

# Identifying Challenges to the Integration of Computer-Based Surveillance Information Systems in a Large City Health Department: A Case Study

JACKY M. JENNINGS, PhD, MPH<sup>a,b</sup>

JEFFREY A. STOVER, MPH<sup>c,d</sup>

MEGAN H. BAIR-MERRITT, MD,  
MSCE<sup>a</sup>

CAROLINE FICHTENBERG, PhD<sup>b,e</sup>

MARY GRACE MUNOZ, MPH<sup>c</sup>

RAFIQ MAZIAD, MD, MPH<sup>c</sup>

SHERRY JOHNSON KETEMEPI,  
MPH<sup>c</sup>

JONATHAN ZENILMAN, MD<sup>f</sup>

## SYNOPSIS

**Objective.** Integrated infectious disease surveillance information systems have the potential to provide important new surveillance capacities and business efficiencies for local health departments. We conducted a case study at a large city health department of the primary computer-based infectious disease surveillance information systems during a 10-year period to identify the major challenges for information integration across the systems.

**Methods.** The assessment included key informant interviews and evaluations of the computer-based surveillance information systems used for acute communicable diseases, human immunodeficiency virus/acquired immunodeficiency syndrome, sexually transmitted diseases, and tuberculosis. Assessments were conducted in 1998 with a follow-up in 2008. Assessments specifically identified and described the primary computer-based surveillance information system, any duplicative information systems, and selected variables collected.

**Results.** Persistent challenges to information integration across the information systems included the existence of duplicative data systems, differences in the variables used to collect similar information, and differences in basic architecture.

**Conclusions.** The assessments identified a number of challenges for information integration across the infectious disease surveillance information systems at this city health department. The results suggest that local disease control programs use computer-based surveillance information systems that were not designed for data integration. To the extent that integration provides important new surveillance capacities and business efficiencies, we recommend that patient-centric information systems be designed that provide all the epidemiologic, clinical, and research needs in one system. In addition, the systems should include a standard system of elements and fields across similar surveillance systems.

---

<sup>a</sup>Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>b</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>c</sup>Office of Epidemiology, Division of Disease Prevention, Virginia Department of Health, Richmond, VA

<sup>d</sup>Department of Epidemiology and Community Health, School of Medicine, Virginia Commonwealth University, Richmond, VA

<sup>e</sup>Baltimore City Health Department, Baltimore, MD

<sup>f</sup>Division of Infectious Disease, Johns Hopkins University School of Medicine, Baltimore, MD

Address correspondence to: Jacky M. Jennings, PhD, MPH, Johns Hopkins Bayview Medical Center, 5200 Eastern Ave., Mason F. Lord Bldg., Center Towers, Ste. 4200, Baltimore, MD 21224; tel. 410-550-4132; fax 410-550-4153; e-mail: <jjennin1@jhmi.edu>.

©2009 Association of Schools of Public Health

Integrated or linked infectious disease surveillance information systems would provide important new surveillance capacities and business efficiencies.<sup>1</sup> Integrated systems would permit, for example, the monitoring of comorbidities, targeting of scarce public health resources for comorbid populations, and limiting of missed opportunities for tracking comorbid individuals. Integrated systems may also decrease duplicative data entry for comorbid individuals and lessen reporting burdens. Given that most surveillance information systems are now computer-based, integrated systems may also provide additional business efficiencies by allowing for the sharing of information technology resources including staff, training, infrastructure, and architecture.

The integration of public health information systems can be achieved in a number of ways. For example, integration may involve enabling linkages between existing information systems or developing a single comprehensive information system that incorporates all the information across different programs. Regardless of the way in which the information systems are integrated, the most critical aspect of integration is the ability to identify the same individual from one information system to the next. This requires a method to link one person's information from one disease control information system to the next, such as linkage by a unique identifier. In the absence of a unique identifier, matching algorithms can be used that rely on a selected set of consistently collected demographic variables such as name, date of birth, social security number, age, and race/ethnicity. The current era of computer-based public health surveillance systems makes the capacity for integration great, given that programs can be written for matching by a unique identifier or matching algorithm. The reality of integration, however, may be different.

