A Study of Actions in Operative Notes

Yan Wang, MS[1], Serguei Pakhomov, PhD[1, 2], Nora E. Burkart, MD[1, 3], James O. Ryan, BA[1], Genevieve B. Melton, MD, MA[1, 3]
[1]Institute for Health Informatics; [2]College of Pharmacy; and [3]Department of Surgery University of Minnesota, Minneapolis, MN

Abstract

Operative notes contain rich information about techniques, instruments, and materials used in procedures. To assist development of effective information extraction (IE) techniques for operative notes, we investigated the sublanguage used to describe actions within the operative report 'procedure description' section. Deep parsing results of 362,310 operative notes with an expanded Stanford parser using the SPECIALIST Lexicon resulted in 200 verbs (92% coverage) including 147 action verbs. Nominal action predicates for each action verb were gathered from WordNet, SPECIALIST Lexicon, New Oxford American Dictionary and Stedman's Medical Dictionary. Coverage gaps were seen in existing lexical, domain, and semantic resources (Unified Medical Language System (UMLS) Metathesaurus, SPECIALIST Lexicon, WordNet and FrameNet). Our findings demonstrate the need to construct surgical domain-specific semantic resources for IE from operative notes.

Introduction

Operative reports are created after every surgical procedure for the purposes of documentation and billing. While some surgeons create operative notes using electronic templates or structured tools[1-3], most operative notes are dictated and transcribed or self-entered, and represent the surgeon's recollection of the details of the procedure[4]. Within the operative report, the 'procedure description' section describes the details of what was observed and performed during the conduct of the procedure.

As an important branch of medicine, surgery is concerned with treatment of injuries or disorders of the body by incision or manipulation. Various factors such as technique used, incision length, supplies used (e.g., mesh type, prosthetic), or pneumoperitoneum pressure[5-8] can affect surgical patient outcomes. Surgeons, who perform surgeries with specialized training in operative procedures, need to determine the best way to perform procedures based on accessible sources of the best evidence available. Extracting information on the techniques, instruments, materials, and other factors surrounding operative procedures from operative reports provides an important opportunity to efficiently access numerous procedure reports and provide the necessary information in a succinct and easily comprehensible fashion for secondary uses like summarization or extraction for high-throughput clinical research.

With widespread electronic health record (EHR) system adoption throughout healthcare, operative reports are increasingly accessible in electronic format and are potential information sources which may be valuable for a wide variety of secondary functions including new medical knowledge development, decision support, and clinical research[9]. While a large amount of work has been directed towards information extraction (IE) from radiology reports, discharge summaries, and other clinical texts[10-20], less work has focused upon operative notes. Effective computerized natural language processing (NLP) systems for operative reports require an understanding of this text and addressing the sublanguage domain-specific features of operative notes. As opposed to other medical sublanguages[21], the 'procedure description' section contains a significant amount of description of actions performed during an operation, as illustrated in the following statements:

(1) "The anterior chamber was filled with viscoelastic expanding the space between the posterior iris and the anterior capsular rim."
(2) "The Murphy hook passed easily along the L5 root and the S1 root and the wound was irrigated with antibiotic solution and hemostasis was verified."

Here, action verbs (e.g., fill, pass, irrigate) encode predicative relations between nominal arguments (e.g., chamber, viscoelastic, Murphy hook, L5 root, antibiotic solution). These predicate arguments convey the important details about actions performed during a procedure. In this study, our goals were to (1) study and characterize actions contained in the 'procedure description' text, (2) utilize knowledge sources in the linguistic and biomedical NLP fields for mapping actions, and (3) provide a resource of procedure-specific actions for other NLP researchers.

Background

Operative reports and 'procedure description' section

Following standards set by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO)[22] and Accreditation Association for Ambulatory Health Care (AAAHC)[23], operative notes are organized into sections that describe the pre-procedure diagnoses, post-procedure diagnoses, name of the procedure, indications for the procedure, findings during the procedure, pathological specimens, type of anesthesia, complications, blood loss, and the detailed description of the procedure. The narrative description of a procedure is the core portion of an operative note and is most often entitled 'description of procedure', 'procedure narrative' or 'procedure description'. In addition to providing specific documentation of what occurred during an operative procedure, the 'procedure description' is rich in information about the materials, equipment or instruments, surgical maneuvers, procedure variations, and the patient-specific anatomy encountered by the surgeon in an operation. As detailed in (3), the following text gives an illustrative example of an excerpt from a 'procedure description' section from an appendectomy in a patient with perforated appendicitis:

(3) "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."

Action description syntactic forms

In linguistics, besides verbs as productive predicates, nominalization, gerunds, and relative nouns are also used to express predicative relations and can take the same arguments as the corresponding verbs[24]. Several syntactic structures have been described for action descriptions, as depicted in Table 1. Part of the motivation of this study was to understand the language (i.e., syntactic structures) of action sentences utilized for depicting actions, including the use of 'activity verbs' to show when a nominal, indefinite verb, or gerund introduces an action.

