

Assessment of Options for Tracking of Drinking Water Contaminants and Relevant Data Sources

Miriam Rotkin-Ellman, Amy D. Kyle PhD, MPH, and
John R. Balmes MD

September 2005

A working paper produced by the

Berkeley Center for Environmental Public Health Tracking
School of Public Health
University of California Berkeley

This publication was supported in part by cooperative agreement number U50/CCU922409 between the Centers for Disease Control and Prevention and the Regents of the University of California. Its contents and the views expressed are solely the responsibility of the authors and do not necessarily represent the views of the funding agency.

Executive Summary

To better understand the links between environmental contaminants and related diseases, the Centers for Disease Control and Prevention (CDC) and many states are working to develop a surveillance system known as the Environmental Public Health Tracking (EPHT) network. This network would provide a means to collect, integrate, analyze, and disseminate data on environmental hazards, human exposures and associated health effects.

There is considerable interest in including human exposure to contaminants in drinking water in the EPHT network.

This report describes conceptually the benefits and limitations of different approaches to tracking drinking water quality and examines the sources of water quality data that could be used to do so. The methods utilized included informal interviews with key informants coupled with relevant background research. Key findings include:

- The application of the tracking conceptual model and data integration methods can provide valuable information on the health threats posed by drinking water resources and how they can be prevented.
- Consideration of the scope and limitations of a study approach is a critical component of tracking method development.
- There is considerable promise for expanding the breadth of a tracking network to include ambient ground and surface water quality data.
- The development of methods for data sharing and integration is a key component of tracking network development.
- Fostering collaboration between and within states that includes the sharing of expertise, ideas, and resources can strengthen state programs and the development of a national environmental public health tracking network.

Introduction

The Pew Environmental Health Commission's 2000 report on the state of environmental public health in the United States highlighted the need for additional information about exposures to environmental contaminants and related diseases. To this end, the commission recommended the establishment of a nationwide Environmental Public Health Tracking Network (EPHT) similar in function to infectious disease surveillance systems already established. (Pew Environmental Health Commission 2000) Since 2002, the Centers for Disease Control and Prevention (CDC) have worked to develop this type of network through implementation of environmental public health tracking projects. The University of California Berkeley Center for Environmental Public Health Tracking has worked with the western states initially involved in the tracking initiative (California, Oregon, New Mexico, Montana, Nevada, Utah, and Washington) to contribute to the development of approaches and methods for environmental public health tracking.

The CDC's model of Environmental Public Health Tracking (EPHT) includes three types of data (about environmental hazards, human exposure to environmental hazards, and health effects) that when combined will provide a basis for reducing the disease burden of environmental contaminants. Data collected in each of these arenas represents points along the causal pathway from a hazardous agent present in the environment to a clinically adverse health effect. (Figure 1) (McGeehin et al. 2004) A national EPHT network, as conceived by CDC, would create a standardized and ongoing means to collect, integrate, analyze and disseminate this data. Data to be incorporated into the network include measurements of contaminants in the environment, contaminants in human bodies (biomonitoring) and the incidence of associated diseases among others.

From the conceptual stages of building a national EPHT network to the awarding of grants for pilot projects, human exposure to contaminants in drinking water has been considered a topic of interest for inclusion in the network. Drinking water was identified as a potential source for agents that cause disease early in public health history during investigations into outbreaks of cholera and other infectious diseases. Concern for chemical and bacteriological contamination of drinking water resources in the United States led to the passage of the Safe Drinking Water Act (SDWA) in 1974 and subsequent amendments in

1986 and 1996. The SDWA establishes health based standards and monitoring schedules for chemicals and bacteria in drinking water that is supplied through public water systems. Public water systems are defined as those that have at least 15 service connections or serve at least 25 people per day for at least 60 days of the year. (US EPA 1999)

In the western region, all states except Wyoming have the authority to implement the SDWA under the auspices of the federal Environmental Protection Agency (EPA) and are responsible for collecting the water quality monitoring data required by the SDWA. The state reports, which contain data on water systems' compliance with the relevant monitoring schedules and achievement of the appropriate standards, are collected by EPA in the Safe Drinking Water Information System (SDWIS) database.

In the Western region, limited groundwater and surface water resources coupled with increasing populations have placed strains on existing drinking water sources. In addition, according to public health officials, changes in federal health based drinking water standards for contaminants such as arsenic have generated a great deal of concern about the potential health threats posed by drinking water across the region. (Layco G, personal communication, Magraw M, personal communication, Flowers L, personal communication) Preliminary investigation among the western states currently involved in the tracking initiative revealed considerable public concern about the safety of drinking water resources. The need to address this significant public concern along with the potential for exposure to disease-causing agents contained in drinking water has generated a great deal of interest in applying the tracking model of data collection, integration and dissemination to describe and prevent diseases associated with drinking water contaminants.

The CDC model of hazards, exposures, and health effects provides a general framework from which an approach could be developed. Given the universe of water quality and health information available on national and state levels, there remain questions of how best to approach tracking for drinking water quality.

The purpose of this report is twofold. First, an analysis of the various project designs currently being implemented or proposed on the state level is used as a means to describe conceptually the benefits and limitations of different approaches to tracking drinking water quality. Secondly, the sources of water quality data that could be used to build a tracking network are examined as a starting point for future projects.

Methods

The information and perspectives presented in this report were gathered from informal interviews conducted in person, over the phone, and over email with key informants from each of the state partners, the Environmental Protection Agency (Region 9), and the Centers for Disease Control and Prevention. A full list of the informants is included in Appendix A. The interviews were structured around common themes and included discussion of the applicability of tracking to drinking water quality data, current and proposed projects, and perceived challenges in establishing a tracking initiative for drinking water. The authors also participated in regional project planning discussions held at a meeting of the seven Western States involved in the environmental public health tracking project in February of 2005 and continued through conference calls and email correspondence. Additional research into the availability of water quality data was conducted through investigation of regulatory requirements, Internet searches, and email correspondence with representatives of agencies with data sources, such as a representative of the United State Geological Survey (USGS).

Our research into the development of a tracking network for drinking water involved two primary areas of investigation. We first report on the information shared by informants on how the general CDC tracking model could be translated into actual projects and studies in participating states. We then examine the sources of water quality data and methods of data integration that could be used to build a tracking network.

Applying the Tracking Model To Drinking Water Data

Discussions with informants about applying the general tracking framework as defined by CDC to the question of drinking water quality and associated health impacts centered on four main themes:

- Is the tracking or surveillance approach useful for understanding links between drinking water contaminants and health effects?
- What are the different approaches to developing a tracking framework either currently being pursued or proposed for future work?
- What are the limitations of applying this tracking framework and factors that states consider when selecting projects?

