INVESTIGATION · SCIENTIFIC AND TECHNICAL INVESTIGATION

Epidemiological Investigation: Tracing the Source of Outbreaks

When people fall ill with the same disease at the same time, epidemiologists investigate to find the source and stop the spread. This article explains how outbreaks are detected, the steps of a field investigation, how the source is identified using descriptive and analytical methods and laboratory evidence, and lessons from landmark outbreaks.

ModQu research article · 2,267 words · 22 references (APA 7th edition)

An outbreak is the occurrence of more cases of a disease than would normally be expected in a particular place, group or time. It may be a cluster of food poisoning after a wedding, a rise in Legionnaires' disease linked to a cooling tower, or the first signs of a new infection spreading in a city. Outbreak investigation is the work of finding out what is causing the illness, how it is spreading and how to stop it. It is one of the oldest forms of scientific investigation in public health, going back at least to John Snow's investigation of cholera in London in 1854 (Snow, 1855).

Outbreak investigation shares much with other kinds of investigation. It starts with a problem that must be defined, gathers evidence from people and records, develops hypotheses about the cause, tests them systematically and acts on the findings. It also has features of its own: it usually takes place under time pressure, while people continue to fall ill, and its results must be strong enough to justify actions such as closing a restaurant, recalling a food product or treating a water supply. Goodman et al. (1990) describe field investigation as a combination of science and judgement, because investigators must often act before every question has been answered.

Detecting outbreaks

Outbreaks are detected in several ways. Public health surveillance, the routine collection and analysis of reports of disease, is the foundation. Langmuir (1963) defined surveillance as the continued watchfulness over the distribution and trends of disease, combined with the timely sharing of the results with those who need to know. Doctors and laboratories are legally required to report certain diseases, and surveillance staff look for increases above the expected level (Thacker & Berkelman, 1988).

Outbreaks are also detected by alert clinicians who notice unusual cases, by members of the public who report that several people became ill after the same event, and increasingly by laboratory networks that compare the genetic fingerprints of bacteria from patients in different places. Informal sources, such as news reports and online posts, can also provide early warning, especially for outbreaks in places with weak official reporting (Brownstein et al., 2009). Internationally, the International Health Regulations require countries to notify the World Health Organization of events that may constitute a public health emergency of international concern (World Health Organization [WHO], 2016).

The steps of a field investigation

Guides to field epidemiology describe a series of steps, which in practice overlap and are often carried out at the same time (Centers for Disease Control and Prevention [CDC], 2012; Gregg, 2008; Reingold, 1998).