Table 1. Action description forms.

Form	Examples	Activity verbs
Action verb	1. The medial edge of the cleft was ***incised*** sharply with a knife. 2. It was ***incised*** just above the level of the bladder flap.	
Activity verb + gerund	1. The scope was removed and the ***curetting*** was ***performed.*** 2. We then ***did*** a ***lengthening*** of the lateral aspect of the quad approximately 5 mm.	perform, carry out, apply, carry, do, fashion, begin, undertake, continue, etc.
Activity verb + verbal nominals	1. The ***incision*** was ***carried*** through the subcutaneous tissue. 2. We ***carried*** the ***dissection*** down through dartos muscle. 3. A sagittal split ***incision*** and subperiosteal ***dissection*** was ***accomplished***.	perform, carry out, apply, carry, do, fashion, undertake, begin, continue, achieve, gain, get, obtain, provide, etc.
Activity verb + Indefinite verb	1. We ***began*** to ***lift*** the gland up and away from essential anatomy. 2. A rongeur was ***used*** to ***remove*** the hyaline cartilage.	begin, continue, use, etc.
Activity verb + deverbal nominals	1. A box ***cut*** was ***made*** to substitute for the PCL. 2. I ***made*** an ***incision*** paralleling the acromioclavicular joint.	make, create, develop, etc.
Activity verb + deverbal nominals	1. General ***anesthesia*** was ***administered***. 2***. Dressings*** were ***applied,*** drapes removed.	administer, apply, etc.

The same action can be narrated with different verb combinations, nominals and voice, as exemplified below. Understanding this better has practical importance in the construction of NLP systems to process those notes.

(4) "Subsequent ***curettement*** of the bone edge was ***performed***."
(5) "Gentle ***curetting*** was ***done***."
(6) "The canal of the humerus was carefully ***curetted***."
(7) "We ***curetted*** the sockets."
(8) "Gentle, sharp ***curettage*** was ***performed***."

Traditionally, most work with medical IE has focused on surface level patterns[25, 26], which can be learned from annotated text or hand built. In the domain of operative reports, such approaches may not be able to achieve high completeness and accuracy for IE due to the linguistic complexity of surgical concepts expression. As shown in examples in Table 1 and (4 - 8), verbs are subject to syntactic variation and nominalization, which can be used to describe the same event. As a consequence, a wide range of syntactic patterns could potentially express the same operation action.

Semantic frames and sublanguages

Due to the similar linguistic complexities to surgical reports, syntactic roles and semantic roles have been used as a potential approach to extract information from scientific text. In the past several decades, several computational linguistics resources such as FrameNet[27], PropBank[28] and VerbNet[29] have been developed to provide semantic frames for predicates, to describe facts or events, along with annotated example sentences. For each frame, a set of frame elements are defined and example surface syntactic realizations of semantic roles are provided. These semantic resources offer an important first step towards deeper text understanding. Tasks requiring semantic processing such as semantic role labeling, question answering, and text categorization have benefited from these resources[30, 31].

Despite ongoing progress of clinical IE methodologies, there has been realization that resources and NLP tools, which may perform well on text from one source, may fail to perform well on text from a new domain. A number of researchers have explored the linguistic differences between different sublanguages associated with clinical and biomedical domains[21, 32-34]. In 2002, Friedman et al.[21] surveyed features of sublanguages, documented two biomedical-domain sub-languages (clinical reports and molecular biology) and discussed the similarities and differences between them. Lippincott et al.[32] showed that rich variation exists across a variety of linguistic dimensions (lexical, syntactic, sentential and discourse features) between subdomains of biomedicine. The authors also stated that an awareness of such variation is necessary when deploying NLP systems for use in single or multiple subdomains. Kilicoglu et al.[33] explored the task of interpretation of nominalizations and developed a set of linguistic generalizations for effective interpretation of a wide range of patterns used to express arguments of nominalization in clinically-oriented biomedical text.

Despite research that has looked at the general topic of sublanguages, limited work has been done examining the sublanguage of surgical procedures. In this study, we aimed to investigate the surface patterns of action descriptions, the action predicates, and distribution of different predicates usage. We also aimed to evaluate the adequacy of existing domain-specific and general English resources to extract action information from procedure descriptions. Finally, we sought to develop a knowledge resource of top action predicates along with the mapping information.

Methods

A total of 362,310 operation narratives obtained from University of Minnesota-affiliated Fairview Health Services, with data from 4 metropolitan hospitals in the Twin Cities including both community and tertiary-referral settings were used for this study. The corpus includes operative reports created by 2,300 surgeons with 4,333 different procedure types defined by Current Procedural Terminology (CPT) codes.

Automated section extraction with locally defined header hierarchy

From the data repository, the 'procedure description' section was first extracted from each note. Most operative notes are organized into sections and sub-sections such as 'Procedure description', 'Pre-operative diagnosis', and 'Anesthesia' and are specific to the type of the note (e.g., an Admission note has sections corresponding to a standard history and physical examination). Typically, each section has a section header string that includes words that provide context for the encapsulated text. For example, a section with a header string 'Procedure description' provides detailed and step-by-step description of a surgery. The text within these sections provides important information about surgeries.