A. Value of EPHT for Drinking Water Data

In general, informants expressed interest and support for the need for a tracking network to enhance the capacity of state health and environment programs to monitor contaminants known or suspected to cause adverse health impacts and build understanding of potential new threats. Additionally, they provided examples of how the application of the principles of tracking to water quality and health effects data offers benefits beyond the specific scope of an individual project and increased the ability to understand the links between environmental exposures and disease. These benefits included increasing the capacity of existing data sources to be used for future investigations, investing in data sharing and database technologies, increasing the depth of investigations by integrating multiple data sources, and fostering collaboration and communication across different agencies.

The focus on working to improve the capacity of existing sources of data to yield new information in the links between environmental exposures and health outcomes contained in the tracking framework appealed to informants. Generating new insights from existing data sources is generally preferred as a way to maximize the limited financial and staff resources found in some state environmental and health agencies (Flowers L, personal communication). Refining or adding additional data to existing databases was cited as a means to increase the ability to conduct future projects. Specifically, Layco (personal communication) and Flowers (personal communication) emphasized the value of “cleaning” and adding geographical

location information (geo-coding) to existing drinking water quality and health effect databases. Addressing and remedying data quality issues while adding the ability to include a spatial element to the data analysis transform the existing databases into powerful tools for further investigations. (Layco G, personal communication and Flowers L, personal communication)

The analysis of multiple different types and/or sources of data required for any type of tracking system is reliant on data management, consolidation, storage, retrieval and integration technology. Further discussion of this technology and approaches to data management are covered in the following section. However, the investments in hardware, software and staff training made towards individual tracking projects can increase the technological capacity of the agency for future projects. Specifically, the tools developed to manage and share databases specifically for water quality and health data could be adapted for other types of data. (Cude C, personal communication)

A key piece of the tracking model is to integrate multiple sources and types of data in order to provide a greater depth of analysis of the link between environmental exposures and health effects. Layco (personal communication) expressed that this depth of analysis conferred benefits beyond the scientific and regulatory communities increased understanding of the relationship between specific exposures and health effects. She cited examples of concerned communities finding study results, achieved through the integration of multiple data sources, to be more legitimate. In particular, these examples pertained to the use of tracking approaches to reexamine areas of concern such as a cancer cluster or areas of known exposures. (Layco G, personal communication)

At an essential level, developing a tracking initiative for drinking water quality involves building collaborations between agencies focusing on environmental quality, human exposure to contaminants in drinking water, and the evaluation of associated health effects. Magraw (personal communication) expressed that these types of collaborations are one of the benefits of tracking because it “gets people to look outside their own little world in health and environment” and improves integration and communication between health and environment programs. (Magraw M, personal communication) State tracking projects have taken different approaches to developing these types of collaborations. Sharing staff resources either through paid staff in other programs (Layco G, personal communication and Flowers L, personal

communication), representation of both programs on the tracking project team (Barck L, personal communication), or joint informational meetings between programs (Magraw M, personal communication) have been found to be helpful in facilitating information transfer beyond the specifics of individual tracking projects. Through stakeholder meetings, planning consortiums, mini-grant programs, and outreach projects, tracking programs have developed collaborations throughout their representative states.

B. Designing an Approach

Discussions with informants surrounding tracking projects currently funded by the CDC tracking initiative or ideas for future studies provided examples of a variety of different approaches to designing a tracking project. One of the challenges to developing a tracking network for drinking water quality was described by Magraw (personal communication) as “defining what to look at.” This gets at the heart of how to organize the large quantity of data available in a useful manner to better understand sources of potential health threats.

Examination of current and proposed tracking projects related to drinking water quality revealed approaches that differed on the point in a continuum, from ambient water quality to health effects, on which they focused. Evaluation of the different foci includes consideration of both the goals of a tracking initiative and methodological limitations stemming from the data available. In order to organize these approaches, we grouped them according to the focus point, although a single project may utilize more than one approach and the approaches may overlap. Four possible foci are health effects, exposure mechanisms, tracking specific contaminants of concern, and geographical area.

Health effects

A tracking project that focuses on health outcomes seeks to combine data on a disease that has an established or suspected environmental etiology, with a measurement or approximation of exposure. In 2002 the CDC Environmental Public Health Tracking Project funded a pilot project in NM that used a health effect-focused approach to developing a tracking system using drinking water quality data. The project intends to explore the ability to track the potential for disease due to exposure to drinking water contaminants by examining the associations between cancer incidence and drinking water arsenic levels. The

selection of this project was driven by concerns over arsenic levels in drinking water, availability of data sets for both cancer incidence (from the state tumor registry) and drinking water concentrations (from the state drinking water monitoring agency), and an established biological link between the exposure and the health effect. (Flowers L, personal communication)

This approach to a tracking project is expected to provide a “ecologic analyses of cancer incidence rates per drinking water arsenic levels.” (Flowers L, personal communication) In addition, the spatial component of this analysis may be able to identify geographical areas experiencing higher rates of cancer and make some suggestions for further evaluation of the associated potential environmental exposures. Flowers (personal communication) also pointed out that the work put into mapping the distribution systems (in order to assign exposure level to census tracts) and geo-coding the tumor registry created powerful tools for further investigations. An additional project using this approach was suggested by Layco (personal communication). She suggested that a project linking bacteriological levels in SDWIS with health outcomes monitored by the National Electronic Disease Surveillance System (NEDSS), an existing national infectious disease surveillance program, might address some of the concerns over bacterial contamination brought up by local health departments. However, Layco (personal communication) emphasized that this was not considered a priority for further work.

There are a number of technical limitations with using this approach for further tracking initiatives. While in NM the investigators were able to gain access to an adequate health effects database, informants from other states expressed frustration at the lack of good health databases. (Barck L, personal communication and Magraw M, personal communication) In New Mexico, the administrative code defines cancer as one of the “Reportable Conditions of Public Health Significance”. (NMTR 2002) This provides a strong legal basis for the collection and sharing of cancer incidence data collected in the New Mexico Tumor Registry. Comprehensive tumor registries are not available in all states, and confidentiality issues have to be addressed to allow for the sharing of available health effect information.

