et al.*, "Development and Validation of a Machine Learning Model to Predict Missed Primary Care Appointments," *Annals of Family Medicine*, vol. 23, no. 1, pp. 24–31, 2025.

[25] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 2020. [Online]. Available: `https://christophm.github.io/interpretable-ml-book/`.

[26] E. Meulemeester, A. van Looveren, E. O'Neil, *et al.*, "Counterfactual explanations for machine learning: A review," 2023. arXiv: `2301.00058`.

[27] R. M. Goffman, S. L. Harris, J. H. May, *et al.*, "Modeling Patient No-Show History and Predicting Future Outpatient Appointment Behavior in the Veterans Health Administration," *Military Medicine*, vol. 182, no. 5–6, e1708–e1714, 2017.

[28] Y.-L. Huang and D. A. Hanauer, "Time Dependent Patient No-Show Predictive Modelling Development," *International Journal of Health Care Quality Assurance*, vol. 29, no. 4, pp. 475–488, 2016.

[29] Y. Li, S. Y. Tang, J. Johnson, and D. A. Lubarsky, "Individualized No-Show Predictions: Effect on Clinic Overbooking and Appointment Reminders," *Production and Operations Management*, vol. 28, no. 8, pp. 2068–2086, 2019.

[30] T. L. Ayalew, "Case cancellations and associated factors on the day of surgery in hospitals of wolaita zone, south ethiopia," *BMC Surgery*, vol. 24, p. 45, 2024. DOI: `10.1186/s12893-024-02330-5`. [Online]. Available: `https://doi.org/10.1186/s12893-024-02330-5`.

[31] J. B. Norris, C. Kumar, S. Chand, H. Moskowitz, S. A. Shade, and D. R. Willis, "An Empirical Investigation into Factors Affecting Patient Cancellations and No-Shows at Outpatient Clinics," *Decision Support Systems*, vol. 57, no. 1, pp. 428–443, 2014.

[32] Y. Peng, E. Erdem, J. Shi, C. Masek, and P. Woodbridge, "Large-scale Assessment of Missed Opportunity Risks in a Complex Hospital Setting," *Informatics for Health and Social Care*, vol. 41, no. 2, pp. 112–127, 2016.

[33] O. Torres, M. B. Rothberg, J. Garb, O. Ogunneye, J. Onyema, and T. Higgins, "Risk Factor Model to Predict a Missed Clinic Appointment in an Urban, Academic, and Underserved Setting," *Population Health Management*, vol. 18, no. 2, pp. 131–136, 2015.

[34] L. G. Kalb, B. Freedman, C. Foster, D. Menon, R. Landa, and K. Kishfy, "Determinants of Appointment Absenteeism at an Outpatient Pediatric Autism Clinic," *Journal of Developmental and Behavioral Pediatrics*, vol. 33, no. 9, pp. 685–697, 2012.

[35] K.-D. Cheng, C.-J. Huang, H.-Y. Tsang, and C.-H. Lin, "Factors Related to Missed First Appointments After Discharge Among Patients with Schizophrenia in Taiwan," *Journal of the Formosan Medical Association*, vol. 113, no. 7, pp. 436–441, 2014.

[36] M. M. Coleman, L. N. Medford-Davis, O. H. Atassi, A. Siler-Fisher, and C. A. Reitman, "Injury Type and Emergency Department Management of Orthopedic Patients Influences Follow-up Rates," *Journal of Bone and Joint Surgery. American Volume*, vol. 96, no. 19, pp. 1650–1658, 2014.

[37] C. E. Mbada, O. Ajayi, O. B. Agbeja, K. A. Mbada, T. O. Awotidebe, and S. N. Oghumu, "Non-Attendance for Out-Patient Physiotherapy: Evaluation, Prediction and Physiotherapists' Perceptions—a Cross-Sectional Study," *African Journal of Physiotherapy and Rehabilitation Sciences*, vol. 5, no. 1, pp. 12–22, 2013.

[38] K. T. Kavanagh, T. R. Smith, G. S. Golden, N. P. Tate, and W. G. Hinkle, "Multivariate Analysis of Family Risk Factors in Predicting Appointment Attendance in a Pediatric Otology and Communication Clinic," *Journal of Health & Social Policy*, vol. 2, no. 3, pp. 85–102, 1991.

[39] P. Kheirkhah, Q. Feng, L. M. Travis, S. Tavakoli-Tabasi, and A. Sharafkhaneh, "Prevalence, predictors and economic consequences of no-show," *BMC Health Services Research*, vol. 16, p. 13, 2016. DOI: 10.1186/s12913-015-1240-5.

[40] A. Kempny, G. Diller, K. Dimopoulos, R. Alonso-Gonzalez, A. Uebing, and W. Li, "Determinants of outpatient clinic attendance amongst adults with congenital heart disease and outcome," *International Journal of Cardiology*, vol. 203, pp. 245–250, 2016. DOI: 10.1016/j.ijcard.2015.10.118.

[41] N. Weingarten, D. L. Meyer, and J. A. Schneid, "Failed appointments in residency practices: Who misses them and what providers are most affected?" *Journal of the American Board of Family Practice*, vol. 10, pp. 407–411, 1997.

[42] E. Kaplan-Lewis and S. Percac-Lima, "No-show to primary care appointments: Why patients do not come," *Journal of Primary Care & Community Health*, vol. 4, no. 4, pp. 251–255, 2013. DOI: 10.1177/2150131913498513.

[43] K. E. Lasser, I. L. Mintzer, A. Lambert, H. Cabral, and D. H. Bor, "Missed appointment rates in primary care: The importance of site of care," *Journal of Health Care for the Poor and Underserved*, vol. 16, pp. 475–486, 2005. DOI: 10.1353/hpu.2005.0050.

