An Analytics Framework for Physician Adherence to Clinical Practice Guidelines: Knowledge-Based Approach

Jaehoon Lee¹,², PhD; Nathan C Hulse¹,², PhD

¹Intermountain Healthcare, Salt Lake City, UT, United States
²Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Nathan C Hulse, PhD
Intermountain Healthcare
36 State Street
Salt Lake City, UT, 84111
United States
Phone: 1 3852976482
Email: nathan.hulse@imail.org

Abstract

Background: One of the problems in evaluating clinical practice guidelines (CPGs) is the occurrence of knowledge gaps. These gaps may occur when evaluation logics and definitions in analytics pipelines are translated differently.

Objective: The objective of this paper is to develop a systematic method that will fill in the cognitive and computational gaps of CPG knowledge components in analytics pipelines.

Methods: We used locally developed CPGs that resulted in care process models (CPMs). We derived adherence definitions from the CPMs, transformed them into computationally executable queries, and deployed them into an enterprise knowledge base that specializes in managing clinical knowledge content. We developed a visual analytics framework, whose data pipelines are connected to queries in the knowledge base, to automate the extraction of data from clinical databases and calculation of evaluation metrics.

Results: In this pilot study, we implemented 21 CPMs within the proposed framework, which is connected to an enterprise data warehouse (EDW) as a data source. We built a Web–based dashboard for monitoring and evaluating adherence to the CPMs. The dashboard ran for 18 months during which CPM adherence definitions were updated a number of times.

Conclusions: The proposed framework was demonstrated to accommodate complicated knowledge management for CPM adherence evaluation in analytics pipelines using a knowledge base. At the same time, knowledge consistency and computational efficiency were maintained.


KEYWORDS
clinical practice guidelines; care process model; visual analytics; clinical decision support

Introduction

Clinical practice guidelines (CPGs) are systematically developed statements that assist clinicians in making decisions about appropriate patient care for specific clinical circumstances [1]. Since the nature of CPGs includes complicated clinical knowledge, it is known to be challenging not only to formulate clinicians’ logics into the form of guidelines, but also to translate and implement these guidelines properly into clinical tasks and processes. Therefore, a number of studies have tried to develop systematic ways to implement CPGs [2,3], including computer-aided clinical decision support (CDS)-based approaches that enable personalized and timely implementation of CPGs [4-6] and knowledge-based approaches to systematically transform complicated knowledge of CPGs into clinical decision and practices [7-9].

Since clinical knowledge within CPGs originated from evidence of best practices, realization of CPGs in practice has had a positive impact on clinical workflow and patient outcome [10-14]. Therefore, it is important to measure and evaluate physicians’ adherence to CPGs in order to understand how providers are following guidelines in the postimplementation phase [15,16]. This evaluation may be an interdisciplinary project involving domain experts, knowledge engineers, and
data analysts, among others. Domain experts should derive evaluation logics and metrics from CPGs and document them by collaborating with knowledge engineers. To perform this evaluation, database engineers should create queries based on the definitions and run them against the clinical database. As a result of data extraction, evaluation outcomes would be delivered to consumers (ie, clinical champions or management leaders), often in the form of Web-based reports.

The problem is that knowledge gaps always exist while transforming logics of CPGs into evaluation-definition documents, developing queries, and generating reports. For example, translating definitions into computationally executable queries may vary by individual knowledge engineers and data developers [17]. Once analytics are delivered, consumers can only see the resulting data, but would lack an understanding of what logics were used to extract the numbers. If analytics are being delivered in a regular manner and evaluation logics are modified, it becomes confusing to know whether a report was made based on an old or new definition. In particular, analytics have become more integrated and automated by integrating data pipelines, report generation, and delivery workflow; this may involve more knowledge translation into the framework and may require a systematic method of management [18,19].

