

Application of Information Technology ■

Development of a Clinical Data Warehouse for Hospital Infection Control

MARY F. WISNIEWSKI, MSN, PIOTR KIESZKOWSKI, BS, BRANDON M. ZAGORSKI, MS,
WILLIAM E. TRICK, MD, MICHAEL SOMMERS, BA, ROBERT A. WEINSTEIN, MD,
FOR THE CHICAGO ANTIMICROBIAL RESISTANCE PROJECT

Abstract Existing data stored in a hospital's transactional servers have enormous potential to improve performance measurement and health care quality. Accessing, organizing, and using these data to support research and quality improvement projects are evolving challenges for hospital systems. The authors report development of a clinical data warehouse that they created by importing data from the information systems of three affiliated public hospitals. They describe their methodology; difficulties encountered; responses from administrators, computer specialists, and clinicians; and the steps taken to capture and store patient-level data. The authors provide examples of their use of the clinical data warehouse to monitor antimicrobial resistance, to measure antimicrobial use, to detect hospital-acquired bloodstream infections, to measure the cost of infections, and to detect antimicrobial prescribing errors. In addition, they estimate the amount of time and money saved and the increased precision achieved through the practical application of the data warehouse.

■ *J Am Med Inform Assoc.* 2003;10:454–462. DOI 10.1197/jamia.M1299.

Antimicrobial resistance among pathogens that cause hospital-acquired infections increases health care costs and patient morbidity and mortality.¹ Strategies to control antimicrobial resistance in hospitals have been outlined by a multidisciplinary group of experts.² If implemented, these strategies are likely to significantly reduce resistance rates. One of the key goals is development of a hospital information system that could recognize and report trends in antimicrobial use and resistance.

Currently, most researchers and quality promotion professionals abstract patient level data manually from patients' medical records and reenter the data into databases. This process is labor-intensive, which limits the population size evaluated, and is error-prone.^{3–5} Current technology makes electronic transfer of data elements into a clinical data warehouse feasible, efficient, and accurate.⁶ We describe how our Infectious Diseases Division developed a relational clinical data warehouse using existing data from our health

care system's clinical computer applications (laboratory, pharmacy, radiology, admission/discharge). We report our successful use of the clinical data warehouse to automate measurement of performance indicators and our initial experience with electronic surveillance for infection control. We outline the steps involved, describe barriers encountered, report our responses (countermeasures) to those barriers, and provide examples of how we have used the data to facilitate performance improvement activities.

Background

Beginning in 1998, we undertook a five-year hospital-based demonstration project focusing on control of antimicrobial resistance, the Chicago Antimicrobial Resistance Project (CARP), under a cooperative agreement with the Centers for Disease Control and Prevention.⁷ CARP undertook a series of discrete quality improvement projects to reduce unnecessary antibiotic prescribing and to reduce the transmission of antimicrobial-resistant pathogens among hospitalized patients within the Cook County Bureau of Healthcare Services, a three-hospital public health care system. Essential (core) measures for CARP interventions included secular trends of hospital-acquired infections, antimicrobial resistance, and antimicrobial utilization. An additional goal of CARP was to determine the costs of treating and preventing hospital-acquired antibiotic resistant infections.

Design Objectives

An essential infrastructure requirement was an information management system designed to detect, track, and report the occurrence of antimicrobial-resistant organisms and to quantify antimicrobial use for the entire inpatient population of 48,000 admissions per year. As is common in health care, a system-wide relational database linking laboratory, pharmacy, and administrative data did not exist.⁴ To manage data

Affiliations of the authors: Department of Medicine, Cook County Hospital, Chicago, Illinois (MFW, PK, BMZ, RAW); Division of Healthcare Quality Promotion, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia (WET); Department of Hospital Information Services, Cook County Bureau of Health Services, Chicago, Illinois (MS); Rush Medical College, Chicago, Illinois (RAW).

This work was supported by The Centers for Disease Control and Prevention Cooperative agreement # U50/CCU515853. The Cook County Bureau of Health Service Institutional Review Board approved this project on December 1, 1998.

Correspondence and reprints: Mary Wisniewski, MSN, Division of Infectious Diseases, Cook County Hospital, 1901 West Harrison Street, Suite 124 Durand, Chicago, IL 60612; e-mail: <wise@aol.com>.

Received for publication: 11/27/02; accepted for publication: 04/19/03.

for the large number of inpatients being evaluated, we built a clinical data warehouse in client/server architecture. The clinical data warehouse was designed to store data collected from nonelectronic sources (e.g., manually abstracted data from patient medical records and surveys scanned using optical character recognition) and to store imported electronic data from existing hospital information systems. To facilitate communication among investigators, we developed a local area network and an intranet Web site that linked investigators and information (Fig. 1).

Methods

The development of the CARP electronic Infectious Disease (eID) clinical data warehouse required several steps (Table 1). Initial activities included translating the CARP objectives into core measures of performance, seeking administrative and technical support, determining which data fields (data elements) were needed, and learning where and how data were stored in each hospital's information systems (Table 2). We learned that each clinical department (e.g., pharmacy, laboratory, medical records) stored data in a separate database server and that each of these servers was interfaced to a single proprietary database that was housed and managed in Pennsylvania as part of the hospitals' information system. This two-tiered configuration enabled us to access data locally from each department's server (Table 3). This was necessary because the hospitals' information system report writers lacked an electronic export function. Obtaining written information about each database's architecture, such as documentation about the database tables' relationships and a data dictionary of encoded elements, was essential for understanding the data and for extracting relevant variables into our own clinical data warehouse.

