ANALYSIS OF HIGH-LEVEL ELECTRONIC OPERATIVE REPORT DOCUMENTATION STRUCTURE

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Genevieve B. Melton-Meaux

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

David Pieczkiewicz, PhD

February 25, 2015

Acknowledgements

It has been challenging to balance family, work, other personal life, and a graduate education. Many people have played important roles in making this dissertation possible, and my acknowledgements are but a small reflection of my appreciation.

First, I would like to recognize the helpful contributions and assistance of my committee members: David Pieczkiewicz, Karen Monsen, Gyorgy Simon, and David Rothenberger, as well as my PhD mentor, Terrence Adam. My long-time collaborators, Serguei Pakhomov and Elizabeth Chen, have also provided great support and friendship.

There are many important informatics colleagues, friends, and mentors that I would like to thank, including George Hripcsak, Adam Wilcox, Frances Morrison, Philip Payne, Peter Embi, James Cimino, Peter Stetson, Michael Chiang, Neil Sarkar, Li Zhou, Hua Xu, Carol Friedman, Suzanne Bakken, Joshua Denny, Chris Lehmann, Michael Kim, Saif Khairat, Edward Shortliffe, Shamim Mollah, Kevin Johnson, Charlie Safran, Don Detmer, Hongfang Liu, Nate Schmiechen, Laël Gatewood, Bonnie Westra, Stuart Speedie, and Connie Delaney.

As a practicing surgeon, I am very fortunate to have outstanding and supportive surgical partners - Robert Madoff, David Rothenberger, Mary Kwaan, Wolfgang Gaertner, and Colon and Rectal Surgeon Associates colleagues. I also have had amazing surgical mentors: Keith Lillemoe, Charlie Yeo, John Cameron, Pamela Lipsett, Julie Frieschlag, David Rothenberger, Robert Madoff, Jon Vogel, Tracy Hull, Feza Remzi, James Church, Victor Fazio, Thomas Magnuson, Stanley Goldberg, Anders Mellgren, Ann Lowry, David Dietz, Scott Steele, Alan Herline, Selwyn Vickers, and Clifford Ko.

My current and previous Health Informatics graduate students have been a great source of motivation, including Yan Wang, Oladimeji Farri, Rui Zhang, Sungrim ("Riea") Moon, Tamara Winden, Nora Burkart, Elliot Arsoniadis, and Zhen Hu. They are complemented by colorectal surgery fellows who challenge and stretch me everyday as an academic surgeon and clinician. I am also grateful for the help and support of staff and colleagues, including Lindsay Bork, Kimberly Marshall, Reed McEwan, Jessica Whitcomb-Trance, Pat Mann, Alexandra Broek, Carol Bigalke, and Faith Goenner.

This work was also made possible by funding support from the American Surgical Association Foundation, Agency for Healthcare Research & Quality, National Institutes of Health, University of Minnesota Regents Scholarship, and other internal support from my departments, Fairview Health Services and University of Minnesota Physicians.

On a personal note, my friends and family have provided support, encouragement, and advice. I would like to acknowledge Justin Dale, Christine Martin, Liza Wick, Terry Harris, Vanderlein Meaux, Di Tran, Guenada Meaux-Lambert, Geraldine Meaux, and Clement Meaux. My parents, Bob and Evelyn Melton, remain a great source of motivation – including their strong values towards education, intellectual curiosity, and livelong personal growth. My "brother of honor", Dennis Melton, has strongly supported and helped me when I have most needed it. Most profoundly, I would like to thank my husband, Antone Melton-Meaux, for his amazing love, faith, patience, support of our family and marriage, help in improving me as an individual, and for leading by example in how to be a better member of the greater communities in which we live.

Dedication

This dissertation is dedicated to Antone, Ava, and Xavier - the "Melton-Meaux" clan.

Their love, support, and encouragement are my greatest source of joy and fulfillment.

Abstract

ANALYSIS OF HIGH-LEVEL ELECTRONIC OPERATIVE REPORT DOCUMENTATION STRUCTURE

Genevieve B. Melton-Meaux

Operative report documentation (ORD) is a fundamental part of surgical practice that has a direct and significant role in quality, medico-legal and billing, and other secondary uses of operative reports. Traditionally, operative reports are created by the primary surgeon through dictation and subsequent transcription of details of the operative procedure after its completion. With the advent of widespread electronic health record (EHR) adoption, there is a potential opportunity to improve aspects of ORD through synoptic reporting and templating of operative reports, as well as to leverage operative report sections and high-level phases described with surgical procedures in operative report. This research seeks to illuminate knowledge about high-level ORD structure including an assessment of attitudes and opinions towards ORD in surgical education, evaluation of structured operative report section names, and appraisal of high-level phases of surgical descriptions in operative reports with the application of automated methods to classify these phases. An electronic survey about ORD teaching and practices was sent to surgical program directors. While most program directors responding to the survey consider ORD teaching a priority, few provide ORD instruction, and significant barriers were perceived by program directors for ORD instruction and ORD using synoptic reporting. To evaluate operative report sections, the HL7 Implementation Guide for Clinical Document

Architecture Release 2.0 Operative Note Draft Standard for Trial Use (HL7-ON DSTU) Release 1 and Logical Observation Identifiers Names and Codes (LOINC®) structured sections were evaluated on 384 unique section headers from 362,311 operative reports. HL7-ON DSTU alone and HL7-ON DSTU with LOINC® section headers covered 66% and 79% of sections headers (93% and 98% of header instances), respectively. Section headers contained large numbers of synonyms, formatting and word form variation, as well as coverage gaps in the current terminology sources. In a third study, high-level phases of the "Surgery Description" from operative reports were identified by surgeons. Automatic classification with support vector machines using topic analysis and information gain for feature selection was then used to identify these high-level phases, followed by assessment of subtopics for phases. Five high-level "Surgery Description" section phases were identified: *Preliminaries, Getting Started, Main Part, Closure,* and *Epilogue*, as well as need for *Observations* to describe findings, events, and other important clinical information. Automated classification of high-level sections performed well on a set of laparoscopic cholecystectomy and random surgery operative reports and was associated with the identification of meaningful subtopics. Overall, this research demonstrates the variability in ORD practices for surgeons nationally and individually with respect to section structure, as well as the value of high-level phases to group content in these important clinical documents. Future work will seek to leverage our understanding of ORD structure to improve information extraction and natural language processing (NLP) techniques for secondary use of operative reports.

Table of Contents

Acknowledgements iii

Dedication v

Abstract vi

List of Tables x

List of Figures xiii

Abbreviations xiiii

Chapter 1: Introduction 1

1.1 Significance 1
1.2 Synoptic Reporting and Templates 3
1.3 Clinical Document Sections 3
1.4 Perioperative knowledge frameworks 4
1.5 Specific Aims 6

Chapter 2: Operative Report Teaching and Synoptic Operative Reports: A National Survey Of Surgical Program Directors 8

2.1 Summary 9
2.2 Introduction 11
2.3 Methods 13
2.4 Results 16
2.5 Discussion 21
2.6 Conclusion 27
2.7 Acknowledgments 28

Chapter 3: Analyzing Operative Report Structure in Development of a Section Header Resource 29

3.1 Summary 30
3.2 Introduction 31
3.3 Materials and Methods 34
3.3.1 Study Overview 34
3.3.2 HL7-ON DSTU Section Header Extraction 35
3.3.3 Operative Report Section Header Evaluation 35
3.4 Results 38
3.4.1 HL7-ON DSTU Section Header Extraction 38
3.4.2 Operative Report Section Header Evaluation 39

3.5 Discussion **42**
3.6 Conclusion **46**
3.7 Acknowledgments **46**

Chapter 4: Modeling the 'Surgery Description' in Operative reports 47

4.1 Summary **48**
4.2 Introduction **50**
4.3 Background **51**
4.3.1 Operative report "Surgery Description" Section **51**
4.3.2 Topic Analysis **53**
4.4 Methods **54**
4.5 Results **57**
4.6. Discussion **66**
4.7. Conclusions **69**
4.8. Acknowledgements **69**

Chapter 5: Conclusions 70

Bibliography 74

List of Tables

Table 2.1. Open-ended questions about operative report documentation teaching 15

Table 2.2. Characteristics of Respondents to Operative Report Documentation Survey 17

Table 2.3. ORD Teaching Modalities 18

Table 2.4. Perceived barriers to ORD Teaching 19

Table 2.5. Perceived weaknesses of synoptic operative reports/templates and their use in resident education 20

Table 2.6. Perceived strengths of synoptic operative reports and their use in resident education 21

Table 2.7. Example responses about synoptic reports for operative report documentation and their role in surgical resident education 25

Table 3.1. Coding for Operative Report Section Headers 36

Table 3.2. Example Operative Report Sections from HL7-ON DSTU 39

Table 3.3. Operative Report Section Header Findings 40

Table 3.4. HL7-ON DSTU Section Header Mappings 41

Table 4.1. Questions about operative report documentation and the "Surgery Description" section 55

Table 4.2. High-level Phases of the "Surgery Description" Section of Operative reports and Descriptions Used for Annotations of the "Surgery Description" section 58

Table 4.3. Inter-rater reliability for annotations of high-level phases 61

Table 4.4. Classification Performance at a Sentence Level for High-Level Phases of the 'Surgery Description' Section using LDA and IG with SVM 63

Table 4.5.a. Subtopics for the phase *Preliminaries* in the “Surgery Description” Section..64

Table 4.5.b. Subtopics for the phase *Getting Started* in the “Surgery Description” Section..64

Table 4.5.c. Subtopics for the phase *Closure* in the “Surgery Description” Section.........65

Table 4.5.d. Subtopics for the phase *Observation* in the “Surgery Description” Section..65

List of Figures

Figure 3.1. Overview of Study..34

Figure 4.1.a. Overall F-measures varied with the number of keywords and topics for laparoscopic cholecystectomy operative reports...62

Figure 4.1.b. Overall F-measures varied with the number of keywords and topics for the random set of operative reports...62

Abbreviations

ACGME - Accreditation Council for Graduate Medical Education

ACS-NSQIP - American College of Surgeons National Surgical Quality Improvement Program

CCD - Continuity of Care Document

CMS - Centers for Medicare & Medicaid Service

CPT - Current Procedure Terminology

E/M - Evaluation and Management documentation

EHR – electronic health record

GALEN - Generalized Architecture for Languages, Encyclopedias and Nomenclatures

H&P - "history and physical" note

HIE - health information exchange

HL7 - Health Level 7

HL7-ON DSTU - HL7 Implementation Guide for Clinical Document Architecture Release 2.0 Operative Note Draft Standard for Trial Use Release 1

IG – information gain

LDA – Latent Dirichlet Allocation

LOINC® - Logical Observation Identifiers Names and Codes

LVG - lexical variation generation

NIC - Nursing Interventions Classification

NLP – natural language processing

ORD – operative report documentation

PNDS - Perioperative Nursing Data Set

SEER - Surveillance Epidemiology and End Results

SVM – support vector machines

UMLS - Unified Medical Language System

Chapter 1: Introduction

1.1 Significance

Surgical procedures represent a highly stressful, costly, and impactful healthcare event. With 45 million surgeries occurring each year in an increasingly elderly population, there is a large and growing need for pre-procedural evaluations[1]; quality improvement in surgical care to increase efficiencies and improve outcomes[2, 3]; and framing patient expectations. Despite the high acuity, cost, and potentially serious consequences associated with surgical care, surgical informatics and particularly clinical research informatics in the surgical domain remain nascent disciplines[4]. While the best means to achieving better surgical quality is not known and likely multi-factorial, it is clear that improving surgical care requires *appropriate and reliable data based upon clinical evidence.*

Operative reports are created with the primary purpose of documentation and billing for a surgical procedure. As such, they have a direct and significant role in surgical quality[5], medico-legal issues[6, 7], and secondary uses such as clinical research[4]. While some surgeons create operative reports using electronic templates or synoptic tools[8-10], operative reports are predominantly and traditionally created by the surgeon with dictation, followed by transcription to create the electronic note that is then stored in an electronic health record (EHR) system.

Operative reports record the surgeon's recollection of the details of the procedure[11] and thus contain a wealth of information on the associated techniques, materials (e.g., supplies or equipment), instruments, and intra-operative events (e.g., complications or unexpected findings) of surgery. Like other clinical texts, operative reports are split into sections where the primary section (*'Surgical Description'* section) describes details of what was observed and performed during the operation.

Effective use of operative reports therefore requires an understanding of surgical domain knowledge including knowledge of the perioperative setting. Multiple studies have demonstrated specific sublanguage features with medical texts compared to general English, along with characteristics of different specific medical sublanguages[12-15]. In contrast to most other medical documents that are collections of facts about the patient and the encounter, operative reports contain detailed descriptions about the actions performed, as well as intra-operative observations of the surgeon[16, 17].

This dissertation brings together survey methodologies and qualitative analysis, knowledge representation and vocabularies/ontologies, machine learning and topic analysis, and surgical domain knowledge. It is important to emphasize that even though operative reports are recognized by the surgical community to be one of the key inputs to support surgical research and quality improvement in surgical care, operative report structure and operative report documentation practices remain a largely unexplored area.

1.2 Synoptic Reporting and Templates

A top-down alternative approach for obtaining information about operations has been advocated by some surgeons, with the use of synoptic operative reports[18, 19]. Synoptic reports, by definition, create a "note" that is structured for documentation. The format of synoptic reports is composed of discrete data fields where each type of desired information has a specific format in the report allowing for standardized collection, storage, and retrieval of this data after creation of the report. Templated notes, in contrast to pure synoptic reports, have a range of formats from more prescriptive, structured, and discrete content that is very similar to synoptic reports to being mostly free-text narrative with only basic structure for document sections.

While synoptic reports are used more widely in pathology and are designed to facilitate documentation in a standard format and the extraction of structured information, dictation with transcription largely predominates for operative report creation in the United States[20]. There are also potentially subjective barriers to the use of synoptic operative reports, which have not been well characterized to date.

