

Public Health Surveillance in the Twenty-First Century: Achieving Population Health Goals While Protecting Individuals' Privacy and Confidentiality

MICHAEL A. STOTO*

TABLE OF CONTENTS

INTRODUCTION	703
I. HISTORY OF PUBLIC HEALTH SURVEILLANCE	705
II. NEW APPROACHES TO DISEASE SURVEILLANCE THAT DEMAND A REEVALUATION OF THE BALANCE BETWEEN PROTECTING INDIVIDUALS' RIGHTS AND MAINTAINING THE EFFECTIVENESS OF DISEASE CONTROL PROGRAMS	708
A. HIV REPORTING: A CASE-BASED APPROACH WITH A STATISTICAL PURPOSE	708
B. SYNDROMIC SURVEILLANCE: COLLECTING INDIVIDUAL-LEVEL DATA TO DETECT DISEASE OUTBREAKS	711
C. SCREENING FOR DIABETES AND OBESITY: CASE REPORTING APPLIED TO NON-COMMUNICABLE DISEASES	715
CONCLUSION	717

INTRODUCTION

Surveillance, a core function of public health, is defined as “ongoing, systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know.”¹ In the context of a session that addresses how data and information can and should inform public health policy and practice, this discussion of surveillance calls attention to the disclosure and use of personal health information. In particular, public health surveillance programs require a careful balance between the development of statistical and epidemiological data and knowledge that are essential to achieving population health goals and the protection of individuals' privacy and confidentiality rights.

* Professor of Health Services Administration and Population Health, Georgetown University School of Nursing & Health Studies. © 2008, Michael A. Stoto.

1. Stephen B. Thacker & Ruth L. Berkelman, *Public Health Surveillance in the United States*, 10 EPIDEMIOLOGIC REVS. 164, 164 (1988) (quoting CTRS. FOR DISEASE CONTROL, COMPREHENSIVE PLAN FOR EPIDEMIOLOGIC SURVEILLANCE ii (1986)).

Public health surveillance, as it is usually defined, includes two very different activities. *Case surveillance* focuses on individuals, or sometimes small groups of individuals, and serves to identify those with certain diseases and takes action to stop disease spread beyond these identified individuals. Historically, case surveillance has been used for communicable diseases capable of causing great harm to the entire population if allowed to spread. The loss of privacy involved with this type of surveillance has been justified in terms of disease averted. In contrast, *statistical surveillance* uses populations to identify differentials and trends that can inform public health policymaking, including the allocation of resources. Individuals need not be identified for the surveillance to serve its purpose, so data can be gathered either anonymously or with promises of confidentiality, thus not violating privacy rights. Both surveillance approaches have roots going back centuries, but it was not until 1963 that Alexander Langmuir of the Communicable Disease Center (now the Centers for Disease Control and Prevention, or CDC) combined them in his classic definition of public health surveillance,² the basis for the CDC's definition quoted above.

Case surveillance and statistical surveillance have different goals and objectives, data sources, and methods. Over time, each approach has resolved the tradeoffs between population benefits and individuals' privacy and confidentiality rights in its own way. In recent years, however, new surveillance programs have been developed that combine, and sometimes confuse, the case and statistical approaches. Individual HIV case reporting, for instance, is advocated as a means of estimating the relative number of cases in different parts of the country in an effort to allocate federal resources.³ In some parts of the country, individually identified hospital emergency room records are transmitted to health departments, which use them in statistical analyses to detect disease outbreaks and covert bioterrorist attacks.⁴ Furthermore, individual case reports are increasingly utilized to monitor obesity, diabetes, and other non-communicable diseases.⁵

2. See Alexander D. Langmuir, *The Surveillance of Communicable Diseases of National Importance*, 268 NEW ENG. J. MED. 182, 182–83 (1963).

3. See INST. OF MED., MEASURING WHAT MATTERS: ALLOCATION, PLANNING, AND QUALITY ASSESSMENT FOR THE RYAN WHITE CARE ACT 12 (2004) (recommending use of either code- or name-based individual case reporting to “improve consistency, quality, and comparability in HIV case reporting for [Ryan White CARE Act] allocation purposes”).

4. See Kenneth D. Mandl et al., *Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience*, 11 J. AM. MED. INFORMATICS ASS'N 141, 143–44 (2004) (“Real-time data streams from . . . emergency department encounters have been established successfully in a number of regions.”); Michael A. Stoto, *Syndromic Surveillance*, ISSUES SCI. & TECH., Spring 2005, at 50 (“This theory [of syndromic surveillance] was turned into a reality when some health departments . . . began to monitor hospital ER admissions and other data streams.”).

5. See, e.g., UNIV. OF ARK. FOR MED. SCI. FAY W. BOOZMAN COLL. OF PUB. HEALTH, YEAR THREE EVALUATION: ARKANSAS ACT OF 2003 TO COMBAT CHILDHOOD OBESITY 27 (2006) (describing Arkansas legislation mandating in-school obesity screening); Thomas R. Frieden, *Asleep at the Switch: Local Public Health and Chronic Disease*, 94 AM. J. PUB. HEALTH 2059, 2059 (2004) (describing New York City's diabetes surveillance program).

Novel approaches to disease surveillance that combine the objectives and approaches of case and statistical surveillance demand a reevaluation of the balance between preserving individuals' rights and the need for effective public health tools. To address these issues, this Essay begins by reviewing the history of public health surveillance to understand the tradeoffs that have been made in the past. The Essay then analyzes the three examples above to illustrate the issues involved and the new tradeoffs that must be addressed. Drawing on these analyses, the final Part proposes some considerations that can help public health find an appropriate balance between population benefits of surveillance and the protection of individuals' privacy and confidentiality rights.

