Toward Best Practice: Leveraging the Electronic Patient Record as a Clinical Data Warehouse

Craig S. Ledbetter, RN, BSIS; Matthew W. Morgan, MD, FRCPC, MSc

ABSTRACT

Automating clinical and administrative processes via an electronic patient record (EPR) gives clinicians the point-of-care tools they need to deliver better patient care. However, to improve clinical practice as a whole and then evaluate it, healthcare must go beyond basic automation and convert EPR data into aggregated, multidimensional information. Unfortunately, few EPR systems have the established, powerful analytical clinical data warehouses (CDWs) required for this conversion. This article describes how an organization can support best practice by leveraging a CDW that is fully integrated into its EPR and clinical decision support (CDS) system. The article (1) discusses the requirements for comprehensive CDS, including on-line analytical processing (OLAP) of data at both transactional and aggregate levels, (2) suggests that the transactional data acquired by an OLTP EPR system must be remodeled to support retrospective, population-based, aggregate analysis of those data, and (3) concludes that this aggregate analysis is best provided by a separate CDW system.

KEYWORDS

• Electronic patient record
• Clinical decision support
• Clinical data warehouse
• Transactional analysis
• Aggregate analysis

Today’s outcomes-focused healthcare environment makes best practice—the delivery of evidence-based, high-quality, cost-effective care—more critical than
ever. And to help clinicians achieve it, healthcare organizations are employing many new information management tools, including

- The electronic patient record (EPR)
- Clinical decision support (CDS) system
- The clinical data warehouse (CDW)

Clearly, when clinicians use EPRs with CDS capabilities to assist in real-time decision making, patient care is improved. And when clinicians, administrators, and researchers use such systems for retrospective decision making, patient care is further improved. Numerous studies have revealed that EPRs with built-in CDS capabilities are effective tools that support best practice and that a critical component of their effectiveness is establishing a powerful CDW.1,2,3

An EPR with robust CDS features requires on-line transactional processing (OLTP) as well as OLAP capabilities.4 During each step in the clinical process, OLAP capabilities can provide CDS from at least two perspectives: (1) transactional patient focus (usually at the point of care) and (2) an aggregate or population-based focus (usually retrospective). Table 1 outlines some of the types of CDS that an EPR with a clinical data warehouse can provide, the focus of the analysis required for that support, and the point in the clinical process where this support can make a difference in the quality of patient care. (For a more comprehensive discussion on the dimensions of CDS see Perreault and Metzger, 1999).5

EPR systems are evolving to meet this wide variety of CDS systems. It is becoming clear that an EPR that hopes to fulfill all of the Institute of Medicine’s requirements for a computer-based patient record6 will have to be created through the integration of the functionality available from both OLTP and OLAP systems.

Requirements of CDS

Defined broadly, CDS systems are software applications that support patient-specific clinical decision making and comprise a set of knowledge-based tools fully integrated with both clinician workflow and a repository of complete, accurate, patient-specific clinical data.7

CDS capabilities are useful in all phases of the clinical process: (1) assessment, (2) planning, (3) intervention, and (4) evaluation. A comprehensive CDS should assist the clinician at every point of care, from rapid access to relevant knowledge bases to interactive criteria-based alerts.

When making healthcare decisions, clinicians consider the health problems and clinical status of specific patients, as well as the expected outcomes of a population of patients with similar health problems and clinical status. Ideally, accepted clinical practices are those shown to be effective for a
population of similar patients. Comprehensive CDS must be able to provide information from both patient-specific and population-based perspectives. These two decision-making perspectives suggest that two levels of information analysis also need to be supported by a comprehensive CDS system. Criteria-based alerts and rules-based protocol orders call for intra-transaction analysis of clinical data. But clinicians, administrators, and researchers require the aggregate analysis of clinical data for retrospective population-based studies, which are essential to the practice of evidence-based medicine. Examples of both point-of-care, transaction-based CDS requirements and retrospective, aggregate-based CDS requirements are provided in Table 1.

**Technical Overview**

It seems there are as many definitions for *EPR, CDS system*, and *CDW* as there are vendors who purport to offer them. A definition of each term, for the purpose of this article, follows.

