Emergency Epidemic Investigation Data Collections
OMB Control No. 0920-1011 (Expiration: 12/31/2025)

Supporting Statement B

Project Officer: Jeffrey M. Zirger, Ph.D.
wtj5@cdc.gov

Information Collection Review Office (ICRO)
Office of Science (OS)
Centers for Disease Control and Prevention
1600 Clifton Road, NE
Atlanta, GA 30333
Telephone: (470) 217-8909

August 6, 2025

Table of Contents

B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Sampling Methods

The respondent universe is the population to which results will be generalized. The respondent universe for each individual Emergency Epidemic Investigation (EEI) will vary, depending on the outbreak or event, but is not a mechanism to collect data that can be generalized to the broader population. The population to which results are generalizable is limited to those potentially exposed to or affected by the health event under investigation. For example, the respondent universe might be defined by identity (e.g., employees at a particular workplace), a behavior (e.g., people who ate at Restaurant X during a certain time period), or geographic characteristic (e.g., residents of Town Z). CDC, in collaboration with external partners, identifies the respondent universe for each Emergency Epidemic Investigation (EEI).

The sampling methods also will vary depending on the outbreak or event. Most investigations of smaller scale outbreaks or events (e.g. where a few to several hundred individuals are involved) include collecting information from all individuals affected by the condition in question and appropriate controls. However, in certain investigations, such as those involving larger numbers of individuals, investigators might choose to collect information from a sample of affected individuals and appropriate controls.

For some investigations there will be multiple study designs, thus multiple sampling methodologies might be used. For example, when little is known about a specific outbreak or event, a hypothesis-generating investigation might be conducted with a convenience sample of individuals. Information from this initial investigation might be used to identify an appropriate respondent universe for the hypothesis-testing phase of the investigation. Additional information then might be collected from either a census or sample, depending on the size of the respondent universe. In case-control studies, different sampling methods might be used to select case-patients and controls. For example, case-patients might be randomly selected from a line list and controls might be selected based on pair-matching (i.e., one or more matching controls selected for each case based on certain characteristics such as age, sex, geographic location, having eaten at a particular restaurant, etc.). CDC, in collaboration with the external partners requesting epidemiological assistance, identifies the sampling methodology for each EEI based on the information needed to identify the agents, sources, modes of transmission, or risk factors.

When sampling methods are used, the number of respondents selected for each investigation will depend upon obtaining a sufficient number of respondents to

determine the agents, sources, modes of transmission, or risk factors to implement prevention and control measures.

When appropriate, power calculations will be conducted to ensure the number of respondents will provide sufficient power for reliable statistical inferences. CDC, in collaboration with the external partners requesting epidemiological assistance will determine the sample size for each Emergency Epidemic Investigation (EEI). Based on data from previous EEIs, on average, the expected number of respondents per investigation is 200.

Procedures for each investigation, including the method and mode of data collection, depend on the nature of the outbreak or event, hypotheses to be tested, time and resources available, number of persons involved, and other circumstances unique to the emergency at hand. EEIs utilize one or more methodological approaches such as: epidemiologic investigation, environmental assessment, and laboratory testing. The components employed will vary depending on the information needed to determine the agents, sources, modes of transmission, or risk factors in order to inform prevention and control measures.

EEIs often include the steps below, though the steps included in each EEI will vary depending on what information already has been determined at the time CDC assistance is requested and the amount of information needed to identify prevention and control measures.

1. Prepare for field work
2. Establish the existence of an outbreak
3. Verify the diagnosis
4. Define and identify cases
5. Describe and orient the data in terms of time, place, and person
6. Develop hypotheses
7. Evaluate hypotheses
8. Refine hypotheses and carry out additional studies
9. Implement control and prevention measures
10. Communicate findings

Study Design

Epidemiologic Investigations

EEI steps 1–6 often are carried out using descriptive study methods. Steps 7–8 often are carried out using analytic study methods. Methods commonly used in epidemiologic investigations are described below, though this list is not exhaustive.[1] Environmental assessments and laboratory testing also are briefly described, though these methods often do not involve respondent burden.

[1] Definitions adopted from *Principles of Epidemiology in Public Health* Practice, 3rd Edition. Developed by: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Office of Workforce and Career Development, Career Development Division, Atlanta, GA 30333. Available: http://www.cdc.gov/osels/scientific_edu/ss1978/SS1978.pdf.

- Descriptive Study

  Definition: A study designed to organize and summarize data regarding the persons effected (e.g., the characteristics of those who became ill), time (e.g., when they become ill), and place (e.g., where they might have been exposed to the cause of illness).

