Environmental Public Health Tracking: Driving Environmental Health Information

Alex E. Charleston, MPH, Holly R. Wilson, MHSE CHES, Peter O. Edwards, MPA, Felicita David, MS, and Shannon Dewitt, BS

Background

In September 2000, the Pew Environmental Health Commission issued a report stating that public health agencies lacked capacity to evaluate and conduct key investigations into the status of the health of their environment. The commission reported that the environmental public health system in the United States was fragmented and ineffective and recommended that agencies with roles in environmental public health improve their capacity to evaluate community exposures and conduct investigations. The report called for the establishment of an environmental public health tracking network (Tracking Network) that would monitor the level of burden from environmentally related disease.

In 2002, Congress appropriated funds to the Centers for Disease Control and Prevention (CDC) to develop this network, and under the stewardship of the National Center for Environmental Health (NCEH), the National Environmental Public Health Tracking Program (Tracking Program) was created. Environmental public health tracking is the “ongoing collection, integration, analysis, and dissemination of data from environmental hazards monitoring, human exposure tracking, and health effect surveillance”. These data sources are essential to a tracking system that will allow the linkage of environmental hazard and potential adverse health effects and improve our understanding of the impact of environmental factors on our health. (Figure 1)

Public health programs should actively collaborate with their partners in the development of information systems and establish standards for content. In 2006, the Tracking Program funded 17 state and local partners to develop and implement the Tracking Network. This followed a period of capacity building among as many as 30 state and local health departments, with funding to build or enhance workforce capacity, enhance informatics capabilities, and to develop partnerships among relevant environmental and public health agencies across the country that had been noted by the Pew Commission as being fragmented, or nonexistent.

The guiding strategy and concept that evolved in 2006, including the data content and selection of measures and indicators for the network, was devised through consensus among the program’s stakeholders. They envisioned the Tracking Network as a Web-based system...
that provided timely access to environmental and health data collected by a variety of agencies, with a parent program that would increase environmental public health capacity at the national, state and local levels.\textsuperscript{5} This consensus process also included participation by academic partner institutions, and a number of public health and national data organizations.

Prior to Tracking, exposure, environmental monitoring, and health surveillance systems existed in separate proprietary domains, and in formats that were generally not compatible with one another. With leadership from CDC and the U.S. Environmental Protection Agency (EPA), state and local recipients of technical assistance funding from the Tracking Program were required to establish liaisons between their local health and environmental agencies, just as CDC and EPA had established a Memorandum of Understanding (MOU) with each other to collaborate on this project. Funding recipients had to demonstrate that collaboration would be forged at the highest levels of each other’s agency, underscoring a desire to break down long-held silos for genuine health-environment collaboration.

For the Tracking Program to be successful, it had to contend with existing data silos. It also needed to leverage the collective know-how of its network of stakeholders to devise a standardized methodology of demonstrating associations among relevant health and environmental data. Additionally, the program had to use the available data to create a suite of nationally consistent reference frames for using the data along with measures relevant to priority health outcomes of interest. Furthermore, the Tracking Program needed to bring together scientific information, technology, and health communication to make the data accessible, usable, and understandable by a variety of users. To ensure these activities were addressed appropriately, CDC’s Tracking Program and its partners established workgroups to identify network content and supporting methodology and technology that would provide integrated health and environmental data in one source.

The National Tracking Network (http://ephtracking.cdc.gov) officially launched in 2009, and as the network matured, an additional nine states have been funded by the program, (Figure 2). Also in 2009, CDC’s Tracking Program leveraged its long-term partnership with the Association of State and Territorial Health Officials (ASTHO) to further expand environmental health capacity through a fellowship program. This collaboration established mentorship opportunities to assist public health agencies not funded through CDC’s Tracking Program improve their ability to utilize environmental and health effects data and enhance their environmental health surveillance capacity.

**Tracking Content**

The Tracking Program developed standardized data requirements for its content in order to include the information in a national network. CDC collaborated with its national, state, and local partners and data stewards to develop national consistent data and measures (NCDMs). NCDMs consist of indicators and measures, as well as the underlying data required to generate them. Additionally, a metadata standard was developed to provide descriptive information about all data on the Tracking Network.\textsuperscript{6}

The Tracking Network utilizes data from a variety of sources. These data sources collect information to meet the data owner’s particular needs and mandates. Some data are collected
at the national level while other data are collected by state organizations. National partners include the EPA, the U.S. Geological Survey (USGS), the National Cancer Institute (NCI), the National Aeronautics and Space Administration (NASA), and the National Oceanic and Atmospheric Administration (NOAA), as well as other programs within CDC.

