The Business Case for Collecting Address Information in Hospital Discharge Databases

September 2009

THE NATIONAL ASSOCIATION OF HEALTH DATA ORGANIZATIONS

NAHDO

NATIONAL ASSOCIATION OF HEALTH DATA ORGANIZATIONS
Improving Health Care Data Collection and Use Since 1986

The Business Case for Collecting Address Information in Hospital Discharge Databases

Executive Summary

Based on a membership survey, the National Association of Health Data Organizations (NAHDO) found that only 19 of the 45 states had full address; the rest had zip code information as the address component in their hospital discharge databases. Based on the survey, it appeared that another four were either considering or had implementation dates for the addition of street address. NAHDO interviewed five state organizations (WA, FL, NJ, NH, WI) to assess issues, considerations, and challenges states face in collecting address information and to assess what uses were made of address information in their state.

The Environmental Public Health Tracking (EPHT) Program is interested in the addition of full address in order to:

- Determine disease prevalence and incidence rates for specific small geographic areas where environmental toxins were released or are present in the environment.
- Connect environmental releases to specific health conditions; for example, monitoring the impact from environmental toxins.
- Overlay environmental releases or toxins with patient residences--which can be a powerful tool to assess impact and provide policy makers with strong visual evidence of the impact.
- Assess the impact on healthcare costs following intervention, e.g., following clean-up of toxic sites.
- Assess the impact of policy changes, such as those related to second-hand smoke reduction—which can be monitored by measuring related health problems in communities adopting smoke-free zones.

448 East 400 South, Suite 301 • Salt Lake City, Utah 84111
Telephone: 801-532-2299 • Fax: 801-532-2228 • E-mail: nahdoinfo@nahdo.org • Web: www.nahdo.org

This document notes the benefits to all stakeholders: it is critical to document the importance of complete address for other stakeholders in order to achieve buy-in to the expansion of the data collection. These benefits include:

- Providing more precise geo-codes—and will reduce 10% of the linkage errors when hospital discharge data is linked with death certificate data.
- Providing greater opportunities for census tract analysis. This could include mapping motor vehicle crash outcomes with hospitalization information.
- For hospitals using discharge data to conduct market share analysis, the addition of street address would allow for focused marketing efforts to specific neighborhoods or census tracts.
- More detailed understanding for the community of how well the overall delivery system is doing in terms of providing appropriate access to care, especially for indigent care.
- Providing greater opportunity to improve data quality in linked databases, such as the cancer registries, which may have not picked up all cancer cases using only zip code for matching.

The barriers/challengers/considerations to adding street address are also identified in this document:

- Legal challenges can occur when states have privacy laws beyond HIPAA which may preclude the collection or release of Personal Health Information.
- Political challenges are related to "opening" the dialogue about the discharge data collection when requesting additional variables—this can be treacherous to the data collection as a whole.
- Security considerations are of utmost importance and the data organization must be able to quell any fears about loss of privacy.
- Cost considerations include the additional efforts to edit and clean the additional variables—which can be burdensome to both the collector and the submitter.

The document also provides Talking Points for use by those advocating for expanded address information.

NATIONAL ASSOCIATION OF HEALTH DATA ORGANIZATIONS

Background

The National Association of Health Data Organizations (NAHDO) is working with the Centers for Disease Control Environmental Public Health Tracking (EPHT) Program partners to improve hospital discharge data by capturing essential data. Specifically, there is interest in examining the impact of adding more address information to hospital discharge data collections. This work is being funded under a Cooperative Agreement with the CDC Tracking Program, *Cooperative Agreements to Develop or Improve Facets of Public Health Information;* through this Agreement efforts are being made to utilize hospital discharge data to extend the environmental health information resources without implementation of additional registries. In order to have more state health data organizations expand their collection of address beyond zip code, it was deemed important to create a business case to convince state health data organizations that additional address information would add value to their existing data elements and benefit the EPHT Program as well as their data stakeholders.