### Table 1. Requirements of Clinical Decision Support

<table>
<thead>
<tr>
<th>Clinical Process</th>
<th>Patient Focus (Point-of-Care)</th>
<th>Population Focus (Retrospective)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Transactional Analysis</td>
<td>Aggregate Analysis</td>
</tr>
<tr>
<td><strong>Assessment</strong></td>
<td>• Risk-factor flags</td>
<td>• Opportunities for improvement</td>
</tr>
<tr>
<td></td>
<td>• Clinical group membership</td>
<td>• At-risk groups</td>
</tr>
<tr>
<td></td>
<td>• Critical value alerts</td>
<td>• Insight into disease processes</td>
</tr>
<tr>
<td></td>
<td>• Assessment templates</td>
<td>• Understanding of current clinical practice</td>
</tr>
<tr>
<td></td>
<td>• Relevant knowledge-base references</td>
<td>• Community health issues</td>
</tr>
<tr>
<td><strong>Plan or Intervention</strong></td>
<td>• Allergy warnings</td>
<td>• Clinical pathway development</td>
</tr>
<tr>
<td></td>
<td>• Drug-to-drug interactions</td>
<td>• Evidence-based practice guidelines</td>
</tr>
<tr>
<td></td>
<td>• Drug-to-procedure interactions</td>
<td>• Protocol development</td>
</tr>
<tr>
<td></td>
<td>• Procedure-to-procedure interactions</td>
<td>• Care standards development</td>
</tr>
<tr>
<td></td>
<td>• Standardized order templates</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Protocol order sets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Criteria-based orders</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Drug cost warnings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Procedure cost warnings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Duplicate drug checks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Duplicate procedure checks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Clinical reminders</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Relevant knowledge-base references</td>
<td></td>
</tr>
<tr>
<td><strong>Evaluation</strong></td>
<td>• Critical value alerts</td>
<td>• Outcomes measures</td>
</tr>
<tr>
<td></td>
<td>• Criteria-based alerts</td>
<td>• Wellness management</td>
</tr>
<tr>
<td></td>
<td>• Variance tracking</td>
<td>• Contract management</td>
</tr>
<tr>
<td></td>
<td>• Relevant knowledge-base references</td>
<td>• Clinical risk adjustment</td>
</tr>
</tbody>
</table>
**EPR.** An EPR is an electronically maintained (computerized) patient record system with point-of-care tools that support clinical care. Ideally, an EPR should support all episodes of care to create a complete longitudinal patient record. The fact that the patient data in the EPR are stored in electronic (digital) format makes possible a number of potential features, including (1) rapid, simultaneous access to the patient record by multiple users, (2) on-line data processing for automating clinical and administrative processes, (3) on-line information processing for clinical and administrative decision support, and (4) integrated access to data from multiple and disparate data sources.

**OLTP.** At the point of care, clinical interventions are transactional in the sense that they are ordered, scheduled, performed, and documented within fairly well-defined clinical processes. These requirements dictate that an EPR system handle a large number of discrete clinical transactions in both data entry and data review modes. On-line transactional processing (OLTP) is the term used to describe this type of processing.

Even in this era of gigahertz processors, OLTP EPR systems must be specifically designed and tuned to deal with the volume of transactional processing required by patient care operations. A transactional system, and in particular its database, is specifically designed for rapid transactional processing, and system performance may suffer if the system is used for a different kind of processing (for example, for retrospective, population-based analyses). A few examples of clinical care processes that are OLTP-transaction-based and require subsecond response times are patient registration, clinical documentation, order entry, results review, and clinical alerting.

**OLAP.** OLTP transaction data have significant analytical value for clinical and administrative decision support as well as research activities. The conversion of transactional data into information that can be used to guide and support operations is termed on-line analytical processing (OLAP).