CDC’s Tracking Program updates data on specific content areas through one of two calls for data. Data submitted during the fall data call include hospitalizations, emergency department visits, and birth defects. Data submitted for the spring data call include community drinking water and childhood lead poisoning. Data are submitted utilizing a mandatory standardized XML schema, which provides a means for an information system to exchange data. Working with grantee workgroups and other network stakeholders, the Tracking Program has established several processes to support the creation, submission, processing, reviewing and publication of data on the Tracking network. CDC Tracking’s data management staff validates accepted data and follows up with state and local partners if they have questions regarding the submitted data. NCDMs are created by aggregating the data including the state and local partner’s data sets. Small cell counts require an additional data processing step called suppression to preserve the confidentiality of data. Suppression involves blanking out data cells with low counts. State and local partners review their data prior to its publishing on the network. The review and validation of data by both CDC Tracking and its state and local partners protect the integrity of the data throughout the process.

The Tracking Program has organized Tracking Network data into three content sections: Environments, Health Effects, and Population Health. The Environments section contains information on environmental issues, such as water quality and air pollution, that may adversely affect human health. The Health Effects section provides information about health conditions that occurred as a result of exposure to environmental contaminants. Finally, the Population Health section provides supplemental data including demographic information and as well as information on health behaviors.

Datasets included on the Tracking Network vary in size and complexity. The largest dataset on the network is housed under the climate change module and allows the users to view 33 years of temperature data, heat index data, and the number of extreme heat events. Information regarding current content is displayed in Table 1. As the Tracking Network has evolved, the program has continued to add new content areas as well as new years of data to add to the robustness of the content. For example, cancers related to exposure to tobacco smoke -- esophageal cancer, larynx cancer, oral cancer, and pancreatic cancer-- were added to the Network in August 2013. The Tracking Network is in the process of adding data on pesticide exposures which will be obtained from the American Association of Poison Control Centers (AAPCC), to the Environments section.

**Tracking Informatics**

Public health informatics is the “systematic application of information and computer science and technology to public health practice, research, and learning that integrates public health and information technology.” Informatics can facilitate such public health actions as improving environmental conditions and enhancing policies that reduce populations at
The Tracking Network integrates public health and information technology by providing a dynamic query system that allows users to interact with data in a number of different ways including charting, mapping, and graphing. Developing a single surveillance system that can accommodate increasing dataset sizes and meet diverse user expectations all while maintaining a user-friendly interface is a complex task. The Tracking Program has addressed these challenges in different ways since its beginning.

Since the launch of the Tracking Network, the amount of health and environment data housed on the network has grown to over 50 gigabytes of data stored in its database. The challenge is to maintain a system that is powerful and robust enough to handle large and complex health and environmental data and still provide meaningful information to the user. To accomplish this, usability and accessibility were considered from the beginning. The Tracking Program has worked to balance these challenges by leveraging technology to maximize the user’s experience. For example, the addition of 33 years of heat events data created the largest single dataset ever added to the Tracking Network. Querying such a large dataset can cause a substantial lag in response, leaving a network user waiting a long time for the data to display. To solve this problem, the Tracking Program leveraged advances in data storage to accommodate the dataset’s large size. Doing so enabled the Network to display meaningful results within a reasonable response time. In addition, the implementation of the current Network interface in 2011 added the ability for a user to query an interactive map which displays data and information for more than 3,000 U.S. counties (Figure 3). This feature was previously unavailable because it took too long for the data to display.

The Tracking Network is designed to meet the needs of diverse groups of users, each of which have different knowledge and skill sets related to data use. The Tracking Program strives to balance the needs of different user groups while keeping the overall design user-friendly. Although the Tracking Network uses standard data visualization styles familiar to scientific, medical, and public health professionals, these types of data displays can be challenging for non-scientific users.

For people without scientific training, or even scientists and public health professionals with limited training, it can be time-consuming and inefficient to sort through data in an online surveillance system. The Tracking Network has conducted usability testing to better understand the varied needs and capabilities of its users. Sometimes a data consumer does not want complex data details and may only require composite information, or a big picture perspective deriving from one or more of the dataset available on the network. At the same time, more advanced Tracking Network users would benefit from expanded options to include items such as using advanced mapping techniques, the ability or option of being able to view contextual social and demographic information within search results, and displaying multiple measures on a single results page.

