Leveraging Health Data and Machine Learning to Predict No-Show Propensity Factors in Glaucoma Patients: A Health Informatics Approach to Improve Clinic Efficiency and Patient Care

Hai-Wei Liang, Ph.D[1], Andrew M. Williams, MD[1]
[1]Department of Ophthalmology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, U.S.

University of Pittsburgh School of Medicine

UPMC LIFE CHANGING MEDICINE

Introduction

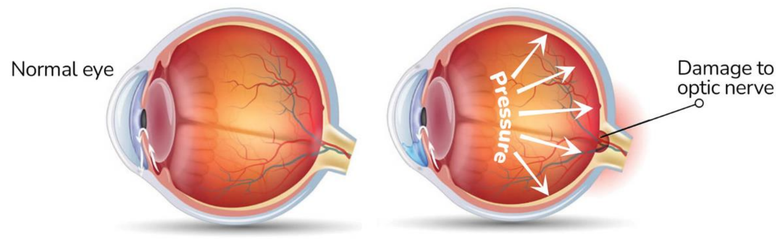

Fig 1. Glaucoma introduction

- Glaucoma is a group of eye conditions that can cause damage to the optic nerve, leading to vision loss or blindness if left untreated
- Missed appointments ("no-shows") disrupt treatment and threaten eye health
- Traditional no-show predictions treat it as a binary outcome (yes/no), useful for hospital resource planning but less informative for individualized patient support
- **The No-Show Propensity Factor (NSPF)** was calculated to account for variability in no-show rates and weighted by appointment frequency, reflecting long-term no-show patterns
- Goal:
  - Target patients needing support
  - Determine effective interventions, such as social services or tailored engagement strategies, to improve care continuity

Objectives

- Develop algorithms to predict no-show propensity for new patients
- Identify key clinical and demographic risk factors informs targeted interventions

Methods

- Study design: Retrospective cohort study
- Time frame: Jan 2013–Dec 2023
- Participants: UPMC EHRs for ophthalmology encounters identified glaucoma patients by billing codes (N=8504)
- Outcome: classified as low or high based on a 31.4% no-show rate reported for resident ophthalmology clinics

$$NSPF = \ln\left[\frac{\frac{(Ns*N_R)+H_N}{Ns+Hc+H_N}}{N_R}\right]$$

N_S = NSPF smoothing factor (set to 10)
N_R = Average no-show rate of the study cohort
H_N = Total number of no-shows
H_C = Total number of completed visits

- Data analysis
  - Prediction:
    - Machine learning models (logistic regression, random forest, SVM, XGBoost) with 10-fold cross-validation assessed performance based on sensitivity

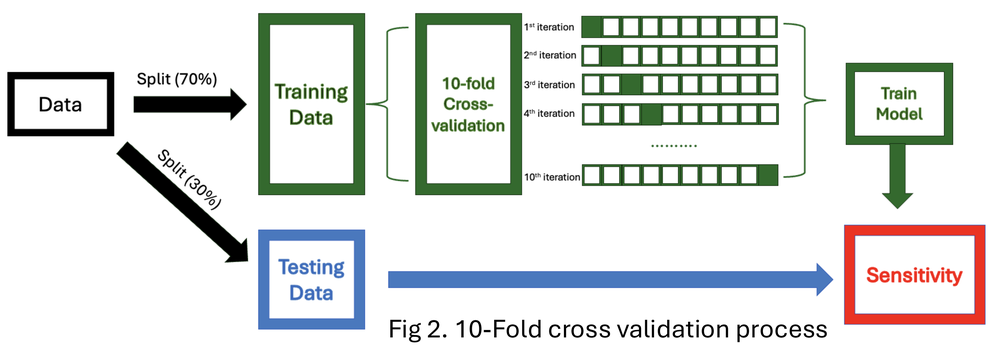

Fig 2. 10-Fold cross validation process

- Statistical analysis:
  - Log-logistic regression were used to calculate relative risks (RR) for baseline characteristics and NSPF, highlighting key factors associated with poor appointment compliance

