Statistical Analysis Plan

Official Title

The MyLungHealth Study Protocol: A Pragmatic Patient-Randomized Controlled Trial to Evaluate a Patient-Centered, Electronic Health Record-Integrated Intervention to Enhance Lung Cancer Screening in Primary Care

Brief Title

Engaging Patients to Enable Interoperable Lung Cancer Decision Support at Scale

ClinicalTrials.gov Identifier

NCT06338592

Unique Protocol ID

UUtah_00153806

IRB Information

University of Utah IRB
IRB #00153806
IRB Approval Date of Current Version: **May 19, 2025**

Sponsor

University of Utah
(Study supported by the Agency for Healthcare Research and Quality)

Principal Investigator

Kensaku Kawamoto, MD, PhD, MHS
Associate Chief Medical Information Officer
Department of Biomedical Informatics
University of Utah

Participating Institutions

- University of Utah Health, Salt Lake City, Utah, United States
- NYU Langone Health, New York, New York, United States

Study Type

Pragmatic, patient-randomized controlled trial

Document Type

Statistical Analysis Plan

Document Date

May 19, 2025

Document Version

IRB-Approved Version

Confidentiality Statement

This document contains the IRB-approved study protocol for the clinical trial identified above and is submitted to ClinicalTrials.gov for regulatory compliance.

Statistical Analysis Plan

Patient Outcomes Assessment in Clinical Trial

Outcomes for patients who are potentially eligible for LCS will be assessed in study 1 and outcomes for patients with documented LCS eligibility at the start of the trial will be assessed in study 2. The primary hypotheses and outcomes for both studies are described below.

Study 1 Primary Hypothesis: Among primary care patients aged 50–79 with uncertain LCS eligibility, MyLungHealth eligibility questionnaires and education will result in increased identification of LCS-eligible patients.

Study 2 Primary Hypothesis: Among primary care patients aged 50–79 with documented LCS eligibility, MyLungHealth education will result in increased LDCT ordering among patients with documented LCS eligibility.

<u>Outcomes:</u>

The primary outcome for Study 1 will be the identification of LCS-eligible patients during the 1-year trial among patients with uncertain LCS eligibility at the start of the trial. Patients will be considered to fulfill this outcome if, at any point during the 1-year trial, the patient's EHR record indicates they meet smoking history eligibility criteria, or a patient affirms they meet eligibility criteria in the patient portal. Secondary outcomes for Study 1 will be LDCT ordering and completion during the 1-year trial among patients with uncertain LCS eligibility at the start of the trial.

The primary outcome for Study 2 will be LDCT ordering during the 1-year trial among patients with documented LCS eligibility as per EHR data at the start of the trial. A secondary outcome for Study 2 will be LDCT completion during the trial. Another secondary outcome for Study 2 will be LCS care-gap closure, defined as the identification and completion of recommended care services among patients eligible for LCS according to the EHR. LCS care-gap closure could be achieved through LDCT completion, other chest CT completion, or documented SDM.

Other planned outcomes include estimated number of lung cancer deaths averted and life-year gains per 1,000 patients; estimated # of lung cancer deaths averted and life-year gains per major complication from screening; intervention use measures (e.g., invitations sent, app launches, viewing of app sections); patient knowledge, preferences, decisional conflict, and perceived SDM quality; smoking history measures (availability of complete and accurate smoking history, history updates); process measures including time spent by providers and patients with provider-facing DP+ and MyLungHealth (obtained from system logs); patient and provider assessment of intervention via SUS and NASA TLX surveys; and, to the extent allowed by governance, operating margin attributable to study patients during the study period.

Data Extraction and Management. Data for analysis will be extracted from the data warehouse. During the clinical trial, patient related data will be transformed into a limited dataset if data needs to be shared across institutional boundaries for data analyses in a HIPAA compliant manner. The sharing of the limited dataset will be from NYU to UUH, where the analyses will be conducted. Content of the limited data set are described in the NYU Data Use Agreement.

To the extent allowed by governance, return on investment analyses will be conducted at UUH and leverage a value analysis tool co-developed by PI Kawamoto to determine the true costs of care.

Covariates. Covariates will include patient demographics, Social Vulnerability Index (SVI), Rural Urban Commuting Area (RUCA), and other patient characteristics including PHR utilization, lung cancer risk factors, and clinic characteristics including health system affiliation, number of providers, and provider specialty.

Covariate Adjustment. The number of units of randomization may not be large enough to guarantee covariate balance. Thus, we will collect covariate data, assess for balance, and adjust for covariates as needed.

Statistical Analysis. The distribution of the primary outcome and other variables for enrolled patients eligible for LDCT screening will be summarized and compared using the chi-squared test for categorical variables or the two sample Student's t-test (or its non-parametric counterpart) for continuous variables. The primary analyses are to compare the population level differences in the primary outcomes between intervention and control arms. We will use generalized estimating equations (GEE) with appropriately selected link function based on the

type of outcome to make the comparison after accounting for key confounders. We may adjust for key covariates using the inverse probability of treatment weighting with propensity score. The difference estimation will be performed under the intention-to-treat principle. The health care system will be included in the model as an effect modifier.

Missing data frequently occur in pragmatic clinical trials with long study periods. If data can be assumed to be missing at completely random, standard GEE will be used. Otherwise, weighted GEE will be considered when a missing at random assumption is more realistic. If observations are neither missing at completely random nor missing at random, a more advanced multiple imputation approach will be used to draw valid inferences. All analyses will be conducted using R and statistical significance will be defined at alpha = 0.05.

Sample Size and Study Power. Empirical EHR data indicate that the number of distinct patients to be included in the primary analyses will be approximately 38,943 patients for Study 1 and 3,153 patients for Study 2. To estimate statistical power for Study 1, we assumed that the rate of identification of screening-eligible patients will be at least 20% in the MyLungHealth arm vs. 5% in the control arm. For Study 2, we assumed the rate of screening among eligible patients will be at least 30% in the intervention group vs. 20% in the control group. Given 1:1 allocation, we found that we would have >99% power to detect the estimated intervention effect with a two-sided test with a significance level of alpha = 0.05 for both Study 1 and Study 2.

Covariate Association Analyses. We will assess the association of covariates with study eligibility and study outcomes. For example, we will assess whether patients with a more disadvantaged background according to the SVI were more likely to not have used the patient portal in the year preceding the start of the study, and therefore were more likely to be excluded from the study. As another example, we will assess whether patients with a more disadvantaged background according to the SVI were less likely to engage with and benefit from the study intervention.

Subset Analyses. In prior research, we and others have found disparities in lung cancer burden and screening among women and racial/ethnic minority populations. Thus, we will conduct subset analyses to assess MyLungHealth's impact among subsets of patients including female and minority patients, as well as patients with different SVI and RUCA characteristics.