I. HISTORY OF PUBLIC HEALTH SURVEILLANCE⁶

The case surveillance approach is fundamental to public health as it uses the police power of the state to control communicable diseases. This approach was used in the Republic of Venice during the fourteenth century, for example, when authorities boarded ships to identify persons with symptoms of bubonic plague and prevent them from disembarking.⁷ Similarly, in 1741, Rhode Island required tavern keepers to report patrons with contagious disease, including smallpox, yellow fever, and cholera, to local authorities.⁸ A postcard reporting format was developed in Massachusetts in 1874, with the resulting information compiled into weekly reports.⁹ In 1878, Congress authorized the forerunner of the United States Public Health Service to collect morbidity data for use in quarantine measures against "pestilential diseases" such as cholera, smallpox, plague, and yellow fever.¹⁰ Compulsory reporting of infectious diseases began on a national basis in Italy in 1881, and in other European countries shortly afterwards.¹¹ At the global level, the World Health Organization (WHO) modified the International Health Regulations in 2005 to require that all countries notify the WHO of all events "which may constitute a public health emergency of international concern."¹² These regulations also require that countries have the core surveillance and response capacities needed to fulfill the international reporting requirements.¹³

All of these aforementioned surveillance programs focused on identifying

6. Much of this Part appeared in substantially similar form in MICHAEL A. STOTO, PUBLIC HEALTH SURVEILLANCE: A HISTORICAL REVIEW WITH A FOCUS ON HIV/AIDS 2–6 (2003), available at <http://www.rand.org/pubs/drafts/2005/DRU3074.pdf>.

7. See Stephen B. Thacker, *Historical Development*, in PRINCIPLES AND PRACTICE OF PUBLIC HEALTH SURVEILLANCE 1, 3 (Steven M. Teutsch & R. Elliott Churchill eds., 2d ed. 2000).

8. See *id.* at 4.

9. See *id.*

10. See *id.* at 3.

11. See *id.*

12. WORLD HEALTH ORG. [WHO], *Revision of the International Health Regulations*, in FIFTY-EIGHTH WORLD HEALTH ASSEMBLY RESOLUTIONS AND DECISIONS 14, 17 (2005), available at http://www.who.int/gb/ebwha/pdf_files/WHA58/WHA58_3-en.pdf.

13. See *id.* at 16, 47–49.

individuals with infectious diseases and taking action amongst the identified individuals to control disease outbreaks. Smallpox, for example, was eradicated in the 1970s using a surveillance-based approach after earlier efforts to immunize nearly 100% of the population failed because of logistical difficulties.¹⁴ The successful “ring strategy” that led to the success of the smallpox eradication campaign relied on intensive surveillance to identify cases, isolate all known cases, and immunize individuals who may have come in contact with cases.¹⁵ Control strategies traditionally include monitoring, contact tracing, treatment, and quarantine¹⁶—indeed, before the development and widespread availability of antibiotics and vaccines in the twentieth century, this is most of what public health and medicine could do. Even after the advent of antibiotics, contact tracing helped to quell re-emerging tuberculosis (TB) in the United States in the 1990s and is still a common and effective public health tool.¹⁷

Although these days there are few cases of “pestilential diseases,” the need for quick action to prevent the spread of infectious diseases remains. One of the main goals of surveillance for diseases such as TB and sexually transmitted diseases (STDs) is to identify infectious individuals before they infect others, thus preventing an exponentially growing epidemic. Case surveillance has received additional prominence with the increasing interest in emerging infections and, since the attacks of September 11, 2001, in bioterrorism.¹⁸ Indeed, case surveillance was a critical tool in controlling SARS in 2003.¹⁹

However, despite its successes, case surveillance may be causing more harm than benefit in some cases. For instance, screening before and during pregnancy and after birth for phenylketonuria, sickle cell disease, neural tube defects, substance abuse, and HIV-infection has been especially problematic.²⁰ The screening methods may be reliable, but effective, acceptable, or affordable means of following up on identified cases are not widely available.²¹ Especially

14. See, e.g., IAN GLYNN & JENIFER GLYNN, *THE LIFE AND DEATH OF SMALLPOX 200–01* (2004) (crediting the successful eradication of smallpox to the combination of vaccination and “surveillance-containment,” also known as “ring strategy”).

15. See *id.* at 201.

16. See Thacker, *supra* note 7.

17. See INST. OF MED., *ENDING NEGLECT: THE ELIMINATION OF TUBERCULOSIS IN THE UNITED STATES 3* (2000) (discussing the resurgence of tuberculosis in the late 1980s and early 1990s, and claiming that “[t]he highest priority is the identification and treatment of infected contacts of individuals with . . . tuberculosis”); INST. OF MED., *MICROBIAL THREATS TO HEALTH: EMERGENCE, DETECTION, AND RESPONSE 159–70* (Mark S. Smolinski et al. eds., 2003) (advocating the continued use and development of surveillance strategies in order to combat future outbreaks).

18. See Mandl et al., *supra* note 4, at 142.

19. See David L. Heymann & Guénaël Rodier, *Global Surveillance, National Surveillance, and SARS*, 10 *EMERGING INFECTIOUS DISEASES* 173, 174 (2004) (describing the case surveillance measures taken by the WHO Global Influenza Surveillance Network and the Global Outbreak Alert and Response Network in Southeast Asia and Canada that led to the identification of SARS and subsequent recommendations for national and international response).