Knowledge gaps are caused by the difficulty of handling too many data points, miscommunication between domain experts and developers, misinterpretation of guidelines, loss of authorship of documents, and revision and update of knowledge sources [20]. This may result in data inconsistency across the analytics pipeline, causing consumers to experience a lack of trust regarding the analytics results. Consumers may be confused and may feel that “these numbers don’t make sense,” but it is difficult to understand the problem with the current analytics process and where it lies in the pipeline.

To address this limitation, we propose an analytics framework whereby data and visualization pipelines are integrated with a knowledge base. A knowledge base is a tool designed to store clinical knowledge content in a systematic way by managing attributes of content authorship, version, and relationship between resources [21,22]. We used a knowledge base as a key component to store CPGs, their adherence definitions, and their executable queries, so that the related documents in their different forms could be managed with metadata and easily shared by domain experts, query developers, and analysts. In addition, we connected data pipelines to executable queries in the knowledge base so that a change of query can be immediately incorporated into the analytics pipelines.

In a pilot study, we adopted locally developed CPGs used in our health care system, resulting in a care process model (CPM). We developed an integrated analytics framework that consists of a knowledge base that employs CPMs in different forms: a commercial data pipelining tool, Alteryx, connected to our clinical database and a commercial visualization tool, Tableau, to generate reports. We built a dashboard that provides views for adherence by physicians to 21 CPMs. Over an 18-month-long proof-of-concept project, we ran a working group to analyze CPM adherence and to manage definition revisions. We investigated how the use of the knowledge base contributed to the filling of knowledge gaps and automating of analytics pipelines with computational efficiency.

**Methods**

**Subject: Care Process Models**

In this study, we used CPMs that were created using locally developed CPGs in our health care organization and that were designed to reduce clinical variation, improve quality, and support local preferences [23]. The CPMs were published and managed by clinical programs, which are made up of clinical expert groups consisting of clinical champions, medical directors, nursing administrators, data managers, and data analysts. Over the last 20 years, clinical programs have developed over 120 CPMs that cover a variety of clinical conditions and procedures, such as hypertension, heart failure, breast cancer, appendicitis, and acute myocardial infarction, among others.

CPMs were originally developed as paper or electronic documents containing descriptions of target problems or procedures, logics of decision-making, and recommended actions. Traditionally, CPM implementation was conducted through the involvement and education of care teams and providers. Over the last four years, we have installed a new enterprise-wide electronic health record (EHR) system in our hospitals and clinics in Utah, USA. We have also started developing computationally executable CPMs inside the EHR system using a variety of decision-support components, including order sets, decision-support rules, and care pathways (ie, decision flow and state-based order recommendation tools), among others.

In addition to CPM implementation, there was a strong need from the clinical leadership to monitor and evaluate how providers are complying with CPMs. Analyzing such data may allow us to understand how well-embedded CPMs are within clinical practices and may allow us to gain insights into how to improve best practices within them. However, it has been challenging to quantitatively measure whether CPMs were used as intended after implementation, since key data points for the evaluation are complicated to define and capture. In addition, time-consuming manual data processing to calculate evaluation metrics was required. To address these problems, clinical programs and informatics specialists initiated an effort to build a framework that creates a systematic approach for data extraction and visual analysis within the evolution cycle of CPM development, implementation, and improvement.

To correctly evaluate adherence to CPMs, three types of information should be defined. First, since CPMs are developed to treat patients with certain conditions, a target population (ie, patient cohort) should be identified. Specifically, a combination of inclusion and exclusion criteria should be defined, including patient demographics, diagnosis, lab results, and medications. Second, metrics to quantitatively measure CPM utilization for the defined target population should be defined. A timeline of when to develop the metrics, typically key concepts extracted from logics and actions in CPMs, should be included since it is unrealistic to capture all of the concepts within the CPMs.
Examples of key concepts include the following: diagnosed with pneumonia, image ordered for d-dimer, or 1-90 days old with a fever (≥ 38°C). Third, an adherence formulation is needed that would be used to count credits given to providers or care units who utilized CPMs as intended for target patients.