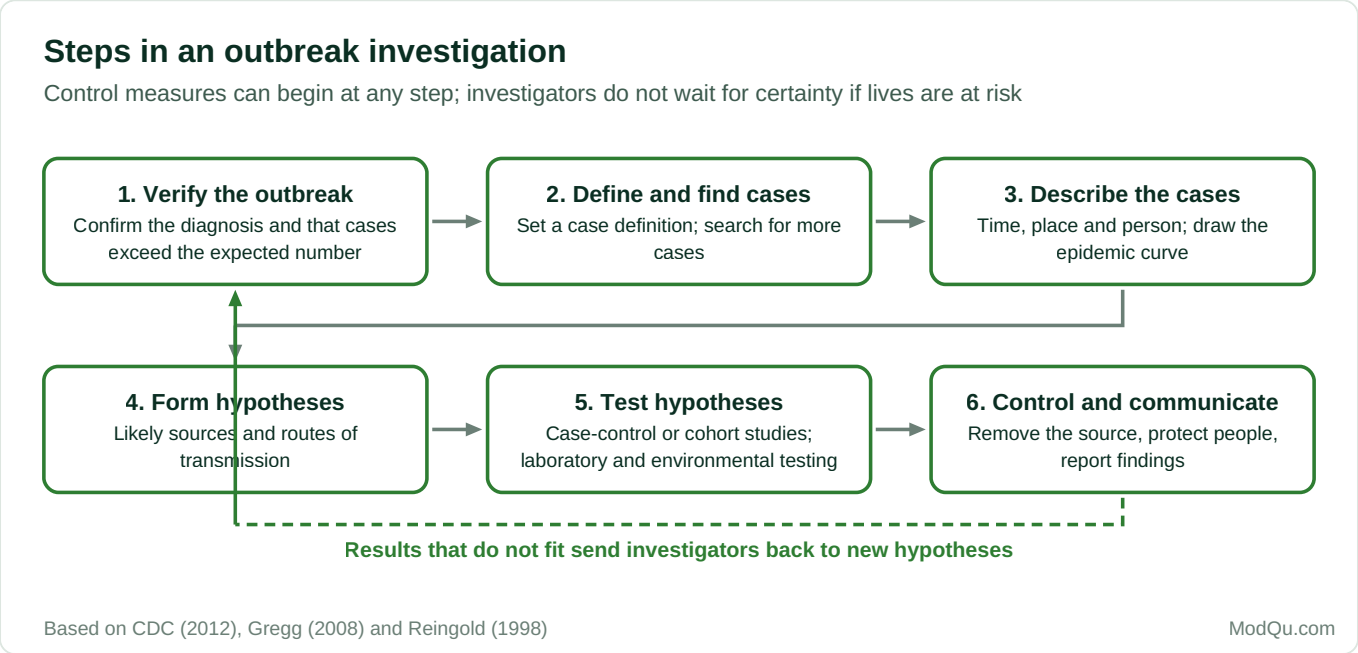


Figure 1. Six core steps of an outbreak investigation. In practice they overlap, and control measures often begin as soon as a likely source is suspected. Diagram by ModQu.

Verifying the outbreak. The first task is to confirm that there really is an outbreak. An apparent increase in cases may result from a change in reporting, a new diagnostic test, more people being tested, or a laboratory error. Investigators confirm the diagnosis, often with laboratory tests, and compare the number of cases with the number normally expected.

Defining and finding cases. A case definition sets out who will be counted as a case, usually using clinical criteria such as symptoms, laboratory results, and limits of time, place and person. Definitions may have levels, such as "confirmed", "probable" and "possible". Investigators then search for further cases by contacting doctors, hospitals and laboratories, and by asking people who attended the same event or used the same service.

Describing the cases. Investigators describe the cases by time, place and person. Time is shown in an epidemic curve, a chart of the number of cases by date or hour of onset. Its shape gives clues about the type of outbreak.

