Cognitive Complexity of the Medical Record
Is a Risk Factor for Major Adverse Events

David Roberson, MD, FACS; Michael Connell, EdD; Shay Dillis, RN, MSN; Kimberlee Gauvreau, PhD; Rebecca Gore;
Elaina Heagerty, MPH; Kathy Jenkins, MD, MPH; Lin Ma, MS; Amy Maurer, MA; Jessica Stephenson; Margot Schwartz, MPH

Abstract
Context: Patients in tertiary care hospitals are more complex than in the past, but the implications of this are poorly understood because “patient complexity” has been difficult to quantify.

Objective: We developed a tool, the Complexity Ruler, to quantify the amount of data (as bits) in the patient's medical record. We designated the amount of data in the medical record as the cognitive complexity of the medical record (CCMR). We hypothesized that CCMR is a useful surrogate for true patient complexity and that higher CCMR correlates with risk of major adverse events.

Design: The Complexity Ruler was validated by comparing the measured CCMR with physician rankings of patient complexity on specific inpatient services. It was tested in a case-control model of all patients with major adverse events at a tertiary care pediatric hospital from 2005 to 2006.

Main Outcome Measures: The main outcome measure was an externally reported major adverse event. We measured CCMR for 24 hours before the event, and we estimated lifetime CCMR.

Results: Above empirically derived cutoffs, 24-hour and lifetime CCMR were risk factors for major adverse events (odds ratios, 5.3 and 6.5, respectively). In a multivariate analysis, CCMR alone was essentially as predictive of risk as a model that started with 30-plus clinical factors.

Conclusions: CCMR correlates with physician assessment of complexity and risk of adverse events. We hypothesized that increased CCMR increases the risk of physician cognitive overload and automated version of the Complexity Ruler could allow identification of at-risk patients in real time.

Introduction
Patients in tertiary care hospitals are more complex than in the past, and complexity varies greatly between patients. Because there are finite limits on the brain’s ability to assimilate and process information, greater patient complexity increases the risk of physician cognitive overload and medical error. Therefore, if one could quantify patient complexity, it is possible that this would provide a useful risk stratification tool.

Cognitive overload is a risk factor for cognitive failure in many domains, and techniques have been created in other fields to quantify complexity. Although there is no perfect method of quantifying complexity, one widely used operational definition is “the minimum amount of information required to fully describe something.” Conceptually, one writes down the most parsimonious but complete description of something (eg, a car, a patient) and counts the number of bits required to store that description.

It is obviously impossible to write a perfect description of a patient and his/her illness or illnesses. We therefore hypothesized that the amount of data recorded in the medical record, which we designated the “cognitive complexity of the medical record” (CCMR), would be a useful surrogate for true patient complexity. Our testable hypothesis was that patients with greater CCMR are at higher risk of adverse events.

The algorithm for calculating CCMR is complex, but conceptually it is fairly simple. We count 1 bit for every character of English language text, and 3 bits for every individual numeric digit recorded in the record. (Although it takes 8 bits to store an English letter in ASCII code, English-language text is highly constrained, and thus 1 bit per character is a reasonable estimate of text complexity.) For images, which require artificially inflated storage space (every chest film in our institution is stored in a 5-megabyte file), we count the number of characters in the dictated report. For headings on standardized charts (eg, “blood pressure”), we count the heading once per 24 hours.

To the best of our knowledge, no prior studies have attempted to quantify the data stored in the medical record, and there is thus no gold standard for CCMR. We therefore tested our tool by 1) comparing its ranking of CCMR to the complexity rankings of experienced clinicians and 2) testing its ability to predict major adverse events in a tertiary care pediatric hospital.
Cognitive Complexity of the Medical Record Is a Risk Factor for Major Adverse Events

Methods

Development of the Complexity Ruler

As discussed in the introduction, we used an estimate of one bit per character for English-language text and 3 bits per digit for numeric data. We counted redundant headings once per 24 hours, and counted the dictated reports rather than the data storage requirements of actual images. We included all data entered into the record, whether by nurse, physician, or other clinician, and included laboratory results, x-ray films, and any other information of any kind. The result is a single number to represent the CCMR for each patient during a given time period. The instructions for the full Complexity Ruler are available online at: www.thepermanentejournal.org/files/Winter2014/cr.pdf.

Face Validity

Three senior physicians rank ordered the complexity (not the acuity) of patients on four different inpatient medical services. We measured the CCMR of five randomly selected inpatients from each of these four services and estimated their lifetime CCMR. We hypothesized that agreement between the Complexity Ruler and expert clinicians would constitute face validity of the Complexity Ruler.