Although CPMs are well-described guidelines, it is often challenging to derive the key information above, as the nature of the CPMs are composed of complicated, domain-specific knowledge and often contain ambiguity. In addition, it is difficult to connect the derived key concepts to data points—often tables or columns in clinical databases—in real-world data sources. Thus, a multidisciplinary collaboration effort may be required in order to transfer knowledge between the various experts below:

1. Domain knowledge expert: an expert, author, or publisher of CPMs. This individual would be responsible for interpreting and defining the highest levels of evaluation criteria.
2. Domain data expert: an individual with both clinical domain and database expertise. This individual would be responsible for translating CPM knowledge and linking concepts to data points in databases.
3. Informatics expert: an expert in clinical knowledge management. This individual would be responsible for communication between domains.
4. Database engineer: a database expert. This individual would be responsible for developing and maintaining database tables for CPM relationship information.
5. Data analyst: a database expert. This individual would be responsible for developing data pipelines for analysis.

**Problem Statement**

The problem we are addressing is that knowledge gaps often exist while translating logics CPMs from domain experts into analytics pipelines. For example, domain experts may define the diagnosis of pneumonia at a high level in the original CPMs. However, the disease should be clearly defined to extract real data from the clinical database, including diagnosis codes in standard terminologies, types of patient visits, clinical context, problem status, and exceptions. Gaps may exist while clarifying such information by mapping incorrect data points, miscommunication, misinterpretation, changes in CPM contents, and changes in data sources. Such inconsistency could result in confusion and distrust of data quality for analytics consumers. In the current analytics environment there is no tool to track and investigate what the gaps are and where they exist.

To address inconsistencies in knowledge translation and improve analytics productivity, we aim to adopt a knowledge base that will help us manage the content across the whole CPM evaluation process. A knowledge base is a component in our EHR system that manages the authoring, review, publication, delivery, and versioning processes that surround clinical knowledge assets for consumption within these frameworks. Typically, clinical knowledge sources may be order sets, decision support rules, nursing protocols, clinical guidelines, and patient education resources. Any knowledge-based content can be a source for a knowledge base, such as concept definitions, formulations, executable queries, and design concepts of visualizations. These assets can be consumed by a human (ie, in the case of narrative clinical care guidelines) or by a computer (ie, rule bases for consumption by inference engines), or they can be aligned for consumption on both ends of the spectrum [21].

There are practical benefits to using a knowledge base for CPM analytics: (1) storing and sharing all the knowledge resources in a centralized place, instead of local storage or exchanging content by emails, to reduce miscommunication and redundancy; (2) being able to manage knowledge resources with unique numeric identifiers and metadata, including author, version, and history of revisions; and (3) making knowledge resources consumable with data analytics tools. In addition, a knowledge base may be useful for end-user data consumers by giving them more contextual information about the data with answers to questions such as “Who defined the definitions of the measures?”, “How were the measures calculated?”, and “What changes have been made?”

### Selection of Care Process Models and Deriving Key Measures

As part of a pilot study to evaluate adherence based on the proposed framework, we selected 21 CPMs. Our selection criteria were as follows: (1) clinical utility and popularity of the CPM, (2) whether the CPM was already implemented within our EHR system, (3) the ease with which the CPM allowed the definition of evaluation metrics, and (4) whether data points related to the CPM were fully or partially collected in our clinical information systems. For the selected CPMs, we worked with CPM publishers in clinical programs to derive the three types of information for adherence evaluation defined in a previous section. The adherence rate (%), defined in the equation below, represents physician utilization of designated CPM components for a group of patients who are intended to be treated by the CPM:

\[
\text{Adherence} (\%) = \frac{\text{number of cases treated using the CPM}}{\text{number of cases intended to be treated by the CPM}}
\]

Since the definition above is highly abstracted, there were diverse details regarding how to practically calculate the adherence. For example, to calculate the numerators, we determined what types of decision-support components were used to implement the CPMs (eg, order set) and which were used to connect key concepts to real data points (eg, order set ordering history). We found that 10 of the 21 CPMs (48%) were implemented as physician order sets, or small groups of order sets, whereas the rest were implemented through combinations of care pathways and decision-support rules.
Table 1. Summary of care process model (CPM) adherence definitions.