[44] R. Pillai, N. Bhangu, M. Narayanan, and W. Yoong, "A demographic study to profile non-attenders at a gynaecology outpatient clinic," *Journal of Obstetrics and Gynaecology*, vol. 32, pp. 156–158, 2012. DOI: 10.3109/01443615.2011.637132.

[45] A. J. Miller, E. Chae, E. Peterson, and A. B. Ko, "Predictors of repeated no-showing to clinic appointments," *American Journal of Otolaryngology*, vol. 36, no. 4, pp. 411–414, 2015. DOI: 10.1016/j.amjoto.2015.01.013.

[46] P. S. Whiting, S. E. Greenberg, R. V. Thakore, V. K. Alamanda, J. M. Ehrenfeld, and W. T. Obremskey, "What factors influence follow-up in orthopedic trauma surgery?" *Archives of Orthopaedic and Trauma Surgery*, vol. 135, no. 3, pp. 321–327, 2015. DOI: 10.1007/s00402-015-2151-8.

[47] K. L. Margolis, N. Lurie, P. G. McGovern, and J. S. Slater, "Predictors of failure to attend scheduled mammography appointments at a public teaching hospital," *Journal of General Internal Medicine*, vol. 8, no. 10, pp. 602–605, 1993. DOI: 10.1007/BF02599771.

[48] M. R. Partin, A. Gravely, Z. F. Gellad, S. Nugent, J. F. Burgess, and A. Shaukat, "Factors associated with missed and cancelled colonoscopy appointments at veterans health administration facilities," *Clinical Gastroenterology and Hepatology*, vol. 14, no. 2, pp. 259–267, 2016. DOI: 10.1016/j.cgh.2015.09.043.

[49] L. Goldman, R. Freidin, E. F. Cook, J. Eigner, and P. Grich, "A multivariate approach to the prediction of no-show behavior in a primary care center," *Archives of Internal Medicine*, vol. 142, pp. 563–567, 1982.

[50] A. Barksdale, J. Hackman, A. Bonham, and M. Gratton, "Cardiology clinic follow-up did not decrease return visits to the ed for chest pain patients," *American Journal of Emergency Medicine*, vol. 32, no. 9, pp. 1208–1211, 2014. DOI: 10.1016/j.ajem.2014.07.003.

[51] A. B. Adeponle, A. O. Obembe, and G. T. Suleiman, "Missed first appointments: Prevalence and associated factors in first-time attendees at an outpatient psychiatric clinic in nigeria," *Mental Health, Religion & Culture*, vol. 10, no. 6, pp. 609–620, 2007. DOI: 10.1080/13674670601034577.

[52] L. Sharp, S. Cotton, A. Thornton, N. Gray, M. Cruickshank, and D. Whynes, "Who defaults from colposcopy? a multi-centre, population-based, prospective cohort study of predictors of non-attendance for follow-up among women with low-grade abnormal cervical cytology," *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 165, no. 2, pp. 318–325, 2012. DOI: 10.1016/j.ejogrb.2012.08.007.

[53] M. E. Menendez and D. Ring, "Factors associated with non-attendance at a hand surgery appointment," *Hand*, vol. 10, no. 2, pp. 221–226, 2015. DOI: 10.1007/s11552-014-9674-4.

[54] A. H. Pang, S. Tso, G. S. Ungvari, H. Chiu, and T. Leung, "An audit study of defaulters of regular psychiatric outpatient appointments in hong kong," *International Journal of Social Psychiatry*, vol. 41, no. 1, pp. 103–107, 1995. DOI: 10.1177/002076409504100203.

[55] L. Traeger, C. O'Cleirigh, M. R. Skeer, K. H. Mayer, and S. A. Safren, "Risk factors for missed hiv primary care visits among men who have sex with men," *Journal of Behavioral Medicine*, vol. 35, no. 5, pp. 548–556, 2012. DOI: 10.1007/s10865-011-9383-2.

[56] J. Collins, N. Santamaria, and L. Clayton, "Why outpatients fail to attend their scheduled appointments: A prospective comparison of differences between attenders and non-attenders," *Australian Health Review*, vol. 26, no. 1, pp. 52–63, 2003. DOI: 10.1071/AH030052.

[57] B. T. Farid and E. Alapont, "Patients who fail to attend their first psychiatric outpatient appointment: Non-attendance or inappropriate referral?" *Journal of Mental Health*, vol. 2, no. 1, pp. 81–83, 1993. DOI: 10.3109/09638239309016959.

[58] A. J. Karter, M. M. Parker, H. H. Moffet, A. T. Ahmed, A. Ferrara, and J. Y. Liu, "Missed appointments and poor glycemic control: An opportunity to identify high-risk diabetic patients," *Medical Care*, vol. 42, no. 2, pp. 110–115, 2004. DOI: 10.1097/01.mlr.0000108744.24921.95.

[59] S. Coodin, D. Staley, B. Cortens, R. Desrochers, and S. McLandress, "Patient factors associated with missed appointments in persons with schizophrenia," *Canadian Journal of Psychiatry*, vol. 49, pp. 145–148, 2004. DOI: 10.1177/070674370404900210.

[60] V. Lotfi and E. Torres, “Improving an outpatient clinic utilization using decision analysis-based patient scheduling,” *Socio-Economic Planning Sciences*, vol. 48, no. 2, pp. 115–126, 2014. DOI: 10.1016/j.seps.2014.02.003.

[61] G. R. Kruse, B. M. Rohland, and X. Wu, “Factors associated with missed first appointments at a psychiatric clinic,” *Psychiatric Services*, vol. 53, pp. 1173–1176, 2002.