20. See *REDUCING THE ODDS: PREVENTING PERINATAL TRANSMISSION OF HIV IN THE UNITED STATES 25–35* (Michael A. Stoto, Donna A. Almaro & Marie C. McCormick eds., 1999).

21. See *id.*

when dealing with vulnerable populations, the stigma and discrimination often associated with being identified as having one of these conditions can overwhelm any potential benefit of the surveillance.²² Furthermore, whether case finding can control disease outbreaks depends on the epidemiological dynamics of the condition—for example, whether cases are infectious before they become symptomatic, as with influenza, or afterwards, as with SARS—and the length of latency period.²³ For many diseases, case surveillance is simply not an effective strategy.

Over the course of the twentieth century, the primary cause of death shifted from infectious to chronic diseases; as a result the focus of surveillance shifted to populations rather than individuals.²⁴ Monitoring populations required statistical analysis of data from birth and death certificates, as well as health surveys based on scientifically chosen sample surveys, such as the National Health Interview Survey (NHIS)²⁵ and Behavioral Risk Factor Surveillance System (BRFSS).²⁶

Registries are another source of data for statistical surveillance programs. The National Cancer Institute (NCI) Surveillance, Epidemiology, and End Results (SEER) system, which operates fifteen population-based cancer registries covering approximately 26% of the U.S. population, uses active surveillance methods to record all incident cases of cancer as well as their treatments and outcomes.²⁷ As a result, NCI is able to estimate cancer incidence and survival rates,²⁸ something that is not possible for most chronic diseases.

Surveillance of occupational morbidity and mortality—developed in concert with new regulations on workplace safety regulation—and injury surveillance became more common in the 1990s as public health turned its attention to intentional and unintentional violence.²⁹ A growing focus on health care quality in the early twenty-first century and attendant concerns about medical errors and iatrogenic injuries in recent years have led to intensified surveillance efforts, along with post-marketing surveillance for adverse effects of drugs and vaccines.³⁰

22. *See id.* at 34–35.

23. *See* Christophe Fraser et al., *Factors That Make an Infectious Disease Outbreak Controllable*, 101 *PROC. NAT'L ACAD. SCI.* 6146, 6146 (2004).

24. *See* STOTO, *supra* note 6, at 5–6.

25. For more information, see National Health Interview Survey, <http://www.cdc.gov/nchs/nhis.htm> (last visited Aug. 23, 2007).

26. For more information, see BRFSS: Turning Information into Health, <http://www.cdc.gov/brfss> (last visited Aug. 23, 2007).

27. *See* National Cancer Institute, SEER: Surveillance, Epidemiology, and End Results Program 2, 7, http://seer.cancer.gov/about/SEER_brochure.pdf (last visited Oct. 13, 2007).

28. *See id.* at 7.

29. *See, e.g.*, INST. OF MED., *REDUCING THE BURDEN OF INJURY: ADVANCING PREVENTION AND TREATMENT* 23, 64–71 (Richard J. Bonnie et al. eds., 1999) (identifying injury as an important public health problem and outlining the major injury surveillance systems).

30. *See* INST. OF MED., *THE FUTURE OF DRUG SAFETY: PROMOTING AND PROTECTING THE HEALTH OF THE PUBLIC* 114–15 (Alina Baciu et al. eds., 2007) (recommending that the FDA “develop and implement

II. NEW APPROACHES TO DISEASE SURVEILLANCE THAT DEMAND A REEVALUATION OF THE BALANCE BETWEEN PROTECTING INDIVIDUALS' RIGHTS AND MAINTAINING THE EFFECTIVENESS OF DISEASE CONTROL PROGRAMS

When case-based surveillance was established as a public health function, the inherent loss of privacy was easily justified in terms of the benefits to those identified and, especially, the population at large. And for statistical approaches to surveillance, individuals' privacy was protected by releasing only aggregate numbers such as averages or proportions, as well as by suppressing small cells—that is, table entries representing fewer than five individuals.³¹ Three recent examples, however, have combined both individual and statistical approaches, upsetting the careful balance between the usefulness of the statistical and epidemiological information and the importance of individuals' privacy and confidentiality rights. To understand the new tradeoffs that must be addressed, it is important to clarify the issues involved in these three examples.

A. HIV REPORTING: A CASE-BASED APPROACH WITH A STATISTICAL PURPOSE³²

Although it is a communicable disease, AIDS cannot easily be controlled by reporting individuals with HIV infection to health departments. By the time the infection becomes apparent, years may have passed in which the individual has already infected many others. Rather, the primary reason for requiring HIV case reporting is statistical, specifically to prepare estimates of the prevalence of the condition to guide the allocation of federal resources.³³ This example raises two issues. The first issue is whether the loss of privacy involved in reporting someone who has HIV infection to public health authorities is justified by the public health benefits. The second issue is whether the statistical estimates derived from this surveillance system are accurate, and thus effective in achieving public health goals. If not, the justification for the loss of individual privacy is further undermined.

Since the beginning of the HIV/AIDS epidemic, surveillance efforts have been critical in determining the number and characteristics of individuals diagnosed with AIDS.³⁴ The current AIDS national surveillance system was

active surveillance of specific drugs and diseases as needed"); INST. OF MED., *TO ERR IS HUMAN: BUILDING A SAFER HEALTH SYSTEM* 9 (Linda T. Kohn et al. eds., 2000) (focusing on human error in medical care and recommending, among other things, "[a] nationwide mandatory reporting system that provides for the collection of standardized information"). See generally INST. OF MED., *CROSSING THE QUALITY CHASM: A NEW HEALTH SYSTEM FOR THE 21ST CENTURY* (2001) (focusing on failures of the American health care system and outlining recommended changes).