<table>
<thead>
<tr>
<th>Clinical program (adherence definitions, n; queries, n)</th>
<th>CPM name</th>
<th>Cohort-defining condition</th>
<th>Decision-support component</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncology, 1, 3</td>
<td>Lung, breast, or colon cancer</td>
<td>Diagnosis and chemotherapy</td>
<td>Order set</td>
</tr>
<tr>
<td>Pediatrics, 1, 1</td>
<td>Febrile infant</td>
<td>Vital sign or a report</td>
<td>Care pathway</td>
</tr>
<tr>
<td>Neuroscience, 1, 1</td>
<td>Acute stroke</td>
<td>Problem and admission time</td>
<td>Order set</td>
</tr>
<tr>
<td>Primary care, 4, 4</td>
<td>Diabetes</td>
<td>Age and diagnosis</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Hypertension</td>
<td>Diagnosis</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Acute sinusitis</td>
<td>Diagnosis and medication</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>High blood pressure</td>
<td>Diagnosis</td>
<td>Care pathway</td>
</tr>
<tr>
<td>Cardiovascular, 1, 3</td>
<td>Heart failure</td>
<td>Age, admission, and diagnosis</td>
<td>Order set</td>
</tr>
<tr>
<td></td>
<td>Acute myocardial infarction</td>
<td>Age, admission, and diagnosis</td>
<td>Order set</td>
</tr>
<tr>
<td></td>
<td>Coronary artery bypass graft</td>
<td>Age, admission, and diagnosis</td>
<td>Order set</td>
</tr>
<tr>
<td>Musculoskeletal, 1, 1</td>
<td>Total hip or knee surgery</td>
<td>Procedure</td>
<td>Order set</td>
</tr>
<tr>
<td>Surgical service, 1, 2</td>
<td>Appendicitis</td>
<td>Procedure</td>
<td>Order set</td>
</tr>
<tr>
<td></td>
<td>Cholecystectomy</td>
<td>Procedure</td>
<td>Order set</td>
</tr>
<tr>
<td>Behavioral health, 3, 3</td>
<td>Depression</td>
<td>CDS(^a) rule and diagnosis</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Suicide prevention</td>
<td>CDS rule</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Mental health integration</td>
<td>Clinical document and clinic visit</td>
<td>Care pathway</td>
</tr>
<tr>
<td>Women and newborn, 1, 2</td>
<td>Jaundice</td>
<td>Newborn</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Neonatal hypoglycemia</td>
<td>Observation, rules, and age</td>
<td>Care pathway</td>
</tr>
<tr>
<td>Intensive medicine, 4, 4</td>
<td>Pneumonia</td>
<td>CDS rule</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Pulmonary embolism</td>
<td>Order, imaging, diagnosis, and care pathway use</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Pediatric sepsis in emergency</td>
<td>CDS rule</td>
<td>Care pathway</td>
</tr>
<tr>
<td></td>
<td>Pediatric minor head trauma</td>
<td>CDS rule</td>
<td>Care pathway</td>
</tr>
</tbody>
</table>

\(^a\)CDS: clinical decision support.

The denominators were used to identify CPM-specific patient cohorts. For example, the denominator for the adherence rate of the hypoglycemia CPM was defined as a combination of (1) whether certain decision-support rules related to hypoglycemia were fired, (2) whether a patient has a history of hypoglycemia according to the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10), and (3) whether specific nursing documents were recorded.

Below are examples of detailed key concepts to be included in the definitions:

1. Cohort (denominator): clinical inclusion and exclusion criteria, patient master index, encounter, facility, care unit, decision-support rule firing criteria, clinical form used, etc.
2. Utilization (numerator): Order set usage, order set title, version, content, orderable items, customized order sets, care pathway title, components in care pathway used, etc.
3. Adherence: adherence definition, aggregation level by patient, encounter, provider, unit and facility, etc.