Cognitive Complexity of the Medical Record versus Major Adverse Events

We conducted a case-control evaluation of the 39 patients who had major adverse events at Boston Children's Hospital, Boston, MA, in 2005 and 2006. A major adverse event was defined as an event of sufficient gravity to require reporting to an outside regulatory agency. Most of these events involved permanent injury to the patient or major complications, for example, unanticipated abdominal or cardiac surgery. Controls were 78 patients who were randomly selected from all admissions to Boston Children's Hospital during 2005 to 2006.

24-Hour versus Lifetime Cognitive Complexity of the Medical Record

The 24-hour CCMR was measured in the 24 hours before the adverse event in the cases (eg, from 3 am the day before to 2:59 am if the event occurred at 3 am). Each control was randomly assigned a date and time of his/her “event,” and 24-hour CCMR was calculated for the 24 hours before the event. Cases and controls whose event occurred in the first 24 hours after admission were excluded from the 24-hour CCMR analysis.

The current Complexity Ruler is too labor-intensive to apply to long periods of time. We therefore estimated the lifetime CCMR by determining their lifetime number of inpatient hospital days and multiplying it by the average CCMR of an inpatient day, stratified by days in the intensive care unit (ICU) vs non-ICU days. All cases and controls were included in the analysis of lifetime CCMR.

Data Analysis

Mean CCMR was compared for cases (reportable event) and controls (no reportable event) using the 2-sample t test with unequal variances. Patient characteristics, admission characteristics, and nursing factors were compared for cases vs controls using the Wilcoxon rank sum test for continuous variables and Fisher exact test for categorical variables. Factors significant at the 0.10 level were included in logistic regression models predicting major adverse events; the odds ratio (OR) and 95% confidence intervals were estimated. The CCMR was dichotomized at a value that maximized the OR for predicting major adverse events.

The c statistic was calculated for each logistic regression model. The c statistic is a measure of the model's ability to discriminate between patients who experienced a major adverse event and those who did not; a value of 0.5 is no better than random, whereas a value of 1.0 means the model predicts outcome perfectly. Multivariate analysis was performed using forward selection; a p value ≤ 0.05 was required for a variable to be retained in the final model. Statistical analyses were performed using SAS Version 9.2 (SAS Institute Inc, Cary, NC).

Figure 1. Mean 24-hour cognitive complexity of the medical record (CCMR) measured by the Complexity Ruler compared with physician rankings of complexity.

CCMR was measured from an average of 5 randomly selected patients on each of 4 inpatient medical services.

MICU = Medical Intensive Care Unit.

Figure 2. Mean 24-hour cognitive complexity of the medical record (CCMR) and estimated CCMR for the preceding 3 years.

Data are for the same patients shown in Figure 1. The average intensive care unit patient had more than 3.5 million bits over the previous 3 years—a staggering amount of information for the human brain to assimilate.

MICU = Medical Intensive Care Unit.
**Results**

**Face Validity**

All 3 senior physicians ranked patients on the short stay service as least complex and patients in the ICU as most complex. Two of the 3 physicians ranked cardiology patients as more complex than general pediatrics patients, and 1 reversed this order. Both 24-hour and lifetime CCMR agreed with the rankings of experienced clinicians (Figures 1 and 2).

**Lifetime Cognitive Complexity of the Medical Record versus Major Adverse Events**

All 39 cases and 78 controls were included in this analysis. The mean estimated lifetime CCMR was much higher for cases (736,033 bits) than controls (119,707 bits, \(p < 0.001\)). Having more than 70,000 bits in the lifetime medical record was associated with an OR of 6.5 for major adverse events \((p < 0.001)\). Compared with more than 30 other risk factors, lifetime CCMR was 1 of 16 statistically significant risk factors for major adverse events and had the third highest OR of these factors (Table 1).

In a multivariate analysis, lifetime CCMR was an independent predictor of adverse events. Patients with a lifetime CCMR greater than 70,000 bits had an OR of 7.7 for adverse events \((p < 0.001; Table 2)\). We compared the discrimination of the univariate model used to evaluate lifetime CCMR (c statistic of 0.72) with the goodness of fit from a multivariate analysis that excluded CCMR \((c\text{ statistic of 0.78})\).

**24-Hour and Lifetime Cognitive Complexity of the Medical Record versus Major Adverse Events**

We hypothesized that 24-hour CCMR and lifetime CCMR might have different predictive values for major adverse events. However, for the analysis of 24-hour CCMR, we could evaluate only those patients whose event happened at least 24 hours after admission and, thus, for whom 24-hour CCMR could be measured (17 cases and 42 controls). We therefore analyzed this group of patients including both 24-hour CCMR and lifetime CCMR as potentially independent predictive variables. Because the datasets are not identical, the lifetime CCMR results are slightly different than in the analysis of all patients (Table 1).