The National Association of Health Data Organizations is a membership organization for health data organizations, including states, hospital associations, and other profit and non-profit organizations engaged in data collection, analysis, and reporting. Most health data organizations collect the full census of hospital discharges representing all-payers in the system. Hospital inpatient discharge data[1] are useful for understanding utilization of hospital care, charges for inpatient stays, the safety and quality of care, and prevalence of disease in the population. Other uses include studying procedural interventions, examining disparities in care and variation in care delivery, and a host of other research and public health uses.

Discharge data systems are most often based upon abstracts of information from standard billing forms (UB-04)—which makes data collection feasible and less costly. Some health data organizations have historically collected address information along with patient identifiers;

[1] Inpatient care is the provision of surgical and non-surgical health care services to individuals admitted to non-Federal acute care hospitals. Records are collected by hospitalization, not by individual, and are represented at the discharge level rather than as aggregated statistics. Inpatient data include all-payer data (including self and uninsured) for all patients admitted to an acute care hospital in the state for a fiscal or calendar year period, or by periods that can be collected into an annual database. Inpatient data generally contain a complete collection of demographic, clinical, and billing data.

others have not collected patient identifiers except for zip code of patient residence. Health data organizations may vary in their capacity to collect address due to statutory restrictions or other reasons.

Study Methods

NAHDO reviewed and updated information available on statewide hospital discharge systems' data elements, including address information. For those data collections where information on address was not available, contact was made with the health data organization to ascertain availability of address data elements.

NAHDO arranged interviews with state health data organizations to determine how address was used, in order to understand and be able to document the business case for address variables. Interviews were conducted with five organizations based on a standardized questionnaire developed by NAHDO consultant Dr. Barbara Rudolph, with input from NAHDO staff. State health data organizations in the states of Florida, Washington, Wisconsin, New Jersey, and New Hampshire were included in the interviews. See Appendix A for a copy of the questions posed to interviewees.

Interviews were conducted with staff members from data collection organizations. These individuals were the knowledgeable parties on what data elements were collected, the political environment impacting on data collection, and the history and policies of the collection agency. They were also knowledgeable on customer uses of the databases, given the need for data use agreements for access to databases and to identifying variables, such as address variables.

In addition, a review of published and gray literature was conducted to ascertain examples of how address information played a role in either research or projects related to hospital care.

The combination of these methods was used to develop this briefing paper on the business case for address variables.

NATIONAL ASSOCIATION OF
HEALTH DATA ORGANIZATIONS

Why Include Address Variables in Hospital Discharge Data Systems?

The state grantees of the CDC EPHT Program have indicated they would like to supplement their existing data systems with hospitalization information from hospital discharge data systems. Additional registry information is becoming burdensome for healthcare providers, who are under pressure to produce information for a variety of payers, and other stakeholders. Resistance to stand-alone public health registries has grown in both state legislatures and provider communities as the needs for more health data continue to expand.

The Environmental Public Health Tracking Program has been working with NAHDO to reduce some of the barriers for using hospital discharge data as a major component of their systems. Incomplete address information is one of the barriers; the EPHT Program would like to use the hospital discharge data in conjunction with other surveillance databases, and to do so may require more detailed address variables. Environmental surveillance information can help investigators identify populations at risk and respond to outbreaks, clusters, and emerging threats. When the population has been identified specific intervention and prevention strategies can be deployed. On the policy side, discharge data added to other surveillance data can be used to estimate the healthcare costs of exposure to environmental toxins and poor air quality, and can also address the impact of interventions and prevention programs. Discharge data with detailed address information would add to the capacity of the EPHT Program.

Address Data Elements Collected by State Health Data Organizations

NAHDO collects information on data elements found in hospital discharge data to assist others (researchers, government officials, consultants) in selecting which databases might meet their needs. States generally collect data elements available on the Uniform Bill 2004 (UB-04), although some states have gone to the 837 Institutional Electronic Claim. When updating the Uniform Bill form, the National Uniform Billing Committee (NUBC) made a number of improvements, including aligning the UB elements with the electronic claim elements. In Table 1 below, Patient Address in the UB-04 contains street, city, state, zip code, and country code. In the earlier versions of the Uniform Bill there was only one field for address; now address

elements are discrete to match the electronic claim. When new data elements are being considered by states, the Uniform Bill is often the source from which states select data elements.

