Improving Coordinate Accuracy for Cancer Cases in Oklahoma

ENVIRONMENTAL PUBLIC HEALTH TRACKING

ASTHO FELLOWSHIP REPORT

Submitted by

Anne Pate, MPH, PhD and Amber Sheikh, MPH

Oklahoma State Department of Health
Chronic Disease Service
Cancer Prevention and Control Program
1000 NE 10th Street
Oklahoma City, OK 73117

Submitted to:

Association of State and Territorial Health Officials
Environmental Public Health Tracking: State-to-State Peer Fellowship Program
2231 Crystal Drive, Suite 450
Arlington, VA 22202

INTRODUCTION

The Association for State and Territorial Health Officials (ASTHO) awarded the epidemiologist for the Oklahoma Cancer Prevention and Control Program (CPCP) at the Oklahoma State Department of Health (OSDH) with funds to complete a fellowship in association with the Centers for Disease Control and Prevention (CDC) National Environmental Public Health Tracking Network (EPHT) program. The fellowship activities were conducted from February 2012 through May 2012, with an extension to July 2012. These activities included attendance and participation in a national workshop, a site-visit to the host state, completion of a project and summarization of the results of that project.

The purpose of the fellowship activities was to provide hands-on experience for states with no currently funded Environmental Public Health Tracking projects or programs. To fulfill this, each fellow was assigned a host state: Oklahoma's host state was New Mexico. New Mexico was chosen in part due to a shared interest in conducting a project looking at arsenic in well water and its associated cancer risk. The site-visit to the host state provided an opportunity for the fellow to learn tracking techniques and develop an understanding of what type of Information Technology (IT) infrastructure would support a tracking portal. The site-visit also assisted in building networks across state lines and further enhanced a national tracking conversation. The attendance at the workshop in Denver provided additional insight into developing a tracking network and program. The fellow was unable to attend the workshop but had a designated individual, another colleague trained in epidemiology, to attend.

Several of the activities originally identified to be completed for Oklahoma's project were altered due to unforeseen circumstances; however the primary objectives of the project were accomplished and are summarized in a later section of this report.

ACTIVITIES REPORT

<u>EPHT Tracking Workshop – Denver, CO</u>

The CDC's National Environmental Public Health Tracking Network conducted a 3-day workshop in Denver, Colorado (April 30-May 3, 2012). Invitation to attend the workshop provided an opportunity to gain insight on EPHT's national workgroups: Content workgroup (CWG), Standards and Network Development (SND) and Program Marketing and Outreach (PMO) workgroup.

Exposure was gained to various environmental health issues and how the data is provided for the public health tracking portals such as pesticide residue in foods, lead poisoning prevention, healthy homes project, and arsenic concentrations in drinking water. It was learned that EPHT would be including new content in the state portals to include expansion of lead poisoning prevention, bio-monitoring and a cancer module.

Focus workgroups such as the geospatial team of CWG brought up challenges such as how much of the tracking data can be or should be provided on the website portals for public usage; how to standardize the data maps for website portals; and how to overcome the challenges of small sample size at county level. New York City, Vermont and New Mexico

showcased their portal websites and the recent changes they had introduced during the PMO session.

Tracking fellows from Fairfax (Lead), Alabama, Idaho (Lead), Nebraska (Radon) and Oklahoma (Arsenic) got an opportunity to network at a breakfast meet and greet. Grantee mentors and phase II grantees shared lessons learned with tracking fellows. The main theme from the discussion was to build strong relationship with IT and build and maintain strong contacts with data stewards for contracts and data sharing agreements. This helps in creating a successful Public Health portal. New York City's advice was to bring in local partners such as daycares, licensing offices etc. and provide reports to them through the portal once it and the associated infrastructure is in place. This helps to increase the buy-in from various stakeholders.

This was a great networking opportunity. Oklahoma received information from Maine (Arsenic in well waters) for the project on how to utilize www.MelissaData.com to find current addresses for rural tracts. All three national workgroups provided reports on the last day, summarizing individual workgroup sessions. Wrap up by former EPHT leads provided valuable insight on EPHT's journey and the challenges that lay ahead.

Host State Site Visit – Santa Fe and Albuquerque, New Mexico

The fellow participated in a site visit with the host state by spending one day in the Santa Fe offices and one day in the Albuquerque offices, with the trip scheduled later than initially planned (05/06/2012-05/08/2012) due to other conflicting travel. The site visit was extremely helpful and successfully provided an overarching perspective of successful implementation of a tracking network in a state and what components made it so successful.