Though this was not the case for arsenic and cancer, an obstacle to focusing a tracking initiative on many health effects is the difficulty in establishing a good biological basis or link

between the data available on contaminant concentrations and health effect incidence (Barck L, personal communication). In Oregon, early discussions of potential tracking project topics included investigating the health effects associated with nitrate concentrations in drinking water due to the substantial data set available from the monitoring of drinking water resources for nitrate contamination. However, the priority of this project was lowered due to concerns about the strength of the biological link to diseases such as methemoglobinemia and cancer. (Barck L, personal communication)

Aside from the limitations posed by the availability of data sources, Macler (personal communication) raised a methodological concern with focusing on health effects associated with exposure to contaminants in drinking water. He noted that the drinking water standards contained in the SDWA are developed with a margin of safety, based on toxicological evidence, to prevent disease occurrence. Additionally, the majority of people in the United States consume water that is subject to and below these standards. Taken together this would result in a very small population exposed to concentrations of contaminants that would be likely to cause adverse health effects making it very difficult to achieve the power necessary in a study to see associations between contaminants in drinking water and adverse health effects. (Macler B, personal communication)

While many of the projects discussed by informants focused on linking disease incidence to environmental exposures, the tracking model also calls for the development of methods for identifying and tracking relevant health effects in the absence of exposure data. The concerns raised by informants about the lack of good data sources on health effects speak to the importance of this endeavor. Projects aimed at developing or improving tumor registries, birth defect surveillance, and other relevant health measures could greatly enhance the capacity of a tracking network for drinking water.

Exposure Assessment

The question of how to assign environmental exposure to drinking water contaminants is an essential component of studies assessing the associations between health effects and drinking water. It is also an important component of the CDC environmental health tracking model. Concerns over the inability to accurately measure or estimate the exposure to drinking water contaminants experienced by individuals have resulted in focusing the design of some

tracking projects in this area. These approaches assess what Wolff (personal communication) referred to as the “human component” that he perceives as missing from methods that focus exclusively on measures of water quality. These projects differ from the approach employed by the NM study because they don’t rely on the water quality data collected as mandated through the Safe Drinking Water Act (SDWA) and stored in the SDWIS database, as the only means of measuring exposure.

Under the SDWA, the monitoring of the quality of public water systems is conducted by sampling the water prior to its distribution to individual residences. This raises concern as to the accuracy of using these measurements to assess individual exposure to contaminants that may be contained in the drinking water. Variability in the quality of the water throughout the distribution system may exist that is not reflected in the results reported to the state. In California, a study was undertaken to investigate this variability in order to better understand how water distribution systems influence the drinking water quality experienced by individuals. The results of this study revealed substantial spatial variation in the quality of the drinking water within large water distribution systems and that the variation was very difficult to predict based on the geography of individual residences. The conclusion drawn from this study was that the data collected for SDWIS alone is not sufficient to describe the exposure to drinking water contaminants and further exposure assessment is required. (Wolff C, personal communication) VanDerslice (personal communication) emphasized that it is not sufficient to assume that data reported to SDWIS is a sufficient measure of exposure without investigating the potential for variation within water systems. In Washington, concern about variability within water distribution systems has led to efforts to delineate the geographical service area of water systems to increase the exposure assessment of further investigations into drinking water quality. (VanDerslice J, personal communication)

The main benefit of focusing a tracking initiative on assessing and describing the exposure to drinking water contaminants is that you will achieve more precision in estimating the potential of a health threat if by taking into factors that influence an individuals exposure, you can better estimate exposure. These factors could be a result of the design of the water distribution system (Wolff C, personal communication) or individual behaviors such as tap water consumption and moving residences. In addition, this approach would allow you to address concerns regarding seasonal variations in water sources and quality. (Magraw M,

personal communication) Wolff (personal communication) added that looking at exposure through a better understanding of the variability within water distribution systems creates the potential for greater participation of water distributors, an important stakeholder, in developing tracking projects as it relies on their expertise and data.

The technological advances of measuring contaminants or metabolites of contaminants directly in humans, known as biomonitoring, provides an additional tool that can be used to measure exposure to contaminants found in drinking water. The Rocky Mountain Biomonitoring Consortium has collaborated with tracking initiatives within its participating states on a project involving measuring arsenic levels in urine. Biomonitoring has also been proposed as a means to validate the exposure estimated by data collected for SDWIS. (Flowers L, personal communication) Additional biomonitoring projects have also been proposed to assess exposure to drinking water contaminants. The Rocky Mountain Biomonitoring Consortium has submitted a grant proposal to EPA for a project comparing arsenic concentrations in drinking water and in urine. In addition to improving the biomonitoring capabilities of the grantees, this project is also aimed at improving the understanding of the degree to which biomonitoring can be used as a measure of exposure in a tracking network. Layco (personal communication) also suggested a project linking lead in drinking water to blood lead levels as feasible but not necessarily a priority due to drinking water not being the primary route of exposure to lead.

While generally there was agreement that data collected for SDWIS only provides an approximate estimation of individual exposure to contaminants in drinking water, the degree to which a focus on exposure assessment is needed was debated. Macler (personal communication) questioned whether the variability in water quality within a water distribution system was large enough to fall outside the safety margins of drinking water standards and therefore translate into significant differences in health outcomes. In addition, the variability in water quality within water systems is a reflection of water distribution system complexity. (Wolff C, personal communication) Smaller water systems and exclusive ground water systems are not thought to have the same level of variability and therefore the SDWIS data might be suitable proxy for exposure. (Flowers L, personal communication) While Wolff (personal communication) emphasized the need to better describe exposure to drinking water contaminants as a key part of developing a tracking system for drinking water, he also

explained that he has observed a hesitancy to include this approach due to the perception that it is too overwhelming and complicated.

Additional tracking projects have been proposed which focus on measuring or describing human exposure to contaminants in drinking water. Wolff (personal communication) and VanDerslice (personal communication) developed a proposal for supplemental funding to address this component of tracking they felt was missing in other projects. This proposal focused on developing methods to better predict the quality of water an individual is exposed to at the tap. Their proposal identified the need for methods to link individuals or small groups to specific water supplies within larger distributional systems and characterizing the variations in water quality within water systems and over time.

Contaminant(s) of concern

While the previous two approaches to designing a tracking project focused on an individual element of the CDC tracking model, some tracking project(s) are focused on a specific contaminant(s) of concern and work across the hazard, exposure, and health effect continuum. Contaminants of concern could be individual compounds highlighted because of regulatory actions such as standard setting (perchlorate), lowering existing standards (arsenic), or suites of compounds that might be considered indicators of the potential for causing poor health outcomes. Multiple tracking projects have focused on exposure to arsenic in drinking water due to the significant population in NM and other western states that has drinking water with arsenic concentrations above what is expected to result in adverse health effects according to the new EPA drinking water standard. (Flowers L, personal communication) Other compounds suggested by informants for tracking projects were: disinfection byproducts, nitrate, bacteria, heavy metals, pesticides, and lead. By focusing on a specific contaminant or suite of contaminants a tracking initiative could provide significant insight into a variety of aspects of drinking water quality and health concerns. Principally, it highlights and prioritizes the increased understanding of contaminants with known or suspected health effects and at the concentrations believed to cause the most risk. As with the example of the NM project, this approach provides a mechanism to respond to new hazard evaluation science for unregulated compounds or the lowering of existing standards.