System Requirements

The eID clinical data warehouse is a relational database built with:

- Relational database management system: Microsoft SQL 7.0 Server (Microsoft Inc., Redmond, WA)* which includes the following applications: object linking and embedding database (OLE DB), Data Transformation Services (DTS), online analytical processing (OLAP), and data mining.^{7,9-10}
- Software applications used to display information: Visual Basic 6¹¹ (Visual Basics, Inc., Cape Town, South Africa), Web enabled Crystal Reports¹² (Crystal Decision, Inc., Edison, NJ), and Microsoft Office Professional.
- Hardware: Two Dell Power Edge 4300 Redundant Base 400 MHz dual Intel Pentium II Servers (Dell Computer, Inc., Round Rock, TX); three Dell Dimension XPS R450MHz Pentium II Minutowers; two Dell Inspiron R450MHz Pentium II Notebooks.
- Operating System: Microsoft Windows NT Server 4.¹³
- Hard disk space: 180 GB redundant array of independent disks (RAID) and external Dell PowerVault 220.
- Tape Backup: Dell Power Vault 110T.

- Power Backup: Smart UPS (undisturbed power supply) 1400 (APC, Inc., Kingston, RI).
- Statistical Software: SAS (SAS Institute, Cary, NC) and SPSS (SPSS Inc., Chicago, IL).
- Scanning: AutoData Systems (Auto Data Systems Design, Inc., San Angelo, TX).

System Connectivity

The Data Transformation Services (DTS)⁸ function of the SQL server enabled open database connectivity (ODBC) linkages to different database platforms (e.g., application databases from Pharmacy and Laboratory) via the Active X scripting engine. This function was programmed to automate data extraction from each primary server to our eID clinical data warehouse. To minimize any potential impact on the primary transactional server, the extraction step was scheduled once every twenty-four hours, between 1:00 AM and 5:00 PM. Data extraction of one day's data takes about two hours. The selected data elements are stored in Microsoft SQL tables. Although extraction could occur more frequently than once a day, currently, no CARP studies require this.

eID Contents

We imported data from 13 different computing and operating systems—pharmacy (3), laboratory (2), radiology (1), medical records (6), and emergency department (1)—in the three CARP hospitals. The clinical data warehouse also stores data generated by CARP studies or surveillance programs (e.g., infection control surveillance). After three years of operation during 1999 to 2002, the database contained 50 GB of information encompassing approximately 130,000 admission/discharge, 22,000,000 laboratory, 6,000,000 pharmacy, 2,000,000 radiology, 1,000,000 emergency department, and 500,000 procedure/diagnosis records (Tables 2 and 3).

Data Conversion

Electronic data sources do not encompass all the data necessary for analysis. Much of the primary clinical data, such as vital signs (e.g., blood pressure or temperature) or clinical assessments (e.g., coma scores, urinary output) are collected manually for specific studies. Additional data, such as outcome evaluations (e.g., infection control surveillance) and investigator determination about appropriateness of antimicrobial therapy, also must be transformed into an electronic format for analysis. We have used an optical character recognition (OCR) program for this transfer. In our studies, there has been $\geq 99\%$ agreement between data collection forms and scanned data. In general, we have used scanning methods rather than keystroke data entries when greater than 3,000 data fields (e.g., 30 data fields \times 100 surveys) require input. In our experience, the amount of time for keystroke data entry for more than 3,000 data fields exceeds the effort needed to develop a scan form and to input and verify data.

Data Accuracy

Validating electronic data for completeness, continuity, and accuracy requires a substantial investment of time and is an ongoing process. Missing or incorrect data from the primary sources, the hospitals' transactional databases, can result from errors in the data entry by care providers, hospital clerks, pharmacy technicians, or laboratory workers.^{5,14} Also,

*Use of trade names and commercial sources is for identification only and does not imply endorsement by the Public Health Service of the U.S. Department of Health and Human Services.

