Preparing the Way for Routine Health Outcome Measurement in Patient Care

Paterson, Grace I.; Zitner, David.

Medical Informatics, Dalhousie University.

Halifax, NS B3H 4H7

e-mail: grace.paterson@dal.ca

Keywords: Health Status; Health Outcomes; Electronic Medical Records; UMLS.

Abstract

In order to interpret the effectiveness of health care activities, it is necessary to measure health status before and after an intervention. From the literature, we can identify outcome measures used in research. Online resources, such as the Cochrane Library and British Medical Journal, use a structured abstract so outcome measures can be readily identified and catalogued. These literature resources strive for relevance by promoting the assessment of healthcare interventions using outcomes that matter to people making choices about health care.

Both general and condition-specific measures are used to collect outcomes information. General measures apply to all systems and are more likely to capture adverse events. Condition-specific measures are more likely to capture benefits of care but could also capture adverse consequences. Patient-centered outcomes measure function, comfort, and likelihood of survival. Measures that are proxies for the patient include objective clinical data, health care activities, and caregiver activities. From a knowledge base of outcome measures, one can generate prompts for health outcomes information to record on the patient record.

Most organizations are able to use some measures to show they provide efficient and effective care. In health care, we do not collect and provide timely, regular and reliable information about either access to care or the results of health care activities. Consequently, we cannot properly allocate resources or manage care. However, from routinely collected patient health status and outcomes of care data, we can generate new knowledge about effective health care. This information can be fed back to caregivers, administrators, and researchers for continuous quality improvement.

1. Introduction

The purpose of health care activities is to maintain or improve health status. The dimensions of health status are comfort, function, and life expectancy. This study examines to what extent the health care research community used improved comfort and function and increased life expectancy as outcome measures in original research and systematic reviews of health care interventions. A recent report quantified the economic value of health gains from medical research [1], and suggests a way to view the benefits of health that communities may wish to consider when allocating scarce resources to health care.

Transfer and uptake of knowledge from the research community to the practice community can be improved through better integration of knowledge and practice. This project sought to catalog and classify the outcome measures recorded in the published research. Through examination of the information captured by researchers as most pertinent for measuring outcomes, we can identify a minimal set of data elements that should be collected by clinicians treating a condition.

In order to do outcomes research in the community setting, we need to harness routinely collected outcomes from primary care physicians and adjust for baseline and case mix variables to determine real world effectiveness of an intervention. This information can be captured in the electronic medical record and aggregated to learn about changes in health associated with care and the link between activities and results of care. We need to ensure we have information technology smart communities in order to improve continuity of care.
In 1994, the Federal/Territorial/Provincial Deputy Ministers of Health made the unanimous commitment that quality of care will be ensured by ongoing monitoring and publication of patient outcomes as changes to the health system are implemented and that evaluation and monitoring of outcomes be integral parts of utilization management [2]. This information is not routinely available now, and Dr. Zitner, an author of the 1994 report, concluded in a 1999 report, *Operating in the Dark: The Gathering Crisis in Canada’s Public Health Care System* [3], that a lack of information gathering has made it difficult to evaluate the system's performance.

Governments have been unable to understand the value of health care expenditures because they have had no way of knowing how many people are improved, harmed or have no change in health following various interventions. Provincial medicare plans have oriented themselves to paying provider claims, rather than generating information that might help in planning and appraising health care. Comparisons of outcomes using administrative datasets are fraught with confounders, and it appears that national health insurance has done nothing to create greater homogeneity in practice styles [4].

1.1 Study Objectives

- to identify the outcomes measures reported in the medical literature of the effects of health care for a range of conditions
- to catalog the variety of condition specific measures that are used for outcomes and identify those which are generic across conditions
- to categorize details of an outcome measure according to what it was measuring and what sign, symptom or finding was required for the measure
- to develop a knowledge engineering methodology that will help define the minimal data set for electronic health records which are meant to be useful for outcome measurement

2. Methodology

2.1 Data Collection

The journal collections were the Cochrane Library of Systematic Reviews, British Medical Journal, and Canadian Family Physician Critical Appraisal reviews. A systematic review is an "application of scientific strategies that limit bias to the systematic assembly, critical appraisal and synthesis of all relevant studies on a specific topic" [5]. All collections use a structured abstract to systematically disclose the main outcome measurements. For manual indexing of outcomes, the full paper and its references were used; for automatic indexing, the structured abstract was used.

The set of conditions for the literature review was developed using a focus group approach. The clinicians, informaticians, and medical students sought a set of conditions where care could be optimized through the increased use of proven therapies. This knowledge was supplemented with an in-house analysis by our Population Health Research Unit of the top reasons that patients use the health care system and a report of the top five diagnoses by inpatient clinical service in a tertiary care hospital.