This is a more inclusive definition of OLAP than is often used within the healthcare informatics industry. Many use the term OLAP to describe a specific software system that has multidimensional reporting capabilities. These OLAP systems typically use some sort of proprietary data structure (often termed cubes) that allows the user to rapidly change the dimensions by which a report is filtered, sorted, and grouped. This provides the click-and-drill functionality considered essential for on-line data analysis or data browsing, but the proprietary nature of these data structures requires special report-writing tools to access the information they contain. Certainly, such systems support analytical processing on-line, but OLAP requirements for the EPR are not limited to this type of aggregate multidimensional analysis. An EPR also requires support for point-of-care transactional OLAP, where analytical processing is carried out using data from real-time transactions to provide a wide variety of criteria-triggered events.

**ROLAP.** Relational on-line analytical processing (ROLAP) is a term used to differentiate software products that support aggregate, multidimensional
analysis without the use of proprietary data structures (cubes). These software products require special denormalized relational structures (for example, star schemas). In place of the proprietary cube files, ROLAP systems use pre-aggregated records in relational tables to provide rapid results to the end user. One advantage of a ROLAP system is the lack of dependence on a proprietary data structure. A variety of relational report-writing tools can be used to output information from the underlying ROLAP structures.

**Transactional and Analytical Requirements.** The type of data processing required to support analysis from a retrospective, population-based focus can tax an EPR system designed primarily for OLTP. The architectural design required for efficient OLTP is fairly inefficient for aggregate analysis and reporting. Aggregate analytical processing on a system designed for OLTP will cause competition for available processing resources, leading to unacceptable delays at the point of care or similar transaction-based clinical care processing situations where subsecond response times should be the norm.

A review of Table 1 confirms that an OLTP system, even one that provides facilities for patient-focused, point-of-care OLAP, cannot efficiently provide the full range of CDS needed for best practice. An EPR’s transactional system is critical for supporting the clinical process at the point of care, including real-time CDS. The transactional data collected by an EPR can and should be leveraged to provide all the types of information needed to support best practice.

Fortunately, it is possible to design a database system to support the complex operations and high-volume disk reads required for the aggregate analysis of specific patient populations: a clinical data warehouse (CDW). Extraction and reorganization of the data collected by the OLTP EPR into a CDW enables efficient population-based aggregate analysis.

**Attributes of a CDW.** Nussbaum and Ault described fairly thoroughly the attributes of a CDW, making the distinction that it is not just a large collection of clinical data. A CDW includes data from the EPR and (potentially) data from other enterprise systems, stored on a separate system and reorganized to support retrospective analysis. The data are filtered and manipulated to provide an integrated data set with common units and conforming dimensions. The data may be stored redundantly at various levels of aggregation to support analysis and enhance performance. The data are often organized into subject-oriented domains (data marts) to support various user populations and simplify security. Specific architectural and technological considerations are also required to optimize a CDW for aggregate analysis.

The CDW does not contain a mirror image of the data in the transactional system but rather a subset of data useful for retrospective aggregate analysis. Data needed only for real-time clinical process management are not needed in the CDW. For example, the OLTP system may contain data required to determine proper work-queue flow and resolve other real-time processing questions. The CDW does not need these data but rather contains the results of the clinical processes that occurred.
The CDW is not a real-time system. Real-time processing should be done on the OLTP system. The “best of breed” clinical systems integrate patient data from multiple clinical systems using real-time interfaces, but these systems still poorly support the need for aggregate OLAP. They also require a separate CDW to support comprehensive analytical processing.

The full range of CDS requirements for an EPR cannot be supported without the establishment of a powerful CDW populated with clinical transaction data.

Case Study: University Health Network

University Health Network (UHN) comprises three large academic teaching hospitals: Toronto General Hospital, Toronto Western Hospital, and Princess Margaret Hospital. In total, there are approximately 1,000 beds that accommodate more than 42,000 inpatients annually. There are more than 70,000 emergency department and 750,000 ambulatory visits annually.

Over the last seven years, UHN has made significant strides to advance its IT infrastructure. The three hospitals share clinical information management systems that (1) manage administrative data, (2) process admissions, discharges, and transfers, (3) support departmental operations, and (4) allow clinicians to enter orders and review results for most diagnostic tests on-line. UHN’s physicians, nurses, and allied health professionals can access important clinical information from more than four thousand networked Pentium computers enterprisewide.