1.3 Clinical Document Sections

Most clinical documents including operative reports are characterized by section headers which provide high level structure and serve as "containers" which provide context to the text within the given section [21]. While some previous research has explored the task of automated classification of sections in documents, including classification with the OpenGALEN project into "tags" of 5 clusters (*Nature, Safety Context, Interpretation, Intention, and Organization*), this was formative work not focusing on operative reports or a specific type of clinical text[21].

In a study analogous to that proposed here for operative reports, Denny et al. developed a terminology of document sections for "history and physical" (H&P) notes and developed an associated section terminology with Logical Observation Identifiers Names and Codes (LOINC®) mappings[22]. In follow-up work, these authors developed an algorithm, "SecTag", to identify and label section headers and section boundaries in H&P notes[23]. Some natural language processing (NLP) tasks have also used predefined sections to aid in important tasks like problem list extraction, named entity recognition, and other tasks[24-26].

1.4 Perioperative knowledge frameworks

There has been limited work on perioperative knowledge representation. Nursing has developed the Perioperative Nursing Data Set (PNDS) by the Association of Perioperative Registered Nurses[27] and Nursing Interventions Classification (NIC)[28, 29], but PNDS and NIC are interface terminologies to document care (e.g., "*Verifies consent for planned procedure"* - PNDS) and not reference terminologies, which represent semantics/meaning for downstream uses. One notable initiative is the Generalized Architecture for Languages, Encyclopedias and Nomenclatures (GALEN) from the European Committee for Standardization[30], which includes a schema for surgical procedures[31]. GALEN includes descriptive logic in GRAIL and OWL[32] of surgical procedures and associated steps (i.e., "MAIN deeds"). While reported to a limited extent previously as a potential resource for operative report generation[33] and procedure coding[34], GALEN is no longer supported and has not been widely disseminated (also not part of Unified Medical Language System - UMLS). GALEN also lacks content from some surgical specialties (e.g., vascular surgery, otolaryngology).

Some formative work has also been done with surgical process modeling[35], but this emphasizes third party observations of surgeons and other individuals in the operating room from an industrial engineering perspective and not the perspective of surgeon-described activities for documentation of cases. Meng et al.[36] created a representation of surgical steps ("deeds") with anatomic concepts as well as procedure concepts to represent radical retropubic prostatectomy steps from operative reports. The investigators then used multiple sequence-alignment to generate a directed graph representing the steps of the case. A large number of errors (62%) were from verbs not mapping (e.g., "*encircled*", "*prepped*"), and insufficient specification of the steps of the procedure. As such, there is an opportunity to further define knowledge frameworks associated with operative reports and the conduct of surgical procedures for better use of operative reports.

1.5 Specific Aims

The goal of this research is broadly to enrich our understanding of the current and ideal state of ORD, including examining associated structure associated with operative report sections and phases used by surgeons in their description of a surgical procedure within operative reports.

The main hypothesis of this research is that fully structuring ORD with synoptic reporting has real and perceived barriers for surgeons and that explicit and implicit high-level structure with operative report sections and phases of surgical procedures can be ascertained and leveraged for better secondary use of operative reports. A corollary to this is that a balance between structured clinical documentation and narrative texts remains. To this end, this research addresses the following objectives:

- Investigate current ORD practices including an assessment of attitudes and opinions towards ORD in surgical education.
- Evaluate formal section header standards and current operative report sections to build an operative report section resource; and
- Identify high-level phases of the "Surgery Description" section of operative reports and evaluate supervised machine-learning techniques to classify high-level phases.

Together, these three studies serve to fill fundamental gaps in our understanding of how surgeons document intraoperative care in the form of operative reports. Findings from this research can inform future work with EHR system development for surgical documentation and help to guide future development of operative report templates and synoptic reports. In addition to its value for NLP applications, the operative report section resource will be helpful for health information exchange (HIE) efforts, specifically efforts of organizations like Health Level 7 (HL7) and its Structured Documents Workgroup, which seeks to develop structured healthcare document standards to promote document and data interoperability. Finally, identification of and the application of high-level phases for the description of procedures will help to further compartmentalize portions of procedures documented in operative reports and topics (or procedure "steps") involved with documenting operations.

Chapter 2: Operative Report Teaching and Synoptic Operative Reports: A National Survey Of Surgical Program Directors

Genevieve B. Melton, MA, MD[a,b], Nora E. Burkart, MD[a,b], Nathan G. Frey[b],

Jeffrey G. Chipman, MD[a], David A. Rothenberger, MD[a], Selwyn M. Vickers, MD[a]

[a]*Department of Surgery, University of Minnesota, Minneapolis, MN, USA*

[b]*Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA*

Published in:

Journal of the American College of Surgeons. 2014 Jan;218(1):113-8.

doi: 10.1016/j.jamcollsurg.2013.09.004. Epub 2013 Nov 6.

*This work was presented in part as an oral presentation at the 8th annual American Surgical Congress, New Orleans, February 2013 (Abstract Number: ASC20130449).

2.1 Summary

***Background*:** Although operative report documentation (ORD) is an essential skill for surgeons and is evolving with electronic health records (EHRs), little is known about current ORD teaching in surgical training.

***Study Design*:** An electronic survey was sent January 2012 to all 1,096 ACGME surgical program directors assessing characteristics of training programs, EHR adoption, ORD education, synoptic or templated report usage for ORD, and attitudes and opinions about ORD education and electronic tools for ORD. Content thematic analysis of qualitative responses was performed iteratively until reaching saturation.

***Results*:** Overall, 441 program directors (40%, 17.9 ± 8.8 years in practice) responded from university-affiliated (383, 87%), community/private (44, 10%), and military (14, 3%) programs. Although most (295, 67%) consider ORD teaching a priority, only 76 (17%) programs provide ORD instruction. Program directors formally trained in ORD were more likely to offer ORD instruction (61% vs. 11%, $p<0.0001$), as were Obstetrics-Gynecology programs (Obstetrics-Gynecology-35% vs. Surgery-18%, Neurosurgery-16%, Ophthalmology-14%, Orthopedics-14%, $p<0.05$ each). While EHR adoption and electronically available operative reports were common (91%), besides Ophthalmology (31%) and Obstetrics-Gynecology (30%) programs, ORD with synoptic reporting was used overall in only 18% of programs. Program directors perceived significant barriers to ORD instruction and synoptic reporting for ORD.

Conclusions**:** While most program directors consider ORD teaching an educational priority, incongruence exists between its perceived value and adoption into surgical training. ORD with synoptic reporting is currently not common in most surgical subspecialties.

2.2 Introduction

During residency, surgical trainees must acquire an extensive knowledge base and a large number of skills to become competent and safe in their practice. As part of this, the information and techniques relevant to surgical procedures and the operating room environment are fundamental and essential. While not traditionally a topic of significant focus in surgical training, effective operative report documentation (ORD) is a critical skill that has a direct role in quality[5], medico-legal and billing[6, 7], and other secondary uses of operative reports[4]. Traditionally, ORD is most commonly performed by the primary surgeon through dictation with the primary surgeon recollecting and providing the details of the operation after completion of the procedure.

With respect to resident training and ORD, the role of residents in ORD remains controversial and not well characterized amongst the many surgical subspecialties. While one study demonstrated similar rates of operative report note completion by both urology residents and experienced surgeons,[37] another study which compared operative reports authored by the attending surgeon to those by the participating resident for the same operation demonstrated significant deficiencies in the content of resident dictations, particularly for complex and multi-faceted procedures.[38] Although objective rates of ORD teaching in most surgical specialties have not been well characterized, Eichholz et. al (2004)[39] reported the results of a survey of Obstetrics and Gynecology program directors and demonstrated for this specialty only 23% of programs at that time provided formal training to residents in how to construct a well dictated operative report.

In addition, several studies examining ORD with dictation as a general practice, including a recent assessment of dictated breast cancer surgery notes[40], have demonstrated high rates of omission of pertinent information related to the conduct of the procedure and with respect to the quality and extent of the procedure. Dictated operative reports may also fail to meet requirements of regulatory agencies such as the inclusion of key operative report section headings specified by the Joint Commission[20].

Electronic health records (EHRs) represent a potential opportunity to change and improve certain aspects of ORD. For instance, EHRs enable synoptic reporting and templating of operative reports, provide for immediate access to electronic documents, and may have the capacity to provide other electronic tools such as decision support to ensure that required elements are included. Synoptic reports, by definition, create documentation in a format with discrete data fields where each type of desired information has a specific format in the report allowing for standardized collection, storage, and retrieval of this data after creation of the report. Templated notes, in contrast to pure synoptic reports, have a range of formats from more prescriptive, structured, and discrete content that is very similar to synoptic reports to being mostly free-text narrative with only basic structure for document sections.

In particular, synoptic reports can be constructed to contain key information elements and potential values or options for these elements that can be used for downstream uses (e.g., quality reporting or clinical research). Synoptic reports have been used extensively by other physician groups such as pathologists[41, 42] and have been proposed as a means to improve the accuracy, compliance, and ease of ORD[43]. Recently, synoptic reporting has also been proposed as a means not only to improve the quality of ORD but also potentially to function as an educational tool for residents in learning both ORD and key aspects of different surgical procedures[44].

Although EHRs are being rapidly adopted in the United States, there is limited information available about current practices with and pedagogy of ORD in surgical training programs and the role of EHRs with ORD. The goal of this study was to gain an updated understanding of educational practices of ORD in residency programs in different surgical sub-specialties. Additional related objectives were to ascertain the role of the EHR in ORD including use of synoptic operative reports and other electronic tools such as templates for documenting operative reports, as well as attitudes and opinions of program directors about ORD teaching for residency education.

2.3 Methods

A survey for surgical program directors was constructed to assess characteristics of training programs and their associated medical institution, information about the program director, current ORD teaching practices, institutional EHR adoption, use of and attitudes about synoptic operative reports, and attitudes towards ORD for residency education. In addition to closed multiple-choice and structured items, there were a number of open-ended questions aimed at assessing current training in ORD, attitudes towards formal ORD training, and use of synoptic reporting and other electronic tools such as templates for resident education.

Items were constructed using an iterative and consensus-based process that included input from all authors, including the surgery program residency director, chairman, and deputy chairman (JGC, SMV and DAR respectively), a surgeon-informatics faculty member (GMM), an informatics graduate student (NGF), and a surgery resident (NEB). Closed ended items included questions about the institution (university/academic, community/private, military), type of program (i.e., ACGME specialty), number of trainees annually, number of years as program director, number of years of practice of program director, use of electronic health records (yes, no, partially used), and specific aides used for operative report documentation (templates, memory aids/cards, electronic procedure-specific templates, other (specify)). "Yes/no" closed items included: availability of operative reports in the electronic health record, formal training of program directors in ORD, formal training for residents in ORD, use of synoptic reports or operative report templates for ORD, if ORD education is a priority, and if they perceive that their hospital loses reimbursement from resident ORD. A number of open-ended questions were also included as part of the survey to gather further information on attitudes and opinions of program directors on these topics (Table 2.1). After all survey items were vetted, several trials of the survey on SurveyMonkey™ were tested, reviewed, and refined to arrive at the final version. To encourage completion of the survey, the entire survey was kept short enough so that it could be completed within 5 to 10 minutes.

Table 2.1. Open-ended questions about operative report documentation teaching
What formal training on documenting operative reports is available in your program? Please comment upon formal training in operative report regulatory requirements. Please comment upon dictating/writing of operative reports by residents. How do you perceive procedure-specific templates (synoptic reports) for resident education?

An email requesting survey participation was sent electronically January 2012 to residency program directors for all ten approved Accreditation Council for Graduate Medical Education (ACGME) surgical training specialties (Colon and Rectal Surgery, Neurological Surgery, Obstetrics and Gynecology, Ophthalmology, Orthopedic Surgery, Otolaryngology, Plastic Surgery, Surgery, Thoracic Surgery, and Urology) using the available contact information from the ACGME website. To maximize the survey response rate, after the initial request was sent, two reminder emails were sent to those who had not yet responded and with corrected contact information for bounced or returned emails.

A content analysis of qualitative responses was performed to arrive at themes for open-ended questions using an iterative approach with three coders (NGF, NEB, GMM) until reaching saturation and responses were categorized at a respondant level. Data are presented as means, medians and ranges. Categorical variables were compared using Chi-squared analysis or Fischer's exact test, where appropriate. All statistical analysis was performed using SAS 9.1.3 (The SAS Institute, Cary, NC). University of Minnesota institutional review board approval was obtained and informed consent waived for this minimal risk study.

2.4 Results

Of 1,096 total surgical programs, 441 program directors (40%, 17.9 ± 8.8 years in practice) responded from Colon and Rectal Surgery (10, 2%), Neurological Surgery (44, 10%), Obstetrics and Gynecology (80, 18%), Ophthalmology (44, 10%), Orthopedic Surgery (56, 13%), Otolaryngology (50, 11%), Plastic Surgery (19, 4%), Surgery (81, 15%), Thoracic Surgery (8 2%), and Urology (39, 9%) training programs (Table 2.2). The programs in this sample reported 4.0 ± 2.5 residents annually and identified affiliations with university (383, 87%), community/private (44, 10%), and military (14, 3%) institutions.

Only 59 (13%) program directors reported formal ORD instruction in their own training. Although most program directors (295, 67%) considered ORD teaching a priority in surgical education and many (196, 44%) believe their hospital loses money from under-reported procedures with resident ORD, only a minority (76, 17%) currently provide ORD instruction. No differences were observed in instruction rates with different hospital-type training programs. Training programs with program directors who themselves were formally trained in ORD were more likely to offer ORD instruction (61% vs. 11%, $p<0.0001$), as were Obstetrics-Gynecology programs (Obstetrics-Gynecology-35% vs. Surgery-18%, Neurosurgery-16%, Ophthalmology-14%, Orthopedics-14%, $p<0.05$ for each).

