EVALUATION OF A NOVEL TERMINOLOGY TO CATEGORIZE CLINICAL DOCUMENT SECTION HEADERS AND A RELATED CLINICAL NOTE SECTION TAGGER

By

Joshua Charles Denny

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Biomedical Informatics

August, 2007

Nashville, Tennessee

Approved:

Associate Professor Anderson Spickard, III

Professor Randolph A. Miller

Associate Professor Kevin Johnson

ACKNOWLEDGEMENTS

I would like to thank my committee, Anderson Spickard, III, Randy Miller, and Kevin Johnson, for their close guidance and encouragement through this project. They each offered unique insights, wisdom, and perspectives that were invaluable in helping me complete this work. Finally, I would like to thank all my evaluators, without whom the quality of this project would be unknown. They enthusiastically devoted many careful hours to make the evaluation a reality.

I would like to thank the National Library of Medicine, whose generous support made this research possible (Training Grant T15 007450-04). I am also thankful to everyone in the Department of Biomedical Informatics for their support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The aims of this project are to build and evaluate a terminology that provides categorization labels (tags) for common segments within clinical documents. For example, “History and Physical Examination” (H&P) notes generally contain *sections*, such as “history of present illness,” “past medical history,” and “physical exam.” Many of these sections can have subsections, such as a “cardiovascular exam” under physical examination or “previous hospitalizations” under past medical history. Development of a tool to parse and label natural-language clinical documents using the new document tagging terminology comprises a key component of this work. The new section header terminology models the common section names, subsection names, and their relationships. The section tagging tool, named SecTag, identifies terminology matches from natural language clinical documents using a combination of linguistic, natural language processing, and machine learning techniques based in part on the KnowledgeMap Concept Identifier (KMCI) program previously developed by the author and colleagues.[1-5] The evaluation study focused on recognizing components of “history and physical examination” (H&P) notes that were generated during hospitalizations and outpatient visits. Clinical domain experts, as external reviewers, rated the SecTag’s ability to identify sections from H&P documents, and judged when SecTag failed to do so.

Electronic health records comprise a rich source of clinical information, including clinical observations, laboratory and imaging reports, and medical diagnoses. Most such records exist in the form of clinical narratives, composed of natural language text created as providers describe interactions with patients. Typical narrative documents include note types such as an admission history and physical, outpatient visit note, or discharge summary[6] and are typically stored electronically; these documents generally expressed in the "boundless chaos of living speech."[7] The primary purpose of generating clinical narratives remains to provide an efficient method of communicating among clinicians; generating structured notes in a "computer-understandable" format remains a distant desideratum due to workflow issues.[8, 9]

Identifying sections and subsections in clinical documentation is an important first step to providing context for understanding the concepts within a document. Much of the information in the text of a large clinical note cannot be easily "understood" by humans or software programs without the contextual clues provided by section headers. For example, it is far easier to disambiguate a term such as "friction rub" if one knows whether it appeared in the "pulmonary auscultation", "cardiac auscultation", "abdominal examination" or "joint examination" segment of an H&P note. In addition, "past medical history" and "family medical history" may both contain lists of identical disease names that, based on context, have very different implications for the patient's health, for use in clinical research, or for decision support.

Recognition of clinical document section tags is not trivial. The specific verbiage used to designate section tags in free-text clinical documents may vary significantly. While a "history of present illness" tag may be present in some form in most initial write-ups of a patient, a clinician may designate this heading by using one of many different acronyms, abbreviations, or synonyms (e.g., "HPI", "history", "history of current illness"). In addition, a given heading may be absent from the document altogether – which may mean that either the reader (and SecTag) must infer the presence of the omitted heading when the corresponding section content is present but unlabeled, or deduce that the section is missing entirely from the document when neither tag nor corresponding content is present. Medical language-processing applications must understand synonymy and the relationships among tags and related content to support inferences regarding the concepts contained within a section.

Providing formal, standardized section tags for unstructured clinical documents offers a number of potential benefits. Section tagging can facilitate applications such as clinical advisors (does the patient have a history of heart failure?)[10, 11], automatic problem list generators[12], systems to support medical education (has a trainee evaluated a patient with pneumonia or ever reported hearing a diastolic murmur?), and potentially aid in construction of structured documentation tools.[13-15] Recent research reports describe use of section identification techniques to generate problem lists[12], extract chief complaints[16], and to answer broader, more generic medical language processing tasks[17]. The current project is one of the first to create a comprehensive terminology for H&P section headers, and it has completed one of the first formal evaluations of a section tagger.

CHAPTER II

BACKGROUND

A Brief History of Clinical Documentation

Clinical documents comprise the record of care providers' encounters with patients. Modern clinical documentation originated in part with Hippocrates (460-370 BC), when he directed physicians of the time to observe and record their findings in a clear and objective manner. He recommended making careful observations of physical exam components such as pulse, fever, movement, and complexion.[18] He also directed doctors to inquire about a patient's environment (a component of a patient's social history) and family medical history. Others, such as Herophilus (280BC) and Vesalius (1514-1564 – considered the father of modern anatomical studies), contributed significantly to the understanding of functional systems as components of patient evaluations. Herophilus separated neurologic function and evaluation into sensory and motor components[19], while Vesalius dispelled prior beliefs about the roles of such major organs as of the liver and the heart and provided detailed anatomical drawings in his *De humani corporis fabrica* that served as a framework for anatomical- and system- based physical evaluations.[20] Sir Thomas Sydenham (1624-1689) has been called the "father of English medicine" and "the first true epidemiologist."[21, 22] He was a keen observer and maintained meticulous notes on his patients' findings and illnesses in order to identify causes and prognosis.[22, 23] After creating the monaural stethoscope, René Laennec (1781-1826) advanced the physical exam of the pulmonary and cardiac systems, including defining such terms as rales, rhonchi, egophony, and the two basic heart sounds ("S1" and "S2", though he misclassified them as atrial and ventricular systole).[24, 25] More recently, Sir William Osler (1849-1919) shaped the training in patient evaluation that current medical students and residents receive. He advocated the importance of the physician's interview, suggesting that "if you listen carefully to the patient, they will tell you the diagnosis."[26] Many others have contributed to this time-honored skill of physicians documenting history and physical examination notes as a component of overall patient evaluation.

Review of textbooks from the last seven decades shows that many of the same common sections headers – such as history of present illness, physical exam, vital signs, and neurological exam – have persisted relatively unchanged in clinical notes over many decades.[27-34] A "Physical Diagnosis" class is common in medical school, which teaches the elements of the patient interview and physical examination, and how to document the findings in a clinical narrative.[35-41] Evaluation of a student's writing and critical assessment skills, via the H&P, is a major component of the clinical instruction of medical trainees. Clinical note generation varies by the setting and patient characteristics. Inpatient visits can generate admission H&Ps, daily progress notes, procedure notes, consultant encounters, and discharge summaries.

The H&P is arguably the most developed form of clinical note, including a number of different common sections and a loose hierarchical organization thereof that have been in use for decades (if not centuries), as shown in Table 1.[27-34] This common organization is echoed in many structured note capture tools.[13-15, 42, 43] Many of these sections (e.g., "physical exam") commonly have subdivisions (e.g., "pulmonary exam"). The organizations of the subdivisions can be based on functional system or anatomical relation. For instance, "jugular venous pulse exam" (a technique that in part estimates the heart's right atrial pressure by visualizing the venous pulsation of the internal jugular vein) can either belong to the "neck physical exam" (an anatomical organization) or occur as part of the "cardiovascular physical exam" (a functional organization).[34]

Table 1: Common Clinical Sections in H&Ps.

Section Title	Subdivisions	Comments
Chief complaint	No	The reason for the encounter, usually as stated by the patient.
History of present illness	No	A brief history of the illness or complaints bringing the patient to medical attention. Can contain elements of a number of different sections (such as relevant past medical history or review of systems), but is not necessarily organized by subsections.
Past Medical History	Sometimes	Includes a record of the patient's past diseases, surgeries, hospitalizations, blood transfusions, and more. Sometimes these are divided into subsections.
Personal/Social History	Common	Often includes work and family environment and substance use.
Health Maintenance	Sometimes	Can contain common immunizations, screening tests results, and markers of disease progress.
Allergies and adverse reactions	No	Contains medication, product, and food allergies; rarely, these could be subsections.
Review of Systems	Yes	A survey of common symptoms that may or may not pertain to the chief complaint.
Physical Examination	Yes	Common subsections include vital signs and a broad range of system-based (e.g., cardiovascular, musculoskeletal) and anatomical-based (e.g., abdominal, oropharynx) exams.
Laboratory data and imaging results	Sometimes	Subsections can include certain categories of individual tests.
Problem list	No	A list of active and prior health concerns of the patient.
Assessment and plan	Depending on complexity	The provider's assessment of the patient's complaints, signs, and symptoms, and recommended plan for medication treatment or workup to further define the diagnosis.
Attestations	Sometimes	At a teaching hospital, can vary from a simple acknowledgement of attending physician review of a given document to a short synopsis of the patient's medical history, assessment, and plan.
Team Members	Sometimes	At a teaching hospital, often includes the service name, attending physician, and any housestaff physicians.

Roles Played by Clinical Documents

Although the structure of clinical documents has remained largely unchanged in recent history, the role of clinical notes, such as H&Ps, has evolved over the centuries. From the time of Hippocrates to Osler, clinical notes primarily served to remind their clinician-author of the clinician's own findings from taking the patient's medical history and the clinician's own care plan, and were usually privately held and maintained by a single individual clinician. These notes on rare occasions provided communication to a small number of other physicians. Today, clinical documents serve as a tool for coordinating a care plan among members of a large patient care team, including physicians, nurses, and other allied health professionals (speech pathologists, pharmacists, social service workers, and others).[44] Review of clinical documentation is often a requirement for billing. The United States Evaluation and Management Coding ("E/M Coding") system determines the level of reimbursement a physician can receive based on the clinical complexity of the case and the extent of documentation.[45] The latter characteristics are determined by counting the number of clinical sections and subsections in the note. For physicians to bill at a certain level, they must include an appropriate number of clinical note sections, such as a history of present illness or review of systems (itself including a given number of subsections, such as "cardiovascular" or "gastrointestinal").[45] Clinical documentation also serves as the legal record of what is known about the patient and what the plan of care is for the patient. Finally, clinical documents serve as data for both case reports and large-scale clinical research. For example, the well-known Harvard Medical Practice study[46, 47], along with others[48, 49], reviewed medical records to estimate the number of hospital deaths.

Methods for Generating Clinical Documents

As the roles of notes have expanded and with the growth of electronic medical records (EMRs), new methods for generating clinical notes, such as H&Ps, have expanded. In Osler's time, clinical notes were short, poorly-structured, hand-written documents. The amount of documentation expanded with requirements for conveying information about the patient to multiple interested parties, with documentation requirements for billing[45] and with the need to serve as the legal record of care.[50] To generate voluminous documentation efficiently, dictation became a commonly used method for note creation in recent decades.[51, 52] Within the past decade, due to the rising cost of transcribing dictated notes, many physicians began using voice recognition systems to replace dictation and transcription.[53, 54] With the advent of EMRs, investigators have developed a number of approaches for direct, physician-generated electronic clinical note capture. Despite much effort to produce structured electronic notes[13-15, 42, 43], most of these systems still produce notes in a free-text format, often dictated or typed using loosely-defined memorized categories or user-defined templates.[14] A few have designed systems that capture structured data; however, these systems remain less common.[13, 15]

Computer-based documentation tools provide a method to capture and store records of patient-provider interaction, allowing the use of tools to potentially improve the quality of documentation by prompting the user for certain information, provide data for clinical research, and serve as a substrate for decision support systems. To be useful, the clinical note data must be structured in a format that is computable. Since most note capture tools, however, still record notes in human-language text, many systems must rely on natural language processing systems to produce structured data for use in computer-based tools.[55-59]

Natural Language Processing and Concept Identification

The converse of capturing a clinical note in a structured format is to allow clinicians to generate notes in traditional “free text” natural language format, and to then, post-facto, apply natural language processing (NLP) systems that contain algorithms to convert human language narratives into machine-readable, coded data.

Some investigators divide NLP into two tasks: “natural language understanding”, which is the process of converting the human-generated text into computer data, and “natural language generation”, which produces human readable text from structured computer data.[60] When used generically, NLP typically refers to the natural language understanding tasks.[16, 17, 55, 57, 61, 62]

Producing a computer-understandable output from a natural language source is a challenging and complex task. The complexity can be divided by syntactic and semantic axes as shown in Table 2 (adapted from Wulfman et al.).[63] Semantic complexity refers to the breadth of concepts covered and the depth of understanding required for the tool to function. Syntactic, or linguistic, complexity refers the degree of “humanness” of the source note: the diversity of sentence structures, words, negation formats, and general organizations (i.e., paragraphs, outline lists, section headers, etc.) used. Tools are needed for each task, and tools addressing more complex problems can build on simpler ones.[63]

For example, a tool to automatically create problem lists (a high syntactic and semantic complexity task) would benefit from a clinical note section tagger (a high semantic complexity but lower syntactic complexity). An early step in NLP systems is often concept identification, whereby a system maps strings to concepts from standardized terminologies such as the UMLS. The NLP approaches also often include components of deeper understanding, such as interpreting certainty of a concept's presence (is it possible, present, or negated?), temporal reasoning, or assigning value attributes to a concept (e.g., the dose and route of a medication).

Table 2: Syntactic and Semantic Complexity of Some Natural Language Processing Tasks

		Syntactic Complexity of Source	
		Low	High
Semantic Complexity of Interpretation	Low	• Form recognition with limited options • Extracting single entities (e.g., code status[64]) from clinical notes via string searching	• Most text classification exercises • Classifying clinical notes into major categories or quality[65] • Identifying drug-drug interactions from notes or abstracts[66]
	High	• Interpreting ECGs[1], echocardiograms, or radiology reports[67] • Identifying section headers from notes	• Full understanding of clinical notes • Identifying Adverse Events from notes[56] • Generating problem lists by screening clinical notes[55]

Beginning with the Sager's Linguistic String Project (LSP) of the 1960s[59], many research groups have furthered the ability of clinical natural language processing applications to generate "computable understandable and processable" renderings of clinical narratives.

The LSP created lexical and syntactic methods for processing text, and added sublanguage grammars that were tuned for a given specialty (such as a chest x-ray report or anatomic pathology report). A study of LSP in 1994 evaluated its recall and precision to identify key treatment concepts from asthma discharge summaries. The LSP researchers found an overall recall of 82.5% and overall precision of 82.1%; on major concepts, the recall and precision improved to 92.5% and 98.6%, respectively.[59]

To aid common understanding of documents, many researchers have developed systems for identifying concepts within clinical and biomedical texts (mapping “mad-cow disease” and “bovine spongiform encephalopathy” to the same concept with a unique identifier). Among many efforts, Cooper and Miller developed and applied three methods to identify concepts from MEDLINE.[68] The NLP systems MetaMap[69], SAPHIRE[70], KMCI[4, 5], a system developed by Nadkarni et al[71], and IndexFinder[72] each have taken unique approaches, using a variety of linguistic tools, to identify UMLS concepts from within text.

The MetaMap system developed by Aronson and colleagues at the National Library of Medicine uses a robust scored-based algorithm, statistical parsing, variant generation, and semantic rules to identify biomedical concepts and has been applied successfully to many document types, including clinical text[55], MEDLINE[73], and patient email messages[74]. By itself, MetaMap does not provide natural language processing components such as negation detection (e.g., “she *denies* chest pain”), syntactic transformations (such as identifying both “arm pain” and “leg pain” from “pain in arm and leg”), or context information (e.g., distinguishing between family medical history and personal history of a disease). Chapman et al. extended MetaMap with a regular expression-based negation detection scheme termed NegEx.[75] The KMCI system provides variant generation and semantic rules similar to MetaMap but adds document-based and proximity-based scoring techniques to classify ambiguous concepts.[4] The KMCI normalization engine uses UMLS components but incorporates an expanded lexicon and lists of prefixes and suffixes. It uses some natural language processing techniques for acronym discovery and semantic phrase regularization of conjunctions (e.g., “pain in arm and leg” maps to the concepts “arm pain” and “leg pain”). Recently, the author and colleagues extended KMCI with negation tagging based on the NegEx algorithm with recall of 0.973 and precision of 0.982 on electrocardiograms.[1] Elkin et al. has developed the Mayo Vocabulary Processor which identifies SNOMED-CT concepts and detects negated concepts.[76, 77]

Many authors have built on concept recognition systems concept recognition to yield greater understanding of natural language text. Using MetaMap, Sneiderman, Rindflesch, and Aronson developed the FINDX program to identify medical findings in text.[78] The SemRep program uses MetaMap combined with natural language processing techniques to identify propositions within medical text.[79] The latter system relies heavily on the logic of the UMLS Semantic Network and the semantic type of concepts (as defined in the UMLS). Fiszman et al. extended SemRep to find hypernymic propositions, the more general proposition in a sentence suggesting two relationships.[62]

However, each of the above systems indexes at a sentence or noun-phrase level and does not explicitly segment documents by section; thus, they do not distinguish between "congestive heart failure" encountered in the family medical history, past medical history, or the assessment sections in a clinical note.

Through the 1990s and 2000s, Carol Friedman and colleagues at Columbia University developed the Medical Language Extraction and Encoding (MedLEE) system, arguably the best and most highly regarded clinical NLP system at present. MedLEE processes notes through five steps: a document preprocessor, text parsing, error recovery from bad parses, phrase regularization, and concept encoding.[57] The document preprocessor segments a document into sections (e.g., "history of present illness"), sentences, and concepts. Sections are labeled by type according to a limited vocabulary and applied to all concepts found within. MedLEE is a natural language processing system that now supports many clinical document types including discharge summaries[56, 57], radiograph reports[67], mammograms[58], and pathology reports[80], among others. More recently, extensions to extract phenotype-genotype associations (BioMedLEE) from biomedical literature found similar performance to experts.[81] Melton and Hripscak evaluated use of MEDLEE for adverse event reporting based on NLP of patients' discharge summaries. combined with a series of adverse event rules.[56] Heinze et al.[17] and Hazlehurst et al.[82] have reported on similar NLP systems; the former involved some degree of string-based section tagging.

Efforts Focusing on Automated Clinical Text Classification

Text classification is the process of automatically assigning categories (or labels) to documents or blocks of text within a document.[65] In general, text classification systems do not attempt to "understand" a document (as an NLP system would) but rather "classify" it as a whole. This reduction simplifies a semantically complex task into a much simpler one by using statistical or machine learning techniques operating on a syntactically complex substrate such as a clinical note. Section header identification is a process of assigning labels to small blocks of text within a document. Most research, however, in text categorization has focused on assigning an entire document to one or more categories[65, 66, 83-91], although some investigators have also developed systems for classifying text by paragraphs or sentences (e.g., a textbook index which would categorize paragraphs or sentences of text as belong to different topics).[92] One well-known example of whole-document classification is MEDLINE, which is manually classified according to the MeSH terminology.[93]

Many authors have applied machine learning methods, such as support vector machines (SVMs) or naïve Bayes techniques, to organize text by different classification scheme.[65, 66, 83, 86, 87, 94-96] Binary classification tasks label text according to two categories, such as "fracture" vs. "no-fracture" or "interesting" vs. "not interesting". Aphinyanaphongs et al. applied SVMs to find high-quality Medline articles.[65] In that study, SVMs performed significantly better than did text-based clinical queries, naïve Bayes algorithms, and text-boosting algorithms. In another binary classification experiment, de Bruijn et al. found that SVMs better identified acute fractures from wrist radiograph reports than three information retrieval techniques and several other machines learning methods.[83] That system had an accuracy of 94%; however, other machine learning methods, including neural networks and naïve Bayes, also performed well with accuracies of 85-88%.[83] Other investigators have used decision trees[97], linear least squares[96], and maximum entropy models[98] with success for binary classification tasks. In addition, the computer science literature is replete with examples of machine learning techniques to classify email as "spam" or important. Investigators have effectively ignored "spam" email using naïve Bayes[85, 87, 89], support vector machines (SVMs)[84, 86, 94], and Boolean information theoretic approaches[91], among other techniques.