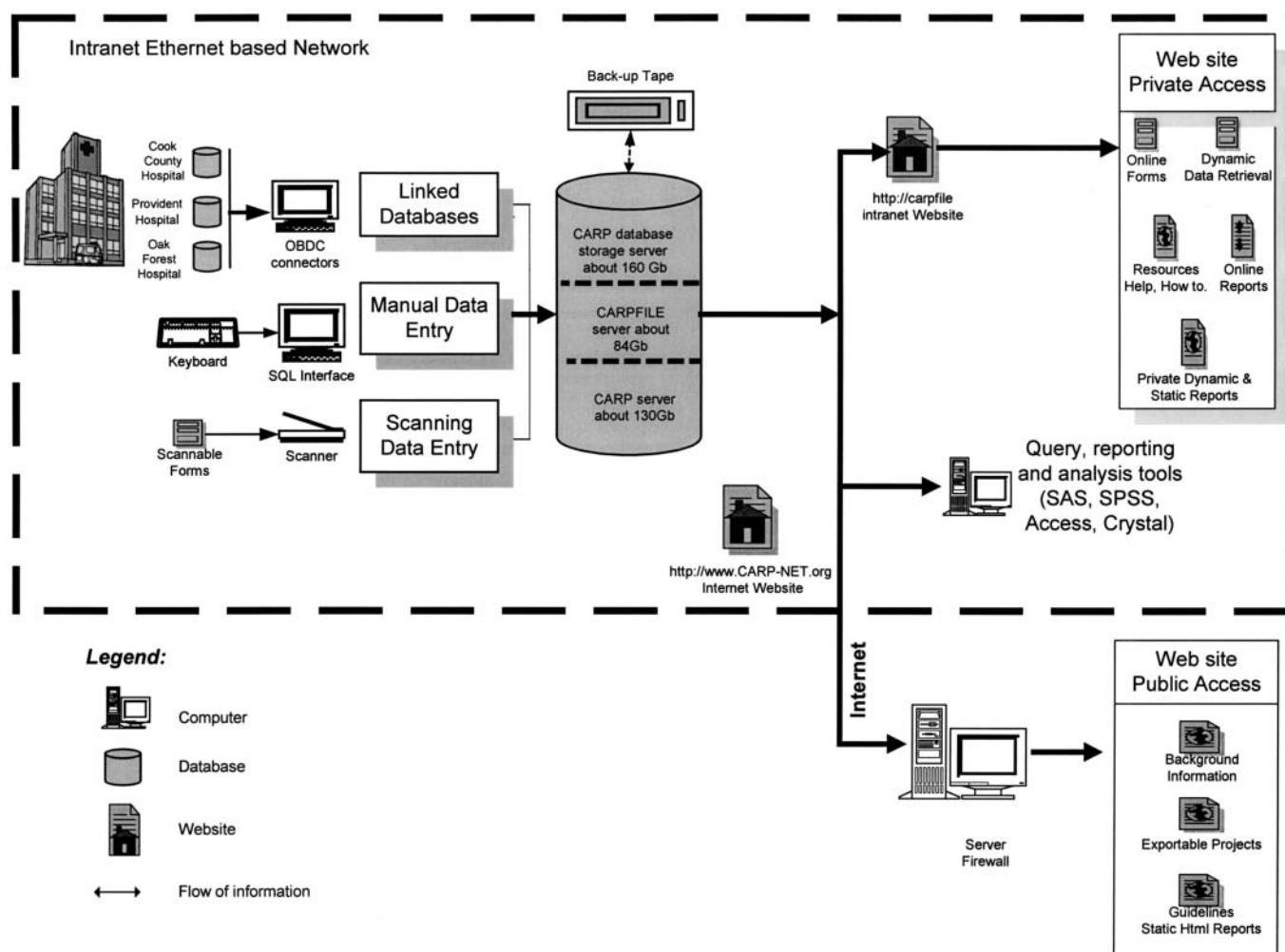


Figure 1. Schematic diagram of the CARP eID Information System shows sources of data, portals of data entry to the CARP database, and flow of information from inception to end user access via private and public access points.

corruption of data can occur during data transmission, storage, or retrieval. By abstracting a sample of our data and manually comparing the data set with the source data, we were able to find missing data and revise the abstraction programming code for each data set. We repeated this process until there was 100% agreement between data sets from our sample. A final check on validating our abstraction and data management for any new report was to compare patient-specific data from three systems, our data warehouse, the hospital information system vendor's application used to view patient results, and the patient medical record. Data from the hospital information system and from our clinical data warehouse matched, but the paper record missed one or more laboratory results 80% of the time. As an ongoing validation, we periodically compare reports programmed from our data warehouse with existing hospital reports (e.g., admission data, pharmacy expenditures, or shoe-leather surveillance programs). This process is performed by the personnel who use the information and by the programmers who developed the applications.

Personnel Resources

The development of the clinical data warehouse took two years and approximately 4,000 hours. The development was

performed by one director-developer (20% of time) who also served as chairperson of CARP's Informatics Subcommittee, one database administrator (70% of time), and one part-time system analyst (20% of time). A physician-administrator serving as the principal investigator of CARP led the staff.

Difficulties Encountered and Lessons Learned

Politics and Regulatory Issues

Two of the biggest challenges in the planning process were accommodating the security and confidentiality mandates of regulatory agencies and obtaining institutional approvals. Our first contact was with our own system's chief information officer, who required that we obtain secondary approval from the chief operating officer at each of the three hospitals and from each of the clinical department administrators (Table 4).

Keepers of data—often self-viewed as “owners” of the data—have a fiduciary responsibility to maintain data in a confidential and secure manner and to ensure that data are interpreted accurately. Allowing a third party direct access to computer servers may be perceived as compromising data-keeper responsibilities. Our request to create a copy of patient

Table 1 ■ Steps in the Development of a Clinical Data Warehouse

Activity	Steps	Description
Business understanding	Determination of essential (core) measures	<ul style="list-style-type: none"> • Transform the project objectives into essential measures of performance
Data understanding	Determination of data sources	<ul style="list-style-type: none"> • Determine necessary data fields (variables) to compute the core measures • Learn where and how data are stored
Data acquisition	Administrative approval	<ul style="list-style-type: none"> • Seek administrative and technical support for the project
	Connection to servers	<ul style="list-style-type: none"> • Obtain appropriate security clearance • Identify database administrator • Obtain software driver • Obtain IP address of the server • Obtain administrative-level read-only sign-on to server • Program ODBC connection
Modeling	Determination of data fields	<ul style="list-style-type: none"> • Gain knowledge of database design, table structure, and data dictionary
	Development of data model	<ul style="list-style-type: none"> • Design table structure and relationships to stored data • Program code to abstract essential data elements • Copy and store elements into the clinical data warehouse
Evaluation	Data integrity	<ul style="list-style-type: none"> • Validate completeness and accuracy • Clean data • Transform data to usable format • Maintain daily data transfer process

Abbreviations: IP, internet protocol; ODBC = open database connectivity.

level data for CARP projects challenged the paradigm of data "ownership." The introduction of the Health Insurance Portability and Accountability Act of 1996 (HIPAA),^{15,16} which addresses confidentiality of patient level data and the safeguards required for security of electronic data transfer, further complicated our task by introducing another level of

scrutiny. HIPAA does not specify how patient data should be protected but does note that institutions and individuals will be held accountable if confidentiality is breached.