The site visit began with a day spent at the Santa Fe offices and learning how New Mexico (NM) implemented the grant and built their tracking network. One of their first steps was to bring together a planning consortium advisory group. This group was broader than an advisory group and was composed of individuals who worked with either environmental or health data. This group identified that geography was the common denominator for these types of data and would be how they were linked.

To identify what would be the priority areas to begin tracking, the NM EPHT staff used local health councils, which in turn conducted local focus groups, to identify community concerns and interests. It was determined that the following were of greatest concern to the citizens of New Mexico:

- Water quality and quantity
- Hazardous waste
- Air quality (outdoor)
- Natural disasters
- Food safety

The next phase in the development of their program was to assess the data that was available and where there were gaps. They conducted a data survey in two phases. This

survey was administered to any health department, environmental department, and any other state or federal agencies that may have any health- or environmentally-related data. Some of the factors that were considered in this survey included:

- Electronic versus hard copy data
- Legislatively mandated
- Ability to share
- Funding source
- Coverage

The program contracted with the School of Law to review laws and legislation to assess if the NM Department of Health had the authority to collect, assess, and analyze health and environmental data.

Once these initial steps were completed: identifying priorities, assessing available data, and determining legal authority, the NM EPHT identified key players to become part of the network, namely:

- Vital Records
- Cancer Registry
- Birth Defects Registry
- Asthma (hospital discharge data)
- Lead
- Hospitals

The primary projects they began with, that followed both community priorities and grant requirements, included: bio-monitoring, drinking water quality, and air quality forecasting.

One of the major components of the NM EPHT program is the web-based data portal. New Mexico created a portal with two aspects, a public aspect and a secure side. The NM EPHT program staff serves as the intermediary between requestors and the data stewards for data requests. While they are not the data stewards themselves, the NM EPHT program is provided with enough information that they can respond to the majority of requests that come through the portal. Furthermore, the program is able to track what aspects of the data are most queried by the public and researchers and identify gaps or areas that could be strengthened. This type of arrangement was made possible by multiple data sharing agreements. The portal has strict suppression rules to ensure security and compliance with the data sharing agreements (denominator < 20 or numerator < 3 will be suppressed). Eventually this portal will become the NM health department data repository. It currently contains non-identifiable record level data as well as environmental point data. The portal can be accessed by a secure route to obtain more detailed information at the census tract or county level. This secure access can also be used to provide researchers with more detailed datasets. One of the biggest challenges with the public portal is keeping up with technology updates and ensuring it continues to function appropriately regardless of the user's software version.

Some of the measurements used for the outcomes on the public portal include age-adjusted incidence rates, unadjusted rates, and prevalence, depending on the outcome of interest. The public portal must be flexible enough to accommodate for many types of users. Part of developing the public portal was the identification and definition of small areas. These geographic areas are smaller than a county but larger than a census tract and were identified through an iterative process which involved multiple parties including health professionals and community members. These small areas were defined and created to meet the need of many data requests that request information related to community level concerns that often encompass unique geographic boundaries. These geographic areas were designed to sufficiently respond to community concerns while simultaneously ensuring security and stability of the calculated rates.

Overall gaps and challenges associated with the program include the following:

- Constant technology updates
- Cleaning and formatting the data
- Keeping the data updated
- Integrating the systems
- Maintaining multiple data submissions

Overall positive aspects associated with the program include the following:

- Provides a flexible enough system to provide specific responses to data requests
- Increases researchers' access to the data
- Helps government develop more personal relationships with the data consumers
- Enhances other capabilities such as data exchange with other surrounding states

The second day of the site visit was spent in the Albuquerque offices. During this time the fellow met with staff including those involved in program outreach, health education and epidemiology. The program staff uses multiple methods to determine what information is of greatest interest to the public. One of these tools is Google analytics, which tracks who has been accessing the public portal and which links are being viewed the most. This analysis helps the staff prioritize and decide which web pages need to be updated or if something such as a mobile app may need to be developed. They have determined that the two primary users of the portal are those individuals newly diagnosed with a disease or researchers/public health professionals

Risk communication is an important aspect of all the products created and released by the tracking program and is embedded in all messages. The website was revamped to ensure it appropriately incorporated risk communication. The health education/health communications coordinator developed an approval flow when creating new materials, thus ensuring that everyone in the program is aware of the materials and has the chance to provide input. In general, the program focuses on interventions and ideally, prevention.

