Analyzing Primary Care Data to Characterize Inappropriate Emergency Room Use

Justin St-MAURICE [a,b,1], M. H. KUO [a]

[a] *School of Health Information Science, University of Victoria, BC, Canada*
[b] *School of Health and Life Sciences, Conestoga College, Ontario, Canada*

Abstract. Primary care data represents the complete biopsychosocial profile of the patient and has recently become available for secondary analysis in Ontario. This study involved extracting de-identified primary care data and analyzing it with natural language processing to extract UMLS codes. These codes were used to statistically relate inappropriate emergency room use to biopsychosocial concepts with logistic regression. The concepts of pain and mental health were statistically significant. This technique demonstrates a creative use of primary care data. It could be used to analyze system use in other health care settings.

Keywords. Primary care, free text data, natural language processing, emergency room use, pain, mental health, biopsychosocial

Introduction

Significant investments have been made to computerize the primary care system in Ontario. Whereas other fields of medicine are focused on a primarily biomedical framework, primary care assesses the complete biopsychosocial profile and provides regular patient-centered assessments. In Ontario, the digitization of primary care has reached over 40 percent and it has become available for secondary analysis.

The primary care data set potentially contains all health events in a person's life. The data should include information about morbidity, treatment, outcomes and health care utilization [1]. Previous studies have examined the question of inappropriate emergency room use through the use of inpatient surveys and questionnaires. With access to primary care data, there are new opportunities to answer the question of inappropriate emergency room use. This study aimed to characterize emergency room use with primary care data.

---
[1] *Corresponding Author. Justin St-Maurice, School of Health Information Science, University of Victoria, BC, Canada/ School of Health and Life Sciences, Conestoga College, Ontario, Canada*

1. Literature Review

Many studies noted that inappropriate use of the emergency room or department was a costly phenomenon that has been increasing [2-11]. The issue of inappropriate emergency room use as a critical question for the long term sustainability of their healthcare system was also highlighted.

Previous studies of inappropriate emergency room relied on patient surveys, interviews or questionnaires [2-8], emergency room registries [9, 10] or both [11]. During a literature review, no studies were found that related data from primary care records to inappropriate usage. Each study [2-11] concluded there were reportable differences in inappropriate emergency room use between the groups they studied.

2. Methods

2.1 Sample Data

Wilcox et al. observed that of 3,348 physicians, only 8% were storing data in structured forms [12]. In contrast, the same study found that 75% of those physicians were consistently entering or dictating encounter notes and medication information. Given its broader use, free text in primary care medical record systems was used for this study.

A clinic in Guelph, Ontario, Canada was selected for the study. 8 physicians provided de-identified records. The physicians had different types of educational backgrounds (5 different universities), age (30 year range), and years in practice (25 year range). The patient population from the physician cohort included 13,878 records. The records were exported in OntarioMD's CDS 3.0 format. Free text content was exported from the sections "MyClinicalNotes", "PastHealth", "PersonalHistory", "ProblemList" and "RiskFactors". The de-identified records were linked to de-identified data from the Guelph General Hospital by using a common health card hash.

2.2 Design and Procedures

In order to analyze the free text data, it needs to be converted into codes. This can be achieved through Natural Language Processing (NLP). MetaMap 2011 was used as the core NLP tool in this study [13]. MetaMap takes sentences and returns annotations from the Unified Medical Language Syntax (UMLS) database. Based on a minimal commitment parser, records are split into chunks and annotated with codes. The parser is based on the notion of a special set of barrier words that indicate boundaries between phrases. These barrier words make it possible to run MetaMap without a training model [14]. The General Architecture for Text Engineering (GATE) tool was used in conjunction to pre-process the free text content [15,16]. The GATE tool then forwarded sentences to MetaMap 2011 through a plugin [17]. Both tools were embedded into a customized software application created for this study.

2.3 Data Analysis

The classification of inappropriate or appropriate ER use was computed with the same method as Field et al. [3] by using the Canadian Triage and Acuity Scale (CTAS). A score of 4 or 5 was considered inappropriate.

The annotation results generated by the custom software were stored in a MySQL database. Binary logistic regression was performed with R [18]. Results were exported from MySQL. Each record formed a row, and concepts formed columns. In R, a for-loop was used to regress each UMLS concept in the model *InappropriateUseYN ~ Age + Gender + Code$_n$*. For each column representing an UMLS code, a zero represented the absence of the code in the patient record and 1 presented the presence of a code in the patient record. To minimize the effects of random error, only UMLS concepts with p-value < 0.001 were used. In addition, to maximize clinical relevance only UMLS concepts with more than a doubling effect on the odds of using the emergency room inappropriately (OR < 0.5 or OR > 2.0) were selected. 417 concepts met these criteria and were exported to a list. The list was presented to the Chief of Staff of the Guelph General Hospital and sorted as either "Biological Symptoms", "Diagnosis", "Psychological", "Social", "Drugs", "Regional Oddities", "EMR Oddities" or "Other".