The foundation of the enterprisewide EPR at UHN is Patient1 (Per-Se Technologies; http://www.per-se.com)—a comprehensive admission-transfer-discharge (ADT), order-entry and results-review system with CDS functionalities. This OLTP-based EPR contains information on more than three million patients and twelve million visits spanning more than a decade. The 160 GB database grows at a rate of 24 GB per year. At peak hours, more than five hundred users are on-line. With no unscheduled downtimes in more than two years, Patient 1 has been available to clinicians twenty-four hours a day at subsecond response times. Patient1’s new Java-based graphic user interface, Vista, is providing UHN with additional features, including improved clinician navigation, enhanced CDS, and the ability to integrate the World Wide Web and other clinical information management tools into a standard clinical desktop.

The Clinical Decision Support Project. UHN is embarking on an ambitious multiyear project to advance best practice by creating significant, measurable improvements in patient care quality and efficiency. The project’s key milestones include the following:

- Migrating to a complete EPR
- Providing comprehensive CDS
- Facilitating integration with other providers
• Improving resource management through scheduling and workflow improvements
• Supporting research strategies

Critical to the overall success of this project is a CDW that takes advantage of Patient 1’s rich transactional data to support clinical and administrative CDS while augmenting clinical research activities.

UHN’s CDS process starts with the identification of quality improvement opportunities (for example, reducing unnecessary laboratory testing or optimizing antimicrobial therapy). Those opportunities are measured, prioritized, and converted to EPR-enabled CDS interventions such as clinical alerts that remind the clinician at the time of ordering that the requested investigation is a duplicate, or (in the case of antimicrobial therapy) recommendations could be based on local epidemiological evidence and alerts that suggest changing from intravenous to oral antibiotics at the appropriate time. The interventions are implemented, tested, and piloted with clinicians, and UHN measures the effectiveness of the interventions on an ongoing basis.

The CDW is crucial in all steps of the process. First, by generating analytical reports that detail clinical practice, the CDW allows UHN to quickly identify and assess CDS opportunities. Without the CDW, UHN would have to extract data from numerous other sources, including the paper chart, to identify opportunities. UHN can also quickly identify the size of the opportunity through analysis to produce data-driven, evidence-based prioritization data, which can then be used to develop interventions (clinical alerts, reminders, clinical pathways, protocol order sets, and guidelines) for Patient 1, the transactional OLAP EPR.

Once the interventions are tested and implemented, the CDW becomes a critical tool for measuring their effectiveness. By monitoring the impact of an alert, for example, the CDW can rapidly identify compliance with interventions and overall effectiveness. Finally, management reports can be generated and provided to senior management that identify the overall impact of the intervention in terms of resource utilization and quality improvement targets.

Project Team. Support from all levels of UHN is a critical success factor in EPR strategy. The CDW project has support from the clinical, administrative, and research leadership at UHN. The information systems group, Shared Information Management Services (SIMS), manages the overall CDW project. The project has a full-time project manager (manager of research informatics) who reports directly to the CIO. The project manager receives approval from the SIMS Project Management Council (SPMC) and reports the status of project milestones back to the SPMC on a regular basis.

In addition, senior medical leadership (the medical advisory committee) monitors the project, as does the research informatics clinical advisory committee. The CDS system clinical advisory committee, a decision-making body made up of clinicians and chaired by a member of senior management, provides additional clinical input.
**Project Plan.** UHN’s CDW strategy began with a prototype CDW built in-house. The details of this project have been described elsewhere. But in summary, transactional data in Patient1 were converted into HL7 messages and exported via an interface engine (STC’s DataGate). The DataGate interface engine copied HL7 messages to files (one file per day of messages). The daily file of HL7 messages was manually transferred from the DataGate server (via FTP) and archived to CD. Due to the large size of the data files (approximately 35–40 MB per day), only registration and diagnostic test order messages were contained within the CDW. A Java program was used to parse data and populate an Oracle database. SPSS statistical analysis software was used to calculate additional fields that were based on data from multiple messages (for example, turnaround time), as was Microsoft Access through an ODBC connection to the Oracle database. This labor-intensive process allowed UHN to identify CDS opportunities focused on diagnostic test-ordering patterns and to measure the impact of the resulting EPR-enabled interventions.