Focusing on specific contaminants can also identify geographical areas where there are high concentrations, so called “hot spots”. Flowers (personal communication) suggested that the geo-coding of the drinking water quality data used for the NM study could be used to provide a geographical distribution of high levels of arsenic. This distributional analysis would be expected to reveal geographical areas where potentially threatened private well sources might be found. (Flowers L, personal communication) The capacity to identify these “hot spots” could be greatly enhanced by expanding the sources of water quality data to include ambient water quality data that could highlight potentially hazardous water sources. In Montana, including ambient environmental water quality data in tracking projects has been used to identify geographic areas of potentially high exposure to arsenic for sampling as part of the bio-monitoring project (Magraw M, personal communication). Identification and recognition of contaminant “hot spots” might also help address the concern brought up by Magraw (personal communication) of the need to identify and reduce the health threats associated with people using private wells in areas of known contamination.

A tracking project focused on an individual contaminant of concern may be limited in its capacity to accurately describe the quality of drinking water or predict the potential for health outcomes by the data and hazard evaluation science available on that individual contaminant. Focusing narrowly on a single or small set of contaminants may also not create a means for further analysis of other contaminants due to the exclusion of data sources that do not contain the contaminant(s) of concern. In response to the limitations posed by tracking individual contaminants, including a suite of contaminants or “indicators” of water quality into a tracking project has been proposed.

The concept of using a suite of contaminants as indicators of the quality of a resource has been well established in the ecological field and in recent years has gained considerable interest in the environmental health arena. The CDC, National Center for Environmental Health, the state of California and other groups have developed a number of different suites of indicators that could be used to evaluate the quality of drinking water resources or identify drinking water that could pose a health risk to consumers. These indicators include direct measurements of the concentrations of chosen contaminants as well as summary measures of exceedence of established levels of concern. (CDC-NCEH 2003, CA DHS-EHIB 2002, VanDerslice J, personal communication) Indicators could incorporate the full list of regulated

compounds in drinking water or be restricted to identify specific health risks. An example of the latter is VanDerslice's (personal communication) proposal for drinking water indicators that would identify Public Water Systems and the populations served where the health based maximum contaminant level goal (MCLG) for any chemical contaminant which is a known or suspected carcinogen (Category A or Category B1) was exceeded.

Just as with the projects focusing on individual contaminants of concern, the use of indicators could be expanded from an analysis solely of drinking water quality to include data on ambient water quality. The inclusion of prevalence data on the "indicator" compounds from ambient monitoring databases would provide a more upstream analysis and might be able to identify geographical areas that could pose a threat if used as drinking water before the threat was realized in a drinking water supply.

Geographical area

A key aspect of establishing an environmental public health tracking network is the integration of geographical identifiers for environmental exposures and causes of health effects. While there is a geographical component to all the other tracking approaches described previously, a slightly different approach to developing a tracking initiative would be to identify geographical areas of concern and create a system for integrating water quality and/or health effect information that describes this area. This geographical area could be delineated by the occurrence of a specific health effect, a hydrologic unit, or political boundaries.

Currently, tracking funds and methods are being used to support small-scale projects in Utah and Montana that focus on specific geographical areas. In Utah the tracking methods of data integration and sharing are being used to conduct a reassessment of a cancer cluster. The project involves the integration of additional data sources including contaminant plume mapping along with data from the initial investigation. (Layco G, personal communication) In Montana, EPHT has funded a county to do a small community based project to sample wells, map results and conduct an educational campaign. (Magraw M, personal communication) In addition, ambient environmental monitoring is often organized on a geographical scale according to hydrologic units. Examples of this include the USGS, National Water Quality Assessment as well as the EPA, Clean Water Act (CWA) 303(d) list of impaired waters.

Focusing on a specific geographic area has advantages from a number of different perspectives. The enhanced capacity to address the concerns of a specific community surrounding potential threats in their area may appeal to community level stakeholders. (Layco G personal communication, Magraw M, personal communication) Magraw (personal communication) found this to be true when using tracking funds and methods of data integration to conduct an environmental health assessment that included community trainings was able to improve the communities trust in the quality of the drinking water. From a research perspective, the ability to integrate multiple data sources has improved the capacity of cancer cluster investigations. (Layco G, personal communication) The development of a tracking network that incorporates ambient environmental quality data could be assisted by using a geographical focus since the geographical scale easily aligns with other established monitoring and data sharing systems.

Focusing on a geographical area may be more applicable in some cases than others. For example, this approach is dependent on the breadth of geo-coded data available to be included in the tracking system. Some states report having adequate coverage of geo-coded water quality data resources while others are still working to achieve this. Even within a state there may be areas that are lacking in data for which this approach would be able to provide minimal information. Finally, incorporation of all the available water quality for a geographical area may not provide adequate information about health risks if the data available is not easily linked to health effects.

A few project ideas were suggested by informants that focused on a specific geographical area. Flowers (personal communication) suggested a comparison of health effects in Albuquerque before and after the city switches from ground water to surface water source of drinking water. VanDerslice (personal communication) was interested in expanding the capacity of cancer cluster investigations by developing software that would “tap into” the Department of Ecology database and other water quality data sources such as USGS to improve the characterization of drinking water quality. This type of approach could also be expanded to include hydrologic units used to describe ambient water quality such as basins, watersheds, etc. or more specifically the water quality in a source water assessment area could be tracked over time. Additionally, the geographical area of concern could be

identified as the hydrologic units (segments) identified by states for listing as impaired under 303(d) of the Clean Water Act (CWA).

C. Limitations of Tracking

While each of the approaches described earlier have specific limitations, there are challenges that were brought up by the informants that are applicable to developing a tracking system for drinking water in general. These challenges include those associated with prioritizing drinking water quality as a hazard, consolidation of conflicting assessment methodologies, and data integration and consolidation.