The Tracking Program developed several tools and features to provide the user a richer understanding of the information available. One example is the timeline tool. Timeline View is a mapping feature that allows a user to view data over time to spot trends in the flow of data. With this feature, users can adjust years using a slide bar on the display page and watch
the maps update dynamically. This feature is available for any single geographic area with multiple years’ worth of data. Unlike a lot of slide maps available on the Internet, the user is able to fully utilize all features, such as zoom and layering, normally unavailable while timeline mode is active.

Another enhancement that seeks to maximize user experience is the Info by Location tool. This tool allows a user to double-click any geographic area within a map and see demographics data provided by the U.S. Census Bureau. It provides the user with information on age groups, gender, race, and ethnicity. In addition, the Info by Location feature was redesigned to better present information about a single location that is easily understood by all users. To accomplish this, the Tracking Program chose an infographic presentation style for the feature. An infographic is a representation of information in a graphic format designed to make the data easily understandable at a glance. People use infographics to quickly communicate a message, to simplify the presentation of large amounts of data, to see data patterns and relationships, and to monitor changes in variables over time.13

Information technology is fast-paced and ever-changing. As the Tracking Program continues to evolve, it is positioned to take advantage of new and emerging technologies. This ensures that the program and the network stay on the cutting edge of technology while providing an informative user experience. The Tracking Program is constantly exploring new techniques and tools to allow the user to make the best use of network information. For example, the Program is currently exploring a display of multiple measures across content areas in ways acceptable and desirable to users on the basis of user experience.

**Informing Users**

Various studies estimate that 50% to 80% of people utilize the Internet to find health information.14, 15, 16 Additional studies have shown that individuals who are inclined to use the Internet to become informed about their health are more likely to seek medical care for conditions.17 To effectively communicate health information online, a system needs to clearly address its audience.18 Failure to do so may result in a user interface that confuses rather than informs. Usability studies have shown that several factors must be addressed in order to more effectively communicate information on the Web. These include site navigation, terminology, and page layout.19 In spite of this, evidence abounds that individuals were often unsuccessful at searching for this information.20 For this reason, the Tracking Network has been designed to help users access data and information in a straightforward, user-friendly manner.

Communication efforts for the Tracking Network encompass a wide range of activities including audience research, plain language and health literacy considerations, outreach, user testing, materials development, technical support, and training. This holistic approach to communication and outreach has contributed to both the usability and increased use of the Tracking Network since its launch. Informed by usability testing, the Program designed the site navigation, terminology, and page layout to ensure effectiveness in its Web communication. In addition, the program ensures that information on the Tracking Network
is easy to understand by all audience groups, accounting for the diversity within the subgroups related to needs, skills, health literacy, technical competency, and the capacity for data use and data interpretation.

The Tracking Network has also used audience research to improve understanding of its users’ needs. Focus groups, interviews, and national surveys with representative members of audience groups informed the writing of Tracking Network content, as well as the development of supplemental materials like video tutorials, animated maps, and infographics. In addition, CDC and program grantees collaboratively explored and developed specific products for several audience groups including environmental health practitioners, policymakers, data stewards, college-level instructors, health educators, nurses, and librarians.

After writing content for the Tracking Network on the basis of findings from the audience research, the Tracking Program conducted multiple rounds of message testing with some of the audience groups to evaluate how well the content is presented. Using feedback from the message testing, CDC adjusted reading levels and incorporated best practices for plain language and Web content writing for the Network’s content. In addition, the program designed supplemental materials to help users better understand the content displayed on the Tracking Network. These materials include animated maps with brief data interpretation statements, illustrations showing the comparative size of unfamiliar units of measure, and infographics highlighting main messages for several different Tracking Network topics.

In addition to developing content for the National Tracking Network, the Tracking Program develops materials for grantees, partners, and others to use in communicating about the network. For example, CDC created communication toolkits for content areas including asthma, birth defects, carbon monoxide, children’s environmental health, climate change, heart health, and women’s health. Each toolkit contains background information about the topic and its relevance for public health surveillance.

Other Tracking Network communication tools include informational podcasts and videos. The program uses podcasts as opportunities to address frequently asked questions about Tracking Network data and to highlight the utility of the Network for nurses and other stakeholders. Tracking Network videos describe the connection between environment and health and showcase how grantee tracking programs are making important contributions to the health of their communities.

Social media can be used by public health to inform and educate individuals about human health. The program has been active on social media since 2010 and maintains a presence on Facebook and Twitter as a way to stay connected with current users and to gain interest from new audiences. The number of likes and followers has increased with each year, confirming that social media use is a relevant tool for Tracking.