3. Results

Over a three year period (Nov. 1, 2008 to Oct. 31, 2011) the patient cohort had a 26% incidence of emergency room usage. During this period, 1931 patients had used the emergency room inappropriately (13.9%). In comparison, in the previous year 798 patients had used the hospital inappropriately (5.7%). Overall, 39.13% of emergency room visits were considered inappropriate.

There were 10,823,636 annotations extracted from the 13,836 records in the cohort. The number of codes per patient varied from 1 to 9031. There were 38,263 distinct UMLS codes. The number of distinct codes per record varied from 1 to 1896. 33,446 distinct codes (87.14%) were only present in 100 different records. Only 4028 (10%) of the total distinct codes were found in at least 1% of the records. Most annotations occurred more than once in a patient record.

14 of the UMLS concepts (41%) that were categorized in the physical symptoms category involved pain related concepts. These results are shown in Table 1. Another common theme in the retrieved concepts were mental health related concepts, such as bipolar (OR = 2.43), anorexia (OR = 2.35), psychiatric problem (OR = 2.42) and Attention Deficit Hyperactivity Disorder (OR = 2.31).

Table 1. UMLS Concepts representing pain.

UMLS Concept ID	UMLS Name	Beta	Probability	Inappropriate Use (with code)	Inappropriate Use (without code)	Odds Ratio
C0000737	Abdominal Pain	0.881	< 0.001	n=188 (36.9%)	n=1743(15.3%)	2.41
C0278140	Severe pain	0.858	< 0.001	n=131 (36.3%)	n=1800(15.6%)	2.36
C1963236	Sinus Pain Adverse Event	0.856	< 0.001	n=50 (37.6%)	n=1881(16.0%)	2.35
C0030193	Pain	0.84	< 0.001	n=1031(25.4%)	n=900 (11.5%)	2.32
C1963242	Stomach Pain Adverse Event	0.811	< 0.001	n=56 (36.1%)	n=1875(16.0%)	2.25

C1962977	Pain NOS Adverse Event	0.772	< 0.001	n=1300(21.3%)	n=631 (10.9%)	2.17
C0150055	Chronic pain	0.746	< 0.001	n=119 (32.5%)	n=1812(15.7%)	2.11
C0016199	Flank Pain	0.74	< 0.001	n=97(33.0%)	n=1834(15.8%)	2.1
C0740418	Chronic back pain	0.714	< 0.001	n=60(32.3%)	n=1871(16.0%)	2.04
C0234254	Radiating pain	0.713	< 0.001	n=48(32.4%)	n=1883(16.0%)	2.04
C0030794	Pelvic Pain	0.707	< 0.001	n=104(32.0%)	n=1827(15.8%)	2.03
C0748706	side pain	0.704	< 0.001	n=41(32.3%)	n=1890(16.1%)	2.02
C0741585	BODY PAIN	0.703	< 0.001	n=50(32.5%)	n=1881(16.0%)	2.02

4. Discussion

As shown in Table 1, the theme of pain as a reason for inappropriate emergency room surfaced through the analysis. This concept is well supported by a variety of literature, as discussed by Pereira et al. [19]. Mental health themes also surfaced as an important dimension and is consistent with literature discussed by Keene and Rodriguez [20]. The themes of pain and mental health are also consistent with anecdotal information from the Guelph community. Though very promising, the methodology and results require further refinement before the results can be interpreted and used for decision making.

Binary logistic regression was used in the analysis to demonstrate the simple use of primary care data in the analysis. However, the model only considered a single component of a broad biopsychosocial profile in each computation. Though significantly more complicated, concepts could be grouped and analyzed together to form "constellations" of terms that would better depict a complete patient profile.

The use of CTAS for the categorization of inappropriate use of the emergency room has been debated. Whereas it is supported by existing literature as a method, it is easily characterized as an oversimplification of a complex issue. There are a variety of legitimate reasons for using the emergency room that would be categorized as inappropriate incorrectly through this study, including suture removals, early morning care or basic, non-critical first aid. The selection of CTAS was based on the availability of data and served as a reasonable proxy measure.

Another limitation is technical; the MetaMap analysis was unable to properly disambiguate some terms, nor was it designed for this specific application. It was not able to properly interpret the difference between the narrative that the 'patient was kicked in the face at the bar' versus that the 'patient was kicked out of the mental health support group'. There were also instances of 'regionalism' in the narrative that were not properly interpreted. For example, the abbreviation of 'GGH', intended to mean 'Guelph General Hospital', was interpreted as a reference to the 'GGH Gene'. This failure to disambiguate abbreviations highlights the geographic nature of natural language technologies and the complexity of their applications in each health care community. These inaccuracies in the natural language processing component of the analysis introduced significant noise which was reduced by limiting the statistical significance of results. Different configuration settings or tools should be used to explore reducing this error in subsequent research aimed at further refining this methodology.