Table 2.2. Characteristics of Respondents to Operative Report Documentation Survey	
	N (%)
Respondents	441 (40%)
Years in practice (mean ± SD)	17.9 ± 8.8
Primary hospital affiliation	
University-based	383 (87%)
Community/private	44 (10%)
Military	14 (3%)
Annual number residents (mean ± SD)	4.0 ± 2.5
Specialty	
Colon and rectal surgery	10 (2%)
Neurological surgery	44 (10%)
Obstetrics and Gynecology	80 (18%)
Ophthalmology	44 (10%)
Orthopedic Surgery	56 (13%)
Otolaryngology	50 (11%)
Plastics Surgery	19 (4%)
Surgery	81 (15%)
Thoracic Surgery	8 (2%)
Urology	39 (9%)

Table 2.3. ORD Teaching Modalities	
Response	%
ORD Teaching Aids	75%
Templates for dictation	
Old operative reports	
Formal lectures	28%
Boot-camp/Intern orientation	
Regular curriculum	
Simulation labs /dictation workshops	
One-on-one formal sessions (mentoring)	13%
Ad-hoc informal one-on-one sessions	29%

ORD, operative report documentation

In qualitative responses, program directors described formal lectures, ORD teaching aids, and one-on-one teaching sessions as ORD instruction modality categories (Table 2.3). Program directors also perceived a number of barriers to ORD training (Table 2.4), many of which relate to billing, reimbursement, and perceived quality issues with resident dictations; time considerations; perceptions that ORD teachings is not necessary; and a lack of established teaching pedagogy for ORD teaching.

Table 2.4. Perceived barriers to ORD Teaching	
Response	%
Residents not permitted to dictate	55%
Medico-legal environment	
Pressure for reimbursement (coding and billing)	
Attendings prefer dictating complex cases	
Perception ORD teaching is not needed	34%
No time for ORD teaching	15%
Deficiencies in teaching methodology	21%
Teaching done mostly on an individual basis	
No previous training	

ORD, operative report documentation

Most training program institutions (91%) had adopted an EHR system with operative reports available electronically. Ophthalmology (31%) and Obstetrics-Gynecology (30%) programs had relatively higher rates of synoptic reporting for ORD, while other surgical specialties had lower overall rates (18%). When asked about synoptic reporting as a teaching modality for ORD, 70 (16%) program directors provided comments, which were mostly negative (49, 70%) towards the use of synoptic reporting generally and for teaching ORD to residents (Table 2.5). Perceived positive aspects of synoptic reporting for ORD related to operative report quality and completeness, time-savings, potential benefit with resident education, and usefulness for documenting common/simple procedures (Table 2.6).

Table 2.5. Perceived weaknesses of synoptic operative reports/templates and their use in resident education

	%
Perception of less educational value	51%
Do not have to demonstrate knowledge and familiarity with procedure	
Encourages trainees not to think/Discourages independent thinking	
Trainees may not recognize deviations	
May develop a dependency upon synoptic operative reports (issue when removed)	
May not capture important information	31%
Can be built poorly or from bad examples	
Miss or have less case- and patient-specific details	
May not include decision-making and findings	
More prone to error or inaccurate recording of case	
May fail in certain, less formulated circumstances	24%
May not work for some specialties due to case mix and variety	
May fail with more complex procedures	
May result in poor quality documentation	45%
Encourages copy and paste	
Encourages adding less or no additional information	
May creates difficult to read documentation	

Table 2.6. Perceived strengths of synoptic operative reports and their use in resident education	
	%
Educational value	32%
Particularly helpful in teaching basic procedure steps for junior residents	
Helps to teach what to expect and to recognize deviation from typical case	
Improved Quality	71%
Completeness	
Works as a check list/reminder	
Compliance	
Better billing and coding	
Saves time, more efficient	35%
Effective for basic/simple/common procedure steps	23%

2.5 Discussion

While some patient care and surgical management knowledge and skills are well-defined expectations of residency education, practical aspects of surgical practice including those surrounding operative reports and other clinical documentation have traditionally been less consistently taught and tested in residency training. This survey of ACGME surgical program directors provides additional insights into the current status of

ORD teaching and describes some of the challenges with ORD in resident education. Despite considering ORD teaching an educational priority, only a fifth of responding programs provide formal ORD instruction to trainees. This is not surprising, in light of the increasing competing demands of postgraduate trainees and their instructors with competency-based training, simulation, and the ever-increasing amount of didactic material required of trainees with more limited work hours. In addition to a variety of approaches enumerated for ORD teaching, program director respondents also described numerous perceived barriers to ORD in residency education.

While synoptic reporting for ORD is increasingly reported in the literature as a means to improve the completeness and efficiency of ORD, our findings suggest that its penetrance remains low and the optimal role of synoptic reporting for ORD in surgical education is unclear. Some of the elements that synoptic reports or the use of electronic templates can ensure include ensuring proper descriptions and inclusion of positioning, preparation, exposure, key anatomy, tissue characteristics, resections, re-approximations, and conditions of the closure, as well as completeness of required operative report sections. As program directors and other surgical educators consider ORD instruction and the idea of improving ORD completeness and quality, these types of elements, as well as other domain-specific data with respect to surgery can be leveraged in synoptic report content.

This study suggests that there are significant opportunities to improve surgical education by focusing both in the art of ORD with dictation and potentially to utilize synoptic reporting as a tool for routine procedures. Our findings concomitantly suggest, moveover, that there remains significant reticence to the use of templates and other aids for ORD, with program directors citing this as a potential means to not learn the more intricate and detailed parts of more complicated cases. It is also clear that program directors view the use of electronic ORD with synoptic reporting as potentially negative to the process of ORD education.

The use of multi-modality education and competency-based learning as opposed to rote didactic lectures is increasingly espoused with medical training[45]. In qualitative feedback of potential ORD teaching modalities, respondents described standard lectures but additionally included the idea of their incorporation outside of the traditional curriculum such as simulation laboratories, dedicated dictation workshops, and boot-camp/intern orientation. While some surgical subspecialties (i.e., neurosurgery and orthopedics) espouse the use of intern "boot-camp" as an opportunity for focused teaching of surgical trainees, our study does not provide objective guidance on the best means for ORD education. Other sources of ORD education identified by program directors include the use of old operative reports, pre-formed templates as examples of operative reports, and informal one-on-one learning from attending surgeons. Interestingly, technology-based media such as electronic resources were not described with the exception of the idea of incorporation of ORD teaching into simulation laboratories. As part of this, an important consideration with these approaches will be the need to address some of the perceived barriers, particularly those making it practically difficult for residents to perform ORD such as billing pressures, the complexity of cases, and medico-legal concerns.

While synoptic reports are increasing being described in the literature as potentially effective for ensuring operative report completeness and extracting structured data from operative reports for secondary uses such as clinical research and surgical registries, our study demonstrates that synoptic reports for ORD are largely viewed negatively in the context of resident education. A large number of responses from program director respondents reflected an unsubstantiated belief that traditional dictation allows for a “replaying” and reinforcement of the steps, knowledge, and skills for a particular operative procedure and that the use of synoptic reports with preformed templates could potentially discourage independent thinking and make trainees dependent upon templates, as well as less likely to recognize variations or subtleties in procedures. Several illustrative comments are included in Table 2.7. Other concerns related more broadly to synoptic operative reports include concerns that templates may do a poor job for complex procedures, fail to meet the needs of certain specialties (e.g., plastic surgery) that may perform a wide variety of non-routine procedures, or may result in documentation that is more difficult to read.

Table 2.7. Example responses about synoptic reports for operative report documentation and their role in surgical resident education
"Thinking it through and reliving the operation with dictation is very helpful for a while [in training]. Templates help with completeness and compliance but can make it automatic and less educational."
"I think that the resident being able to "run" the case through their head as they are dictating demonstrates knowledge and familiarity with the procedure. Synoptic reports can be a good crutch from a compliance/quality of document standpoint, but not necessarily from the education/learning standpoint."
"For procedures which are routine probably helpful as a reminder, I am concerned that templates always lead to an implication that the template is the ultimate complete needed information so what is not on the template is not necessary. i.e. rather than raising the bar, it lowers the bar"
"Synoptic reports encourage them to be lazy. If there is any deviation from the procedure described in the template, they rarely make note of it or alter the template. On the other hand, I have seen operative reports that have obviously been cut and pasted from things found on Internet searches that in no way represent what occurred in our OR..."
"The templates help residents learn what to expect, and they use these not only for dictation but also when they are reviewing how to do a case they may not be familiar with. I think it is helpful to have a standard dictation for them to learn one way to do something they can then learn when to deviate."

In considering the large proportion of negative perceptions by program directors towards electronic documentation with synoptic reports or electronic templates, it is unclear how many of these opinions were grounded in fact as opposed to opinion or negative perception towards these technologies. Many of the provided opinions require further evidence to truly validate. The surgical literature has consistently demonstrated value and significant benefits from synoptic operative reports including completeness, quality and research benefits for secondary use, billing benefits, and educational benefits. But clearly, this survey effectively demonstrates that significant adaptive issues remain in there being more widespread introduction and use on a larger scale.

There was significant variation in responses about ORD education practices between surgical subspecialties. For instance, Obstetrics-Gynecology programs were in most cases twice as likely compared to other surgical specialties to formally provide ORD to residents. A related finding linked to surgical specialty was the observed higher rate of ORD teaching to residents if the program director was formally trained in ORD. While the reported use of synoptic reporting for ORD remains low overall (less than 20%), a similar pattern of specialty differences with synoptic operative reports was seen, with Obstetrics-Gynecology and Ophthalmology programs having greater rates of use. We speculate that these two specialties have a greater tradition of ORD teaching and that these surgical specialties have a relatively greater proportion of routine procedures with minimal variation lending well to the use of synoptic reporting or electronic procedure templates.

A number of important limitations should be noted. While a response rate of 40% is considered good for survey data, this leaves room for bias in our results. Importantly, this survey was not a validated instrument with psychometric or externally validated calibration. Instead, items were constructed using a consensus-based and iterative process with content experts in surgical training, qualitative analysis, and health informatics. In our analysis of looking at the potential best means of performing electronic documentation, our survey did not make significant distinctions between electronic templates and synoptic reports, which would be important to better understand subtleties, strengths, and weaknesses of these different approaches. Future work is needed to both systematically characterize current evidence as well as provide original evidence into understanding the optimal role of synoptic reporting for ORD including its role in surgical education.

2.6 Conclusion

While most surgical program directors consider ORD teaching an educational priority, incongruence exists between its perceived value and adoption into surgical training. The efficacy of different ORD pedagogical methods remains unclear. While ORD with synoptic reporting is currently not common in most surgical specialties, it will likely have greater future importance with EHR adoption and the need for secondary use of information from operative reports.

2.7 Acknowledgments

The author would like to acknowledge support from the American Surgical Association Foundation (GM, NB) and Institute for Health Informatics Seed Grant (GM).

Chapter 3: Analyzing Operative Report Structure in Development of a Section Header Resource

Genevieve B. Melton, MD[a, b], Yan Wang, MS[b], Elliot Arsoniadis, MD[a, b], Serguei V.S. Pakhomov, PhD[b,c], Terrence J. Adam, PhD[b,c], Mary R. Kwaan, MD[a], David A. Rothenberger, MD[a], Elizabeth S. Chen, PhD[d]

[a]*Department of Surgery, University of Minnesota, Minneapolis, MN, USA*

[b]*Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA*

[c]*College of Pharmacy, University of Minnesota, Minneapolis, MN, USA*

[d]*Center for Clinical and Translational Science, University of Vermont, Burlington, VT, USA*

Submitted December 2014 to *MEDINFO Proceedings 2015* for consideration as a manuscript in *Studies in Health Technology and Informatics.*

3.1 Summary

Operative reports contain essential details of surgical procedures and are an important form of clinical documentation. Sections within operative reports segment and provide high level note structure. We evaluated the HL7 Implementation Guide for Clinical Document Architecture Release 2.0 Operative Note Draft Standard for Trial Use (HL7-ON DSTU) Release 1 as well as Logical Observation Identifiers Names and Codes (LOINC®) section names on 384 unique section headers from 362,311 operative reports. Overall, HL7-ON DSTU alone and HL7-ON DSTU with LOINC® section headers covered 66% and 79% of sections headers (93% and 98% of header instances), respectively. Section headers contained large numbers of synonyms, formatting variation, and variation of word forms, as well as smaller numbers of compound sections and issues with mismatches in header granularity. Robust operative report section mapping is important for clinical note interoperability and effective use of operative reports by natural language processing systems. The resulting operative report section resource is made publicly available.

Keywords: Surgical Procedures, Operative; Vocabulary, Controlled; Medical History Taking/methods; Quality and Safety; Documentation; Electronic Health Records

3.2 Introduction

Operative reports are traditionally created at the completion of a surgical procedure by the primary surgeon who recalls the procedure details and dictates these into a narrative that is subsequently transcribed. Effective operative report documentation is important for assessing surgical quality [5], billing and medico-legal issues [6, 7], and other secondary uses of operative reports [4]. With increasing adoption of electronic health record (EHR) systems, operative reports and other clinical documents are increasingly generated and immediately available. EHR systems also enable other mechanisms for note generation, including voice to text software, typed notes, synoptic reporting, and templated notes.

Synoptic reports are used to create documents with discrete data fields whereby desired information from the note can be collected, stored, and retrieved in a standardized fashion. In contrast, templated notes range in the amount of structure that they contain, including some having highly prescriptive and structured formats to others having mostly free-text narrative with primarily document section structure alone. While dictation and transcription remains the most common mechanism for operative report creation, synoptic reports and templates are increasingly used in surgery for operative report creation and appear to encourage improved completeness of these documents [43, 46]. Independent of the mechanism used to create the document, section headers in operative reports provide high level structure and serve as “containers” which provide context to the text within the given section [21].