31. See Michael A. Stoto, *Statistical Issues in Interactive Web-based Public Health Data Dissemination Systems* 21, 32–33 (RAND Health, Working Paper No. WR-106, 2003), available at http://www.rand.org/pubs/working_papers/2005/RAND_WR106.pdf.

32. Much of this Part originally appeared in STOTO, *supra* note 6, at 21–25.

33. See INST. OF MED., *supra* note 3, at 12.

34. See Lawrence O. Gostin, John W. Ward & A. Cornelius Baker, *National HIV Case Reporting for the United States: A Defining Moment in the History of the Epidemic*, 337 *NEW ENG. J. MED.* 1162, 1162 (1997).

implemented prior to the identification of HIV as the etiologic agent of AIDS and the development of an antibody test to determine HIV infection.³⁵ Each state requires that all patients diagnosed with AIDS be reported by name to their local, state, or territorial health department.³⁶ These reports are then forwarded (without names but with unique identifiers) to the CDC, where a national surveillance database is updated and analyzed, providing uniform data on trends and distribution of individuals diagnosed with AIDS.³⁷ Standard records for each case include information on age at diagnosis, sex, race and ethnicity, state of residence (and metropolitan area, if relevant), mode of exposure to HIV, month of AIDS diagnosis, date reported, and other information.³⁸ Statistical analysis of these surveillance data established, for instance, that HIV was transmitted sexually and through blood products and identified a series of risk factors (homosexual sex, multiple partners, intravenous drug use, and so on) that were useful in developing early prevention strategies.³⁹

The AIDS surveillance system evolved over time by changing its case definition to reflect the growing clinical understanding of the disease and the availability of appropriate clinical diagnostic tests while maintaining its focus on AIDS rather than the underlying HIV infection.⁴⁰ The basic reporting responsibilities and procedures, however, remain unchanged. And until the development of potent antiretroviral therapies in the 1990s, AIDS case reporting, although imperfect, provided a relatively accurate picture of trends in HIV infection, especially relative prevalence of HIV in groups defined by geography, race and ethnicity, and primary mode of infection.⁴¹ Estimates of HIV incidence and prevalence were made by statistical techniques such as calculating backward from reported AIDS cases according to well-established patterns of disease progression.⁴² Since the 1990s, however, developments in therapy for HIV and AIDS have decoupled HIV infection and its progression to AIDS.⁴³ As a result, the timing of the progression from HIV infection to AIDS and from AIDS to death is increasingly difficult to predict, making HIV incidence and prevalence estimates based on AIDS cases much less accurate, and AIDS case reports are no longer adequate to monitor trends in new HIV infections.⁴⁴

HIV case surveillance—not to be confused with AIDS case surveillance—

35. *See id.* at 1162–63 (examining the history of HIV/AIDS surveillance).

36. *See id.* at 1163.

37. *Id.*

38. *Id.*

39. *See* INST. OF MED., *supra* note 3, at 1.

40. *See* Ctrs. for Disease Control and Prevention, *Guidelines for National Human Immunodeficiency Virus Case Surveillance, Including Monitoring for Human Immunodeficiency Virus Infection and Acquired Immunodeficiency Syndrome*, 48 MORBIDITY & MORTALITY WKLY. REP., Dec. 10, 1999, at 1–2 [hereinafter CDC, *Guidelines*].

41. *See id.*

42. *See* RON BROOKMEYER & MITCHELL H. GAIL, AIDS EPIDEMIOLOGY: A QUANTITATIVE APPROACH (1994).

43. *See* CDC, *Guidelines*, *supra* note 40, at 1–2.

44. *See id.*

started in some states in the 1980s for contact tracing and to link people to care. This approach was not common at that time, however, since no effective treatment was available and also due to the difficulties with protecting the privacy and confidentiality of those with HIV. In response to concerns about the limitations of the current AIDS surveillance system in providing accurate information about trends in the HIV epidemic, CDC now recommends that all states and territories extend their AIDS surveillance activities to include case reporting of HIV infection.⁴⁵ As of February 2007, forty-seven states, the District of Columbia, and five U.S. dependent areas had implemented HIV case surveillance using the same confidential system for name-based case reporting currently used for AIDS cases.⁴⁶

HIV case reporting is said to have a number of benefits relative to AIDS case reporting. In its official guidance on HIV surveillance, the CDC maintains that HIV case reporting will produce “a more realistic and useful estimate of the resources needed for patient care and services than does AIDS prevalence alone.” Accounts of HIV case reporting in the popular press sometimes suggest that such a system will identify a larger number of infected individuals, and thus lead to greater federal funding for states who adopt such a system. The possibility that federal treatment funds might be allocated according to the number of individuals living with HIV rather than AIDS was brought to attention during Congressional debate about the Ryan White Care Act reauthorization in 2000, leading to great concern in areas such as San Francisco with “mature” epidemics.⁴⁷