Although, as discussed earlier, informants reported high levels of public interest and concern about drinking water quality and health effects, this concern was not considered uniform. VanDerslice (personal communication) noted that the level of concern over drinking water quality varied geographically such that the concern in rural areas served by smaller water systems is often higher than urban areas with large water systems. Other informants expressed that exposure to potential health hazards through drinking water is often considered by health and environment agencies to be a lower priority for investigation. (Wolff C, personal communication, Macler B, personal communication, Barck L, personal communication, and Fox-Williams K, personal communication) Reasons for this ranged from the belief that overall the drinking water standards (MCLs) are low enough to not result in adverse health effects (Macler B, personal communication) to feeling that other routes of exposure to contaminants such as air pollutants are a greater priority where attention should be focused. (Fox-Williams K, personal communication)

According to Macler (personal communication), projects that are based on the integration of health and environment data pose challenges associated with the different assessment methodologies used in the health and environmental fields. He stressed the importance of understanding the difference between the level of risk reflected in the toxicological methodologies used to set drinking water standards and those that can be detected by the epidemiology used to measure health effects. Relatively low levels of risk allowed by drinking water standards mean that studies could have severely limited the power to see results unless they are very large. The challenge to adapt existing epidemiological and

toxicological methodologies to perform effective surveillance is faced by all projects but does not preclude the development of new methods.

Regardless of the focus of a tracking network, it is built upon multiple sources of data resulting in limitations and challenges associated with data integration. First and foremost the major concern was finding data sources that could be used for tracking. Multiple informants recognized that addressing the lack of availability of health effect data is a challenge to all tracking projects. (Barck L, personal communication and Magraw M, personal communication) In addition to locating sources of data, some of the existing data sources need considerable investment to make integration with other data sources possible. While data exists on water collected from private wells in Utah, much of it exists as paper records and the logistics of transcribing it would be difficult generating little “excitement” about looking at this data.

Once databases have been identified, there are challenges associated with integrating data from data collected under the auspices of different programs and having different “owners” (Cude C, personal communication, Layco G, personal communication, and Flowers L, personal communication). The North West Water Quality Exchange was designed specifically to enable multiple sources of data to be accessed from a single platform without the individual users giving up control over the data. Cude (personal communication) sees this technology as key to tracking initiatives because they enable participation from a wide range of sources while addressing the concerns of the individual data owners. This approach however does not address issues associated with the quality of the data. Other states have approached the data quality problem by requesting data sets from other entities to then be utilized by the tracking project. This is the approach utilized by Flowers (personal communication) in the New Mexico study of arsenic and the tumor registry, where the drinking water quality data was provided by the Drinking Water Bureau of the Environment Department to be analyzed by the tracking project. Layco (personal communication) explained that offering to clean up the data can be used as a powerful incentive to encourage these type of data transactions. However, this approach is very labor and resource intensive and can be prohibitive to making new linkages. The NW Water Quality Exchange model of a unifying platform for disparate data sources also faces the challenge of developing methods for integrating the many different analyte and method names used by different data providers.

The NW Water Quality Exchange addressed this through use of an abstracting service which allows for searches on multiple forms of an analyte name. However, this process is cumbersome and a more sustainable approach would be a national discussion on data stewardship and the adoption of consistent names and methods. (Cude C, personal communication)

The integration of the data into a resource that is useful for tracking, and the analysis of the data require significant technological investment. One of the limitations brought up by many informants was access to both technology and the expertise to facilitate the types of projects they envisioned. (Layco, G, personal communication and Magraw M, personal communication) Specifically Layco (personal communication) highlighted technological assistance as a barrier for states to be able to move beyond pilot projects to data linkages that could be ongoing. Cude (personal communication) in his work with the Northwest Water Quality Exchange, has found the currently available technology to be effective and useful but resource intensive in terms of both the cost and the expertise required. He identified the development of more affordable and accessible technology as key to enable the large scale integration of multiple data sources. On an institutional level, Cude (personal communication) identified the centralization of the information technology (IT) governance as a critical asset in Oregon and Washington that enabled the development of the NW Water Quality Exchange.

Collaboration within states, between states, and with federal agencies offers significant advantages when tackling the types of data challenges discussed above. As part of the CDC tracking planning grants the states have conducted inventories of databases available in their state. These inventories have identified significant sources of data on environmental water quality information and generally less robust sources of health information. Regional collaboration has the potential to utilize the expertise and technological resources of other states. Federal agencies such as the EPA and USGS also have extensive data resources and expertise that could be utilized in a tracking project. Magraw (personal communication) reported that a tracking project related to water quality was suggested as a good topic for collaboration between Montana and the USGS. While collaborations both within states and between states offers means of overcoming some limitations, Magraw (personal communication) pointed out that the differences in infrastructure between states with greater

access to state funding versus those reliant almost entirely on federal funds can create barriers to collaborations between states.

Beyond the details of the source, quality and logistics of the data are issues pertaining to the distribution and comprehensiveness of the data that will be incorporated into the tracking system. Layco (personal communication) pointed out that while the drinking water quality data provided by SDWIS covers the majority of the population in Utah, there is limited monitoring of water quality for the minority of the population in rural areas. She added that many of the rural health departments have expressed frustration about the lack of information on water quality in their areas. While the distribution of water quality information varies between states and between contaminants, generally there is more data from larger water systems than smaller water systems because the monitoring requirements are more extensive and less from private wells for which monitoring is not typically required. Increasing concern about the water quality in rural areas reliant on small water systems and private wells could be addressed by expanding the scope of a tracking project beyond drinking water quality data reported to SDWIS. This can assist in the identification of areas of known contamination but as Magraw (personal communication) pointed out, inadequate and inaccurate information on the location of private wells can result in exposure to contaminants even in areas of known contamination. Although the incorporation of ambient water quality data for surface and ground waters adds a greater breadth to a tracking network, the issue of unequal data distribution may persist and merits consideration.

Discussions with informants revealed significant interest in pursuing the development of a tracking network for drinking water quality along with recognition of the challenges faced in this endeavor. Informants also described varying approaches to the developing such a network currently being undertaken or proposed by state EPHT programs. Evaluation of these different approaches provides critical insight into the breadth of study and project designs that could be incorporated in the EPHT network. In addition, this type of analysis can provide the framework from which an evaluation of the data sources that could be used to build the network is conducted.

Building a Tracking Network of Water Quality Data

The tracking model is based on the linear causal pathway starting with introduction of a hazardous agent into the environment resulting in an exposed population and culminating in an adverse health impact. (McGeehin et al. 2004) When applying this pathway to the assessment of drinking water quality, the sources of data available can be divided into categories according to the scope and rationale for the monitoring. The monitoring programs of many different entities, ranging from national regulatory and research agencies to small-scale volunteer monitoring programs, collect data on water quality. These monitoring programs include measuring levels in the environment at large, levels in areas delineated as sources for drinking water systems under the Safe Drinking Water Act (SDWA), and water supplied to consumers through a water system. (Figure 2) Some of the data produced as a result of this monitoring has been organized into accessible databases, which could be incorporated into a tracking initiative, on many different scales. In this section we will outline the different types of water quality databases and highlight the issues associated with incorporating these different types of data into a tracking system. By integrating multiple sources of data, the EPHT network has the potential to serve as a powerful tool for identifying and preventing negative health outcomes associated with drinking water quality.