Previous research has examined section headers in clinical notes, with some work exploring the task of automated classification of sections in documents. As an initiative with the OpenGALEN project, Mori et al. utilized "tags" in 5 clusters (*Nature, Safety Context, Interpretation, Intention, and Organization*) and evaluated this approach to classify 600 section headings [21]. Denny et al. reported on a terminology of document sections in "history and physical" (H&P) notes and developed an associated section terminology with Logical Observation Identifiers Names and Codes (LOINC®) mappings [22]. These authors later developed an algorithm, "SecTag", to identify and label section headers and section boundaries in H&P notes. Similarly, others have utilized predefined sections to aid in a number of natural language processing tasks, such as problem list extraction and named entity recogition [24-26].

In the United States, H&P note formats are largely governed by Evaluation and Management (E/M) documentation, which provides guidelines for assessing adequacy of documentation for each patient encounter resulting in a "level of service" and justification for a patient bill [47]. Operative reports, in contrast, are not covered by E/M, and the Joint Commission [48] and Centers for Medicare & Medicaid Services (CMS) [49] have specified criteria for operative reports including information on suggested contents and note sections. Overall, Joint Commission designates eleven required elements for operative reports: name(s) of primary surgeon/ physician and assistants, pre-operative diagnosis, post-operative diagnosis, name of the procedure performed, findings of the procedure, specimens removed, estimated blood loss, date and time recorded, indications for the procedure, intra-operative complications, and a full description of the procedure.

The Health Level 7 (HL7) Structured Documents Workgroup seeks to develop structured healthcare document standards to promote document and data interoperability. This group has created implementation guides for clinical documents including one for operative reports, the *Implementation Guide for HL7 Clinical Document Architecture (CDA) Release 2.0, Operative Note Draft Standard for Trial Use Release 1* (HL7-ON DSTU) [50]. This specification includes Level 1 (header constraints), Level 2 (section level constraints of the structuredBody of the ClinicalDocument), and Level 3 (entry level constraints within a section; specifying only the Plan section of Operative reports) requirements. The HL7-ON DSTU was created using a variety of data sources and expert opinions including subject matter expert input, summary statistics from sample operative reports, Joint Commission Operative Note Requirements: Standard IM.6.30 [48], and CMS Operative Report Requirements [49]. Where possible, the HL7-ON DSTU utilizes existing clinical statement entries and Continuity of Care Document (CCD) elements, and other Implementation Guide templates. As such, some items considered clinical statement entries in other contexts are treated as sections. The HL7-ON DSTU also maps section headings using LOINC® where available.

To improve available resources and tools for clinical natural language processing[1] specifically for operative reports [16, 17] and using our experience with clinical standards evaluation [51-53], we sought to use the HL7-ON DSTU and LOINC® section codes to represent operative report section headers and to develop a resource for operative report section headers.

[1] http://healthinformatics.umn.edu/research/nlpie-group

3.3 Materials and Methods

3.3.1 Study Overview

Figure 3.1 provides a high level summary of this study. The HL7-ON DSTU was examined and section headers with associated LOINC® codes were collected along with potential document section headers from LOINC®. All operative reports over a 4-year period from University of Minnesota-affiliated Fairview Health Services, which includes an academic medical center, five community hospitals, and three ambulatory surgery centers, were collected from a full range of general surgery and surgical subspecialties, and section headers were extracted. Headers were mapped and coded to eliminate non-section headers, assess section header variation, and identify granularity issues with mapping to structured sections. The section headers and mappings were combined into a resource for future use. Institutional review board approval was obtained and informed consent waived for this minimal risk study.

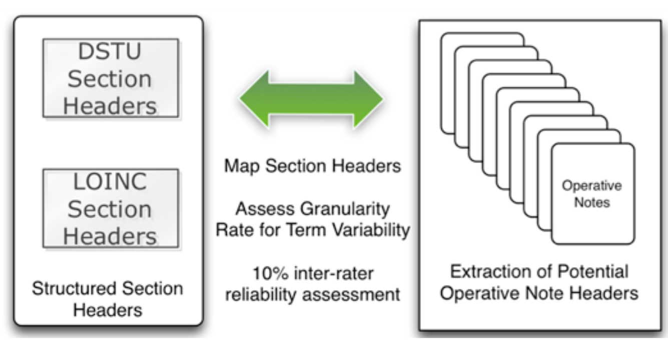


Figure 3.1. Overview of Study

3.3.2 HL7-ON DSTU Section Header Extraction

Level 1 and Level 2 HL7-ON DSTU constraints were used in this evaluation. Level 1 header elements included information commonly contained in sections for operative reports (e.g., "Primary Performer" - typically referred to as the "Surgeon" or "Primary Surgeon" in operative reports is a header element). Required and optional operative report section names were used along with LOINC® section mappings, section descriptions and suggested information about each item. Level 3 constraints were excluded from the analysis as were Level 1 header elements not related to section headers (such as elements to encode the overall operative report specification).

In addition to the HL7-ON DSTU section headers, LOINC® section header names, codes, and descriptions were collected by extracting entries of "DOCUMENT_SECTION", resulting in 121 distinct sections from LOINC® Version 2.42. Unmapped terms with this list were also mapped to "CLASS" entries of "H&P.HX", "H&P.HX.LAB", "H&P.PX", or "H&P.SURG PROC" with free text search in a second step.

3.3.3 Operative Report Section Header Evaluation

Potential operative report section headers were automatically extracted using heuristic rules including the use of capitalization, semi-colons and hyphens, and line-breaks. Frequencies of each potential header were calculated and a cut-off of 100 was used in coding headers. The eliminated set of headers accounted for less than 2% of overall entries.

Table 3.1. Coding for Operative Report Section Headers	
Coding and Explanation	**Example(s)**
Not a Section Header: *term is not a known section header or document title*	"SEE RADIOLOGY REPORT FROM"; "OPERATING ROOM"
Document Title: *term is a document title*	"BRIEF OPERATIVE NOTE"; "OPERATIVE REPORT"
Document Header Information: *term is other document information*	"PATIENT IDENTIFICATION"; "DEPT"
Correct Section Header: *term is preferred section in* HL7-ON DSTU	"ANESTHESIA"; "COMPLICATIONS"; "SURGERY DESCRIPTION"
New Section Header: *term is section not in* HL7-ON DSTU	"CROSS-CLAMP TIME"; "PREOPERATIVE HISTORY"
Synonym: *term is alternate synonymous section name (new or known section)*	"SURGERY DESCRIPTION" (*preferred)* vs. "OPERATION DESCRIPTION"
White-Space, Formatting, Misspelling: *white-space, formatting, or misspelling*	"POST-OPERATIVE" vs. "POSTOPERATIVE"
Word Form Variant: *term is word form variant to preferred or synonym term*	"PREOPERATIVE DIAGNOSIS" vs. "PREOPERATIVE DIAGNOSES"
Abbreviation: *term is an abbreviation*	"EBL" vs. "ESTIMATED BLOOD LOSS"
Compound Section Header: *two or more sections designated*	"OPERATIVE INDICATIONS AND CONSENT"

Table 3.1 (continued) Coding for Operative Report Section Headers	
Coding and Explanation	**Example(s)**
Same Granularity: *term has same granularity as mapped header*	"SPECIMENS" vs. "SPECIMENS REMOVED"
Less Granularity: *term has less granularity than mapped header*	"DIAGNOSES" vs. "POSTOPERATIVE DIAGNOSIS" or "PREOPERATIVE DIAGNOSIS"
More Granularity: *term has more granularity than mapped header*	"ARTHROSCOPIC FINDINGS" vs. "FINDINGS"

Each header was then manually designated as a document title, a potential section header, or not a potential section header by two coders. Potential section headers were then mapped to a HL7-ON DSTU standard section name and the provided LOINC® section specification, if available. Table 3.1 provides an overview and examples of section header codings.

Each extracted potental section header name was then coded to designate if the header was: the preferred section header name; a new section header from the HL7-ON DSTU; a synonym; had white space, misspelling, or formatting variation; a word form variant; an abbreviation; a compound section header; or a header with additional granularity compared to the HL7-ON DSTU suggested section header specifications. Mapped entries were compared to the section name and coded according to their granularity as: equal, greater, or less granularity. Finally, if a section header could not be mapping to the HL7-ON DSTU, the section header was mapped to LOINC® [54].

Approximately 10% of all mappings were evaluated by both coders (a surgeon and informatician (GM) and a surgeon and informatics graduate student (EA)) in order to assess inter-rater agreement. Percent agreement and Kappa were calculated for mappings to document titles, non-section headers, and section headers; coding for HL7-ON DSTU section headers; and assessment for entry variation (e.g., word forms and synonyms).

The section headers and subsequent mappings were used to create a resource of operative report section headers to improve the reuse of these notes. The resource contains section header terms, term mapping to HL7-ON DSTU and LOINC, and information granularity of mappings.

3.4 Results

3.4.1 HL7-ON DSTU Section Header Extraction

Operative report section names and LOINC® section mappings for all designated sections in the HL7-ON DSTU were collected. This included 3 main header elements related to operative report sections which resolved to 6 potential section elements, 12 section names (8 required), and 4 subsection names (all options). Table 3.2 contains example entries for these 18 elements from the HL7-ON DSTU.

Table 3.2. Example Operative Report Sections from HL7-ON DSTU		
Section	**LOINC Code**	**Component Name**
Consent (O, H)	N/A	CONSENT FOR SURGICAL PROCEDURE
Anesthesia (R, Sec)	8724-7	SURGICAL OPERATION NOTE ANESTHESIA
Indications (O, Sec)	10217-8	SURGICAL OPERATION NOTE INDICATIONS
Implants (O, Sub)	55122-6	SURGICAL OPERATION NOTE IMPLANTS

R:required;O:optional;H:header;Sec:Section; Sub:Subsection

3.4.2 Operative Report Section Header Evaluation

Automated extraction of headers from 362,311 operative report section resulted in 2,999,414 entries. Removal of entries with a frequency of less than 100 (n=52,054) resulted in 2,947,360 (98.3%) total entries and 476 unique entries.

Initial coding demonstrated that 8 headers (6,975 instances) were document titles, 7 headers (15,525 instances) were document header information, and 77 headers (26,189 instances) did not represent valid potential section headers (Table 3.3). Of the remaining 384 section headers (2,898,771 instances), 66% section headers (93% instances) mapped to the DSTM and after including LOINC® sections for the remaining elements, successful mappings were obtained for 79% of headers (98% of instances). We also observed large numbers of synonymous terms, normalized variants and other formating associated with section headers.

Table 3.3. Operative Report Section Header Findings

	Headers N (%)	Instances N (%)
Overall	476 (100)	2,947,360 (100)
Document Title	8 (2)	6,875 (0.2)
Header Information	7 (1)	15,525 (0.5)
Not a Section Header	77 (16)	26,189 (0.9)
Section Header	384 (81)	2,898,771 (98)
Map to HL7-ON DSTU	255 (66)	2,735,563 (93)
Granularity		
Same Granularity	179 (70)	2,132,446 (78)
Greater Granularity	65 (25)	594,605 (22)
Less Granularity	11 (4)	8,512 (0.3)
Variation in Terms		
Normalized Word Form	63 (25)	328,090 (12)
Formatting Variation	18 (7)	318,233 (12)
Synonyms	177 (69)	1,203,053 (44)
Abbreviation	18 (7)	256,103 (9)
Map to HL7-ON DSTU or LOINC®	304 (79)	2,833,094 (98)
Mapping Failure	80 (21)	65,677 (2)
Multiple Sections	22 (28)	10,607 (16)
No Mapping	58 (72)	55,070 (84)

Table 3.4. HL7-ON DSTU Section Header Mappings

Total Headers	Unique Headers	Section Name	Same Granularity
211,303	4	Anesthesia	99.8%
93,408	6	Complications	100%
16,898	10	Disposition	58.5%
115,307	5	Estimated Blood Loss	100%
2,945	7	Implants	100%
197,127	20	Indications	100%
18,192	20	Operative Note Fluids	79.9%
565,596	33	Operative Note Surgical Procedure	100%
22948	11	Plan	88.8%
283	1	Planned Procedure	100%
327,151	13	Postoperative Diagnosis	100%
333,307	14	Preoperative Diagnosis	99.7%
203,494	21	Primary Performer	0%*
383,174	31	Secondary Performer	0%*
32,530	15	Specimens Removed	99.6%
100,374	28	Surgery Description	100%
1,100	2	Surgical Consent	100%
22,208	2	Surgical Drains	100%
86,415	11	Surgical Operation Note Findings	99.5%
484	2	Surgical Date of Procedure	100%
1319	2	Surgical Procedure Duration	100%

Table 3.4 summarizes mappings to the HL7-ON DSTU including numbers of terms mapping to different headers and the proportion of terms that mapped with equal granularity. There was significant variability in expression for many HL7-ON DSTU section headers, and differences in granularity particularly for section headers for primary performer and secondary performer. An analysis of the 30 most common section terms mapped to the HL7-ON DSTU in all but one case, and the remaining header was a LOINC® section mapping (data not pictured).

In the overlap coding of 50 entries, percent agreement and Kappa for the initial mapping of document titles, non-section headers, and section headers was 100% and 1.00; the HL7-ON DSTU mapping agreement for section headers mapping was 92% and 0.94, respectively.

A number of section headers did not map to the HL7-ON DSTU or LOINC® section headers. A number of these appeared to be unique to operative reports such as "Tourniquet Time", "Sponge and Needle Counts", "Bypass Time", "Preoperative Antibiotics", and "Preoperative Status".