Mean 24-hour CCMR was higher for cases (31,323 bits) than controls (14,454 bits), a significant difference \((p = 0.008)\). Inclusion of more than 15,000 bits in 24 hours was associated with an OR of 5.3 for a major adverse event \((p = 0.008)\). Compared with more than 30 other risk factors for adverse events, both 24-hour and lifetime CCMR were among the 10 statistically significant risk factors, and lifetime CCMR had the second highest OR of all factors (Table 3).

In a multivariate analysis after controlling for other factors, 24-hour and lifetime CCMR were not statistically significant. The c statistic for goodness of fit for the most predictive multivariate analysis was 0.87. When 24-hour CCMR was forced into the model, greater than 15,000 bits in 24 hours had an OR of 9.7 for a major adverse event \((p = 0.02, c\text{ statistic = 0.86})\). When lifetime CCMR was forced into the model, greater than 70,000 bits had an OR of 19.3 for a major adverse event \((p = 0.02, c\text{ statistic = 0.88; Table 4})\).

**Discussion**

Many authors have commented on the phenomenon of increasing patient complexity and hypothesized that complexity is a risk factor for errors and major adverse events. However, the impact of complexity on safety and quality has been difficult to study because it has not been quantifiable. We hypothesized that the data stored in the medical record (what we have designated the CCMR) would be a useful surrogate for true patient complexity and would therefore predict the risk of major adverse events.
Cognitive complexity is not the same as medical acuity. Although very sick patients will often be very complex, there are patients with cognitively straightforward problems (e.g., carbon monoxide poisoning) who are critically ill, and patients with very complex illnesses (long-standing diabetes) who are—at the moment—physically well.

**Robust Prediction of Risk**

Using a number of assumptions, we developed a tool to measure CCMR.\(^8\) We demonstrate in this project that both 24-hour and lifetime CCMR are risk factors for major adverse events. Naturally, our tool is preliminary and is subject to revision and improvement. However, even in its first iteration, it was very robust at predicting the risk of major adverse events. Most importantly, use of either 24-hour or lifetime CCMR alone was essentially equal in predictive power as multivariable models built on more than 30 clinical factors. Also important, use of 24-hour or lifetime CCMR alone was nearly as useful in predicting the risk of major adverse events as was a multivariable analysis starting with more than 35 clinical variables.

To implement risk reduction strategies, high-risk patients must be identified in real time, and risk stratification systems that incorporate multiple clinical variables are inherently difficult to implement in real time. On the other hand, CCMR could be readily “built into” the electronic medical record and be available continuously in real time.

If patients with high CCMR were identified in real time, potential risk-reduction strategies could be tested. Since cognitive capacity increases with experience,\(^3\) one hypothesis is that very complex patients should be managed by physicians with greater seniority. More complex patients might be provided with additional physician or nurse staffing. Determining which risk reduction strategies might be effective is far beyond the scope of this study, but it is impossible to test strategies scientifically unless one can identify the patients at high risk.

**Limitations**

The Complexity Ruler and this project both have many important limitations. Since there is no gold standard for cognitive complexity, we had to construct the Complexity Ruler de novo using multiple assumptions. Although the first version of the Complexity Ruler is predictive of adverse events, we are certain that it can be modified and improved. For example, different weightings of English-language text and digits might more accurately predict the likelihood of a major adverse event.

The Complexity Ruler measures the total CCMR of the chart, including physician notes, laboratory tests, radiographs, nursing notes, flowsheets, and all other items. We did measure the different sections of the medical record separately but did not find any increased predictive value from looking at these as separate variables (data not shown). This could simply reflect the small size of our sample. Future studies will examine whether there are specific sections of the record whose CCMR is more predictive of major adverse events.

Because the Complexity Ruler measures the amount that clinicians write in the chart, it is possible that it overestimates the CCMR of patients whose clinicians are more verbose. This is a particular concern because clinicians can “cut and paste” sections of previous notes into current notes in the electronic medical record, increasing the amount of text without adding new information. There are computational tools to reduce or eliminate the redundancy from text (“compression algorithms”); therefore, it will be possible in future studies to determine whether verbosity and repetition are significant issues when measuring CCMR.

These data demonstrate that high CCMR is a risk factor for major adverse events but do not establish a causative link. We believe that the mechanism is cognitive overload of the physician, but additional studies will be needed to test that hypothesis.

Because we studied only very serious adverse events, we do not know if CCMR as measured by our tool is a risk factor for less serious adverse events.

