Spatial Data Enrichment for Historical Analysis of Non-Occupational Exposure to Asbestos

Vicky Tam, Patrick Emedom-Nnamdi, Madhura Gurav, Frances K. Barg, Douglas J Wiebe
University of Pennsylvania

Penn SRP Center

Perelman School of Medicine University of Pennsylvania | CEET

Introduction

Asbestos exposure is most widely studied as an occupational phenomenon. However, there is emerging evidence of asbestos-related diseases among non-occupationally exposed individuals. Using Ambler, Pennsylvania (PA), a community with substantial occupational and community exposure to asbestos, we aimed to characterize non-occupational exposure to asbestos and its resultant mortality within a spatial context.

Methods

Using historical Sanborn Fire Insurance Maps from 1930, we reconstructed a digital historical narrative of Ambler. We leveraged the spatial capabilities of geographic information system (GIS) to digitally map property boundaries and building outlines. The data captured includes address information and details on building construction materials along with square footage. Using the data collected, we were able to approximate property level socioeconomic status within the community.

We then used publically available Census records to identify individuals living in Ambler within the same timeframe. We extracted names, address, gender, race, occupation and industry. Occupational exposure was defined on the basis of an individual's occupation and listed industry. Paraoccupational exposure was defined as having the same address as an individual with occupational exposure. We then spatially enabled these Census records by linking the data with the existing digital map of our study area. We were then able to describe associations among exposure and socioeconomic status.

What is GIS?

Geographic Information Systems, or GIS, is a system of computer software, hardware, and data. GIS can be used to create, analyze, and visualize digital layers of spatial information. GIS analysis and maps can reveal relationships, patterns, and trends not visible through tables, charts, or other means. In this study, we used Esri's ArcGIS, a mapping and analytics software.

Acknowledgements

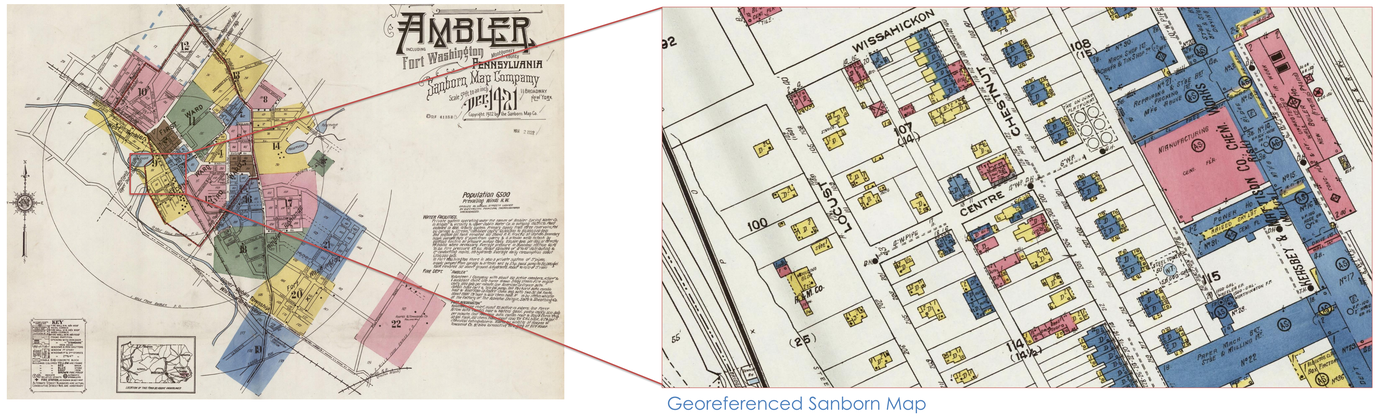
Research reported in this poster was supported by National Institute Environmental Health Sciences of the National Institutes of Health under award number P42 ES023720 Penn Superfund Research Program (SRP) Center Grant, and Supplement (3P42ES023720-01S1). This research reflects work done through Project 3 of the P42 SRP grant and the Biostatistical Research Support Core (Core D).

Penn Arts & Sciences | CEET CENTER OF EXCELLENCE IN ENVIRONMENTAL TOXICOLOGY | Penn SRP Center | Perelman School of Medicine University of Pennsylvania

Results

1. Georeferencing

Georeferencing is the process of aligning an image, a map for example, to a known coordinate system. Prior to georeferencing, a digital image of a map has no spatial reference to the earth's surface. Georeferencing places the map on the correct "place" on the earth's surface. The mechanics of georeferencing the Sanborn Fire Insurance maps in ArcGIS includes shifting, rotating, scaling, and orthorectifying the image using existing spatial data and landmarks.

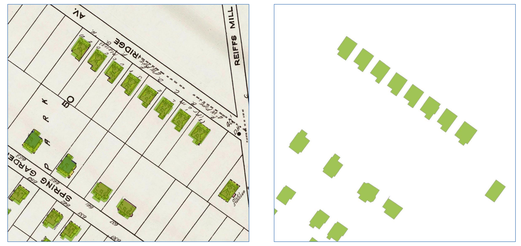


Georeferenced Sanborn Map

2. Digitizing

Digitizing is the process of converting geographic features on an georeferenced image of a map into digital format. The x, y coordinates of point, line and polygon features are recorded and stored as the spatial data. Using the Sanborn maps as a basemap, we used a combination of heads-up digitization and editing current day parcel and building boundary layers to recreate the historical landscape of Ambler in 1930.

Georeferenced Sanborn Map + Digitized Parcels

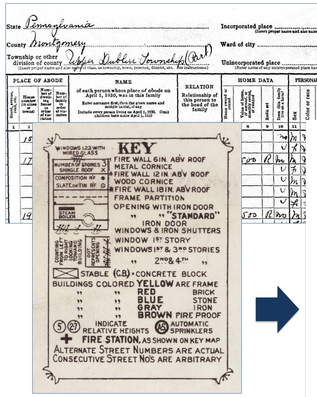


Georeferenced Sanborn Map + Digitized Buildings

Digitized Parcels + Buildings

3. Database Development

Tabular attribute information is the basis of geographic features, allowing for the visualization, query, and analysis of spatial data. We captured parcel level address, building and road information from the Sanborn maps. In addition, we incorporated household level data from the 1930 Census to further enrich our spatial database. Using this information, we were able to describe the associations among exposure and socioeconomic status.



ADDR1	ADDR3	LAND_SF
120 MATTISON AVE	AMBLER PA 19002	3750
206 MATTISON AVE	AMBLER PA 19002	3900
208 MATTISON AVE	AMBLER PA 19002	3900
218 MATTISON AVE	AMBLER PA 19002	3900
220 MATTISON AVE	AMBLER PA 19002	6623
330 MATTISON AVE	AMBLER PA 19002	22410
342 MATTISON AVE	AMBLER PA 19002	15000
300 PARK AVE	AMBLER PA 19002	7900

Spatial Database