Results

- Among 8,504 unique glaucoma patients with 163,385 appointments, 7,400 (4.5%) were no-shows.
- Baseline Characteristics:
  1. Majority were >60 years old (78.6%), White (71.7%), and female (53.2%)
  2. Nearly half were not in a committed relationship (48.4%) and publicly insured (56.5%)
  3. 30% lived in areas with a high Distressed Community Index
- Baseline Clinical Characteristics:
  1. Most were diagnosed with Primary Open-Angle Glaucoma (85%)
  2. 47.8% had moderate to severe glaucoma
  3. 10% were blind (logMAR ≥1.0)

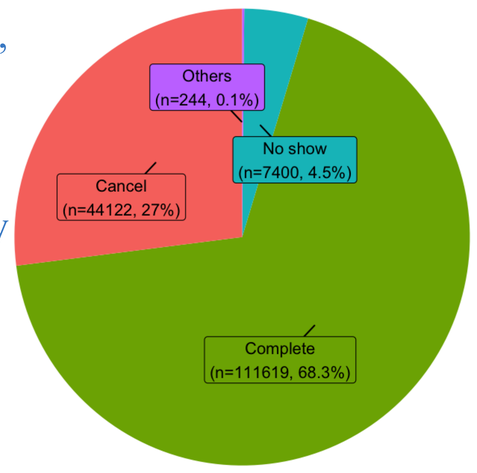

Fig 3. Pie chart of appointment status

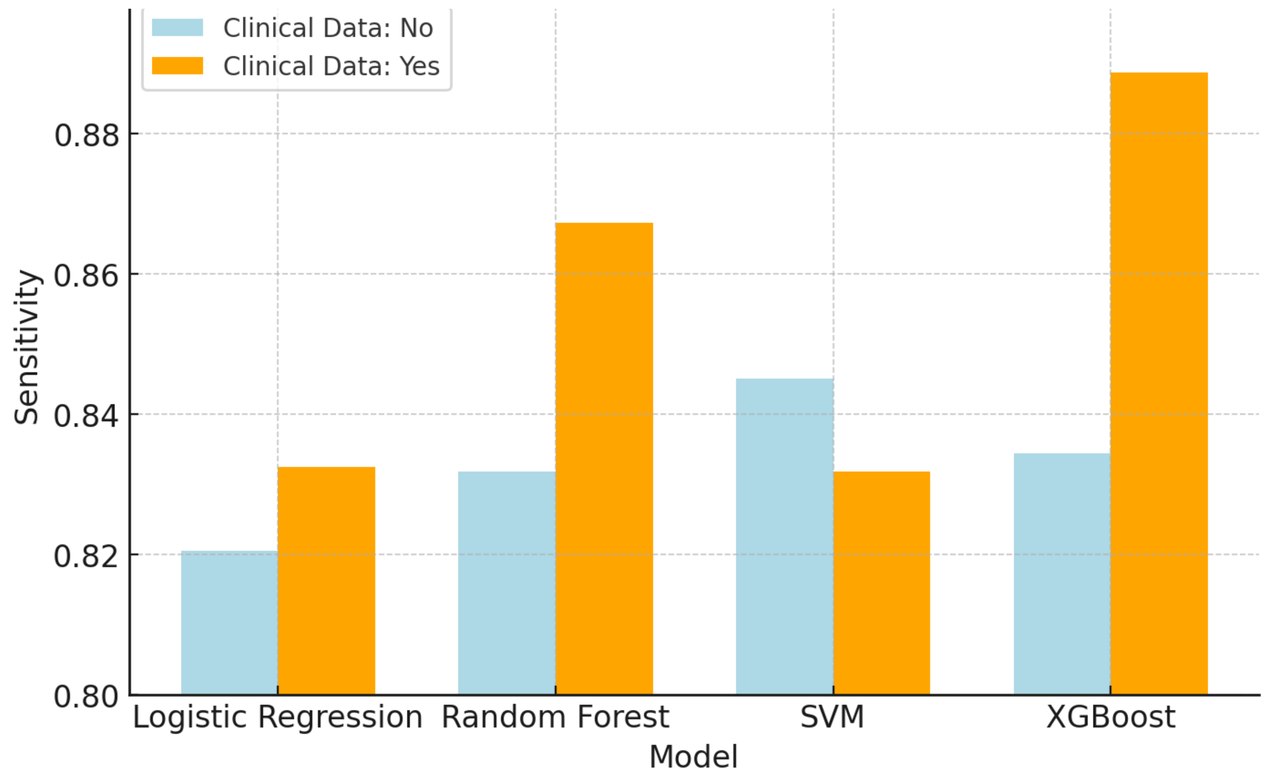

Fig 4. Sensitivity by Models and by with or without clinical data