Multiclass classification, such as the current task of section classification, involves predicting among many categories, and is computationally more complex.[99] Pakhomov et al. implemented a naïve Bayesian classifier to assign diagnostic codes based on clinical notes.[95] Among 35,676 possible diagnostic codes, the system achieved a recall and precision of greater than 98%. Lee et al. implemented a naïve Bayes classifier to predict the section topics for 15,000 OMIM articles among 25 different categories.[100] Naïve Bayes models have also been used to predict syndromes.[101] Cheng et al. found a naïve Bayes model boosted with chi-square feature selection methods superior to SVMs using alignment only for predicting protein subfamilies.[102] Using informational retrieval techniques, the author and colleagues achieved sensitivities of 0.78-1.00 and specificities of 0.85-0.96 identifying four broad topics from medical education documents using UMLS hierarchies on the topic of interest.[3]

The above methods focused on classifying an entire document into categories; section identification involves classifying segments of the document into one of many classes. Relevant to this work, Berrios et al. created tools to index textbooks at the level of their sentences or paragraphs.[92, 103-106] The MYCIN II system allowed annotation of sentences with concepts and different templates (e.g., “amoxicillin treats organism”) that effectively classified the sentences of the document into different templates for latter queries.[104] The Internet-based Semi-automated Indexing of Documents (ISAID) system used a UMLS concept identifier to help automate this process by suggesting possible categories to the user.[92]

Standardized Terminologies: Development Principles and Design Goals

Major stakeholders in US healthcare have long recognized that the lack of interoperability between EMR systems is a major impediment to health care quality.[44, 107] Development of common terminologies to represent and communicate health information offers the potential to improve communication and share tools to improve healthcare across different clinical applications and health care systems.[108-111] Initial standardization work for clinical applications focused on development of terminologies to support administrative and billing functions such as the International Classifications for Diseases (ICD), but has extended to support clinical domains, research, and natural language processing through such efforts as the UMLS, SNOMED-CT, and RxNorm.[76, 112-114]

The principles of terminology development informed development of the section header terminology for its data representation, specified goals, and organization. Terminologies consists of groups of strings (words, phrases, or other characters), called terms, that are often conceptually grouped by a common unique identifiers. Creators of terminologies often prespecify the goals for the terminology and the processes by which new terms are to be created, naming schema, and metadata to be included for each term. In Cimino's "Desiderata for Controlled Terminologies," he argues that controlled vocabularies should have unique, unambiguous concepts linked by hierarchical relationships (e.g., "right ventricular myocardial infarction" as a child of "myocardial infarction").[115] He also argued that terminology developers should reject catch-all "not elsewhere classified" terms since these terms only have an ever-changing definition by exclusion of existing terms. While each concept should be unambiguous, one must distinguish between different contextual usages of the concept ("he had a myocardial infarction" versus "he had a family history of myocardial infarction") and semantic ambiguity, such as "cold" to refer to an upper respiratory infection, a temperature, or a subjective feeling. Cimino argued that controlled terminologies should support *polyhierarchies*, such that "hepatorenal syndrome" can be a child of both "hepatic diseases" or "renal disease." SNOMED-CT supports multiple hierarchies; nearly 28% of its concepts have more than one parent.[116] Gene Ontology (GO) also supports polyhierarchy using a directed acyclic graph structure.[117]

Cimino argues that the biggest criticism of any terminology is the completeness of its content for its domain.[115] To this end, terminologies should provide formal methods to expand their content, instead of haphazard expansion. Chute suggested terminologies should seek complete domain coverage and integrate with other terminologies when they are lack comprehensiveness.[118] In the 1990s and 2001, the International Standards Organization (ISO) specified that terminology developers should define their purpose, integrate with existing terminologies, and quantify the terminology's domain coverage.[119-121]

Terminologies are grouped by their purpose. Reference terminologies support storage, retrieval, and interoperability as "deep" internal representations.[112, 122] Administrative terminologies, such as ICD-9, can support communication to external agencies for billing. Clinical interface terminologies are a systematic collection of terms to supports entry of information into computer systems expecting structured data.[112, 123] Often, they are mapped to reference terminologies. The goal of an interface terminology is to represent formal concepts with the colloquialisms of the terminology's users and allow users to construct concepts quickly. Thus, interface terminologies need broad synonymy.[112] Similarly, a provider would need an interface terminology for section headers in constructing a note (to easily select the section headers or to recognize them after the note was written) and computer systems would need an interface terminology for section headers to store information by sections, share across systems, and infer conclusions from notes.

The UMLS is an example of a terminology that functions as both an interface and a reference terminology, supporting broad-scale synonymy, concept orientation, and polyheirarchy.[124] Instead of producing a new "standard" vocabulary, UMLS curators assembled a variety of existing terminologies that have been linked by common identifiers.[111] This aggregate terminology is named the UMLS Metathesaurus and contains more than 100 component vocabularies. Terms in the Metathesaurus are grouped into concepts by concept unique identifiers (CUIs). Each unique term (i.e., string) has a unique identifier (a SUI). Each concept (i.e., a CUI) is unique and unambiguous. For instance, the string "cold" as a single SUI that links to multiple CUIs: the upper respiratory virus, a sensation of feeling cold, and a temperature, among others. Each string contains a variety of metadata, including the string type (a "preferred term", a "suppressible synonym", etc), its source language, and other information. This information is useful in natural language processing applications, for instance, as "suppressible synonyms" are often ignored.[4, 69] Each concept is assigned one or more of 135 different "semantic types" (e.g., "chronic obstructive lung disease" is a "Disease or Syndrome"; "penicillin" is both an "Antibiotic" and an "Organic Chemical"). The UMLS maintains the hierarchical relationships inherited from its component vocabularies, translating the structure of each into a common relationship format, such as parent-child. To support natural language processing tools, the UMLS contains the SPECIALIST lexicon, a collection of core biomedical terms with their part of speech, inflectional forms, common acronyms and abbreviations. The Semantic Network, the third major component of the UMLS, contains relationships between the conceptual semantic types, such as "Pharmacologic Substance – treats – Disease or Syndrome." It also contains a hierarchy relating the semantic types (e.g., an "Antibiotic" IS-A "Pharmacological Substance").

Terminologies Useful for Section Tagging

Currently available structured terminologies were not designed to represent the hierarchy or vast synonymy of H&P sections in clinical notes. Despite the growth of the UMLS, review of the expressed design goals of its component terminologies reveals that none of them were purposefully designed to represent for clinical note section headings.[125] For example, in the UMLS, the review of systems or physical exam subsection "Head, Ears, Eyes, Nose and Throat" exists as concept C1512338. The concept includes a synonym "HEENT" but no clinical section header children or parent concepts. Similarly, while many individual sections headers (such as "review of systems" or "physical exam") exist in the UMLS, the source hierarchies do not contain robust, organized set of subsections. A notable exception may be the Logical Observation Identifiers Names and Codes (LOINC), which was designed "to facilitate the exchange and pooling of results" by providing "universal identifiers for laboratory and other clinical observations."[126] LOINC has grown to include the major section headings of history and physical, discharge summary, and operative note reports.[126] While LOINC provides a number of individual sections and subsections for each of these note types, it does not provide synonymy (that "history present illness" and "HPI" indicate the same concept) or a detailed, multi-level hierarchy (e.g., that a "HEENT" exam contains an eye exam which can contain a more detailed level of exam including a fundoscopic exam, a scleral exam, a slit lamp exam, and more).

Among its many other components, the LOINC document section header terminology includes 310 canonical terms; 186 of these are unique. It formally represents one degree of hierarchy in a using a “component name” and a “system” specifier (e.g., component=“Physical findings”, system=“genitalia”). A third level is informally represented for a few concepts using period notation (e.g., “physical findings.shoulder shrug” with system “shoulder.bilateral”). In general, LOINC does not include concepts synonyms (or contains very few synonyms for some concepts), although one can find some synonyms by using LOINC’s integration with other vocabularies in the UMLS. LOINC also includes other attributes about each concept, such as method of formulation (reported or observed), properties, and times over which they can be measured (most H&P findings have “point” time values, meaning they occur at a particular time).

The Quick Medical Reference (QMR)® Knowledge Base is another source of section header names and findings. QMR was the result of 35 person-years of development effort to build an evidence-based diagnostic support engine. It organizes more than 4,000 common clinical findings into a multilevel hierarchy of 525 elements (the “findings hierarchy”); it contains many headers recognizable as H&P section headers, such as “Review of Systems” or “Family History.”[127-129] Since QMR’s hierarchy is based on the clinical findings used to construct the QMR KnowledgBase, its organization is limited to the findings in its vocabulary instead of clinical practice. However, QMR includes elements from the patient’s history, past diagnoses, physical exam findings, and laboratory, imaging, and pathology results; its breadth approximates the majority of common clinical practice. More than half of its hierarchy represents laboratory, imaging, or pathology tests; 173 represent a hierarchy of common physical examination or patient history elements. A complete list of the QMR hierarchy is available in Appendix A.

Many of these 173 headers represent common elements one would expect to find labeled in a history and physical, such as “Vital Signs” or “Cardiovascular Exam.” Some of QMR’s headers are disease- or symptom- based, such as “Pain Back” (which groups types of back pain), rather than representing an H&P section header; these items are not appropriate as a clinical document section headers.

LOINC and QMR have both been well studied[127-134], but not as interface terminologies for identifying and manipulating sections of clinical notes. LOINC has enjoyed widespread adoption as a reference terminology for laboratory and imaging reporting.[135] However, since both are reference terminologies, the utility of their document section header hierarchies as interface terminologies for clinician to use in daily practice may be limited. Neither, for example, represents many, if any, synonyms for its section concepts, making application to clinical text difficult given that “free text” in clinical documents contains frequent use of ad hoc abbreviations, acronyms, and synonyms. The concept string construction of QMR and LOINC, while exacting and appropriate for a reference terminology, is often awkward for a user; a user is not likely to use the header “Review of Diseases of Congenital or Inherited Nature History” (QMR) or “Physical findings.sensation” (LOINC) in their note.

Other vocabularies, such as SNOMED-CT or Medical Subject Headings (MeSH), contain section headers but the have incomplete coverage and their organization is incomplete. For example, children of the concept “physical examination” in MeSH include a number of generic methods of examination such as “auscultation” and “palpation” (which includes only the child “digital rectal examination,” a tenuous child concept at best) combined with a only a few functional sections (“neurological examination” and “muscle strength”) but no anatomical groupings.[136] SNOMED-CT, for instance, includes a number of concepts relating to “cardiovascular exam” such as “cardiovascular examination and evaluation,” “full CVS examination,” and “brief examination of cardiovascular system,” but not any clinical subsections such as “cardiac auscultation” or “jugular venous pulse assessment.”[136, 137]

The HL7 Clinical Documentation Architecture

Recognizing the need for a common representation for these clinical documents, such as H&Ps, the Health Language 7 (HL7) effort has created the Clinical Document Architecture (CDA). The CDA provides a common structure (i.e., an “empty shell” with rules for how elements may be added, but which does not specify any content) and semantics for human-readable clinical documentation that promotes machine-readability and interoperability across multiple platforms and institutions. This common structure defines a method to represent clinical narratives, clinical section headers (e.g., “physical exam”) and their subsections (“cardiovascular exam”), and detailed, computer-understandable interpretation of the clinical narrative (i.e., the output of an NLP system). The CDA defines common methods for representing these section headers and detailed computer output using standard terminologies as LOINC and SNOMED CT. The first version (Release 1) was unveiled in 2000 and was the first specification derived from the HL7 reference information model (RIM)[138]; Release 2 became public in 2005.[139] Both are American National Standards Institute (ANSI) approved. The CDA has found widespread acceptance for application in electronic environments worldwide[140-142] and in the United States[139, 143, 144]. Both CDA releases[138, 139] utilize HL7 version 3[145] and are represented using eXtensible Markup Lanugage (XML).

Figure 1 shows a hypothetical parsed document fragment. The structured document is contained within the “structuredBody” element, which contains the original narrative text (as text elements) subdivided into multiple section elements (e.g., “chief complaint”, “review of systems”). Each identified section contains the code identifying the section name, the codeSystemName (typically this would be LOINC for sections), and the codeSystem, a numerical code to identify the terminology and version. Sections can be nested, so that one can specify the “cardiovascular exam” of the “physical exam.” The nested section organization provides the context for the elements and concepts defined in the narrative block. The section title element is the author-label for the section. The section “text” element contains the author-derived narrative block for the section as originally written, though optionally, XML references to identified concepts can be inserted into the text. The final element, “entry,” contains identified concepts and attributes from the text block, potentially the output of a concept identifier or natural language processing tool. The entry element defines specifications for concept representation, relationships between concepts (e.g., observation1-CAUS-observation2), the time of an event (“cholecystectomy in 1980”), and the value of a concept (“temperature is 98.6F”). All elements except for the narrative text are optional.[138]

Original Narrative:	**CDA Representation:**
Chief Complaint: shortness of breath History of Present Illness: Mrs. Smith is a 31yowm with CHF (EF 20%), COPD, and LAM who presents with acute onset dyspnea while…	<ClinicalDocument> …CDA Header… <structuredBody> <section> <code code="10154-3" codeSystem="" codeSystemName="LOINC"></code> <title>Chief complaint</title> <text>shortness of breath</text> <entry>[coded observations and attributes of the original text] </entry> </section> <section> <code code="10164-2" codeSystem="" codeSystemName="LOINC"></code> <title>HPI</title> <text>Mrs. Smith is a 31yowm with CHF (EF 20%), COPD, and LAM who presents with acute onset dyspnea while… </text> <entry>…</entry> </section> </structuredBody> <ClinicalDocument>

Figure 1: CDA release 2 encoding of a clinical note fragment.

Algorithms for Detecting Clinical Document Sections

Few researchers have formally studied algorithms for automated identification of sections within clinical documents. Meystre and Haug created a natural language processing system to generate problem lists from clinical documentation.[12, 55] The first step involved a document parser that identified sections within clinical documents Their parser used string-matching techniques to recognize specific strings in documents as identifying section headers. Strings between 3 and 52 characters long preceded by a blank line or ending with a colon or newline character, depending on the document title, were matched against a list of candidate section headers. Their system categorized all text from the beginning of one section header to the start of the next recognized section header as belonging to the first section header.[146] The system terminology mapped 539 titles to 20 canonical section names but did not contain a hierarchy of these section concepts. The terminology was mapped to LOINC identifiers, where possible. Many strings were "nonsection" strings, such as "and throat" or "at"; these strings did not match sections but were ignored when they are encountered to improve the algorithm's specificity. The system developed by Meystre and Haug also ignored subsections, such as "cardiovascular exam" of a "physical exam" section.

Other groups have also used string-matching techniques to identify section tags, such as the medical language processing tools MedLEE.[57, 61] MedLEE's document preprocessor uses string matching to recognize some major section headers, such as discharge diagnoses, history of present illness, or hospital course, and applies the recognized label to the adjacent block of text. The MEDLEE system also encodes section-type information on a sentence level, using syntactic and semantic information to parse phrases such as "status post heart transplant" to understand that the heart transplant was a past event. Lifecode[17] and MediClass[82] also recognize conceptual context by parsing sentences, though it is not clear that they identify section headers in documents to tag blocks of text.

Finally, few groups have reported algorithms designed to identify section headers that were not explicitly tagged in the clinical document of interest by the author as the author generated the document. For example, an "implicit" tag is present in the text, "mother has a history of coronary disease” (i.e., implicit tag is “mother’s medical history” as part of “family medical history”). Similarly, few attempts have been made to disambiguate explicit or implicit section tags based on expected order of appearance of sections within a document – e.g., determining that the paragraph following the “chief complaint” is likely the “history of present illness,” even if the author did not include a label “History of Present Illness” there.

Summary

Clinical documentation has evolved as the core way of communicating a patient’s changing medical history, key findings as they unfold over time, and varying plans for diagnosis and treatment. Providers typically write clinical narratives using natural language text but follow a common format that divides the text into sections and subsections, which provide context and understanding to the concepts contained within them. While much work has been done in natural language processing and text categorization, the process of section header identification and section header terminology development for clinical documents has not been formally evaluated. Existing systems map the most common sections using simple string-matching techniques of labeled sections in documents. Existing terminologies lack domain completeness and synonymy to allow efficient, detailed parsing of sections from clinical notes. No current section parsing algorithms were designed to identify sections that have not been labeled by the author.

CHAPTER III

DEVELOPMENT OF THE SECTION HEADER TERMINOLOGY AND RELATED SECTION TAGGER DESIGN

Overview

As a first step to constructing a section header tagging application, the author developed a section header terminology to serve as both an interface and reference terminology, attempting to model all clinically-relevant section tags in H&P documents. Initial terminology construction involved manual review of a large number of past and current physical diagnosis textbooks, examination of existing H&P "templates" used by a variety of clinical subspecialty groups, and direct manual review of a large sample of H&Ps. The H&Ps were selected randomly from the institutional electronic medical record system. Those H&P documents were not used in the evaluation of the system described in Chapter IV. The resulting clinical document section header terminology contained 1109 concepts with 4,332 terms.

The second stage of terminology system development created the SecTag application, which identifies sections tags from natural language text using the terminology developed in the initial step. The tagger identifies both those labels specifically appearing in the document as section headers, and the tagger deduces the presence of implicit section headers (i.e., those not labeled by the author – for example, "40 pack-year history of smoking” implies presence of a “tobacco use history” section header, even if it is not preceded by the overt section tag “Tobacco use history:”). The SectionTagger can function as either a standalone application or a preprocessor for other applications, such as the KnowledgeMap Concept Identifier (KMCI) or other natural language processing system.

The evaluation of SecTag and the clinical document section header terminology is described in Chapter IV. The evaluation focused on SecTag’s performance identifying sections over a broad range of de-identified “history and physical” documents, focusing on SecTag’s recall and precision on identified sections and the sensitivity for identifying “major” sections, whether labeled or unlabeled, in target documents. This study was approved by the Vanderbilt University Medical Center Institutional Review Board, #070129.

H&P Corpus for Terminology and SecTag Development and Evaluation

To develop the section header terminology and provide a testing and training corpus of documents, the author created a corpus of electronic H&P documents, divided into a training set and an evaluation set. The training set was used for development of the terminology and SecTag; it was also used to train the machine learning component of SecTag. The evaluation set was not used until the evaluation (Chapter IV).

Figure 2 shows the steps whereby the author selected the random H&Ps. Since the type of note was not identified in the EMR, an automated program randomly selected 25,000 notes from the entire note set of EMR, restricting to those notes whose titles contained "H&P", "admission", or "history." This "candidate H&P set" represented a combination of typed and dictated notes. These notes were de-identified of the 18 HIPAA safe harbor provisions using DE-ID, a commercially-available scrubbing software, and other post-processing refinements.[147] DE-ID was originally developed for pathology reports but has been expanded to other forms of clinical documentation. After optimization, an internal analysis of 200 records found it removed 5378 out of 5472 (98.3%) identifiers, with an aggregate error rate – which includes any potential error, including non-HIPAA items, partial items and not inherently identifying items – of 1.7% (95% CI 1.4% to 2.1%).

The candidate H&P set contained 166 different note titles. The author filtered these notes by title to identify more precisely all H&Ps by looking through randomly chosen documents from the training set for each note title. In some cases, all documents with a given title were categorized individually by author review. The goal was to exclude short outpatient clinic notes, short attending attestations (primarily used for billing purposes), and other non-history type documents (e.g., brief notes on a new admission or note addendums). The resulting history and physical document set was randomly divided by an automated program into an evaluation set (n=9567) and a training set (n=1200), maintaining similar proportions of individual document titles.

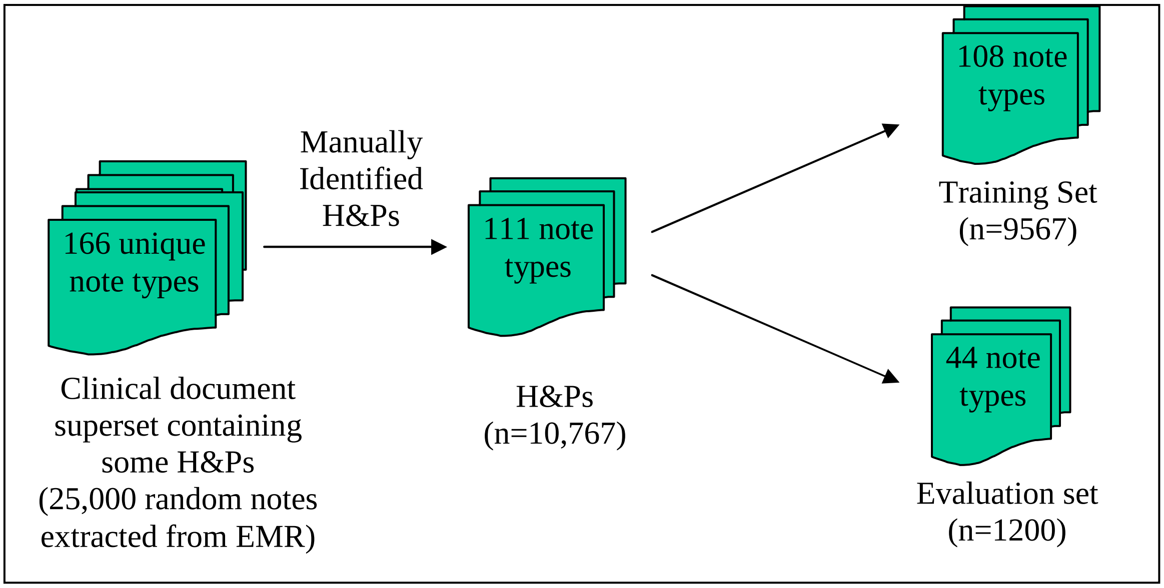

Figure 2: Shared Note Corpus Flow Chart
The Training Set was used for terminology construction and training and development of SecTag. The Evaluation set was used only for the evaluation (Chapter IV).

Section Header Terminology Construction

Overview

As previously described, the author created a section header terminology to serve as both an interface and reference terminology for the major section headers in H&Ps. The terminology was created using concepts derived from LOINC, QMR, history and physical exam textbooks, and review of H&Ps from the training set (Figure 2). The terminology data model was based loosely on the organization of the UMLS, maintaining a concept-oriented structure. Finally, the author created a medical word dictionary to serve as input for the SecTag's spelling correction algorithm.

Section Header Terminology Creation

The goal of the newly developed section tagging terminology is to provide a list of concepts and synonyms that can function as both a reference and interface terminology[112] for clinical section headings and their subsections. The author defined a “section” as a clinically meaningful grouping of symptoms, history, findings, results, or clinical reasoning that is not based on a diagnostic concept. A valid clinical document section (segment) header would include words that provide context for the encapsulated text but whose words themselves do not add specific clinical information, such as a diagnosis or symptom. For example, “back pain” is not a valid section tag because it is the name of a symptom that may or may not originate in the back (e.g., a perinephric abscess and acute pancreatitis can present with back pain), while the word “back” in the phrase “back: pain on flexion” would indicate the anatomical region and would be a valid tag in that it could give a location to the word “pain” it could contain. Similarly, a “past medical history” (a valid section) of “back pain” provides context and timing to “back pain.” Removal of the concept “past medical history” does not alter the presence or absence of any given finding in the note. Using these understandings as a guide, the author sought to develop a terminology to adequately represent H&Ps. This provides an initial step toward representing clinical documentation sections on a broader scale, including clinic notes, discharge summaries, and progress notes, which share many basic elements with history and physical notes.