[62] M. Grunebaum, P. Luber, M. Callahan, A. C. Leon, M. Olfson, and L. Portera, “Predictors of missed appointments for psychiatric consultations in a primary care clinic,” *Psychiatric Services*, vol. 47, no. 8, pp. 848–852, 1996. DOI: 10.1176/ps.47.8.848.

[63] M. M. N. El-Din, F. N. Al-Shakhs, and S. S. Al-Oudah, “Missed appointments at a university hospital in eastern saudi arabia: Magnitude and association factors,” *Journal of the Egyptian Public Health Association*, vol. 83, no. 5-6, pp. 415–433, 2008.

[64] L. M. Guzek, W. F. Fadel, and M. R. Golomb, “A pilot study of reasons and risk factors for no-show in a pediatric neurology clinic,” *Journal of Child Neurology*, vol. 30, no. 11, pp. 1295–1299, 2015. DOI: 10.1177/0883073814555188.

[65] E. Ahmadi, A. Garcia-Arce, D. T. Masel, E. Reich, J. Puckey, and R. Maff, “A metaheuristic-based stacking model for predicting the risk of patient no-show and late cancellation for neurology appointments,” *IISE Transactions on Healthcare Systems Engineering*, vol. 9, pp. 272–291, 2019.

[66] M. Alloghani, A. J. Aljaaf, D. Al-Jumeily, A. Hussain, C. Mallucci, and J. Mustafina, “Data science to improve patient management system,” in *Proceedings of the 2018 11th International Conference on Developments in eSystems Engineering (DeSE)*, Cambridge, UK, pp. 27–30. DOI: 10.1109/DeSE.2018.00013.

[67] G. Lee, S. Wang, F. Dipuro, *et al.*, “Leveraging on predictive analytics to manage clinic no show and improve accessibility of care,” in *Proceedings of the 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Tokyo, Japan, pp. 429–438. DOI: 10.1109/DSAA.2017.93.

[68] H. McLeod, G. Heath, E. Cameron, G. Debelle, and C. Cummins, “Introducing consultant outpatient clinics to community settings to improve access to pediatrics: An observational impact study,” *BMJ Quality & Safety*, vol. 24, no. 6, pp. 377–384, 2015. DOI: 10.1136/bmjqs-2014-003593.

[69] R. E. Mark, P. L. Klarenbeek, G. J. M. Rutten, and M. M. Sitskoorn, "Why don't neurosurgery patients return for neuropsychological follow-up? predictors for voluntary appointment keeping and reasons for cancellation," *Clinical Neuropsychologist*, vol. 28, no. 1, pp. 49–64, 2014. DOI: 10.1080/13854046.2013.869510.

[70] H. G. Dove and K. C. Schneider, "The usefulness of patients' individual characteristics in predicting no-shows in outpatient clinics," *Medical Care*, vol. 19, no. 8, pp. 734–740, 1981. DOI: 10.1097/00005650-198108000-00004.

[71] J. V. Dervin, D. L. Stone, and C. H. Beck, "The no-show patient in the model family practice unit," *Journal of Family Practice*, vol. 7, no. 6, pp. 1177–1180, 1978.

[72] V. J. Lee, A. Earnest, M. I. Chen, and B. Krishnan, "Predictors of failed attendances in a multi-specialty outpatient centre using electronic databases," *BMC Health Services Research*, vol. 5, p. 51, 2005.

[73] N. L. Ma, S. Khataniar, D. Wu, and S. S. Y. Ng, "Predictive analytics for outpatient appointments," in *Proceedings of the 2014 International Conference on Information Science & Applications (ICISA)*, Amman, Jordan, pp. 1–4. DOI: 10.1109/ICISA.2014.6847410.

[74] E. S. Gromisch, A. P. Turner, S. L. Leipertz, J. Beauvais, and J. K. Haselkorn, "Who is not coming to clinic? a predictive model of excessive missed appointments in persons with multiple sclerosis," *Multiple Sclerosis and Related Disorders*, vol. 38, p. 101 513, 2020. DOI: 10.1016/j.msard.2020.101513.

[75] M. Aladeemy, L. Adwan, A. Booth, M. T. Khasawneh, and S. Poranki, "New feature selection methods based on opposition-based learning and self-adaptive cohort intelligence for predicting patient no-shows," *Applied Soft Computing*, vol. 86, p. 105 866, 2020. DOI: 10.1016/j.asoc.2019.105866.

[76] T. N. Lehmann, A. Aebi, D. Lehmann, M. B. Olivet, and H. Stalder, "Missed appointments at a swiss university outpatient clinic," *Public Health*, vol. 121, no. 10, pp. 790–799, 2007. DOI: 10.1016/j.puhe.2006.12.008.

[77] A. Al-Khadra, G. Magbool, L. Wosornu, S. Al-Awdah, H. Qutub, and R. Al-Khatib, "Why do cardiology out-patient appointments fail in saudi arabia?" *International Journal for Quality in Health Care*, vol. 4, no. 4, pp. 305–310, 1992. DOI: 10.1093/intqhc/4.4.305.

[78] F. J. Spikmans, J. Brug, M. M. Doven, H. M. Kruizenga, G. H. Hofsteenge, and V. Bokhorst-van der Schueren, "Why do diabetic patients not attend appointments with their dietitian?" *Journal of Human Nutrition and Dietetics*, vol. 16, no. 2, pp. 151–158, 2003. DOI: 10.1046/j.1365-277X.2003.00431.x.

[79] R. Lalloo and J. M. McDonald, "Appointment attendance at a remote rural dental training facility in australia," *BMC Oral Health*, vol. 13, p. 33, 2013. DOI: 10.1186/1472-6831-13-33.