Data from existing HIV case reporting systems, however, are incomplete in several important ways. In contrast to the AIDS case reporting system, which is relatively complete, the HIV reporting system collects data only from persons who choose to be tested and who do so at a non-anonymous testing site (i.e., where the HIV test result is linked with identifying information, including patient and provider names). Thus, HIV case reporting data exclude individuals who are infected but have not been tested as well as those who utilize anonymous testing sites or home collection test kits. Because of this selectivity, HIV case reporting by name is unrepresentative of the larger population of infected persons. Further, because reported HIV cases could represent infections that are anywhere from a few weeks to a few years old, the data would reflect the time that individuals chose to be tested rather than when the individual became infected. As a result, HIV case reporting data provide only partial information about HIV prevalence, rather than information about HIV incidence, that is, new HIV infections.⁴⁸

45. See CDC, *Guidelines*, *supra* note 40, at 3.

46. See CDC, HIV Infection Reporting, *available at* <http://www.cdc.gov/hiv/topics/surveillance/reporting.htm>.

47. See INST. OF MED., *supra* note 3, at 8–9.

48. INST. OF MED., NO TIME TO LOSE: GETTING MORE FROM HIV PREVENTION 17–18 (Monica S. Ruiz et al. eds., 2001); see also Mira Johri et al., *New Approaches to HIV Surveillance: Means and Ends: A*

In its report *No Time To Lose*, the Institute of Medicine

concluded that a new surveillance system focused on HIV incidence is needed in order to more effectively guide HIV prevention planning, resource allocation, and evaluation decisions at the national, state, and local levels. To the extent possible, the system would provide estimates at the state and local level and for the population groups at highest risk for HIV infection.⁴⁹

In particular, the Institute recommended that CDC

create a surveillance system that can provide national population-based estimates of HIV incidence. The recommended surveillance system would estimate new HIV infections using blinded [blood samples collected for other purposes from] well-characterized sentinel populations (e.g., drug users in treatment, people attending sexually transmitted disease clinics and tuberculosis clinics, clinics serving women of reproductive age), surveys that characterize the populations served by those sites, and advanced testing technologies that are able to identify recent HIV infections.⁵⁰

In conclusion, the value of the additional information that reporting individually identified HIV cases might provide either for the individual or in terms of more accurate statistical data or funding allocations is less than some would anticipate. In this context, the loss of privacy and confidentiality in reporting individual HIV cases to health departments may not be justified.

B. SYNDROMIC SURVEILLANCE: COLLECTING INDIVIDUAL-LEVEL DATA TO DETECT DISEASE OUTBREAKS⁵¹

Heightened awareness of the risks of bioterrorism since the September 11th attacks, coupled with a growing concern about naturally emerging and re-emerging diseases such as West Nile, SARS, and pandemic influenza, have led public health policymakers to realize the need for early warning systems. The sooner health officials know about an attack or a natural disease outbreak, for instance, the sooner that they can treat those who have already been exposed to the pathogen to minimize the health consequences, vaccinate some or all of the population to prevent further infection, and identify and isolate cases to prevent

Summary Report of the Law, Policy and Ethics Conference on HIV Surveillance 2–3 (CIRA Working paper, Vol. II, Issue 2, 1998), available at http://cira.med.yale.edu/law_policy_ethics/newapproaches.doc.

49. INST. OF MED., *supra* note 48, at 21.

50. *Id.* at 22.

51. Much of this section appears in substantially similar form in Michael A. Stoto, *Syndromic Surveillance in Public Health Practice*, in *INFECTIOUS DISEASE SURVEILLANCE AND DETECTION* 63 (2007). Reprinted with permission from the National Academies Press, Copyright 2007, National Academy of Sciences.

further transmission.⁵² Responding to this need, many health departments have developed “syndromic surveillance” systems, in which individually identified hospital emergency room records are analyzed statistically to detect possible disease outbreaks and covert bioterrorist attacks.⁵³

Syndromic surveillance, however, requires public health agencies to acquire large amounts of routine, individually identified health data before there is any indication of a disease outbreak. If an outbreak were known to be occurring there would be no argument about the need for these data, but most of the time, of course, there is no outbreak. The problem is that the legal structures that balance public health requirements with the protection of privacy and confidentiality do not contemplate surveillance systems that can be justified only in retrospect, that is, if they detect an outbreak. To understand the tradeoffs, it is important to consider both the likely efficacy of syndromic surveillance and the privacy and confidentiality risks involved.

Traditional public health surveillance programs monitor disease using pre-specified case definitions and employ manual data collection, human decision making, and manual data entry.⁵⁴ In contrast, current electronic surveillance systems employ sophisticated information technology and statistical methods to gather and process large amounts of data and display the information for decision makers in a timely way. For instance, syndromic surveillance systems assume that during an attack or a disease outbreak, people will first develop symptoms, then stay home from work or school, attempt to self-treat with over-the-counter (OTC) products, and eventually see a physician with nonspecific symptoms, all days before they are formally diagnosed and reported to the health department.⁵⁵ To identify such behaviors, syndromic surveillance systems regularly monitor existing data for sudden changes or anomalies that might signal a disease outbreak.⁵⁶ Syndromic surveillance systems have been developed to include data on school and work absenteeism, sales of OTC products, calls to nurse hotlines, and counts of hospital emergency room (ER) admissions or reports from primary physicians for certain symptoms or complaints.⁵⁷

The possibility “of earlier detection and more rapid response to a bioterrorist event has tremendous intuitive appeal,”⁵⁸ but there are practical concerns about the use of these systems in state and local public health practice. In statistical terms there is a relatively narrow window between what can be detected in the first few days and what is obvious.⁵⁹ As a result, the statistical value of

52. See James W. Buehler et al., *Syndromic Surveillance and Bioterrorism-Related Epidemics*, 9 EMERGING INFECTIOUS DISEASES 1197, 1197–98 (2003).