The QMR findings hierarchy and LOINC were key enabling reference vocabularies for this project. Vanderbilt has been given permission by the University of Pittsburgh to use the QMR Knowledge Base for research purposes, and the QMR vocabulary per se (as opposed to the knowledge base) has been declared to be in the public domain. The author used the basic organizational structure of the QMR Findings Hierarchy for the initial hierarchical structure of the new section terminology, keeping approximately 150 patient history and physical headers and 160 laboratory, imaging, and pathological headers. The author then revised the hierarchy by incorporating all relevant LOINC headers (approximately 155 unique strings), modifying the structure as appropriate. The author expanded and revised the section hierarchy, concepts, and synonym lists based on the review of several clinicians and general and subspecialty clinical textbooks from across many decades.[27-34] The author obtained and incorporated the list of section terms created by Meystre and Haug.[12] It contained 539 strings, many of which were mapped to LOINC terms and to a common "concept name." Combining these three elements resulted in the first section header terminology draft.

To revise the terminology based on actual clinical notes, the author examined H&Ps and H&P templates selected from sample EMRs from the training set. In the Vanderbilt EMR, users can create H&Ps via several mechanisms: a template-based "notewriter," dictation (which may or may not be template-based, with variable templates), or hand-written notes (which are later scanned into the EMR; estimated at <5% of outpatient encounters and virtually no inpatient H&Ps). Users can also type documents without a template or upload documents written in Microsoft Word® into the EMR.

First, the author searched through all active EMR notewriter H&P templates (n=82) for strings that appeared to be section titles. The author processed these templates with SecTag and marked all strings that were possible section titles for review. Focusing on sensitivity, this list included any string containing at least one letter and less than 55 characters that matched any of the following:

- Contains multiple capital letters anywhere in the sentence (including strings with boundaries between uppercase and nonuppercase characters such as “OP clear” or “TEMP 97”).

- Ended in a colon, dash, or period. Strings ending in periods must start with a capital letter at the beginning of the line.

- Matched a concept in the terminology after phrase and sentence filtering (see below).

This process resulted in 1045 candidate section strings; of these, the author added 301 new synonyms or concepts added to the evolving section header hierarchy. The author also manually reviewed a number of subspecialty clinic notes to assure captured of very detailed elements of the physical exam and past medical history that may pertain only to certain subspecialties, such as neuro-ophthalmology or rheumatology.

The third step was evaluation of the then-current section header terminology against the training corpus of H&Ps extracted from the EMR (Figure 2). The author wrote a program that processed the training set of documents using the same rules above to look for possible section strings not currently in the terminology. This resulted in 10,138 additional unique strings. The author manually reviewed all tags with more than 20 occurrences in the document corpus, resulting in only 13 additions to the terminology. Through manual review of several hundred other documents, additional terms were also added.

The final terminology used for the study contained 1109 concepts and 4332 synonyms with a maximum depth of 10 levels. Appendix B includes a partial list of concepts in the terminology.

Data Representation Model

The section header terminology data-representation model supports conceptual-orientation, polyheirarchy, links to external vocabularies, and support for concept and term attributes. Each section concept is distinct and has a unique numerical "concept identifier" (CID) and unique string name (the "concept name") to which multiple strings may be mapped through unique "string identifiers" (SID) in a many-to-many fashion, much like the organization of the UMLS's SUIs (for strings) and CUIs (for concepts) identifiers. The unique concept name is composed without spaces. For example, the concept "physical_examination," whose CID is 545, is mapped to 34 strings, each with a unique SID (see Table 3).

Table 3: Partial List of String-Concept Mappings for the “physical_examination.”

CID	Concept Name	SID	String	String Type	String Source	Source ID
545	physical_examination	4715	physical examination	PT		
545	physical_examination	4256	PE	SS		
545	physical_examination	928	physical examination as compared to admission	PT		
545	physical_examination	1235	external examination	PT		
545	physical_examination	1521	physical exam compared admission	PT		
545	physical_examination	1758	physical exam	PT		
545	physical_examination	2099	physical exam as compared to admission	PT		
545	physical_examination	2114	exam	PT		
545	physical_examination	2331	My key findings of this patient's physical exam are	PT		
545	physical_examination	2557	admission physical exam	PT		
545	physical_examination	3117	examination on discharge	PT		
545	physical_examination	3369	examination on discharge compared to admission	PT		
545	physical_examination	3459	physical examination by organ systems	PT	LOINC	11384-5
545	physical_examination	3590	physical findings	PT	LOINC	29545-1
545	physical_examination	3873	examination	PT		

CID = concept identifier. SID = string identifier. PT=preferred term (usually from either a vocabulary input or a clinical note). SS = suppressible synonym.

Concepts are organized in a hierarchical structure with parent-child relationships. For instance, a “shoulder exam” is a child of “musculoskeletal exam,” and “Family and Social History” is composed of (via parent-child relationships) “Family Medical History” and “Social History.” Some concepts can have multiple parents. Each concept with multiple parents has a primary parent-child relationship and “alternate” relationship(s). The nearest regional anatomic parent is preferentially chosen as the primary relation, when multiple categorizations are possible. For example, “jugular_venous_pulse_exam” is a child of both “neck_exam” and “cardiovascular_exam”; its primary parent is taken by the above heuristic to be “neck_exam” since this is the closest anatomical “container”. In other cases, the author assigned relationships according to categorizations in textbooks or the medical literature (Figure 3). Finally, if no teleological assignment could be made, the author used the most typical parent concept as the primary relationship.

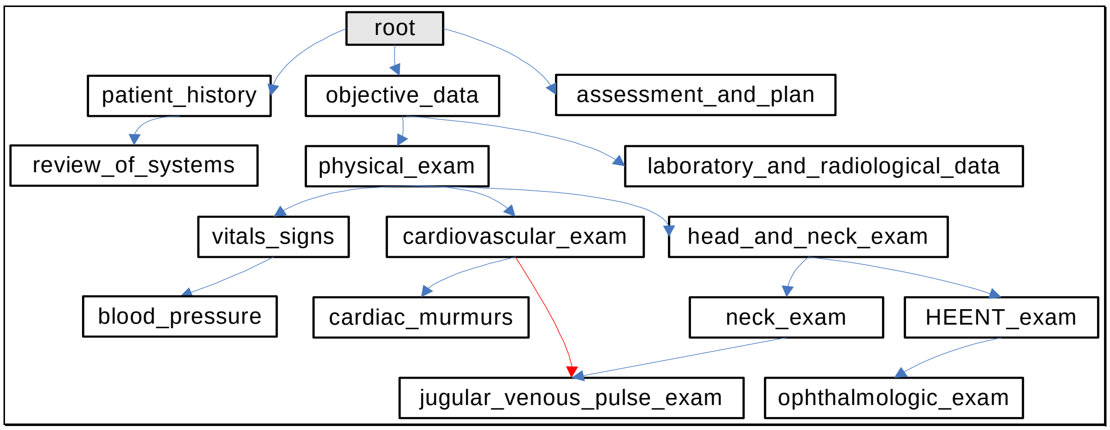

Figure 3: Partial diagram of the section terminology.
The red link from cardiovascular_exam to jugular_venous_pulse_exam is an "alternate" parent-child relationship, while the primary parent is neck_exam.

To support interaction with other, potentially existing, applications, the section header terminology data model retains source terminologies identifiers (such as LOINC IDs or UMLS CUIs) so that one could restrict section matching to those concepts belonging to a given external terminology or use external concept identifiers (e.g., LOINC IDs) instead of the section terminology's CIDs. Since a concept may reside in many vocabularies, the author mapped section concepts to concepts from other terminologies in a many-to-many fashion. Strings can also have a source vocabulary and identifier. Figure 4 represents the database schema.

Specifying certain attributes for each concept and string can improve section header concept matching. String attributes include a string type ("concept name", "preferred term," "suppressible synonym," "normalized," and "normalized without stop words"). The latter two strings are generated strings to speed matching. The "concept name" is the unique name of the concept in the terminology. Preferred terms are the bulk of string names, including those from other vocabularies, and include many abbreviations. Suppressible synonyms are strings that are sometimes found in documents but are relatively nonspecific. Suppressible synonyms include single letter strings (e.g., "A" for the concept "Assessment" or "T" for "Temperature"), common words that rarely represent specific sections (e.g., "patient" for "patient name", "time" and "date"), and some abbreviations (e.g., "PT" for "prothrombin time," which is often used to mean "patient" in the context of a document).

Concepts have several attributes: a type, a data type, and a "next section" attribute. Concept types can be either "atomic" or "composite." Composite concepts represent combinations of atomic concepts, such as "hematologic-lymphatic-oncologic." Very common primary groupings, such as "head, eyes, ears, nose, throat" are defined as "atomic concepts." The component concepts of composite concepts often overlap with other composite concepts or may cross hierarchies of atomic concepts. The concept data type can be "default/prose", "short", "date/time", "title", or "numeric." This refers to the type of information contained within the section. A "default/prose" section can be many lines or paragraphs. Finally, the "next section" attribute is a rule that defines what section typically follows a short, date, or numeric section. For example, the "next section" attribute for "chief compliant" is "History of Present Illness."

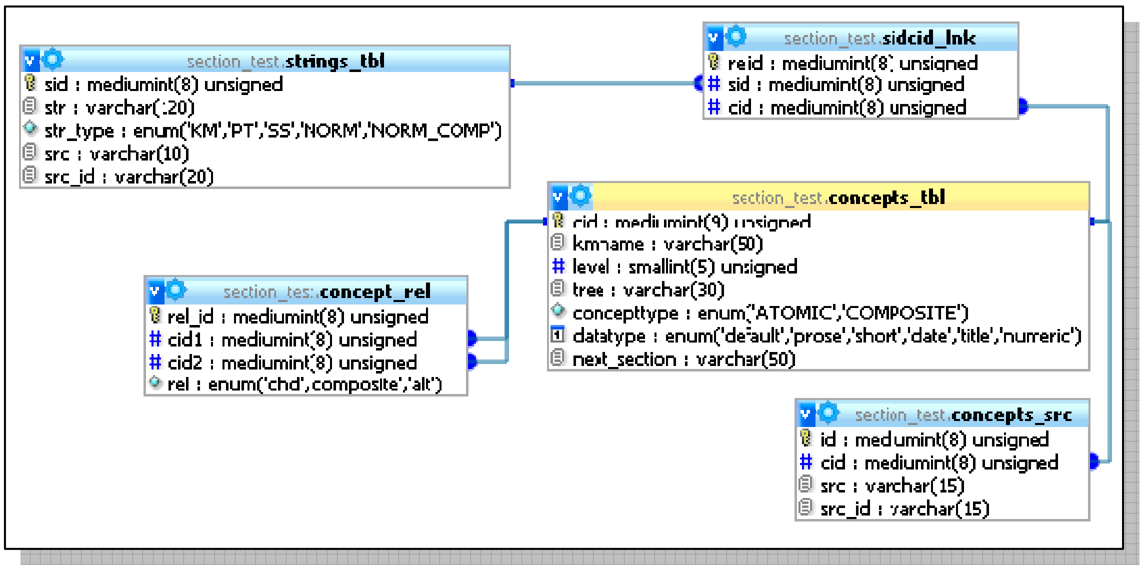

Figure 4: Database Model for the Section Terminology.

Medical Word List Creation for Spelling Correction

To create a list of medical terms for automatic spelling correction, the author used the strings from the UMLS Metathesaurus (2005AC version), KM Lexicon (which is based on the SPECIALIST lexicon), the open source Medical Words project[148], and words extracted from the templates of a clinical documentation tool. From the Metathesaurus, the author extracted all words from all English language strings, ignoring all strings that were not designated as suppressible synonyms. Since Aspell, the project's target spell-checker, does not allow words with non-alphabetical characters in them, the author excluded all words containing non-alphabetical characters. The author used a program to extract all non-recognized words from a total of 3,285 clinical note templates, representing pre-designed forms for physician documentation of outpatient encounters, inpatient notes, and consultation notes as well as a number of ancillary services. Three physicians reviewed all words in the templates that did not match a word from the Metathesaurus or KM Lexicon and were less than 4 characters long to determine validity. The resulting vocabulary contained a total of 366,613 words.

SecTag Development

Overview

The section tagging application, named SecTag, identifies concepts from the section header terminology in natural language clinical documents. It is designed to identify within clinical documents all document-labeled section headers and also to predict names and placement of section headers that are not explicitly labeled in the document but whose related content clearly appears in a given segment of the document. The SecTag output is a structured, XML-tagged version of the original note with identified section headers.

The SecTag algorithms process documents in five major steps (Figure 5): (a) identify sentences and lists (e.g., “1. Congestive heart failure”); (b) identify all candidate sections using lexical tools, spelling correction, and natural language processing techniques; (c) calculate the Bayesian probabilities that each sentence belongs to any given section; (d) determine the most likely section for each sentence, using the “exactly matched” sections to help disambiguate unclear sections; and (e) discard “bad” section matches. To develop and iteratively improve the algorithm, the author used the training set from the shared H&P note corpus. This corpus was also used to calculate the prior probabilities for the Bayesian predictor.

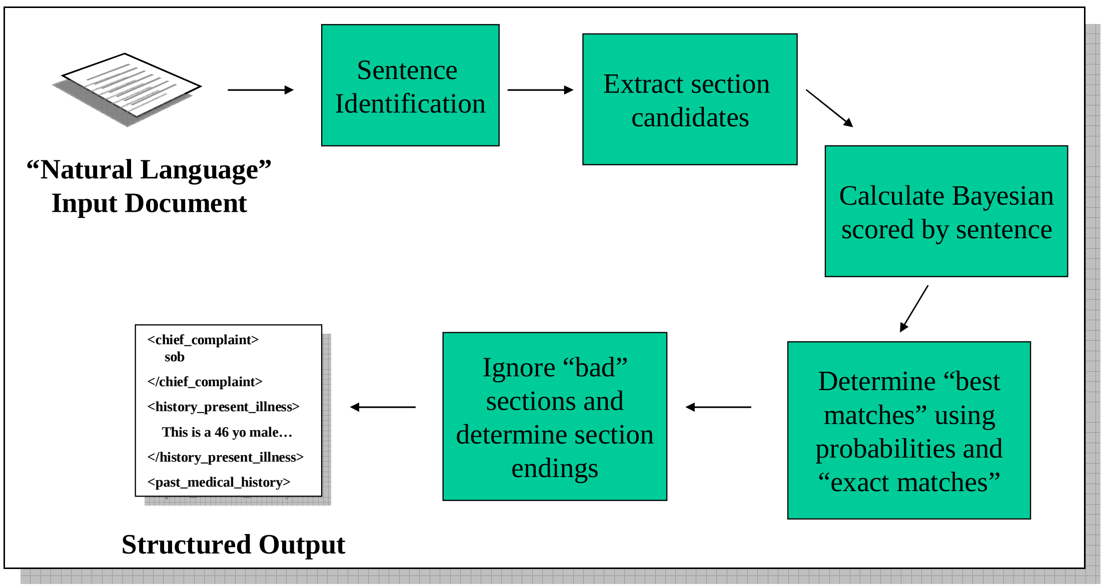

Figure 5: Flow Chart of KnowledgeMap Section Tagger application major steps.

Preprocessing of Documents

Clinical notes, often hastily generated by providers, are often not well-formatted documents, especially if typed directly by the provider or dictated without a template. They may lack proper capitalization at the beginning of sentences, having extra newline characters in the middle of a sentence, or have other types of formatting incongruences. Even dictated notes, which tend to have more consistent formatting since they are transcribed by commercial firms specializing in medical transcription, may be word-wrapped with newline characters. For these reasons, SecTag first identifies sentences before attempted to recognize section header tags. The algorithms that SecTag uses for this task are based loosely on those employed in the KnowledgeMap concept identifier (KMCI).[4] In addition, the SecTag algorithm separates out individual list items (e.g., "1. congestive heart failure… 2. hypertension") identified by dashes or numerical progressions. The document preprocessor also removes common headers and footers inserted into documents, such as page numbers and some transcription-inserted information in dictated notes (such as a transcriptionist's initials and document code numbers, transcription dates and times).

After completing sentence identification, SecTag next identifies groups of sentences that appear to be part of a list. Each block of sentences beginning with ordinals are labeled with a "list number" and the element of the list (in the above example, "CHF" and "hypertension" would be labeled as different elements of list *X*). This information aids in section identification: a group of sentences within a list element is likely to either belong all to the same section or, possibly, a subsection of that section. For example, if list *X* was preceded by "Assessment and Plan," all items in the list are likely "assessment_and_plan" elements or belong to a subsection of "assessment_and_plan," such as "cardiovascular_plan."

Identifying Candidate Sections and Subsections

After SecTag demarcates the sentences within a document, it then processes the document, sentence by sentence, to find all possible section headers therein. First, the algorithm attempts to identify any explicitly-labeled section tags in the document by searching for strings that begin lines (sentences) and are less than 55 characters, consist of only capital letters, or ends in a dash, colon, or period. To find section tags within in a "sentence," the algorithm looks for uppercase-nonuppercase word boundaries and recursively processes strings embedded in sentences ending in colons or dashes (e.g., identifying "temperature" from "He is in no acute distress, Temp: 97F"). Only strings matching terminology header entries are kept as possible tags. Only those lines that SecTag matches in their entirety, and begin and end in a colon or dash, and which directly match a terminology term or its synonym (and the match is not a "suppressible synonym," such as "patient" or "A") are marked as "exact matches." The matching process is improved by SecTag normalizing all words. The normalization process uses the KMCI normalization algorithm, which has as its base vocabulary an augmented version of the UMLS' SPECIALIST Lexicon. Like KMCI, the normalization algorithm also utilizes a list of common prefixes and suffix-translation rules. SectionTagger removes common "stop words" (such as prepositions, determinants, and pronouns), with the exception of the letter "A" when not occurring as part of a sentence.

To further increase lexical sensitivity, the system employs three techniques, in sequence, to match strings that do not directly match terminology entries:

- **Derivational and semantic variant generation**: Using tables and algorithms originally developed for KMCI, SecTag generates derivational (e.g., "intestine" → "intestinal") and semantic (e.g., "lungs" → "pulmonary") word variants for all words not in the terminology. The algorithm also generates all possible form variants using 156 suffix-based "form-rules," which allow interconversion from one lexical variant to another (such as "appendix" to "appendiceal"). This leads to a list of "alternate forms" for each word in the possible tag.

- **Spelling correction**: The author integrated the open-source spell checker Aspell into SecTag.[149] Aspell has widespread acceptance in the open-source community, functions in many platforms, and achieves good performance in medical spelling correction tasks.[150, 151] The Aspell medical word dictionary used for this project was derived from the UMLS, KnowledgeMap lexicon, and clinical note templates (see above). SecTag applies Aspell to words not recognized by its normalization routines. For words that appear misspelled, the Aspell's top ten suggested alternatives are added to the list of "alternate forms."

- **Modifier extraction**: SecTag removes certain modifiers, such as possessive words, numbers (written as either a word or a number), anatomical references (e.g., "right", "superior", "bilateral"), and other common words (e.g., "recent", "other") if doing so allows a section header terminology match.

After generation of all alternate forms, the SecTag creates the set union of all possible candidate section tags and scores each candidate tag. Candidate sections tags receive a point for each exactly-matched normalized word, 0.95 points for spelling-corrected words and derivational forms, 0.9 for semantically-related forms, and 0.85 for variants generated with form-rules. Candidate tags' scores are penalized for extra words in target phrases. Candidate terminology matches with scores greater than 80% of the maximum possible score are kept as possible candidate tags for the document currently being analyzed.

Noun phrase processing to identify unlabeled sections. Finally, SecTag employs some natural language processing schemes to detect section tags within a sentence, such as "He is here for a chief complaint of SOB." Before normalization, SecTag identifies parts-of-speech using the same rule-based library from Cogilex, Inc that KMCI uses.[152] SecTag looks for strings matching section headers in all noun phrases of sentence fragments and sentences with linking verbs. During formative testing throughout the application's development, the author found that most section tags found in predicates of action verbs were false positives identifications (e.g., "EKG" in the sentence "We will order an EKG" is not a section header for EKG results). Thus, to favor specificity, SecTag does not process predicates of action verbs for section tags. When identifying noun phrases, the system considers numbers as parts of noun phrases but keeps numerical ranges (joined by a dash or preposition) grouped together. Thus, SecTag finds "cranial nerve" from the phrase "cranial nerve 2 through 12 were within normal limits" but finds "cranial nerve 11" from the phrase "cranial nerve 11 is intact." The system also ignores noun phrases matching to suppressible synonyms (such as “patient” or “date” or single letters such as “T”, which can mean “temperature”). After identification of candidate noun phrases, SecTag employs the same normalization and variant generation techniques described above if necessary.

Adding common ancestor concepts to possible section tags. If multiple matched tag concepts exist for a given document entry after the above steps, SecTag looks to see if there is a parent concept that can explain all of them. For example, for the sentence “Mother and father both had heart disease,” the system would pull concepts “mother_medical_history” and “father_medical_history” (among others, depending on the location in the document). In this step, it would add “parent_medical_history” as the closest common ancestor for both “mother_medical_history” and “father_medical_history” (Figure 6). Since nearly all sections could be related by some common ancestor, SecTag restricts possible common ancestors such that the cumulative length of the path connecting the leaf nodes (through a common ancestor) is less than 4, a distance chosen because it allows SecTag to span most common subtrees, such as related physical exam elements, but it will not tie together elements across very different subtrees, such as a medication list and physical exam components. Thus, in Figure 8, “substance_use” is a common ancestor for “tobacco_use” and “ethanol_use,” but there is no common ancestor for “tobacco_use” and “cardiovascular_review” since the cumulative path length between “tobacco_use” and “cardiovascular_review” (through “patient_history”) is 5.