The fellow also met briefly with the geospatial epidemiologist who provided some pointers for improving the results of the Oklahoma project. The recommendations included:

- Review the text in the cases
- Look for an online converter (Melissa Data)
- Look for alternative address
- Link with DMV data
- Use other databases as available, such as hospital discharge data
- Use death certificates
- Use the tax assessor file to assign census tract

Fellow Project

Project Summary

The Oklahoma Central Cancer Registry (OCCR) geocodes its database each year with high quality address data, as required by the grant from the CDC's National Program of Cancer Registries (NPCR). The OCCR uses a variable called census tract certainty to categorize the quality of the geocoding of each case. Only those cases with the most accurate census tract certainty were incorporated into the fellow's previous research study assessing the geographic relationship between arsenic in drinking water and incidence of urinary tract cancers. Due to the large rural population in Oklahoma, many of the addresses associated with cases in the OCCR database are rural routes and post office boxes which cannot be geocoded with a high level of confidence. This resulted in excluding 474 cases from the study. Furthermore, these excluded cases were primarily located in more rural counties, where the probability of being on well water and consequently potential exposure to arsenic was higher.

The proposed project was in collaboration with the University of Oklahoma, College of Public Health (OUCOPH) to have a team of students go out into the field and use Global Positioning System (GPS) units to collect the coordinates of the addresses that were geocoded to their zip code only. The project was to begin with the 292 cases excluded from the fellow's previous research study and expand to the rest of the OCCR database.

Specific Aims

The specific aims and expected outcomes of this project included the following:

- Establish collaboration with the OUCOPH which will provide students with field experience and provide the OCCR with higher quality data.
- Provide the OCCR with the ability to conduct environmental epidemiologic investigations with more accurate geographic information associated with the cancer cases.
- Host an in-state data stewards meeting at the conclusion of the project including the following potential attendees: OUCOPH students and professors, CPCP staff, USGS staff, and Health Care Information staff.
- Analyze the data from the previous research project to assess if the results were impacted by the improved data.

- Write and submit a scientific paper on the results of the previous study and the impact of the increased accuracy of the geographic data to a peer-reviewed journal.

Benefits and significance to Oklahoma and EPHTN

This project resulted in multiple benefits for Oklahoma. The fellow identified a group of graduate students and a professor from OUCOPH who leads the group who is willing to participate in public health activities and field work. The quality of the geographic information associated with the cancer cases was improved significantly which resulted in an improved ability of researchers to identify a potential association between exposure and disease, even in an ecological study. This project has provided an opportunity to ensure that any GPS equipment purchased in the future will meet security standards. The project has provided an opportunity for the fellow to learn the process of working through a potential security issue with IT and implement alternative techniques to complete a project.

Research design/Methods/Key Personnel

The initial proposal for this application was to collaborate with the OUCOPH and have a team of students go out into the field and use Global Positioning System (GPS) units to collect the coordinates of the addresses that were geocoded to their zip code only. The OSDH had ten GPS units that were initially thought to be available for use on this project. Ultimately, the use of GPS units was put on hold due to internal OSDH IT security concerns. Due to this delay the fellow found an alternative to using the GPS units that was offered as an option during the EPHT Workshop in Denver. The project began with 292 of the cases that were excluded from a previous ecologic study conducted by the fellow due to low geocoding accuracy.

This project was submitted to and approved by the OSDH Institutional Review Board (IRB) for approval prior to initiation. The fellow and a colleague tested the initial methods by attempting to find rural routes in the field by visiting a local post office and driving around to see if any old signs were left indicating rural routes. Unfortunately, this field experiment was unsuccessful. Due to several hurdles in implementation of the project including difficulty in finding physical locations of the rural routes in the field as well as a limited supply of GPS units (only one was approved for use), the fellow pursued other avenues to obtain the same information. Another option for obtaining coordinates that was learned about at the EPHT workshop in Denver was the use of a website called www.melissadata.com that has the ability to convert rural route addresses to physical addresses that can then be geocoded. The fellow proceeded to use this application as an alternative to using GPS units. Furthermore, the fellow learned of other options from the site visit to the host state, New Mexico. Those cases that could not be found at www.melissadata.com were matched with hospital discharge data and death certificate data as well as reviewed to see if they had multiple primaries and if there was any geographic location that could be used from the other primary. The fellow updated the OCCR database as coordinates were identified through these multiple options.

