

Emerging Infections Program (0920-0978)
Revision
February 2015

SUPPORTING STATEMENT PART B: Statistical Methods

Amy McMillen, MPH
Centers for Disease Control and Prevention
National Center for Emerging and Zoonotic Infectious Diseases
Office of the Director
1600 Clifton Rd
Atlanta GA 30333
404-639-1045
auh1@cdc.gov

TABLE OF CONTENTS

B. Statistical Methods

1. Respondent Universe and Sampling Methods
2. Procedures for the Collection of Information
3. Methods to Maximize Response Rates and Deal with Nonresponse
4. Test of Procedures or Methods to be Undertaken
5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Sampling Methods

ABCs conducts population-based surveillance and includes all cases in a defined catchment area. No sample selection is involved in this surveillance study. Therefore, the data collection covers the entire target population. See 2012 Protocol for Active Bacterial Surveillance (Attachment 19) for the populations under surveillance for each pathogen and area as of January 2012. Because ABCs personnel submit the disease surveillance forms as a part of their job to perform a public health service, the response rate is expected to be 100%.

FoodNet conducts active population-based surveillance for nine pathogens and one syndrome among all residents of our catchment area. The population under surveillance is 47 million persons and represents 15% of the U.S. population. We identify approximately 19,000 reports of illness (all pathogens combined) each year. We work with health departments in ten states to collect basic demographic and lab data on all cases but only some cases are interviewed (the number of cases interviewed depends on each state health department).

The All Age Influenza Hospitalization Surveillance (Flu Hosp) project covers a population about 23 million residents who have the potential of being hospitalized with laboratory-confirmed influenza. Between 2003-04 and 2010-11 the number of laboratory-associated influenza hospitalizations has ranged from approximately 1,000 to 6,500 people. Based on these figures and the improved efficiency of the Flu Hosp project surveillance procedures, it is anticipated that there is still no need for a sampling scheme. A standardized case report form is completed on all persons that meet the case definition, however to ease the burden on sites, five readily available variables (site-assigned unique case number, age or date of birth, sex, hospital admission date, and positive influenza test result) are submitted to CDC as soon as possible. Although timely completion of the remainder of the case report is encouraged, sites have until September 30 to complete medical chart reviews and data abstraction.

Ten EIP sites participate in the pediatric and adult influenza hospitalization project and represent 12 metropolitan areas and approximately 7% of the US population. All hospitals that accept adult and/or pediatric admissions in the catchment areas under surveillance are included for active public health surveillance so accurate population-based rates can be calculated.

Age-specific rates of laboratory-confirmed influenza-associated hospitalizations and influenza-associated severe complications are calculated using population denominators from the most recent census data available for pediatric and adult populations. Hospitalization rates are routinely estimated for the following age groups: < 6 months, 6-23 months, 2-4, 5-17, 18-49, 50-64, and ≥ 65 years. Additional group-specific rates are calculated as needed for groups disproportionately affected by influenza-associated hospitalizations.

Interim analyses of aggregate data are conducted to estimate hospitalization rates and monitor factors associated with severe influenza throughout the influenza season. Final analysis includes a season summary of the epidemiologic characteristics of hospitalized cases using standard descriptive statistics. Where appropriate, univariate and multivariate analyses are conducted to evaluate factors associated with serious influenza-associated complications.

All analyses are conducted using SAS. Aggregate results are regularly shared with relevant CDC programs, including the ACIP, and with the public and scientific community via scientific publications.

The HAIC-A conducts population-based surveillance and includes all cases in defined catchment areas in each of the 10 EIP sites. No sampling is involved for resistant Gram-negative bacilli surveillance. Sampling is performed in some instances in CDI surveillance to reduce the burden of data collection, as follows:

Application of CDI case sampling at selected EIP sites: In sites with catchment area populations greater than 1,500,000 persons, because of the relatively high volume of positive *C. difficile* toxin specimens, sites will apply a stratified random sampling scheme, based on age and gender, after determining which positive *C. difficile* toxin specimens qualify as a CDI case for the surveillance. CDI cases will be categorized into strata based on age and gender. There will be a total of 8 strata; two gender groups (i.e., males, females) and four age categories (i.e., 1-17, 18-44, 45-64, >64). All CDI cases in the youngest age group (i.e., 1-17), regardless of gender, will be sampled. For the other 6 age/gender strata, the EIP sites will randomly sample 1:3 of the CDI cases.