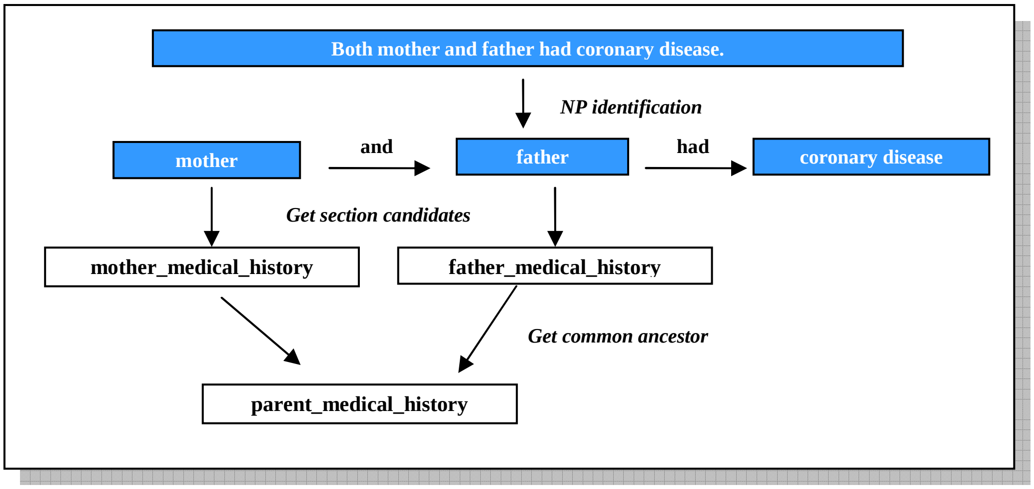

Figure 6: Extraction of Possible Section Candidates Using Noun-Phrase Processing and Common Ancestors.
In this example, many possible section candidates for "mother" and "father" would be extracted and kept as possibilities until disambiguation, along with the new section "parent_medical_history."

Naïve Bayes Section Predictions

Overview. The SecTag application uses naïve Bayesian scores to predict unlabeled section headers, disambiguate among possible section header candidates for a given segment of text, and as a "figure of merit" score that is used to discard poor section header matches. Bayesian scores for all section headers candidates are calculate for each sentence in a document. The prior probabilities for the section headers presence and the probability that a given word occurs in each section were calculated using the frequency of occurrences of given words in each section in the training set.

Building word and section prior probabilities. Calculating the Bayesian prior probabilities between document words and the section headers under which they commonly appear involves two successive applications of SecTag to the training set documents. The first application of SecTag does not use frequency-based scoring techniques or Bayesian prediction to disambiguate among vague, partially matched candidate section tags or to predict section headers for unlabeled sections. During this initial step, the system stores the "intra-document" location of each section in each document and its frequency of appearance in that location. The "intra-document location" is, roughly speaking, the decile segment of the document, taken from beginning to end, in which the candidate matched tag occurs. Pragmatically, this is calculated using the ordinal sentence number of the sentence containing the tag of interest divided by the total number of sentences in the document, rounded to the nearest tenth (i.e., 0.1). It is a number between 0 and 1. When processing each training set document, SecTag stores location probabilities for a section for each sentence in the section, thus treating short sections differently than long ones.

During SecTag's second iteration processing the training set, it uses frequency and location scoring information (but not Bayesian scores) to predict exact matches for previously ambiguous tags and to help to ignore bad matches (i.e., to recognize that a candidate tag that might match "history of present illness" but which appears at the end of the H&P document is not likely to be that section header). Having completed this pass of disambiguation for all training set documents, the system updates the frequency and location information for all sections based on the newly disambiguated results. SectionTagger also records data on relative patterns of section ordering (the proportion of times that section *X* occurs before section *Y*) and records the frequency with which individual words (or their stems) appear in each section. When sections are nested (e.g., a cardiac murmurs exam within a cardiovascular exam within a physical exam), the words count for all sections to which they belong. To improve sensitivity, the words are first normalized (e.g., changing “was” to “be”) and then stemmed using Porter stemming.[153] In addition, SectionTagger aggregates certain types of strings by their type:

- Roman numerals
- Integers
- Floating point numbers
- Dates and times
- Numbers written with a maximum range (e.g., “strength 3/5”, “III/VI systolic murmur”, or “III out of VI”).
- Weights (any number followed by a unit of weight)
- Lengths (any number followed by a unit of length)
- Single and double quotation marks. These are more likely to occur in specific sections, such as the chief complaint or history of present illness.

- Presence of list elements (e.g., “1.”, “2.”, “I”, “II”) in the section, which are identified by the document preprocessor.

The author reviewed the 60 most common words and then manually selected words to ignore in Bayesian scoring (i.e., “stop words”). The latter was accomplished by looking at each word’s individual section predictive value; for example, integers, the most common entry, are highly predictive for the physical exam and laboratory data and thus were retained as a meaningful word (i.e., not classified as a stop word) while “be” and “has” were discarded as stop words since they were not usefully predictive of section headers.

Word-based Bayesian prediction. When processing new documents, SecTag calculates Bayesian probabilities for the sentence’s section tag sentence-by-sentence using the prior probabilities calculated from the training set of H&Ps. The system uses a naïve Bayes algorithm, assuming conditional independence, which markedly simplifies computational complexity. Although this assumption is incorrect for clinical notes (the presence of one word (e.g., “systolic”) influences the presence of words following it (e.g., “murmur”), prior research has shown that this assumption does not significantly affect performance in practice.[154] The naïve Bayes algorithm is computationally tractable over many classes (in this case, more than 1,000 different sections are possible). It is relatively easy to implement, fast to calculate, and allows the potential for real-time updating as new documents are processed.

Bayes formula for probability of a section header being present in a document, given a set of words from a candidate document segment (e.g., from a sentence), is:

$$P(Section_i \mid \overline{words}) = \frac{P(Section_i) \cdot P(\overline{words} \mid Section_i)}{P(\overline{words})}$$

where P(Section) is the probability that the given candidate document segment occurs in documents at that location. Using the assumption of conditional independence (for naïve Bayes) reduces calculation complexity to linear with respect to the vector words by approximating of P(words|Section) and P(words) to yield the following:

$$P(Section_i \mid \overline{words}) = P(Section_i) \cdot \prod_{j=1}^{|\overline{words}|} \frac{P(word_j \mid Section_i)}{P(word_j)}$$

Since the probability of a word occurring in a given section $P(word_j|Section_i)/P(word_i)$ is difficult to calculate for text copra, investigators often estimate it using the "m-estimate," where each word is assumed an equal prior probability (estimated by 1/|Vocabulary|) and the size of the class (i.e., $Section_i$) is estimated by the total number of words in all instances of $Section_i$.[88, 155] Using the m-estimate yields the following:

$$P(Section_i \mid \overline{words}) = P(Section_i) \cdot \prod_{j=1}^{|\overline{words}|} \frac{n_j + 1}{n_{Section_i} + |Vocabulary|}$$

where $n_{sectioni}$ is the total number of words in the all occurrences of the section and n_j is the count of $word_j$ in $section_i$. This estimate has been used for classifying high quality Medline articles,[65] email classification (spam vs. not),[90] and classifying news articles based on interest.[85] It is a good estimate for the task of whole-document classification (where the entire document belong to a single class), since the size of the document is related to the prior probability of the document. However, since many section headers occur within a document (thus the document correctly "belongs" to many "classes"), and the boundaries of each section in the document are initially unknown to SecTag, the system must make predictions on the subdocument level. The author designed SecTag to calculate probabilities for each sentence in the document. Using the number of words in the entire section corpus ($n_{section}$) overly biases toward small sections. Thus, the author determined by experimentation that scaling $n_{section}$ by taking its square root yielded the best results. (This converts the probability into a score rather than a true probability, since the value may exceed 1.)

Section-based Bayesian Prediction. Following calculation of the Bayesian score for each sentence based on its words, SecTag then adjusts the probability of each section based on the exact-matched sections (i.e., sections labeled in the document with all capital letters or started a sentence and ended with a colon or a dash) that have occurred before it. For example, if "history_present_illness" has already occurred as an exact-matched section, the probability that a following section could be "chief_complaint" is diminished. This calculation includes subsections as well; "vital_signs" will tend to follow "physical_examination" but not "assessment_and_plan." Combining this probability with the Bayesian section-word score, the final calculation is:

$$P(Section_i \mid \overline{words}) = P(Section_i) \cdot \prod_{k=1}^{m} P(Section_i \mid priorSection_k) \cdot \prod_{j=1}^{|\overline{words}|} \frac{n_j + 1}{\sqrt{n_{Section_i}}}$$

where $Section_k$ represents the exact-matched sections preceding $Section_i$.

Predicting sections. When processing a given document, the SecTag Bayesian calculation ranks all 1,109 theoretically possible section headers. Only the four topmost-ranked section headers are considered when attempting to predict a label for a given document section. If no section is active (i.e., an identified and confirmed section that is potentially applicable for the current sentence), SecTag will make the best-ranked section active if it is much more likely than next-best section candidate. If there is an active section, the system uses a number of rules to decide whether to keep the active section, add a child section, or terminate the section (see below). Since all sections are ranked, the top several matches are often related to each other via parent-child relationships. The author found that the best-match child section of an active parent is often a valid match, even if the parent section had a slightly better rank; this is partially due to counting words for both the section and subsections when building the probability tables. Thus, if a child concept of an active section is in the top 4 and the score differential is small, SecTag marks the child concept as being present (the parent remains active as well such that the sentence belongs to both the child and the parent section). An example of this is seen with a sentence like "no murmurs, rubs, or gallops." If this phrase appeared in a document segment that SectionTagger determined should be labeled as "physical_examination," SectionTagger would add the label "cardiovascular_exam" to encapsulate this phrase (because of its Bayesian score), in addition to retaining the (parent) header "physical_examination." For lists, if the list is enclosed within a section with an identified header, only subsection headers for that particular header are considered. Otherwise, the system favors grouping all sentences of a list element within the same tree of section concepts.

Disambiguation of Ambiguous Section Labels

Many sections strings can map to multiple different concepts; for example, “cardiovascular” can refer to either “cardiovascular_exam,” “cardiovascular_plan,” “cardiovascular_system_review,” “cardiovascular_hospital_course,” or “cardiovascular_family_history.” SecTag uses several techniques to disambiguate these possibilities. First, while extracting possible section tags, it keeps track of the “active” sections at each level, and checks ambiguous tags with all currently “active” tags. Thus, if the active section was “physical_examination” at tree level 2, “cardiovascular” would be interpreted as “cardiovascular_exam” since it is a child concept in the tree of “physical_examination,” and “cardiovascular_exam” would be added to the active section list at level 3. If the next section encountered was labeled “abdominal”, the system would first check to see if it was a child concept in the tree of “cardiovascular_exam”, and then replace it with “abdominal_exam” since it is a child of “physical_examination.” If the next section did not exist as a child concept of any of the active sections in the list (e.g., “electrocardiogram”), SecTag terminates all active sections headers at the sentence before the new section, and then would proceed to disambiguate the possible candidate tags for the new section (in this case, “electrocardiogram” is a unique section header).

SecTag then processes the document using each identified tag in it, scanning for section endings and determining if a section should be ignored. The system can be used with or without naïve Bayesian probabilities. The individual word and section probabilities for the Bayesian scoring methods were determined by automated tagging of the training set.

Discarding poor matches. Since SecTag's candidate section header identification process is designed to be sensitive, not specific, many of the "possible" matches should actually be ignored (see Figure 7). SectionTagger considers all non-exact matches as possibly incorrect. As previously noted, candidate tags that are determined to be child concepts of active section tags at that position in the document are retained.

Conversely, SectionTagger discards a non-child, non-exact candidate match if:

- The candidate tag was predicted (via NLP or other means) but no sentences are assigned to the section (i.e., it is an "empy section"). For example, the word "patient," possible representing a "patient name" section header, may appear on a line by itself but instead be the last word of a sentence.

- The Bayesian rank of the best candidate section is lower than 4^{th}.

- The sentence is part of a list and is not a child of the section concept that begins the list.

- The section was designated by a space and found within a sentence.

- It is a predicted match and an exact match exists elsewhere in the document.

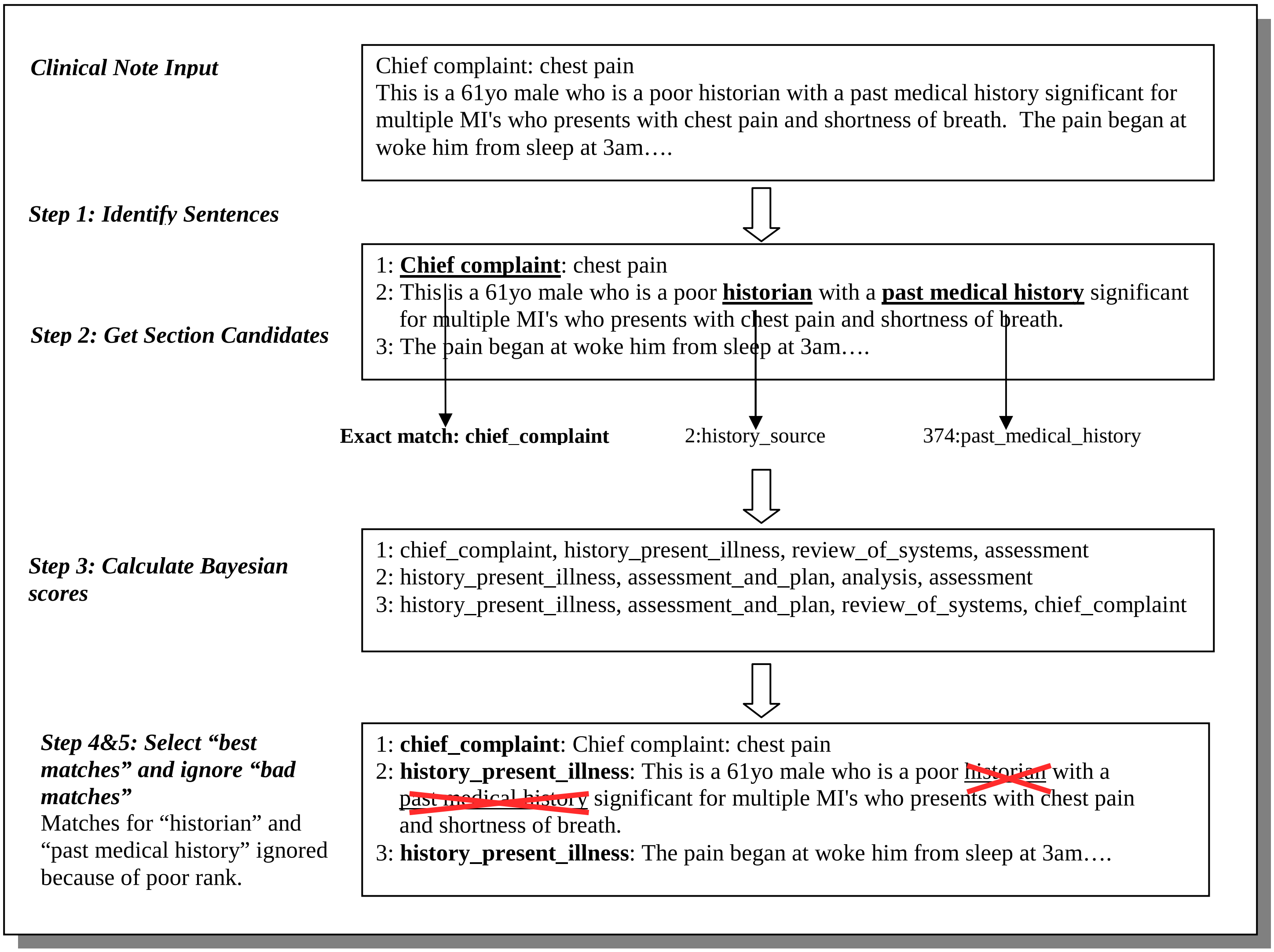

Figure 7: SecTag Processing of a Section of Text.
The "Chief complaint" tag is an exact match since it starts a line, matches a string exactly, and ends in a colon. The Bayesian score of the next sentence highly favors the "history_present_illness" because it follows the chief complaint, occurs toward the beginning of the document, and contains words common for this section. Thus, "historian" and "past_medical_history" are ignored as possible tags and the section is labeled as "history_present_illness."

Scoring ambiguous matches. SecTag scores non-ignored candidate sections using attributes and frequency metrics that are summed together. The attributes, all with a maximum score of 1, include the presence of children concepts, whether it could be child of a prior section, and the normalized level of concept in the tree (favoring those concepts that occur higher in the hierarchy). Sections appearing as exact matches in other locations in the document are penalized. SecTag also adds to this metric the string matching method: those concepts that matched via an exact "preferred term" match receive a score of 1, while normalized string matches or filtered matches receive lower scores. It then calculates the path length (see Figure 8) in the tree between each candidate concept and the nearby section concepts before and after it (excluding ambiguous concepts). It does this until it reaches the end of a "block" of sentences (as determined by whitespace or list elements) or a level 1 or higher header (so chosen as these are typically major sections in the terminology, such as "physical_examination" or "history_present_illness"). The probability of the section appearing in that position in the document, and the probability that the section would appear in the document are both included in the score. For the evaluation, the Bayesian score of each of the candidates is also added into the score but weighted twice the value of the other components.

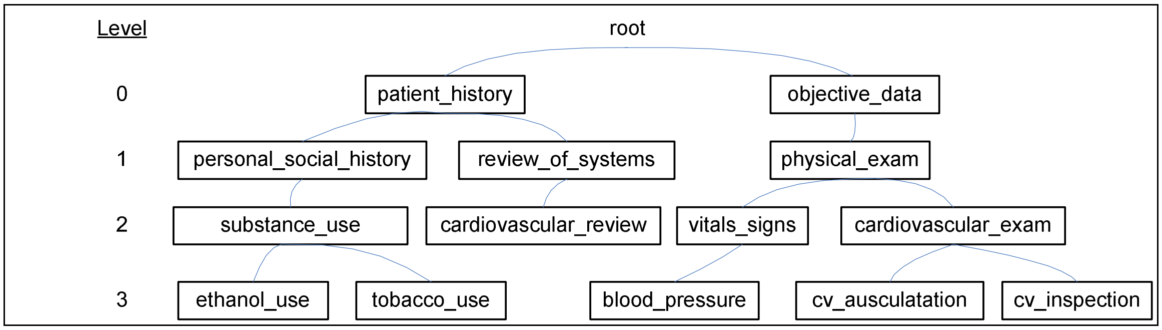

Figure 8: Diagram of fragment of the section terminology.
The path length between any nodes is the distance by counting edges between the two nodes, with the exception that edges connecting to “root” are weighted 5 as it connect mostly unrelated trees. The path length between vital_signs and cardiovascular_exam is 2; the path length between cardiovascular_review and cardiovascular_exam is 14.

Determining Section Endings

Active sections are terminated through a combination of rule-based and probabilities, if Bayesian scores are used. An active section is terminated when SecTag encounters the next “valid” section that is not a child. Rule-based termination also acts on active sections that are numerical, date/time, short, or document title data types. These sections are terminated if:

- The next line is blank

- There is already text in the section, and the next line begins with a capital letter, and there is text before the next section. For example, in a string like “Chief complaint: wheezing<newline>He presents with…”, this rule allows the sentence beginning “He presents with” to be considered a new section (i.e., the history of present illness).

- If section is a date/time format, text following a matching date/time is not included and is eligible to become another section. For example, SecTag terminates the date section of "Date: January 5, 2004 3:00pm Attending physician…" after "3:00pm."

- If the section is a document title (e.g., "Inpatient History and Physical"), the section is closed immediately. (Document titles are defined as zero-length sections in the terminology.)

Another rule-based termination operates on sections occurring within a list. If a prior list item contains an active section when SecTag encounters the next list item, the prior active section is terminated. For example, if a list item began "2. Cardiovascular:" and then describes the cardiovascular plan, the application would terminate the "cardiovascular_plan" section once it encountered the next list item ("3") unless the Bayesian score of that text predicted "cardiovascular_plan." In addition to the rule-based termination, the Bayesian probabilities are used to determine section endings. These proceed differently for sections that are predicted versus those that are author-labeled. For predicted sections, the system favors specificity by ending the active section once its rank falls out of the top four or the gap between the a better-ranked parent section and the predicted section widens. SecTag terminates document-labeled or NLP-derived sections if: 1) A nested subsection of another active subsection at a higher level has a better Bayesian score; for example, if the current active sections are "physical_exam → neurologic_exam → cranial_nerve_exam" and the Bayesian probability for neurologic_exam was greater than that of cranial_nerve_exam, SecTag will end the cranial_nerve_exam section and revert back to neurologic_exam. 2) There is a large score differential between the current active section and the best predicted match, which is not a subsection, and the current active section is not within a list. Finally, after finishing processing all sentences, SecTag terminates any remaining active sections at the end of the document. The output is an XML-tagged document, divided into sections with section identifiers and concept names. The system can also produce multiple HTML outputs, such as the evaluation output seen in Figure 9.

CHAPTER IV

EVALUATION

Overview

In the previous chapter, the author described the SecTag application, a tool for identifying section headers in clinical notes. This tool includes a terminology and an algorithm to categorize notes sections based on that terminology, accounting for synonyms, misspellings, and sections that are not explicitly labeled in the document. The primary goals of the presently described evaluation were to 1) assess SecTag's recall and precision in identifying H&P section headers that are labeled in the document and 2) assess the system's recall and precision in identifying all major sections, whether the document contained an explicit section header label or not. The evaluation subjects were board-eligible physicians unfamiliar with SecTag who evaluated how well the system labeled segments of randomly selected H&Ps from the project's evaluation H&P set (Figure 2), which had been themselves randomly selected from the electronic medical record system. The previously mentioned IRB review and approval covered this evaluation study.

Evaluation Methods

First, an automated program fed all H&P notes from the evaluation set into SecTag to create an identified set of documents. Another computer program randomly divided the evaluation set into groups of about 100 documents, which were randomly assigned to the project's physician reviewers to evaluate. All reviewers were board-eligible internal medicine or pediatric clinicians; some had additional subspecialty training. None were familiar with the section tagging software or the section header terminology prior to recruitment. They were compensated for their time.

The author created a web interface to present the documents to evaluators. On a single display screen, SecTag-identified sections and subsections were highlighted with different colors (see Figure 9) on the left half of the display, and the original document, in the original format, was presented on the right half. The evaluators were given a hierarchical list of all section tags that were found in the training document set, formatted such that more common sections appeared larger than less common ones (see Figure 10). All evaluators began work with the same five documents initially, representing a "standardized" training set. Afterwards, during the evaluation, every eighth document was taken from the common document set that all evaluators reviewed, in order to judge inter-rater agreement.

