

## How Disease Surveillance Systems Can Serve as Practical Building Blocks for a Health Information Infrastructure: the Indiana Experience

Shaun J. Grannis, M.D., M.S.<sup>1,2</sup>, Paul G. Biondich, M.D., M.S.<sup>1,2</sup>, Burke W. Mamlin, M.D.<sup>1,2</sup>  
Greg Wilson, M.D.<sup>2</sup>, Linda Jones, RN, MSN<sup>3</sup>, J. Marc Overhage M.D., Ph.D.<sup>1,2</sup>

<sup>1</sup>The Regenstrief Institute, Inc., <sup>2</sup>Indiana University School of Medicine, <sup>3</sup>Indiana State Department of Health, Indianapolis, IN

### Abstract:

*Although many organizations are beginning to develop strategies to implement and study regional and national health information exchanges, there are few operational examples to date. The Indiana Network for Patient Care (INPC) is an example of a currently operational Regional Health Information Organization (RHIO) built upon a foundation of open, robust healthcare information standards. Having demonstrated the scalability of this design, the Indiana State Department of Health (ISDH) contracted with the Regenstrief Institute to implement a statewide disease surveillance system incorporating encounter data from all 114 Indiana hospitals with emergency departments. We describe the 4-year implementation plan, including our design rationale and how we plan to address the specific implementation challenges of data collection, connectivity in diverse environments and current hospital buy-in. To date, 42 hospitals are in various stages of engagement, with 33 hospitals actively providing real-time surveillance data. We will discuss how this project creates the foundation for a potential statewide health information exchange.*

### Introduction

The federal government has committed to developing a national health information network (NHIN) over the next ten years.<sup>1</sup> Momentum continues to build as many organizations begin to assess the potential of regional health information infrastructures. These organizations include the AHRQ,<sup>2</sup> e-Health Initiative and Foundation,<sup>3</sup> Markle foundation,<sup>4</sup> and the current presidential administration.<sup>5</sup> Further, the Department of Health and Human Services has established the office of the National Coordinator for Health Information Technology.

Part of the government's growing interest in NHIN initiatives is fueled by recent estimates of sizeable health care cost savings. The Center for Information Technology Leadership (CITL) recently projected that health information exchange would save \$78 billion annually.<sup>6</sup> Expected cost savings come from reduced health care information management labor costs, fewer duplicate tests and procedures, reduced fraud and abuse, improved

service delivery efficiency, and reduced medical errors.

Most who work within the healthcare industry envision bright futures with widely implemented electronic health record (EHR) systems, eager in anticipation of specialized functionalities such as physician order entry and intelligent decision support. But, in order for these EHRs to have any chance of success, we believe that substantial energies and resources must *first* focus on building a robust information sharing infrastructure, which includes reliable connectivity, standardized clinical messaging and data repositories. An information sharing infrastructure is an absolute prerequisite to interoperable EHR systems because such mechanisms provide the all-important data that "drives" clinical decision making.<sup>7</sup>

Clinical and public health arenas face similar challenges when considering the development of regional information exchanges. Health care information is scattered across many independent databases and systems as separate data islands with different patient and provider identifiers, concept identifiers, and location identifiers. This is true for data collected *within* an institution and for data collected about the same patient at *different* health care institutions or public health organizations. These pervasive realities create layers of complexity in health care information aggregation efforts for both public health and clinical care uses. On the public health front, the Centers for Disease Control (CDC) is promoting standards and specifications to ensure a consistent and coherent public health information network can be built to serve the nation's public health information needs.<sup>8</sup>

Because public health initiatives have to cross all of the separate silos and other facets of healthcare, public health and clinical medicine have shared interests in a NHIN. Public health is in fact part of the comprehensive health care ecology both from a patient care perspective and from a health care IT perspective. Data generated in typical clinical workflow, such as immunization records and reportable laboratory results, are just two examples of information having dual use in both clinical medicine and public health. Similarly, routinely collected point of care emergency department encounter data can be

of great value to public health syndromic surveillance efforts.

Syndromic (statistical) surveillance focuses on the use of early indicators of disease to identify outbreaks before definitive diagnoses are made. Results from several studies demonstrate that data from emergency department encounters, hospital admissions, and retail pharmaceutical sales can signal the onset and evolution of disease outbreaks earlier than traditional surveillance methods.<sup>9,10</sup> Many states are developing surveillance networks for their entire populations.<sup>11</sup> These statewide efforts can serve public health and clinical care needs simultaneously.

