

Modeling Risk of Progression to Lung-RADS 4 from a Benign Lung Cancer Screening CT using Imaging and Clinical Data

 **Gunvant Chaudhari**, Masha Bondarenko, Anand Innanje, Terrance Chen, Jae Ho Sohn

BACKGROUND

Lung cancer screening with low-dose chest CT scans (LDCTs) is evidence-based in reducing lung cancer mortality. Lung-RADS is a widely used guideline for recommending clinical workup (e.g. biopsy) based on imaging findings. However, 90% of patients have benign LDCTs (Lung-RADS ≤ 2), but their recommended interval of next screening is standardized to one year. Personalizing screening interval using AI-based prognostication of lung cancer risk in these low-risk patients could reduce costs, radiation exposure, unnecessary workup, and delay to lung cancer diagnosis.

GOAL

Create AI models that can predict which low-risk patients will progress to Lung-RADS 4 in the next 2 years.

METHODS

1. Build statistical models for clinical/demographic data and a 3D Convolutional Neural Network for full-volume of CT
2. Train models on multi-institutional dataset (National Lung Screening Trial) consisting of 15,910 patients
3. Test models' generalizability to UCSF lung screening data from 583 patients

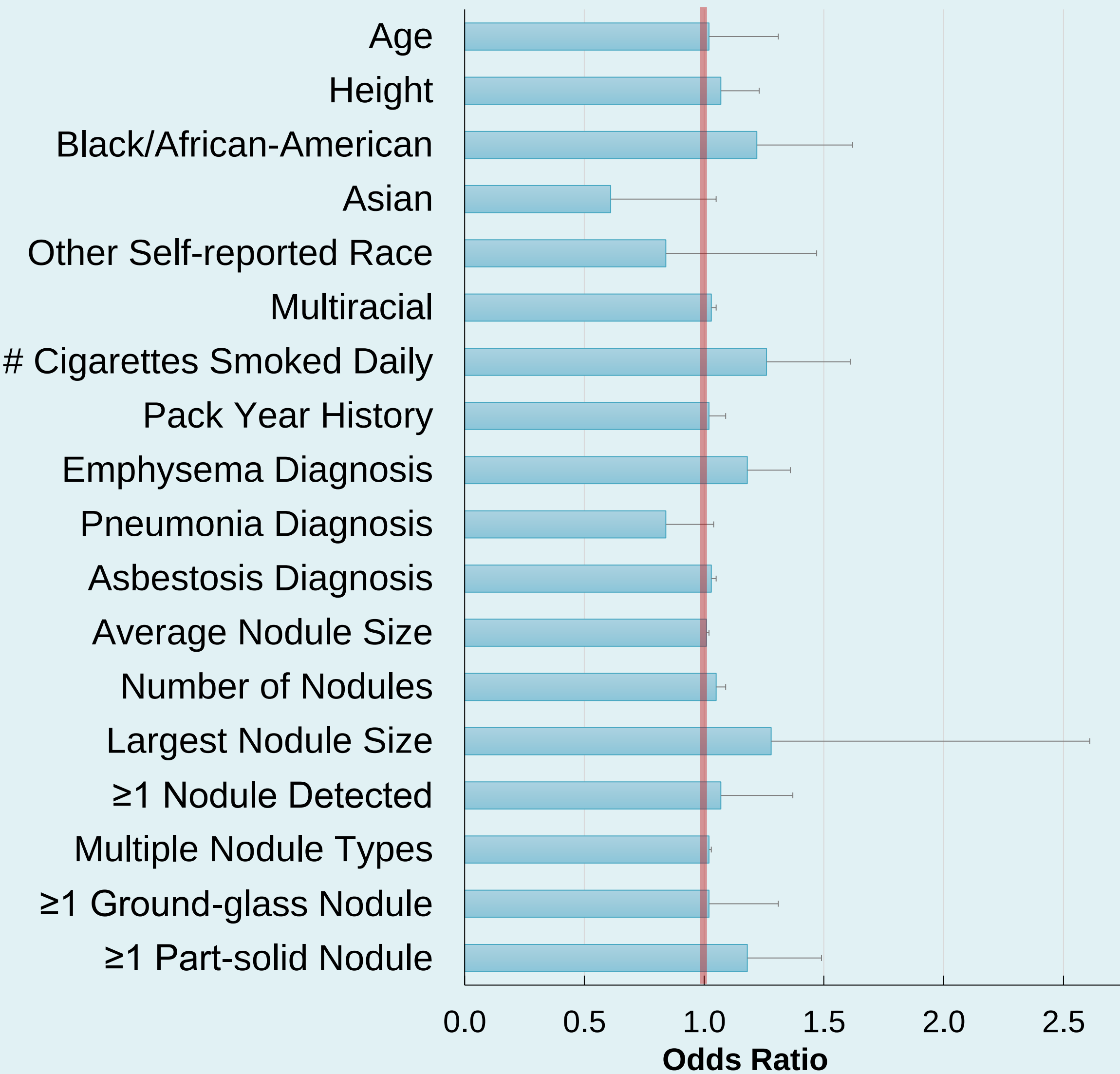


AI can personalize and better prognosticate lung cancer risk using clinical features and a baseline screening CT.

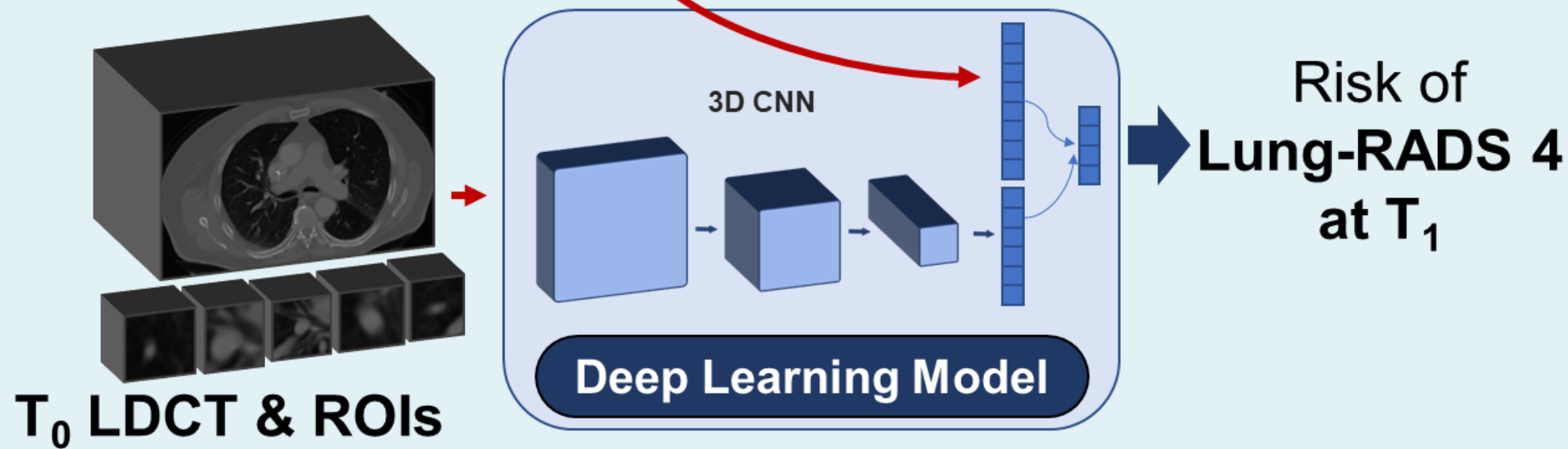
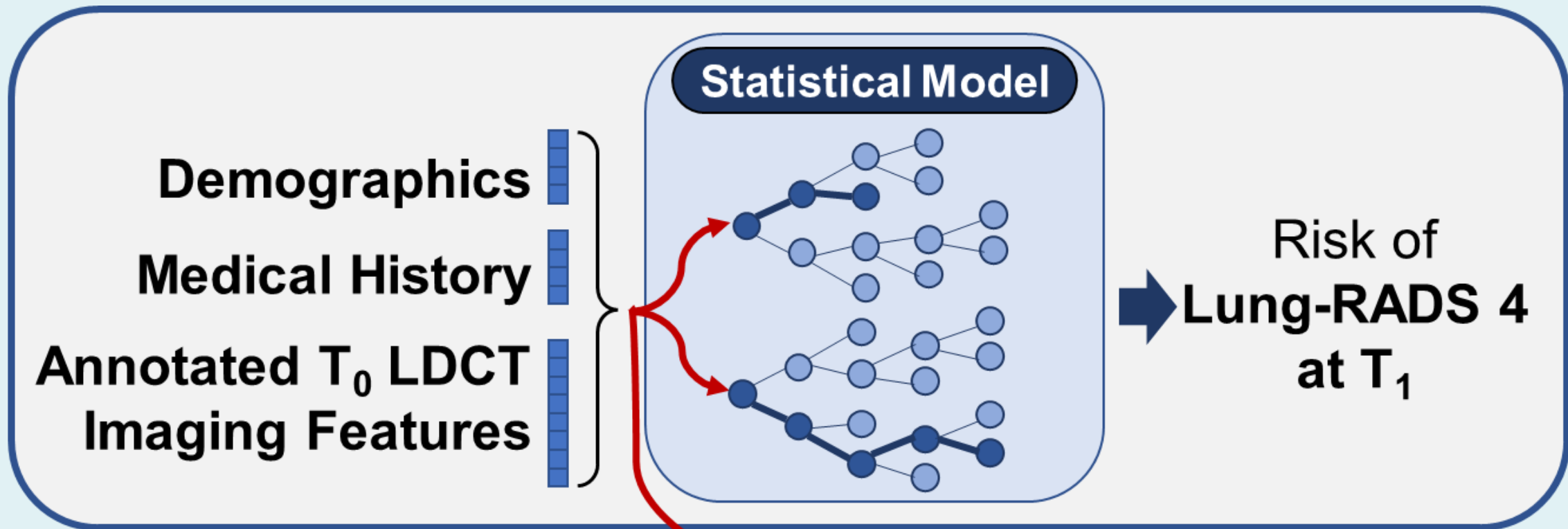
RESULTS