While clinical notes are typically organized into standard sections, clinicians often label sections of documents with non-standardized terms based on use of acronyms, abbreviations or synonyms. Review of operative notes showed that the 'procedure description' section could be labeled as 'Procedure details', 'technique procedure', 'OP report', 'incisions', 'case details', etc. We also found that occasionally a section like 'procedure description' may not be associated with a section header. In this case, the existence of a specific section can be inferred by the semantic content of the text. In this study, we examined the section headers and sections in a subset of operative notes and developed a tool to extract 'procedure description' sections using a large set of operative notes.

Potential sections headers were extracted from the data repository using a random set of 3,000 operative notes. One of two surgeons (GM and NB) reviewed the 300 most used section headers along with headers from relevant note templates and grouped them into a hierarchy of headers. We developed a tool based on this hierarchy to extract the 'procedure description' section by combining features such as header string matching, header format pattern, section length, and section-specific terms. An evaluation of 200 operative notes with 1,594 sections demonstrated an accuracy of 95% for correct extraction of 'procedure description' section.

Sentence categorization

All sentences within 10 random selected 'procedure description' sections were reviewed, revealing that sentences could be classified into three categories (Perception/Report, Action, and Other) based on semantic content of the event described (Table 2). In our dataset, most sentences fall into the action category.

Table 2. Sentence Categories.

Category	Examples
Perception/Report	1. I could feel no full thickness tear. Visualized no full thickness tear. 2. Sponge and needle counts were reported as correct. 3. There appeared to be a simple cyst within.
Action	1. We placed a double stranded Mersilene tape around the coracoid. 2. A box cut was made to substitute for the PCL. 3. We continued mobilization up to the hepatic flexure.
Other	1. This array of components allowed for full extension with minimal recurvatum and easy flexion. 2. She wanted to proceed with the right knee.

Adapted Stanford parser

Syntactic parsing, especially full syntactic parsing, is a very important step toward natural language understanding[35], and has been applied in a range of tasks such as semantic role labeling and question answering. Full syntactic parsing of texts provides deep linguistic features like predicate and POS tag of predicate, voice, phrase type, position, and path, which have been shown to perform considerably better than surface-oriented features for IE. The Stanford constituency Probabilistic Context Free Grammar (PCFG) parser[36] used in this study is one of best performing parsers trained on Penn Treebank[37]. As an unlexicalized parser, it uses probabilities associated with syntactic categories instead of word–word dependencies, which means that the Stanford parser is less sensitive to specific domain text used for training. We then extended the Stanford parser lexicon to adapt the parser to our local surgical narrative.

Initial experiments on full parsing of procedure descriptions with a Newswire-derived lexicon showed a low parsing performance. The errors in parsing output were examined. We observed that many verbs used in medical text do not exist in the Stanford lexicon. For example, in the following sentences, the words '*re-draped*', '*prepped*' and '*exsanguinated*' do not exist in Stanford lexicon. Since the parser cannot tell from the local context that these are verbs, the sentences are wrongly parsed as shown.

(9) "He was ***re-draped*** and ***prepped***."
(ROOT (S (NP (PRP He)) (VP (VBD was) (ADJP (JJ re-draped) (CC and) (JJ prepped)))))

(10) "The limb was ***exsanguinated*** and tourniquet inflated to 350 mmHg."
(ROOT (S (NP (DT The) (NN limb)) (VP (VBD was) (ADJP (JJ exsanguinated) (CC and) (JJ tourniquet)) (S (VP (VBN inflated) (PP (TO to) (NP (CD 350) (NN mmHg))))))))

To improve parsing performance, the Stanford parser was expanded with the SPECIALIST Lexicon[38]. SPECIALIST has been successfully used to augment the Stanford parser lexicon for better noun phrase identification in a previous study by Huang and authors[11]. Since the Stanford statistical parser requires a relative frequency of each tag, which is absent from the SPECIALIST Lexicon, the approach used by Huang and in our study utilized unambiguous SPECIALIST Lexicon entries and converted these into Penn Treebank tags based on predefined mappings.

In addition to unambiguous verbs, we also included verbs with more than one POS tag into the Stanford lexicon. Our hypothesis here is that the Stanford parser can resolve the correct tag for a verb in a sentence based on the local context if it has information that the word can potentially be a verb. Since we did not have *a priori* distribution statistics for each verb, we assigned phrase tags and frequencies to each verb based on the statistics collected from

the original Stanford parser lexicon. Our experiments show that, based on the local context, the Stanford parser will often assign an appropriate tag to a verb. For example, the word '*extubated*' is in the original Stanford lexicon. With the original lexicon, the parser has an error (parsing A). After adding entries of tag VBN and VBD of the word '*extubated*' to the Stanford parser lexicon, the parser gives a correct parsing (parsing B), as shown in (11):

(11) "The patient was awakened, ***extubated*** and alert."