The medical students read the literature to identify the outcome measures used, located the instruments used for rating scales and questionnaires, and entered the outcomes information into the database. The outcomes information was composed of the condition, specific outcomes, the measurement instrument, each item in the instrument, the signs, symptoms and findings associated with each specific outcome or item in an instrument, and the classification of the outcome measure.

A controlled vocabulary was used during data entry. The medical students logged onto the Unified Medical Language System (UMLS) Knowledge Server to retrieve the unique concept identifiers for conditions and signs, symptoms and findings. A natural language processing system was implemented to determine if we could speed up this process using automatic indexing of concepts in the outcome measures section of an article abstract.

2.2 Catalog and Categorization of Outcome Measures
The system developed a catalog of general and condition-specific health outcome measures. The articles were from evidence-based sources guided by editorial policies that ensure that a reference to the definitive version of the instrument was cited in the references and generally available in published literature. One of the guidelines for developing health measurements is that "the definitive version of the instrument must be made readily available to users" [6]. We did not purchase any measurement instruments.

General health status and quality of life outcome measures were identified as those that were applied across conditions. The system captured information on the details of a measurement instrument, so our analysis could determine the signs, symptoms, and findings gathered across conditions.

A scheme for classifying the outcome measures recorded by researchers was developed. For each question in a measurement instrument, the associated sign, symptom or finding was categorized as a direct patient measure or as a proxy measure for the patient. The dimensions for direct patient measures were function, comfort, and life expectancy. Proxy measures were subdivided as physiological function, patient-oriented proxy, and activity-oriented proxy. In some cases, the outcome measure was categorized as contextually ambiguous. The worksheet used for inter-rater reliability testing is shown in Figure 1.

<table>
<thead>
<tr>
<th>Classification of Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument ID: ____________________________</td>
</tr>
<tr>
<td>Direct Measures: Function, Comfort, LifeSpan, Physiological</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question</th>
<th>Function</th>
<th>Comfort</th>
<th>LifeSpan</th>
<th>Physiological</th>
<th>Proxy</th>
<th>Contextually Ambiguous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Number</td>
<td>S/S/ UML S Code</td>
<td>Physical Function</td>
<td>Mental Function</td>
<td>Social Function</td>
<td>Physical Comfort</td>
<td>Mental Comfort</td>
</tr>
</tbody>
</table>

FIGURE 1. Worksheet for Categorization of Health Outcome as Patient-Oriented or Proxy Measure

2.3 Collaborative Approach to System Development

We gathered together a team of medical and computer science faculty and students to design and implement the system. Technical reports and bi-weekly project plans were shared with our principal sponsor, the Global Information Networking Institute (GINI). Additional sponsorship came from the Canadian Cochrane Network Dalhousie Site.

Generally, the people developing information systems are divorced from those who actually use them. In our setting, the medical students worked in close proximity with the computer science students during system development and implementation. The computer science students were enrolled in a co-op program, so had an opportunity to build on their prior work experience. The medical students had been taught a reasoning process called case-oriented problem-stimulated (COPS) learning in a tutorial setting, and were able to seek out the information and enter it into the computer database with minimal training.

Medical informatics is multidisciplinary and uses collaborative approaches to solve problems. Members of the team had training and experience in informatics, family medicine, medical quality assurance, health services administration, health services accreditation, population health, medical education, computer science, information retrieval, and communication technologies.

Methods used for system development were entity-relationship data model, relational schema for Oracle database management system (DBMS), web-based data entry, MS Access Reports, and knowledge discovery using Rosetta,
rough sets analysis. The programming languages included Perl, Java Database Connectivity, JavaScript, SQL, and DBI.

3. Results

3.1 Identification of Outcome Measures

Thirty-three systematic reviews from the Cochrane Library of Systematic Reviews 1999 Edition Version 2 were collected from online version, http://cochrane.kfinder.com. Forty-seven papers published by the BMJ Publishing Group and listed as collected resources for a condition were included from online version, http://www.bmj.com. Eight critical appraisal reviews from the Canadian Family Physician were included from online version, http://www.cfpc.ca. In addition, we included articles published on Dalhousie-based health outcomes research in asthma and cardiovascular disease. The total number of articles was 100. A subset of articles addressed co-morbid conditions.

3.2 Catalog of Condition-Specific and General Measures

There were 225 specific outcome measures and 20 multiple-item outcome instruments ranging from the 136-item Sickness Impact Profile to the 5-item Health Practices Index. Five of the multiple-item scales were generic, so items were not associated with a condition in Table 1. The text of each item was entered into the system and associated sign, symptoms and findings UMLS-coded.