Initially, water quality databases were created to store the data collected for a single purpose, such as ecological monitoring required under the Clean Water Act (CWA). The data in these databases was generally collected using the same methodology and with a scope defined by the study or regulatory requirement. Examples of this type of database include small scale monitoring of a single water resource to statewide mandated reporting of water quality from private wells upon sale of a property in Oregon. The development of new computing and mapping technology expanded upon these individual study or regulatory program specific databases and facilitated the development of databases that could describe the water quality within a state or for a specific type of water resource. Database inventories conducted during the planning stages of state tracking projects have the potential to locate both the single purpose and integrated types of databases. Collaboration and outreach to different types of regulatory and non-regulatory entities is required to make a thorough inventory of the types of data available to be included in a tracking network.

Regulatory requirements differ between surface and ground water and often result in the development of separate databases. Generally speaking, surface water data collection is mandated through federal laws while groundwater protection regulations are primarily state undertakings. Examples of this can be found in states throughout the Western Region. The Ground-Water Information Center (GWIC), maintained by the Montana Bureau of Mines and Geology, is a database describing the state's ground water resources including water quality and quantity. (MTECH 2005) The California Ground water ambient monitoring assessment program was developed in response to state legislative action and includes state health and environmental agencies, USGS, and national laboratory partners. (USGS 2005) State initiatives to consolidate permit related data has also resulted in new sources of data.

These integrated databases are examples of the most common type of water quality database, known as a data warehouse, which consists solely of measurements taken from a variety of sources and made available in a common searchable format. The advantage of this type of database for tracking is the ability to access large amounts of data through a single search mechanism. However, the integration of large amounts of data can result in a loss of understanding regarding the intent, purpose, or method of the original sampling event. This can make it very difficult to assess the relevance of the data for use in tracking or surveillance. In addition, little information is provided on the degree to which the data represent the water quality of the geographical area it was taken from. This makes it difficult to assess the comprehensiveness of the data.

These data warehouses use different methods of integrating data from disparate sources and provide differing levels of detail on the source and quality of the data presented. For example, the Environmental Protection Agency's (EPA) main water quality databases known as STORET (short for STOrage and RETrieval), initially required that states use the STORET database platform or else submit output from state databases in text files. Recent technological advancements have enabled states to link or map to STORET while maintaining an individual technological environment. (Cude C, personal communication) The search for technology which would allow for the exchange of data without the need for a shared computing environment led to the development of the NW Water Quality Exchange. Unlike STORET, the data providers in the NW Water Quality Exchange maintain control over their

own data but make it available for use in the regional setting through National Environmental Information Exchange Network Nodes.

The concentrations of different parameters collected in data warehouses alone do not provide information on the potential for adverse health impacts. The measured concentrations must be interpreted through comparison to established standards. The results of these assessments have also been consolidated into searchable databases, where the output is a qualitative description of the water unit. The advantage of these types of databases is that by definition they are designed to integrate and collect data from a variety of different sources and they could be used to highlight areas of greater concern for more detailed investigation. However, the significance of the description of the water unit is completely reliant on the design parameters, definitions, and standards that govern the assessment.

An example of this is the Water Quality Assessment Database developed by the EPA to consolidate the different types of water quality assessment reporting required by the Clean Water Act. The database includes a description, on a watershed or by water-body basis, of the uses assessed, attainment status for each use, and likely source of impairment. While the list of water segments that are impaired due to failure to attain the drinking water use could signal potential areas of concern, there are significant barriers to being able to use this data for tracking. Principally, the assessments included are designed as a snapshot of the quality of the state's water in time. (US EPA 2005) It would be difficult to track attainment status over time because different areas are assessed. Additionally, the lack of consistency in the definitions of designated uses between states and over time, would limit regional collaborative efforts. The source water assessment program (SWAP) of the EPA similarly integrates water quality, hydrological, and spatial information into a determination of the vulnerability and susceptibility of a drinking water resource. The scope of these assessments vary according to the size of the water system and between state programs making integration of the assessment difficult. To create a drinking water system that was able to track water quality over time and space, the data behind these assessment databases is more useful than the qualitative determination of attainment or vulnerability.

The other distinction between sources of water quality data that merits consideration is the differences between data collected to satisfy regulatory requirements versus as part of research. The Clean Water Act, the Safe Drinking Water Act, and multiple state regulations

contain provisions for the monitoring and reporting of water quality data. However, non-regulatory entities such as the United States Geological Survey (USGS) and research institutions such as universities, can represent resources for substantial water quality data. The USGS conducts many different types of water quality investigations and surveillance around the country. As a research agency the water quality information collected is not directly linked to regulatory requirements, parameters, or standards and reflects both ongoing surveillance and snapshots of water quality at a specific time and location. Both continuous and discrete water quality data is stored in a data warehouse type of database known as the USGS National Water Information System (NWIS) and available to the public through a web interface, NWISWeb. (USGS 2005) Since the USGS is not as concerned with the attainment of water quality standards, it is probably the most comprehensive source for low level monitoring data available for areas throughout the country. The consolidation of the data collected from around the country into the single database (NWIS) provides some site location information but does make it difficult to access descriptions of the reason for sampling. Any tracking initiative aimed at integrating sources of water quality information would benefit from the development of partnerships with the USGS in order to gain meaningful access to the large amount of information available.

Conclusions

The perspectives and experience shared by informants provide compelling evidence that there is considerable interest and promise in the development of a tracking network for drinking water. The application of the tracking conceptual model and data integration methods can provide valuable information on the health threats posed by drinking water resources and how they can be prevented. Additionally, the initiation and execution of tracking projects has been shown to enhance the capacity of health and environment agencies through increased data management capabilities, collaboration across agencies and specialty areas, and study depth achieved.

Beyond the value of developing a tracking network for drinking water quality, our review of state tracking projects suggests that the consideration of the scope and limitations of a study approach be a critical component of tracking method development. While all the approaches we considered shared some limitations, each one also contributed unique insights and subsequent challenges to the understanding of the health hazards of drinking water exposures. The diversity of data sources available for inclusion in a tracking network provides the opportunity for the development and evaluation of multiple approaches and methods. Limiting tracking projects to a specific approach, such as ecological linking studies between concentrations in drinking water and health effects, fails to take advantage of the full spectrum of capabilities a tracking network offers.