  Common Uses: Descriptive studies will be employed by almost all EEIs where the person, place, and time of a particular outbreak or event are not already understood. This information is critical for identification of the population at risk, development of the case definition, case finding, and hypothesis generation. For example, in the Arizona Guillain-Barre Syndrome investigation described in Supporting Statement A, descriptive data were collected via medical record abstraction and interviews with patients, family, and caregivers to verify the diagnosis and confirm the existence of an outbreak.

- Cross-sectional Study

  Definition: A study in which a sample of persons from a population is enrolled and their exposures and health outcomes are measured simultaneously; a survey.

  Common Uses: Cross-sectional studies often are used when sufficient information is not available to define the case or exposure. In this case, information on exposures and illness can be gathered from a subpopulation to generate hypotheses about agents, sources, modes of transmission, or risk factors. For example, in the Arizona Guillain-Barre Syndrome investigation described in Supporting Statement A, a food exposure questionnaire and open-ended interviews were conducted with case-patients to generate hypotheses about potential sources of the outbreak.

- Cohort Study

  Definition: An analytic study in which enrollment is based on status of exposure to a certain factor or membership of a certain group. Populations are then compared according to whether they experienced disease, death, or other health-related outcome. In retrospective cohort studies, the exposures and outcomes already have occurred at the time of the investigation.

  Common Uses: Cohort studies are commonly used in investigations where 1) the exposure is rare or can be precisely defined, 2) the population at risk is a well-defined group (such as workers at a particular factory or attendees of a wedding), or 3) the disease is common. For example, in a 2012 investigation of *Campylobacter jejuni* infections in 3 states, interviews with case-patients revealed each had a connection with a specific restaurant (Restaurant A). A retrospective cohort study of patrons of Restaurant A was conducted to identify specific foods eaten and illness

status. Results of this investigation, along with environmental sampling and laboratory testing, identified the implicated food source.

- Case-control Study

  Definition: An analytic study that enrolls one group of persons with a certain disease, chronic condition, or type of injury (case-patients) and a group of persons without the health problem (controls) and compares differences in exposures, behaviors, and other characteristics to identify and quantify associations, test hypotheses, and identify causes.
  Common Uses: Case-control studies often are used when 1) the disease or outcome is rare, 2) persons with the disease or outcome can be readily identified, 3) multiple exposures are under investigation or the exposure is common, or 4) the population at risk is unknown. For example, in the Arizona Guillain-Barre Syndrome investigation described in Supporting Statement A, a case-control study, was conducted among city residents to compare multiple potential exposures between case-patients and controls and ultimately identify the likely source of the outbreak.

Environmental Assessment and Laboratory Testing
Many EEIs employ environmental assessment and laboratory testing in addition to epidemiologic investigations. These investigation components can be critical in verifying diagnoses, identifying cases, and developing hypotheses or verifying hypotheses. Environmental assessments can provide important source and exposure information. Laboratory testing can provide important information about diagnoses or source of disease. For example, if descriptive data indicate all case-patients ate at a particular restaurant, an environmental assessment and laboratory testing might identify a specific contaminated food item, eliminating the need for additional data collection. In another example, descriptive data might not yield sufficiently conclusive results to inform an environmental assessment or laboratory testing. Instead, hypotheses might be generated and tested, and then environmental assessment and laboratory testing might be used to confirm or verify results of the analytic study. EEIs might involve laboratory testing of environmental or biospecimen samples.

Data Collection Mode

Data collection modes commonly used by EEIs are described below; other modes also might be used as appropriate. Many EEIs use multiple modes to collect data. Ultimately, the type of mode(s) used will be determined based on the location and availability of respondents, the speed with which information is needed, and the specific information needed to identify the agents, sources, modes of transmission, or risk factors so that effective prevention and control measures can be implemented.

- Survey Modes

- Face-to-Face Interview

  Definition: An in-person interview in which a trained interviewer administers a questionnaire to the respondent.

  Common Uses: Face-to-face interviews are a common mode used in EEIs because they allow for rapid data collection and are conducive to open-ended responses. This makes them particularly useful during the hypothesis-generating stage of the investigation. They also are useful when the information to be collected is particularly complex and interviewer prompts or explanation might be needed. Face-to-face interviews are most feasible when respondents are centrally located.
- Telephone Interview

  Definition: A telephone interview in which a trained interviewer administers a questionnaire to the respondent over the phone.

  Common Uses: Telephone interviews are a common mode used in EEIs because they allow for rapid data collection and are conducive to open-ended responses. This makes them particularly useful during the hypothesis-generating stage of the investigation. They also are useful when the information to be collected is particularly complex and interviewer prompts or explanation might be needed. Telephone interviews are often used instead of face-to-face interviews when respondents are not centrally located or when resources do not allow for face-to-face interviews.
- Self-administered Paper-and-Pencil Questionnaire

  Definition: A paper questionnaire that is completed by the respondent.