53. *Id.*

54. See *supra* Part I.

55. See Mandl et al., *supra* note 4, at 143.

56. See *id.* at 142–43.

57. See Stoto, *supra* note 4, at 49–50.

58. *Id.* at 51.

59. *Id.* at 54.

syndromic surveillance for detecting bioterrorist attacks has not yet been demonstrated.⁶⁰ In addition, syndromic surveillance's success "depends on local health departments' ability to respond effectively."⁶¹

When a syndromic surveillance system sounds an alarm, health departments typically wait a day or two to see if the number of cases continues to remain high or if a similar signal is found in other data sources. Doing so, of course, reduces both the timeliness and sensitivity of the original system. If the health department decides that an epidemiological investigation is warranted, it may begin by identifying those who are ill and talking to their physicians. If this does not resolve the matter, additional tests must be ordered and clinical specimens gathered for laboratory analysis. Health departments might also choose to initiate active surveillance by contacting physicians to see if they have seen similar cases.⁶²

Arthur Reingold provocatively noted the challenges confronting syndromic surveillance: characterizing the types of bioterrorist events that it is likely to detect, determining whether it can help identify the population at risk in a more timely way during a bioterrorist event, determining the appropriate response to apparent increases in illnesses signaled by syndromic surveillance, and ultimately demonstrating that it can reduce morbidity and/or mortality following a bioterrorist event.⁶³ Reingold also questioned whether syndromic surveillance will produce useful information about naturally occurring diseases and the importance of identifying the circumstances under which it is likely to strengthen local and state public health departments.⁶⁴

Moreover, since the development and implementation of syndromic surveillance systems began in recent years, success in gaining access to personal health data has been mixed. Varying interpretations of the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule⁶⁵ are at the heart of the problem. Although some argue that the Privacy Rule permits data owners to disclose protected health information to public health authorities,⁶⁶ covered entities⁶⁷

60. *See id.*

61. *Id.* at 51.

62. *Id.*

63. *See* Arthur Reingold, *If Syndromic Surveillance Is the Answer, What Is the Question?*, 1 *BIOSECURITY & BIOTERRORISM: BIODEFENSE STRATEGY, PRAC., & SCI.* 77, 78–81 (2003).

64. *See id.*

65. *See* Standards for Privacy of Individually Identifiable Health Information, 45 C.F.R. §§ 160, 164 (2005); *see also* Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (codified as amended in scattered sections of 26 U.S.C., 29 U.S.C., 42 U.S.C.).

66. *See* 45 C.F.R. § 164.512(b) (2005) (permitting disclosure of protected health information to public health authorities for certain purposes, such as controlling disease and reporting child abuse or neglect); *see also* Marie C. Pollio, *The Inadequacy of HIPAA's Privacy Rule: The Plain Language Notice of Privacy Practices and Patient Understanding*, 60 *N.Y.U. ANN. SURV. AM. L.* 579, 591 (2004) (asserting that the Privacy Rule requires no authorization for uses and disclosures required by law or made as part of public health activities).

cite the rule in refusing to provide data to researchers and health departments.⁶⁸ In addition to HIPAA, a variety of federal, state, and local public health laws enable, restrict, and otherwise influence the ability to share data for public health surveillance purposes.⁶⁹ Concerns about protecting proprietary data also influence data sharing for public health purposes.

While the HIPAA Privacy Rule may allow covered entities to provide data to public health authorities, it does not require them to do so.⁷⁰ As a result, covered entities may feel exposed to liability in lawsuits: because the release is not mandatory, arguably the entity has a choice about releasing the data. Such concerns reflect an apparent misunderstanding of the distinction between HIPAA and disease reporting laws. HIPAA does not provide for disease reporting mandates—such mandates have their basis in state public health laws. Rather, HIPAA privacy rules include exemptions that allow for disease-reporting under state public laws,⁷¹ and these laws typically mandate reporting not only of specific diseases but also of clusters of disease or unusual health events that may indicate a public health threat.⁷²

The crux of the problem is that syndromic surveillance requires the collection of large amounts of data before there is any indication of a disease outbreak. Indeed, the purpose of gathering the data is to identify when such an outbreak may be occurring. Routine data of this sort are generally not covered by existing public health reporting laws, which focus on specific diseases.⁷³ One solution to this ethical and legal dilemma may be found in distinguishing between the need for statistical and individual-level information. In practice, health departments would gather only aggregate data—which is anonymous—for statistical detection of possible events. In addition, an informatics system would allow public health agencies to go back to identified source records if and when evidence of a possible event emerges. At the point when there is evidence of an emerging event of concern, reporting personal information would be justified under existing public health reporting laws. In addition to resolving the legal dilemma, this approach would help to control the costs of false positives in syndromic surveillance, since initial investigations could be performed electronically rather than by sending teams out to the field. While this would require more sophisti-

67. Covered entities include health care plans, health care providers, and health care clearinghouses. 45 C.F.R. § 160.102 (2005).