The added public health benefit of integrated infectious disease information systems is in part dependent on the extent to which comorbidities exist at the point of integration. At one large city health department, previous work had been conducted suggesting that there was considerable overlap in patient populations among the city's infectious disease surveillance programs including acute communicable diseases (ACDs), human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS), sexually transmitted diseases (STDs), and tuberculosis (TB).<sup>2-4</sup> For example, among African American patients referred to the TB clinic, 13.6% had a history of syphilis and 16.5% had at least one documented visit at a city STD clinic.<sup>3</sup> This study and other work suggest that, at least at this large city health department, integrated infec-

tious disease information systems have the potential to improve the effectiveness and efficiency of public health surveillance.

The idea of integrating public health surveillance information is not new.<sup>1</sup> At the Centers for Disease Control and Prevention (CDC), efforts toward integration began as early as the 1990s, with initiatives aimed at creating common data standards and infrastructure across surveillance information systems. For example, the development of the National Electronic Telecommunications System for Surveillance (NETSS) in 1990 defined a standard case report format and set of variables for disease reporting across programs.<sup>5,6</sup> Subsequently, in 1999 the National Electronic Disease Surveillance System (NEDSS) promoted a more integrated architecture for public health surveillance information by allowing for Internet-based reporting of public health, laboratory, and clinical data.<sup>7</sup> Other integration initiatives followed at CDC, including the Public Health Information Network in 2004 and the creation of the National Center for Public Health Informatics in 2006.<sup>8,9</sup> These initiatives showed an increasing recognition by CDC that surveillance systems across disease control programs shared many common practices and reflected a broader effort to move from stand-alone solutions to networked, integrated solutions.<sup>1,8</sup>

We conducted a case study at a large city health department of the primary computer-based infectious disease surveillance information systems during a 10-year period to identify the major challenges for information integration across the systems.

## METHODS

We assessed the computer-based surveillance information systems in one large city health department among four infectious disease control programs including ACDs, HIV/AIDS, STDs, and TB. The assessments took place in December 1998 and approximately 10 years later in June 2008. The assessments included in-depth interviews with key informants as well as database extraction and review.

### Interviews

We conducted interviews, at a minimum, with the director of each local disease control program. If a different individual managed the surveillance database, that individual was interviewed. A standard interview was used with free-form dialogue to express the exchange of information. The interviews took one to three hours with additional follow-up in person or by telephone to clarify or provide additional information.

The interviewer collected supporting documentation including manuals, data dictionaries, data reporting, and flowcharts.

The interviews first established which and how many computer-based and noncomputer-based surveillance information systems were in use within each of the four disease control programs. Second, the primary computer-based surveillance information system in each program was defined using the following criteria: (1) full name and alias of the surveillance system; (2) number of records captured; (3) development application; (4) description of the contents of the database; (5) time period of the data captured; (6) reporting agencies; and (7) number of duplicative surveillance systems in use—i.e., the number of information systems being used in addition to the primary surveillance system.

#### Database extraction and review

Following the interview in 1998, data were extracted from each of the primary surveillance databases (except for HIV/AIDS data due to confidentiality-related restrictions). The data extraction was conducted to compare a selection of demographic variables and their coding conventions across the systems. The selected variables were those that were frequently utilized in matching algorithms to identify the same individual across surveillance information systems. The data extraction was also conducted to verify or supplement information obtained in the key informant interviews. In 2008, in lieu of the extraction process, we conducted a data review, including a comparison of the demographic variables, during the interview.

## RESULTS

### Background research and interview information

Following are descriptions of the primary computer-based surveillance information systems utilized by the four infectious disease control programs in one city health department in 1998 and 2008. (Figure 1 summarizes these findings.)