3.5 Discussion

As the demand for the extraction of meaningful information from more challenging clinical data sources such as clinical texts becomes an area of greater focus, operative reports and other clinical documents will be reused for a variety of purposes. These efforts aid quality improvement, research, and ultimately clinical data interoperability. This study examines a structured document standard for operative reports, which includes Level 1 (Header) and Level 2 (Section) specifications. Section headers from the HL7-ON DSTU and additional LOINC® sections headers were evaluated on headers extracted from a large number of operative reports from an integrated healthcare delivery system. The standards covered most header instances although amongst unique headers, about 20% did not map. We also observed a large amount of variation in the section header term expression including many synonyms, formatting variations, variation in word forms, and compound section headers within the corpus.

While the HL7-ON DSTU provides eight required section names and a small number of main header items that are conventionally sections in operative reports, our study demonstrates the wide variability in expressions of these elements, the frequent use of optional sections, including the 8 section/subsection elements designated in the HL7-ON DSTU as well as 49 section headers designated in LOINC® and not in the HL7-ON DSTU and 58 section headers unique to both the HL7-ON DSTU and LOINC®. As the HL7-ON DSTU authors note, the base specification for an operative report, like other clinical documents, is the HL7 CDA, Release 2.0, allowing for other sections not present in the HL7-ON DSTU to occur in operative reports. Further, despite the significant challenges with variability in expression of section headers present in our corpus, the "exact text of the section names are not mandated" by the HL7-ON DSTU.

Several of the unique sections that did not map to the HL7-ON DSTU or LOINC®, including "Tourniquet Time", "Sponge and Needle Counts", "Bypass Time", "Preoperative Antibiotics", and "Preoperative Status" may be valid optional section headers. Some of these are important elements for operative reports for certain subspecialities (e.g., cardiac surgery, transplant surgery, or vascular surgery). A further assessment of operative report sections in other centers may also be helpful for establishing the generalizability of the results of our study.

We also observed issues with respect to both granularity as well as variability in expressing different section headers. In particular, the section header "Disposition" which is standard to the HL7-ON DSTU had a number of entries that were more granular than the general header, including "Postoperative Condition" or "Prognosis". Similarly, both "Primary Performer" and "Secondary Performer" had wide amounts of variability to the amount of detail expressed. "Primary Performer" included more granular terms such as "Attending Neurosurgeon" and less granular terms such as "Physician", which is under specified enough that it is unclear whether this represents a Primary Performer or not. Similarly, "Secondary Performer" had mostly more granular terms, many of which were trainees including residents and fellows, as well as the designation of assistants and other providers involved with procedures.

With respect to variability in section header expression, many terms including "Operative Note Surgical Procedure" and "Surgery Description" had many terms to express the same section header. This was similarly the case with "Primary Performer" and "Secondary Performer", as just described. Surprisingly, sections like "Indications" and "Operative Note Fluids" also had wide variability with 20 different section terms for these two section headers each. We observed that while "Operative Note Fluids" is the section recommended for operative reports, surgeons were sometimes to describe the more significant resuscitative elements like blood products and colloid administration and instead used ad hoc section headers like “Components Used”.

While the majority of the section headers were fully specified by their name, there were some section headers where the content of the associated section was ambiguous. For instance, the section header “Procedure” or “Procedure(s)” in most cases designates “Operative Note Surgical Procedure”, which lists the procedure(s) performed by the surgeon, similar to the “Surgical Procedure” (Header) which provides coded enumaration of the procedures performed. However, in some cases, the section “Procedure” can be the section most commonly labeled with the operative report section “Surgery Description”, which described the procedure in detail. The disambiguation of these headers may be addressed in future studies at the semantic level with the contents of operative reports using machine-learning or other automated approaches. This also points out the large amount of variability in expression and practice with operative report composition.

There are several limitations to this study. Section headers were extracted from operative reports using a set of deterministic section segmenting rules, and some were likely missing in our analysis. Additionally, the study was conducted within a single regional hospital setting, and its findings could be further strengthened with future studies examining this question at other sites.

3.6 Conclusion

Structured document standards and well-formed section header designations are important for interoperability of clinical documents and natural language processing systems that consume these documents. We evaluated the HL7-ON DSTU specification for operative report section headers and LOINC®. Our findings confirm that most section headers are covered by the HL7-ON DSTU and LOINC®. However, there is a large amount of variability in section header expression, and a number of section headers specific to operative reports not currently present in these reseouces. These findings should be considered for future HL7-ON DSTU iterations and possibly for addition to LOINC®. The resulting section header resource can also be used for section header mappings for natural language processing systems.

3.7 Acknowledgments

This research was supported by National Institutes of Health (R01LM011364), Agency for Healthcare Research & Quality (R01HS022085), American Surgical Association Foundation Fellowship, and University of Minnesota Clinical and Translational Science Award (8UL1TR000114). The authors thank Fairview Health Services for support of this research.

Chapter 4: Modeling the 'Surgery Description' in Operative reports

Genevieve B. Melton, MA, MD[a,b], Yan Wang, MS[b], Serguei Pakhomov, PhD[b,c], Elliot Arsoniadis, MD[a,b], Elizabeth Lindeman, BA[b], David A. Rothenberger, MD[a], Mary R. Kwaan, MPH, MD[a], Terrence Adam, MD, PhD[b,c], Rui Zhang, PhD[a,b]

[a]*Department of Surgery, University of Minnesota, Minneapolis, MN, USA*

[b]*Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA*

[c]*College of Pharmacy, University of Minnesota, Minneapolis, MN, USA*

In preparation for submission for peer review.

4.1 Summary

Introduction: Like other clinical documents, operative reports are split into sections, which provide high-level structure to these texts. The "Surgery Description" section in operative reports provides information about the surgical procedure and its main details. We hypothesized that there are high-level parts of the "Surgery Description" section, each with different types of information.

Methods: A group of surgeons examined and discussed the process of operative report documentation; identified high-level phases of the "Surgery Description" section of operative reports; and established broad definitions and annotation guidelines. Following this, two reviewers annotated a set of similar operative reports (100 laparoscopic cholecystectomy) and a set of 300 random operative reports to test generalizability. Feature selection methods with Latent Dirichlet Allocation (LDA) topic analysis and information gain (IG) were used with support vector machines (SVM) to classify these high level phases. Subtopics were analyzed to determine if they encompassed clinically important features for these high-level phases.

Results: Surgeon consensus identified five high-level phases to the "Surgery Description" section: *Preliminaries, Getting Started, Main Portion, Closure,* and *Epilogue*, as well as need for *Observations* to describe findings, unexpected events, and other important clinical information. Performance using LDA and IG with SVM for classifying the high-level phases had overall F-measures of 0.968 and 0.962 on laparoscopic cholecystectomy notes and 0.806 and 0.767 on random operative reports,

respectively. An initial subtopic analysis of the laparoscopic cholecystectomy notes demonstrated clear steps associated with each high-level phase.

***Conclusions*:** Surgeons defined high-level phases of the “Surgery Description” section, which could then be classified with automated techniques that could be leveraged for information extraction and other secondary uses of operative reports.

4.2 Introduction

An operative report is a clinical document created at the conclusion of a surgical procedure that details information about the performers (i.e., surgeon or other proceduralist), diagnoses, and name of procedure(s); the techniques, instruments, and materials used; and the associated findings and intraoperative complication(s). Traditionally, following a surgical procedure, surgeons use voice dictation followed by transcription into text to create an operative report. Operative reports are also generated using electronic templates and synoptic reporting[43, 46], and represent core documents within electronic health record (EHR) systems with the primary purpose of documenting surgical care and billing the surgical procedure.

In the United States, history and physical examination notes (H&P) and progress notes are billed according to Evaluation and Management (E/M) documentation guidelines by Centers for Medicare & Medicaid Services (CMS), which assess adequacy of documentation for each patient encounter resulting in a "level of service" and justification for a patient bill[47]. In contrast, operative reports and other procedures are not billed using E/M and instead billing for procedures uses Current Procedure Terminology (CPT)[2] with associated case modifiers as appropriate for professional billing of surgeon services.

In addition, Joint Commission Operative Note Requirements (Standard IM.6.30)[48] and CMS Operative Note Requirements[49] have recommended criteria for operative reports including information on suggested contents and note sections. In particular, the

---

[2] http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt.page (accessed December 24, 2014).

Joint Commission has eleven required elements for operative reports: name(s) of primary surgeon/ physician and assistants, pre-operative diagnosis, post-operative diagnosis, name of the procedure performed, findings of the procedure, specimens removed, estimated blood loss, date and time recorded, indications for the procedure, intra-operative complications, and a description of the procedure.

The “Surgery Description” section of operative reports contains detailed content about the particular surgical procedure. Information extraction from this section is highly dependent on successful use of clinical natural language processing (NLP). Surprisingly, only a small number of NLP studies have focused on characterizing this text more deeply, including work identifying action verbs and semantic frames in the “Surgery Description” section[16, 17]. Our goals with this study were to explore the text of potential high-level phases of the “Surgery Description” section in operative reports. Furthermore, we sought to explore the efficacy of automated techniques to classify these high-level phases to further segment this key document section within operative reports as a potentially important intermediate step to characterizing and utilizing operative reports more effectively for secondary analysis including information extraction and NLP.

4.3 Background

4.3.1 Operative report “Surgery Description” Section

Unlike other clinical texts and sections of the operative report, the “Surgery Description” section describes a set of observations and actions performed on the patient.

The following excerpt from an operative report describing an appendectomy performed for perforated appendicitis is illustrative of this section:

> *"The patient was brought to the Operating Room. After the induction of suitable general anesthesia the abdomen was sterilely prepped and draped. An oblique right lower quadrant (incision) was made. The peritoneal cavity was entered in a muscle splitting fashion. There was turbid yellow fluid in the right lower quadrant and this was aspirated. The cecum and appendix were mobilized into the wound. The appendix was grossly suppurative with gangrenous changes at its base. The appendiceal blood supply was clamped, divided, and ligated. The base of the appendix was crushed and ligated. The appendix was removed…. Dr. XXX was scrubbed and present for the entire procedure."*

As exemplified in this text, the "Surgery Description" section can be conceptualized and modeled as a chronological set of actions and observations that comprise the procedure. In this example, the section includes information about bringing the patient into the operating room, starting anesthesia, and preparing the operative site. The text then gives details about the incision and entry into the abdomen (the "peritoneal cavity") followed by the observation of finding infected "turbid" fluid. The note then details standard described maneuvers to dissect the appendix, followed by more description of the appearance of the perforated appendix, and finally dividing and excising the appendix from the colon.

Efforts aimed at supporting interoperability of clinical documents include the Health Level 7 (HL7) Structured Documents Workgroups, which are aimed at expanding the Clinical Document Architecture (CDA) specification and provide structured document guidelines including header and section specifications with recommended contents for different types of documents. One of these workgroups has provided an implementation guide for operative reports (*Implementation Guide for HL7 CDA Release 2.0, Operative Note Draft Standard for Trial Use Release 1* (HL7-ON DSTU))[50], which provides consensus operative report structure that takes into account Joint Commission and CMS Operative Note Requirements. With the HL7-ON DSTU, the "Surgery Description" section is defined as follows:

> The Operative Note Surgery Description section records the particulars of the surgery with an extensive narrative and may include surgical site preparation, pertinent details related to sedation/anesthesia, measurements and markings, waiting times, incisions, surgical approach, instrumentation, sponge counts, tissue manipulation, wound closure, sutures used, and vital signs and other monitoring data. Complications may be recorded in this section. Local practice often identifies the level and type of detail required based on the procedure or specialty.

The HL7-ON DSTU also has a number of conformance statements in defining the "Surgery Description" section, including required and optional recommendations for the contents of this section:

- CONF-OP-32: The Surgery Description section SHALL include a statement regarding whether or not a sponge and needle count was completed.
- CONF-OP-33: The Surgery Description section SHOULD include a statement regarding whether or not an instrument count was completed.
- CONF-OP-34: If the Operative Note Fluids section is NOT present, there MAY be a statement in the Surgery Description section providing details of the fluids administered or explicitly stating there were no fluids administered.
- CONF-OP-35: If the Surgical Drains section is NOT present, there MAY be a statement in the Surgery Description section providing details of the drains placed or explicitly stating there were no drains placed.

4.3.2 Topic Analysis

There are a number of methods used in automated topic analysis, which are established methods aimed at extracting thematic structures from documents[55, 56]. In most cases, topic modeling will identify ranked lists of words (topics or subtopics) based on a pre-defined number of topics. While topic modeling has been used in general English NLP extensively, it has been used to a limited extent with clinical NLP[57-61] and not specifically to the "Surgery Description" section of operative reports.

A widely used and well-established technique is Latent Dirichlet Allocation (LDA), which is based upon a statistical model[62]. LDA assumes that a given text can have multiple topics, which can be represented as probability distributions over words. LDA is also a generative graphic model, and it can discover underlying topic structures of texts. The resulting keywords for each topic from the topic-term matrix generated from LDA represent good feature candidates, which are helpful in discovering latent topics, as well. Information gain (IG), which tests entropy changes of the system (such as with topic classification), is also commonly used as a method for feature selection.

4.4 Methods

Surgeons were asked to participate in a session examining the process of surgical procedures and operative report documentation to identify high-level phases to the "Surgery Description" along with associated definitions and annotation guidelines of what types of content belonged in each high-level phase of the procedure. This was done in two iterations with the four surgeons. The core assessment consisting of questionnaire composed of 5 items with the first session (Table 4.1). Also, 15 random operative reports were also provided to the surgeons as examples of how operative reports can vary by surgeon author and by procedure. A summary of responses after the first session was compiled along with potential high-level phases. The second iteration was focused on defining which types of contents and text belonged to each phase – which were then used to form annotation guidelines.

Table 4.1. Questions about operative report documentation and the "Surgery Description" section.
1. How do you author and create operative reports in your practice? 2. What high level phases or groups of actions do you include in the "Surgery Description" section in documenting a surgical procedure? 3. How could you potentially group these high level items further in Item 2? 4. How would you define the content of the high level groups defined in Item 3? 5. Are there any additional items that a surgeon documents in the "Surgery Description" section not included in previous items, and what are they?