UHN has used its initial CDW for retrospective analysis of physician diagnostic test-ordering practices, evaluation of the effectiveness of Patient1’s OLTP-based clinical alerts, and identification of quality improvement opportunities. The diagnostic test-ordering analysis included determining the most used diagnostic tests in all clinical environments, the extent of duplicate diagnostic testing practices, the average number of diagnostic tests ordered per visit site, and the priority of test ordering. The results in some cases were surprising.

For example, five tests—complete blood count (CBC), actuated partial thromboplastin time (APTT), prothrombin time (PT), blood film review, and fibrinogen—constituted 95 percent of all inpatient hematology tests ordered. At least 10 percent and often more than 25 percent of commonly ordered diagnostic laboratory tests were redundant, based on widely accepted time periods. UHN’s analysis of Patient1 clinical alerts revealed that displaying guidelines at the time of order entry could significantly reduce the number of inappropriate orders. Such data, contained in ongoing reports, have allowed UHN to identify clinical areas such as testing where there is noncompliance with alerts. As a result, UHN can quickly assess the problem, promote awareness of the issues, and implement appropriate educational maneuvers. In addition, UHN used the CDW to identify opportunities for improving the operational efficiency and effectiveness of diagnostic labs, especially with respect to turnaround times. Many areas for improvement were identified using these data.

This initial success provided a strong case for the implementation of a comprehensive CDW. UHN chose Decision1, Per-Se Technologies’ CDW, as its solution.

Decision1 provides several benefits over the prototype, including

- A greatly expanded data schema
- A vendor-maintained extraction, transformation, and loading process
• A utility for adding additional data elements (that is, elements not supported in the vendor-supplied schema) to the extraction process
• Specialized data warehousing structures such as star schemas, views, and summary tables to support multidimensional report writing
• A set of predefined management report workbooks
• A simplified end-user reporting environment
• A metadata repository with predefined reports and an ad hoc reporting capability

**The Decision1 Model.** The development of best practices is an iterative process. A feedback mechanism that gives the clinician information on important measures of performance is essential for best practice development. A clinical committee cannot simply decide to initiate change in a clinical practice pattern and then implement those changes in the operational EPR system. The organization is obligated to monitor any important outcomes those changes may affect.

The transactional EPR contains data necessary for this feedback mechanism but holds it in a database that is structured and tuned for transactional processing. Per-Se Technologies’ solution extracts data from the Patient1 hierarchical database, transforms and loads it into the Decision1 relational database, and uses data warehousing technologies to remodel the data to support multidimensional aggregate reporting.

In the diagram of the Decision1 Model (see Figure 1), it is easy to see what would be lost without the feedback mechanism provided by the aggregate processing system.

Information obtained from the CDW should be integrated into the EPR in several ways:

• Rules-based clinical alert criteria evaluated during transactional OLAP can be based on information obtained through aggregate analysis using the CDW. In this way, real-time alerts and reminders will address specific opportunities for improvement discovered in the clinical data. The effectiveness of those rules can be monitored in the same way.

• Clinical pathways, protocol order sets, and other standards of care utilized during the clinical care process can be developed and monitored using aggregate information from the CDW. This represents an evidence-based, iterative approach to achieving best practice.

• Relevant aggregate information (epidemiological reports, and so on) from the CDW can be presented to the clinician at the point of care to support specific types of clinical decisions. Clinical factors that change slowly, such as the prevalence of certain diseases in the patient population or the current antibiotic resistance patterns for a particular microorganism, can be constantly updated and available to the clinician.
Patient1 handles the processing of clinical transactions in Per-Se Technologies’ EPR solution, and its integrated rules engine can perform criteria-based analysis on the data captured during those transactions. The system can deliver real-time alerts and information to the clinician at the point of care or initiate protocol orders based on predefined criteria. Additionally, both orders and results can be routed to clinical and management queues for action or review.