[80] L. Livianos-Aldana, M. Vila-Gomez, L. Rojo-Moreno, and M. A. Luengo-Lopez, "Patients who miss initial appointments in community psychiatry? a spanish community analysis," *International Journal of Social Psychiatry*, vol. 45, no. 3, pp. 198–206, 1999. DOI: 10.1177/002076409904500305.

[81] W. Dickey and J. I. Morrow, "Can outpatient non-attendance be predicted from the referral letter? an audit of default at neurology clinics," *Journal of the Royal Society of Medicine*, vol. 84, no. 10, pp. 662–663, 1991.

[82] K. R. Mathu-Muju, H. F. Li, J. Hicks, D. A. Nash, A. Kaplan, and H. M. Bush, "Identifying demographic variables related to failed dental appointments in a university hospital-based residency program," *Pediatric Dentistry*, vol. 36, no. 3, pp. 296–301, 2014.

[83] N. J. Wang and G. Aspelund, "Children who break dental appointments," *European Archives of Paediatric Dentistry*, vol. 10, no. 1, pp. 11–14, 2009. DOI: 10.1007/BF03262665.

[84] L. F. Dantas, S. Hamacher, F. L. C. Oliveira, S. D. Barbosa, and F. Viegas, "Predicting patient no-show behavior: A study in a bariatric clinic," *Obesity Surgery*, vol. 29, no. 1, pp. 40–47, 2019. DOI: 10.1007/s11695-018-3493-2.

[85] S. AlMuhaideb, O. Alswailem, N. Alsubaie, I. Ferwana, and A. Alnajem, "Prediction of Hospital No-show Appointments Through Artificial Intelligence Algorithms," *Annals of Saudi Medicine*, vol. 39, no. 6, pp. 373–381, 2019.

[86] C. Elvira, A. Ochoa, J. C. Gonzalvez, and F. Mochon, "Machine-learning-based no-show prediction in outpatient visits," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, p. 29, 2018, Verify issue and full pagination.

[87] K. Topuz, H. Uner, A. Oztekin, and M. B. Yildirim, "Predicting pediatric clinic no-shows: A decision analytic framework using elastic net and bayesian belief network," *Annals of Operations Research*, vol. 263, no. 1–2, pp. 479–499, 2018. DOI: 10.1007/s10479-017-2704-3.

[88] X. Ding, Z. F. Gellad, C. Mather, *et al.*, "Designing risk prediction models for ambulatory no-shows across different specialties and clinics," *Journal of the American Medical Informatics Association*, vol. 25, no. 7, pp. 924–930, 2018. DOI: 10.1093/jamia/ocy007.

[89] H. B. Harvey, C. Liu, J. Ai, *et al.*, "Predicting no-shows in radiology using regression modeling of data available in the electronic medical record," *Journal of the American College of Radiology*, vol. 14, no. 10, pp. 1303–1309, 2017. DOI: 10.1016/j.jacr.2017.05.012.

[90] S. Carrillo González, J. Lorduy Gómez, V. Castro Bocanegra, Á. Quintana Salcedo, R. Muñoz Baldiris, and M. Ramos Herrera, "Factors influencing the cancellation of scheduled surgeries: A literature review," *Repertorio de Medicina y Cirugía*, vol. 33, no. 1, pp. 27–32, 2024.

[91] R. A. Bush, V. M. Vemulakonda, S. T. Corbett, and G. J. Chiang, "Can we predict a national profile of non-attendance paediatric urology patients: A multi-institutional electronic health record study," *Informatics in Primary Care*, vol. 21, no. 3, pp. 132–138, 2014.

[92] V. Chariatte, P. Michaud, A. Berchtold, C. Akré, and J. Suris, "Missed appointments in an adolescent outpatient clinic: Descriptive analyses of consultations over eight years," *Swiss Medical Weekly*, vol. 137, no. 33-34, pp. 677–681, 2007.

[93] M. J. McMullen and P. A. Netland, "Lead time for appointment and the no-show rate in an ophthalmology clinic," *Clinical Ophthalmology*, vol. 9, pp. 513–516, 2015.

[94] B. A. Majeroni, T. Cowan, J. Osborne, and R. P. Graham, "Missed appointments and medicaid managed care," *Archives of Family Medicine*, vol. 5, pp. 507–511, 1996.

[95] W. Hamilton, A. Round, and D. Sharp, "Patient, hospital, and general practitioner characteristics associated with non-attendance: A cohort study," *British Journal of General Practice*, vol. 52, no. 477, pp. 317–319, 2002.

[96] L. A. Adams, J. Pawlik, and G. M. Forbes, "Nonattendance at outpatient endoscopy," *Endoscopy*, vol. 36, no. 5, pp. 402–404, 2004. DOI: 10.1055/s-2004-814320.

[97] S. Hampton-Robb, R. C. Qualls, and W. C. Compton, "Predicting first-session attendance: The influence of referral source and client income," *Psychotherapy Research*, vol. 13, no. 2, pp. 223–233, 2003. DOI: 10.1093/ptr/kpg021.

[98] M. Matas, D. Staley, and W. Griffin, "A profile of the noncompliant patient: A thirty-month review of outpatient psychiatry referrals," *General Hospital Psychiatry*, vol. 14, no. 2, pp. 124–130, 1992. DOI: 10.1016/0163-8343(92)90025-C.

[99] R. Weinerman, V. Glossop, R. Wong, L. Robinson, K. White, and R. Kamil, "Time of day influences nonattendance at urgent short-term mental health unit in victoria, british columbia," *Canadian Journal of Psychiatry*, vol. 48, no. 5, pp. 342–344, 2003. DOI: 10.1177/070674370304800517.