We have been tasked to build a surveillance network for Indiana. The Regenstrief Institute has a 30-year legacy of successful, sustained health information initiatives.<sup>12,13</sup> Based on this legacy and the longstanding collaboration between the ISDH and Regenstrief,<sup>14</sup> the state has contracted with us to develop and deploy a statewide infrastructure for syndromic surveillance that will capture encounter data from all 114 Indiana emergency departments (ED's). The four-year rollout began in June, 2004. Here is how we're doing it.

#### Methods

In 2003 the Indiana State Department of Health (ISDH) distributed a state syndromic surveillance plan to all Indiana hospital CEO's and CIO's. This plan described the vision to develop a statewide infrastructure for electronic transfer and analysis of data from all Indiana agencies collecting health care data. The mandate for this program was contained in a new state law entitled, "Counterterrorism symptom and health syndrome data collection."<sup>15</sup>

In conjunction with state legislature and the Regenstrief Institute, the ISDH identified the syndromic surveillance dataset. Table 1 lists the minimum data elements collected for each ED encounter. While additional data could be collected, the ISDH elected to begin with a manageable subset. The hospital name identifies each institutional data source. Basic patient identifying data including name, medical record number, sex, and date of birth provides a link for follow-up in the event of an outbreak. The patient's chief complaint (CC) reflects the primary symptoms motivating a patient to seek treatment and is used to determine which syndrome the patient might have. Day and time of visit provide a timestamp for time-based outbreak detection algorithms. Address, city, state, zip, and county provide supplemental information for outbreak detection algorithms. For example, a group of 15 respiratory cases noted on a given day may not be cause for alarm but 15 cases on the same day and

**Table 1: Data elements collected for each ED visit**

|                                   |
|-----------------------------------|
| Hospital name                     |
| Patient name                      |
| Medical record number             |
| Birth date                        |
| Sex                               |
| Address, City, State, Zip, County |
| Date/time of encounter            |
| Chief complaint                   |

street likely signal an event of public health significance.

Although the HIPAA final rule permits disclosure of protected health information (PHI) for public health surveillance,<sup>16</sup> some note that the final rule uses the verb "may disclose", rather than "must disclose". Consequently, many stakeholders interpret the rule as requiring additional specific legislative or regulatory permission. To cover this possibility, Indiana enacted legislation in March, 2004 requiring syndromic surveillance using ED data.

We do not want to disrupt hospital workflow or increase their work loads, so we limited the surveillance data elements to those already captured during patient registration by most emergency departments and can be transmitted in a standard Health Level Seven (HL7) registration message,<sup>17</sup> which many hospitals are capable of exporting from their registration system. We elected to receive HL7 messages directly from each hospital. In unusual cases where HL7 transmission is not possible, we are prepared to accept other formats.

We were uncertain of hospitals' capabilities to collect and transmit the minimum dataset. Before starting this project, ISDH and Regenstrief surveyed hospitals' IT capabilities in a number of areas. We asked questions about existing electronic messaging capability for securely transmitting encounter data such as chief complaint and demographics, their preferred data transmission method (i.e. batch FTP, email, real-time), data format options (i.e. HL7, flat-file, delimited), and types of data systems including patient registration, ED charting, and laboratory results.

Indiana is divided into 10 public-health preparedness (PHP) districts based on geography, population, and distribution of public health resources. (Figure 1) The ISDH wanted to include at least 2 hospitals in each district during the first year of implementation. For the first year's implementation we chose hospitals that could 1) collect and transmit chief complaint data electronically, 2) transmit HL7 registration messages, and 3) satisfy the two hospitals per district requirement. Because of preexisting data sharing relationships, we also included all hospitals

participating in the Indiana Network for Patient Care (INPC)<sup>13</sup> in the first year group.

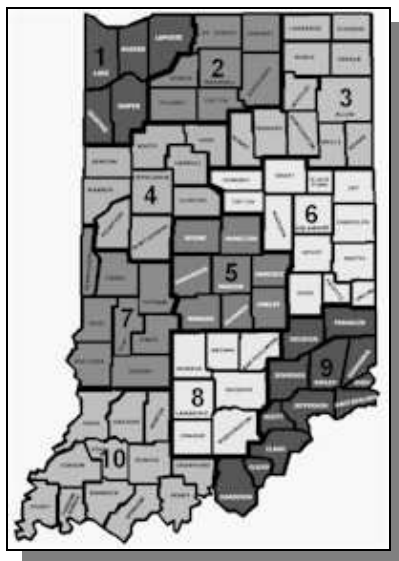


Figure 1: Indiana public health preparedness districts

With legislation in place, data elements established, the first-year hospitals chosen, and public awareness raised through a press release, representatives of Regenrief and ISDH visited each hospital for face-to-face meetings during which we described the rollout plan and answered hospital questions. Presentation attendees varied from hospital to hospital, but often included CIO's, ED managers, clinicians, application software managers, interface engineers and networking specialists. At the initial meetings we emphasized the value of standardized messaging and data re-use. We pointed out precisely where surveillance data elements fit in the HL7 message. We asked for their preferences for secure network connectivity. We also encouraged hospital

leadership to sign the data sharing agreements.