The initial set of phases and descriptions derived from the surgeon sessions composing annotation guidelines for operative reports was then modified after performing annotation on a small set of 10 operative reports with two annotators (EL and GM). The two corpora were composed of 100 laparoscopic cholecystectomy notes and 300 random operative reports. The "Surgery Description" section was annotated for each high-level phase at the sentence level. If a sentence or statement contained more than one phase, it was labeled with all of the appropriate phases. The remaining notes were annotated and sentences were reviewed and consensus established by the two annotators on any sentences where there was a question on an individual annotation. Inter-annotator agreement was assessed with an overlap of two annotators (EA and EL) on a set of ten laparoscopic cholecystectomy and 30 random operative reports at a sentence level, assessing percentage agreement and Kappa statistic. Since sentences could have more than one annotation, Kappa was calculated for each phase and percent agreement was calculated both at the overall sentence level and for each high-level phase.

Topic modeling of the “Surgery Description” section was performed. Since topic modeling assumes that a document (in this case the “Surgery Description” section) is a mixture of topics and a topic contains words, our goal was to ascertain common high-level topics across operative reports. Each sentence was treated as an individual document. We removed stopwords[63] and used lexical variation generation (LVG)[64] to lexically normalized word tokens.

LDA with Gibbs Sampling (iteration = 1000) was implemented using the Stanford Topic Model Toolkit (TMT-0.4.0)[65]. One of the parameters to train the model was the number of topics, which was varied from 200-800. IG was also used to reduce word features from all words (baseline). IG was then used to rank topic words and to filter top numbers of keywords (varied from 100-900) as features for classification.

In addition to the topic word features with LDA and the filtered keywords with IG, the relative position in tenths (e.g., 1/10 of the whole length of the section) was included as a feature for sentences in the section of a given document. Supervised machine-learning with support vector machines (SVM) was then implemented using 10-fold cross-validation on laparoscopic cholecystectomy and random operative reports. The performance of SVMs was reported using precision, recall and F-measure for each high-level phase and overall, defined as:

Precision = TruePositive / (TruePositive + FalsePositive)
Recall = TruePositive / (TruePositive + TrueNegative)
F-measure = 2 × Precision × Recall / (Precision + Recall)

Finally, for different high-level phases of the section, LDA was implemented to investigate subtopics on the laparoscopic cholecystectomy notes, and a surgeon reviewed topic words manually for each of the high-level phases. University of Minnesota institutional review board approval was obtained and informed consent waived for this minimal risk study.

4.5 Results

Surgeons identified five high-level phases to the "Surgery Description" section: *Preliminaries, Getting Started, Main Portion, Closure,* and *Epilogue*. While the names of these were different for several of the surgeons, the definitions were similar. With discussion, these five informal names were agreed to along with the definitions, as summarized in Table 4.2. In addition to the high-level phases, all four surgeons also agreed that the "Surgery Description" contains information about events, anatomy, or findings observed during the procedure, which can transcend the phases of the surgery and may occur within any of the phases. Thus, *Observations* was a component of the "Surgery Description" section describing findings, unexpected events, and other important clinical information.

Table 4.2. High-level Phases of the "Surgery Description" Section of Operative reports and Descriptions Used for Annotations of the "Surgery Description" section

Phase	Description
Preliminaries	• The first part of an operative report, describing any activity in the operating room prior to incision or preparation for the surgery. • These sections usually detail the type of anesthesia used, how the patient was identified, what position the patient is in, and if the patient was prepped in a sterile manner. • This section ends before any major incision has been made or once the patient is under anesthesia. • This section does not include preparation and exposure that happens once the patient is under anesthesia. For example, *"…the lid speculum was placed"* are not included here.
Getting Started	• This phase includes preparation work and the beginning of the case that happens once patient is under anesthesia, usually including the first major incision. • Sometimes, preliminary information is included in this phase if there is not a separation in sentences. For example, "Once patient was prepped in the usual sterile fashion and under the benefits of general anesthesia, ….. (description of beginning of procedure)" • Observations sometime occur in this phase

Table 4.2. Continued	
Phase	Description
Main Portion	• This phase goes from right after the exposure is obtained after first incision is made up until the closure of the wound. • Usually the largest portion of the annotation and includes all description of the procedure after the entrance and exposure and before closure of the incision. • This contains many of the procedure steps and details of technique. • Observations are common in this phase
Closure	• This phase describes the end of the procedure, from when the wound is closed. • This does not include information about the patient after the wound has been closed, such as condition, sponge counts, or movement of the patient • If a wound does not need to be closed to end surgery, this may include brief information about what was done to end the procedure, such as “a cotton ball was placed in the left ear”
Epilogue	• This phase includes information about the patient after the wound has been closed. This includes sponge counts, patient condition, where the patient is moved to, and other information regarding the procedure that may not directly deal with patient. • This can sometimes have intermixed observations

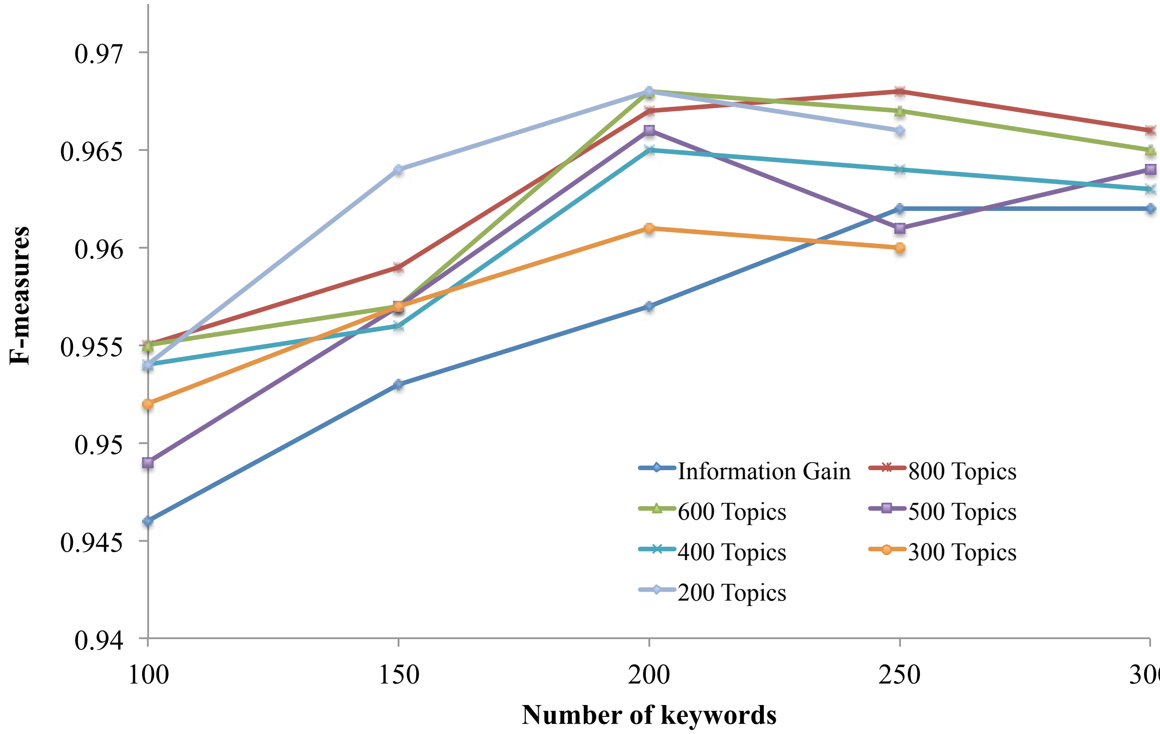
Table 4.2. (Continued)	
Phase	Description
Observations	• These statements include information about the patient that are important descriptions for documenting the procedure. • These statements typically do not include an action from the physician, although an action following the described observation may occur. • These statements can occur between other phases or within phases, the later often occurring with the Main Portion. • Examples: "Good hemostasis was noted throughout." "The tympanic membrane appeared retracted with an effusion."

The corpora of 100 laparoscopic cholecystectomy notes and 300 operative reports from a random selection of procedures contained 1412 and 3225 sentences, respectively. Inter-annotator assessment on ten laparoscopic cholecystectomy notes and 30 random operative reports showed 98.77% and 93.40% agreement and 0.93 and 0.69, as summarized in Table 4.3. IG (not pictured), performed best at approximately 500 words for the laparoscopic cholecystectomy notes and 1000 words for the random set of operative reports. We observed that the performance with LDA for correctly classifying high-level sections improved with increasing the numbers of topics, and had better than the performance than IG (Figure 4.1.a and 4.1.b). For the laparoscopic cholecystectomy notes, it had a peak performance at approximately 200 keywords (Figure 4.1.a). For the random set of operative reports, optimal performance was observed with approximately

500 keywords, although this varied with different numbers of topics (Figure 4.1.b). Table 4.4 summarizes performance of SVM with each set of features and dataset with laparoscopic cholecystectomy operative reports performing better than the random set of operative reports with LDA.

Table 4.3. Inter-rater reliability for annotations of high-level phases.

	Laparoscopic Cholecystectomy Operative Report Set		Random Operative Report Set	
Phase	**% Agreement**	**Kappa**	**% Agreement**	**Kappa**
Preliminaries	100.00	1.00	97.20	0.89
Getting Started	98.00	0.90	94.20	0.66
Main Portion	99.30	0.99	89.70	0.73
Closure	100.00	1.00	97.00	0.84
Epilogue	100.00	1.00	94.20	0.69
Observations	95.30	0.70	88.10	0.30
Average	98.77	0.93	93.40	0.69



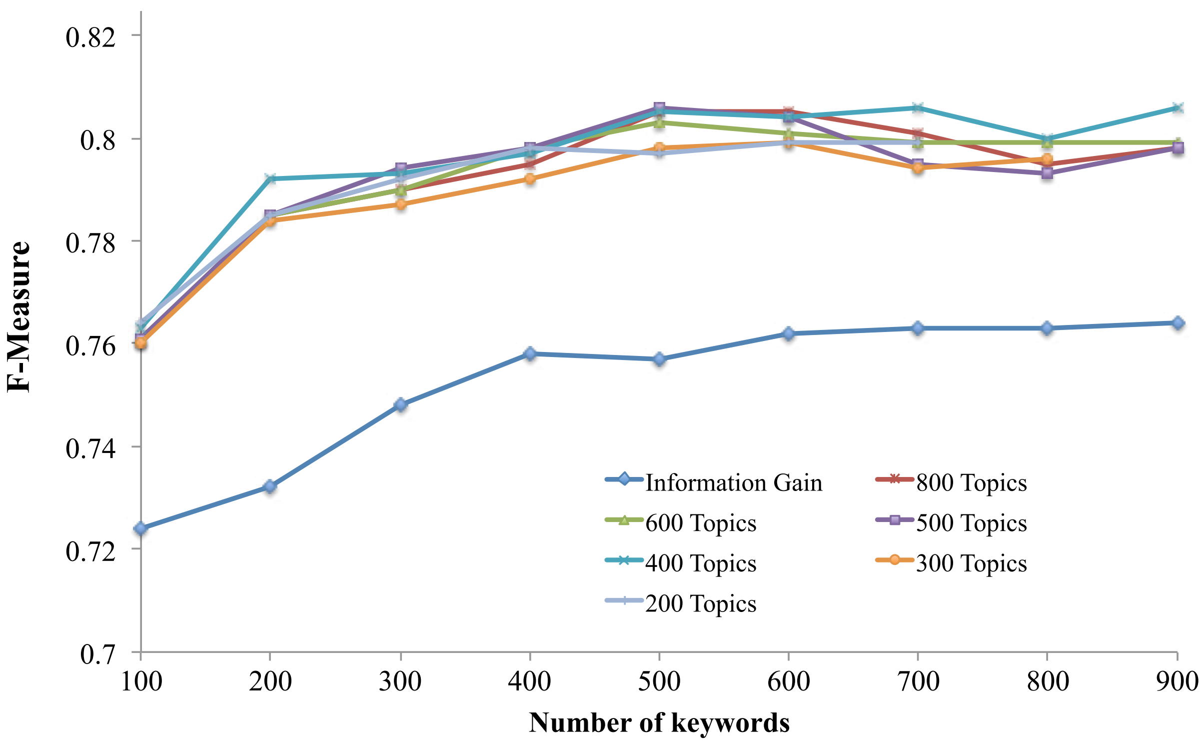


Figure 4.1.a and 4.1.b. Overall F-measures varied with the number of keywords and topics for the a) laparoscopic cholecystectomy and b) random set of operative reports.

Table 4.4. Classification Performance at a Sentence Level for High-Level Phases of the 'Surgery Description' Section using LDA and IG with SVM*.

		Laparoscopic Cholecystectomy Operative Report Set			Random Operative Report Set		
Features	**Phase**	**Precision**	**Recall**	**F-Measure**	**Precision**	**Recall**	**F-Measure**
IG**	Preliminaries	0.988	0.943	0.965	0.856	0.847	0.852
	Getting Started	0.972	0.956	0.964	0.645	0.624	0.634
	Main Portion	0.968	0.973	0.971	0.766	0.847	0.804
	Closure	0.983	0.980	0.982	0.799	0.680	0.735
	Epilogue	0.958	0.948	0.953	0.880	0.789	0.832
	Observations	0.816	0.866	0.840	0.627	0.510	0.562
	OVERALL	0.962	0.962	0.962	0.769	0.770	0.767
LDA#	Preliminaries	0.988	0.955	0.971	0.851	0.907	0.878
	Getting Started	0.978	0.967	0.972	0.712	0.636	0.672
	Main Portion	0.968	0.980	0.974	0.821	0.870	0.845
	Closure	0.989	0.980	0.985	0.822	0.776	0.799
	Epilogue	0.959	0.979	0.969	0.914	0.879	0.896
	Observations	0.865	0.856	0.860	0.617	0.502	0.554
	OVERALL	0.968	0.968	0.968	0.805	0.810	0.806

* SVM, support vector machine; LDA, latent Dirichlet allocation; IG, information gain;
** IG, information gain, 500 features for laparoscopic cholecystectomy operative reports, and 1000 features for random set of operative reports
# LDA, 200 features selected for laparoscopic cholecystectomy operative reports, and 500 features for random set of operative reports

Table 4.5.a. Subtopics for the phase *Preliminaries* in the "Surgery Description" Section

Subtopic Number	Keywords	Examples
1	Draped Sterile Prepped	• Abdomen was prepped and draped in the usual sterile fashion. • Abdomen widely prepped and draped in the usual sterile fashion. • The patient was brought to the operating room put under general anesthesia abdomen widely prepped and draped in the usual sterile fashion.
2	Induction Anesthesia Benefits	• After induction of general endotracheal anesthesia #NAME# abdomen was prepped and draped in the usual sterile fashion. • Under the benefits of general endotracheal anesthesia left upper quadrant Veress needle was inserted. • The patient was brought to the operating room put under general anesthesia.