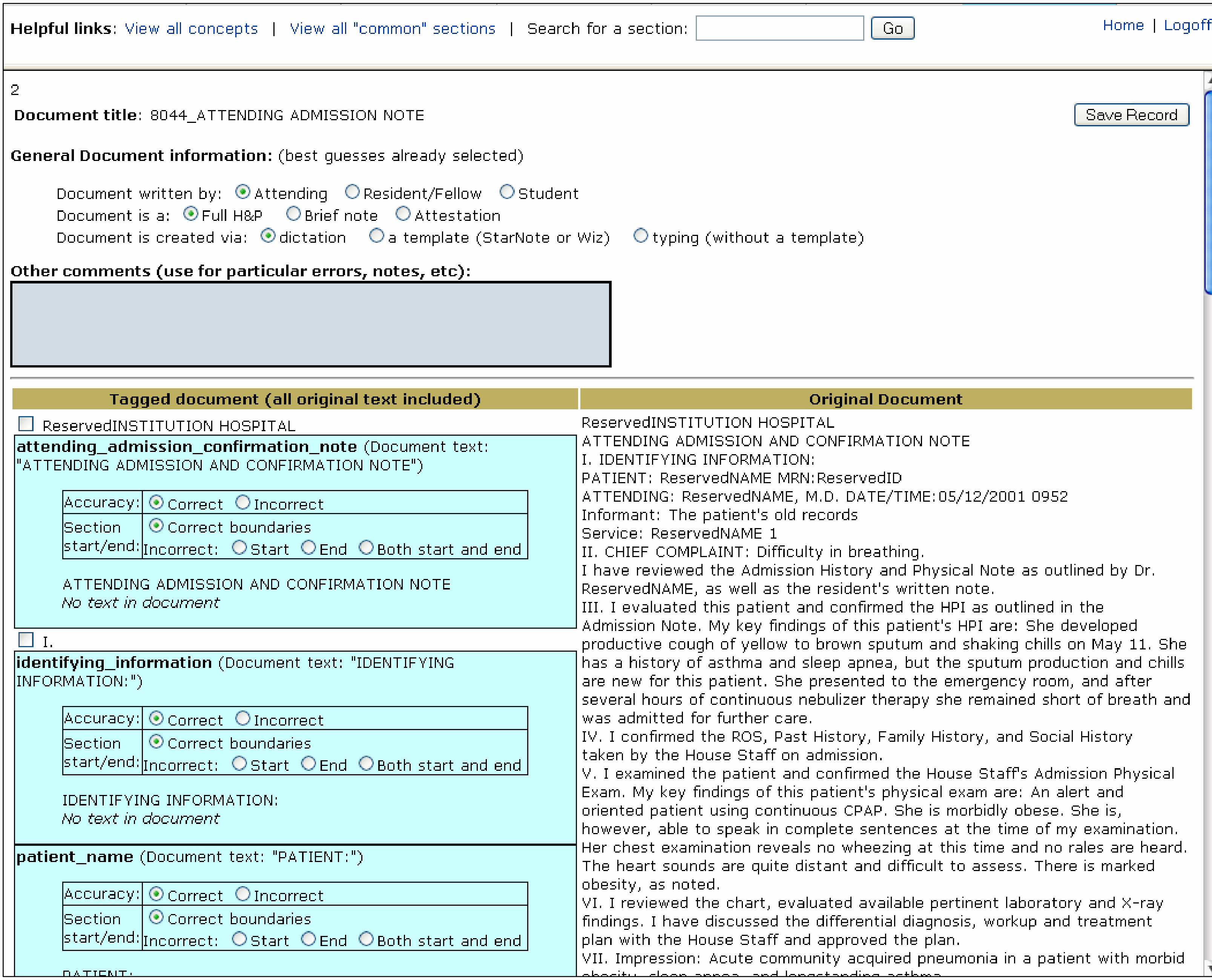

Figure 9: HTML Scoring Interface.

chief_complaint (Tree:5.26, Frequency: 0.7601)
reason_for_consult (Tree:5.27, Frequency: 0.0007)
history_present_illness (Tree:5.28, Frequency: 0.9225)
subjective (Tree:5.28.57, Frequency: 0.0032)
brief_history (Tree:5.28.59, Frequency: 0.0058)
other_issues (Tree:5.28.61, Frequency: 0.0073)
reason_for_study (Tree:5.29, Frequency: 0.0002)
risk_factors (Tree:5.30, Frequency: 0.0083)
cardiac_risk_factors (Tree:5.30.64, Frequency: 0.0185)
cerebral_vascular_risk_factors (Tree:5.30.68.62, Frequency: 0.0002)
home_risk_factors (Tree:5.30.73.64, Frequency: 0.0007)
work_risk_factors (Tree:5.30.73.65, Frequency: 0.0005)
changes_to_admission_note (Tree:5.31, Frequency: 0.0002)
hospital_course (Tree:5.32, Frequency: 0.0049)
hospital_course_by_system (Tree:5.32.77, Frequency: 0.0002)
general_course (Tree:5.32.77.70, Frequency: 0.0003)
derm_course (Tree:5.32.77.71, Frequency: 0.0005)
heent_course (Tree:5.32.77.72, Frequency: 0.0005)
neck_course (Tree:5.32.77.72.12.4, Frequency: 0.0020)
ent_course (Tree:5.32.77.72.13, Frequency: 0.0005)
ophthalmologic_course (Tree:5.32.77.72.14, Frequency: 0.0007)
lymphatic_course (Tree:5.32.77.73, Frequency: 0.0010)
cardiovascular_course (Tree:5.32.77.75, Frequency: 0.0015)
gastrointestinal_course (Tree:5.32.77.78, Frequency: 0.0003)
genitourinary_course (Tree:5.32.77.79, Frequency: 0.0003)
musculoskeletal_course (Tree:5.32.77.81.23, Frequency: 0.0020)
arm_course (Tree:5.32.77.81.23.14.16, Frequency: 0.0012)
endocrine_course (Tree:5.32.77.84.25, Frequency: 0.0002)
pulmonary_course (Tree:5.32.77.87, Frequency: 0.0012)
fluid_electrolyte_nutrition_course (Tree:5.32.77.88, Frequency: 0.0002)
fluid_course (Tree:5.32.77.88.27, Frequency: 0.0008)
electrolyte_course (Tree:5.32.77.88.28, Frequency: 0.0002)
transport_history (Tree:5.33, Frequency: 0.0002)
family_and_social_history (Tree:5.34, Frequency: 0.0549)
personal_and_social_history (Tree:5.34.78, Frequency: 0.7317)
education_history (Tree:5.34.78.90, Frequency: 0.0056)

Figure 10: Common Matching H&P Sections.
A printed form of all matching sections was given to all physician reviewer. Frequency data is from the training set of H&Ps.

When scoring documents, the author instructed the reviewers to use the document author's label as a guide. For instance, if the document included the section label "Head, Eyes, Ears, Nose, Throat" but then the document's author provided only an ear exam in that section, the reviewers were instructed to mark either the broader ("HEENT exam") or more specific ("ear exam") concept. Likewise, a single sentence in the "History of Present Illness" could be attributed to another section (even though not labeled as such) if appropriate; for example, "He has a 40pack-year history of smoking" can be accurately tagged as "tobacco use" even if the author had placed it in the history of present illness section.

The web-based scoring interface asked reviewers to mark each SecTag-identified section as correct or incorrect and whether the SecTag-identified section boundaries were correct or had an incorrect starting boundary, ending boundary, or both. For SecTag errors, reviewers could indicate whether the system labeled a section where there was none, state that it had mislabeled the section and then explain what the evaluators themselves thought a better section tag would be, and/or indicate whether the SecTag provided section tag was too specific, too general, or an erroneous homonym (e.g., SecTag selected "pulmonary_review" instead of "pulmonary_exam" for a document label "pulmonary").

To maximize the sensitivity of evaluators for finding false-negatives, the author provided evaluators with a list of clinically-important "major" section headers and subsection headers (see Appendix C). The author asked reviewers to identify any "major" sections from this list not identified by SecTag (even if not labeled by the document's author).

Evaluators labeled the missing section using a checkbox next to each sentence that allowed them to add the missing information. Finally, the author asked the reviewers to mark any sections labeled by the authors that were not tagged by SecTag. For the purposes of this study, the author defined a section label as any fragment of text on line by itself or followed by a colon. However, the author instructed them to identify only clinically-relevant labels that were not a finding ("back pain") or diagnosis (e.g., "myocardial infarction").

The author manually interacted with evaluators as they each scored the first three common documents of the training set. This was done to encourage each reviewer to conform to a consensus methodology during the evaluation. The training documents included a brief admission note with unlabeled major sections, a dictated but poorly formatted attending attestation note with few sections labeled by the document's author (an attending physician), and a comprehensive resident H&P note containing many explicit section labels.

Finally, evaluators categorized documents as to whether they appeared to be from an attending physician, a housestaff physician, or a medical student. Evaluators classified documents as being one of: "full history and physical," a "brief admission note," or an "attestation" of a housestaff note. If a note matched more than one category (such as a resident note appended with an attending attestation), the author instructed reviewers to classify it by the majority author and most complete appropriate classification (full history physical > attestation > brief admission note).

SecTag Evaluation Measurements

The primary outcome was recall and precision on all "major" sections. The evaluation also calculated overall recall and precision on all identified sections. Study definitions included: (1) MAJOR TRUE POSITIVE (MTP): a correctly-identified section that is a major section (e.g., "History of Present Illness"); (2) TRUE POSITIVE (TP): any correctly-identified section; (3) MAJOR FALSE POSITIVE (MFP): a misidentified section that is a major section (e.g., selecting the section "pulmonary_exam" when the section header actually was something else); (4) FALSE POSITIVE (FP): any misidentified section given a label by SecTag; (5) OMITTED SECTIONS: those sections identified by evaluators but not tagged by SecTag; (6) MAJOR SECTION RECALL: the number of MTPs divided by the total number of major sections (determined by adding the omitted major sections + MTP); (7) MAJOR SECTION PRECISION: the number of MTPs divided by total meaningful term attempts (MTP/[MTP + MFP]); (8) OVERALL RECALL: TP/(TP + omitted sections); (9) OVERALL PRECISION: the ratio of the number of correctly matched sections to the total number of identified (i.e., proposed) sections (TP/ [TP + FP]).

SecTag Evaluation Statistical Analyses and Sample Size calculation

The author examined a large database of tagged notes using a simpler version of the section terminology. In this setting, the author found using an automated program that there were approximately 18 sections per document for over 22,444 notes in a medical school documentation database. To determine with 90% power an accuracy of 0.9 within a 95% confidence interval of 0.05 would require 471 sections (or about 26 documents).

To assess SecTag's ability to find certain major sections, a larger sample size is needed. Determining a major section within a 10% confidence interval with 90% power requires 137 instances of the section. If each major section appears within at least 70% of the documents, then approximately 195 documents are needed.

The evaluation calculated interrater agreement via Cohen's Kappa. The author used Wilcoxon rank sum test to compare nonparametric data (expressed as median and interquartile range). The author used Student's t-test for parametric data (expressed as mean ± standard deviation) after verifying the data had a normal distribution. Distribution data was compared via the χ^2 statistic. Confidence intervals were calculated using binomial exact method for ranges approaching 100%. All statistical analyses were performed with Stata, version 9.2 (StataCorp LP, College Station, TX).

CHAPTER V

EVALUATION RESULTS

SecTag Recall and Precision

The physician-evaluators scored a total of 319 individual documents; 308 were scored by a single reviewer and 11 were common documents. Of the 319 documents, 66 were written by attendings and 252 by housestaff; only one note was classified as a medical student note. Reviewers classified 88% of the notes as full H&Ps, 6% as attending attestations, and 6% as brief admission notes. These notes contained a total of 16,036 sections (median 52 sections/document). There were 355 unique sections identified in the evaluation set. Table 4 shows the comparison of the evaluation and training set. The two groups of documents were similar except for the distribution of document titles and the frequency of sections predicted through natural language processing. This difference is largely due to one document type (a dictated attending attestation) that was more common in the training than the evaluation set. When comparing the evaluation set to the set of all H&P titles, there was no statistical different among the top 5 document frequencies.

Table 4: Comparison of Training and Evaluation Sets.

Numbers in parenthesis are the 25%-75% interquartile range for each.

	Training Set	Evaluation Set	p
Number of documents	5881	319	
Unique document titles	78	25	<0.01*
Median word length (IQR)	825 (541-1090)	875 (575-1117)	0.18
Median section count (IQR)	54 (32-74)	52 (31-70)	0.11
Labeled	34 (16-52)	34 (18-52)	0.18
Bayes predicted	5 (3-8)	6 (3-8)	0.60
NLP-predicted	12 (7-16)	7 (0-10)	<0.01

*The top five note categories were in the same order in each group and roughly similar percentages, accounting for more than 75% of the total notes.

Table 5 shows the overall recall and precision for all section concepts across all documents with and without labels present in the document. SecTag was more effective identifying labeled than unlabeled sections (p<0.001). Table 6 shows the recall and precision for major sections. Recall was slightly better on non-major sections than major sections and precision slightly better on major sections, though this difference is not likely clinically significant (recall 99.3% vs. 98.6%, p<0.001; precision 95.0% vs. 96.2%, p<0.001). Since reviewers did not identify unlabeled sections that were not major sections (e.g., they would not identify an untagged medical record number that was not labeled in the document as an error since it is not a major section), the difference in the overall recall and the recall for major sections is likely exaggerated.

Table 5: Overall Recall and Precision.

	Label in document	No label in document	Total
Number tagged correctly	11,353	3,976	**15,329**
Number tagged incorrectly	103	604	**707**
Number where SecTag omitted correct tag	20	140	**160**
Recall	**99.8%** (99.7 – 99.9)	**96.6%** (96.0 – 97.2)	**99.0%** (98.8 – 99.1)
Precision	**99.1%** (98.9 – 99.3)	**86.8%** (85.8 – 87.8)	**95.6%** (95.3 – 95.9)

Table 6: Major Section Recall and Precision.

	Label in document	No label in document	Total
Number tagged correctly	6,250	1,310	**7,560**
Number tagged incorrectly	71	227	**298**
Number where SecTag omitted correct tag	4	107	**111**
Recall	**99.9%** (99.9 – 1.00)	**92.4%** (91.1 – 93.8)	**98.6%** (98.3 – 98.8)
Precision	**98.9%** (98.6 – 99.1)	**85.2%** (83.5 – 87.0)	**96.2%** (95.8 – 96.6)

Table 7: Recall and Precision on Each Major Sections.

Section name	N	Num labeled (%)		Recall		Precision	
Chief complaint	283	280	(99%)	100%	(98 - 100)	100%	(98 - 100)
History present illness	353	281	(80%)	99%	(98 - 100)	93%	(90 - 96)
Past medical history	296	282	(95%)	99%	(98 - 100)	99%	(97 - 100)
Family medical history	255	250	(98%)	100%	(99 - 100)	98%	(95 - 99)
Parental medical history	192	7	(4%)	95%	(90 - 0.98)	90%	(84 - 94)
Sibling medical history	38	3	(8%)	87%	(71 - 95)	97%	(84 - 100)
Children medical history	3	1	(33%)	100%	(29 - 100)	100%	(29 - 100)
Health maintenance	92	91	(99%)	100%	(96 - 100)	100%	(96 - 100)
Personal and social history	267	252	(94%)	100%	(98 - 100)	99%	(96 - 100)
Substance use	254	138	(54%)	94%	(91 - 97)	98%	(96 - 100)
Medications	282	254	(90%)	100%	(98 - 100)	99%	(96 - 100)
Allergies and adverse reactions	254	249	(98%)	100%	(99 - 100)	100%	(99 - 100)
Review of Systems	462	437	(95%)	100%	(98 - 100)	95%	(92 - 97)
Physical examination	336	307	(91%)	100%	(99 - 100)	99%	(97 - 100)
Vital signs	333	201	(60%)	99%	(97 - 100)	92%	(89 - 95)
General	268	211	(79%)	99%	(97 - 100)	100%	(98 - 100)
Dermatologic	216	172	(80%)	99%	(97 - 100)	95%	(92 - 98)
Lymph nodes/Heme	142	131	(92%)	99%	(96 - 100)	99%	(96 - 100)
HEENT	767	595	(78%)	98%	(97 - 100)	98%	(96 - 99)
Cardiovascular	293	235	(80%)	100%	(98 - 100)	98%	(96 - 99)
Gastrointestinal	295	225	(76%)	99%	(98 - 100)	97%	(94 - 98)
Chest	374	291	(78%)	99%	(97 - 100)	98%	(96 - 99)
Genitourinary	138	116	(84%)	99%	(96 - 100)	94%	(89 - 97)
Neurological	320	244	(76%)	97%	(94 - 100)	95%	(91 - 97)
Psychological	67	62	(93%)	100%	(95 - 100)	99%	(92 - 100)
Musculoskeletal	82	51	(62%)	95%	(87 - 99)	92%	(84 - 97)
Extremity exam	314	210	(67%)	97%	(95 - 99)	94%	(90 - 96)
Lab, imaging, and pathology results	393	246	(63%)	98%	(96 - 99)	88%	(84 - 91)
Analysis, assessment and plan	600	503	(84%)	98%	(96 - 99)	96%	(94 - 97)
Total	**7969**	**6325**	**(79%)**	**98.6%**	(98.3 - 98.8)	**96.2%**	(95.8 - 96.6)

Evaluators identified 160 sections that SecTag missed (omitted). These were either major sections (labeled in the document or not) or labeled clinically important non-major sections. Of these missing sections, all but 11 were found in the terminology: four laboratory or radiology findings, three plan subdivisions, and several alternative groupings of existing sections (e.g., “nose and ear exam” [without throat], facial exam as a separate component of head exam).

Table 7 shows the recall and precision for each major section. SecTag effectively identified labeled and non-labeled major sections. Overall, document authors rarely provided section labels for first degree relative family history (only 5% were labeled), but SecTag was still able to identify these sections, primarily using noun phrase processing.

Table 8: Accuracy of Section Boundary Detection for Correctly-labeled Sections.

Boundaries	**Label in document (%)**	**No label in document (%)**	**Total (%)**
Correct	10983 (94.8%)	3220 (85.9%)	14,203 (92.7%)
Incorrect Start	197 (1.7%)	20 (0.5%)	217 (1.4%)
Incorrect Ending	344 (3.0%)	502 (13.4%)	846 (5.5%)
Incorrect Start and End	56 (0.5%)	6 (0.2%)	62 (0.4%)
Total	11580	3748	15,328

Table 9: Type of Boundary Errors.
Sections that were “too long” occurred when SecTag included clinical content that was not relevant to the section; sections “too short” excluded important content from the section. Non-clinical content includes outline headers, dates, and generic non-clinical statements (e.g., page headers or transcription labels).

	Too long	Too short	Non-clinical content
Incorrect Start	4%	13%	3%
Incorrect Ending	27%	40%	11%
Incorrect Start and End	0%	3%	1%

SecTag identified the correct start and end boundaries for 92.7% of the correctly identified sections (Table 8). The system better predicted the labeled boundaries than unlabeled boundaries ($p<0.001$). The most common error was an ending error, meaning the ending of the section either ended too early (failed to include relevant content) or included content that did not belong to that section. Ending errors occurred 5.5% of the time; they were more common for both labeled and unlabeled document sections. Unlabeled sections were more likely to have an incorrect ending than labeled sections. A failure of analysis of 112 randomly selected incorrectly boundaries noted that 40% of the boundary errors were in starting or ending the section too early (Table 9). Approximately 15% of the boundary errors were due to nonclinical content in the section (e.g., outline headers, attestation statements, or transcription-inserted information); in general, reviewers were instructed to mark boundaries correct if the additional content was not clinically important.

Precision of Section Identification Techniques

Table 10 shows the precision of each of SecTag's component algorithms. The system accurately identified sections with document labels. It also accurately identified unlabeled sections by using NLP methods, removing modifiers such as anatomical locations, and detecting labels within a "sentence" (e.g., identifying both "temperature" and "weight" from "temp: 97.5 weight: 155lb"). Predicting sections via the Bayesian score was accurate for 81% of the predictions. Spelling correction was problematic, with a precision of 62%. Correct matches included multiword and single-word matches (e.g., "cheif complaint", "labarotory"); incorrect matches were all single word matches in the source phrase. Eight spelling correction errors were incorrect mapping of a source document acronym into an incorrect acronym in the section terminology. For example, "UDS" (meaning "urine drug screen") was mapped to the acronym "ID" (meaning "infectious disease"). Of all spelling correction errors, 8 were due to incorrectly disambiguating between the possible sections for an accurate spelling correction (e.g., "ucolor," meaning urine color, became "skin color" since urine color is not a defined section), 4 were abbreviation/acronyms not present in the terminology (including a person's initials), and 6 were a result of the wrong spelling correction chosen. In two of these cases, they were a result of Aspell's algorithm, which does not allow words that contain numbers and letters, and thus assumes that these are misspellings.

Table 10: Precision of SecTag Component Methods to Identify Sections.

Method	Count	(%)	Number Correct	Precision	(95%CI)
Labeled sections					
Exact or normalized match	11,221	(70.0%)	11,123	99%	(98.9 - 99.3)
Variant generation	130	(0.8%)	110	85%	(77-90)
Unlabeled Sections					
Bayesian prediction	1,867	(11.6%)	1,503	81%	(79-82)
Next-section rules	29	(0.2%)	27	93%	(77-92)
NLP	2,112	(13.2%)	1,939	92%	(91-93)
Both Labeled and Unlabeled Sections					
Spelling correction	53	(0.3%)	33	62%	(48-75)
Labels within a sentence*	471	(2.9%)	444	94%	(92-96)
Modifier removal	153	(1.0%)	150	98%	(94-100)
Totals	16,036		15,329	96%	(95-96)

Using the Bayesian scoring technique, SecTag predicted a total of 129 different section headers, with a precision of 81% (Table 11). The most common predicted section was the plan (153 occurrences), followed by assessment (96), substance use (76), and laboratory data (74). The median prediction occurred 6 times; 79 different sections were predicted less than 10 times in the evaluation set. The algorithm did especially poorly predicting individual vital signs (though accurate for overall vital signs), certain laboratory tests, and some physical exam elements. The most common errors were on electrocardiograms, cardiovascular plan, and individual family member medical histories. The precision of the Bayesian was not dependent on the number of times the section was predicted or the frequency of the section in the training set.