**ACDs.** In 1998, the Maryland Electronic Reporting & Surveillance System (MERSS)<sup>10</sup> was the primary computer-based surveillance information system utilized by the city health department's ACD program to gather information and investigate outbreaks of diseases including (but not limited to) hepatitis B, hepatitis C, salmonella, shigella, and Lyme disease. MERSS was a Microsoft® Access-based system<sup>11</sup> designed and implemented in 1998 by the Maryland State Health Department Division of Communicable Disease Surveillance. MERSS, which contained approximately 10,840

records with information from 1989 through 1999, was a case-centric system, meaning each record represented one reported case of disease. The database for the city resided physically at the state health department, and information was remotely entered at the city health department via dial-in access. Because of confidentiality and data security concerns, access to the system for data entry, management, or analysis was restricted to one computer at the city health department. The system was capable of generating reports including line listings and frequencies by event date, and reports were provided regularly to the city and state health departments as well as CDC (via the state).

In April 2006, the primary computer-based surveillance information system transitioned from MERSS to the National Electronic Disease Surveillance System Base System (NBS), a system designed and supported by CDC.<sup>12</sup> In contrast to the case-centric nature of MERSS, the NBS is patient-centric and Internet-based, allowing for data exchange using established CDC and industry data standards.<sup>13</sup> The database remains physically at the state health department and is populated with approximately 46,137 cases including information from 2006 to June 2008. At the 2008 follow-up assessment, MERSS was still being maintained as an archival database to permit the examination of temporal disease-specific trends with data prior to 2006. In addition, the city ACD program maintains a separate Microsoft® Excel<sup>14</sup> database that predates NBS and serves the specific function of tracking all reported ACD outbreaks in the city.

**HIV/AIDS.** In 1998 and 2008, the HIV/AIDS Reporting System (HARS) was the primary computer-based surveillance information system in use by the HIV/AIDS program at this city health department.<sup>15</sup> HARS is a CDC-supported system developed in PRODA.<sup>16</sup> HARS supports the collection of demographic, risk, clinical, and laboratory data on people diagnosed with either HIV infection or AIDS. The identifying information and modules used for the data collection have changed over time for this program. These changes were a result of a switch in 2007 from unique-identifier reporting for HIV to name-based reporting. The database in 1998 resided physically at the city health department; however, the database was moved to the state health department in 2008. The shift reflected a larger change occurring around 2000 when the state (vs. local jurisdictions) began to manage all HIV/AIDS surveillance.

HARS is designed as a patient-centric system, although prior to the name-based reporting of HIV, HIV cases were handled in a case-centric way. In 2008, HARS at this city health department included

**Figure 1. A description of the primary computer-based surveillance information systems utilized by four infectious disease control programs in a large city health department in 1998 and 2008**

Primary surveillance system	Infectious disease control programs							
	ACDs		HIV/AIDS		STDs		TB	
	1998	2008	1998	2008	1998	2008	1998	2008
Primary surveillance system	Maryland Electronic Reporting & Surveillance System <sup>a</sup>	National Electronic Disease Surveillance System <sup>b</sup>	HIV/AIDS Reporting System <sup>c</sup>	HIV/AIDS Reporting System <sup>c</sup>	STD Management Information System <sup>d</sup>	STD Management Information System <sup>d</sup>	Tuberculosis Information Management System <sup>e</sup>	Tuberculosis Information Management System <sup>e</sup>
Alias/acronym	MERSS <sup>a</sup>	NBS <sup>b</sup>	HARS <sup>c</sup>	HARS <sup>c</sup>	STD*MIS <sup>d</sup>	STD*MIS <sup>d</sup>	TIMS <sup>e</sup>	TIMS <sup>e</sup>
Number of records	10,840	46,137	13,882	16,400	7,500	>250,000	NA	1,029
Application	Access <sup>f</sup>	Internet-based	PRODAS <sup>g</sup>	PRODAS <sup>g</sup>	xBase++ <sup>h</sup>	xBase++ <sup>h</sup>	PowerBuilder <sup>i</sup>	PowerBuilder <sup>i</sup>
Number of variables	NA	NA	550	550	650	650	850	850
Description	Case-centric; reported cases of ACDs (hepatitis, salmonella, shigella, Lyme disease, meningitis, etc.)	Patient-centric; reported cases of ACDs (hepatitis, salmonella, shigella, Lyme disease, meningitis, etc.)	Patient-centric; reported individuals with HIV (unique identifier) and/or AIDS (name-based)	Patient-centric; reported individuals with HIV and/or AIDS (name-based)	Case-centric; reported cases of STDs (syphilis only)	Case-centric; reported cases of STDs (chlamydia, gonorrhea, syphilis, and HIV)	Case-centric; reported cases of verified TB	Case-centric; reported cases of verified TB
Database time period	1989–1998	2006–2008	1981–1998	1981–2008	1996–1998	1996–2008	1998	1998–2008