A group of experts with different backgrounds reviewed these concepts, documented the definitions, and built queries. Table 1 shows a summary of the derived CPM adherence definitions.

Results

Development of Framework

We developed the analytics framework by integrating several homegrown and commercial tools. For the knowledge base, we used Intermountain’s knowledge repository that our organization has used over the last 10 years. We adopted Alteryx as a tool for data analytics pipelines and Tableau as a tool for visual transformation and Web-based dashboard development. We used our enterprise data warehouse (EDW) as the data source, which is an integrated clinical data repository for research and quality improvement that stores over 100 billion records tied to encounters, lab observations, diagnoses, procedures, medications, and billing over 20 years. We also used CPM-centered database tables in our EDW that were developed by clinical programs, which store condition-specific patient cohorts, quality metrics, and outcomes for clinical research and quality improvement.

Figure 1 depicts the architecture and flow of knowledge in the framework. Domain experts in clinical programs translate paper-based CPMs to adherence definition documents. Structured Query Language (SQL) developers create queries
based on the documents and deploy them into the knowledge base on the Web (see Figure 2). Alteryx imports the SQL information from the knowledge base to run against the EDW; Alteryx then exports the cohort and metrics as a Tableau-specific data file to a Tableau server so they can be visually transformed by Tableau to automatically generate charts on the Web. In the dashboard, links are embedded in the charts, enabling users to view the original CPM documents and queries used in the knowledge base.

Figure 3 shows a CPM in various forms during knowledge transformation: human-readable PDF document (top left); logics of the CPM are implemented as an order set in a computerized physician order entry (CPOE; bottom left); adherence definition document (top right); and computerized adherence logic as a SQL (bottom right). Since these knowledge content components in different forms originated from one source, they are semantically linked to each other with authorship, translation record, and version history.

Figure 1. Architecture diagram of the knowledge-based analytics framework for care process model (CPM) adherence. EDW: enterprise data warehouse; SQL: Structured Query Language.
Development of Dashboard

We developed a Tableau-based dashboard in a production environment that is accessible to clinicians and researchers in our organization through secured user access. It employs five detailed views for each representative CPM adherence in different contexts and scales.

1. Main view: this view provides an overview of adherence trends at the highest level. It consists of a bar chart representing the average cumulative rates of CPM utilization and a line chart showing the monthly average over time (see Figure 4). Each CPM is marked with different color.

A filter to the right of the bar graph narrows down the data by facility. Average adherence rates range from 0% to 100%, with 0% indicating that no physician used any CPM for the specific condition-based cohort and 100% indicating that all physicians used CPMs for all associated relevant cases.

2. Facility view: this view is used to monitor the monthly average adherence rate by hospitals. This view consists of a dual chart that represents the percent adherence rate by month and the number of encounters where CPM was used or not used.

3. Provider view: this view shows the summary of CPM utilization by individual providers. Users can select a
Data Extraction and Early Usage Pattern

Data pipelines were scheduled to run regularly, with different refresh frequencies depending on CPMs and data sources. As of July 1, 2018, the number of patients eligible for 21 CPMs was 230,669 and the number of encounters was 377,507. For those target patients, 7895 providers utilized CPMs at least once. Total adherence rate across all the CPMs during the period was 8.8%.

Figure 4. Main view of the Tableau–based care process model (CPM) utilization dashboard (screenshot).

Figure 5. Provider view of the care process model (CPM) utilization dashboard (screenshot).