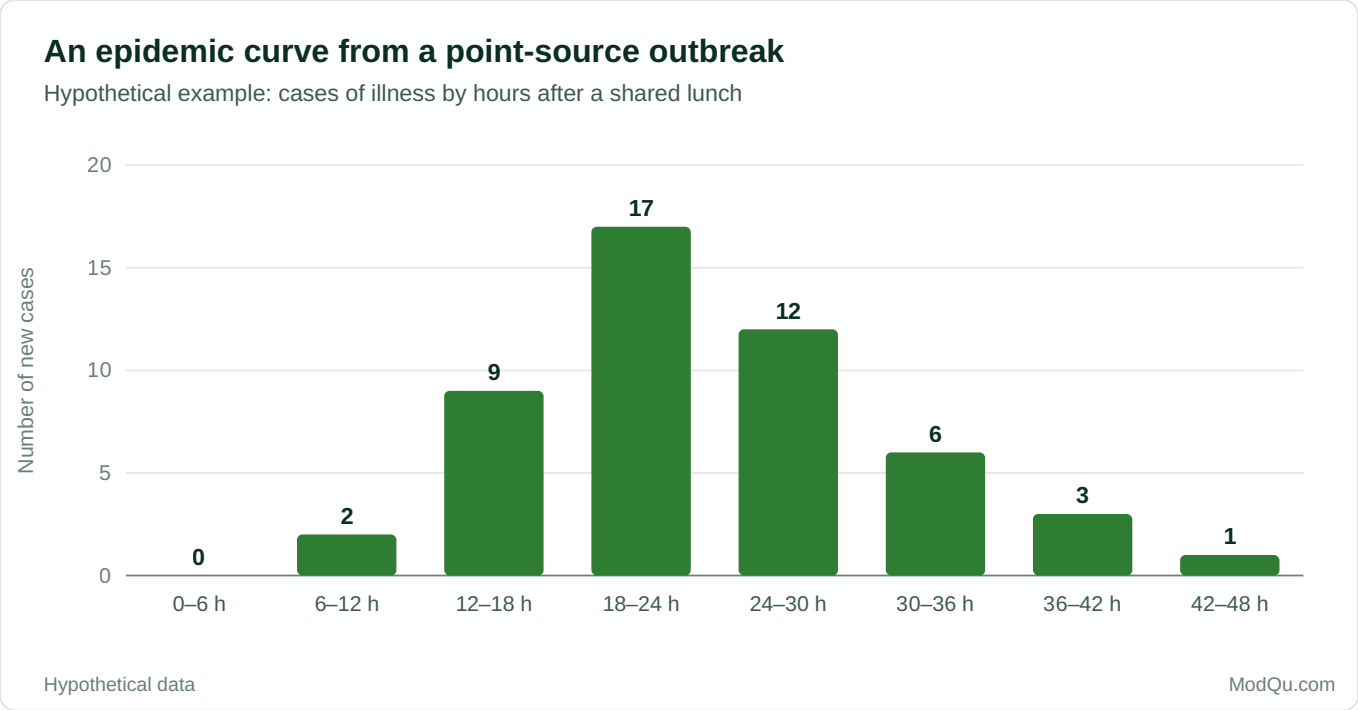


Figure 2. When everyone is exposed to the same source at the same time, cases rise and fall in a single peak over about one incubation period. A curve with repeated or continuing peaks suggests a continuing source or spread from person to person. Chart by ModQu.

A single sharp peak suggests a point source, such as one contaminated meal, and the timing of the peak can indicate the likely incubation period and therefore the likely organism. A plateau suggests a continuing source, and a series of increasing peaks suggests spread from person to person. Mapping cases can reveal geographic patterns, such as clustering near a water source or along a delivery route. Describing who is affected, by age, sex, occupation and other characteristics, can point to the exposures they share.

Forming hypotheses. Using these descriptions, interviews with cases and knowledge of the disease, investigators form hypotheses about the source and the route of transmission. Detailed "hypothesis-generating" interviews with a small number of cases, asking about everything they ate, drank and did, often reveal a common exposure.

Testing hypotheses. Hypotheses are tested with analytical studies that compare people who became ill with people who did not. In a cohort study, used when a defined group was exposed, such as guests at a wedding, investigators calculate the rate of illness among those who did and did not eat each food. In a case-control study, used when the population at risk is not clearly defined, investigators compare the exposures of cases with those of similar people who did not become ill (Rothman et al., 2008). A food or exposure that was much more common among cases than among controls, or that was associated with a much higher attack rate, is a strong suspect.

Laboratory and environmental evidence. Laboratories identify the organism in patients and, where possible, in suspected sources such as food, water or equipment. Matching the strain found in patients to the strain found in the source provides strong confirmation.

Control and communication. Control measures, such as removing a food from sale, closing a facility, issuing advice to boil water or vaccinating contacts, are taken as soon as the evidence justifies them. Investigators then report their findings to the public, the authorities and the scientific community so that similar outbreaks can be prevented.

Judging cause

Epidemiological evidence shows associations: certain exposures are more common among people who fell ill. Deciding whether an association reflects cause requires judgement. Hill (1965) proposed a set of considerations, including the strength of the association, its consistency across studies, its specificity, the correct time order, a dose-response relationship, biological plausibility, and agreement with other evidence. In outbreaks, the combination of a strong statistical association, a plausible route of transmission and a laboratory match between patients and source usually provides convincing evidence.

Molecular epidemiology

Laboratory methods have transformed outbreak investigation. PulseNet, a network of public health laboratories in the United States established in 1996, compares the DNA fingerprints of bacteria such as Salmonella and E. coli from patients across the country, allowing investigators to link cases that would otherwise appear unrelated (Swaminathan et al., 2001). Whole genome sequencing, which determines the complete genetic code of each isolate, now allows much finer distinctions between strains and has become the standard method for foodborne disease surveillance in many countries (Gerner-Smidt et al., 2019). It can show that cases scattered across many regions share an identical strain, and it can link patients to a food or factory with high confidence.

Landmark investigations

Several outbreak investigations have shaped the field.

Legionnaires' disease, 1976. After a convention of the American Legion in Philadelphia, more than 180 people developed pneumonia and 29 died. Epidemiologists found that illness was associated with spending time in or near the convention hotel, and suspected airborne spread (Fraser et al., 1977). Months later, laboratory scientists identified a previously unknown bacterium, later named Legionella pneumophila (McDade et al., 1977). The bacterium was later found to grow in water systems such as cooling towers, leading to new rules for maintaining them.

The first reports of AIDS, 1981. In June 1981, the Centers for Disease Control published a report of a rare form of pneumonia in five previously healthy young men in Los Angeles (Centers for Disease Control, 1981). This short report, produced by clinicians who noticed an unusual pattern, marked the beginning of the recognition of AIDS and showed the value of alert clinicians and rapid reporting.