<table>
<thead>
<tr>
<th>Condition (n=11)</th>
<th>Multiple-item Instruments</th>
<th>Outcome Items</th>
<th>Sign, Symptom, Finding % in File</th>
<th>% UMLS coded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asthma</td>
<td>4</td>
<td>203</td>
<td>25.3</td>
<td>29.1%</td>
</tr>
<tr>
<td>Cardiovascular Disease</td>
<td>1</td>
<td>7</td>
<td>7.9</td>
<td>100%</td>
</tr>
<tr>
<td>Ulcerative Colitis</td>
<td>5</td>
<td>5</td>
<td>0.5</td>
<td>80%</td>
</tr>
<tr>
<td>Crohn's Disease</td>
<td>4</td>
<td>43</td>
<td>6.5</td>
<td>95.3%</td>
</tr>
<tr>
<td>Depression</td>
<td>5</td>
<td>97</td>
<td>18.7</td>
<td>96.9%</td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td>69</td>
<td>13.5</td>
<td>91.3%</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td></td>
<td>15</td>
<td>2.9</td>
<td>86.7%</td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
<td>13</td>
<td>1.4</td>
<td>92.3%</td>
</tr>
<tr>
<td>Inflammatory Bowel Disease</td>
<td>1</td>
<td>35</td>
<td>6.5</td>
<td>93.1%</td>
</tr>
<tr>
<td>Ischemic Heart Disease</td>
<td></td>
<td>142</td>
<td>16.7</td>
<td>82.4%</td>
</tr>
</tbody>
</table>

Table 1. Conditions, Instruments, Outcome Items and their Signs, Symptoms and Findings

We developed an automatic indexing system that mapped concept strings in the "main outcome measures" paragraph of a structured abstract to the UMLS concept unique identifiers (CUI). For a set of 33 Cochrane articles, there are 441 terms and related CUIs retrieved from automatic indexing and 143 terms from manual indexing. Only 39 terms reside in automatic and manual indexing [7]. Only 40% of the specific and multiple-item outcome measures in our catalog could be coded to UMLS concepts. Some factors cause it. First, an outcome vocabulary is not available in UMLS so that the natural language processing is not able to perform as an effective an index as people do. Second, some terms generated from automatic indexing are senseless but they cannot be removed properly.

The UMLS is an applied work—it seeks to develop content and applications for a stated objective. As such, there are two areas for evaluation: (1) the quality of the UMLS content around outcomes vocabulary, and (2) the ability of the natural language application to link to UMLS content for text strings [8].

3.3 Categorization of an Outcome Measure as Patient-Oriented or Proxy Measure

Patient-oriented outcomes accounted for 55% of measures (comfort 38%, function 4%, life expectancy 3%). Proxy measures accounted for 44% of measures (physiological function 16%, patient-oriented proxy 21%, activity-oriented proxy 7%). Only 1% of outcomes were categorized as contextually ambiguous.

3.4 System Development and Knowledge Engineering Methodology
The medical and computer science students were supervised by medical faculty, attended bi-weekly project meetings with the sponsor of the project, the Global Information Networking Institute (GINI), and prepared technical reports. Evaluation at the end of the project indicated that the team achieved a shared understanding sufficient to reach a collaborative consensus around the data model, reports, and classification system. The students co-op reports dealt with related topics that needed further study [9,10].

The initial data entry system was cumbersome, and the medical students requested that computer science students address the need for speed. Our GINI sponsors suggested we test one of the new technologies, a JDBC link to Oracle Database Management System. This was accomplished using collaboration between our computer science students and the technical support community on the web.

The system includes a knowledge engineering methodology that was tested to determine if a minimal data set for an electronic health record would emerge using a rough sets data analysis of the signs, symptoms, and findings used for outcomes measurement. The dataset was too sparse for the analysis.

4. Discussion

This project provided an opportunity to understand the "information gap" between what research identifies as measurable results of care, and what administrators gather to manage the health care system. The community health boards need better information to support decision making around the benefits of care, and clinicians need to know what information should be routinely gathered to support outcomes measurement. Community physicians take a general measure of the patient's health status at each visit, but this is not routinely recorded, so we do not have a way to aggregate the results of care and link the activities of care to the results of care. It has been stated that the universal, perfect index for health measurement can never exist [6]. Clinicians in our Queen Elizabeth II Health Sciences Centre have agreed to state the outcome of care on the discharge summary report that is sent to the community physician. Cardiology and Internal Medicine now measure health status along the dimensions of comfort and function.

5. References