Table 1: UB-04 Patient Address Codes

Form Locators	Description	Line	Type	Size	
FL 09	Patient Address-Street	1a	AN	40	Discreet
FL 09	Patient Address -City	2b	AN	30	Discrete
FL 09	Patient Address - State	2c	AN	2	Discrete
FLF 09	Patient Address - ZIP	2d	AN	9	Discrete
FL 09	Patient Address – Country Code	2e	AN	3	Discrete

This paper responds to interest in determining which state health data organizations had address information as part of the data that is collected. NAHDO updated their information to provide details on this. There are nineteen states with confirmed street address information; two states (New Hampshire and Virginia) will be collecting full address when new administrative rules take effect; and NAHDO could not acquire information about address data elements for two states (Illinois and Delaware) (see Appendix B for listing of states). In addition, the state of Washington will likely include detailed address information in their Emergency Department (ED) Data Pilot; if this is successful, expanded address information in the inpatient data may follow.

Challenges Associated with Address Information

While seemingly a simple variable, address information can potentially result in political battles within states between those desiring greater geographic detail and those advocating for increased privacy of health information, or for less reporting burden. Hospital discharge data systems, while present in 48 states, do have very idiosyncratic histories, missions, and stakeholders. Some hospital discharge data systems have been collected by state agencies (insurance, budget control, public health, healthcare financing); others operate under the state mandate but are independent organizations; still others are hospital associations without state mandate. Each of these organizational circumstances are faced with challenges when implementing hospital

discharge data collections. These challenges can impact upon which variables are collected—especially those that can be used to identify individual patients—such as address. In New Hampshire, waves of privacy issues occurred in the legislature during the time the state was requesting address variables. The legislature made statutory changes to use of Social Security Number (SSN) in response to privacy advocacy efforts. In Washington State, the addition of the ED data collection to the hospital discharge created an opportunity to pilot test collecting more detailed address information. If the new ED data collection goes well, it will be easier to add complete address to the inpatient data.

Some organizations historically decided to collect and use zip code information or county level information only in their state hospital discharge data collections (Wisconsin and Florida are examples). Health data organizations functioning within these limited address variables can still produce valuable healthcare information, but there are limits to the type of geographic analysis that is possible when only either zip code or county is collected. Small-area analysis with either county or zip code as the reference can create results where differences within a zip code or county are masked. Extremely high income and low income persons may inhabit the same zip code or county, and this may lead to incorrect assumptions regarding average income levels under analysis. For example, if attempting to determine where either free health care or public health services are located, a higher average income might mask the need for low cost services. Zip codes can also change in non-census years, making trending of data and linking with other time periods more difficult.

Generally, those states with full address information (street address, city, state, zip code), can extend the utility of the database to include more accurate geographic surveillance of conditions, procedures, and events—providing more detailed population health information within small geographic areas., The additional information also can assist by improving the reliability of linkages across data systems. However, additional detail in databases can also lead to errors when address is incorrectly supplied. For example, hospitals often collect “mailing address” for purposes of billing, rather than necessarily home address (from interview with New Hampshire). These errors can generally be addressed through software programs designed to clean address

variables or by hospitals retrieving patient home address—but this may add time or complexity to the processing of data. Some hospitals may resist editing protocols when errors are frequent and not necessarily important for their purposes; this happened in New Hampshire when the topic arose during discussions with hospitals on revised administrative rules. These arguments reflect the inevitable tension between data users and data providers over the content of data sets.[2]

Added Value of Complete Address Information to Environmental Tracking

The value of adding complete address information to the Environmental Tracking program is significant. The complete address can be used in the following components of Tracking:

1. Determination of disease prevalence and incidence rates for specific small geographic areas where environmental toxins were released or are present in the environment. For example, the Maine EPHT program tracked the presence of ambient ozone and asthma-related emergency department visits. They utilized zip code because more complete address was not known. They were able to estimate the change in ED visits when ozone levels were high.