Results

After determining that alternative methods would need to be used rather than the field work using GPS units, the fellow pursued a variety of other methods. One was implemented after learning about it from the EPHT workshop in Denver, www.melissadata.com. The second method was identified after discussing the missing information with the epidemiologist with the New Mexico tracking program and they provided some other options to find address information that could be geocoded. The following steps were taken with 292 cases that had a census tract certainty of 4, indicating they were geocoded based on the zip code of residence at diagnosis.

1. Use www.MelissaData.com, an address verification website that converts rural routes and other unmatched addresses to physical street addresses. While it did not match all of the unknown addresses, the fellow was able to match 51% on the first round, leaving 134 that did not match.
2. Those cases that did not match were reviewed in more detail and some addresses were identified by reviewing other tumors diagnosed in the same individual in the same time period and had a legitimate address associated. Others were matched with hospital discharge data or death certificate data or their other primary was appropriate to use the address information from. Using these alternative methods, another 43 cases were matched, leaving 92 unmatched.

With the combined techniques of online matching and manual case review and linkages with other databases, 68.5% of the cases were assigned a latitude/longitude, leaving less than one third to be manually geocoded with the GPS units once they are approved for use by IT.

Discussion

One of the biggest setbacks for this project was unforeseen problems with the GPS units. In planning the project it was not anticipated that there would be any issues using the existing units that OSDH owned. In the process of preparing them to be taken into the field it was realized that the software was extremely out-of-date as well as IT administrative procedures severely limited the use of the particular models that the agency had bought many years prior. The units had to attach to a hand-help PDA to be read, which are not considered secure and are prohibited from being connected to the network due to recently implemented IT security protocols. The GPS units were outdated, therefore several programs, including the fellow, worked together to identify what newer model would be appropriate to purchase. The new units were deemed necessary not just for this project, but for other program area projects as well. Since the beginning of this project, a newer model GPS has been identified, purchased and is currently in IT being reviewed to ensure it will meet all IT security administrative rules and procedures.

Reviewing the expected outcomes, one was met while several were only partially met. The challenges met while working on this project were not expected and working through them provided the fellow a valuable experience in pursuing alternative options to complete a project as well as possible with available resources and delays in implementation.

PLANNED ACTIVITIES

Action Plan for Future Tracking Activities

The fellow plans on re-analyzing the data including the more accurate coordinate data to determine if more significant associations can be detected. These updated results will be compiled and submitted for publication in a peer-reviewed journal.

In addition, the fellow plans on continuing to obtain more accurate geographic information for the cancer cases. Based on lessons learned from this current project, the fellow will ensure the database is matched at least annually with other databases that will provide more accurate address information, such as hospital discharge data and death certificate data. The fellow will continue to work with the OUCOPH to engage students in research projects involving the OCCR database and to encourage environmental health projects.

In addition, the fellow will seek out collaborations both internally and externally, such as the blood lead screening program, to expand the ability of the OSDH to conduct tracking activities.

The fellow will continue to collaborate with Health Care Information to ensure OCCR data is available online at www.health.ok.gov/ok2share for public queries and to encourage the mapping component to be developed and implemented.

Specific Aims and Long-Term Goals

The fellow will continue to apply for environmental health funding opportunities and will work to expand the OSDH's tracking activities. Ultimately, the fellow would like to see Oklahoma funded to implement the full developed tracking activities. Until that time, the fellow will continue to participate in and encourage these activities as well as demonstrate their importance to senior leadership.

CONCLUSION

This project provided unique opportunities for the fellow. Not only did the visit to the host state offer an excellent perspective on implementation and maintenance of a highly successfully tracking network, but also ended up providing the fellow with additional tools that ensured the project was completed more successfully. This project provided an opportunity for the fellow to implement alternative methods to achieve completion, in spite of challenges. It allowed the fellow to work more closely with IT and learn their project management processes as well as how to purchasing and implement new equipment. Experiences such as those dealt with during this project have provided the fellow with additional knowledge to continue to pursue increasing environmental health tracking activities in Oklahoma.