Application of sampling of “Healthcare Facility Onset” (HCFO) CDI cases at all EIP sites: Among CDI cases classified as HCFO, 1 in 10 will be randomly selected for case report form completion. Some sites with low numbers of HCFO cases may choose to complete a case report form on all HCFO CDI cases. In contrast, a case report form will be completed on all community-onset CDI cases. From these data, community-onset CDI cases will be classified as either “community-onset, healthcare facility-associated” or putative “community-associated” CDI cases. Interviews are conducted for patients with putative community-associated CDI.

2. Procedures for the Collection of Information

Case finding in ABCs is active and laboratory-based. As positive laboratory reports are essential to the case definition, the microbiology laboratories in acute care hospitals and reference laboratories processing sterile site specimens for residents of the surveillance area are the most efficient sites for case identification. In addition, some of the data of interest on cases of invasive bacterial disease is readily accessible in the microbiology laboratory. However, most data that are essential for describing the population-based epidemiology of these diseases (e.g., age, residence within the surveillance area, outcome) may not be available in many microbiology laboratories. Therefore, a standard case report is completed on all identified cases through medical record review. The standard case report form contains questions on basic demographics, underlying conditions, vaccinations and risk factors for infection. Data collection is done differently in each surveillance area; for example, through the cooperation of on-site hospital personnel (e.g., Infection Control Practitioners or Medical Records personnel), through medical record review or clinician interview by county health department personnel, or through medical record review by surveillance personnel.

To assure complete timely reporting and collection of data, contact with microbiology

laboratories must be frequent. In hospitals without computerized microbiology data, surveillance personnel should call designated microbiology laboratory contacts regularly to identify new cases and request isolate submission. Where microbiology data are computerized, electronic listings of all isolates of the pathogens of interest from normally sterile sites should be obtained on a monthly basis. If enrollment into special studies due to slow reporting falls below 90% or isolate collection falls below 85% of surveillance cases, regular calls to microbiology labs should be instituted to ensure that delayed reporting of cases does not have an adverse effect on enrollment rates into special studies or isolate collection rates.

Each area must determine what means will be used for collection of data that are unavailable in the clinical microbiology laboratory. It is essential that the method(s) selected are detailed in writing and shared with CDC and the other surveillance areas, to permit assessment of the comparability of data collection. In addition, problems with proposed methods for data collection should be identified promptly and new methods substituted and changes documented when appropriate. In addition to formal audits of the surveillance systems, surveillance areas regularly assess the completeness of information collected for each case. If any core variables (e.g. outcome) are frequently incomplete, the data collection method should be revised to correct the problem. CDC should be notified regarding changes in data collection methods as these occur.

In FoodNet, sampling method is not used in our data collection. Rates are calculated from the number of cases divided by the total population (census data). Trends over time are calculated using a negative binomial regression model to account for the change in catchment area (from 5 sites to 10 sites) and the variability in incidence between pathogens and sites. Rates are calculated overall, by pathogen, by state, by age groups, and by species or serotype.

The Flu Hosp project conducts active public health surveillance for laboratory-confirmed influenza hospitalization cases in all age groups within selected catchment areas in 10 states (*See project flowchart Attachment 20*). Sites prospectively identify cases by reviewing hospital laboratory, admissions, infection control practitioner databases/logs, or reportable conditions databases. This involves active contact with hospital laboratories, admissions departments, and infection control practitioners, or review of reportable condition databases. Methods may vary slightly among surveillance areas or among hospitals within an area depending on the availability of laboratory and admissions databases. For hospitals with computerized viral laboratory data, computerized listings of all influenza positive cases in all age groups are obtained on a weekly basis throughout the influenza season. In an effort to minimize burden for hospitals without computerized laboratory data, surveillance personnel contact designated laboratory contacts in each health care facility approximately every two weeks to identify potential new cases. Influenza admissions also may be tracked by infection control professionals or other hospital staff serving hospital wards where influenza cases might be admitted. For hospitals in states where hospitalized influenza cases are a reportable condition, infection control practitioners review laboratory results and admission logs. For all potential cases identified, medical charts are reviewed by state health department appointed surveillance officers to determine whether case definition inclusion criteria are met.

Once there is verification of positive influenza test and confirmation that patient meets the case definition and inclusion criteria, sites conduct medical and laboratory chart review and data

abstraction to collect detailed clinical and epidemiologic information contained in the standardized case report form (Attachment 9). To obtain as complete an influenza vaccine history as possible sites will use the following sources, in order of priority, to collect this information: 1) review the patient's medical chart, 2) consult the state vaccination registry, 3) contact the patient's provider via fax or telephone and/or 4) contact the patient or their proxy. If providers and/or patients or proxies need to be contacted, a standardized interview will be used to obtain influenza vaccination history.