2. Connecting environmental releases to specific health conditions, for example, monitoring the impact from environmental toxins. In Wisconsin, the EPHT worked with the US Environmental Protection Agency Region 6 to analyze local factory emissions of trichloroethylene (a solvent associated with liver and lung damage), and estimate community cancer risks. As a result, the manufacturer agreed to change manufacturing processes, eliminating the release of this solvent, even though the release was within the legal limit.[3] With complete address information you could also analyze hospital discharge data in the area where release occurred to assess the actual hospitalization costs of the release of this solvent.

[2] Iezzoni, L.I. (1997) *Assessing Quality Using Administrative Data*. Annals of Internal Medicine, Part 2, 15 October, 127:666-674.
[3] Charleston, A., Bannerjee, A., and Carande-Kulis, V.G. (2008) *Measuring Success: The Case for Calculating the Return on Investment of Environmental Public Health Tracking*. Journal of Public Health Management and Practice, Vol.14, No.6: 600-604.

3. Many different types of studies today utilize geo-coding as a critical component for providing a visual tool for determining patterns—e.g., overlaying environmental releases or toxins with patient residences can be a powerful tool to assess impact and provide policy makers with visual evidence of the impact. For example, geo-coding can be used to map cancer clusters along with known toxins, to assess impact of releases.

4. Assessment of impact on healthcare costs following intervention, e.g., clean-up of toxic sites. Address information assists in assessing reductions in healthcare conditions or utilization of hospital services. For example, New York City collects street address for the rat portal—this type of information could be linked to discharge information to help understand the impact of rat clean-up programs.

5. The impact of policy changes related to second-hand smoke reduction can be monitored by measuring related health problems in communities adopting smoke-free zones.

These various examples of how complete address within hospital discharge data could be used in conjunction with other environmental information from the EPHT Program reflect the importance of adding these critical data elements to discharge data for improving the nation's health.

Value Equation for Complete Address on Other Stakeholders

It is also critical to document the importance of complete address for other stakeholders with differing interests from the EPHT Program. These other stakeholders must also benefit for a state health data organization to take on the effort to change statutory requirements. Having local support may be critical to overcome privacy concerns that often are raised when changes to data elements bring the dataset into the public eye. Hospitals are less likely to resist the additional burden if their own needs can also be met by adding the data elements.

The examples in this section come from three sources: interviews conducted with state health data organizations; print literature; and the gray literature.

- In New Hampshire, the data agency is looking forward to implementing street address and expects to use this information in a variety of ways, including more detailed geocoding. They anticipate that the new address information will provide more precise geo-codes—and will reduce the 10% linkage error rate when hospital discharge data is linked with death certificate data. The new address variables will become available with discharges beginning January 1, 2010.
- In Washington State, the health data agency links both birth and death certificate data with the hospital discharge data. They would like to be able to get census tract level information, which they cannot do with just zip code, but could do with complete address information. They would also like to be able to link with the trauma registry and traffic crash outcomes. Again, more complete address provides opportunities for census tract analysis.
- In New Hampshire (and other states), the Department of Health EPHT recently completed a study on Heart Attacks Tied to the Environment, where they examined trends in heart attack hospitalizations over time, and evaluated geographic differences. This was part of a national effort by the EPHT Program.
- Hospitals use discharge data to conduct market share analysis—this includes a variety of areas that can be examined. For instance, hospitals could asses the need for added service lines or where to site hospital affiliated clinics or specialty services. They could also look at existing patient demographic information to analyze which segments of the market are being underserved. Analyzing where they receive transfers from and those areas that don't transfer patients to them would allow for focused marketing efforts to specific neighborhoods or census tracts.

- States with address variables can easily map patient location for "ambulatory care sensitive conditions" along with geographic location of primary care physicians. This

could assess how well the overall delivery system is doing in terms of providing appropriate access to care. One could also calculate the potential savings if the population were cared for in the community as opposed to in the ED or on medical floors.

- Detailed address also becomes important when examining potential cancer clusters or other health conditions. This information can be mapped with exposures from industry or contaminants in the environment. Clean-up sites may be determined from this type of information.

- Florida currently uses zip codes and county codes in mapping information from their hospital discharge data on utilization of services, health conditions by county, etc. If more detailed address information were available, more detailed maps could be done by census tract or voting districts.