A): (ROOT (S (NP (DT The) (NN patient)) (VP (VBD was) (ADJP (JJ awakened) (, ,) (JJ extubated) (CC and) (JJ alert)))))

B): (ROOT (S (NP (DT The) (NN patient)) (VP (VBD was) (VP (VBN awakened) (, ,) (VBN extubated) (CC and) (JJ alert)))))

The frequency of VBD and VBN entries were also adjusted for each verb whether the verb existed within the original Stanford parser lexicon or was a new entry. We observed that the actions are most often narrated in a passive voice; so, the frequency of VBN tag assignment for each verb was adjusted to be more than the VBD tag assignment for the same verb. In addition to verbs, adjective and adverb entries from the SPECIALIST Lexicon were added. Phrase tags and frequencies were collected the same way as above for verb POS tags. Some commonly used words were found to have different frequencies compared to general texts like the Wall Street Journal. In these cases, tag frequencies were also adjusted. For example, the word '*alert'* can be a verb, an adjective or a noun. In general English, it is mostly used as a verb. However, in 'procedure description' text, it is mostly used as an adjective (e.g., "*The patient was transferred to the PICU when awake and alert*").

Categorization of actions, expansion of nominals, and distributions in operative notes

Parsing results of the adapted Stanford parser for 'procedure description' text were used to collect the most frequently used verbs. For each parsed sentence, the top-level main verbs of each sentence were collected based on the syntactic tree. A random set of 50 notes (964 sentences) was used to evaluate the accuracy of the approach. A trained linguist and an informatisist annotated the main verbs of each sentence for the entire evaluation set (JR, YW). Kappa statistic indicates reasonable inter-rater agreement (0.78) and proportion agreement (0.94). The approach demonstrated a recall of 90.2% for detecting main verbs from all 13,095 tokens in the evaluation set.

For the entire set of 'procedure description' sections, verbs, including phrasal verbs, and their frequency were collected. We focused on verbs providing coverage for over 92% of all the top-level verbs in the corpus (Section 'Results'). These verbs were categorized into action verbs, activity verbs, and verbs with a perception/report or other non-action usage (Table 2). Since gerunds and other nominals derived from an action verb are also used to describe actions, potential nominals of each verb were collected through automatic and manual approaches from existing resources including the SPECIALIST lexicon, the WordNet lexicon,[39] New Oxford American Dictionary[40], and Stedman's Medical Dictionary[41]. From this, the incidence of verbs and their nominals used to describe actions were collected from the overall corpus.

Semantic and domain knowledge resource coverage evaluation

Since semantic resources derived from general English, lexical resources, and domain-knowledge play an important role in IE, the adequacy of existing resources to facilitate the usage of verb predicates and their nominals was evaluated with the UMLS, SPECIALIST Lexicon, WordNet, and FrameNet.

In the biomedical and clinical domains, the UMLS Metathesaurus is a large, multi-purpose database built from over 100 different terminology sources in patient care, health services billing, public health statistics, and biomedicine. It is designed to support a broad range of biomedical research and includes rich information. For example, the UMLS concept '[C0677554] Anastomosis – action' has a semantic type 'Therapeutic or Preventive Procedure' and the entry provides detailed definition of the action from several sources like 'CHV/PT | surgical connection between two hollow organs'. The SPECIALIST lexicon includes the syntactic, morphological, and orthographic information for each lexicon term and is, as previously described, a resource for improving the performance of NLP tasks.

WordNet and FrameNet are two notable general English semantic resources repeatedly used in biomedical and clinical research. WordNet is a repository of hierarchically organized English words that are organized into sets of synonymous terms (verbs, nouns, adjectives, and adverbs), called synsets, each of which represents one lexical concept. The database contains about 150,000 lexical items organized in over 115,000 synsets. The Berkley FrameNet project is an online resource for general English semantics. As introduced in the background section, it is an essential lexical semantic resource providing predicate frames that can aid in natural language understanding.

Results

Application of the Stanford parser on the 'procedure description' section demonstrated that the 200 most frequent top-level verbs in the entire corpus covered 92% of all top verbs in the 'procedure description' section. To test the coverage of the verbs selected in several related surgical domains, we created separate datasets for Prostatectomy, Colectomy, and Total Abdominal Hysterectomy, each with 1,000 randomly selected operative notes with corresponding CPT codes. The 200 verbs demonstrated 89%, 90%, and 92% coverage of top verbs in the three datasets respectively. Each verb was individually examined and 147 verbs were classified as action verbs, while 15 were activity verbs. Table 3 shows a partial list of these action verbs:

Table 3. Action verb examples.

place	drape	bring	dissect
take	close	identify	open
remove	give	divide	tolerate
prep	tie	cauterize	transect

Using the SPECIALIST lexicon, WordNet, New Oxford American Dictionary, and Stedman's Medical Dictionary, a total of 97 unique nominals (median 0, range (0-2)) were extracted, several of which are listed in Table 4. Table 5 shows the distribution of verbs, gerunds, and nominals used to describe actions. As shown in Table 5, physicians tend to use verbs to describe actions and prefer using a passive voice, as in: *"A computer plan was developed for placement of 75 palladium-103 seeds"*. The predominant use of the passive voice was also true for nominal action predicates.