continued on p. 43

**Figure 1 (continued). A description of the primary computer-based surveillance information systems utilized by four infectious disease control programs in a large city health department in 1998 and 2008**

	Infectious disease control programs								
	ACDs		HIV/AIDS		STDs		TB		
	1998	2008	1998	2008	1998	2008	1998	2008	
Regular reports generated for	City, State	City, State, CDC	City, State, CDC	State, CDC	City, State, CDC	City, State, CDC	City, State, CDC	City, State, CDC	City, State, CDC
Number of duplicative surveillance systems	None	Two (Excel <sup>k</sup> and MERSS databases)	None	One (Access <sup>f</sup> database)	Three (clinic [two] and registry databases)	Two (clinic database and registry databases)	One (Epi Info <sup>TM</sup> database)	Two (Excel <sup>k</sup> and Epi Info <sup>TM</sup> databases)	

<sup>a</sup>Maryland Department of Health and Mental Hygiene. Maryland Electronic Reporting & Surveillance System. Baltimore: Maryland Department of Health and Mental Hygiene; 1998.

<sup>b</sup>Centers for Disease Control and Prevention (US). Public Health Information Network. National Electronic Disease Surveillance System: NEDSS Base System (NBS) [cited 2009 Jul 3]. Available from: URL: <http://www.cdc.gov/phn/activities/applications-services/needs/nbs.html>

<sup>c</sup>CDC (US), National Center for HIV, STD, and TB Prevention. HIV/AIDS Reporting System (HARS) [cited 2009 Jul 3]. Available from: URL: <https://team.cdc.gov/team/cdc/dispatch.cgi/w:hars>

<sup>d</sup>CDC (US), Division of STD Prevention. Sexually transmitted diseases: STD\*MIS [cited 2008 Jul 7]. Available from: URL: <http://www.cdc.gov/std/std-mis>

<sup>e</sup>CDC (US), Division of Tuberculosis Elimination. Tuberculosis Information Management System (TIMS) [cited 2008 Jul 7]. Available from: URL: <http://www.cdc.gov/tb/programs/tims.htm>

<sup>f</sup>Microsoft Corp. Access<sup>®</sup> 1997. Redmond (WA): Microsoft Corp.; 1997.

<sup>g</sup>Conceptual Software. Professional Database Analysis System (PRODAS): Version 3.2A. Houston (TX): Conceptual Software; 1987.

<sup>h</sup>Alaska Software, Inc. xBase++ [cited 2009 Jul 3]. Available from: URL: <http://www.alaska-software.com/products/xpp/xpp.shtm>

<sup>i</sup>Sybase. PowerBuilder [cited 2009 Jul 3]. Available from: URL: <http://www.sybase.com/products/modelingdevelopment/powerbuilder>

<sup>j</sup>STD\*MIS as utilized in a case-centric manner at this city health department

<sup>k</sup>Microsoft Corp. Excel<sup>®</sup> 2007 [cited 2009 Jul 3]. Available from: URL: <http://office.microsoft.com/en-us/excel/default.aspx>

<sup>l</sup>CDC (US), Division of Integrated Surveillance Systems and Services (DISSS). Epi Info<sup>TM</sup> [cited 2009 Jul 3]. Available from: URL: <http://www.cdc.gov/epiinfo>

ACD = acute communicable disease

HIV = human immunodeficiency virus

AIDS = acquired immunodeficiency syndrome

STD = sexually transmitted disease

TB = tuberculosis

NA = not applicable

CDC = Centers for Disease Control and Prevention

information on approximately 16,400 AIDS cases from as early as 1981. HIV/AIDS information is collected from physicians, hospitals, laboratories, and death certificates. HARS has the capability to produce statistical analyses and local tabulations for generating reports to the city and state health departments and to CDC. The city health department maintains a duplicative database of HIV and AIDS patients so that it can have immediate local access to data that were otherwise controlled and managed at the state level. At both assessment time points (1998 and 2008), there was no reported use of a module that facilitated linking between CDC's HIV and STD information systems.