<table>
<thead>
<tr>
<th>Provider</th>
<th>Position</th>
<th>Avg Utilization (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen J. DO, BREN T J. (Physician · Surgeon)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary
- Patient: 83
- Hospital encounter: 82
- Total encounter: 82

Monthly Utilization Chart
- 2017/01: 10%
- 2017/02: 15%
- 2017/03: 20%
- 2017/04: 25%
- 2017/05: 30%
- 2017/06: 35%
- 2017/07: 40%
- 2017/08: 45%
- 2017/09: 50%
- 2017/10: 55%
- 2017/11: 60%
- 2017/12: 65%
- 2018/01: 70%
- 2018/02: 75%
- 2018/03: 80%
- 2018/04: 85%
- 2018/05: 90%
- 2018/06: 95%
- 2018/07: 100%
We analyzed user sessions of the dashboard using a Tableau system monitoring report, as represented in Figure 6. Test and administration users were excluded from the analysis. Overall, the usage pattern was stable with minor seasonal effects and spikes.

**Data Provenance: Tracking Revisions of Care Process Model Adherence Definitions**

Since implementation, we set up monthly meetings involving clinical programs, CPM implementation leadership, data analysts, and knowledge engineers. These meetings were meant to (1) monitor CPM adherence data, (2) review current adherence definitions and discuss room for improvement, and (3) discuss ways to encourage providers to use CPMs. During the pilot study period, a number of revisions to the definitions were made (see Table 2). Many of the revisions included amendments to the definitions, while some included the addition of new data sources to the EDW or resolving data-quality issues. Definitions were updated with revisions in the knowledge base and the links to the Alteryx pipelines were automatically renewed.

As seen in the bottom portion of Figure 4, the adherence rate steadily increased since implementation. We believe there are two practical reasons for this increase. One is that clinical programs have encouraged their clinicians to utilize CPMs, through CPOE training, physician education, etc. The other reason is that some clinical programs revised their adherence definitions. For example, the oncology program expanded CPM-designed order sets, which resulted in an increase of the numerator. The intensive medicine program revised the definitions of their target patients by narrowing them down to specific facilities, which resulted in a decrease of the denominator.

**Discussion**

**Comparison With Prior Work**

Several studies have used knowledge engineering tools for translation of CPG logics to implement them into practice for the purpose of knowledge standardization and computational automation [4-9]. Unlike in those studies, we used knowledge engineering tools for managing knowledge transformation within analytics processes. Compared with prior work, our original results included (1) validating the usefulness of knowledge management tools within analytics processes, (2) validating the feasibility of integrating knowledge management tools and an analytics framework, and (3) demonstrating the proposed approach using empirical clinical data from local EHR systems.
Limitations

Although the main contribution of this study is the use of knowledge management, we did not quantitatively analyze the improvement in the consistency of knowledge in transformation or the productivity of analytics pipelines. Rather, we demonstrated it qualitatively, including determining which functionalities of the knowledge base were able to support the consistency of CPM-related knowledge in the development and maintenance phases. We will conduct further analyses in the future as we collect additional data. This will include adding more CPMs with complex clinical settings and practices, including chronic conditions and comorbidities that span multiple encounters, locations, and providers.

In this study, we simplified the adherence definitions to include whether a designated CPM component is used or not used for a target patient, although there may be many variations. We will continue to add more detailed definitions of adherence, including how specific order items or content within CPM components are used. By doing so, we can investigate the mechanism of CPM utilization (ie, use of standard clinical guidelines) and how this can change patient care or clinical workflow.

Conclusions

This case study demonstrated that the proposed analytics framework could accommodate complicated knowledge management and data pipelining for CPM evaluation using a knowledge base, while maintaining computational efficiency. It is expected that the benefits of using a knowledge base will be more significant as we add complicated clinical guidelines into the analytics framework.

Authors' Contributions

JL developed the framework, built the dashboard, performed data analysis, and designed the study. NCH helped design the study, provided knowledge base consultation, collaborated with clinicians, and validated the feasibility of the architecture and the results of data analysis. The authors do not have any funding or financial support associated with this project.

Conflicts of Interest

None declared.

References


Abbreviations

CDS: clinical decision support
CPG: clinical practice guideline
CPM: care process model
CPOE: computerized physician order entry
EDW: enterprise data warehouse
EHR: electronic health record
ICD-10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems
SQL: Structured Query Language
information, a link to the original publication on http://biomedeng.jmir.org/, as well as this copyright and license information must be included.