Table 4. Action verbs and their nominals.

Verb	Nominals
anaesthetize	anaesthetization, anesthesia
anastomose	anastomosis
cannulate	cannulation
curette	curettage, curettement
drain	drainage
debride	debridement
expose	exposure

Table 5. Predicate distributions.

Predicate form	Total	Passive voice	Active voice
Verb	3,808,845 (94,4%)	3,306,300 (86.8%)	502,545 (13.2%)
Gerund	13,425 (0.3%)	12,820 (95.3%)	605 (4.7%)
Nominal	211,102 (5.2%)	184,509 (87.4%)	26,593 (12.6%)

Table 6 shows the distribution of several top-, middle-, and low-incidence actions and nominals of each action verb. Most actions are expressed using verb predicates with the exception of '*incision*' and '*dissection*', which were frequently described with the pattern of 'verb + nominal'.

As summarized in Figure 1, which shows coverage of the top 147 actions, the SPECIALIST Lexicon had very good coverage of both verb predicates (89.9%) and nominal predicates (100%), although it missed some phrasal verbs (e.g., '*bring back*', '*dissect out*', '*carry down*'). WordNet also had good coverage for predicates, specifically 89.9% for verbs and 93.8% for nominals. Since it is a resource addressing general English, WordNet missed some domain-specific terms like '*curette*', '*exsanguinate*', '*extubate*', and '*free up*', etc. The UMLS Metathesaurus, which contains important domain knowledge, covered only 11.5% of action verb predicates and 58.8% of nominal predicates. As a semantic resource, the FrameNet also had poor coverage of nominals (36.1%) and fair coverage for verbs (64.2%). The SPECIALIST Combined with WordNet had a coverage of nominal (100%) and verbs (93.8%).

Table 6. Usage of verbs, gerunds, and nominals to describe surgical actions.

Action	Total action mentions	Categorized action mentions			Nominals
		Verb	Gerund	Nominals	
place	431,576	430,871 (99.84%)	0	703 (0.16%)	placement
close	260,450	253,243 (97.23%)	14 (0.01%)	7,193 (2.76%)	closure
drape	176,439	174,883 (99.12%)	13 (0.01%)	1,543 (0.87%)	drape
take	167,126	167,125 (100.00%)	0	0	-
prep	165,522	164,459 (99.36%)	21 (0.01%)	1,427 (0.86%)	prep
incise	163,007	37,032 (22.72%)	0	125,973 (77.28%)	incision
remove	156,487	156,078 (99.74%)	0	408 (0.26%)	removal
bring	129,445	129,444 (100.00%)	0	0	-
irrigate	92,689	90,171 (97.28%)	0	2517 (2.72%)	irrigation
dissect	82.450	52,185 (63.2%)	5	32,260 (36.7%)	dissection

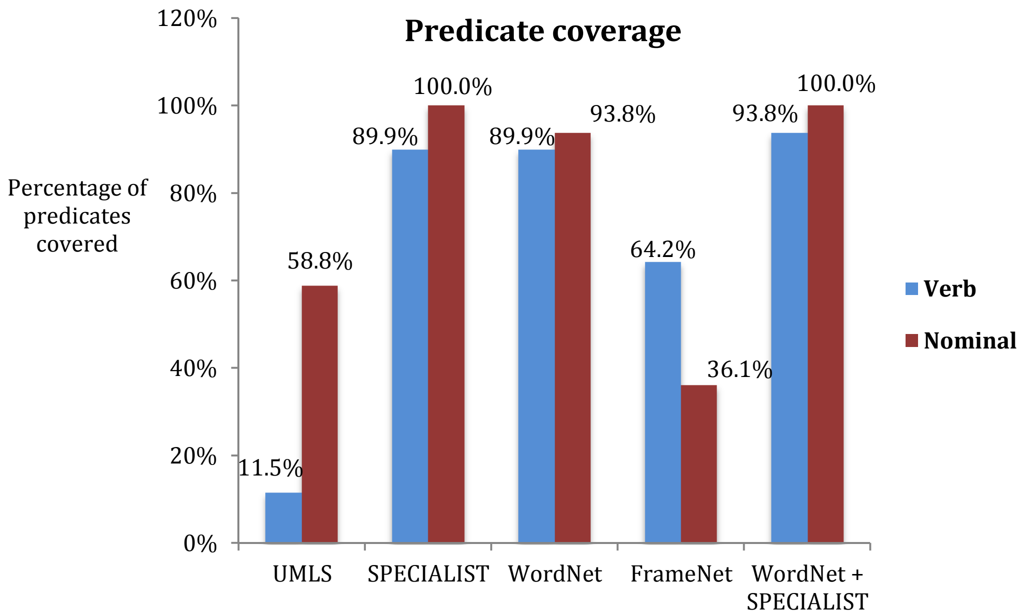

Figure 1. Coverage of operative note verbs and nominals by semantic resources.