**STDs.** In 1998 and 2008, the city STD program utilized STD\*MIS as its primary computer-based surveillance information system to collect demographic, risk, and clinical information on confirmed case reports of STDs (including HIV in some cases).<sup>17</sup> STD\*MIS is a case-centric, CDC-supported application developed in xBase++<sup>18</sup> and provided to state and local health departments. The application is intended to address common data management issues facing STD programs nationally, including the electronic transfer of morbidity data to CDC via NETSS and the management of STD field investigations. At this city health department, STD\*MIS was implemented in 1996 and was initially only used to track confirmed syphilis cases. In 1998, the system contained 7,500 syphilis case records. In 1999, the use of STD\*MIS was expanded to track confirmed cases of chlamydia, gonorrhea, and HIV (for HIV, only when an STD program clinic, outreach activity, or affiliate conducted the testing). By 2008, more than 250,000 cases were in the system. Data were collected from the state health department, public STD clinics, outreach counseling and testing programs, disease intervention and case management specialists, laboratories, and private physicians, and were regularly reported to the city and state health departments and to CDC.

Prior to the use of STD\*MIS and in duplicative use today, the city STD program operates three other computer-based STD surveillance information systems. The information captured by the systems overlaps in part with the STD\*MIS information and with the other systems. One duplicative information system was a locally developed registry. The registry included information on all confirmed cases and negative test results of patients tested for gonorrhea and chlamydia. The registry was fed information from two public health STD clinics, which had two separate information systems, including the registry information as well as all clinic encounter data for all individuals tested for STDs. In 2008, a new application called Insight<sup>19</sup> replaced the clinic-based information systems.

**TB.** The city health department uses the Tuberculosis Information Management System (TIMS) to manage TB cases and to track and report TB-control program activities.<sup>20</sup> TIMS was developed using the computer application development system PowerBuilder.<sup>21</sup> Use of TIMS began in 1998, and it was still in use in 2008 (although TIMS has been retired as of January 1, 2009). TIMS automated the administration of TB prevention, surveillance, and control programs, and provided electronic reporting capability through NETSS. TIMS tracked, reported, and verified cases of TB by name with active follow-up. In 2008, there were 1,029 cases in the system.

The application was a six-module design with the following modules: (1) client, (2) surveillance, (3) patient management, (4) program evaluation, (5) daily program operations, and (6) system. In 1998, only the client module was implemented; by 2008, the surveillance module was also in use. Data for the two modules were collected from clinics, laboratories, pharmacies, and private providers. The city TB program operated two duplicative systems. The first was TIMS loaded onto a separate server to enable the tracking of latent TB cases, and the second was a Microsoft® Access<sup>11</sup> database, which was analyzed using Epi Info™ to permit the tracking and analysis of specific subgroup populations, such as refugees.<sup>22</sup>

#### Database extraction and review information

**Variable data collection and coding structures.** Although most of the primary computer-based surveillance information systems used by the city health department programs collected what appeared to be the same demographic variables, there were many differences in how these variables were collected and coded. The included demographic variables (Figure 2) were selected for comparison because often these are variables used to link individuals among systems.

The collection of name data varied from three fields in NEDSS and TIMS to two fields in STD\*MIS and only one field in HARS. STD\*MIS differed from the other three systems by not collecting the middle initial and by collecting nicknames. The collection of date of birth data showed similarity among the systems except for HARS, which, in contrast to the other three systems, collected only a two-digit year. The collection of gender data was also similar among the systems except that STD\*MIS included a category for unknown gender.