  Common Uses: Self-administered paper-and-pencil questionnaires are often used when interviews are not feasible due to resource limitations and the information to be collected can be captured using straight-forward questions with fixed response options.
- Self-administered Internet Questionnaire

  Definition: A self-administered questionnaire that is completed by the respondent on a computer or other electronic device via an Internet connection.

  Common Uses: Internet questionnaires are used less frequently in EEIs. Due to the rapid nature of the response required in an EEI there often is little time for developing a web-based data collection tool. However, internet questionnaires can be used when respondents have internet and computer access. They are particularly useful when respondents are not centrally located.

- Medical Record Abstraction

  Definition: Abstraction of data from existing medical or laboratory records.

Common Uses: Medical record abstraction can provide important information about medical history, symptoms, and diagnoses. Medical record abstraction is particularly useful when case-patients or controls are no longer available for data collection, detailed medical information is needed that might be difficult to recall, information for case definitions or case confirmation is needed, or descriptive data are needed for hypotheses-generation. Medical record abstraction relies on data that already have been collected and recorded and therefore presents no burden to the individual about whom information is collected.

- Biospecimen Sample
  Definition: A sample of material, such as urine, blood, tissue, cells, DNA, RNA, and protein from humans, animals, or plants.[2]
  Common Uses: Biospecimen samples are often used to determine case definitions, verify case status, or identify the infectious agent or source.

- Environmental Sample:
  Definition: A sample of any material that is collected from an environmental source.
  Common Uses: Environmental samples are often used to determine the agents, sources, or modes of transmission of the health problem.

2. Procedures for the Collection of Information

Because of the acute nature of the outbreaks or events to be investigated, periodic data collection is not employed in most EEIs. Data collection to identify agents, sources, modes of transmission, or risk factors occur over a period of no greater than 90 days to allow for rapid implementation of effective prevention and control measures.

- Statistical method for stratification and sample selection

  When statistical methods are employed in the collection of information, CDC provides statistical assistance relating to sampling methodology and selection of controls.

- Estimation procedure

  Data analysis is conducted under the advice of statisticians or data analysts from CDC or the requesting organization and will involve descriptive statistics. Additional bivariate and multivariate analyses are conducted as needed to identify the agents, sources, modes of transmission, or risk factors

[2] Definition from the NCI Dictionary of Cancer Terms available: http://www.cancer.gov/dictionary?cdrid=561324.

so that effective prevention and control measures can be implemented.

- Degree of accuracy needed for the purpose described in the justification

  The purpose of the EEI is to collect information rapidly to identify unknown information (e.g., agents, sources, modes of transmission, or risk factors) necessary for instituting effective prevention and control measures. Quality control procedures will be implemented in each EEI to the extent possible given the rapid nature of the data collection. For example, when possible, instruments and methodology will be pilot tested. In each investigation, the goal is to collect the best data possible in the timeliest manner possible to direct public health action.

- Unusual problems requiring specialized sampling procedures

  CDC does not expect unusual problems requiring specialized sampling.

- Any use of periodic (less frequent than annual) data collection cycles to reduce burden

  Because of the acute nature of the outbreaks or events to be investigated, periodic data collection is not employed. Data collection to identify agents, sources, modes of transmission, or risk factors occur over a period of no greater than 90 days to allow for rapid implementation of effective prevention and control measures.

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

In EEIs, nonresponse bias typically is expected. For example, case-patients who have experienced a negative health event will have more interest in responding than will controls. Often, controls (unaffected by the health event under investigation) will be oversampled to account for this potential bias. In another example, case-patients who are most ill might be least likely to respond. This might be countered by follow-up procedures with non-respondents. Also, for each EEI, response rates are maximized by informing potential respondents of the critical nature of the outbreak or event and the importance of collecting information to identify effective prevention and control measures. Before collecting information, investigators inform respondents that participation is voluntary, that respondents are not personally identified in any published reports of the study, and that their privacy will be protected to the extent allowed under federal law. Study designs and epidemiologic methods are chosen to minimize the effect of non-response bias and EEI investigators will acknowledge when and how it might impact their results.

4. Test of Procedures or Methods to be Undertaken

Though each data collection instrument is tailored to the needs of each specific outbreak or event, questions from instruments employed in previous investigations are used when possible. In this way, data collection instruments that are refined over time can be utilized. A data collection instrument library (**Appendices**) is maintained by archiving the final data collection instruments administered in EEIs under this generic ICR.

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

CDC EEI investigators are trained in biostatistics and epidemiology. In most cases, investigators collaborate extensively with health officials of the external partners requesting assistance. All investigations are supervised by CDC's experienced epidemiologists. Expert statistical resources are provided by CDC.