Discussion

Information contained in operative notes is critical to better understanding and improving surgical clinical practice and potentially has many secondary uses for surgical research and quality improvement. Traditionally, many of the decisions made by surgeons about how to optimally perform a particular surgical procedure are made on the basis of the clinician's clinical experience, opinions from colleagues, or available case series reports. These information sources are often limited to small groups, and, unfortunately, randomized controlled trials in surgery are rare and difficult to conduct due to ethical and financial barriers. Likewise, manual review of large numbers of operative reports is not a scalable solution. With the accumulation of large volumes of machine-readable operative notes, there is an opportunity for developing tailored clinical NLP methods to extract and provide necessary information from these narratives. The features specific to the surgical domain sublanguage in operative notes have important implications for the development of effective computerized NLP systems for operative note IE. In this work, we studied and characterized the 'procedure description' sublanguage for surgical actions to lay the groundwork for accurate and effective IE from operative notes. Specifically, we studied prevalent predicates, including 147 action verbs and 97 action nominals. We believe that the set is broad enough (92% coverage of verbs in our operative note

repository) to support IE from operative notes and deep enough to deal with the syntactic variability that exists in the sublanguage.

Overall, actions within operative notes in the 'procedure description' section were mostly verb-based predicates along with associated semantic arguments. Nominal predicates were uncommonly used with some notable exceptions, like 'incision' and 'dissection'. Also, a great majority of action-verb predicates (86.8%) were found to be in the passive voice. Similarly, verb nominals, which occurred less often (5.2%), were also predominantly (87.4%) in the passive voice. Only a very small portion of actions was described by gerund predicates. Our coverage evaluation demonstrated that the SPECIALIST lexicon had entries for all the nominal predicates and 89.9% of the verbs, with the exception of some phrasal verbs (e.g., '*free up*'). As a general English resource, WordNet misses some medical terms (e.g., '*extubate*'). Despite this, WordNet still covered 93.8% of the nominal predicates and 89.9% of the verbs.

Interestingly, the domain resource (UMLS) and the semantic resource (FrameNet) showed low coverage for both verb predicates and nominal predicates. Consisting of different types of biomedical vocabularies, the UMLS encompasses terms and codes in a wide range of categories including diagnosis, procedures, disease, anatomy, drugs, genetics, nursing and others. In the 2011AB version, the UMLS includes 215,327 Therapeutic or Preventive Procedure concepts, 31,826 Diagnosis Procedure concepts, 8,175 other Health Care Activity concepts and 451 Daily or Recreational Activity concepts. It is somewhat unexpected that such a large vocabulary covered only 11.5% percent of the action verbs. The evaluation of the mapping results shows that all the phrasal verbs like '*take down*' or '*free up*' were not covered by the UMLS. Also, a large number of prevalent and domain specific verbs such as '*incise*', '*expose*' and '*close*' were also not defined. Nominal predicates, on the other hand, had fair coverage (58.8%) by the UMLS. Since the UMLS provides linkage to biomedical terminologies and FrameNet had potential for semantic processing with frames, improvement and expansion of both resources for the surgical domain is a neccessary step in future system development.

Besides actions within the main clauses of sentences (e.g., '*The 20-French rigid cystoscope with blade was removed and an attempt was made to place a 24-French rigid resectoscope*'), phrases also contain actions as in the following:

(12) "The patient was taken to the operating room ***where general anesthetic was administered***",
(13) "***After the successful induction of spinal anesthesia***, she was placed supine on the operating table",
(14) "***Prior to removing the trocar***, cystoscopy was again performed".
(15) "An attempt was made ***to place the 24-French rigid resectoscope"***.

Although these phrases were not systematically analyzed in this study, we did observe that actions expressed in phrases tended to have fewer semantic arguments compared with the actions described in a main clause. Additionally, in some cases phrases can be used to describe an event that may or may not be an actual action performed in a procedure. For example in (15), it is difficult to determine if the action was performed or not. Due to the large syntactic variability of the sentence structures of these phrases, in this work we focused on the verbs in the main clauses. However, we realize that for many NLP tasks or applications, such as procedure summarization, it will be critical to effectively extract these actions as well.

One important discussion point surrounds our use of the Stanford parser and its augmentation with the SPECIALIST lexicon. As we presented in the 'Methods' Section, the analysis of actions in this work was based on the deep parsing output of 'procedure description' text with the Stanford parser expanded with the addition of the SPECIALIST lexicon. Since the Stanford parser was trained on a general English corpus, the parser's grammar statistics are collected from a much different text than the operative report text that we are interested in. Consequently, the adapted parser may not be capable of resolving many of the complex or unusual sentences found in the 'procedure description' section. It is also possible that better parsing accuracy can be achieved by retraining the parser on an annotated corpus from the medical domain. However, we found that the parsing output from the current adapted Stanford parser showed good parsing accuracy on 'procedure description' text in this study.