To overcome the barriers associated with ownership and confidentiality concerns, we obtained senior administrative-

Table 2 ■ Sources of Data Included in the CARP Clinical Data Warehouse Listed by Location, Type, Vendor, and Time Period

Cook County Bureau of Health Services		Time Period		
Location	Application Type	1999–2000	2001–2002	2003
Cook County Hospital	Patient demographic	SMS*	SMS	Cerner
	Encounter discharge summary		SoftMed†	
	Census-bed information		SMS	
	Laboratory		SCC Soft‡	
	Pharmacy	← SMS →		
	Radiology	← ER LOG§ →		
	Emergency room visits			
Provident Hospital of Cook County	Patient demographic	← SMS →		
	Encounter discharge summary	SMS	SoftMed	
	Census-bed information			
	Laboratory	N/A¶	SCC Soft	
	Pharmacy	← SMS →		
	Radiology	← N/A →		
	Emergency room visits			
Oak Forest Hospital	Patient demographic	SMS	SMS	
	Encounter discharge summary		SoftMed	
	Census-bed information		SMS	
	Laboratory	N/A	SCC Soft	
	Pharmacy	← SMS →		
	Radiology	← N/A →		
	Emergency room visits	← N/A →		

*SMS, Siemens Medical Solutions, formerly known as Shared Medical Systems, available at <www.smed.com>.

†Softmed Systems, Inc., available at <www.softmed.com>.

‡SCC Soft Computer, available at <www.softcomputer.com>.

§Emergency Room Department at Cook County Hospital, available at <www.ccbh.org>.

||Cerner Corp, available at <www.cerner.com>. Unlike previous systems, Cerner data architecture is one database with all application data sets contained in individual tables. Additional clinical data elements will be available when the system is implemented, including device use, bedside clinical observations, and clinical documentation.

¶N/A, data set not available.

Table 3 ■ Core Measures, Data Sources, Tables, and Variables

Core Measures	Source of Data	Warehouse Tables	Examples of Select Variables
Antimicrobial resistance	• Laboratory	• Microbiology	Test ordered, source of specimen, date/time of collection & final tests results, organism, and susceptibility (quantitative and qualitative results)
Antimicrobial utilization	• Pharmacy	• Pharmacy	Generic name of medication, dose, frequency, route, order number, date/time of transaction
Resource utilization and cost*	• Laboratory	• Clinical labs	Diagnostic test and order numbers, medication transactions, procedures (CPT codes), diagnosis (ICD-9 codes), and price assigned to each line item resource
	• Radiology	• Microbiology lab	
	• Pharmacy	• Radiologic procedures	
		• Transactions and drugs dispensed	
	• Medical records	• Surgical procedures	
		• Diagnosis	
Patient demographics	• SMS invasion	• Bed information	Patient visit number,† patient medical record number, date/time of admission/discharge, hospital room number for each day, date of birth, sex, race, level of care type, diagnosis, procedures
	• Medical records	• Discharge diagnosis	
		• Procedures	
Hospital-acquired infections	• Surveillance surveys	• Infection control surveillance	Organism culture results, and classification of infection or colonization, primary or secondary, hospital- or community-acquired, catheter-related infection

NOTE. A complete description of the contents and data dictionaries of the clinical data warehouse can be found on our Web site <www.carp-net.com>.

*Used to determine actual costs from a hospital perspective.

†Patient level data were linked across applications with use of two primary keys: the patient visit number and patient medical record number.

level endorsement of the project. In addition, we informed administrators and data keepers that use of the information by the infectious diseases and infection control divisions of our system's hospitals would help meet regulatory mandates for infection control and improve patient care. Administrative acceptance included approval and endorsement from the chief of the Cook County Bureau of Health Services, hospital attorneys, and department directors. We implemented a security program that included policies, procedures, and technological safeguards to protect patient-level data and to assure compliance with our institutional review board (IRB) requirements. Establishment of the data warehouse was determined to be exempt from IRB review.

Technical Capability

The administration of the hospitals' information system data repository is a contracted service through an extramural vendor. This operational practice limited the development of technical expertise in database management within the hospitals' information services staff and restricted their ability to assist us in establishing a connection to the application servers. We learned that staff from each clinical department (e.g., pharmacy, laboratory, medical records, radiology) assisted with the design and management of data in its own server. Understanding the design of the information system software architecture and determining data flow of the two-tiered server configuration were critical to successfully build our clinical data warehouse.