Table 11: Most Common Sections Predicted Using Bayesian Score.

Section name	Correct	Incorrect	Total	PPV
plan	126	27	153	82%
assessment	84	12	96	88%
substance_use	75	1	76	99%
laboratory_data	66	8	74	89%
oropharynx_exam	60	10	70	86%
extremity_exam	47	21	68	69%
abdominal_exam	54	13	67	81%
vital_signs	63	2	65	97%
past_surgical_history	31	21	52	60%
general_exam	44	0	44	100%
pulmonary_exam	43	1	44	98%
history_present_illness	25	14	39	64%
cardiovascular_plan	19	17	36	53%
laboratory_and_radiology_data	27	7	34	79%
chest_xray	28	5	33	85%
head_ent_exam	33	0	33	100%
neurological_exam	21	9	30	70%
pupil_exam	26	4	30	87%
strength_exam	15	15	30	50%
analysis	28	1	29	97%
gastrointestinal_exam	22	4	26	85%
jugular_venous_pulse	15	10	25	60%
medications	21	4	25	84%
coagulation_panel	11	12	23	48%
mother_medical_history	13	10	23	57%
musculoskeletal_extremity_exam	18	5	23	78%
ophthalmologic_exam	23	0	23	100%
cardiovascular_exam	20	0	20	100%
electrocardiogram	7	13	20	35%
infectious_disease_plan	19	1	20	95%
neck_exam	17	3	20	85%
Total	1503	364	1867	81%

Discarded Section Candidates

SecTag generated a total of 1,664 possible sections headers for which it considered the "best" candidate section header a poor match and thus discarded it. A poor Bayesian score was the most frequent reason to discard a possible section label (Table 12). The author evaluated 20 random notes to determine if the poor matches were appropriately discarded. Manual review suggested that 93% of the poorly-matched section headers were incorrect; 7% could have been retained rather than discarded, but none were major sections.

Table 12: Reasons for Discarded Section Candidates

	Count	(%)
Exact match elsewhere	287	17%
Bayesian score did not match tag	962	58%
Empty non-labeled section	111	7%
Match duplicates prior exact match	304	18%
Total	1664	

LOINC Concept and String Coverage

LOINC represented 86% of the concepts in the major sections and 77% of all sections tagged (TABLE 13). Its string representation was poor (20% of the labeled strings matching after normalization). The most common major sections missing from LOINC were family medical history entries for first degree relatives, grouped physical exam subcomponents (e.g., "musculoskeletal and extremity exam"), and major sections matched in more granular ways (e.g., "jugular venous pulse" instead of "neck exam" or "cardiovascular exam"). LOINC had slightly better coverage of labeled sections that unlabeled ones; LOINC contained concepts representing 72% of labeled concepts and 79% of unlabeled section concepts (p<0.001).

Table 13: LOINC Coverage in Identified Sections.

	All sections (%)		Major sections (%)	
Concept matches				
LOINC concept	12407	(77%)	6739	(86%)
Not in LOINC	3629	(23%)	1119	(14%)
String matches				
Matched a LOINC string	3246	(20%)	2535	(32%)
Did not match a LOINC string	12790	(80%)	5323	(68%)

Interrater Reliability

The interrater reliability on accuracy between all reviewers was good (Kappa = 0.70, p<0.0001). Reviewers agreed less on section boundaries; assessment for correctness was better (Kappa = 0.49, p<0.0001) than for the type of error (Kappa 0.43).

CHAPTER VI

DISCUSSION

The current study is one of the first large-scale efforts to formally evaluate a section header terminology and a related clinical note section header (tag) identifier. The author found that SecTag effectively identifies common sections in a wide variety of general H&P documents. The system uses a combination of NLP methods, concept matching with variant generation, and a score-based algorithm including a naïve Bayes classifier to effectively match section labels in documents and predict unlabeled sections. Incorrectly identified and omitted sections were rarely due to concept absence in the target vocabulary, suggesting the section terminology underlying SecTag sufficiently represented the documents in this corpus.

Accurate section identification is a key first step toward greater document understanding. To be useful, section identification should be coupled with more in-depth natural language processing or concept identification tools. Without identifying section-level context, a system cannot distinguish between identical diseases appearing in the family medical history or in the past medical history. Such understanding is crucial to allow for decision support or research. Furthermore, more detailed parsing than just "family medical history" is needed. Decision support systems, operating on contextual understanding of concepts within a note, could suggest that a patient with a family history of colon cancer in a first degree relative or with a past medical history of Crohn's disease needs early colorectal cancer screening.[156] Similarly, knowledge that a patient has a history of alcohol use or diabetes may trigger an alert to give early pneumococcal vaccination, instead of waiting until age 65.[157]

Accurate knowledge that a block of text belongs to a certain section may improve concept matching and name-value pairing, much in the same way as the Linguistic String Project found improved "understanding" by programming specific sublanguage grammars.[59] For example, the acronym "BS" in the respiratory/chest exam section of the physical exam likely means "breath sounds" but means "bowel sounds" in the abdominal exam section; likewise, such adjectives as "normal", "nontender" or "pain", "nonenlarged" may occur within many sections but indicate distinctly different clinical entities and evoke different differential diagnoses based on their contexts. "Subsection" parsers may be tuned for a higher prior probability, for instance, of negation within the review of systems than the history of present illness and employ different parsing schemes. For example, if in the vital signs section, one could assume any floating point number between 35.0 and 40.0 is likely a temperature, or a number following by a percentage is likely an oxygen saturation, especially if preceded by a number that could be a respiratory rate. Such ranges and probabilities could be built automatically, using documents that were tagged at the more specific levels (i.e., temperature, oxygen saturation).

The created section header terminology used in this study performed well, with only 11 new sections identified in this study should be present in the terminology. LOINC, the current standard for section terminology as adopted in the HL7 CDA, represented the concepts of major sections well though had inadequate synonymy to serve as a interface vocabulary for real-world clinical sections tags. Its performance among the major section of clinical interest was better than its overall performance, though still 14% of major sections by hierarchy were not labeled with a LOINC concept. Most of these were detailed physical exam subsections, assessment and plan systems, or unusual combinations of systems (“skin/breast” or “eye, ear, nose, throat”). In some cases, the LOINC match was less than perfect; the author mapped some terms from specialized components of LOINC to serve a more general purpose, such as “Treatment Plan” from a psychological corpus, to the general “plan” section. An effective parser must map these more granular or multi-section concepts to a LOINC concept. While LOINC does provide some synonymy of its concepts, its strings poorly covered the expressivity of section titles identified in the note. LOINC is designed as a reference terminology and thus the brief synonymy provided is more to ease lookup than enable LOINC as a NLP terminology.

Section tagging can be an enabling tool for clinical research and medical training. The author and colleagues have developed a companion application to the EMR that collects all trainees’ patient notes and stores them as their “experience log.” Many organizations, such as Association of American Medical Colleges and the Accreditation Council for Graduate Medical Education, include exposure and experience tracking as a key component of competency assessment.[158, 159] A full log of patient exposure, via their clinical documentation, could allow rapid reports of a student’s exposure by chief complaint and past medical history. An evaluator could assess, automatically, whether students have covered key physical exam elements or had exposure to important diseases and diagnoses. Complex questions, such as has a student obtained competency with a particular diagnosis or chief complaint, may also be aided by such a system. For instance, competency for a chief complaint of back pain may involve assessing certain physical exam elements and vital signs; asking the presence of saddle anesthesia, incontinence, weakness, and weight loss; and considering appropriate diagnoses in the context of the patients past medical history.

Immediate future directions for SecTag include making the application compatible with the HL7 CDA, which would allow compatibility with a large range of applications. The application could be useful as a web service, though HIPAA concerns would limit external use. To be useful, the application should also be extended to other document types, such as discharge summaries. SecTag stores prior probabilities and frequency data for sections by document type (selectable as a run-time parameter); likely, optimal performance would require training for new document types. Despite this limitation, the probabilities derived from H&Ps may have a broad applicability across many different document types that follow the same basic format, such as progress, clinic, and consultation notes. Discharge summaries and procedure notes are likely to require additional vocabulary development and training. The current training method, whereby the system "automatically" trains itself by iteratively processes a corpus of documents makes training on new corpora simpler. However, this method assumes a relatively well-tagged set of training documents and a complete terminology for the document set.

The failure analysis revealed several areas for possible improvement to enhance SecTag's performance. Some of the errors were due to the scrubbing software. It removed certain disease eponyms, sometimes confusing SecTag. Also, certain section labels that were acronyms were removed, ostensibly because the initials matched one of the other names in the document (e.g., the patient, a physician, or a nurse). This makes section prediction more difficult since a sometimes short sentence is then paired with a name, which often occurs in other sections (e.g., "attending physician" or "patient name").

A second cause of failure was the spelling correction algorithm, which performed especially poorly. Because Aspell does not support words containing numbers, it assumed these needed "correcting" to words without numbers. The author adjusted the algorithm to skip these words from spell-correction. A few words were missing from the spell-check vocabulary, which have been added. Despite these errors, its performance would have been significant improved if the possible section candidates chosen via spelling correction were compared against the Bayesian "figure of merit" score for the sentence. Instead, for some of the suggestions, the "best" candidate match was chosen and assumed to be correct without a "figure of merit" score validation.

The failure analysis suggests several other possible improvements to the software. Dictated documents often contained various formats of patient names, medical record numbers, and page numbers as new page headers. This format routinely caused errors in section tagging by correctly creating a "medical_record_number" tag that terminated the active section. While many times the algorithm did correctly return to the last section through Bayesian scoring, sometimes other, incorrect, sections were predicted or no section was predicted at all. The author attempted to handle this by removing these strings before processing by SecTag; however, the algorithm did not correctly process all possible formats of this string. A more efficacious strategy may be to allow the application to return to the last active section if the score is sufficiently high and a "short" data type section (e.g., "medical record number" or "patient name") caused its termination. Second, many laboratory/physical exam boundaries were not correctly detected. In this setting, the algorithm should be more lenient to accepting the Bayesian prediction. Also, a number of section boundary errors due to imperfect sentence parsing, especially with floating point numbers, such as temperatures and some drug dosages.

The Bayesian algorithm predicted unlabeled sections with an accuracy of 81% over 631 different possible sections from which to choose from. This represents a significant gain over chance selection (1 out of 1109), but was not as accurate as other methods. However, the Bayesian method was critical in the step of discarding erroneous sections, greatly improving their accuracy. One cause of error in the Bayesian prediction is imperfection of the gold standard. It was derived automatically from iterative tagging of the training corpora. The assumption was that with a low error rate that favored specificity over sensitivity (when the Bayesian prediction is turned off), the errors in classification would be deemphasized with regard to correct classifications. However, some sentence constructions NLP errors (e.g., identifying "electrocardiogram" from a plan to get an EKG as well as the procedure result) can systematically introduce errors for words that appear as a possible section in multiple parts of the note.

While the naïve Bayes approach performed acceptably, more sophisticated algorithms, such as support vector machines, may perform better. One method to reduce computational complexity of more robust methods, such as SVMs, may be to nest the classification decisions within known sections. First deciding which section at level *X* makes the most sense, then restrict the next classification to level *X+1* sections occurring within the selected parent section. For example, if “physical_exam” was chosen over “review_of_systems”, the next classification would ignore all children of “review_of_systems” (e.g., “cardiovascular_review”) from the next decision process.

Limitations

The study results must be interpreted in light of a few limitations. The author used H&Ps from a single medical center; formatting, styles, and section names may be different in this setting versus others. The author attempted to mitigate this bias by forming the bulk of the terminology from external, national sources such as the published literature and use of standard vocabularies such as LOINC and QMR. In addition, the author processed a variety of formats of documents from CBD templates to output from multiple dictation companies. Second, the gold standard was derived from automatically tagged documents, which allowed quick derivation of a large tagged corpus accurate for most sections. A manually tagged corpus, while potentially more accurate, may be infeasible since the section terminology can change frequently as new concepts are added. Since the non-probabilistic tagging performs better on key sections than subsections (as subsections are more likely to be predicted than tagged), this biases the predictions towards parent concepts.

The subcategory performance may be overestimated in some areas, specifically in laboratory and imaging results. These sections are not well-represented in the terminology, and reviewers were specifically told that this was not a goal of this study. Two reasons motivated this: first, these categories are already well represented in existing terminologies such as LOINC, SNOMED, and QMR, and, second, these results are often readily available from most EMRs and thus a natural language processing tool that identifies them has little use. Furthermore, the performance among other subcategorizations or labeled sections that were not components of the "key" sections may suffer from an information bias since reviewers may have been more attentive for key sections instead.

This study presents an initial evaluation of performance of a novel terminology of section tags on general and subspecialty H&Ps. The author has not validated its performance on other document types or on documents from other institutions. While the author designed the terminology to represent progress notes, consultation notes, and clinic notes, these were not generally included (with a few exceptions due to document mistitling) in the evaluation document set. Currently, the author is beginning to extend the terminology to contain sections from discharge summaries. Procedure (medical or surgical) and other note types may not be covered adequately at the current time.

Conclusion

This work is one of the first formal evaluations of a section header terminology and related section tagger. The system accurately identified both labeled and unlabeled sections in clinical H&Ps. Although LOINC contained most of the major sections in notes, the terminology's performance exceeded it, especially as an interface terminology allowing NLP on clinical notes. The section terminology contained the vast majority of labeled sections in clinical notes. To be of most use, the system needs to be coupled with a robust concept identification or NLP system to truly "understand" the content within a section. More research is also needed to extend the terminology and algorithms to other documents types and to study methods to use the section tagger to improve the efficacy of NLP tasks.

APPENDIX A

QMR FINDINGS HIERARCHY

01. Patient History
01.01. Demographics
01.01.01. Age
01.01.02. Race/Ethnic Background
01.01.03. Sex
01.02. Social History
01.02.01. Environmental Exposure History
01.02.01.01. Infectious Disease Exposure History
01.02.01.02. Insect and Animal Exposure History
01.02.01.03. Occupational History
01.02.01.04. Toxic Substances Exposure History
01.02.01.05. Miscellaneous Environmental Exposure History
01.02.02. Substance Abuse History
01.02.03. Level of Activity History
01.03. Past Medical History
01.03.01. Allergic Disorder History
01.03.01.01. Treatment History
01.03.01.01.01. Previous Surgery History
01.03.01.01.02. Current or Recent Drug Administration History
01.03.01.01.02.01. Current or Recent Antibiotic Administration History
01.03.01.01.02.02. Cytotoxic or Immunosuppressive Medication Administration His
01.03.01.01.02.03. Current or Recent Miscellaneous Medication Administration Hi
01.03.01.01.03. Blood Products Administration History
01.03.01.01.04. Radiation Therapy or Exposure History
01.03.01.01.05. Miscellaneous Therapy History
01.03.01.01.06. Invasive Diagnostic Procedure History
01.03.01.02. Trauma History
01.03.01.03. Residence or Travel History
01.04. Family History
01.05. Review Of Systems
01.05.01. Review of General Symptoms
01.05.02. Review Of Integumentary Symptoms
01.05.03. Review of Head Eyes Ears Nose And Throat Symptoms
01.05.04. Review of Lymphatic Hematopoeitic and Clotting Symptoms
01.05.04.01. Lymphatic System Symptoms
01.05.04.02. Hematopoeitic System Symptoms
01.05.04.03. Clotting Abnormality or Bleeding Disorder History
01.05.05. Review Of Cardiovascular Respiratory and Thoracic Symptoms
01.05.05.01. Dyspnea
01.05.05.02. Pain Chest
01.05.05.03. Miscellaneous Pulmonary History and Symptoms
01.05.05.04. Hypertension History
01.05.05.05. Miscellaneous Cardiovascular History and Symptoms
01.05.06. Breast Diseases History and Symptoms

01.05.07. Review Of Gastrointestinal and Dietary History and Symptoms
01.05.07.01. Previous Gastrointestinal Disorders History
01.05.07.02. Previous Gastrointestinal Surgery or Trauma History
01.05.07.03. Pain Abdomen
01.05.07.04. Diet and Appetite History
01.05.07.05. Bowel Habits
01.05.07.06. Dysphagia
01.05.07.07. Jaundice History
01.05.07.08. Nausea/Vomiting/Regurgitation
01.05.08. Review of Genitourinary Reproductive and Obstetrical Symptoms
01.05.08.01. Urinary Tract Disorders History and Symptoms
01.05.08.02. Female Urogenital Menstrual And Obstetrical History and Symptoms
01.05.08.03. Male Genitourinary and Reproductive History and Symptoms
01.05.08.04. Sexual and Venereal Diseases History and Symptoms
01.05.09. Review Of Neuropsychiatric Symptoms
01.05.09.01. Previous Neuropsychiatric Disorder or Therapy History
01.05.09.02. Neuropsychiatric Medication Administration History
01.05.09.03. Headache
01.05.09.04. Neurological Deficits History
01.05.09.05. Sleep Disturbances
01.05.09.06. Auras/Seizures History
01.05.09.07. Mental Status History
01.05.09.08. Miscellaneous Neuropsychiatric Symptoms
01.05.10. Review of Disorders of Musculoskeletal System
01.05.10.01. Pain Neck
01.05.10.02. Pain Back
01.05.10.03. Joint Symptoms
01.05.10.04. Pain or Discomfort Extremities
01.05.10.05. Miscellaneous Musculoskeletal Symptoms
01.05.11. Review of Diseases of Metabolism or Endocrine System
01.05.11.01. Endocrine Disorders History Symptoms or Therapy History
01.05.11.02. Metabolic Disorders History
01.05.12. Review of Infectios Diseases History or Symptoms
01.05.13. Review of Diseases of Congenital or Inherited Nature History
01.05.14. Review of Diseases of Neoplastic or Malignant Nature History
02. Physical Examination
02.01. General Appearance
02.02. Vital Signs
02.02.01. Temperature
02.02.02. Heart Rate
02.02.03. Respiratory Rate and Pattern of Respiration
02.02.04. Weight
02.02.05. Height
02.02.06. Blood Pressure
02.03. Inspection And Palpation Skin Hair and Nails
02.03.01. Inspection Hair and Nails
02.03.02. Skin and Mucosa Pigmentary or Color Changes
02.03.03. Skin Superficial Lesions or Palpable Deep Lesions
02.03.04. Inspection Skin Rashes
02.03.05. Skin Texture Temperature and State of Hydration
02.04. Head Eyes Ears Nose And Throat Exam
02.04.01. Inspection And Palpation Face And Neck
02.04.01.01. Cranial Arteries Exam
02.04.01.02. Examination Salivary and Lacrimal Glands
02.04.02. Ears Exam
02.04.03. Eyes Exam

02.13.05. Neurologic Exam Sensory
02.13.06. Neurologic Exam Muscle Strength and Tone
02.13.07. Neurologic Exam Coordination and Gait
02.13.08. Neurologic Exam Tremor Chorea or Extrapyramidal Signs
02.13.09. Neurologic Exam Speech
02.13.10. Neurologic Exam Observation of Seizures
02.13.11. Neurologic Exam Signs of Meningeal Irritation
02.13.12. Miscellaneous Neurologic Exam
02.14. Musculoskeletal Exam
02.14.01. Extremities and Joints Inspection and Palpation
02.14.01.01. Bone Joints or Tendon Abnormalities Extremities
02.14.01.02. Joints Range of Motion Extremities
02.14.01.03. Joints Pattern of Involvement
02.14.01.04. Miscellaneous Exam Extremities
02.14.02. Ribs Sternum and Costochondral Junctions Exam
02.14.03. Back and Spine Exam
02.14.04. Inspection and Palpation of Skeletal Muscles
03. Laboratory Tests
03.01. Body Fluid or Body Substance Analysis
03.01.01. Blood Analysis
03.01.01.01. Hematologic Studies
03.01.01.01.01. Hemoglobin and Hematocrit
03.01.01.01.02. Complete Blood Count And Peripheral Smear
03.01.01.01.03. Platelet Count
03.01.01.01.04. Reticulocyte Count
03.01.01.01.05. Rbc Indices
03.01.01.01.06. Hemolysis Studies
03.01.01.01.07. Iron Metabolism Related Studies
03.01.01.01.08. Leukocyte Enzyme Assays
03.01.01.01.09. Hemoglobin Electrophoresis
03.01.01.01.10. Red Blood Cell Mass Determination
03.01.01.01.11. Coagulation Tests
03.01.01.01.12. Miscellaneous Hematologic Studies
03.01.01.02. Blood Microbiological Studies
03.01.01.02.01. Blood Culture or Isolation of Microbiological Organism
03.01.01.02.02. Blood Smear For Parasites
03.01.01.02.03. Blood Immunological Tests Indicating Exposure to Infectious Age
03.01.01.02.04. Miscellaneous Blood Microbiologic Studies
03.01.01.03. Blood Immunologic Studies
03.01.01.03.01. Antibodies Autoimmune
03.01.01.03.02. Serum Complement Studies
03.01.01.03.03. Histocompatibility Antigen Determination
03.01.01.03.04. Serum Immunoelectrophoresis and Quantitative Immunoglobulins
03.01.01.03.05. Miscellaneous Blood Immunologic Studies
03.01.01.04. Blood Biochemical Analysis
03.01.01.04.01. Serum Electrolytes Routine
03.01.01.04.02. Glucose Blood
03.01.01.04.03. Blood Urea Nitrogen and Serum Creatinine Levels
03.01.01.04.04. Calcium/Phosphate/Magnesium Serum
03.01.01.04.05. Bilirubin Serum
03.01.01.04.06. Lipids Serum
03.01.01.04.07. Ketones Serum
03.01.01.04.08. Lactate Serum
03.01.01.04.09. Serum Protein and Enzymatic Components Analysis
03.01.01.04.09.01. Serum Phosphatases
03.01.01.04.09.02. Serum Amylase and Lipase