Examination of the most prevalent action predicates and their usage in operative reports also gave certain insights into knowledge sources for frame semantics. Analysis of operative note predicates revealed that existing common used resources are not fully adequate for effective IE from operative notes. This is an important consideration for future work that could build upon semantic frames in operative note summarization. Our results also indicate that further work may be needed for creating new frames and adapting existing frames, as the frame resource (FrameNet) had significant coverage gaps to action predicates. Moreover, in operative notes some predicates are used for a different meaning than in general English. For example, the phrasal verb '*come across*' means '*meet*' in general English, but in the following example, the phrasal verb means '*go through*'.

(16) “We ***came across*** the liver parenchyma using the Helix device.”

This example demonstrates the need for large annotated corpora for both semantic frame generation and also for the related and subsequent semantic role labeling process. Besides expansion and adaption of current frames and operative note annotations, we anticipate needing to build robust algorithms to define how to transform the relevant constituents of a surface sentence to the semantic arguments in frames. To facilitate sophisticated text mining applications, a lexicon that describes real, observed usage of predicates and other domain terms in operative notes and a domain knowledge resource that provide domain information are also required.

The overall work of this study gives insight into the language used by surgeons to communicate action events in the operating room. This study provides an understanding of the relative variability of action expressions. The action verbs, their nominals, and mappings are available to other researchers on request. Our next step is to work towards development of new frames and extension of existing frames such as those in the extended Propbank in the Clinical Text Analysis and Knowledge Extraction system (cTAKES) with a pilot study to assess the feasibility of this as a methodology for operative note IE related to surgical techniques.

Acknowledgements

The author would like to thank Fairview Health Services and support from American Surgical Association (GM), National Library of Medicine #R01 LM009623-01 (SP), and Institute for Health Informatics Seed Grant (GM/SP).

References

1. Park J, Pillarisetty VG, Brennan MF, Jarnagin WR, D'Angelica MI, DeMatteo RP, et al. Electronic Synoptic Operative Reporting: Assessing the Reliability and Completeness of Synoptic Reports for Pancreatic Resection. J Am Coll Surg. 2010;211(3):308-15.
2. 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.
3. DeOrio JK. Surgical templates for orthopedic operative reports. Orthopedics. 2002;25(6):639-42.
4. Donahoe L, Bennett S, Temple W, Hilchie-Pye A, Dabbs K, MacIntosh E, et al. Completeness of dictated operative reports in breast cancer-the case for synoptic reporting. J Surg Oncol. 2012.
5. Slim K BJ, Kwiatkowski F, Lescure G, Pezet D, Chipponi J. Effect of CO2 gas warming on pain after laparoscopic surgery: a randomized double-blind controlled trial. Surg Endosc. 1999:13:1110–4.
6. Vezakis A DD, Gibson J S, Moore M R, Shah H, Larvin M, McMahon M J. Randomized comparison between low-pressure laparoscopic cholecystectomy and gasless laparoscopic cholecystectomy. Surg Endosc. 1999:13:890–3.
7. Kum C K EE, Aljazari A, Troidl H. Randomized comparison of pulmonary function after the “French” and “American” techniques of laparoscopic cholecystectomy. Br J Surg. 1996:83:938–41.
8. Dwyer PJ, McGregor JR, McDermott EW, Murphy JJ, Higgins NJ. Patient recovery following cholecystectomy through a 6 cm or 15 cm transverse subcostal incision: a prospective randomized clinical trial. Postgrad Med J. 1992;68(804):817--9.
9. Lamiell JM, Wojcik ZM, Isaacks J. Computer auditing of surgical operative reports written in English. Proc Annu Symp Comput Appl Med Care. 1993.
10. Mamlin BW, Heinze DT, McDonald CJ. Automated extraction and normalization of findings from cancer-related free-text radiology reports. AMIA Annu Symp Proc. 2003;420-4.
11. Huang Y, Lowe HJ, Klein D, Cucina RJ. Improved Identification of Noun Phrases in Clinical Radiology Reports Using a High-Performance Statistical Natural Language Parser Augmented with the UMLS Specialist Lexicon. J Am Med Inform Assoc. 2005;12(3):275-85.
12. Friedlin J, McDonald CJ. A natural language processing system to extract and code concepts relating to congestive heart failure from chest radiology reports. AMIA Annu Symp Proc. 2006;2006:269-73.
13. Jiang M, Chen Y, Liu M, Rosenbloom ST, Mani S, Denny JC, et al. A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries. J Am Med Inform Assoc. 2011;18(5):601-6.
14. Halgrim SR, Xia F, Solti I, Cadag E, Uzuner Ö. A cascade of classifiers for extracting medication information from discharge summaries. J Biomed Semantics. 2011;2(Suppl 3):S2.
15. Jagannathan V, Mullett CJ, Arbogast JG, Halbritter KA, Yellapragada D, Regulapati S, et al. Assessment of commercial NLP engines for medication information extraction from dictated clinical notes. Int J Med Inform. 2009;78(4):284-91.
16. Lakhani P, Langlotz C. Automated Detection of Radiology Reports that Document Non-routine Communication of Critical or Significant Results. J Digit Imaging. 2010;23(6):647-57.