- Improved data reliability and validity can be garnered through data linkages with full address information. For example, cancer registries benefit by assessing whether they received all appropriate cases from hospitals in their registry by comparing with the hospital discharge data.

As indicated, there are many different uses of address information; states relying on zip code or county only are missing opportunities for more detailed analysis of health conditions, healthcare utilization and services, uncompensated care issues, and environmental impact on health. They are also missing opportunities to improve: their data linkages; the reliability and validity of their registry data; measurement of response to policy initiatives; and evaluation of prevention and intervention programs. While potential costs or threats are associated with revamping data collections, there are also costs associated with not providing the additional detail to healthcare data customers. Users of the data are unable to get at the detail needed for today's information environment, and as a consequence support for the existing data collections may wane.

NATIONAL ASSOCIATION OF
HEALTH DATA ORGANIZATIONS

Summary

State health data organizations now have a federal partner interested in expanding the address variable information in hospital discharge databases. The business case for this, however, extends beyond the EPHT Program; it includes the data providers and other users of the data. The capacity for using detailed address continues to grow—as Geographic Information Systems (GIS) become more sophisticated. Web-enabled GIS can rapidly extend the value of the hospital discharge database with detailed address information. Customers can examine links between healthcare utilization and the environment or measure effectiveness of prevention and intervention programs, as well as policy initiatives. Other data stewards can validate their data through connections to the hospital discharge data. State health data agencies have been good stewards of patient information—collectively processing millions of patient records annually without releasing information that could identify individual patients. We cannot afford to continue to limit address information to zip code or county.

NATIONAL ASSOCIATION OF HEALTH DATA ORGANIZATIONS

Improving Health Care Data Collection and Use Since 1986

Talking Points for State Initiatives to Attain Address Variables

The National Association of Health Data Organizations and the CDC Environmental Public Health Tracking Program both recommend the addition of full address (street, city, state, zip code, county code) to state hospital discharge data. This addition would assist states, federal partners, hospitals, and other data users in more detailed examinations of hospital utilization, health conditions, preventable hospitalizations, tracking of environmental exposures, and targeting public health prevention efforts.

Examples of the ways in which this addition to current data collections would provide added value are:

- Determination and mapping of resulting diseases from environmental exposures can assist in surveillance activities at both the state and federal levels.

- Estimates of the healthcare costs associated with exposure to environmental toxins and poor air quality can be identified in detail and mapped. Mapping of this type of information is an effective tool for communicating to the public and policy makers.

- Geo-coding at the more detailed level of street address can provide new insight into preventable hospitalizations related to specific conditions such as asthma. Zip code only analysis can do this, but at a cruder level than if done at street level. Using zip code as the geographic unit of analysis can mask variation in race, income and education levels, and insurance coverage—all of which can impact on preventable hospitalizations.

- More complete address can assist in monitoring the impact of public health interventions on hospital related healthcare costs, such as the establishment of public clinics to address diabetes within neighborhood settings.

- Complete address information can be used by data experts when linking other databases to the hospital discharge data—address can be used to weed out improper linkages. Many state agencies link the following databases with hospital discharge: motor vehicle traffic accident information, birth files, death files, Medicaid files, and cancer registry files. Because some files may not contain patient name for privacy reasons, probabilistic matching is used and that requires verification when match probability is low.

- More detail on health conditions found in specific voting districts is useful to state policy makers when assessing the need for additional services, such as increased primary care or school-related clinics to prevent excess hospitalization.

- States, with Federal funding, have used address as an important tool to examine whether there were geographic differences in heart attack rates.

448 East 400 South, Suite 301 • Salt Lake City, Utah 84111
Telephone: 801-532-2299 • Fax: 801-532-2228 • E-mail: nahdoinfo@nahdo.org • Web: www.nahdo.org

- Hospitals also use discharge databases to examine their place in the market and to determine whether new service lines should be added. Outpatient clinics may be sited according to patient demographics in certain neighborhoods or areas.

While some concerns arise regarding privacy of patient information (like address), those concerns can be addressed through data management approaches currently in place in health data organizations.