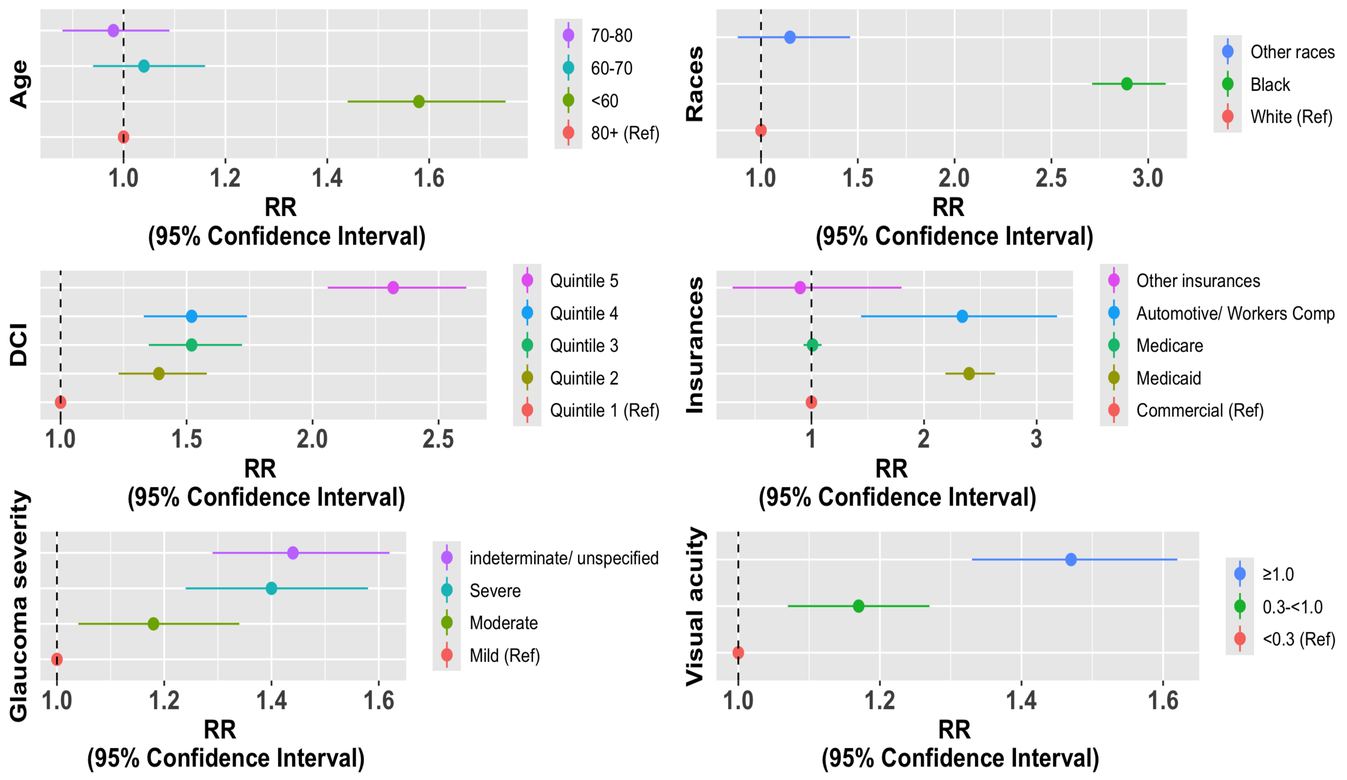

Fig. 5. Forest plot of the relative risk associated with baseline characteristics on the No Show Propensity Factor

Conclusion

This study identifies high-risk patients and informs targeted interventions to reduce no-shows, enhance appointment adherence, and improve care continuity while optimizing resource allocation. Findings may guide future efforts to improve patient outcomes

Reference

Kamdar A, Yohannes GB, Swaminathan SS. Association Between Sociodemographic Risk Factors and No-Show Propensity in a Glaucoma Population Before and During Covid-19 Pandemic. J Glaucoma. 2025 Feb 11. doi: 10.1097/IJG.0000000000002550. Epub ahead of print. PMID: 39927886.