03.01.01.04.09.03. Serum Transaminases
03.01.01.04.09.04. Myocardial and Skeletal Muscle Enzymes Serum <non transamina
03.01.01.04.09.05. Serum Protein Electrophoresis
03.01.01.04.09.06. Renin Plasma
03.01.01.04.09.07. Miscellaneous Serum Protein and Enzymatic Components
03.01.01.04.10. Vitamins And Minerals Assays
03.01.01.04.11. Uric Acid Serum
03.01.01.04.12. Blood Toxicological Studies
03.01.01.04.13. Miscellaneous Blood Biochemical Studies
03.01.01.05. Serum Osmolality
03.01.02. Csf Analysis
03.01.02.01. Csf Routine Exam
03.01.02.02. Csf Microbiological Studies
03.01.02.03. Csf Special Studies
03.01.03. Nasal Mucosa or Discharge Microbiological Studies
03.01.04. Oral Lesions Microbiological Studies
03.01.05. Transtracheal Aspiration
03.01.06. Sputum Analysis
03.01.06.01. Sputum Exam Routine
03.01.06.02. Sputum Culture
03.01.06.03. Sputum Exam Special Procedures
03.01.07. Bile Analysis
03.01.07.01. Bile Microbiological Studies
03.01.08. Feces Analysis
03.01.08.01. Feces Exam Routine
03.01.08.02. Feces Microbiology
03.01.08.03. Feces Special Studies
03.01.09. Urinalysis
03.01.09.01. Urinalysis Routine And Microscopic
03.01.09.02. Urine Microbiology
03.01.09.03. Urine Simple Biochemical Analysis
03.01.09.04. Urinalysis Special Procedures
03.01.10. Semen Analysis
03.01.11. Prostatic Fluid Analysis
03.01.12. Urethra and Cervix Microbiological Studies
03.01.13. Skin Lesion Microbiological Studies
03.01.14. Serous Fluid Analysis
03.01.14.01. Ascitic Fluid or Peritoneal Aspirate Analysis
03.01.14.01.01. Ascitic Fluid Routine Examination
03.01.14.01.02. Ascitic Fluid Special Analysis
03.01.14.01.03. Ascitic Fluid Microbiological Studies
03.01.14.01.04. Peritoneal Aspirate Analysis
03.01.14.01.04.01. Peritoneal Aspirate Routine Studies
03.01.14.01.04.02. Peritoneal Aspirate Microbiological Studies
03.01.14.02. Joint Fluid Analysis
03.01.14.02.01. Joint Fluid Routine Examination
03.01.14.02.02. Joint Fluid Microbiological Studies
03.01.14.02.03. Joint Fluid Special Studies
03.01.14.03. Pericardial Fluid Analysis
03.01.14.03.01. Pericardial Fluid Routine Studies
03.01.14.03.02. Pericardial Fluid Microbiological Studies
03.01.14.04. Pleural Fluid Analysis
03.01.14.04.01. Pleural Fluid Routine Studies
03.01.14.04.02. Pleural Fluid Microbiological Studies
03.01.14.04.03. Pleural Fluid Special Studies
03.01.14.05. Miscellaneous Serous Fluid Analysis

03.01.15. Sweat Analysis
03.01.16. Skin Testing
03.01.16.01. Skin Anergy Panel
03.01.16.02. Skin Tests for Exposure to Microorganisms
03.01.16.03. Miscellaneous Skin Tests
03.02. Toxicological Studies
03.03. Microbiological Studies
03.03.01. Smears and Stains for Infectious Agents
03.03.01.01. Smears and Stains for Bacteria
03.03.01.01.01. Smears and Stains for Positive Organisms
03.03.01.01.01.01. Smears and Stains for Gram Positive Cocci
03.03.01.01.01.02. Smears and Stains for Gram Positive Rods
03.03.01.01.02. Smears and Stains for Gram Negative Organisms
03.03.01.01.02.01. Smears and Stains for Gram Negative Cocci
03.03.01.01.02.02. Smears and Stains for Gram Negative Rods
03.03.01.02. Smears and Stains for Nocardia and Actinomyces
03.03.01.03. Smears and Stains for Spirochetal Infections
03.03.01.04. Smears and Stains for Legionella Infections
03.03.01.05. Smears and Stains for Fungi
03.03.01.06. Smears and Stains for Mycobacteria
03.03.01.07. Smears and Stains for Mycoplasma Rickettsia And Chlamydia
03.03.01.08. Smears and Stains for Viruses
03.03.01.09. Smears and Stains for Protozoa
03.03.01.10. Smears and Stains for Parasitic Organisms
03.03.02. Culture
03.03.02.01. Cultures for Bacteria
03.03.02.01.01. Cultures for Gram Positive Organisms
03.03.02.01.01.01. Cultures for Gram Positive Cocci
03.03.02.01.01.02. Cultures for Gram Positive Rods
03.03.02.01.02. Cultures for Gram Negative Organisms
03.03.02.01.02.01. Cultures for Gram Negative Cocci
03.03.02.01.02.02. Cultures for Gram Negative Rods
03.03.02.02. Cultures for Nocardia and Actinomyces
03.03.02.03. Cultures for Spirochetes
03.03.02.04. Cultures for Legionella Infections
03.03.02.05. Cultures for Fungi
03.03.02.06. Cultures for Mycobacteria
03.03.02.07. Cultures for Mycoplasma Rickettsia And Chlamydia
03.03.02.08. Cultures for Viruses
03.03.03. Immunological Tests for Infectious Agents
03.03.03.01. Immunological Tests for Bacteria
03.03.03.01.01. Immunological Tests for Gram Positive Organisms
03.03.03.01.01.01. Immunological Tests for Gram Positive Cocci
03.03.03.01.01.02. Immunological Tests for Gram Positive Rods
03.03.03.01.02. Immunological Tests for Gram Negative Organisms
03.03.03.01.02.01. Immunological Tests for Gram Negative Cocci
03.03.03.01.02.02. Immunological Tests for Gram Negative Rods
03.03.03.02. Immunological Tests for Nocardia and Actinomyces
03.03.03.03. Immunological Tests for Spirochetes
03.03.03.04. Immunological Tests for Legionella Infections
03.03.03.05. Immunological Tests for Fungi
03.03.03.06. Immunological Tests for Mycobacteria
03.03.03.07. Immunological Tests for Mycoplasma Rickettsia And Chlamydia
03.03.03.08. Immunological Tests for Viruses
03.03.03.09. Immunological Tests for Protozoa
03.03.03.10. Immunological Tests for Parasitic Organisms

03.04. Tests for Malignant Neoplastic Disorders
03.04.01. Cytological Examination for Malignant Neoplastic Disorders
03.04.02. Serological Markers Consistent With Malignant Neoplastic Disorders
03.04.03. Biopsies Consistent with Malignant Neoplastic Disorders
03.05. Tests Of Physiological and/or Organ System Function <NON-Imaging>
03.05.01. Endocrine Function Tests
03.05.01.01. Pituitary Function Tests
03.05.01.02. Thyroid Function Tests
03.05.01.03. Parathyroid Function Tests
03.05.01.04. Adrenal Function Tests
03.05.01.05. Pancreas Endocrine Function Tests
03.05.01.06. Gonadal Function Tests
03.05.01.07. Miscellaneous Endocrine Function Tests
03.05.02. Kidney Function Tests
03.05.03. Liver Function Tests
03.05.04. Cardiovascular Function Tests <NON-Imaging>
03.05.04.01. EKG
03.05.04.02. Pressure Central Venous
03.05.04.03. Prolonged Cardiac EKG Monitoring
03.05.04.04. Cardiovascular Stress Tests
03.05.04.05. Cardiac Output and Arteriovenous Oxygen Difference Measurement
03.05.04.06. Plasma Volume Determination
03.05.04.07. Cardiac Catheterization Pressure and Flow Measurements
03.05.04.08. Noninvasive Peripheral Vascular Studies
03.05.04.09. Miscellaneous Cardiovascular Function Tests
03.05.05. Pulmonary Function Tests and Pulmonary Monitoring
03.05.05.01. Pulmonary Sleep Monitoring
03.05.05.02. Arterial Blood Gases and pH
03.05.05.03. Pulmonary Function Tests
03.05.06. Gastrointestinal Function Tests
03.05.06.01. Esophagus Manometry
03.05.06.02. Hepatobiliary Function Tests
03.05.06.03. Pancreas Exocrine Function Tests
03.05.06.04. Tests of Gastrointestinal Absorption
03.05.06.05. Hepatic Vein Wedge Pressure
03.05.06.06. Schilling Test
03.05.06.07. Miscellaneous Gastrointestinal Function Tests
03.05.07. Neurologic and Musculoskeletal Function Tests
03.05.07.01. Central Nervous System Electrophysiological Monitoring
03.05.07.01.01. EEG
03.05.07.01.02. Visual and Auditory Evoked Potentials
03.05.07.01.03. Electronystagmogram and Vestibular Function Tests
03.05.07.02. Electromyography
03.05.07.03. Nerve Conduction Velocities
03.05.07.04. Miscellaneous Neurologic and Musculoskeletal Function Tests
03.05.08. Ophthalmological Function Tests
03.05.08.01. Eyes Tonometry
03.05.08.02. Eyes Slit Lamp Examination
03.05.08.03. Eyes Formal Visual Field Testing
03.05.09. Physiological or Pharmacological Challenge Tests
03.06. Imaging Techniques <Indirect>
03.06.01. Routine Xray Radiographic Studies
03.06.01.01. Xray Chest
03.06.01.01.01. Xray Lung Fields
03.06.01.01.02. Xray Pleura
03.06.01.01.03. Xray Mediastinum

03.06.01.01.04. Xray Heart And Great Vessels
03.06.01.01.05. Xray Ribs and Bony Thorax
03.06.01.01.06. Miscellaneous Xray Chest
03.06.01.02. Xray Abdomen Plain Film
03.06.01.02.01. Xray Abdomen Intraperitoneal Contents
03.06.01.02.02. Xray Kidneys
03.06.01.02.03. Xray Abdominal Aorta
03.06.01.02.04. Miscellaneous Xray Abdomen Plain Film
03.06.01.03. Xray Joints And Bony Structures
03.06.01.03.01. Xray Bones General Features
03.06.01.03.02. Xray Skull
03.06.01.03.03. Xray Cervical Thoracic and/or Lumbosacral Spine
03.06.01.03.04. Xray Pelvis
03.06.01.03.05. Xray Long Bones Extremities
03.06.01.03.06. Xray Hands or Feet
03.06.01.03.07. Xray Joints
03.06.01.04. Xray Soft Tissues
03.06.01.04.01. Xray Intracranial Sinuses
03.06.01.04.02. Xray Neck Soft Tissues
03.06.01.04.03. Xray Extremities Soft Tissues
03.06.01.04.04. Mammography
03.06.01.04.05. Miscellaneous Xray Soft Tissues
03.06.02. Ultrasonography
03.06.02.01. Ultrasonography of Heart And Great Vessels
03.06.02.02. Ultrasonography Plethysmography and Doppler Flow Studies Blood Ves
03.06.02.03. Ultrasonography Abdomen
03.06.02.03.01. Ultrasonography Kidneys and Retroperitoneum
03.06.02.03.02. Ultrasonography Liver And Biliary Tract
03.06.02.03.03. Ultrasonography Pancreas
03.06.02.03.04. Ultrasonography Abdominal Vessels
03.06.02.03.05. Ultrasonography Pelvis
03.06.02.03.06. Miscellaneous Ultrasonography
03.06.02.04. Computerized Axial Tomograpy and Magnetic Resonance Imaging
03.06.02.04.01. Computerized Tomography or MRI Head
03.06.02.04.01.01. Computerized Tomography or MRI Head
03.06.02.04.01.02. Computerized Tomography or MRI Face and Bony Skull
03.06.02.04.02. Computerized Tomography or MRI Chest
03.06.02.04.03. Computerized Tomography or MRI Abdomen
03.06.02.04.04. Computerized Tomography Pelvis
03.06.02.04.05. Computerized Tomography or MRI Spine
03.06.02.04.06. Miscellaneous Computerized Axial Tomograpy and MRI
03.06.03. Radiographic Contrast Studies
03.06.03.01. Bronchography
03.06.03.02. Angiocardiography
03.06.03.03. Pulmonary Arteriography
03.06.03.04. Gastrointestinal Barium Contrast Studies
03.06.03.04.01. Barium Swallow or Cine-Esophogram
03.06.03.04.02. Upper GI Series Barium Meal
03.06.03.04.03. Small Bowel Follow Through
03.06.03.04.04. Barium Enema
03.06.03.05. Pancreatography Retrograde
03.06.03.06. Cholecystography
03.06.03.07. Intravenous Cholangiography
03.06.03.08. Percutaneous Cholangiography
03.06.03.09. Endoscopic Retrograde Cholangiopancreatography
03.06.03.10. Pyelography

03.06.03.11. Cystography
03.06.03.12. Myelography
03.06.03.13. Angiography <extra cardiac>
03.06.03.13.01. Cerebral Angiography
03.06.03.13.02. Thoracic <non cardiac> Angiography
03.06.03.13.03. Abdominal Angiography
03.06.03.13.04. Renal and Adrenal Angiography
03.06.03.13.05. Extremities Angiography
03.06.03.14. Lymphangiography
03.06.03.15. Hysterosalpingography
03.06.03.16. Miscellaneous Radiographic Contrast Studies
03.06.04. Radionuclide Imaging
03.06.04.01. Brain Radioisotope Scan
03.06.04.02. Thyroid Radioisotope Scan
03.06.04.03. Cardiac Radioisotope Scan
03.06.04.04. Lung Radioisotope Scan
03.06.04.05. Hepatobiliary Radioisotope Scan
03.06.04.06. Kidney Radioisotope Scan
03.06.04.07. Bone Radioisotope Scan
03.06.04.08. Miscellaneous Radionclide Imaging
03.07. Endoscopic Visualization Procedures
03.07.01. Respiratory Tract Endoscopy
03.07.01.01. Laryngoscopy
03.07.01.02. Bronchoscopy or Bronchial-Alveolar Lavage
03.07.02. Gastrointestinal Endoscopy
03.07.02.01. Esophagoscopy Gastroscopy And/Or Duodenoscopy
03.07.02.02. Sigmoidoscopy And/Or Colonoscopy
03.07.03. Peritoneoscopy
03.07.04. Cystoscopy and Cystometrogram
03.07.05. Culdoscopy and Culdocentesis
03.07.06. Arthroscopy
03.07.07. Miscellaneous Endoscopic Visualization Procedures
03.08. Biopsies and/or Histopathological Studies
03.08.01. Biopsy Nervous System
03.08.01.01. Brain Biopsy
03.08.01.02. Peripheral Nerve Biopsy
03.08.02. Respiratory Tract Histopathological Studies
03.08.02.01. Oropharynx Biopsy
03.08.02.02. Upper Respiratory Tract Biopsy
03.08.02.03. Bronchial Washings Or Brush Biopsy
03.08.02.04. Lung Biopsy
03.08.02.04.01. Endobronchial or Transbronchial Biopsy
03.08.02.04.02. Open Lung Biopsy
03.08.02.04.03. Lung Biopsy unspecified
03.08.02.05. Pleura Biopsy
03.08.03. Breast Biopsy/Aspirate
03.08.04. Endocrine Organ Biopsy
03.08.04.01. Pituitary Biopsy
03.08.04.02. Thyroid Biopsy
03.08.04.03. Parathyroid Biospy
03.08.04.04. Adrenal Biopsy
03.08.04.05. Gonadal Biopsy
03.08.05. Cardiovascular Biopsy
03.08.05.01. Myocardial Biopsy
03.08.05.02. Pericardial Biopsy
03.08.05.03. Blood Vessel Biopsy

03.08.06. Gastrointestinal Tract and Peritoneum Histopathological Studies
03.08.06.01. Esophagus Biopsy
03.08.06.02. Stomach Biopsy
03.08.06.03. Duodenum or Small Intestine Biopsy
03.08.06.04. Colon Biopsy
03.08.06.05. Rectum or Anus Biopsy
03.08.06.06. Hepatobiliary Histopathological Studies
03.08.06.06.01. Liver Biopsy or Aspirate
03.08.06.06.02. Biliary Tract Biopsy
03.08.06.07. Pancreas Biospy
03.08.06.08. Peritoneum Biopsy
03.08.06.09. Miscellaneous Gastrointestinal Biopsies
03.08.07. Genitourinary Histopathological Studies
03.08.07.01. Kidney Biopsy
03.08.07.02. Ureter or Urinary Bladder Biopsy
03.08.07.03. Uterus Biopsy or Curettage
03.08.07.04. Prostate Biopsy or Aspirate
03.08.07.05. Miscellaneous Genitourinary Histopathological Studies
03.08.08. Hematopoetic And Reticuloendothelial Histopathological Studies
03.08.08.01. Bone Marrow Biopsy
03.08.08.01.01. Bone Marrow Aspirate and Biopsy Routine Studies
03.08.08.01.02. Bone Marrow Microbiological Studies
03.08.08.01.03. Bone Marrow Special Studies
03.08.08.02. Lymph Node Biopsy or Aspirate
03.08.08.03. Splenic Aspirate or Biopsy
03.08.09. Skin Biopsy
03.08.10. Musculoskeletal Histopathological Studies
03.08.10.01. Skin-To-Muscle Biopsy
03.08.10.02. Muscle Biopsy
03.08.10.03. Bone Biopsy
03.08.10.04. Synovium Biopsy
03.08.10.05. Miscellaneous Musculoskeletal Histopathological Studies
03.08.11. Miscellaneous Biopsies and/or Histopathological Studies

APPENDIX B

PARTIAL LIST OF SECTION TERMINOLOGY UNIQUE CONCEPT IDENTIFIERS

Below is a list of all terms in the Section Terminology with levels higher than level 3. There are 556 concept terms in this list (out of a total of 1109). Tree numbers are the corresponding frequencies represents the frequency of the section in the training set. Those concepts that are "composite concepts" – not a primary grouping – are designated as such.

- source_and_reliability (Tree:1, Frequency: 0.0000)
  - history_source (Tree:1.1, Frequency: 0.4462)
  - reliability (Tree:1.2, Frequency: 0.0000)
- document_types (Tree:2, Frequency: 0.0000)
  - general_history_and_physical (Tree:2.3, Frequency: 0.0000)
    - inpatient_history_and_physical (Tree:2.3.1, Frequency: 0.0000)
    - clinic_history_and_physical (Tree:2.3.2, Frequency: 0.0000)
    - attending_admission_confirmation_note (Tree:2.3.3, Frequency: 0.0000)
  - discharge_summary (Tree:2.4, Frequency: 0.0000)
  - clinic_note (Tree:2.5, Frequency: 0.0000)
  - progress_note (Tree:2.6, Frequency: 0.0000)
  - procedure_note (Tree:2.7, Frequency: 0.0000)
    - epidural_procedure_note (Tree:2.7.4, Frequency: 0.0000)
  - consultation_note (Tree:2.8, Frequency: 0.0000)
  - operative_notes (Tree:2.9, Frequency: 0.0000)
    - post_operative_note (Tree:2.9.5, Frequency: 0.0000)
    - brief_operative_note (Tree:2.9.6, Frequency: 0.0000)
- providers (Tree:3, Frequency: 0.0049)
  - providers_by_type (Tree:3.10, Frequency: 0.0000)
    - physician (Tree:3.10.7, Frequency: 0.0071)
      - dictating_physician (Tree:3.10.7.1, Frequency: 0.3353)
      - requesting_physician (Tree:3.10.7.2, Frequency: 0.0000)
      - private_physician (Tree:3.10.7.3, Frequency: 0.0000)
      - surgeon (Tree:3.10.7.4, Frequency: 0.0029)
      - pediatrician (Tree:3.10.7.5, Frequency: 0.0000)
      - obstetrician (Tree:3.10.7.6, Frequency: 0.0002)

primary_physician (Tree:3.10.7.7, Frequency: 0.2530)
other_housestaff (Tree:3.10.7.8, Frequency: 0.3183)
attending_physician (Tree:3.10.7.9, Frequency: 0.7065)
anesthesiologist (Tree:3.10.7.10, Frequency: 0.0000)
additional_attending_physician (Tree:3.10.7.11, Frequency: 0.0000)
additional_resident (Tree:3.10.7.12, Frequency: 0.0000)
intern (Tree:3.10.7.13, Frequency: 0.0075)
resident (Tree:3.10.7.14, Frequency: 0.0143)
fellow (Tree:3.10.7.15, Frequency: 0.0019)
pathologist (Tree:3.10.7.16, Frequency: 0.0007)
radiologist (Tree:3.10.7.17, Frequency: 0.0002)
infectious_disease_attending (Tree:3.10.7.18, Frequency: 0.0000)
cardiologist (Tree:3.10.7.19, Frequency: 0.0048)
nonphysician (Tree:3.10.8, Frequency: 0.0000)
midwife (Tree:3.10.8.20, Frequency: 0.0000)
orthodontist (Tree:3.10.8.21, Frequency: 0.0000)
nurse_practitioner (Tree:3.10.8.22, Frequency: 0.0003)
providers_by_role (Tree:3.11, Frequency: 0.0000)
assistant (Tree:3.11.9, Frequency: 0.0000)
first_assistant (Tree:3.11.9.23, Frequency: 0.0000)
second_assistant (Tree:3.11.9.24, Frequency: 0.0000)
referral (Tree:3.11.10, Frequency: 0.0008)
referring_physician (Tree:3.11.10.25, Frequency: 0.0020)
additional_referring_physician (Tree:3.11.10.26, Frequency: 0.0000)
consultant (Tree:3.11.11, Frequency: 0.1339)
operating_room_consultation (Tree:3.11.11.27, Frequency: 0.0000)
endoscopist (Tree:3.11.12, Frequency: 0.0000)
perfusionist (Tree:3.11.13, Frequency: 0.0000)
standard_coding_systems (Tree:4, Frequency: 0.0000)
cpt_code (Tree:4.12, Frequency: 0.0017)
icd_code (Tree:4.13, Frequency: 0.0000)
patient_history (Tree:5, Frequency: 0.0987)
demographics (Tree:5.14, Frequency: 0.0002)
age (Tree:5.14.14, Frequency: 0.0114)
estimated_gestational_age (Tree:5.14.14.28, Frequency: 0.0000)
race (Tree:5.14.15, Frequency: 0.0000)
gender (Tree:5.14.16, Frequency: 0.0005)
address (Tree:5.14.17, Frequency: 0.0015)
home_address (Tree:5.14.17.29, Frequency: 0.0000)
work_address (Tree:5.14.17.30, Frequency: 0.0000)
emergency_contact (Tree:5.14.18, Frequency: 0.0000)
insurance (Tree:5.14.19, Frequency: 0.0005)
phone_number (Tree:5.14.20, Frequency: 0.0034)
home_phone_number (Tree:5.14.20.31, Frequency: 0.0003)
cell_phone_number (Tree:5.14.20.32, Frequency: 0.0022)
work_phone_number (Tree:5.14.20.33, Frequency: 0.0010)
patient_name (Tree:5.14.21, Frequency: 0.7799)
mrn (Tree:5.14.22, Frequency: 0.5371)
gravida (Tree:5.14.23, Frequency: 0.0112)
para (Tree:5.14.24, Frequency: 0.0063)
clinic (Tree:5.14.25, Frequency: 0.0127)
room (Tree:5.14.26, Frequency: 0.0071)
unit (Tree:5.14.27, Frequency: 0.0394)
ssn (Tree:5.14.28, Frequency: 0.0000)
case_number (Tree:5.14.29, Frequency: 0.0144)
admission_date (Tree:5.15, Frequency: 0.0000)

