Host site: Epidemiology Section/ Occupational and Environmental Epidemiology Branch
NC Department of Health and Human Services/ Division of Public Health

Assignment Location: Raleigh, NC

Primary Mentor: Kim Gaetz MSPH, PhD

Secondary Mentor: Chandrika Rao, MS., PhD.

Agency Description:

As part of the Division of Public Health, we provide data to support public health decisions and policy efforts across North Carolina. We promote a people-focused, team-oriented, transparent, and joyful workplace. Our agency is located in Raleigh, NC, part of the Research Triangle Area, a top region for environmental science, engineering and health research. The Occupational and Environmental Epidemiology Branch quantifies exposures to environmental and occupational contaminants; conducts health risk assessments and risk communication; provides medical evaluation and surveillance for adverse health effects; and provides health-based guidance on levels of exposure to such contaminants. In addition, we provide surge capacity support to the Epidemiology Section for outbreak and disaster response activities. North Carolina Central Cancer Registry (CCR) is one of the five branches with in the State Center for Health Statistics within the NC Division of Public Health, NC Department of Health & Human Services. The CCR is a legislatively mandated cancer surveillance program which is responsible for collecting cancer diagnosis from all 100 counties. Data collected includes patients first course of treatment, stage at diagnosis, and demographic information. CCR's priority is to collect timely, complete and accurate data electronically, use the data to monitor cancer burden across the state, identify priority populations with late-stage cancer diagnosis, address racial disparities to achieve health equity.

Assignment Description:

The main responsibilities of the CSTE Informatics Fellow will be to work with NC Division of Public Health employees in the Occupational and Environmental Epidemiology Branch and the NC Central Cancer Registry to create, test, and implement new data pipelines. The fellow may work with the information technology (IT) team to store, integrate, and label the data in the NC DHHS data repository using Amazon Web Services for rapid retrieval and analyses to uncover insights not previously visible from the siloed datasets. They will work with the data architect to define the data pipeline logic and the frequency of updates needed from each data source. The IT team will create a business plan for storage
and automation of climate, environmental contaminant, and cancer data into the Environmental Health Data Dashboard (EHDD) and will provide a proof of concept in a test environment. The fellow's work will be supervised and guided by the OEEB Epidemiology Supervisor, the Central Cancer Registry Director, and the IT project manager. The fellow will have the opportunity to work on an interdisciplinary team including epidemiologists, statisticians, IT staff, GIS analysts, health educators, and industrial hygiene consultants as well as clinicians and laboratory staff.

**Preferred Background & Skills:**

Experience with data flows as a biomedical data programmer, conducting data processing and transformation, data automation.

**What can the fellow expect to gain from 2 years at this host site?**

After 2 years at our agency, the fellow will have a solid understanding of the complexity of public health data systems, the partnerships between public health agencies and laboratories, and the necessary checks to ensure data quality and completeness at every step of the data flow. The fellow will also gain an understanding of the end product that public health professionals need to inform disease prevention and be responsive to the needs of the general public and policy makers as well as researchers and other stakeholders.

**Potential Projects include:**

*Host sites have listed up to 5 projects*

**Project 1: Create an Internal Dashboard to filter and monitor cancer surveillance data.**

The objective is to create a user-friendly interactive data dashboard with calculating, sorting and filtering capability. There are several data tracking mechanisms used to monitor the progress for reported data from hospitals, medical facilities, and other data sources. This dashboard can serve as an interactive, universal site to see and customize data filters and timelines to improve on data reporting quality and timeliness.

This dashboard will better serve the data collection and data quality operations of the cancer registry, albeit thru data monitoring, predictability, and facility recruitment.

**Project 2: Automating Data Flows for Timely Updates to the Environmental Health Data Dashboard**

This project has three major objectives:
1. Increase epidemiologic, statistical, data science, and geospatial capacity in the Division of Public Health, Occupational and Environmental Epidemiology Branch (OEEB), North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT), and the State Center for Health Statistics (SCHS).

2. Leverage IT infrastructure built during the COVID-19 response to integrate climate, environmental, and occupational data and related health outcomes with existing health data on the internal cloud-based data repository for secure and rapid retrieval.

3. Automate data from the internal repository into both internal and public facing dashboards in partnership with the SCHS based on, but not limited to, the indicators required for the National Environmental Public Health Tracking Network.

The main deliverables of this project will be automated data flows between the Vital Statistics, Prenatal Outcomes, and Cancer Registry data, NC Poison Control carbon monoxide and pesticide poisoning event data, and the Environmental Health Data Dashboard.

Much more efficient and effective reporting of data to stakeholders through updated and interactive data visualizations; better understanding and ability for OEEB and SCHS staff to monitor trends in data to more quickly investigate issues of concern; Increased availability of timely data to the general public, community advocates, and policy makers to support initiatives related to environmental health, climate resilience, and disaster recovery.

Project 3: Process Automation of Death Certificate Matching with Central Cancer Registry Case Data

With the recent implementation of the electronic death registration system in the Vital Records section, we are planning an interface between the cancer registry and the Death Certificate data. This will allow registry staff to find missing cancer cases incrementally as opposed to annually. This data will also assist the statisticians within the registry to make more rapid distribution to reporting sources. Depending on interest, the program could be distributed to other groups in need of Death Certificate data.

By retrieving death certificate information more frequently, more current data can be utilized in publications and distributions of information to lawmakers and the public.

Project 4: Electronic Pathology Reporting for Cancer Cases

Implement cancer electronic pathology and biomarker reporting (ePath) by utilizing the Association of Public Health Laboratories (APHL) Informatics Messaging Services (AIMS) cloud platform. Support DMI project initiative activities by validating cancer messages in HL7 2.5.1 format, onboarding new local/national pathology laboratories, participate in DMI work group meetings.

Receiving data electronically from path labs will enhance timely case ascertainment of cancer cases.

**Additional information about the placement:**

There is also the potential to help with development of data flows from reporting laboratories into a new Maven based COVID-19 Wastewater Surveillance system, projected to be developed in 2023.