Trust Relationships

To create links to local servers, we found it essential to use established, or to develop new, relationships with the clinical

department directors and staff who managed each local server. In exchange for the time provided by these staffers, we helped them access data. For example, we abstracted pharmacy data about unit-level drug dispensing practices, information the pharmacy needed to plan implementation of automated drug dispensing machines.

Database Content Knowledge

Selecting data elements from each database required knowledge of the database model and the data dictionary and clinical content expertise. Ideally, we would have had the documentation supporting each application, but often this was not available. Once we could view the database contents, clinically trained staff helped interpret the data variables and relational table design. The pharmacy database was relatively simple in design and required abstracting drug data from only one table. In addition, one of the pharmacy directors was able to provide important assistance by directing us to tables containing census information that we needed to report the antimicrobial use per patient days. In contrast, the laboratory database was complicated, and we required assistance from the software vendor's application specialist. The microbiology laboratory database had over 600 tables, only a fraction of which contained data elements relevant to the CARP study. We created one microbiology table by linking ten different tables; some of these tables contained only two variables (e.g., tables that contained encoded data descriptors). After data aggregation and validation, we learned that some laboratory results were archived to other data tables. Much effort was required to review the large number of tables, to determine the meaning of encoded variables, to locate stored variables,

Table 4 ■ Barriers and Responses (Countermeasures) to the Development of eID Clinical Data Warehouse

Focus	Difficulties Encountered	Countermeasures
Politics and regulatory issues	Lack of cultural acceptance of innovative strategies Proposed legislation restricting release of confidential electronic health care information	Know your leverage points Enlist high-level administrative support Be informed of regulatory issues
Trust relationship	Lack of trust	Build teams Share knowledge Use data only for stated purposes
Technical capability	Lack of technical expertise	Keep asking questions Educate staff in needed skills Network Use consultants when necessary
Database content knowledge	Lack of information	Seek assistance Team up with information application specialists and clinical staff
Usability of data	Corrupted data Data stored in unconventional format Data stored in non-standardized nomenclature	Limit number of data handoffs Transform unusable data into new variables
Information technology	Limited or absent resources	Purchase upgradable equipment Network
Evolution	Changing vendors brings new data types, formats, servers, and owners	Expect and plan for change Be flexible
Data analysis	Analytic and technical skill expensive	Start slow Develop expertise Contract for specific project skills

and to learn how to link tables and to abstract complete data sets. In addition to the complicated table structure, the microbiology data were partitioned into time-dependent preliminary and final results, and each of the final pathogens had a panel of 1 to ≥ 10 antimicrobial susceptibility results.

Usability of Data

Few published studies report on the accuracy or usability of clinical warehouse data. Because data accuracy is highly variable,¹⁴ methods should be applied for continuous monitoring. In addition, data may be accurate but not in a usable format for computation. For example, in our system's pharmacy tables, the field containing the dose of a medication could be expressed in three different ways: the dose field could be a number representing the actual dosage (e.g., 250 mg of medication); the dose field could represent the volume of fluid to be infused to administer one dose (e.g., 10 mL of medication); or the dose field could indicate the number of units to be administered (e.g., two pills). To address this problem, we had to create a new dose field by having clinical staff manually transform the free text into uniformly comparable data. Once the initial mapping of the 3,247 unique antibiotic dosing combinations in our pharmacies' formularies was completed (approximately 80 person-hours of labor), the ongoing update of new lines of data was minimal but repeated every three months. Another usability problem occurred when the clinical data warehouse encountered nonstandardized data formats. For example, in our pharmacy tables, at times, the drug name also included the preparation strength, which made simple queries by medication name nearly impossible. A third set of usability issues resulted from storage of data as text. The microbiology table contained the results of infrequently performed tests in a free-text comment field that was not abstracted easily. Also, the radiology text-based report files were compressed, and access

to them required proprietary extraction software and natural language processing software to interpret these reports.¹⁷

Evolution

Changes in application software and vendors are inevitable, and this requires clinical data warehouse developers to be continually in "development and refinement" mode. Table 2 provides an overview of our hospitals' information systems, the data sets that we have been able to obtain, and the time frame for new systems. In 2002, the Cook County Bureau of Health Services contracted with a new clinical information system vendor. This evolution in the information systems will require that CARP repeat the data acquisition process (Tables 1 and 3) or access these data from a data repository function planned by the Cook County Bureau of Health Services Information Systems Department.

Data Analysis

Having information stored in a database and analyzing these data to answer hypotheses are two separate domains. Although programming simple queries to provide descriptive statistics was performed readily, the desire for applying business intelligence models or analyzing and displaying patient level data in a complex arrangement of variables required more advanced programming and analytic skills. We considered purchasing an Application Service Provider, i.e., a software program developed, marketed, and sold for specific applications, but we found none to meet our specific objectives. We developed business specifications outlining our requirements and solicited bids from four consulting firms. The four responses provided plans to use our existing clinical data warehouse and to program the data into information about the care received by patients with infectious diseases. We assessed these proposals and determined that purchasing analytical services was expensive, that intellectual property would be jointly owned, that the final product (programming

code) would be proprietary, and that any product modifications or enhancements would require an ongoing contractual relationship. Entering into long-term contractual agreements was deemed too expansive for the scope of the project and assumed that the consultant could ensure data integrity and provide ongoing product support. In addition, the consultant required the study investigators with content expertise to work intensely with contractors in the design of the application. As an alternate approach, we chose to develop internal resources by employing a statistician and recruiting a master's-level graduate student.