AI Model Version	NLST ROC-AUC (n=18,696)	UCSF ROC-AUC (n=583)
Statistical Models (Clinical Feature Based)		
Logistic Regression	0.648	0.750
Support Vector Classifier	0.653	0.752
XGBoost	0.667	0.733
Deep Learning Model (CT Based CNN)	0.627	0.555
Statistical + Deep Learning Models		
Multilayer Perceptron	0.695	0.686
Logistic Regression	0.691	0.721
XGBoost	0.704	0.724

Performance of various AI models on a held-out NLST test set and on the UCSF test set. The CT-based deep learning model is not able to generalize to the UCSF LCDT data given the low AUC. However, the statistical and combined models do generalize well.



Results of logistic regression on clinical, demographic, and imaging features during iterative feature selection. Shown are odds ratios and the upper limit of the 95% confidence interval.



Methodology for developing the clinical and demographic feature based statistical model and the CT based deep learning model. T0 refers to the baseline scan and T1 refers to interval scan in 1 year. ROIs are the automatically identified nodule “region of interests.”

DISCUSSION

An AI model that takes in clinical features and baseline screening CT can predict progression to a higher risk of lung cancer. The CT-based model has limited generalizability, likely due to differences in LDCT acquisition over time. Training on more recent scans and more validation is needed prior to clinical use.

ACKNOWLEDGEMENTS

Support in part by a grant from the RSNA.



University of California
San Francisco