The Tracking Program also contributes to public health workforce development through training. In addition to raising awareness about use of the Tracking Network, the program has created training products for professionals and students. Web-based training courses developed for professionals describe the major components of environmental public health
tracking and how local health departments and community-based organizations can work with grantee tracking programs. “Keeping Track, Promoting Health: Classroom Modules” is designed for college-level instruction on environmental public health using the Tracking Network. Future projects and products under consideration include the development of an e-book, one of the first for CDC, conducting additional usability testing for Tracking Network content and functionality, and expanding the number and variety of communication tools and resources available on the Tracking Network.

Impact and Discussion

Since its inception in 2002, the Tracking Program has worked to combine health and environmental data streams where few or no links existed previously. Five years after its launch, the Tracking Network has emerged as a robust tool for environmental public health that allows a variety of users the opportunity to learn about how their environment impacts their health. It provides surveillance data for public health practitioners and policy makers and serves as a unique platform of health and environment data that academicians and other researchers can use for hypothesis generation.

Having established the informatics and data visualization framework for the Network, the Program’s focus has expanded to include data linkage and utilization projects to generate hypotheses and demonstrate relationships between the environment and certain health outcomes. For example, the Tracking Program has partnered with several academic partners to enhance utilization of tracking data. Projects have included a linkage study examining the relationship between drinking water contaminants and adverse birth outcomes as well as an ecological study of ambient air quality and childhood lead exposures. In addition, tracking data are being utilized to inform programmatic activities. Health and environmental agencies and other interested stakeholders have used data from the Tracking Network to support and drive public health actions at the state and local levels. Successful applications have included targeting outreach and prevention activities, identifying communities at risk, informing policies, and impacting city or state planning.

The Network will continue to change and improve over time by adding new data and by using emerging technologies. Maintaining a quality user experience on the Network will also continue to be a priority as the technology evolves. Future enhancements for the Network will make the information display more intuitive, helping all users to better understand the data. As the Network further matures, the data will continue to fuel scientific research and the findings it produces will support public health activities and interventions.

References


10. O’Carroll, PW.; Yasnott, WA.; Ward, ME.; Ripp, LH.; Martin, EL.; Ross, DA. Public Health Informatics and Information Systems. NY: Springer; 2002.


Figure 1.
Environmental Public Health Tracking
Figure 2.
Map of National Tracking Network Grantees, 2014
Figure 3.
Tracking Network Query Response for Extreme Heat Days and Events.
### Table 1

**Tracking Network Content Area**

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Number of Nationally Consistent Data and Measure</th>
<th>Years available</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Environments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate Change</td>
<td>25</td>
<td>1979–2011 *</td>
</tr>
<tr>
<td>Community Design</td>
<td>32</td>
<td>2000–2012 *</td>
</tr>
<tr>
<td>Homes</td>
<td>23</td>
<td>2000–2011 *</td>
</tr>
<tr>
<td>Air Quality</td>
<td>16</td>
<td>2001–2011 *</td>
</tr>
<tr>
<td>Community Water</td>
<td>61</td>
<td>1999–2012 *</td>
</tr>
<tr>
<td><strong>Health Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma</td>
<td>13</td>
<td>2000–2011 *</td>
</tr>
<tr>
<td>Birth defects</td>
<td>26</td>
<td>1998–2011 *</td>
</tr>
<tr>
<td>Cancer</td>
<td>62</td>
<td>2001–2009</td>
</tr>
<tr>
<td>Carbon Monoxide Poisoning</td>
<td>9</td>
<td>2000–2011 *</td>
</tr>
<tr>
<td>Childhood Cancers</td>
<td>16</td>
<td>2001–2009</td>
</tr>
<tr>
<td>Childhood Lead Poisoning</td>
<td>14</td>
<td>2000–2011 *</td>
</tr>
<tr>
<td>Heart Attack</td>
<td>3</td>
<td>2000–2011</td>
</tr>
<tr>
<td>Reproductive and Birth Outcomes</td>
<td>16</td>
<td>2000–2010</td>
</tr>
<tr>
<td><strong>Population Health</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population Characteristics</td>
<td>9</td>
<td>2000–2011 *</td>
</tr>
<tr>
<td>Biomonitoring: Population Exposures</td>
<td>14</td>
<td>1999–2008 *</td>
</tr>
<tr>
<td>Children's Environmental Health</td>
<td>88</td>
<td>2000–2011 *</td>
</tr>
<tr>
<td>Health Behaviors</td>
<td>6</td>
<td>2000–2012</td>
</tr>
</tbody>
</table>

* Not every measure in this content area may have the data for all available years