Review of the multiple sources of water quality data reveals considerable promise in expanding the breadth of a tracking network to include ambient ground and surface water quality data. Inclusion of this data could provide an upstream analysis of drinking water quality that might include the spatial distribution of contaminants of concern or surveillance of important geographical areas such as source water areas. A tracking network which included both ambient environmental data and drinking water quality data, could potentially providing an evaluation of whether ambient monitoring could provide advanced warning of threats before realized in drinking water provided to consumers.

In addition to the evaluation of study approaches and data sources, another key component to tracking network development raised by informants and evidenced by our study of water quality databases is the development of methods for data sharing and integration. The

approach to data sharing piloted by the NW Water Quality Exchange provides a powerful tool for integrating water quality data from many sources. By retaining data ownership and control with the data providers, this method enables a sustainable and comprehensive resource that would be available for use by many different projects. The Exchange also provides the most substantial description of the data sources, which is crucial for evaluating the quality and relevance of the data sources for different types of projects.

The mechanics of developing a tracking network, some of which we touched on in this paper along with others addressed through the many facets of the Environmental Public Health Tracking Project, are numerous and complex. However, an analysis of methods and approaches should also include the discussion of the degree to which they are capable of providing the most insight into the question of concern. The unequal distribution of data resources both within and between states cited both by informants and evident in the scope of water quality and health effect monitoring programs, may limit the capacity of a tracking network to identify the people who may be at the highest risk. An assessment of the degree to which a tracking project is able to identify the people who may be at the highest risk is a necessary component of method development.

The process of interviewing informants and participating in the meetings of western states tracking grantees revealed the enormous potential for advancement, both in the development of tracking methods and in the results of individual projects, gained by increasing collaboration between and within states. Fostering collaborations that include the sharing of expertise, ideas, and resources can strengthen state programs and the development of a national environmental public health tracking network. Hopefully, our synthesis of the experiences and ideas of state tracking projects can act as a stimulus for further discussions and collaboration into the development of a tracking network for drinking water quality.

References

CA DHS-EHIB. 2002. California Environmental Health Indicators. California Department of Health Services Environmental Health Investigations Branch.

CDC-NCEH. 2003. Environmental Public Health Indicators. Centers for Disease Control and Prevention National Center for Environmental Health, Division of Environmental Hazards and Health Effects.

McGeehin MA, Qualters JR, Niskar, AS. 2004. National Environmental public health tracking program: Bridging the information gap. Environmental Health Perspectives. 112: 1409-1413. <http://www.ehponline.org/docs/2004/7144/abstract.html>.

MTECH (Montana Tech of the University of Montana) 2005. Ground-Water Information Center. Montana Bureau of Mines and Geology Available: <http://mbmeggwic.mtech.edu/> [accessed September 30, 2005]

NMTR (New Mexico Tumor Registry) 2002. 27 Years of Cancer Experience in NM Available: <http://hsc.unm.edu/epicpro/27%20Years%20of%20Cancer%20Experience.html> [accessed September 30, 2005]

Pew Environmental Health Commission 2000 America's Environmental Health Gap: Why the Country Needs a Nationwide Health Tracking Network September 2000.

U.S. EPA. 1999. Understanding the Safe Drinking Water Act. EPA 810-F-99-008. Washington, DC: US. Environmental Protection Agency

U.S. EPA (U.S. Environmental Protection Agency) 2005. The 2002 National Assessment Database Factsheet. Available: (http://www.epa.gov/waters/305b/NAD_factsheet.html) [accessed September 30, 2005]

USGS (United States Geological Survey) 2005. GAMA: Ground-Water Ambient Monitoring and Assessment. Available: (<http://ca.water.usgs.gov/gama/WhatIsGAMA.htm>) [accessed September 30, 2005]

USGS (United States Geological Survey) 2005. Water-Quality Data for the Nation. Available: (<http://waterdata.usgs.gov/nwis/qw>) [accessed September 30, 2005]