Status Report: Use of the eID Clinical Data Warehouse—Surveillance and Measurement

Through timely access to reports about antimicrobial use and resistance trends and the rapid tracking of microorganisms in individual patients, the clinical data warehouse has facilitated assessment of, and actions to promote, quality health care; has automated the examination of resource utilization; and has provided the ground work to replace labor-intensive "shoe-leather" surveillance with electronic processes. The surveillance, quality improvement, and cost accounting uses of the clinical data warehouse are summarized in Table 5, and four sets of examples are described here.

Redundant Antibiotics

Because antibiotic combinations with redundant antimicrobial spectra are a potentially remediable source of excessive antimicrobial use, we wrote a computer program to detect combinations of antibiotics with overlapping spectrum of activity (i.e., two antibiotics that target the same or very similar bacteria).²¹ Using Web-based access to the computer program, a clinical pharmacist was able to query the database using drop-down menus, select a specific day, and obtain a list of patients who were receiving two or more antibiotics with overlapping activity. The patient list, sorted by unit location, noted all antimicrobials prescribed and highlighted the redundant combinations. The patient information could be printed with an attached data collection instrument to record follow-up observations. During a five-week test period, 71% of eligible cases flagged by the computer were judged to be truly redundant based on a clinical review, physicians readily accepted this unsolicited advice in 98% of these cases, and drug cost savings were \$4,500.²¹ Before the clinical data warehouse, surveillance for this type of medication error was not performed routinely. During the development for this intervention, a manual review of one day's antimicrobial use in our hospital required 10 person hours (PH) just for case identification. Automated real-time identification of this potential medication error allowed the clinical pharmacist to shift his focus from reviewing patient charts for errant medication orders to evaluating the appropriateness of therapy and conferring with the prescribing physician.

Antimicrobial Measures

The measurement of antimicrobial utilization was programmed to trend specific drugs by predefined daily dose, duration of therapy in days, and number of courses of antimicrobial therapy.²⁰ Utilization was stratified by administration route (intravenous or oral), patient location, and time (monthly aggregation). Using the inpatient census data, rates of use were normalized (e.g., rates per 1,000 patient-days) for comparison. Automated trend reports of anti-

microbial utilization are available at the investigators' desktop computers, through the study's intranet Web site. Without the clinical data warehouse, 0.37 full-time equivalent (FTE) or 770 PH per year would be required to quantify antimicrobial use; with the clinical data warehouse, after the initial development of programming code (160 PH), clinical staff have "as needed" access to reports offering additional levels of unit-based analysis. Previous improvement activities undertaken by study investigators required manual tabulation of antimicrobial utilization.²⁴

Surveillance for Infections and Trending Antibiotic-resistant Organisms

Use of the clinical data warehouse has automated the daily identification of patients with new positive cultures. Infection control personnel use a Web-based program to select a particular day (or days) and create a line listing of patients with positive cultures. This method has replaced the 0.44 FTE (915 PH/yr) previously required to manually review printed laboratory reports and to transcribe cases to a traditional report log and has detected approximately 6% more cases.

Our clinical data warehouse also has been used to create "unit-based antibiograms" (i.e., bacterial susceptibility patterns for infections occurring on a specific patient care unit) that investigators can use to guide selection of empiric antimicrobial therapy and that CARP staff use to evaluate the effectiveness of their interventions aimed to reduce rates of antibiotic resistance. Before the clinical data warehouse, 33 PH per year were used to compile the annual hospitalwide antibiograms; with the clinical data warehouse, we needed a one-time expenditure of 30 PH to write code and use 4 PH for each annual update. Moreover, using the clinical data warehouse, we are able to combine microbiology results with administrative data to create antibiograms by hospital unit, length of stay, or other variables. The clinical data warehouse also allows application of programming rules to simplify removal of "duplicate" isolates (repeated cultures obtained from a single patient).

Programming the clinical data warehouse also has allowed quantification and trending of resistant infections by anatomic site (e.g., positive cultures of sputum or urine) and by bacterial species detected. We display normalized rates (e.g., number of isolates per 1,000 patient days). Before the clinical data warehouse, 0.2 FTE (416 PH/yr) were required to trend the five most common antibiotic resistant organisms; with the clinical data warehouse, after the initial development of programming code (40 PH), clinical staff have on-demand access to reports. This level of retrieval and detail was not available to the clinical staff before development of the clinical data warehouse.

Surveillance for Hospital-acquired Infections

Determining whether positive culture results reflect infections acquired before hospitalization or in the hospital is an important aspect of infection control programs and traditionally requires time-consuming manual data collection. We computed bloodstream infection rates electronically at our health care facilities by developing computer algorithms and applying them to our clinical data warehouse. Before the clinical data warehouse, 0.22 FTE (458 PH/yr) were required to perform chart review of patients with

Table 5 ■ Examples of Chicago Antimicrobial Resistance Project Electronic Surveillance Applications: Focus and Value