The collection of address data was similar for NEDSS, HARS, and STD\*MIS with a few exceptions. Most notably, HARS did not collect street address information. The collection of race and ethnicity data showed differences by program. Three programs—

**Figure 2. A comparison of selected demographic variables across the primary computer-based surveillance information systems utilized by four infectious disease control programs in a large city health department, 2008**

Primary surveillance system	Infectious disease control programs			
	ACDs	HIV/AIDS	STDs	TB
	National Electronic Disease Surveillance System (NEDSS) <sup>a</sup>	HIV/AIDS Reporting System (HARS) <sup>b</sup>	STD Management Information System (STD*MIS) <sup>c</sup>	Tuberculosis Information Management System (TIMS) <sup>d</sup>
Name	Last _____ First _____ MI _____ (three fields)	Last, First, MI _____  (one field)	Last name _____ First name & nicknames _____ (two fields)	Last _____ First _____ MI _____ (three fields)
Date of birth	__ __ __ __ mo day year	__ __ __ __ mo day year	__ __ __ __ mo day year	__ __ __ __ mo day year
Sex	__ M Male __ F Female (two codes)	__ 1 Male __ 2 Female (two codes)	__ 1 Male __ 2 Female __ 9 Unknown (three codes)	__ 1 Male __ 2 Female (two codes)
Address	Address _____ City _____ State _____ County (code) _____ County _____ ZIP code _____ (six fields)	City _____ County _____ State _____ ZIP code _____ (four fields)	Street _____ City _____ County (code) _____ State _____ ZIP code _____ (five fields)	Street _____ City _____ State _____ ZIP code _____ (four fields)
Ethnicity	Y Hispanic N Non-Hispanic U Unknown (three codes)	1 White (not Hispanic) 2 Black (not Hispanic) 3 Hispanic 4 Asian/Pacific Islander 5 American Indian/Alaska Native 9 Not specified	1 Hispanic 2 Non-Hispanic (two codes)	1 Hispanic 2 Non-Hispanic (two codes)
Race	W White B Black O Other U Unknown (four codes)	1 American Indian/Alaska Native 2 Asian/Pacific Islander 3 Black 4 White 9 Other/unknown (five codes)	1 American Indian/Alaska Native 2 Asian/Pacific Islander 3 Black 4 White 9 Other/unknown (five codes)	1 White 2 Black 3 American Indian or Alaska Native 4 Asian or Pacific Islander specify: (optional) _____ (four codes)

continued on p. 46

**Figure 2 (continued). A comparison of selected demographic variables across the primary computer-based surveillance information systems utilized by four infectious disease control programs in a large city health department, 2008**

	Infectious disease control programs			
	ACDs	HIV/AIDS	STDs	TB
Social security number	Not collected	Not collected	-----	-----

<sup>a</sup>Centers for Disease Control and Prevention (US). National Electronic Disease Surveillance System [cited 2009 Jul 3]. Available from: URL: <http://www.cdc.gov/NEDSS>

<sup>b</sup>CDC (US), National Center for HIV, STD, and TB Prevention. HIV/AIDS Reporting System (HARS) [cited 2009 Jul 3]. Available from: URL: <https://team.cdc.gov/team/cdc/dispatch.cgi/w.hars>

<sup>c</sup>CDC (US), Division of STD Prevention. Sexually transmitted diseases: STD\*MIS [cited 2008 Jul 7]. Available from: URL: <http://www.cdc.gov/std/std-mis>

<sup>d</sup>CDC (US), Division of Tuberculosis Elimination. Tuberculosis Information Management System (TIMS) [cited 2008 Jul 7]. Available from: URL: <http://www.cdc.gov/tb/programs/tims/tims.htm>

ACD = acute communicable disease

HIV = human immunodeficiency virus

AIDS = acquired immunodeficiency syndrome

STD = sexually transmitted disease

TB = tuberculosis

MI = middle initial

mo = month



NEDSS, STD\*MIS, and TIMS—all collected ethnicity separately from race, whereas HARS collected ethnicity together with race in one variable. The collection of information about race/ethnicity also differed by program. With regard to racial/ethnic categories, for example, the NEDSS program used four categories of race including white, black, other, and unknown, in contrast to the STD\*MIS program, which used five categories of race including white, black, Asian/Pacific Islander, American Indian/Alaska Native, and other/unknown. In addition, the coding for race/ethnicity differed by program. For example, in the NEDSS program, ethnicity and race were coded as character variables, while the HARS, STD\*MIS, and TIMS programs used numeric coding for race/ethnicity. Two of the four systems collected social security number (SSN) data.