[100] K. L. Hon, T. F. Leung, Y. Wong, K. C. Ma, and T. F. Fok, "Reasons for new referral non-attendance at a pediatric dermatology center: A telephone survey," *Journal of Dermatological Treatment*, vol. 16, no. 2, pp. 113–116, 2005. DOI: 10.1080/09546630510032430.

[101] A. H. Zailinawati, C. J. Ng, and H. Nik-Sherina, "Why do patients with chronic illnesses fail to keep their appointments? a telephone interview," *Asia-Pacific Journal of Public Health*, vol. 17, no. 2, pp. 10–15, 2005. DOI: 10.1177/101053950501700203.

[102] P. H. Dyer, C. E. Lloyd, R. J. Lancashire, S. C. Bain, and A. H. Barnett, "Factors associated with clinic non-attendance in adults with type 1 diabetes mellitus," *Diabetic Medicine*, vol. 15, no. 5, pp. 339–343, 1998. DOI: 10.1002/(SICI)1096-9136(199805)15:5<339::AID-DIA572>3.0.CO;2-#.

[103] S. B. Cashman, J. A. Savageau, C. A. Lemay, and W. Ferguson, "Patient health status and appointment keeping in an urban community health center," *Journal of Health Care for the Poor and Underserved*, vol. 15, no. 3, pp. 474–488, 2004. DOI: 10.1353/hpu.2004.0045.

[104] K. H. Carlsen, K. M. Carlsen, and J. Serup, "Non-attendance rate in a danish university clinic of dermatology," *Journal of the European Academy of Dermatology and Venereology*, vol. 25, no. 11, pp. 1269–1274, 2011. DOI: 10.1111/j.1468-3083.2010.03924.x.

[105] M. Gordon, K. M. Antshel, L. Lewandowski, and D. Seigers, "Economic grand rounds: Predictors of missed appointments over the course of child mental health treatment," *Psychiatric Services*, vol. 61, no. 6, pp. 657–659, 2010. DOI: 10.1176/ps.2010.61.6.657.

[106] S. Snowden, P. Weech, R. McClure, S. Smye, and P. Dear, "A neural network to predict attendance of paediatric patients at outpatient clinics," *Neural Computing and Applications*, vol. 3, pp. 234–241, 1995. DOI: 10.1007/BF01411359.

[107] M. U. Ahmad, A. Zhang, and R. Mhaskar, "A predictive model for decreasing clinical no-show rates in a primary care setting," *International Journal of Healthcare Management*, vol. 12, no. 1, pp. 1–7, 2019. DOI: 10.1080/20479700.2019.1617206.

[108] M. K. Daniels and S. Jung, "Missed initial appointments at an outpatient forensic psychiatric clinic," *Journal of Forensic Psychiatry & Psychology*, vol. 20, no. 6, pp. 964–973, 2009. DOI: 10.1080/14789940903174162.

[109] L. A. Siminoff, L. R. Hausmann, and S. Ibrahim, "Barriers to obtaining diagnostic testing for coronary artery disease among veterans," *American Journal of Public Health*, vol. 98, no. 12, pp. 2207–2213, 2008. DOI: 10.2105/AJPH.2007.125377.

[110] M. O. Guay, S. Tanzi, M. T. Arregui, G. Chisholm, M. de la Cruz, and E. Bruera, "Characteristics and outcomes of advanced cancer patients who miss outpatient supportive care consult appointments," *Supportive Care in Cancer*, vol. 22, no. 11, pp. 2869–2874, 2014. DOI: 10.1007/s00520-014-2275-7.

[111] D. Giunta, A. Briatore, A. Baum, D. Luna, G. Waisman, and F. G. de Quiros, "Factors associated with nonattendance at clinical medicine scheduled outpatient appointments in a university general hospital," *Patient Preference and Adherence*, vol. 7, pp. 1163–1170, 2013. DOI: 10.2147/PPA.S49006.

[112] E. Y. Yoon, M. M. Davis, J. Van Cleave, S. Maheshwari, and M. D. Cabana, "Factors associated with non-attendance at pediatric subspecialty asthma clinics," *Journal of Asthma*, vol. 42, no. 7, pp. 555–559, 2005. DOI: 10.1081/JAS-200067650.

[113] M. J. Storrs, H. M. Ramov, and R. Lalloo, "An investigation into patient non-attendance and use of a short-message reminder system at a university dental clinic," *Journal of Dental Education*, vol. 80, no. 1, pp. 30–39, 2016.

[114] A. D. Cohen, J. Dreiher, D. A. Vardy, and D. Weitzman, "Nonattendance in a dermatology clinic–a large sample analysis," *Journal of the European Academy of Dermatology and Venereology*, vol. 22, no. 10, pp. 1178–1183, 2008. DOI: 10.1111/j.1468-3083.2008.02757.x.

[115] M. A. Praveena, J. S. Krupa, and S. SaiPreethi, "Statistical analysis of medical appointments using decision tree," in *Proceedings of the 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, vol. 1, Chennai, India, pp. 59–64. DOI: 10.1109/ICONSTEM.2019.8918914.

[116] S. R. Devasahay, S. Karpagam, and N. L. Ma, "Predicting appointment misses in hospitals using data analytics," *Mhealth*, vol. 3, p. 12, 2017.

[117] K. J. Glowacka, R. M. Henry, and J. H. May, "A hybrid data mining/simulation approach for modelling outpatient no-shows in clinic scheduling," *Journal of the Operational Research Society*, vol. 60, no. 8, pp. 1056–1068, 2009. DOI: 10.1057/palgrave.jors.2602626.