Surveillance activities	
Rates of antimicrobial resistance	<ul style="list-style-type: none"> • Focus: Automated quantification of antibiotic-resistant isolates stratified by hospital, patient type, and time of acquisition • Units of measure: Number of resistant isolates by month (using two different duplicate isolate rules) and by time of onset • Reports: On-demand trend report of resistant isolates, expressed per 1,000 patient days • Value: Before clinical data warehouse (CDW) 400 PH/yr required to trend five most common resistant organisms hospitalwide; with CDW, after the initial development of programing code (40 PH), clinical staff have on-demand access to reports
Potential infections	<ul style="list-style-type: none"> • Focus: Automated identification of epidemiologically important positive microbiology data • Units of measure: Daily list of individual patients including date of culture collection, organism, susceptibility, source of specimen, and patient location • Reports: On-demand daily line listing retrievable from intranet Web site • Value: Before CDW, 915 PH/yr required to create daily line listing; with CDW, after the initial development of programing code (80 PH), clinical staff have on-demand access to reports
Blood stream infections	<ul style="list-style-type: none"> • Focus: Applied computer algorithms to categorize blood culture isolate as infection versus contamination and whether due to an infected intravascular device¹⁸ • Units of measure: Number of patients with catheter-related or other bloodstream infection, expressed per 1,000 patient days • Reports: Incidence trend of these infections • Value: Before CDW, 452 PH/yr required to perform chart review of patients with positive blood cultures; with CDW, computer algorithm performs this decision-making task
Central venous catheters (CVC)	<ul style="list-style-type: none"> • Focus: Identification of patient with central venous catheters (CVCs) using data mining software to predict patients by regression or decision tree analytic methods¹⁹ • Units of measure: Number of patients with CVCs, patient characteristics • Report: Manuscript • Value: Demonstration of feasibility of electronic surveillance to predict presence of a CVC
Antimicrobial consumption	<ul style="list-style-type: none"> • Focus: Automate quantification of antimicrobial use and stratified by hospital, unit, and route of administration • Unit of measure²⁰: Defined daily dose, therapy days, duration of therapy, and starts of new antibiotic courses • Report: On-demand access to trend of antibiotic use reports retrievable from intranet Web site • Value: Before CDW, 779 PH/yr would be expended to quantify use; with CDW, after the initial development of programing code (120 PH), clinical staff have on-demand access to reports offering unit-based analysis
Quality improvement/patient safety activities	
Antimicrobial redundant prescribing practices	<ul style="list-style-type: none"> • Focus: Automated identification and access to information for clinical pharmacy staff to a daily listing of patient receiving two or more antibiotics classified as redundant (i.e., overlapping antimicrobial spectra)²¹ • Unit of measure: Daily list of individual patients prescribed antibiotics; duplicate spectra antimicrobials highlighted • Report: On-demand daily line listing retrievable from intranet Web site • Value: Before CDW, surveillance for this type of medication error was not performed routinely; however, a pilot study of manual surveillance of one day's antimicrobial use took 10 PH (yearly projected time, 1.75 FTE) for case identification for one hospital; with CDW, 160 PH were used for design and development of application, which was then available on demand
Cost accounting activities	
Cohort identification	<ul style="list-style-type: none"> • Focus: Identified cohort of patients with >5 discharge diagnoses (surrogate of severity of illness)^{22,23} • Unit of measure: Population of discharge patients per year and count of individual patients' diagnoses • Report: List of patients meeting inclusion criteria • Value: Before CDW, we were unable to identify a cohort of patients with >5 discharge diagnoses (study inclusion criteria); with CDW, we are able to perform criterion-based queries in less than 15 minutes
Resource utilization	<ul style="list-style-type: none"> • Focus: Identified line item use of resources for individual patients • Unit of measure: Procedures, diagnostic tests, medications, patient bed occupancy • Report: Electronic list of patient encounters and resources utilized • Value: Estimated time saving of electronic vs manual abstraction of 1,500 medical records was 1,750 person hours

positive blood cultures and to determine when and where the infection was acquired; with the clinical data warehouse, the computer algorithm performs this task rapidly, and the results are well correlated with rates determined by the infection control professional.¹⁸

Conclusions

We believe that a relational data warehouse should be a component of every hospital information system. Because our vendor did not provide an easily accessed clinical data warehouse for research and quality improvement purposes,

we developed our own. The computer hardware and software technology for developing clinical data warehouses are available and relatively affordable. However, clinical and technical expertise, personnel time, administrative support, and substantial work are required to develop and implement a clinical data warehouse, especially across a network of hospitals. CARP, through this demonstration project, was able to successfully access microbiology, pharmacy, and related data; to overcome the barriers to create a clinical data warehouse; and to build an infection control information system that led to savings of time and money and that allowed personnel to redirect their efforts from acquiring data to implementing infection control interventions.

The availability of information systems-based data for our entire inpatient population has provided close to real-time, desktop access to information for our clinicians and investigators. With this information, we measure performance, monitor infection rates and antimicrobial use, and calculate costs of patient care.¹⁸⁻²⁵ The next steps are to expand our electronically facilitated patient care interventions,^{18,19} to address new infection control challenges,^{27,28} to further automate disease surveillance and electronic reporting to public health agencies, to assess the value of data mining techniques^{18,26} for infection control risk factor identification, to explore additional sources and types of data as our health care system begins to develop an electronic patient medical record, and to evaluate the ability of our system to meet challenges of bioterrorism surveillance.²⁹ The health care information system industry should add this level of function—electronic surveillance, data transfer, and reporting—into its hospital applications to remain competitive and to provide information that will give health care purchasers a greater return on their investment.