Appendix A

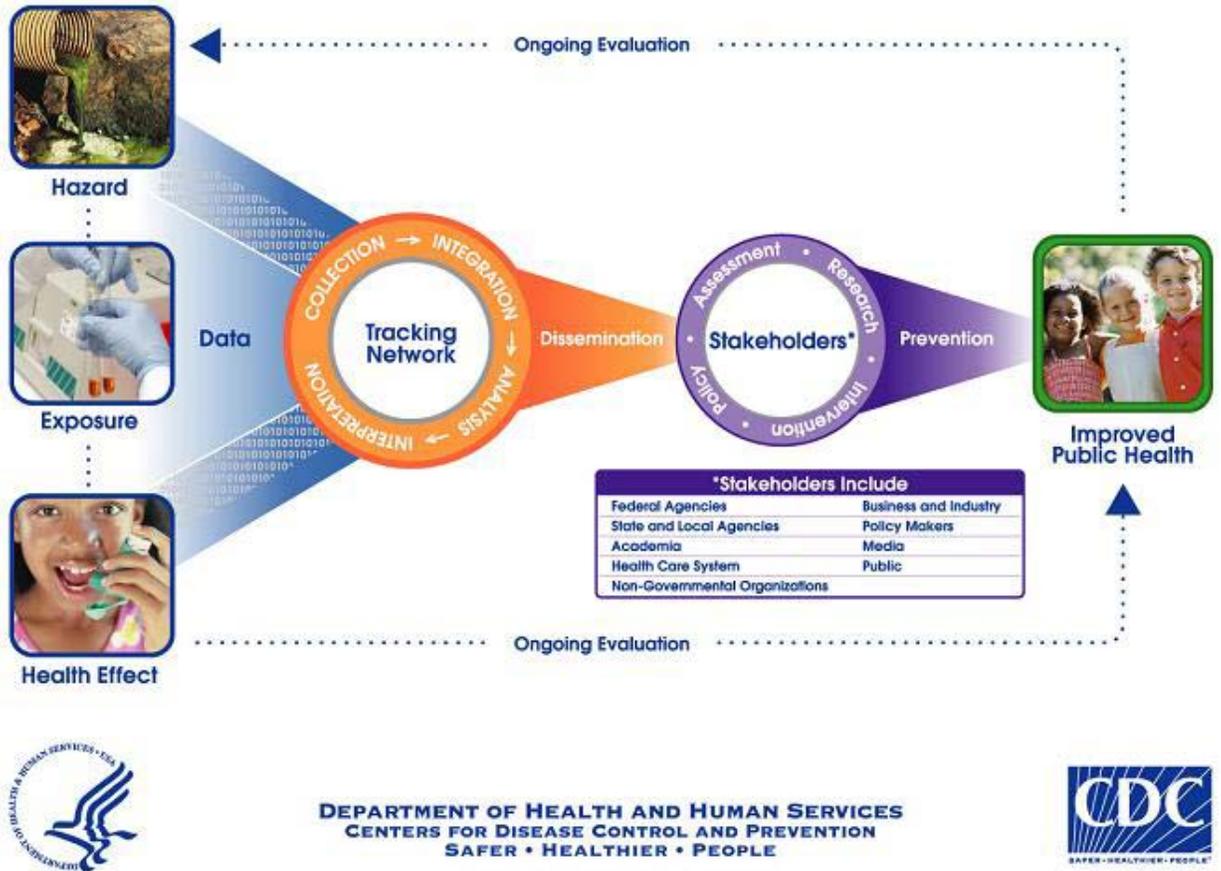
Informants

Barck, Lorena	Program Manager Environmental Public Health Tracking Environmental and Occupational Epidemiology Oregon DHS/Health Services	OR
Brundage, Thomas	Research Analyst Environmental Public Health Tracking Environmental and Occupational Epidemiology Oregon DHS/Health Services	OR
Cude, Curtis	WQ Data Exchange Specialist Business Systems Development Oregon Dept. of Environmental Quality	OR
Flowers, Len	Chief, Environmental Health Epidemiology Bureau New Mexico Department of Health	NM
Fox-Williams, Kathleen	Environmental Scientist Nevada Division of Environmental Protection	NV
Garland, Rodney	Research Analyst Environmental Public Health Tracking Environmental and Occupational Epidemiology Oregon DHS/Health Services	OR
Layco, Gambrelli	Manager, Environmental Public Health Tracking Program Utah Department of Health Environmental Epidemiology Program	UT
Macler, Bruce	Toxicologist Drinking Water Program US EPA Region 9	US EPA
Magraw, Marjean	Coordinator, Environmental Public Health Tracking Project MT DPHHS	MT
Rezai, Susan	Public Health Advisor Environmental Health Tracking Branch, DEHHE National Center for Environmental Health Centers for Disease Control & Prevention	CDC

Stoker, Yvonne	Hydrologist U.S. Geological Survey Office of Water Quality	USGS
Thompson, Catherine	Epidemiologist Environmental Public Health Tracking Environmental and Occupational Epidemiology Oregon DHS/Health Services	OR
VanDerslice, James	Senior Epidemiologist Office of Environmental Health Assessments Washington State Department of Health	WA
Wolff, Craig	IT/GIS Manager – Environmental Health Tracking California Department of Health Services	CA

Figure 1

ENVIRONMENTAL PUBLIC HEALTH TRACKING



Source: McGeehin et al. 2004

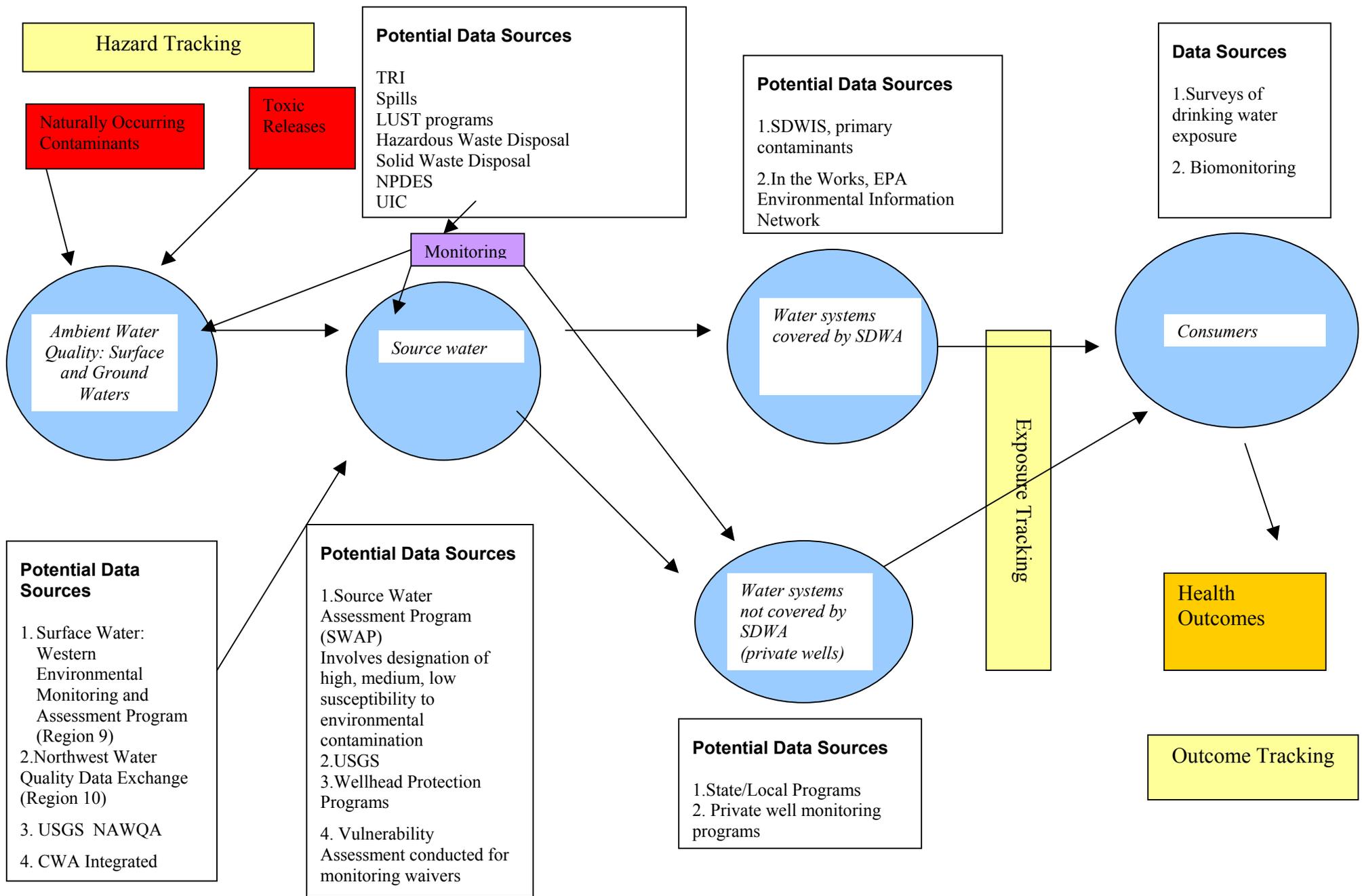


Figure 2. Tracking framework and data sources.