[118] X. Qu, R. L. Rardin, L. Tieman, *et al.*, “A statistical model for the prediction of patient non-attendance in a primary care clinic,” in *Proceedings of the IIE Annual Conference*, Orlando, FL, USA, p. 1.

[119] K. Pehr, “No show: Incidence of nonattendance at a dermatology practice in a single universal payer model,” *Journal of Cutaneous Medicine and Surgery*, vol. 11, no. 1, pp. 53–56, 2007. DOI: `10.2310/7750.2006.00024`.

[120] J. Dreiher, A. Goldbart, J. Hershkovich, D. A. Vardy, and A. D. Cohen, “Factors associated with non-attendance at pediatric allergy clinics,” *Pediatric Allergy and Immunology*, vol. 19, no. 6, pp. 559–563, 2008. DOI: `10.1111/j.1399-3038.2007.00701.x`.

[121] I. Mohammadi, H. Wu, A. Turkcan, T. Toscos, and B. N. Doebbeling, “Data Analytics and Modeling for Appointment No-show in Community Health Centers,” *Journal of Primary Care & Community Health*, vol. 9, pp. 1–11, 2018.

[122] A. T. Machado, M. A. Werneck, S. D. Lucas, and M. H. Abreu, “Who did not appear? first dental visit absences in secondary care on a major brazilian city: A cross-sectional study,” *Ciencia & Saude Coletiva*, vol. 20, no. 1, pp. 289–298, 2015. DOI: `10.1590/1413-81232015201.21162013`.

[123] S. Srinivas and A. R. Ravindran, “Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: A prescriptive analytics framework,” *Expert Systems with Applications*, vol. 102, pp. 245–261, 2018. DOI: `10.1016/j.eswa.2018.02.022`.

[124] R. J. Mieloszyk, J. I. Rosenbaum, P. Bhargava, and C. S. Hall, “Predictive modeling to identify scheduled radiology appointments resulting in non-attendance in a hospital setting,” in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju Island, Korea, pp. 2618–2621. DOI: `10.1109/EMBC.2017.8037338`.

[125] S. L. Chua and W. L. Chow, “Development of predictive scoring model for risk stratification of no-show at a public hospital specialist outpatient clinic,” *Proceedings of Singapore Healthcare*, vol. 28, no. 2, pp. 96–104, 2019. DOI: `10.1177/2010105818809029`.

[126] A. D. Cohen, D. M. Kaplan, M. Kraus, E. Rubinshtein, and D. A. Vardy, “Nonattendance of adult otolaryngology patients for scheduled appointments,” *Journal of Laryngology & Otology*, vol. 121, no. 3, pp. 258–261, 2007. DOI: `10.1017/S0022215106004492`.

[127] D. Kogan-Liberman, Y. Rivas, J. Thompson, and G. Tomer, "Improving nonattendance at outpatient pediatric endoscopy unit of a tertiary center," *Journal of Pediatric Gastroenterology and Nutrition*, vol. 61, no. 2, pp. 234–237, 2015. DOI: 10.1097/MPG.0000000000000773.

[128] E. S. Kazarian, F. S. Carreira, N. W. Toribara, and T. D. Denberg, "Colonoscopy completion in a large safety net health care system," *Clinical Gastroenterology and Hepatology*, vol. 6, no. 4, pp. 438–442, 2008. DOI: 10.1016/j.cgh.2007.12.023.

[129] D. M. Blumenthal, G. Singal, S. S. Mangla, E. A. Macklin, and D. C. Chung, "Predicting non-adherence with outpatient colonoscopy using a novel electronic tool that measures prior non-adherence," *Journal of General Internal Medicine*, vol. 30, no. 5, pp. 724–731, 2015. DOI: 10.1007/s11606-014-3145-y.

[130] M. Dashtban and W. Li, "Deep learning for predicting non-attendance in hospital outpatient appointments," in *Proceedings of the 52nd Annual Hawaii International Conference on System Sciences (HICSS)*, Maui, HI, USA, pp. 3731–3740. DOI: 10.24251/HICSS.2019.452.

[131] J. D. Ross, A. McIver, A. Blakely, J. Dalrymple, W. Peacock, and C. Wallis, "Why do patients default from follow-up at a genitourinary clinic? a multivariate analysis," *Genitourinary Medicine*, vol. 71, pp. 393–395, 1995.

[132] L. Bofill, D. Waldrop-Valverde, L. Metsch, M. Pereyra, and M. A. Kolber, "Demographic and psychosocial factors associated with appointment attendance among hiv-positive outpatients," *AIDS Care*, vol. 23, no. 10, pp. 1219–1225, 2011. DOI: 10.1080/09540121.2011.564112.

[133] M. Deina, C. Opara, D. Tsapepas, T. Ahuja, and J. S. Ancker, "Decision Analysis Framework for Predicting No-shows to Appointments Using Machine Learning Algorithms," *BMC Health Services Research*, vol. 24, no. 1, p. 37, 2024.

[134] S. Chakraborty, K. Muthuraman, and M. Lawley, "Sequential Clinical Scheduling with Patient No-Shows and General Service Time Distributions," *IIE Transactions*, vol. 42, no. 5, pp. 354–366, 2010.

[135] J. Feldman, N. Liu, H. Topaloglu, and S. Ziya, "Appointment Scheduling Under Patient Preference and No-Show Behavior," *Operations Research*, vol. 62, no. 4, pp. 794–811, 2014.

[136] C. Zheng and P. Zhao, "Cost-Sensitive Hierarchical Classification for Imbalanced Classes," *Applied Intelligence*, vol. 50, no. 5, pp. 1485–1498, 2020.