## DISCUSSION

The assessments of the four computer-based primary infectious disease surveillance systems in a large city health department during a 10-year period showed a number of challenges to information integration across the systems. One challenge was the existence of duplicative data systems within each of the programs at each assessment time point. The existence, extent, and persistence of duplicative systems suggest that no one information system fulfills all the needs of the local system. Thus, identifying one primary system for integration would be difficult, but also might have limited utility.

Other common challenges were the differences in the variables used for collecting similar information, such as race/ethnicity, and the inconsistencies in the collection of specific variables across systems, such as SSN. We highlight SSN because it is a unique identifier and, if available and consistently collected, could be used to match individuals across disease control information systems. The variability creates real barriers to de-duplicating individuals among the systems.

The systems also differed in their basic architecture, such as the case- or patient-centric nature of the system. Transforming from a case- to patient-centric system represents a considerable effort, as it first requires the de-duplication of data within the case-centric systems to create patient-centric systems. A case- vs. patient-centric system also represents a fundamentally different approach to surveillance and, thus, may present additional barriers to integration.

Many of the identified barriers are symptomatic of the fact that categorical disease control programs have traditionally used a silo-based approach to disease

surveillance. Although these systems were developed with similar goals—i.e., to facilitate epidemiologic assessment of disease trends and program management for a particular jurisdiction—historically, there was not the funding or full recognition of the importance of data integration.<sup>23,24</sup> Thus, common data standards or elements were not implemented across the systems, resulting in part in program systems with different data-variable coding structures and poor compatibility, and limiting the capacity to address public issues such as comorbidities.<sup>23,25</sup> Therefore, integrating data across systems for patient management, or for analysis, is difficult without substantial investment in recoding variables and programming resources. Furthermore, development of the systems was historically restrained by the technological capacity available at the federal and local level, within the context of a public health environment that has limited resources. Notably, the costs associated with developing and maintaining public health information systems, including IT staff, remain a major challenge.

## CONCLUSIONS

If evidence continues to show that integrated public health surveillance information systems would provide important new surveillance capacities and increase business efficiency, then this case study of the challenges associated with information integration of the computer-based infectious disease surveillance systems in one large city health department highlights a number of recommendations. Programs, such as these four infectious disease control programs, should attempt in cooperation with CDC to design information systems that are patient-centric and provide all the epidemiologic, clinical, and research needs (to the extent possible) in one system. A single system would limit or eliminate the need for duplicative information systems. A standard system of elements and fields should be implemented across similar surveillance systems. The variables should be consistent and identifiable as similar among the systems. In addition, CDC might consider providing sustainable applications to match datasets among surveillance systems, which would allow public health departments to at least conduct comorbidity assessments on a regular basis.

## REFERENCES

1. Centers for Disease Control and Prevention (US), Agency for Toxic Substances and Disease Registry. Integrating public health information and surveillance systems: a report and recommendations from the CDC/ATSDR Steering Committee on Public Health Information and Surveillance System Development. Atlanta: CDC/ATSDR; 1995.