After obtaining the hospital's agreement to participate, we worked on two kinds of tasks. The first was an analysis of their HL7 messages to 1) assess their compliance with the HL7 standard, and 2) analyze their internal codes for the surveillance data elements. Because HL7 does not specify standard codes for fields such as gender, we defined mappings from the local codes to our project-wide codes. All hospitals adapted their messages to send 'M' and 'F' for gender codes. To identify the message source, hospitals send unique application/facility codes assigned by Regenrief. The state wanted their internal county ID included in each message. We preprocess messages lacking this code and map the code using city and zip. Some ED's collect minimal data when first encountering patients using a process called "quick-registration". They capture chief complaints electronically *after* triaging patients. Quick-registration processes generate multiple HL7 registration messages that we aggregate at the receiving end. The first message may contain patient name and address, while the follow-up message may contain the chief complaint. After reviewing these issues, we configure Regenrief's message processors to accommodate new incoming streams.

The second task is establishing network connectivity, which has many challenges. We encounter hospitals with varying experience in clinical messaging. Many have not sent registration messages outside their institutions. We've encountered hospitals with varying levels of control over their network infrastructure. Hospitals using offsite datacenters have less flexibility to select technology for external data flows. Variations in network security policy require different technical solutions. Many hospitals have a corporate policy mandating virtual private networking (VPN) for secure connectivity. Still others fear the

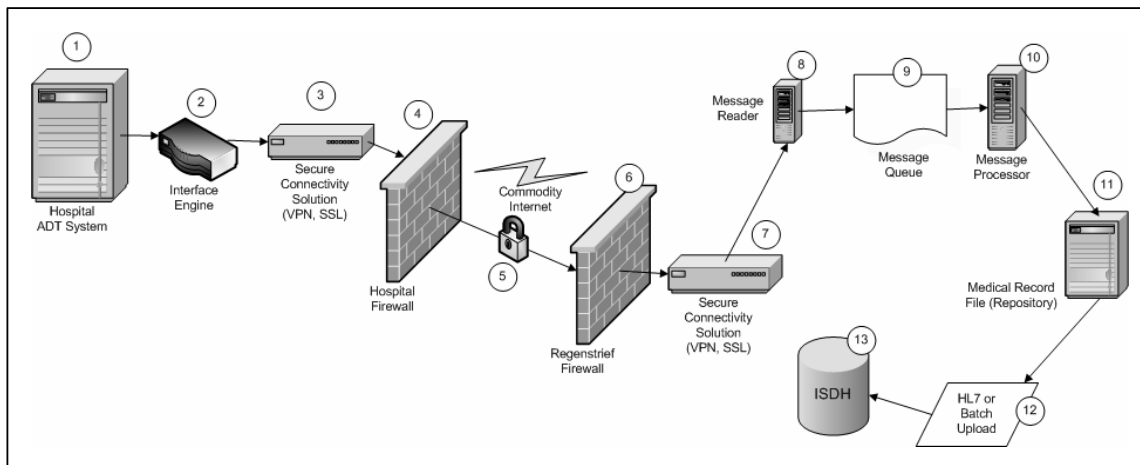


Figure 2: ED encounter data flow

consequences of connecting two networks through a VPN, and prefer the restricted features of Secure Sockets Layer tunneling (SSL). In this diverse and challenging environment there is need for flexible, secure and robust network expertise on the receiving end. We can receive data using Virtual Private Network (VPN) technology, Secure Sockets Layer Tunneling (SSL), and encrypted FTP (sFTP) transfers.