Table 4.5.b. Subtopics for the phase *Getting Started* in the "Surgery Description" Section

Subtopic Number	Keywords	Examples
1	Trocar Xiphoid Vision	• A 5 mm trocar was placed. • Under direct vision a 5 mm trocar was placed in the right lateral abdomen below the xiphoid.
2	Veress Needle Supraumbilical	• Under general anesthesia the abdomen was insufflated through a vertical infraumbilical incision using a Veress needle. • We entered the abdomen through a small supraumbilical incision using a Veress technique without difficulty.
3	Skin Open Port	• Infraumbilical skin incision made open technique used Hasson port placed abdomen insufflated ports placed in routine position and technique. • An open technique used to place Hasson port. • The abdomen was insufflated ports were placed in routine position technique.
4	Liter Dissection Location	• Liters were placed with low pressures. • Dissection carried down.
5	Pneumoperitoneum Hypochondrium	• With the Hasson technique in an infraumbilical location the abdomen was entered and pneumoperitoneum was established. • Three additional mm trocars were placed along the right hypochondrium.

Table 4.5.c. Subtopics for the phase *Closure* in the "Surgery Description" Section

Subtopic Number	Keywords	Examples
1	Removed Ports Gas Suctioned	• Gas and trocars were removed. • Ports were removed. • Saline was suctioned out.
2	Sterile Pneumoperitoneum	• SteriStrips and sterile dressing applied • Pneumoperitoneum was released and the fascia of the infraumbilical port was closed with multiple interrupted Vicryl suture.
3	Incision Blocked Carterthomason	• The incisions were all blocked with Marcaine. • Skin incisions were closed with Vicryl and SteriStrips • The CarterThomason suture was used to pass Vicryl suture in figure-of-eight fashion to close this.
4	Trocar Visualization Bleeding Vision	• The trocars were then removed under direct visualization without evidence of bleeding. • Trocar sites were then infiltrated with % Marcaine with epinephrine and the trocars were removed under direct visualization.
5	Subcuticular Fascial	• Fascial defect closed with Vicryl skin with subcuticular. • The fascia was closed with interrupted Vicryl skin with subcuticular.

Table 4.5.d. Subtopics for the phase *Observation* in the "Surgery Description" Section

Subtopic Number	Keywords	Examples
1	Uterus Evidence Leakage	• Unharmed uterus noted. • No evidence of bile leakage noted. • There was no evidence of bile leak.
2	Bed Dry Clips Intact Liver	• The clips were intact and the bed was dry. • The initial trocar and Veress needle site was examined and it was dry. • Hemostasis was noted at the liver bed. • The clips were in good position.
3	Surveyed Pathology	• The abdomen was surveyed and no other pathology seen.

We then used LDA for different phases of laparoscopic cholecystectomy notes to identify subtopics. As exemplified in Tables 4.5.a-d, pertinent subtopics with *Preliminaries, Getting Started, Closure,* and *Observations* high-level phases were identified with LDA along with example sentences for each. These subtopics often corresponded to steps and associated tools and techniques used in performing the particular step.

4.6. Discussion

Operative reports serve as the primary form of documentation from surgical procedures and have significant value for improving surgical care and acute inpatient care in general. While there has been extensive research in the area of clinical NLP aimed at information extraction from clinical texts generally, there have been only limited studies examining clinical NLP for operative reports. This is largely due to the paucity of surgeon informatics stakeholders[4] and limited study of operative reports by clinical NLP groups. The current study combined expert opinion to describe potential high-level phases of the "Surgery Description" section and automated methods for identification of these phases using topic analysis techniques. We observed that both LDA and IG were effective at identifying these phases, and that there were subtopics in these phases with laparoscopic cholecystectomy notes, roughly corresponding to different sets of actions typically performed with laparoscopic cholecystectomy procedures.

The approach of multiple sessions with surgeons to gain consensus was effective in eliciting the main phases one for this small group of surgeons and had face validity in its findings. However, the surgeons were at a single institution, which may limit the generalizability of these findings. The use of paper and flexibility of when and where the initial questionnaire was administered gave the respondents good flexibility in their participation, as well as with the second iteration of collated responses. While the annotation guidelines and the use of topic analysis was the focus of these results, further validation of these findings with expert consensus might be obtained using a Delphi technique[66-68] or other validated consensus method.

There were several important observations gained in the process of annotation and assessing inter-annotator agreement. Agreement was highest with the more uniform laparoscopic cholecystectomy operative reports compared to the set of random operative reports. Additionally, annotators sometimes found it difficult to differentiate sentences with *"Observations"*, which were often found within longer, more complex sentences within the *"Main Portion"* phase. Also, some procedures did not have an incision or exposure associated with it (e.g., colonoscopy or bronchoscopy), and these procedures went almost directly to the *"Main Portion"* of the operative report.

The use of topic analysis and information gain for classification was effective as a means to identify high-level phases of "Surgery Description" sections with LDA performing better than IG in these experiments, although not very markedly. Most of the high-level phases performed well in our analysis, with the *Observations* classification performing less consistently well. Not surprisingly, the very similar laparoscopic cholecystectomy notes had higher performance than the more varied random set of operative reports. We speculate that the differences in performance could be attributed to the significant variation in the main portions of the procedure and that certain portions like *Preliminaries* and *Epilogue* being very similar even between procedures. We also observed that *Observations* often had anatomic terms and events or findings (e.g., ‘bleeding’, ‘adhesions’, ‘injury’) as subtopic words.

Sentence position as a feature was also important for classifying the high-level phase, likely since the content of some phases (e.g., *Preliminary* and *Closure*) is highly correlated to the sentence location within the section. When comparing the performance of feature sets excluding sentence positions to that using it as a feature, we observed about a 5% drop in F-measure compare to the LDA with position features for the random set of operative reports and a 1% decrement in performance for laparoscopic cholecystectomy operative reports.

A number of important limitations and potential next steps should be noted. First, the set of notes used in this study was relatively small. Additionally, further work will need to be done to validate these high level phases, which correspond to the intuitive phases of an operation that surgeons were able to identify. It is also unclear if more complex, inter-disciplinary, multi-team operative procedures will have similar characteristics. Future work is needed also to understand the value of this work in improving information extraction performance in operative reports, including its value in developing automated frame-based semantic approaches for operative report NLP systems such as the task of classifying predications expressed in operative report sentences. Ultimately, the development and maturation of these automated methods has the potential to benefit operative report quality and compliance, surgical billing, and research needs for surgery.

4.7. Conclusions

Consensus with a small group of surgeon experts and automated analysis of the "Surgery Description" section demonstrates high-level phases with meaningful parts. This preliminary work with a set of uniform laparoscopic cholecystectomy operative reports and more varied operative reports, show that these phases could be identified using supervised machine learning leveraging LDA and IG for feature extraction. These techniques may ultimately improve operative report NLP and information extraction for secondary applications in surgical clinical research and quality improvement.

4.8. Acknowledgements

This research was supported by National Institutes of Health (R01LM011364), Agency for Healthcare Research & Quality (R01HS022085), American Surgical Association Foundation Fellowship, and University of Minnesota Clinical and Translational Science Award (8UL1TR000114). The authors thank Fairview Health Services for support of this research.

Chapter 5: Conclusions

Despite the high impact of surgical care, there is a critical and unmet need to develop evidence-based informatics support mechanisms to guide and assist surgeons with clinical research and delivery of optimal surgical care[4]. Many clinical decisions in surgery continue to be based upon anecdotal clinical experience, interaction with mentors and peers, and case series (typically small single-institutional cohort studies) published in the literature. Traditional clinical studies in surgery are retrospective in nature with manual abstraction of patient records or small prospective studies[69].

While randomized controlled trials are gold-standard studies for evidence-based medicine, in surgery often they are financially impractical or unethical to conduct[70]. Health services researchers exploit large databases most often based upon insurance claims, Medicare data, and other administrative coding -- allowing researchers to answer questions on large patient populations[71]; however, they rely upon error-prone claims data and do not have sufficient granularity to answer important clinical questions, particularly for complex and specialty-specific populations[72, 73]. Disease or procedure-based "registry" initiatives including American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP)[74] and National Cancer Institute's Surveillance Epidemiology and End Results (SEER)[75] provide detailed, validated, and high quality clinical data for surgical and cancer outcomes, respectively. But these programs are costly, placing increasing burden on an already financially strained healthcare system; are limited in comprehensiveness (e.g. ACS-NSQIP's cyclical and sampling data collection);

and are not inherently flexible for adding clinical variables. Also, some surgical registries provide limited operative details, focusing instead on postoperative outcomes.

Improved use of operative reports for secondary uses like clinical research and surgical quality improvement represents an important opportunity for surgical care. As demonstrated in this research, while ORD with synoptic reporting might provide a "top-down" way to generate structured documentation and operation data, there are a number of barriers at present to its wider use. The findings of our survey show that some surgical specialties perform ORD education with formal teaching for surgical resident trainees. However, it appears that most surgical training programs treat this skill as an experiential one based in apprenticeship for residents instead of a more formal didactic competency. Moreover, while synoptic operative reports are used by surgeons in some disciplines, particularly for routine procedures, it appears from our survey of training institutions that their use remains limited in most surgical specialties.

Evaluation of structured sections headers from the HL7-ON DSTU and LOINC was performed with a large operative report corpus. Our study also examined coverage of these resources and potential gaps where current specifications might be expanded. A publicly available resource of operative report sections and mappings is an important result of this work that may used by others for NLP applications with operative and procedure notes, as well as potentially for data interoperability with documents. For this later application, while structured sections for documents are currently rarely used in practice for document exchange and reuse, standards organizations working in this area, such as HL7 promote the use of structured documents including high-level section headers. Furthermore, future roadmaps look to increase document inter-operability as an important part of health information exchange.

In assessing the “Surgery Description” section of operative reports in the final study, we identified five high-level phases, as well as need for *Observations* to describe findings, unexpected events, and other important clinical information. We found that text from the “Surgery Description” could be classified into these high-phases with reasonable performance using supervised machine-learning techniques with the application of topic analysis techniques for feature selection. Performance of our algorithms was predictably better for operative reports that were more uniform (i.e., the set of more uniform laparoscopic cholecystectomy operative reports) as compared to the performance on operative reports describing a diverse set of procedures.

We also observed that breaking a surgical procedure into smaller components revealed subtopics. We believe that surgical NLP and information extraction tasks will be easier but subgrouping the large ‘Surgery Description’ section into more manageable parts. It may also be feasible to leverage subtopics into identification of “sub-steps” within these high-level phases. For instance, we anticipate applying frame-based semantics with predicate argument structures to categorize surgical actions and their arguments. We also anticipate leveraging unsupervised and semi-supervised techniques to group operative report “steps” and label them in future studies and potentially to create standard ways to describe these steps that can be replicated – potentially so that standard terminologies, an “interface terminology” of sorts, for surgery actions and steps.

Overall, this work demonstrates the variability in ORD practices as reported at a national level and empirically observed at the individual surgeon level with respect to section structure and high-level phases to these important clinical documents. Our future work will seek to leverage our understanding of ORD structure to improve the application of information extraction and natural language processing (NLP) techniques for secondary use of operative reports. Effective resources and algorithms for operative reports will ultimately help surgeons to have powerful and scalable tools that could be leveraged for surgical clinical research and an improved surgical evidence base. Together, these tools will help front line surgeons provide better surgical care.

Bibliography

1. Kozak LJ, DeFrances CJ, Hall MJ. National hospital discharge survey: 2004 annual summary with detailed diagnosis and procedure data. Vital Health Stat 13. 2006 Oct(162):1-209.

2. Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: an evaluation of all participating hospitals. Ann Surg. 2009 Sep;250(3):363-76.

3. Steinberg SM, Popa MR, Michalek JA, Bethel MJ, Ellison EC. Comparison of risk adjustment methodologies in surgical quality improvement. Surgery. 2008 Oct;144(4):662-7; discussion -7.

4. Melton GB. Biomedical and health informatics for surgery. Advances in surgery. 2010;44:117-30.

5. Singh R, Chauhan R, Anwar S. Improving the quality of general surgical operation notes in accordance with the Royal College of Surgeons guidelines: a prospective completed audit loop study. J Eval Clin Pract. 2011 Jun;18(3):578-80.

6. Flynn MB, Allen DA. The operative note as billing documentation: a preliminary report. Am Surg. 2004 Jul;70(7):570-4; discussion 4-5.

7. Lefter LP, Walker SR, Dewhurst F, Turner RW. An audit of operative notes: facts and ways to improve. ANZ J Surg. 2008 Sep;78(9):800-2.