Case finding in HAIC-A population-based surveillance is active and laboratory-based. As positive laboratory reports are essential to the case definitions for both CDI and for resistant Gram-negative bacilli, the microbiology laboratories in acute care hospitals, reference laboratories, and other healthcare facilities (e.g., long term care facilities, dialysis center referral laboratories, etc.) processing specimens for residents of the surveillance areas are the most efficient sites for case identification. In addition, some of the data of interest on cases of *C. difficile* or resistant Gram-negative bacilli are readily accessible from the microbiology laboratory. However, most data that are essential for describing the population-based epidemiology of these infections (e.g., age, residence within the surveillance area, outcome) may not be available in many microbiology laboratories. Therefore, a standard case report is completed on all identified cases through medical record review. In the case of putative community-associated CDI cases, a standard interview may also be conducted to gather additional exposure and risk factor information not typically available in the medical record (such as information on household contacts, food exposures, etc.). The standard case report form includes questions on basic demographics, underlying conditions, and risk factors for infection. Data collection may be performed differently in each surveillance area; for example, through the cooperation of on-site hospital personnel (e.g., Infection Control Practitioners or Medical Records personnel), or through medical record review by EIP site personnel.

To assure complete timely reporting and collection of HAIC-A data on CDI and resistant Gram-negative bacilli, contact with microbiology laboratories must be frequent. EIP sites must demonstrate to CDC project staff a comprehensive understanding of all laboratories within their catchment areas that are performing testing for CDI or for resistant Gram-negative bacilli, to ensure complete case capture. This entails EIP site personnel communicating regularly (e.g., annually) with all healthcare facilities and providers in their catchment areas (e.g. through telephone inquiries, email communications or mailings) to ensure that they know the laboratories serving those facilities and providers (including laboratories such as large regional reference laboratories that may be located outside the catchment area) and the type(s) of microbiological testing for CDI and resistant Gram-negative bacilli performed in those laboratories. In hospitals without computerized microbiology data, surveillance personnel communicate regularly with designated microbiology laboratory contacts to identify new cases and request isolate submission. Where microbiology data are computerized or where queries of laboratory automated testing instruments can be programmed, electronic listings of all isolates of the pathogens of interest identified from the body sites under surveillance (e.g., stool for *C. difficile*) should be obtained on at least a monthly basis. Regular interactions of EIP site personnel with microbiology laboratory staff members ensures that case reporting is complete and timely, and that isolate submission rates to CDC are acceptable.

Each EIP site must determine what means will be used for collection of HAIC-A data that are

unavailable in the clinical microbiology laboratory. All sites used standardized case report forms to collect these data. It is essential that the method(s) selected by the sites are shared with CDC and the other EIP sites to permit assessment of the comparability of data collection. In addition, problems with proposed methods for data collection should be identified promptly and new methods substituted and changes documented when appropriate. In addition to formal audits of the surveillance systems, surveillance areas regularly assess the completeness of information collected for each case. If any core variables (e.g. outcome) are frequently incomplete, the data collection method should be revised to correct the problem. CDC should be notified regarding changes in data collection methods as these occur.

3. Methods to Maximize Response Rates and Deal with No response

The state public health laboratories and partnering academic institutions submit the disease surveillance forms as a part of their job to perform a public health service therefore, the response rate is expected to be 100% for ABCs.

FoodNet calculates performance standards overall and for each site twice a year to gauge progress on data completeness (see Attachment 21). Data elements that are less than 80% complete are not included in analysis. Periodic review of the performance standards is conducted and discussions are held with sites who do not meet performance standards to develop plans for improved performance.

The Flu Hosp project does not have a method to deal with non-response because the state public health laboratories and partnering academic institutions submit the disease surveillance forms as a part of their job to perform a public health service. Therefore, the response rate is expected to be 100%. However, some responses will require the surveillance officer to contact patients to obtain vaccination status information. Based on data collected between 2007-08 and 2010-11, approximately 8.7% of cases had incomplete vaccination history because the patient or proxy could not be interviewed.

Contact information will only be required in some circumstances when the patient's influenza vaccination history is not noted in the medical record, hospital database or state vaccination registry. To obtain as complete an influenza vaccine history as possible sites will use the following sources, in order of priority, to collect this information: 1) review the patient's medical chart, 2) consult the state vaccination registry, 3) contact the patient's provider via fax or telephone and/or 4) contact the patient or their proxy. If providers and/or patients or proxies need to be contacted, a standardized interview will be used to obtain influenza vaccination history. Attempts will be made to contact a patient up to 3 times to obtain this information.