Once we establish data format and network connectivity, the data flow as illustrated in figure 2. The registration system generates an HL7 message and sends it to the hospital's interface engine. The interface engine directs the message to their secure connectivity solution (VPN, SSL tunnel) and the message is encrypted. The encrypted data passes through the hospital's firewall to the Internet. Data passes through Regenstrief's firewall where we decrypt the message and pass it to the message reader. The message queue holds messages for processing. The message processor extracts and formats message content. Regenstrief collates and transfers surveillance data every three hours via secure connection to the ISDH. We monitor message flows from each hospital. When flows drop below a pre-specified threshold, an email alert is sent to system administrators. The ISDH currently forwards aggregate deidentified data to the ESSENCE syndromic surveillance system.<sup>18</sup>

## Results

Of the 114 Indiana hospitals with ED's, 108 (95%) returned pre-implementation surveys. Sixteen registration system vendors were represented among 99 respondents. Siemens Medical Solutions and Meditech were the most prevalent, both with 21 installations. We found inaccuracies in some survey results. Forty-four hospitals reported collecting chief complaint on paper only, but follow-up phone interviews revealed that at least 23 electronically captured chief complaint. Some were unfamiliar with the term chief complaint, but knew it by other names such as "reason for visit" or "patient states". Although 95 hospitals reported ability to generate HL7 registration messages, in a separate question 83 hospitals reported they could not generate HL7. Follow-up phone interviews revealed that inconsistencies arose when multiple personnel or personnel with insufficient knowledge completed surveys. Not surprisingly, face-to-face meetings with hospitals are the most accurate source of information.

We began engaging hospitals in October, 2004 and currently have 42 signed data use agreements. Although we offer several modes for data transmission, we currently receive all data via secure point-to-point real-time messaging. The

infrastructure uses a variety of connectivity solutions, including dedicated ISDN, leased T1 lines, and encrypted connections over commodity Internet. We receive an average of 3,200 encounters per day requiring a daily bandwidth of 4.5 MB. Six hospitals perform quick-registrations with multiple messages per encounter. The remaining hospitals send a single message per ED encounter. Occasionally more than one hospital shares the same interface engine, allowing multiple institutions to send data through the same connection.

## Discussion

We've shown that statewide health information exchange is possible, but – not surprisingly – a discovery process. Encouraging use of the HL7 standard benefits, makes the process smoother, but requires additional education for the participants.

We've identified three rate-limiting steps in this process. They include securing a signed data use agreement, coordinating and attending face-to-face meetings with hospitals, and establishing physical connectivity.

*Initial buy-in and securing a signed data use agreement.* Most hospitals sign the data sharing agreement in a timely fashion. However, until the agreement is signed, we cannot begin analyzing hospital data because it contains protected health information.

*Attending face-to-face meetings with hospitals.* The logistical challenge of coordinating face-to-face meetings requires synchronizing stakeholder schedules from the state, Regenstrief, and hospitals. Although Regenstrief and the ISDH are based in centrally located Indianapolis, trips to many distant hospitals require full-day trips. We consistently note hospitals are more receptive to the project presentation when delivered in person rather than via teleconference. Discussions tend to be longer and hospitals ask more questions. Given the foundational nature of this initiative, we believe it is time well invested.

*Establishing physical connectivity.* We conclude that real-time data flows are preferred to batch transfers because real-time messaging can easily leverage flow control (ACK/NACK) built into the HL7 messaging protocol. SSL tunneling is the optimal approach to connectivity for point-to-point data connections because it combines real-time data flows, data encryption, and can be installed behind the hospital's firewall in a straightforward fashion. Further, there are several inexpensive, robust, open-source SSL tunneling implementations.<sup>19,20</sup> While VPN technology can be robust, secure, and flexible, it has the following drawbacks: installation and configuration can be complex (one hospital spent

over 2 weeks working with their vendor's tech support to troubleshoot their VPN connection), different vendors' products occasionally fail to interoperate, the protocol allowing VPN endpoints to function behind firewalls (IPSec/UDP) is still evolving,<sup>21</sup> and improperly configured VPN's can create address space collisions. FTP transfers are fraught with multiple points of failure, are not real time, and lack robust flow control capabilities.

By working through these rate-limiting steps, we are laying the groundwork upon which future services can build. Once reliable HL7 streams are in place, adding content is relatively trivial. In fact, this centralized model of data sharing serves as the basis for INPC's citywide RHIO.<sup>13</sup> With the infrastructure already in place, we could expand services to include electronic delivery of reportable disease results to the ISDH and local providers, chronic disease surveillance, and exchange of data such as immunization records or death events. As we add services, the additional re-use of data lowers costs for everyone. With the state's backing and cooperation from hospitals, this infrastructure could lay the foundation for the country's first statewide RHIO.