Not having been designed as a generic CDW solution, Decision1 is closely integrated with Patient1. Code native to the Patient1 database performs extremely efficient data extraction and transformation. Customers can augment the data stream through add-on extraction reports written with the embedded Patient1 reporting tool. The relational database is implemented using Oracle 8i Enterprise Edition, which has substantial support for data warehousing
operations. Once loaded into the Oracle database, the extracted data are further reorganized into a number of specialized data structures (fact tables, dimension tables, views, cross-tab tables, and so forth) to support multidimensional reporting. These structures are relational objects, not proprietary data structures (for example, data cubes, as implemented by some multidimensional reporting solutions). The advantage of this ROLAP design is that most commercially available relational report-writing tools can also access these specialized structures. “Business areas” (data sets that support the reporting requirements of specific clinical and management areas) are then created with these specialized multidimensional structures using Oracle’s Discoverer, a relational multidimensional analysis tool. Discoverer provides metadata, security, and other administrative services to support a maintainable and customizable set of reporting applications. Access to Decision1 or to specific reports generated by the CDW system is available directly from the Patient1 user environment.

**Future Goals.** UHN’s CDW will support the deployment of numerous CDS interventions—such as rules, protocols, aggregate information, and quality indicators—with clinicians at the point of care. Some of these include the following.

**Using Rules.** Rules are used for a variety of purposes:

- Flagging high-risk patients within a patient population based on risk factors discovered using the CDW (for example, risk for morbidity or mortality, risk for cost outlier, risk for length-of-stay outlier). Early, focused intervention might improve outcomes in these high-risk cases.
- Preventing medication errors with allergy checks, drug interaction checks, drug-procedure interaction checks, and duplicate order checks, then watching aggregate error rates to evaluate effectiveness.
- Offering cost advisement on antibiotics when lower-cost alternatives exist, then watching aggregate antibiotic cost-per-case to identify opportunity and to evaluate effectiveness.
- Providing alerts to eliminate inappropriate tests or procedures, then using aggregate analysis to identify the opportunity and verify the improvement.
- Providing clinical reminders to help clinicians comply with various protocols (for example, pediatric immunization schedules, screening mammography guidelines, flu vaccine recommendations), which can improve hospital-, payer-, or government-monitored indicator rates.

**Using Protocols, Pathways, Care Plans.** Two interventions of this type will be (1) employing clinical pathway variance analysis reports for the patient’s specific pathway and (2) employing clinician practice comparison for patients on the overall population’s pathway.

**Using Relevant Aggregate Information.** Two interventions will be (1) using reports showing antibiotic sensitivities and culture rates of common
pathogenic microorganisms by patient or specimen type (for example, from sputum cultures from community acquired pneumonia patients), and (2) using epidemiological reports showing incidence of certain diseases, such as influenza, by age group and zip code over time (for example, to predict seasonal rates and current risk of disease).

**Using Quality Indicators, Special Studies, Variance Analysis.** Two interventions will be (1) using current outcome report for a treatment protocol currently being considered by a physician for a patient and (2) using percentile rankings for functional outcome scores (for example, geriatric assessment, childhood development) to make it easy to evaluate how a given patient’s score compares to a similar patient population.

**Conclusion**

Truly comprehensive CDS requires analytical data processing at both the transactional and aggregate levels. Organizations that realize the power of this approach can take advantage of transactional data acquired by their EPR to support real-time transactional, patient-based clinical care processing and retrospective, population-based aggregate analysis. An EPR-leveraged CDW is essential to achieving this goal. The current healthcare environment, with its focus on outcomes and improved quality of care, requires that the question no longer be, Do we need a CDW? but rather, How can we achieve a CDW? Whether an organization chooses to start with data marts and grow, build a solution from the ground up, or enlist the help of an experienced vendor, a CDW is a necessity for any organization that wants to provide its clinicians with the ability to pursue best practice.

**References**


**About the Authors**

Craig S. Ledbetter, RN, BSIS, is a senior healthcare operations analyst for Per-Se Technologies, Inc. He has fifteen years of experience as a registered nurse in the acute care setting and a bachelor’s degree in information systems.

Matthew W. Morgan, MD, FRCPC, MSc, is assistant professor, Department of Medicine, University of Toronto, and a general internist at University Health Network. Previously the director of clinical informatics at UHN, he recently joined Per-Se Technologies, Inc., as director of healthcare informatics.