If necessary, participating sites will interview patient or proxy by phone to obtain vaccination history. Sites employ the following methods to try to locate patients' families: 1) medical charts, 2) laboratory records, or 3) directory assistance ("411"), and 4) internet phone/address searches (including name and address/reverse directories). If a proxy is needed, sites try to

identify the family member who is most familiar with the patient's medical history during the phone interview.

Once a correct phone number is identified, sites make multiple attempts to reach the family member. To minimize non-response because of unusual work or life schedules, these attempts include calling during different daytime and early evening periods during the week and weekend. Sites stop trying to call a patient or proxy if they cannot locate a correct phone number after using the search methods listed above or if successful contact is not made after multiple attempts at what appears to be a correct number.

A primary limitation of this activity is that case ascertainment may not be complete. To identify all laboratory-confirmed cases, all laboratories would need to be audited, not just hospital laboratories; however, because the majority of influenza positive cases will not require hospitalization, the workload in determining which of the positive cases required hospitalization would be impractical.

Another limitation of performing surveillance for laboratory-confirmed influenza is that not all patients with influenza will receive influenza diagnostic testing and not all those that are tested will be positive, even if they have influenza, due to the timing of viral shedding and specimen collection. However, because the clinical presentation of influenza is similar to that of many other illnesses, we have limited our case definition to individuals with laboratory-confirmed evidence of influenza.

For the HAIC-A, EIP site epidemiology staff submit the case report forms as a part of their job to perform a public health service, and therefore, the response rate is expected to be 100%. Performance measures that are tracked for the EIP sites include measures related to the completeness and timeliness of case report form completion and isolate submission. For the interviews of community-associated CDI cases, previous data have shown the response rate to be approximately 68%.

4. Tests of Procedures or Methods to be Undertaken

For ABCs and the Flu Hosp project, the data being collected represents standard clinical and demographic information. No tests of procedures or questions were performed.

For FoodNet, except for HUS surveillance, FoodNet does not use a standardized case report form. Each state uses their own state-specific forms from which data elements are extracted and sent to CDC. If FoodNet would like to collect new data elements, these will be reviewed with sites and the feasibility of collecting such data will be evaluated and discussed.

For the HAIC-A, pilot CDI and resistant Gram-negative bacilli surveillance efforts were conducted: CDI surveillance in two EIP sites, and resistant Gram-negative bacilli surveillance in three EIP sites.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

CDC conducts a conference call with site surveillance officers to discuss ABCs-related issues monthly. CDC also organizes two annual meetings: the ABCs Steering Committee meeting with attendance by the ABCs Principle Investigators and one surveillance officer from each site, and the ABCs Surveillance Officers meeting with attendance by at least two surveillance officers from each site. Londell McGlone (jo13@cdc.gov; 404-639-0729) compiles the data that is sent from individual sites on a monthly basis. Biannual reports are produced by Tracy Pondo (dio2@cdc.gov; 404-639-8243) and Melissa Lewis (bmj4@cdc.gov; 404-639-3778) and reviewed by Gayle Langley (fez7@cdc.gov; 404-639-8092). Other members of the ABCs team at CDC or EIP sites can perform additional analyses after proposals are cleared by committees.

For FoodNet, staff at state health departments collects the data and an extract is sent to CDC. Jennifer Huang (uzo0@cdc.gov; 404-639-3955) compiles the data at CDC and produces monthly and yearly reports. Olga Henao (dot8@cdc.gov; 404-639-3393) and Stacy Crim (dex2@cdc.gov; 404-639-2257) are responsible for trend analysis. Any member of the FoodNet team at CDC, sites or federal partners can perform additional analysis.

The following identifies individuals who are consulted for Influenza statistical and data analysis: Shikha Garg (bev8@cdc.gov; 404-639-6142) Influenza Division, National Center for Immunization and Respiratory Diseases (NCIRD), CDC; principal investigator. Other staff in the Influenza Division is consulted as needed. Each EIP site analyses and reports their data, as needed.

For the HAIC-A, statistical consultation has been provided by Jonathan Edwards, MStat (JREdwards@cdc.gov, 404-639-4711) and Yi Mu, PhD (hrb3@cdc.gov, 404-639-4223). Data will be collected by EIP personnel and by local facility staff, as described previously. Identification of the specific EIP surveillance officers and local facility staff members who will participate in training and data collection activities is at the discretion of the EIP site or the facility, respectively. Analyses are prioritized by the HAIC-A Steering Group; major analyses are typically performed by CDC staff, while site-specific analyses or special multi-site analyses may be performed by CDC or EIP site staff.