17. Denecke K, Bernauer J. Extracting Specific Medical Data Using Semantic Structures Artif Intell Med. 2007;4594:257-64.
18. Dang PA, Kalra MK, Blake MA, Schultz TJ, Halpern EF, Dreyer KJ. Extraction of Recommendation Features in Radiology with Natural Language Processing: Exploratory Study. AJR Am J Roentgenol. 2008 August 1, 2008;191(2):313-20.
19. Ware H, Mullett CJ, Jagannathan V. Natural Language Processing Framework to Assess Clinical Conditions. J Am Med Inform Assoc. 2009 July 1, 2009;16(4):585-9.
20. Turchin A, Kolatkar NS, Grant RW, Makhni EC, Pendergrass ML, Einbinder JS. Using Regular Expressions to Abstract Blood Pressure and Treatment Intensification Information from the Text of Physician Notes. J Am Med Inform Assoc. 2006 November 1, 2006;13(6):691-5.
21. 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.
22. Joint Commission on Accreditation of Healthcare Organizations 2012; Available from: http://www.jointcommission.org/.
23. Accreditation Association for Ambulatory Health Care. 2011; Available from: http://www.aaahc.org/eweb/dynamicpage.aspx?webcode=about_accred.
24. Cohen KB, Palmer M, Hunter L. Nominalization and Alternations in Biomedical Language. PLoS ONE. 2008 09;3(9):e3158.
25. Alphonse E, Aubin S, Bisson G, Hamon T, Lagarrigue R, Nazarenko A, et al. Event-based information extraction for the biomedical domain: the Caderige project. Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and its Applications. 2004:43-9.
26. Riloff E. Automatically generating extraction patterns from untagged text. Proceedings of the thirteenth national conference on Artificial intelligence - Volume 2. 1996:1044-9.
27. Baker CF, Fillmore CJ, Lowe JB. The Berkeley FrameNet Project. Proceedings of the 17th international conference on Computational linguistics - Volume 1. 1998:86-90.
28. Kingsbury P, Palmer M, Marcus M. Adding Semantic Annotation to the Penn TreeBank. In Proceedings of the Human Language Technology Conference. 2002.
29. Surdeanu M, Harabagiu S, Williams J, Aarseth P. Using predicate-argument structures for information extraction. Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1; Sapporo, Japan: Association for Computational Linguistics 2003. p. 8-15.
30. Giuglea A-M, Moschitti A. Semantic role labeling via FrameNet, VerbNet and PropBank. Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics; Sydney, Australia: Association for Computational Linguistics; 2006. p. 929-36.
31. Gildea D, Hockenmaier J. Identifying semantic roles using Combinatory Categorial Grammar. Proceedings of the 2003 conference on Empirical methods in natural language processing: Association for Computational Linguistics; 2003. p. 57-64.
32. Lippincott T, Séaghdha DÓ, Sun L, Korhonen A. Exploring variations across biomedical subdomains. Proceedings of the 23rd International Conference on Computational Linguistics; Beijing, China: Association for Computational Linguistics; 2010. p. 689-97.
33. Kilicoglu H, Fiszman M, Rosemblat G, Marimpietri S, Rindflesch TC. Arguments of nominals in semantic interpretation of biomedical text. Proceedings of the 2010 Workshop on Biomedical Natural Language Processing; Uppsala, Sweden: Association for Computational Linguistics; 2010. p. 46-54.
34. Wang Y, Melton GB, Pakhomov S. It's about This and That: A Description of Anaphoric Expressions in Clinical Text. AMIA Annu Symp Proc. 2011;2011:1471-80.
35. Gildea D, Palmer M. The necessity of parsing for predicate argument recognition. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics; Philadelphia, Pennsylvania: Association for Computational Linguistics; 2002. p. 239-46.
36. Stanford Parser. http://nlpstanfordedu/software/lex-parsershtml.
37. Marcus MP, Marcinkiewicz MA, Santorini B. Building a large annotated corpus of English: the penn treebank. Comput Linguist. 1993;19(2):313-30.
38. SPECIALIST Lexicon. http://wwwnlmnihgov/pubs/factsheets/umlslexhtml.
39. Miller GA. WordNet: A Lexical Database for English. Communications of the ACM. 1995;38(11):4.
40. The New Oxford American Dictionary. 2nd ed: Oxford University Press; 2005. p. 2096.
41. Stedman's Medical Dictionary. 28th ed: Lippincott Williams & Wilkins; 2005. p. 2100.