[137] D. A. Pereira, E. L. Costa, and C. N. J. Silla, "Resampling Strategies for Imbalanced Datasets in Hierarchical Classification," *Data Mining and Knowledge Discovery*, vol. 35, pp. 1950–1984, 2021.

[138] J.-N. Eckardt, R. Schmitz, C. Trautwein, and S. Heßling, "Semi-supervised Learning in Cancer Diagnostics," *Frontiers in Oncology*, vol. 12, p. 960 984, 2022.

[139] T. Oikonomidi, G. Norman, L. McGarrigle, J. Stokes, S. N. van der Veer, and D. Dowding, "Predictive Model-Based Interventions to Reduce Outpatient No-Shows: A Rapid Systematic Review," *Journal of the American Medical Informatics Association*, vol. 30, no. 3, pp. 559–569, 2022.

[140] Y. Tarabichi, J. Higginbotham, N. Riley, D. C. Kaelber, B. Watts, *et al.*, "Reducing Disparities in No Show Rates Using Predictive Model-Driven Live Appointment Reminders for At-Risk Patients: a Randomized Controlled Quality Improvement Initiative," *Journal of General Internal Medicine*, vol. 38, no. 13, pp. 2921–2927, 2023.

[141] Y.-L. Huang and D. A. Hanauer, "Patient No-Show Predictive Model Development Using Multiple Data Sources for an Effective Overbooking Approach," *Applied Clinical Informatics*, vol. 5, no. 3, pp. 836–860, 2014.

[142] M. R. M. Habibi, F. M. Abadi, H. Tabesh, *et al.*, "Evaluation of No-Show Rate in Outpatient Clinics With Open Access Scheduling System: A Systematic Review," *Health Science Reports*, vol. 7, no. 7, e2160, 2024.

[143] P. G. Shekelle, M. M. Begashaw, I. M. Miake-Lye, M. Booth, B. Myers, and A. Renda, "Effect of Interventions for Non-Emergent Medical Transportation: A Systematic Review and Meta-Analysis," *BMC Public Health*, vol. 22, p. 799, 2022.

[144] G. Leibner, S. Brammli-Greenberg, and A. Israeli, "To Charge or Not to Charge: Reducing Patient No-Show," *Israel Journal of Health Policy Research*, vol. 12, no. 27, 2023.

[145] D. L. Nguyen, B. R. Graham, M. D'Amico, J. T. Heywood, and B. Littenberg, "Missed Appointments in Resident Continuity Clinic: Patient–Provider Relationship Discontinuity and Its Consequences," *Journal of Graduate Medical Education*, vol. 3, no. 3, pp. 350–355, 2011.

[146] Y.-L. Huang, A. H. Bryce, T. Culbertson, *et al.*, "Alternative Outpatient Chemotherapy Scheduling Method to Improve Patient Service Quality and Nurse Satisfaction," *Journal of Oncology Practice*, vol. 14, no. 2, e82–e91, 2018.

[147] A. M. Chen, "Socioeconomic and Demographic Factors Predictive of Missed Appointments in Outpatient Radiation Oncology: An Evaluation of Access," *Frontiers in Health Services*, vol. 3, p. 1 288 329, 2023.

[148] A. G. Bean and J. Talaga, "Appointment breaking: Causes and solutions," *Journal of Health Care Marketing*, vol. 12, no. 4, pp. 14–25, 1992.

[149] E. Kaplan-Lewis and S. Percac-Lima, "Barriers to appointment keeping in patients with cancer: Perspectives of underserved populations," *Journal of General Internal Medicine*, vol. 28, no. 5, pp. 667–673, 2013. DOI: `10.1007/s11606-012-2301-3`.

[150] D. J. Rumala, "How You Split Matters: Data Leakage and Subject Characteristics in Longitudinal MRI Analysis," 2023. arXiv: `2309.00350`.

[151] H. Harutyunyan, H. Khachatrian, D. C. Kale, G. Ver Steeg, and A. Galstyan, "Multitask Learning and Benchmarking with Clinical Time Series Data," *Scientific Data*, vol. 6, no. 1, p. 96, 2019.

[152] F. Gao, P. J. Wallace, H. Wang, and E. P. Xing, "Hierarchical Structured Model for Predicting Childhood Obesity Using Electronic Health Records," *Journal of Biomedical Informatics*, vol. 94, p. 103 184, 2019.

[153] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Belmont, CA: Chapman and Hall/CRC, 2017, ISBN: 9780534980542.

[154] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[155] D. Holzmüller, L. Grinsztajn, and I. Steinwart, "Better by Default: Strong Pre-Tuned MLPs and Boosted Trees on Tabular Data," in *Advances in Neural Information Processing Systems*, A. Globerson, L. Mackey, D. Belgrave, *et al.*, Eds., vol. 37, 2024, pp. 26 577–26 658.

[156] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLOS ONE*, vol. 10, no. 3, e0118432, 2015.

[157] E. Choi, M. T. Bahadori, L. Song, W. F. Stewart, and J. Sun, "Doctor AI: Predicting Clinical Events via Recurrent Neural Networks," in *Proceedings of the 1st Machine Learning for Healthcare Conference*, ser. Proceedings of Machine Learning Research, vol. 56, 2016, pp. 301–318.

[158] Z. C. Lipton, D. Kale, C. Elkan, and R. Wetzel, "Learning to Diagnose with LSTM Recurrent Neural Networks," 2016. arXiv: `1511.03677`.

[159] A. Bagnall and G. C. Cawley, "On the Use of Default Parameter Settings in the Empirical Evaluation of Classification Algorithms," 2017. arXiv: 1703.06777.

[160] S. Kapoor and A. Narayanan, "Leakage and the Reproducibility Crisis in ML-based Science," *Patterns*, vol. 4, no. 9, p. 100 804, 2023. arXiv: 2207.07048.