Milwaukee water supply. An estimated 403,000 people in Milwaukee became ill with diarrhoea caused by the parasite Cryptosporidium, which had passed through one of the city's water treatment plants. Investigators used surveys of residents, reports from laboratories and records of water quality to link the outbreak to that plant (Mac Kenzie et al., 1994). It remains one of the largest waterborne outbreaks recorded in the United States and led to changes in water treatment standards.

Walkerton, 2000. In the Ontario town of Walkerton, contamination of the drinking water with E. coli O157:H7 and Campylobacter made around 2,300 people ill, and seven died. The investigation and public inquiry found that manure from a farm had entered a shallow well after heavy rain, and that operators had failed to chlorinate the water adequately and had falsified records, while provincial oversight had been weakened (O'Connor, 2002; Hrudey et al., 2003). The case shows how outbreak investigation can expose institutional as well as technical failures.

Contact tracing

For infections that spread between people, investigators trace the contacts of each case so that they can be tested, treated, vaccinated or asked to isolate. Eames and Keeling (2003) showed with mathematical models how effective contact tracing can control the spread of infection, and that its effectiveness depends on how quickly contacts are found and how much transmission occurs before symptoms appear. During the COVID-19 pandemic, Ferretti et al. (2020) estimated that much transmission occurred before symptoms and argued that manual tracing would be too slow on its own, prompting the development of digital tracing apps. Heymann (2015) gives standard guidance on contact management for many specific infections.

Pitfalls in outbreak investigation

Investigators face several recurring difficulties. Recall is one: people asked what they ate several days ago often cannot remember accurately, and those who became ill may search their memories harder than those who did not, which can create false associations. Using menus, receipts, loyalty card data and photographs can help. Small numbers are another: in a small outbreak, chance alone can produce apparent associations, so investigators look for consistency between epidemiological, laboratory and environmental evidence rather than relying on a single statistical result.

A third difficulty is that the obvious suspect may be wrong. Foods that are commonly blamed, such as shellfish or chicken, are not always the cause, and less obvious items such as salads, sprouts, spices or ice can carry infection. An investigation that confirms its first hypothesis without testing alternatives risks closing the wrong business while the real source continues to make people ill. Finally, by the time an outbreak is recognised, the suspected food may have been eaten or discarded, so investigators often must rely on indirect evidence such as supplier records and the testing of other batches.

Communication and ethics

Outbreak investigators must decide what to tell the public and when. Announcing a suspected source too early can harm a business that turns out not to be responsible, and can lead people to ignore the true source. Announcing too late can allow more people to fall ill. Most public health agencies therefore act on the balance of evidence, issuing precautionary advice while making clear what is confirmed and what is still being investigated, and updating the public as evidence develops.

Investigations also raise questions of privacy. Case interviews collect detailed personal information, including where people went and whom they met. Public health law usually authorises this collection, but investigators must protect the information, use it only for disease control and avoid publishing details that could identify patients. Trust matters: people who fear that information will be used against them, for example in relation to their immigration status or sexual behaviour, may not cooperate, and the investigation will suffer.

A short example

The following hypothetical case illustrates the method. Thirty of the 120 guests at a wedding reception develop vomiting and diarrhoea. The local health department verifies the outbreak, confirms norovirus in samples from several patients, and defines a case as any guest who developed vomiting or diarrhoea within three days of the reception.

The epidemic curve shows a single peak about 36 hours after the meal, consistent with a point source. Investigators send a questionnaire to all guests about what they ate. Attack rates are similar for most dishes, but 60 per cent of guests who ate the salad became ill, compared with 5 per cent of those who did not. Inspection of the caterer's kitchen finds that one food handler had been ill with vomiting the day before the wedding and had prepared the salad by hand. Samples from the food handler contain the same norovirus strain as the patients.

The health department advises the caterer on excluding ill staff for 48 hours after symptoms stop, and on hand hygiene and the preparation of ready-to-eat food. The investigation report is shared with other local caterers.

Conclusion

Outbreak investigation is a disciplined form of inquiry that moves from verifying a problem, through defining and describing cases, to forming and testing hypotheses about the source, and finally to control and communication. It combines interviews, statistics, laboratory science and environmental inspection, and it increasingly relies on genetic sequencing to link cases with high precision. Landmark investigations from Legionnaires' disease to Walkerton show that the same method that identifies a contaminated salad can uncover new diseases and institutional failures. Its lasting value lies in turning the experience of the people who fell ill into knowledge that protects others.

References

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection—Harnessing the web for public health surveillance. *The New England Journal of Medicine, 360*(21), 2153–2157. https://doi.org/10.1056/NEJMp0900702

Centers for Disease Control and Prevention. (2012). *Principles of epidemiology in public health practice: An introduction to applied epidemiology and biostatistics* (3rd ed.). U.S. Department of Health and Human Services.