Adding this additional information to the discharge data is an effective way to better meet the needs of government, the public, and the hospital community.

Appendix A

Interview Questions Regarding Address Information

For all states

- **What components of address do you collect?** (Street, city, county, zipcode, geocode?)

- **How long have you been collecting these elements?**

- **Does the address information make your data more valuable? In what way?**

- **Are there any studies/reports you couldn't do if you did not have address information?**

- **Generally, what components of address are required for most studies? Are you aware of how CDC tracks health conditions related to the environment using address? (asthma, lead, water quality).**

- **Do you use address variables for linkage of hospital discharge to other databases? (e.g., vital records, cancer registries, asthma registries?)**

- **Do you geocode the data for tracking? Or does CDC Tracking geocode the information you provide? If so, do you know what kinds of issues they have looked at? (asthma? Lead? COPD? Lung Cancer?)**

- **How difficult is it for Public Health entities to access your address information?**

    - Not able to access?
    - Time-consuming process of negotiation?
    - Sign data use and quickly access data?

- **If Public Health entities can access your data with the address variable—what kinds of studies are they doing?**

- **Any other comments on using address?**

Barbara Rudolph (May 15, 2009)

Appendix B:
States Collectng Full Street Address and/or Zip Code Data

State	Health Data Organization	Street Address	Zip
Alabama			
Alaska			
Arizona	Arizona Dept of Health Services	yes	yes
Arkansas	Arkansas Department of Health and Human Services	yes	yes
California*	Office of Statewide Health Planning and Development		yes
Colorado	Colorado Hospital Association		yes
Connecticut*	CT Office of Health Care Access		yes
Delaware	Delaware	UNK	yes
District of Columbia	DC Hospital association		
Florida*	FL ACHA		yes
Georgia	Georgia Hospital Association	yes	yes
Hawaii	Hawaii Health Information Corporation	yes	yes
Idaho			
Illinois	Illinois Dept of Public Health	UNK	yes
Indianna	Indiana Hospital&Health Association		yes
Iowa	Iowa Hospital Association		yes
Kansas	Kansas Hospital Association		yes
Kentucky	Kentucky Cabinet for Health and Family Services		yes
Louisana	Louisiana Dept of Health & Hospitals	yes	yes
Maine*	Maine Health Data Organization		yes
Maryland*	Maryland Health Cost Review Commission		yes
Massachusetts*	Massachusetts Division of Health Care Finance & Policy	yes	yes
Michigan	Michigan Health & Hospital Association		yes
Minnesota	Minnesota Hospital Association		yes
Mississippi			yes
Missouri*	Hospital Industry Data Institute	yes	
Montana			yes
Nebraska	Nebraska Hospital Association	yes	
Nevada	Center for Health Information Analysis	yes	yes
New Hampshire*	NH Dept Health and Human Services		yes
New Jersey*	NJ Department of Health and Senior Services	yes	yes
New Mexico*	NM Health Policy Commission	yes	yes
New York*	New York State Department of Health	yes	yes
North Carolina	North Carolina/UNC		yes
North Dakota	North Dakota department of health		yes
Ohio	Ohio Hospital Association		
Oklahoma	Oklahoma State Department of Health	yes	yes
Oregon*	Office for Oregon Health Policy and Research		yes
Pennsylvania*	PHC4	yes	yes
Rhode Island	Rhode Island Dept of Health		yes
South Carolina	South Carolina Office of Research and Statistics*	yes	yes
South Dakota	South Dakota		yes
Tennessee	Tennessee	yes	yes
Texas	Texas Dept. of State Health Services	yes	yes
Utah*	Office of Health Care Statistics		yes
Vermont	VT BISHCA		yes
Vermont	Vermont Health and Hosp Assn		yes
Virgina	Virginia Health Information+	Eff. 7/9/2009	yes
Washington*	Washington State Dept of Health	yes	yes
West Virginia	West Virginia Health Care Authority	yes	yes
Wisconsin*	DHFS/DPH/Bureau of Health Information and Policy		yes
Wyoming			yes
Totals		19/20	45

* Indicates Tracking State Grantee