References ■

- Jarvis W. Selected aspects of the socioeconomic impact of nosocomial infections: morbidity, mortality, cost, and prevention. *Infect Control Hosp Epidemiol.* 1996;17:552-7.
- Goldmann D, Weinstein R, Wenzel R, et al. Strategies to prevent and control the emergence and spread of antimicrobial-resistant microorganisms in hospitals. A challenge to hospital leadership. *JAMA.* 1996;275:234-40.
- Overhage J, Suico J, McDonald C. Electronic laboratory reporting: barriers, solutions and findings. *J Public Health Manag Pract.* 2001;7:60-6.
- Thacker S, Stroup D. Future directions for comprehensive public health surveillance and health information systems in the United States. *Am J Epidemiol.* 1994;140:383-97.
- Gordis L. *Epidemiology* (ed 2). Philadelphia, PA: Saunders, 2000.
- Peterson T. Microsoft® SQL Server 2000 Data Transformation Services. Indianapolis, IN: Sams Publishing, 2001.
- The Chicago Antimicrobial Resistance Project. <<http://www.carp-net.org>>. Accessed October 10, 2002.
- de Ville B. Microsoft® Data Mining: Integrated Business Intelligence for e-Commerce and Knowledge Management. Boston, MA: Digital Press, 2001.
- Spenik M, Sledge O, Fernandez K, et al. Microsoft® SQL Server 7.0 DBA Survival Guide. Indianapolis, IN: Sams Publishing, 1999.
- Sturm J. Microsoft® Data Warehousing with SQL Server 7.0. Redmond, WA: Microsoft Press, 2000.
- Jennings R. Database Developer's Guide with Visual Basic® 6. Indianapolis, IN: Sams Publishing, 1999.
- Seagate Software. I. Crystal Reports™ Version 8—User's Guide. Scotts Valley, CA: Seagate Software, 2000.
- Minasi M, Anderson C, Creagan E. Mastering Windows NT Server 4 (ed 4). Alameda, CA: Sybex Inc., 1997.
- Hogan W, Wagner M. Accuracy of data in computer-based patient records. *J Am Med Inform Assoc.* 1997;4:342-55.
- Health Insurance Portability, and Accountability Act of 1996. <http://www.access.gpo.gov/su_docs/aces/aces140.html>. Accessed January 18, 2003.
- Hodge J Jr, Gostin L, Jacobson P. Legal issues concerning electronic health information: privacy, quality, and liability. *JAMA.* 1999;282:1466-71.
- Trick W, Chapman W, Wisniewski M, et al. Electronic interpretation of chest radiograph reports to detect central venous catheters. *Infect Control Hosp Epidemiol.* Nov 2003; in press.
- Trick W, Zagorski B, Tokars J, et al. Comparison of Computer Algorithms (CA) to Traditional Surveillance (TS) to Calculate Bloodstream Infection (BSI) Rates at Three Healthcare Facilities. Presented at the 40th Annual Meeting of the Infectious Diseases Society of America (IDSA), Chicago, IL, October 24-27, 2002.
- Zagorski B, Trick W, Cordell R, Solomon S, Weinstein R. Comparison of predictive models to detect central venous catheters (CVC) using data stored in the hospital information system (HIS). Presented at the 12th Annual Meeting of the Society for Healthcare Epidemiology of America (SHEA), Salt Lake City, UT, April 6-9, 2002.
- Zagorski BM, Trick WE, Schwartz DN, et al. The effect of renal dysfunction on antimicrobial use measurements. *Clin Infect Dis.* 2002;35:1491-7.
- Glowacki RC, Schwartz DN, Itokazu GS, Wisniewski MF, Kieszowski P, Weinstein RA. Antibiotic combinations with redundant antimicrobial spectra: clinical epidemiology and pilot intervention of computer-assisted surveillance. *Clin Infect Dis.* 2003;37:59-64.
- Roberts RR, Scott RD II, Cordell R, et al. The use of economic modeling to determine the hospital costs associated with nosocomial infections. *Clin Infect Dis.* 2003;36:1424-33.
- Roberts R, Cordell R, Scott RD I, et al. The actual cost of healthcare-associated infection. Presented at the 11th Annual Scientific Meeting of the Society for Healthcare Epidemiology of America (SHEA), Toronto, Ontario, April 2001.
- Schiff G, Wisniewski M, Bult J, et al. Improving inpatient antibiotic prescribing: insights from participation in a national collaborative. *Joint Comm J Qual Improve.* 2001;27:387-402.
- Schwartz D, DeMarais P, Steele L, Trick W, Weinstein R. Beneficial effect of a training intervention based on IDSA and SHEA guidelines for antibiotic (AB) use in long term care (LTC) Facilities. Presented at the 39th Annual Meeting of the Infectious Diseases Society of America (IDSA), San Francisco, CA, October 2001.
- Silver M, Sakata T, Su H, et al. Case study: how to apply data mining techniques in a healthcare data warehouse. *J Healthc Inf Manag.* 2001;15:155-64.
- Yasnoff W, Overhage J, Humphreys B, et al. A national agenda for public health informatics. *J Public Health Manag Pract.* 2001;7:1-21.
- Prevent Antimicrobial Resistance, 2002. <www.cdc.gov/drugresistance/healthcare>. Accessed October 12, 2002.
- Barthell E, Cordell W, Moorhead J, et al. The Frontlines of Medicine Project: a proposal for the standardized communication of emergency department data for public health uses including syndromic surveillance for biological and chemical terrorism. *Ann Emerg Med.* 2002;39:422-9.