8. Park J, Pillarisetty VG, Brennan MF, Jarnagin WR, D'Angelica MI, DeMatteo RP, Coit DG, Janakos M, Allen PJ. Electronic Synoptic Operative Reporting: Assessing

the Reliability and Completeness of Synoptic Reports for Pancreatic Resection. Journal of the American College of Surgeons. 2010;211(3):308-15.

9. Gur I, Gur D, Recabaren JA. The Computerized Synoptic Operative Report A Novel Tool in Surgical Residency Education. Arch Surg. 2012;147(1):71-4.

10. DeOrio JK. Surgical templates for orthopedic operative reports. Orthopedics. 2002;25(6):639-42.

11. Donahoe L, Bennett S, Temple W, Hilchie-Pye A, Dabbs K, MacIntosh E, Porter G. Completeness of dictated operative reports in breast cancer—the case for synoptic reporting. Journal of Surgical Oncology. 2012.

12. Friedman C, Kra P, Rzhetsky A. Two biomedical sublanguages: a description based on the theories of Zellig Harris. J Biomed Inform. 2002 Aug;35(4):222-35.

13. Friedlin J, McDonald CJ. A natural language processing system to extract and code concepts relating to congestive heart failure from chest radiology reports. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2006:269-73.

14. Mamlin BW, Heinze DT, McDonald CJ. Automated extraction and normalization of findings from cancer-related free-text radiology reports. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2003:420-4.

15. Jiang M, Chen Y, Liu M, Rosenbloom ST, Mani S, Denny JC, Xu H. A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries. Journal of the American Medical Informatics Association : JAMIA. 2011 Sep-Oct;18(5):601-6.

16. Wang Y, Pakhomov S, Burkart NE, Ryan JO, Melton GB. A study of actions in operative notes. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2012;2012:1431-40.

17. Wang Y, Pakhomov S, Melton GB. Predicate argument structure frames for modeling information in operative notes. Studies in health technology and informatics. 2013;192:783-7.

18. Moore L, Churley-Strom R, Singal B, O'Leary S. Laparotomy operative note template constructed through a modified Delphi method. Am J Obstet Gynecol. 2009 May;200(5):e16-7.

19. Gerancher JC, Viscusi ER, Liguori GA, McCartney CJ, Williams BA, Ilfeld BM, Grant SA, Hebl JR, Hadzic A. Development of a standardized peripheral nerve block procedure note form. Reg Anesth Pain Med. 2005 Jan-Feb;30(1):67-71.

20. Laflamme MR, Dexter PR, Graham MF, Hui SL, McDonald CJ. Efficiency, comprehensiveness and cost-effectiveness when comparing dictation and electronic templates for operative reports. AMIA Annu Symp Proc. 2005:425-9.

21. Mori AR, Consorti F, Galeazzi E. A tagging system for section headings in a CEN standard on patient record. Proceedings / AMIA Annual Symposium AMIA Symposium. 1998:755-9.

22. Denny JC, Miller RA, Johnson KB, Spickard A, 3rd. Development and evaluation of a clinical note section header terminology. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2008:156-60.

23. Denny JC, Spickard A, 3rd, Johnson KB, Peterson NB, Peterson JF, Miller RA. Evaluation of a method to identify and categorize section headers in clinical documents. Journal of the American Medical Informatics Association : JAMIA. 2009 Nov-Dec;16(6):806-15.

24. Meystre S, Haug PJ. Automation of a problem list using natural language processing. BMC medical informatics and decision making. 2005;5:30.

25. Meystre S, Haug PJ. Natural language processing to extract medical problems from electronic clinical documents: performance evaluation. J Biomed Inform. 2006 Dec;39(6):589-99.

26. Ritchie MD, Denny JC, Crawford DC, Ramirez AH, Weiner JB, Pulley JM, Basford MA, Brown-Gentry K, Balser JR, Masys DR, Haines JL, Roden DM. Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. American journal of human genetics. 2010 Apr 9;86(4):560-72.

27. Dreger V. The Perioperative Nursing Data Set (PNDS): a new language for documenting nursing care. Insight. 2003 Jan-Mar;28(1):15-7.

28. Park HT, Lu DF, Konicek D, Delaney C. Nursing interventions classification in systematized nomenclature of medicine clinical terms: a cross-mapping validation. Computers, informatics, nursing : CIN. 2007 Jul-Aug;25(4):198-208; quiz 9-10.

29. Moorhead S, Delaney C. Mapping nursing intervention data into the Nursing Interventions Classification (NIC): process and rules. Nursing diagnosis : ND : the official journal of the North American Nursing Diagnosis Association. 1997 Oct-Dec;8(4):137-44.

30. Rector AL, Nowlan WA. The GALEN project. Computer methods and programs in biomedicine. 1994 Oct;45(1-2):75-8.

31. Trombert-Paviot B, Rodrigues JM, Rogers JE, Baud R, van der Haring E, Rassinoux AM, Abrial V, Clavel L, Idir H. GALEN: a third generation terminology tool to support a multipurpose national coding system for surgical procedures. International journal of medical informatics. 2000 Sep;58-59:71-85.

32. Rodrigues JM, Trombert Paviot B, Martin C, Vercherin P. Integrating the Modelling of EN 1828 and Galen CCAM Ontologies with Protege: towards a Knowledge Acquisition Tool for Surgical Procedures. Studies in health technology and informatics. 2005;116:767-72.

33. Wagner JC, Rogers JE, Baud RH, Scherrer JR. Natural language generation of surgical procedures. Int J Med Inform. 1999 Feb-Mar;53(2-3):175-92.

34. Rodrigues JM, Trombert-Paviot B, Baud R, Wagner J, Meusnier-Carriot F. Galen-In-Use: using artificial intelligence terminology tools to improve the linguistic coherence of a national coding system for surgical procedures. Studies in health technology and informatics. 1998;52 Pt 1:623-7.

35. Neumuth T, Jannin P, Strauss G, Meixensberger J, Burgert O. Validation of knowledge acquisition for surgical process models. Journal of the American Medical Informatics Association : JAMIA. 2009 Jan-Feb;16(1):72-80.

36. Meng F, D'Avolio LW, Chen AA, Taira RK, Kangarloo H. Generating models of surgical procedures using UMLS concepts and multiple sequence alignment. AMIA Annu Symp Proc. 2005:520-4.

37. Hoffer DN, Finelli A, Chow R, Liu J, Truong T, Lane K, Punnen S, Knox JJ, Legere L, Kurban G, Gallie B, Jewett MA. Structured electronic operative reporting: comparison with dictation in kidney cancer surgery. Int J Med Inform. 2012 Mar;81(3):182-91.

38. Novitsky YW, Sing RF, Kercher KW, Griffo ML, Matthews BD, Heniford BT. Prospective, blinded evaluation of accuracy of operative reports dictated by surgical residents. Am Surg. 2005 Aug;71(8):627-31; discussion 31-2.

39. Eichholz AC, Van Voorhis BJ, Sorosky JI, Smith BJ, Sood AK. Operative note dictation: should it be taught routinely in residency programs? Obstet Gynecol. 2004 Feb;103(2):342-6.

40. Donahoe L, Bennett S, Temple W, Hilchie-Pye A, Dabbs K, Macintosh E, Porter G. Completeness of dictated operative reports in breast cancer--the case for synoptic reporting. J Surg Oncol. 2012 Jul 1;106(1):79-83.

41. Srigley JR, McGowan T, Maclean A, Raby M, Ross J, Kramer S, Sawka C. Standardized synoptic cancer pathology reporting: a population-based approach. J Surg Oncol. 2009 Jun 15;99(8):517-24.

42. Baskovich BW, Allan RW. Web-based synoptic reporting for cancer checklists. J Pathol Inform. 2011;2:16.

43. Mack LA, Dabbs K, Temple WJ. Synoptic operative record for point of care outcomes: a leap forward in knowledge translation. Eur J Surg Oncol. 2010 Sep;36 Suppl 1:S44-9.

44. Gur I, Gur D, Recabaren JA. The computerized synoptic operative report: a novel tool in surgical residency education. Arch Surg. 2012 Jan;147(1):71-4.

45. Mainiero MB, Lourenco AP. The ACGME core competencies: changing the way we educate and evaluate residents. Med Health R I. 2011 Jun;94(6):164-6.

46. Melton GB, Burkart NE, Frey NG, Chipman JG, Rothenberger DA, Vickers SM. Operative report teaching and synoptic operative reports: a national survey of surgical program directors. Journal of the American College of Surgeons. 2014 Jan;218(1):113-8.

47. Department of Health and Human Services. Centers for Medicare & Medicaid Services. Medicare Learning Network. Evaluation and Management Services Guide. http://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/eval_mgmt_serv_guide-ICN006764.pdf. Accessed November 10, 2014.

48. Joint Commission Operative Note Requirements: Standard IM.6.30, Elements of Performance for IM.6.30. http://www.jointcommission.org/standards_information/jcfaqdetails.aspx?StandardsFAQId=215&StandardsFAQChapterId=43 and http://www.jointcommission.org/NR/rdonlyres/A032623D-02AF-4955-AF7C-08F3D5802E64/0/06_obs_im.pdf. Accessed November 10, 2014.

49. Centers for Medicare & Medicaid Services (CMS) Operative Note Requirements: Survey Protocol, Regulations and Interpretive Guidelines for Hospitals: A-0396 §482.51. http://www.cms.gov/Regulations-and-

Guidance/Guidance/Manuals/downloads/som107ap_a_hospitals.pdf. Accessed November 10, 2014.

50. Implementation Guide for HL7 Clinical Document Architecture (CDA) Release 2.0, Operative Note Draft Standard for Trial Use (DSTU) Release 1. March 2009. http://www.hl7.org/search/viewSearchResult.cfm?search_id=644266&search_result_url=%2Fdocumentcenter%2Fpublic%2Fstandards%2Fdstu%2F2008SEP%2FCDAR2%5FOPNOTE%5FR1%5FDSTU%5F2009MAR%2Ezip. Accessed November 10, 2014.

51. Chen ES, Carter EW, Winden TJ, Sarkar IN, Wang Y, Melton GB. Multi-source development of an integrated model for family health history. Journal of the American Medical Informatics Association : JAMIA. 2014 Oct 21.

52. Chen ES, Manaktala S, Sarkar IN, Melton GB. A multi-site content analysis of social history information in clinical notes. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium. 2011;2011:227-36.

53. Rajamani S, Chen ES, Akre ME, Wang Y, Melton GB. Assessing the adequacy of the HL7/LOINC Document Ontology Role axis. Journal of the American Medical Informatics Association : JAMIA. 2014 Oct 28.

54. LOINC 2.48 and RELMA 6.6. https://loinc.org/downloads/files/loinc-and-relma-complete-download/gotoCopyrightedFile. Accessed November 10, 2014.

55. Lin CY, Hovy E. The automated acquisition of topic signatures for text summarization. Proceedings of the 18th conference on Computational linguistics. 2000:495-501.

56. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. the Journal of machine Learning research. 2003;3:993-1022.

57. Arnold C, Speier W, editors. A topic model of clinical reports. Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval; 2012: ACM.

58. Halpern Y, Horng S, Nathanson LA, Shapiro NI, Sontag D, editors. A comparison of dimensionality reduction techniques for unstructured clinical text. ICML 2012 Workshop on Clinical Data Analysis; 2012.

59. Zhang R, Pakhomov S, Gladding S, Aylward M, Borman-Shoap E, Melton GB, editors. Automated Assessment of Medical Training Evaluation Text. AMIA Annual Symposium Proceedings; 2012: American Medical Informatics Association.

60. Cohen R, Elhadad M, Elhadad N. Redundancy in electronic health record corpora: analysis, impact on text mining performance and mitigation strategies. BMC bioinformatics. 2013;14(1):10.

61. Perotte AJ, Wood F, Elhadad N, Bartlett N, editors. Hierarchically supervised latent dirichlet allocation. Advances in Neural Information Processing Systems; 2011.

62. Blei DM, Ng AY, Jordan MI, Lafferty J. Latent Dirichlet allocation. Journal of Machine Learning Research 3 2003:993-1022.

63. Stopword list. http://www.textfixer.com/resources/common- english-words.txt.

64. LVG. http://www.nlm.nih.gov/research/umls/new_users/online_learning/LEX_004.htm.

65. Stanford Topic Model Toolkit. http://www-nlp.stanford.edu/software/tmt.

66. Brown BB. Delphi process: A methodology used for the elicitation of opinions of experts: DTIC Document1968.

67. Dalkey NC, Brown BB, Cochran S. The Delphi method: An experimental study of group opinion: Rand Corporation Santa Monica, CA; 1969.

68. Sackman H. Delphi assessment: Expert opinion, forecasting, and group process: DTIC Document1974.

69. Chang EY, Pannucci CJ, Wilkins EG. Quality of clinical studies in aesthetic surgery journals: a 10-year review. Aesthet Surg J. 2009 Mar-Apr;29(2):144-7; discussion 7-9.

70. Kao LS, Tyson JE, Blakely ML, Lally KP. Clinical research methodology I: introduction to randomized trials. J Am Coll Surg. 2008 Feb;206(2):361-9.

71. Riley GF. Administrative and claims records as sources of health care cost data. Med Care. 2009 Jul;47(7 Suppl 1):S51-5.

72. Preen DB, Holman CD, Lawrence DM, Baynham NJ, Semmens JB. Hospital chart review provided more accurate comorbidity information than data from a general practitioner survey or an administrative database. J Clin Epidemiol. 2004 Dec;57(12):1295-304.

73. Peabody JW, Luck J, Jain S, Bertenthal D, Glassman P. Assessing the accuracy of administrative data in health information systems. Med Care. 2004 Nov;42(11):1066-72.

74. Ingraham AM, Richards KE, Hall BL, Ko CY. Quality improvement in surgery: the American College of Surgeons National Surgical Quality Improvement Program approach. Adv Surg. 2010;44:251-67.

75. National Cancer Institute, Surveillance Epidemiology and End Results. 2011 [January 20, 2011]; Available from: http://seer.cancer.gov.