[161] T. Oikonomidi, G. Norman, L. McGarrigle, J. Stokes, S. N. van der Veer, and D. Dowding, "Predictive Model-Based Interventions to Reduce Outpatient No-Shows: A Rapid Systematic Review," *Journal of the American Medical Informatics Association*, vol. 30, no. 3, pp. 559–569, 2023.

[162] J. Liu, J. Xie, K. K. Yang, and Z. Zheng, "Effects of Rescheduling on Patient No-Show Behavior in Outpatient Clinics," *Manufacturing & Service Operations Management*, vol. 21, no. 4, pp. 780–797, 2019.

[163] S. R. Finkelstein, N. Liu, B. Jani, and D. Rosenthal, "Appointment reminder systems and patient preferences: Patient technology usage and familiarity with other service providers as predictive variables," *Health Informatics Journal*, vol. 19, no. 2, pp. 79–90, 2013. DOI: 10.1177/1460458212458429.

[164] P. Hasvold and R. Wootton, "Use of Telephone and SMS Reminders to Improve Attendance at Hospital Appointments: A Systematic Review," *Journal of Telemedicine and Telecare*, vol. 17, no. 7, pp. 358–364, 2011.

[165] J. F. Steiner, M. R. Shainline, J. Z. Dahlgren, A. Kroll, and S. Xu, "Optimizing number and timing of appointment reminders: A randomized trial," *American Journal of Managed Care*, vol. 24, no. 8, pp. 377–384, 2018. DOI: 10.37765/ajmc.2018.377.

[166] K. Werner, S. A. Alsuhaibani, R. F. Alsukait, *et al.*, "Behavioural economic interventions to reduce health care appointment non-attendance: A systematic review and meta-analysis," *BMC Health Services Research*, vol. 23, no. 1136, 2023. DOI: 10.1186/s12913-023-10059-9.

[167] S. J. Shah, P. R. Cronin, C. S. Hong, *et al.*, "Targeted Reminder Phone Calls to Patients at High Risk of No-Show for Primary Care Appointment: A Randomized Trial," *Journal of General Internal Medicine*, vol. 31, no. 12, pp. 1460–1466, 2016.

[168] J. Car, C. Gurol-Urganci, T. de Jongh, I. Vodopivec-Jamsek, and R. Atun, "Mobile Phone Messaging Reminders for Attendance at Health Care Appointments," *The Cochrane Database of Systematic Reviews*, no. 12, p. CD007458, 2013.

[169] K. Toker, K. Atas, A. Mayadağlı, Z. Görmezoğlu, I. Tuncay, and R. Kazancıoğlu, "A solution to reduce the impact of patients' no-show behavior on hospital operating costs: Artificial intelligence-based appointment system," *Healthcare*, vol. 12, no. 21, p. 2161, 2024. DOI: 10.3390/healthcare12212161.

[170] E. P. Greenup and D. Best, "Systematic review and meta-analysis of no show or non-attendance rates among telehealth and in-person models of care," *BMC Health Services Research*, vol. 25, p. 663, 2025. DOI: 10.1186/s12913-025-12826-2.

[171] S. Ganeshan, A. W. Liu, A. Kroeger, *et al.*, "An electronic health record–based automated self-rescheduling tool to improve patient access: Retrospective cohort study," *Journal of Medical Internet Research*, vol. 26, e52071, 2024. DOI: 10.2196/52071.

[172] Y. M. AlSerkal, N. M. Ibrahim, A. S. Alsereidi, *et al.*, "Real-time analytics and ai for managing no-show appointments in primary health care in the united arab emirates: Before-and-after study," *JMIR Formative Research*, vol. 9, e64936, 2025. DOI: 10.2196/64936.

[173] A. Shetty, H. Groenevelt, and V. Tilson, "Intraday dynamic rescheduling under patient no-shows," *Health Care Management Science*, vol. 26, no. 3, pp. 583–598, 2023. DOI: 10.1007/s10729-023-09643-6.

[174] S. Atta, R. B. Brown, L. M. Wasser, *et al.*, "Effect of a patient portal reminder message after no-show on appointment reattendance in ophthalmology: A randomized clinical trial," *American Journal of Ophthalmology*, vol. 263, pp. 93–98, 2024. DOI: 10.1016/j.ajo.2024.02.026.

[175] S. Yu, V. G. Kulkarni, and V. Deshpande, "Appointment scheduling for a health care facility with series patients," *Production and Operations Management*, vol. 29, no. 2, pp. 388–409, 2019. DOI: 10.1111/poms.13117.

[176] C. DuMontier, K. Rindfleisch, J. Pruszynski, and J. J. I. Frey, "A multi-method intervention to reduce no-shows in an urban residency clinic," *Family Medicine*, vol. 45, no. 9, pp. 634–641, Oct. 2013.

[177] A. B. Senderey, T. Kornitzer, G. Lawrence, *et al.*, "It's How you Say It: Systematic A/B Testing of Digital Messaging Cut Hospital No-Show Rates," *PLOS ONE*, vol. 15, no. 6, e0234817, 2020.

[178] M. Alturbag, "Effectiveness of Personalised Phone Calls and Short Message Service Reminders in Improving Patient Attendance at a Radiology Department," *Cureus*, vol. 16, e53640, 2024.

[179] P. K. Betancor, D. Boehringer, J. F. Jordan, *et al.*, "Efficient Patient Care in the Digital Age: Impact of Online Appointment Scheduling in a Medical Practice and a University Hospital on the No-Show-Rate," *Frontiers in Digital Health*, vol. 3, p. 117, 2025.