68. See ASS'N OF ST. & TERRITORIAL HEALTH OFFICIALS, THE IMPACT OF THE HIPAA PRIVACY RULE ON SYNDROMIC SURVEILLANCE 3 (2004), available at http://biotech.law.lsu.edu/cdc/astho/29724_ASTHO.pdf; Daniel Drociuk, J. Gibson & J. Hodge, Jr., *Health Information Privacy and Syndromic Surveillance Systems*, 53 MORBIDITY AND MORTALITY WKLY. REP. (SUPP.) 221, 221–25 (2004), <http://www.cdc.gov/mmwr/preview/mmwrhtml/su5301a40.htm>;

69. See ASS'N OF ST. & TERRITORIAL HEALTH OFFICIALS, *supra* note 68, at 3.

70. See 45 C.F.R. § 164.512(b) (2005).

71. *Id.*

72. See, e.g., ARIZ. REV. STAT. ANN. § 36-785 (2007); 20 ILL. COMP. STAT. 2305/2.1 (2007); IOWA CODE § 135.145 (2007).

73. See, e.g., ARIZ. REV. STAT. ANN. § 36-785 (2007); 20 ILL. COMP. STAT. 2305/2.1 (2007); IOWA CODE § 135.145 (2007).

cated information technology than current public health reporting systems typically possess, it provides a way to balance public health needs and individual rights to privacy and confidentiality.

C. SCREENING FOR DIABETES AND OBESITY: CASE REPORTING APPLIED TO
NON-COMMUNICABLE DISEASES

Although the focus of public health surveillance was originally on infectious diseases, population-level chronic disease surveillance has a long history. The analysis of vital statistics by cause of death was pioneered in the nineteenth century by William Farr in England and Lemuel Shattuck in the United States.⁷⁴ And as indicated above, population-based surveys such as the National Health Interview Survey (NHIS) and the Behavioral Risk Factor Surveillance System (BRFSS) and registries such as the NCI's SEER system for cancer surveillance are long established. This sort of surveillance system provides statistical data for the entire population as well as groups defined by demography, socio-economic status, geography, and other factors.⁷⁵

Recent developments in chronic disease surveillance, however, have focused on individuals rather than populations. Rather than identifying trends and differentials between increasingly fine-grained populations, screening efforts seek to identify individuals with undetected chronic diseases such as diabetes.⁷⁶ The anonymity and confidentiality traditionally used in the collection of statistical data on chronic diseases is no longer possible, raising questions about whether the benefits to the individuals concerned and to public health generally—which depend on the reliability of the screening programs and the interventions that follow—justify the loss of their privacy.

New York City's health commissioner Thomas Frieden, for instance, wrote that “[l]ocal health departments do a good job of monitoring and controlling conditions,” such as infectious diseases, “that killed people 100 years ago,” but have not kept pace with the epidemiologic transition to the non-communicable diseases which are now responsible for 80% of deaths in the United States.⁷⁷ One way to address non-communicable diseases, as Frieden notes, is to establish disease registries such as those already in place for tuberculosis and cancer.⁷⁸ Registries can both improve management of patients and track the effectiveness of community-based interventions.⁷⁹ New York has also conducted telephone surveys gathering information at the local level about diagnosed diabetes, self-reported obesity and other risk factors, as well as a health and nutrition examination survey to gather actual physical and laboratory

74. See Thacker, *supra* note 7, at 4.

75. See *supra* notes 29–30 and accompanying text.

76. See Stoto, *supra* note 6, at 5.

77. Frieden, *supra* note 5, at 2059.

78. See *id.*

79. See *id.*

measurements.⁸⁰

In December 2005, the New York City Board of Health adopted a diabetes surveillance program that includes mandatory reporting of glycosylated hemoglobin to a registry established by the city's Department of Health and Mental Hygiene.⁸¹ Laboratories are required to report glycosylated hemoglobin levels, a measure of the degree to which an individual's diabetes is under control, along with the identity of the patient and the physician who ordered the test.⁸² The resulting information "will be used to map the epidemiology of diabetes and to monitor the epidemic."⁸³ The registry is also intended to help improve the treatment of individual cases by providing physicians with lists of patients with poor glycemic control, treatment recommendations, and suggestions about advising patients regarding diabetes management.⁸⁴ In the future, patients may receive a letter directly from the health department if their glycosylated hemoglobin value is above a level to be determined.⁸⁵ Patients will be allowed to opt out, but their data will remain as part of the registry.⁸⁶

Although "the endeavor has aroused concern about patients' privacy and . . . the role of health departments," Dr. Frieden responds that the surveillance program aims "to respond to an epidemic of a chronic disease with the type of surveillance and other tools that health departments routinely use to prevent and control communicable diseases."⁸⁷ Three other potential problems suggested by the HIV surveillance example discussed above have apparently not been addressed. First, the registry will only contain information on those who have been tested, and will contain no information on individuals with undiagnosed diabetes. Second, undoubtedly some of those who are tested will not have a regular source of health care, so the link to the physician who ordered the test will be meaningless. Finally, it seems likely that the registry will contain duplicative reports on the same individuals, perhaps linked to different physicians or using different patient identifiers.

Other programs attempt to identify individuals with an elevated risk of developing such diseases, such as obese and overweight children. For instance, as a result of a state law passed in 2003 ("Act 1220"), public schools in Arkansas measure students' body mass index (BMI), and on this basis send annual confidential reports to parents of children who are obese or at risk of obesity.⁸⁸ Act 1220 also mandates improved access to healthier foods in schools

80. *See id.*

81. *See* Robert Steinbrook, *Facing the New Diabetes Epidemic—Mandatory Reporting of Glycosylated Hemoglobin Values in New York City*, 354 *NEW ENG. J. MED.* 545, 545 (2006).

82. *See id.* at 546.

83. *Id.* at 546.

84. *See* Lawrence O. Gostin, *Law as a Tool To Facilitate Healthier Lifestyles and Prevent Obesity*, 297 *JAMA* 87, 88 (2007).

85. *See* Steinbrook, *supra* note 81, at 547.

