Harmonizing Multi-Level Public Health Data to Guide Lung Cancer Screening in New York City

Maya Lipsman, MPH, Kimberly Burke, MPH, Parisa Tehranifar, DrPH, Jialin Mao, MD, PhD, Meghan Cahill, MPH, Brooke O'Neill, NP, Phoenix Matthews, PhD

HERBERT IRVING COMPREHENSIVE CANCER CENTER
Columbia · NewYork-Presbyterian

Background

- Only 1 in 5 people eligible in New York state get lung cancer screening.
- Access to screening across the state and New York City (NYC) is uneven.
- A partnership between Columbia University's and Weill Cornell's Cancer Centers allowed the purchase of a mobile lung cancer screening van to serve NYC.
- A mobile van addresses access barriers by bringing services directly to high-need communities.
- Given uneven lung cancer burden, persistent inequities, and the large geographic area covered by this van, this project aimed to bring together 3 datasets to identify highest priority NYC census tracts for the van to visit.

Solutions and Methods Implemented

We used 3 datasets:
1. Adult smoking prevalence by census tract (BRFSS, 2022)
2. Cancer incidence by aggregated census tract (NY State Cancer Registry, 2017-2021)
3. Cancer mortality by community district (NYCDOH EpiQuery, 2021)

- Incidence and mortality data were converted into census tract-level estimates for consistency.
- Equal interval classification
  - To determine cut points, each metric was divided by 4 to create 4 risk score groups.
  - Census tracts were scored from 1-4 (higher number, higher burden).
- For each census tract, the 3 metric scores were summed to generate a composite "risk score" (range 3-12). Census tracts were ranked based on this score.

Goals and Metrics

Goal: integrate multiple datasets to create a census tract-level composite risk score for lung/bronchus cancer in NYC.

Objectives:

- **Compile data** on adult current smoking prevalence, lung/bronchus cancer incidence, and lung/bronchus cancer mortality.
- **Harmonize data across different geographic units** (census tract, aggregated census tract, community district) into a consistent census tract-level dataset.
- **Create a composite scoring system** to identify areas of highest need for lung cancer screening outreach.

Outcome

Census Tracts by Risk Score

3.5% or 79 tracts scored 9
1.3% or 29 tracts scored 10
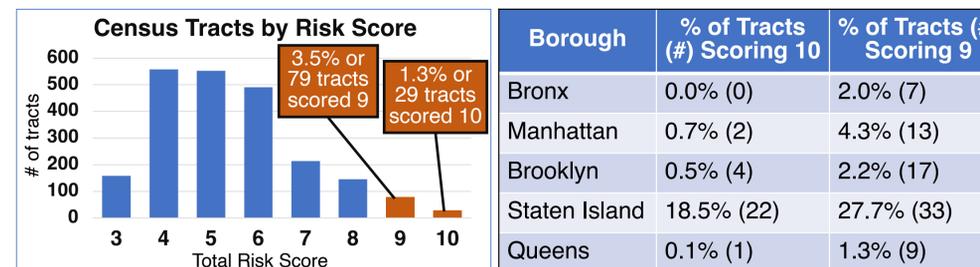
Borough	% of Tracts (#) Scoring 10	% of Tracts (#) Scoring 9
Bronx	0.0% (0)	2.0% (7)
Manhattan	0.7% (2)	4.3% (13)
Brooklyn	0.5% (4)	2.2% (17)
Staten Island	18.5% (22)	27.7% (33)
Queens	0.1% (1)	1.3% (9)

- The analysis produced a scoring system covering all 2,227 NYC tracts.
- Tracts with the highest composite scores were identified as high-priority areas for mobile lung cancer screening.
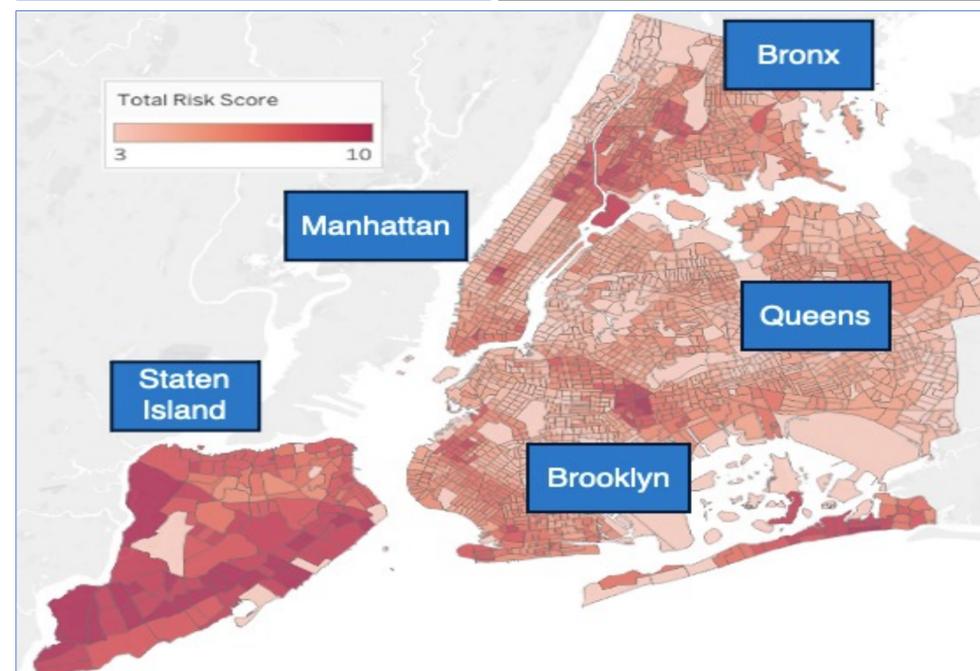
Lessons Learned

We demonstrated the feasibility of integrating heterogeneous datasets across geographic scales to generate transparent and actionable, census tract-level prioritization for cancer education and screening outreach.

Future work:
- Planning tailored outreach activities
- Refining the methodology by exploring dynamic updates as new surveillance data become available and community feedback is received.