Centers for Disease Control. (1981). Pneumocystis pneumonia—Los Angeles. *Morbidity and Mortality Weekly Report, 30*(21), 250–252.

Eames, K. T. D., & Keeling, M. J. (2003). Contact tracing and disease control. *Proceedings of the Royal Society B: Biological Sciences, 270*(1533), 2565–2571. https://doi.org/10.1098/rspb.2003.2554

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., & Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science, 368*(6491), Article eabb6936. https://doi.org/10.1126/science.abb6936

Fraser, D. W., Tsai, T. R., Orenstein, W., Parkin, W. E., Beecham, H. J., Sharrar, R. G., Harris, J., Mallison, G. F., Martin, S. M., McDade, J. E., Shepard, C. C., & Brachman, P. S. (1977). Legionnaires' disease: Description of an epidemic of pneumonia. *The New England Journal of Medicine, 297*(22), 1189–1197. https://doi.org/10.1056/NEJM197712012972201

Gerner-Smidt, P., Besser, J., Concepción-Acevedo, J., Folster, J. P., Huffman, J., Joseph, L. A., Kucerova, Z., Nichols, M. C., Schwensohn, C. A., & Tolar, B. (2019). Whole genome sequencing: Bridging one-health surveillance of foodborne diseases. *Frontiers in Public Health, 7*, Article 172. https://doi.org/10.3389/fpubh.2019.00172

Goodman, R. A., Buehler, J. W., & Koplan, J. P. (1990). The epidemiologic field investigation: Science and judgment in public health practice. *American Journal of Epidemiology, 132*(1), 9–16. https://doi.org/10.1093/oxfordjournals.aje.a115647

Gregg, M. B. (Ed.). (2008). *Field epidemiology* (3rd ed.). Oxford University Press.

Heymann, D. L. (Ed.). (2015). *Control of communicable diseases manual* (20th ed.). APHA Press.

Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine, 58*(5), 295–300. https://doi.org/10.1177/003591576505800503

Hrudey, S. E., Payment, P., Huck, P. M., Gillham, R. W., & Hrudey, E. J. (2003). A fatal waterborne disease epidemic in Walkerton, Ontario: Comparison with other waterborne outbreaks in the developed world. *Water Science and Technology, 47*(3), 7–14. https://doi.org/10.2166/wst.2003.0146

Langmuir, A. D. (1963). The surveillance of communicable diseases of national importance. *The New England Journal of Medicine, 268*(4), 182–192. https://doi.org/10.1056/NEJM196301242680405

Mac Kenzie, W. R., Hoxie, N. J., Proctor, M. E., Gradus, M. S., Blair, K. A., Peterson, D. E., Kazmierczak, J. J., Addiss, D. G., Fox, K. R., Rose, J. B., & Davis, J. P. (1994). A massive outbreak in Milwaukee of Cryptosporidium infection transmitted through the public water supply. *The New England Journal of Medicine, 331*(3), 161–167. https://doi.org/10.1056/NEJM199407213310304

McDade, J. E., Shepard, C. C., Fraser, D. W., Tsai, T. R., Redus, M. A., & Dowdle, W. R. (1977). Legionnaires' disease: Isolation of a bacterium and demonstration of its role in other respiratory disease. *The New England Journal of Medicine, 297*(22), 1197–1203. https://doi.org/10.1056/NEJM197712012972202

O'Connor, D. R. (2002). *Report of the Walkerton Inquiry: Part one. The events of May 2000 and related issues*. Ontario Ministry of the Attorney General.

Reingold, A. L. (1998). Outbreak investigations—A perspective. *Emerging Infectious Diseases, 4*(1), 21–27. https://doi.org/10.3201/eid0401.980104

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed.). Lippincott Williams & Wilkins.

Snow, J. (1855). *On the mode of communication of cholera* (2nd ed.). John Churchill.

Swaminathan, B., Barrett, T. J., Hunter, S. B., & Tauxe, R. V. (2001). PulseNet: The molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerging Infectious Diseases, 7*(3), 382–389. https://doi.org/10.3201/eid0703.017303

Thacker, S. B., & Berkelman, R. L. (1988). Public health surveillance in the United States. *Epidemiologic Reviews, 10*, 164–190. https://doi.org/10.1093/oxfordjournals.epirev.a036021

World Health Organization. (2016). *International Health Regulations (2005)* (3rd ed.). WHO.