2. Golden MR, Rompalo AM, Fantry L, Bein M, Perkins T, Hoover DR, et al. Early intervention for human immunodeficiency virus in Baltimore sexually transmitted diseases clinics. Impact on gonorrhea incidence in patients infected with HIV. *Sex Transm Dis* 1996;23:370-7.
3. Peek M, Zenilman JM. Sexually transmitted diseases in patients attending a Baltimore tuberculosis clinic. Assessment of use of multiple categoric services. *Sex Transm Dis* 1997;24:8-10.
4. Thomas DL, Zenilman JM, Alter HJ, Shih JW, Galai N, Carella AV, et al. Sexual transmission of hepatitis C virus among patients attending sexually transmitted diseases clinics in Baltimore—an analysis of 309 sex partnerships. *J Infect Dis* 1995;171:768-75.
5. National Electronic Telecommunications System for Surveillance—United States, 1990–1991. *MMWR Morb Mortal Wkly Rep* 1991;40(29):502-3.
6. CDC (US). National Electronic Telecommunications System for Surveillance [cited 2008 Jul 7]. Available from: URL: <http://www.cdc.gov/ncphi/diss/nndss/netss.htm>
7. Progress in improving state and local disease surveillance—United States, 2000–2005. *MMWR Morb Mortal Wkly Rep* 2005;54(33):822-5.
8. Preface: 60 years of public health science at CDC. *MMWR Morb Mortal Wkly Rep* 2006;55(Suppl 2):1.
9. Institute of Medicine. The future of the public's health in the 21st century. Washington: National Academies Press; 2003.
10. Maryland Department of Health and Mental Hygiene. Maryland Electronic Reporting & Surveillance System. Baltimore: Maryland Department of Health and Mental Hygiene; 1998.
11. Microsoft Corp. Microsoft® Access. Redmond (WA): Microsoft Corp.; 1997.
12. CDC (US). National Electronic Disease Surveillance Base System (NBS): Release 1.1.1. Atlanta: CDC; 2003.
13. CDC (US). Public Health Information Network. National Electronic Disease Surveillance System: NEDSS Base System (NBS) [cited 2009 Jul 3]. Available from: URL: <http://www.cdc.gov/phinf/activities/applications-services/nedss/nbs.html>
14. Microsoft Corp. Microsoft® Excel 2007 [cited 2009 Jul 3]. Available from: URL: <http://office.microsoft.com/en-us/excel/default.aspx>
15. CDC (US), National Center for HIV, STD, and TB Prevention. HIV/AIDS Reporting System (HARS) [cited 2009 Jul 3]. Available from: URL: <https://team.cdc.gov/team/cdc/dispatch.cgi/w.hars>
16. Conceptual Software. Professional Database Analysis System (PRO-DAS): Version 3.2A. Houston (TX): Conceptual Software; 1987.
17. Centers for Disease Control and Prevention (US), Division of STD Prevention. Sexually transmitted diseases: STD\*MIS [cited 2008 Jul 7]. Available from: URL: <http://www.cdc.gov/std/std-mis>
18. Alaska Software, Inc. xBase++ [cited 2009 Jul 3]. Available from: URL: <http://www.alaska-software.com/products/xpp/xpp.shtml>
19. Netsmart Technologies. Insight: Version 6.1 [cited 2009 Jul 3]. Available from: URL: [http://www.ntst.com/products/Insight6\\_1.asp?show=yes](http://www.ntst.com/products/Insight6_1.asp?show=yes)
20. CDC (US), Division of Tuberculosis Elimination. Tuberculosis information management system (TIMS) [cited 2008 Jul 7]. Available from: URL: <http://www.cdc.gov/tb/programs/tims/tims.htm>
21. Sybase. PowerBuilder [cited 2009 Jul 3]. Available from: URL: <http://www.sybase.com/products/modelingdevelopment/powerbuilder>
22. CDC (US), Division of Integrated Surveillance Systems and Services (DISSS). Epi Info™ [cited 2009 Jul 3]. Available from: URL: <http://www.cdc.gov/epiinfo>
23. Thacker SB, Stroup DF. Future directions for comprehensive public health surveillance and health information systems in the United States. *Am J Epidemiol* 1994;140:383-97.
24. Thacker SB, Stroup DF. The future of national public health surveillance in the United States. *J Public Health Manag Pract* 1996;2:1-3.
25. Spitalny KC. Learning to design new systems: communicable disease surveillance. *J Public Health Manag Pract* 1996;2:40-1.