#### References

1. Thompson T, Brailer D. The Decade of Health Information Technology: Delivering Consumer-centric and Information-rich Health Care. Washington, DC: Department of Health and Human Services: National Coordinator for Health Information Technology; 2004. Available from: <http://www.hhs.gov/onchit/framework/hitframework.pdf> (accessed March 16, 2004).
2. AHRQ Health Information Technology Programs. 2005 [cited March 16, 2005]; Available from: <http://www.ahrq.gov/research/hitfact.htm>
3. Connecting Communities for Better Health. 2005 [cited March 16, 2005]; Available from: <http://ccbh.ehealthinitiative.org/about/default.msp>
4. Diamond C, Garrett D, Lumpkin J, Marchibroda J, Pardes H. Achieving Electronic Connectivity In Healthcare: A Preliminary Roadmap from the Nation's Public- and Private-Sector Healthcare Leaders: The Markle Foundation; 2004. Available from [http://www.connectingforhealth.org/resources/cfh\\_aech\\_roadmap\\_072004.pdf](http://www.connectingforhealth.org/resources/cfh_aech_roadmap_072004.pdf) (accessed March 16, 2005).
5. Bush G. State of the Union Address. 2004 [cited March 16, 2005]; Available from: <http://www.whitehouse.gov/news/releases/2004/01/20040120-7.html>
6. Walker J, Pan E, Johnston D, Adler-Milstein J, Bates D, Middleton B. The Value Of Health Care Information Exchange And Interoperability. Health Aff 2005.
7. McDonald C, Overhage J, Mamlin B, Dexter P, Tierney W. Physicians, Information Technology, and Health Care Systems: A Journey, Not a Destination. J. Am. Med. Inform. Assoc 2004;11:121-124.
8. CDC. Public Health Information Network standards, functions, and specifications: Version 1.2. 2003 [cited March 16, 2003]; Available from: <http://www.cdc.gov/phn/architecture/index.htm>
9. Hogan WR, Tsui FC, Ivanov O, Gesteland PH, Grannis S, Overhage JM, et al. Detection of pediatric respiratory and diarrheal outbreaks from sales of over-the-counter electrolyte products. J Am Med Inform Assoc 2003;10(6):555-62.
10. Reis BY, Mandl KD. Syndromic surveillance: the effects of syndrome grouping on model accuracy and outbreak detection. Ann Emerg Med 2004;44(3):235-41.
11. Bravata D, McDonald K, Smith W, Rydzak C, Szeto H, Buckneridge D, et al. Systematic Review: Surveillance Systems for Early Detection of Bioterrorism-Related Diseases. Ann Intern Med 2004;140:910-922.
12. McDonald CJ, Overhage JM, Tierney WM, Dexter PR, Martin DK, Suico JG, et al. The Regenstrief Medical Record System: a quarter century experience. Int J Med Inform 1999;54(3):225-53.
13. Overhage J, McDonald CJ, Suico JG. The regenstrief medical record system 2000: Expanding the breadth and depth of a community wide EMR. Proc AMIA Symp 2000:1173.
14. Overhage J, Suico J, McDonald C. Electronic laboratory reporting: barriers, solutions and findings. J Public Health Manag Pract 2001;7(6):60-6.
15. IC 16-19-10-8: Counterterrorism symptom and health syndrome data collection. In: IC16-19-10-8; 2004.
16. Department of Health and Human Services, Office of the Secretary. The Health Insurance Portability and Accountability Act of 1996, Standards for Privacy of Individually Identifiable Health Information. 45 CFR § 164.512 (b). 2000;65(250):82668-78.
17. Health Level Seven (HL7) version 2.3.1. April 14, 1999 [cited March 16, 2005]; Available from: <http://hl7.org/about/hl7about.htm#v231>
18. Burkhom H, Elbert Y, Feldman A, Lin J. Role of data aggregation in biosurveillance detection strategies with applications from ESSENCE. Syndromic Surveillance: Reports from a National Conference, 2003. MMWR 2004;53(Suppl):67-73.
19. Hatch B. Stunnel: Universal SSL Wrapper. 2005 [cited March 16, 2005]; Available from: <http://www.stunnel.org/>
20. Parkinson I. Custom SSL for advanced JSSE Developers. 2002 [cited March 16, 2005]; Available from: <http://www-128.ibm.com/developerworks/java/library/j-customssl/>
21. Huttenen A, Swander B, Volpe V, DiBurro L, Stenberg M. UDP Encapsulation of IPsec ESP Packets: Internet Engineering Task Force; 2005. Report No.: RFC 3948.