### Table 3. Univariate analysis: 17 cases and 42 controls*\(^a\)

<table>
<thead>
<tr>
<th>Clinical factor</th>
<th>Odds ratio</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>70,000 Bits of lifetime CCMR</td>
<td>16.7</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>≥ 1 G-tube medication</td>
<td>19.2</td>
<td>0.02</td>
</tr>
<tr>
<td>≥ 1 Nebulized medication</td>
<td>15.5</td>
<td>0.007</td>
</tr>
<tr>
<td>≥ 4 Medications</td>
<td>11.0</td>
<td>0.001</td>
</tr>
<tr>
<td>≥ 1 Assistive device</td>
<td>7.3</td>
<td>0.003</td>
</tr>
<tr>
<td>≥ 2 Oral medications</td>
<td>6.6</td>
<td>0.03</td>
</tr>
<tr>
<td>Intubated</td>
<td>4.8</td>
<td>0.01</td>
</tr>
<tr>
<td>≥ 2 IV medications</td>
<td>4.3</td>
<td>0.03</td>
</tr>
<tr>
<td>Elective admission</td>
<td>3.6</td>
<td>0.04</td>
</tr>
</tbody>
</table>

* Cases and controls included only those for whom 24-hour CCMR data were available. The following 28 risk factors did not have statistically significant results: median age at admission, patient type, day of the week, month of the year, primary service, location of service, median length of admission, any ongoing medical conditions, median number of ongoing conditions, any serious ongoing medical condition, purpose of admission, condition at admission, any diagnostic evaluation, any medical treatment, any nonsurgical invasive treatment, any surgical treatment, English not first language of patient, time of event, condition before event, level of assistance with daily living, level of nursing monitoring required, communicative ability of patient, treated with continuous oxygen, treated with noninvasive continuous positive airway pressure/biphasic positive airway pressure, treated with nebulizers, frequent respiratory treatment, median number of dressing changes, and parent/caregiver present.

### Table 4. Models of multivariable analysis: Limited dataset*\(^a\)

<table>
<thead>
<tr>
<th>Models</th>
<th>Odds ratio</th>
<th>95% Confidence interval</th>
<th>p value</th>
<th>c statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Most predictive model</td>
<td></td>
<td></td>
<td></td>
<td>0.87</td>
</tr>
<tr>
<td>&gt; 4 Medications</td>
<td>21.6</td>
<td>(18.25-65.6)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>&gt; 1 Nebulized medication</td>
<td>15.2</td>
<td>(1.2-200.3)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Model 2: Most predictive model that forces in 24-hour CCMR</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 15,000 Bits of 24-hour CCMR</td>
<td>9.7</td>
<td>(1.4-69.9)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>&gt; 1 Nebulized medication</td>
<td>9.8</td>
<td>(1.0-93.7)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Model 3: Most predictive model that forces in lifetime CCMR</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 70,000 Bits of lifetime CCMR</td>
<td>19.3</td>
<td>(1.5-240.5)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>&gt; 1 Nebulized medication</td>
<td>16.3</td>
<td>(1.3-207.3)</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

* The limited dataset includes only the 17 cases and 42 controls for which we have 24-hour CCMR. CCMR = cognitive complexity of the medical record.
Finally, our datasets were relatively small, and thus it was not possible to determine whether 24-hour CCMR and lifetime CCMR are independent risk factors for major adverse events.

Conclusion

A new instrument, the Complexity Ruler, was designed to measure the cognitive complexity of the medical record. In this study, the measured CCMR correlated with complexity as assessed by experienced clinicians, and it was a risk factor for major adverse events. Although unproven, it may be that the mechanism linking CCMR to major adverse events is cognitive overload.

An automated version of the Complexity Ruler is being developed. Automated, real-time information about complexity may have value in identifying patients at high risk of adverse events.

Disclosure Statement

Dr Connell is Chief Executive Officer of Knowledge Design Inc; Boston, MA. Boston Children’s Hospital, (Boston, MA) has applied for a patent for the method of determining patient complexity reported in this article. Draper Laboratory (Cambridge, MA) has sponsored this study with an unrestricted research grant. The authors have no other competing interests to disclose.

Both studies were approved by the Boston Children’s Hospital institutional review board.

Acknowledgments

The authors gratefully acknowledge the support of numerous clinicians at Boston Children’s Hospital, who helped us understand the many sources of information accumulating in the medical record. Kathleen Louden, ELS, of Louden Health Communications provided editorial assistance.

References


Besieged City

The physician should look upon the patient as a besieged city and try to rescue him with every means that art and science place at his command.

—Alexander of Tralles, c.525-c.605, one of the most eminent of ancient physicians