The potential uses of this methodology are numerous. Further refined, the methodology could be used to characterize the use of community health services, mental health clinics, social services or specialists. The technique would also prove

very interesting to public health agencies who are seeking to use primary care data for secondary purposes.

5. Conclusions

Despite some limitations, this study has demonstrated a novel use of primary care data to characterize the inappropriate use of emergency rooms in Canada. Free text was converted into UMLS concepts through a custom software application. Through the analysis, the themes of mental health and pain were associated with inappropriate emergency room use. These results are consistent with existing literature.

Future work surrounds the use of the current data which can be used to provide additional insight regarding emergency room use by investigating other relationships between CTAS levels, physicians and biopsychosocial concepts.

References

[1]   St-Maurice J. Primary Care Data?: Gold or Pyrite? A Literature Review. Journal of Health Information Management 2011; 25 (2): 27-31
[2]   Carret M, Fassa A, Kawachi I. Demand for emergency health service: factors associated with inappropriate use. BMC health services research 2007; 7:131
[3]   Field S, Lantz A. Emergency department use by CTAS Levels IV and V patients. Canadian journal of emergency medical care 2006; 8(5): 317-22
[4]   David M. Emergency outpatient services in the city of Berlin: Factors for appropriate use and predictors for hospital admission. European journal of emergency medicine: official journal of the European Society for Emergency Medicine 2006; 13(6): 352-7
[5]   Selasawati H, Naing L, Wan Aasim W,Winn T, Rusli B. Factors Associated with Inappropriate Utilisation of Emergency Department Services. Asia-Pacific Journal of Public Health 2007;19(2):29-36
[6]   Loría-Castellanos J, Flores-Maciel L, Márquez-Ávila G, Valladares-Aranda MA. Frequency and factors associated with misuse of hospital emergency services. Cir Cir 2010; 78(6): 505-510
[7]   J Tsai JC, Liang YW, Pearson WS. Utilization of emergency department in patients with non-urgent medical problems: patient preference and emergency department convenience. Journal of the Formosan Medical Association 2010; 109(7): 533-42
[8]   Northington WE, Brice JH, Zou B. Use of an emergency department by nonurgent patients? The American Journal of Emergency Medicine 2005; 23(2): 131-137
[9]   De Vos P, Vanlerberghe V, Rodríguez A, García R, Bonet M, Van der Stuyft P. Uses of first line emergency services in Cuba. Health policy (Amsterdam, Netherlands). 2008; 85(1): 94-104
[10]   Selasawati HG, Naing L, Wan Aasim WA, Winn T, Rusli BN. Inappropriate utilization of emergency department services in Universiti Sains Malaysia hospital. The Medical journal of Malaysia 2004; 59(1): 26-33
[11]   Siminski P, Bezzina AJ, Lago LP, Eagar K.. Primary care presentations at emergency departments: rates and reasons by age and sex. Aust Health Rev. 2008; 32(4):700-9.
[12]   Wilcox A, Bowes W, Thornton SN, Narus SP, Narus S. Physician use of outpatient electronic health records to improve care. AMIA Symposium 2008; pp. 809-13,
[13]   Aronson AR, Lang FM. An overview of MetaMap: historical perspective and recent advances. Journal of the American Medical Informatics Association 2010; 17(3): 229-36
[14]   N. Kang, van Mulligen E, Kors J. Comparing and combining chunkers of biomedical text. Journal of biomedical informatics 2011; 44(2):354-60
[15]   Cunningham H, Maynard D, Bontcheva K, Tablan V. GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. in Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02); 2002.
[16]   http://gate.ac.uk/releases/gate-6.1-build3913-ALL/tao.pdf (last accessed May 2012)
[17]   Gooch P, Roudsari A. A tool for enhancing MetaMap performance when annotating clinical guideline documents with UMLS concepts. IDAMAP Workshop at 13th Conference on Artificial Intelligence in Medicine (AIME'11); 2011.
[18]   Kabacoff RI. R in Action. Shelter Island, NY: Manning Publications Inc; 2011.
[19]   Pereira S, Oliveira e Silva A, Quintas M, Almeida J, Marujo C, Pizarro M, Angélico V, Fonseca L, Loureiro E, Barroso S, Machado A, Soares M, da Costa AB, de Freitas AF. Appropriateness of emergency department visits in a Portuguese University hospital? Annals of Emergency Medicine 2001; 37(6):580-586
[20]   Keene J, Rodriguez J. Are mental health problems associated with use of Accident and Emergency and health-related harm? European journal of public health 2007; 17(4):387-93