86. *See* Gostin, *supra* note 84, at 88; Steinbrook, *supra* note 81, at 547.

87. Steinbrook, *supra* note 81, at 545.

88. *See* *ARK. CODE ANN.* §§ 20-7-134 to -135 (2007).

and the creation of local committees to promote physical activity and nutrition.⁸⁹

The rationale for such programs is that early detection of disease and interventions to change behavioral risk factors can prevent severe consequences later in life. The ethical question is whether the health benefits of this screening—to the individual and to society—justify the violation of privacy and confidentiality. The magnitude of the health benefits depends on the efficacy of early detection and intervention. Although well established in theory, empirical evidence about the efficacy of programs that track individuals with diabetes, obesity, and other risk factors is lacking in practice. For instance, an evaluation three years after the Arkansas program began found that the major policy changes at the district and school levels regarding food availability and physical activity, as well as the BMI measurement and reporting process, have been accepted by schools, parents, and children.⁹⁰ There is little evidence so far, however, about changes in diet and activity patterns at home, or obesity levels.⁹¹

CONCLUSION

A basic principle of public health is to employ the least restrictive option that achieves population health goals. From this perspective, surveillance programs must strike a balance between the potential population benefits of the statistical and epidemiological data and knowledge, on the one hand, and protecting individuals' privacy and confidentiality rights on the other. The examples discussed here suggest four considerations that can help find an appropriate balance.

First, policy makers must evaluate whether the proposed public health intervention is likely to achieve its public health goals. If not, the challenges to privacy and confidentiality need not be addressed. As illustrated in the examples above, public health interventions are sometimes not as effective as promised.

Second, policy makers should clearly determine the public health need for individual rather than aggregate statistical data, and opt for the latter if it serves public health needs. The HIV case reporting examples indicate, for instance, that individual-level data may not be needed for public health purposes. HIV case reports do not, as the analysis shows, actually provide more accurate data on HIV incidence and prevalence than existing alternatives.⁹²

Third, policy makers should consider intermediate solutions. Individual-level syndromic surveillance data are not needed until aggregate data suggest a disease outbreak may be underway. Developing the information technology to

89. *See id.* § 20-7-135.

90. UNIV. OF ARK. FOR MED. SCI., FAY W. BOOZMAN COLL. OF PUB. HEALTH, *supra* note 5, at 27.

91. *See id.* at 28.

92. *See* INST. OF MED., *supra* note 54, at 19–21 (“[A] system of population-based HIV incidence estimation will provide the most accurate and timely data for these objectives.”).

gather aggregate data and then link back to the records that generated an alert when necessary would be preferable. With respect to diabetes and its risk factors, it may be more effective to identify the population groups that are highly affected and target risk reduction and treatment strategies at them, rather than at individuals. Identifying schools with a high proportion of overweight children and enhancing their physical education programs and altering their food availability may be more effective and less stigmatizing in the long run than notifying the parents of children who are overweight.

Fourth, policy makers should clarify the circumstances under which population goals may override protection of individuals' privacy and confidentiality rights. Although no single factor is ever definitive, the following should be considered:

The extent to which the disease in question is transmissible from person to person. This consideration would argue in favor of surveillance for HIV infection and syndromic surveillance for bioterrorism or pandemic influenza, but not for obesity or diabetes screening.

Availability of reliable screening methods. One of the problems with identifying overweight children is that, while obesity is certainly a risk factor for diabetes later in life, it is not clear that childhood obesity is a good predictor of adult diabetes risk.

The extent to which identification and reporting of cases benefits those with the condition. A minimum requirement here is the availability of an effective treatment. The development of new and effective HIV medications in the 1990s, for instance, altered the balance of benefits and risks associated with HIV testing. It is less clear that reporting individuals with HIV infection to the local health department actually leads to better access to needed care.⁹³ With respect to diabetes and obesity, losing weight does reduce the risk of diabetes, but does identifying obese children lead them to lose weight and maintain the loss long enough to significantly reduce the risk of diabetes?

The extent to which identification and reporting of cases is effective in controlling the outbreak in the population. Epidemiological dynamics—whether cases are infectious before or after they become symptomatic, the length of a pathogen's latency period and similar factors—are important. Case surveillance was able to help control the SARS outbreak in Toronto and Asia in 2003 because those infected developed symptoms before they were contagious.⁹⁴ The opposite is true of influenza—people may infect others a day or two before they develop symptoms themselves—making case surveillance less effective.⁹⁵ Obesity is clearly not contagious in the usual sense, but a population dynamic is possible: if fitness becomes the norm, individuals may lose weight because

93. See *id.* at 18 (“[C]ontact with a health department after testing HIV-infected was not associated with receipt of timelier care.”).

94. See Fraser et al., *supra* note 23, at 6146.

95. *Id.* at 6150.

obesity is socially unacceptable.

Vulnerability of the affected population. Stigmatization and discrimination against those identified through a screening program can be worse when those identified are part of a vulnerable population, and this should be considered when considering a surveillance program that identifies individuals. For example, one could ask whether, if statistics reveal that a disproportionate number of obese children are African American or Hispanic, these groups are harmed.

Ultimately, surveillance is a double-edged sword. The information from surveillance programs is critically needed to inform and guide public health policy and manage public health programs. However, all surveillance data derive from individuals' personal health information, meaning that their privacy and confidentiality are at risk. Therefore, before a surveillance program is initiated, a careful, case-by-case analysis balancing the benefits of the information for public health purposes and the rights of the individuals who are the subjects of the data is needed.