service (Tree:5.15.31, Frequency: 0.2472)
admitting_attending (Tree:5.15.32, Frequency: 0.0003)
consultation_attending (Tree:5.15.33, Frequency: 0.0000)
date_time (Tree:5.16, Frequency: 0.3995)
date (Tree:5.16.35, Frequency: 0.0156)
date_of_birth (Tree:5.16.35.34, Frequency: 0.0042)
date_of_discharge (Tree:5.16.35.35, Frequency: 0.0156)
date_of_death (Tree:5.16.35.36, Frequency: 0.0000)
date_of_injury (Tree:5.16.35.37, Frequency: 0.0000)
date_of_surgery (Tree:5.16.35.38, Frequency: 0.0000)
date_of_examination (Tree:5.16.35.39, Frequency: 0.0000)
date_of_autopsy (Tree:5.16.35.40, Frequency: 0.0000)
date_of_procedure (Tree:5.16.35.41, Frequency: 0.0000)
date_of_request (Tree:5.16.35.42, Frequency: 0.0000)
date_of_service (Tree:5.16.35.43, Frequency: 0.0612)
date_of_admission (Tree:5.16.35.44, Frequency: 0.2850)
report_date (Tree:5.16.35.45, Frequency: 0.0000)
duration_of_recording (Tree:5.16.35.46, Frequency: 0.0000)
due_date (Tree:5.16.35.47, Frequency: 0.0000)
conception_date (Tree:5.16.35.48, Frequency: 0.0000)
date_dictated (Tree:5.16.35.49, Frequency: 0.1220)
date_transcribed (Tree:5.16.35.50, Frequency: 0.1732)
time (Tree:5.16.36, Frequency: 0.0029)
time_of_arrival (Tree:5.16.36.51, Frequency: 0.0003)
time_of_injury (Tree:5.16.36.52, Frequency: 0.0000)
time_of_birth (Tree:5.16.36.53, Frequency: 0.0000)
patient_summary (Tree:5.17, Frequency: 0.0002)
identifying_information (Tree:5.18, Frequency: 0.0110)
code_status (Tree:5.19, Frequency: 0.0184)
living_will (Tree:5.20, Frequency: 0.0007)
livingwill_codestatus (Tree:5.20-28, Frequency: 0.0000) , composite concept
condition (Tree:5.21, Frequency: 0.0071)
admission_condition (Tree:5.21.37, Frequency: 0.0002)
discharge_condition (Tree:5.21.38, Frequency: 0.0008)
diagnoses (Tree:5.22, Frequency: 0.0139)
principal_diagnosis (Tree:5.22.39, Frequency: 0.0027)
secondary_diagnoses (Tree:5.22.40, Frequency: 0.0037)
diagnosis_at_death (Tree:5.22.41, Frequency: 0.0002)
other_diagnosis (Tree:5.22.42, Frequency: 0.0000)
cytologic_diagnosis (Tree:5.22.43, Frequency: 0.0000)
admission_diagnosis (Tree:5.22.44, Frequency: 0.0138)
discharge_diagnosis (Tree:5.22.45, Frequency: 0.0010)
postprocedure_diagnosis (Tree:5.22.46, Frequency: 0.0003)
preprocedure_diagnosis (Tree:5.22.47, Frequency: 0.0005)
final_diagnosis (Tree:5.22.48, Frequency: 0.0000)
psychiatric_diagnostic_classifications (Tree:5.22.49, Frequency: 0.0000)
axis_i (Tree:5.22.49.54, Frequency: 0.0003)
axis_ii (Tree:5.22.49.55, Frequency: 0.0003)
axis_iii (Tree:5.22.49.56, Frequency: 0.0003)
axis_iv (Tree:5.22.49.57, Frequency: 0.0003)
axis_v (Tree:5.22.49.58, Frequency: 0.0024)
procedures (Tree:5.23, Frequency: 0.0112)
secondary_procedures (Tree:5.23.50, Frequency: 0.0003)
principal_procedures (Tree:5.23.51, Frequency: 0.0024)
surgical_procedures (Tree:5.23.52, Frequency: 0.0037)
diet (Tree:5.24, Frequency: 0.0129)

- discharge_diet (Tree:5.24.53, Frequency: 0.0003)

- birth_history (Tree:5.25, Frequency: 0.0000)
  - birth_weight (Tree:5.25.54, Frequency: 0.0002)
  - birth_length (Tree:5.25.55, Frequency: 0.0000)
  - birth_headcircumference (Tree:5.25.56, Frequency: 0.0000)
- chief_complaint (Tree:5.26, Frequency: 0.7601)
- reason_for_consult (Tree:5.27, Frequency: 0.0007)
- history_present_illness (Tree:5.28, Frequency: 0.9225)
  - subjective (Tree:5.28.57, Frequency: 0.0032)
  - history_of_exposure (Tree:5.28.58, Frequency: 0.0000)
  - brief_history (Tree:5.28.59, Frequency: 0.0058)
  - current_pregnancy (Tree:5.28.60, Frequency: 0.0000)
  - other_issues (Tree:5.28.61, Frequency: 0.0073)
  - pertinent_clinical_findings (Tree:5.28.62, Frequency: 0.0000)
  - pain_history (Tree:5.28.63, Frequency: 0.0000)
    - pain_temporal_pattern (Tree:5.28.63.59, Frequency: 0.0000)
    - pain_alleviating_factors (Tree:5.28.63.60, Frequency: 0.0000)
    - pain_initiating_event (Tree:5.28.63.61, Frequency: 0.0000)
- reason_for_study (Tree:5.29, Frequency: 0.0002)
- risk_factors (Tree:5.30, Frequency: 0.0083)
  - cardiac_risk_factors (Tree:5.30.64, Frequency: 0.0185)
  - gi_risk_factors (Tree:5.30.65, Frequency: 0.0000)
  - cancer_risk_factors (Tree:5.30.66, Frequency: 0.0000)
  - derm_risk_factors (Tree:5.30.67, Frequency: 0.0000)
  - neurological_risk_factors (Tree:5.30.68, Frequency: 0.0000)
    - cerebral_vascular_risk_factors (Tree:5.30.68.62, Frequency: 0.0002)
    - epilepsy_risk_factors (Tree:5.30.68.63, Frequency: 0.0000)
  - congenital_risk_factors (Tree:5.30.69, Frequency: 0.0000)
  - pulmonary_risk_factors (Tree:5.30.70, Frequency: 0.0000)
  - trauma_risk_factors (Tree:5.30.71, Frequency: 0.0000)
  - abuse_risk_factors (Tree:5.30.72, Frequency: 0.0000)
  - psychological_risk_factors (Tree:5.30.73, Frequency: 0.0000)
    - home_risk_factors (Tree:5.30.73.64, Frequency: 0.0007)
    - work_risk_factors (Tree:5.30.73.65, Frequency: 0.0005)
- changes_to_admission_note (Tree:5.31, Frequency: 0.0002)
- hospital_course (Tree:5.32, Frequency: 0.0049)
  - clinical_trend (Tree:5.32.74, Frequency: 0.0000)
  - hospital_course_by_problem (Tree:5.32.75, Frequency: 0.0000)
  - hospital_course_by_location (Tree:5.32.76, Frequency: 0.0000)
    - emergency_department_course (Tree:5.32.76.66, Frequency: 0.0000)
    - nursery_course (Tree:5.32.76.67, Frequency: 0.0000)
    - nicu_course (Tree:5.32.76.68, Frequency: 0.0000)
    - micu_course (Tree:5.32.76.69, Frequency: 0.0000)
  - hospital_course_by_system (Tree:5.32.77, Frequency: 0.0002)
    - general_course (Tree:5.32.77.70, Frequency: 0.0003)
    - derm_course (Tree:5.32.77.71, Frequency: 0.0005)
    - heent_course (Tree:5.32.77.72, Frequency: 0.0005)
    - lymphatic_course (Tree:5.32.77.73, Frequency: 0.0010)
    - hematology_course (Tree:5.32.77.74, Frequency: 0.0000)
    - cardiovascular_course (Tree:5.32.77.75, Frequency: 0.0015)
    - vascular_course (Tree:5.32.77.76, Frequency: 0.0000)
    - thorax_course (Tree:5.32.77.77, Frequency: 0.0000)
    - gastrointestinal_course (Tree:5.32.77.78, Frequency: 0.0003)
    - genitourinary_course (Tree:5.32.77.79, Frequency: 0.0003)
    - neuro_psych_course (Tree:5.32.77.80, Frequency: 0.0000)
    - musculoskeletal_rheumatological_course (Tree:5.32.77.81, Frequency: 0.0000)

immunologic_course (Tree:5.32.77.82, Frequency: 0.0000)
allergies_course (Tree:5.32.77.83, Frequency: 0.0000)
endocrine_metabolic_course (Tree:5.32.77.84, Frequency: 0.0000)
infectious_disease_course (Tree:5.32.77.85, Frequency: 0.0000)
oncology_course (Tree:5.32.77.86, Frequency: 0.0000)
pulmonary_course (Tree:5.32.77.87, Frequency: 0.0012)
fluid_electrolyte_nutrition_course (Tree:5.32.77.88, Frequency: 0.0002)
fluid_electrolyte_nutrition_gastrointestinal_cours (Tree:5.32.77.88-33, Frequency: 0.0000) , composite concept
transport_history (Tree:5.33, Frequency: 0.0002)
family_and_social_history (Tree:5.34, Frequency: 0.0549)
personal_and_social_history (Tree:5.34.78, Frequency: 0.7317)
infectious_disease_exposure_history (Tree:5.34.78.89, Frequency: 0.0000)
education_history (Tree:5.34.78.90, Frequency: 0.0056)
occupational_history (Tree:5.34.78.91, Frequency: 0.0241)
occupational_environmental_history (Tree:5.34.78.91-18, Frequency: 0.0000) , composite concept
environmental_history (Tree:5.34.78.92, Frequency: 0.0000)
substance_use (Tree:5.34.78.93, Frequency: 0.3246)
family_enviroment (Tree:5.34.78.94, Frequency: 0.0773)
sexual_activity_history (Tree:5.34.78.95, Frequency: 0.0012)
habits (Tree:5.34.78.96, Frequency: 0.0073)
diet_history (Tree:5.34.78.97, Frequency: 0.0000)
travel_history (Tree:5.34.78.98, Frequency: 0.0003)
religious_history (Tree:5.34.78.99, Frequency: 0.0000)
psychological_stressors (Tree:5.34.78.100, Frequency: 0.0003)
family_medical_history (Tree:5.34.79, Frequency: 0.7004)
family_history_by_relationship (Tree:5.34.79.102, Frequency: 0.0019)
family_history_by_category (Tree:5.34.79.103, Frequency: 0.0002)
past_medical_history (Tree:5.35, Frequency: 0.7992)
past_medical_history_and_physical_examination (Tree:5.35-24, Frequency: 0.0576) , composite concept
chronic_illnesses (Tree:5.35.80, Frequency: 0.0000)
cardiac_history (Tree:5.35.81, Frequency: 0.0048)
hemetologic_oncologic_history (Tree:5.35.82, Frequency: 0.0000)
oncologic_history (Tree:5.35.82.104, Frequency: 0.0032)
hematologic_history (Tree:5.35.82.105, Frequency: 0.0000)
gi_history (Tree:5.35.83, Frequency: 0.0000)
past_surgical_history (Tree:5.35.84, Frequency: 0.1551)
past_anesthesia_history (Tree:5.35.84.106, Frequency: 0.0020)
past_gi_surgery (Tree:5.35.84.107, Frequency: 0.0000)
blood_products (Tree:5.35.85, Frequency: 0.0032)
radiation_therapy_or_exposure_history (Tree:5.35.86, Frequency: 0.0000)
invasive_diagnostic_procedure_history (Tree:5.35.87, Frequency: 0.0000)
trauma_history (Tree:5.35.88, Frequency: 0.0000)
injury_history (Tree:5.35.89, Frequency: 0.0022)
prior_hosptilization (Tree:5.35.90, Frequency: 0.0192)
reproductive_history (Tree:5.35.91, Frequency: 0.0039)
obstetric_history (Tree:5.35.91.108, Frequency: 0.0042)
gynecologic_history (Tree:5.35.91.109, Frequency: 0.0019)
genetic_diseases (Tree:5.35.92, Frequency: 0.0002)
outpatient_history (Tree:5.35.93, Frequency: 0.0000)
growth_development (Tree:5.35.94, Frequency: 0.0005)
social_development (Tree:5.35.94.110, Frequency: 0.0000)
fine_motor_development (Tree:5.35.94.111, Frequency: 0.0000)
gross_motor_development (Tree:5.35.94.112, Frequency: 0.0002)
speech_development (Tree:5.35.94.113, Frequency: 0.0000)
prenatal_development (Tree:5.35.94.114, Frequency: 0.0000)

- childhood_disease_history (Tree:5.35.95, Frequency: 0.0000)
- congenital_disease_history (Tree:5.35.96, Frequency: 0.0000)
- cognitive_history (Tree:5.35.97, Frequency: 0.0000)
- seizure_history (Tree:5.35.98, Frequency: 0.0014)
- blood_type (Tree:5.35.99, Frequency: 0.0003)
- psychiatric_history (Tree:5.35.100, Frequency: 0.0008)
- vascular_medical_history (Tree:5.35.101, Frequency: 0.0000)

health_maintenance (Tree:5.36, Frequency: 0.0197)
- immunizations (Tree:5.36.102, Frequency: 0.0525)
  - childhood_immunizations (Tree:5.36.102.115, Frequency: 0.0000)
- screening_tests (Tree:5.36.103, Frequency: 0.0019)
  - cancer_screening_tests (Tree:5.36.103.116, Frequency: 0.0000)
  - diabetes_screen (Tree:5.36.103.117, Frequency: 0.0000)
  - lipid_screen (Tree:5.36.103.118, Frequency: 0.0000)
  - osteoporosis_screen (Tree:5.36.103.119, Frequency: 0.0005)

medications (Tree:5.37, Frequency: 0.6850)
- medication_history (Tree:5.37.105, Frequency: 0.0005)
- medications_by_situation (Tree:5.37.106, Frequency: 0.0000)
  - medications_outside_hospital (Tree:5.37.106.120, Frequency: 0.0000)
  - medications_at_transfer (Tree:5.37.106.121, Frequency: 0.0014)
  - admission_medications (Tree:5.37.106.122, Frequency: 0.0053)
  - inpatient_medications (Tree:5.37.106.123, Frequency: 0.0007)
  - outpatient_medications (Tree:5.37.106.124, Frequency: 0.0150)
  - discharge_medications (Tree:5.37.106.125, Frequency: 0.0024)
  - current_medications (Tree:5.37.106.126, Frequency: 0.0435)
  - postoperative_medications (Tree:5.37.106.127, Frequency: 0.0000)
  - labor_medications (Tree:5.37.106.128, Frequency: 0.0000)
- medications_by_type (Tree:5.37.107, Frequency: 0.0000)
  - immunosuppressants_medications (Tree:5.37.107.129, Frequency: 0.0000)
  - chemotherapeutics (Tree:5.37.107.130, Frequency: 0.0031)
  - antibiotics (Tree:5.37.107.131, Frequency: 0.0073)
  - alternative_therapies (Tree:5.37.107.132, Frequency: 0.0000)

allergies_and_adverse_reactions (Tree:5.38, Frequency: 0.6811)
- medication_allergies (Tree:5.38.108, Frequency: 0.0173)
- food_allergies (Tree:5.38.109, Frequency: 0.0000)
- environmental_allergies (Tree:5.38.110, Frequency: 0.0000)
- contrast_allergies (Tree:5.38.111, Frequency: 0.0000)
- latex_allergy (Tree:5.38.112, Frequency: 0.0000)

review_of_systems (Tree:5.39, Frequency: 0.7062)

ros_pmh_familyhx_socialhx (Tree:5.39-34, Frequency: 0.0024) , composite concept
- general (Tree:5.39.113, Frequency: 0.3269)
- skin (Tree:5.39.114, Frequency: 0.3441)
- heent (Tree:5.39.115, Frequency: 0.0608)
  - head_neck (Tree:5.39.115.133, Frequency: 0.0034)
  - ent (Tree:5.39.115.134, Frequency: 0.2066)
  - eye_ear_nose_throat (Tree:5.39.115.134-20, Frequency: 0.0003) , composite concept
  - ophthalmologic (Tree:5.39.115.135, Frequency: 0.2085)
  - oropharynx (Tree:5.39.115.136, Frequency: 0.0002)
  - face (Tree:5.39.115.137, Frequency: 0.0015)
- lymphatic (Tree:5.39.116, Frequency: 0.0005)
- hematology (Tree:5.39.117, Frequency: 0.0503)
- hematologic_lymphatic_immunologic (Tree:5.39.117-17, Frequency: 0.0000) , composite concept
- hematologic_lymphatic (Tree:5.39.117-14, Frequency: 0.1293) , composite concept
- hematologic_oncologic (Tree:5.39.117-10, Frequency: 0.0003) , composite concept
  - coagulation_history (Tree:5.39.117.138, Frequency: 0.0095)
- cardiovascular (Tree:5.39.118, Frequency: 0.2829)

        chest_pain_ros (Tree:5.39.118.139, Frequency: 0.0000)
      vascular (Tree:5.39.119, Frequency: 0.0328)
      thorax (Tree:5.39.120, Frequency: 0.0002)
        chest (Tree:5.39.120.140, Frequency: 0.0528)
        back (Tree:5.39.120.141, Frequency: 0.0008)
        breast (Tree:5.39.120.142, Frequency: 0.0053)
      gastrointestinal (Tree:5.39.121, Frequency: 0.3364)
        abdomen (Tree:5.39.121.144, Frequency: 0.0049)
        bowel_habits (Tree:5.39.121.145, Frequency: 0.0073)
        dysphagia (Tree:5.39.121.146, Frequency: 0.0219)
        jaundice_history (Tree:5.39.121.147, Frequency: 0.0000)
        nausea_vomiting_regurgitation (Tree:5.39.121.148, Frequency: 0.0000)
      genitourinary (Tree:5.39.122, Frequency: 0.3026)
        urinary_tract (Tree:5.39.122.149, Frequency: 0.0000)
        genital (Tree:5.39.122.150, Frequency: 0.0015)
      neuro_psych (Tree:5.39.123, Frequency: 0.0002)
        psychiatric (Tree:5.39.123.151, Frequency: 0.1808)
        neuropsychiatric_medication_administration_history (Tree:5.39.123.152, Frequency: 0.0000)
        headache (Tree:5.39.123.153, Frequency: 0.0574)
        neurological (Tree:5.39.123.154, Frequency: 0.2622)
        sleep_disturbances (Tree:5.39.123.155, Frequency: 0.0000)
        mental_status (Tree:5.39.123.156, Frequency: 0.0109)
        speech (Tree:5.39.123.157, Frequency: 0.0014)
        memory (Tree:5.39.123.158, Frequency: 0.0005)
      musculoskeletal_rheumatological (Tree:5.39.124, Frequency: 0.0002)
        musculoskeletal (Tree:5.39.124.160, Frequency: 0.2800)
        rheumatological (Tree:5.39.124.161, Frequency: 0.0002)
      immunologic (Tree:5.39.125, Frequency: 0.0865)
      extremities (Tree:5.39.126, Frequency: 0.0209)
      allergies (Tree:5.39.127, Frequency: 0.0119)
      allergy_lymphatic (Tree:5.39.127-13, Frequency: 0.0000) , composite concept
      endocrine_metabolic (Tree:5.39.128, Frequency: 0.0000)
        endocrine (Tree:5.39.128.163, Frequency: 0.1893)
        metabolic (Tree:5.39.128.164, Frequency: 0.0012)
      infectious_disease (Tree:5.39.129, Frequency: 0.0049)
      congenital (Tree:5.39.130, Frequency: 0.0000)
      oncology (Tree:5.39.131, Frequency: 0.0002)
      pulmonary (Tree:5.39.132, Frequency: 0.3196)
      renal (Tree:5.39.133, Frequency: 0.0073)
objective_data (Tree:6, Frequency: 0.0155)
   physical_examination (Tree:6.40, Frequency: 0.9438)
      general_exam (Tree:6.40.135, Frequency: 0.6136)
      vital_signs (Tree:6.40.136, Frequency: 0.5601)
        temperature (Tree:6.40.136.165, Frequency: 0.3239)
        heart_rate (Tree:6.40.136.166, Frequency: 0.2659)
        respiratory_rate (Tree:6.40.136.167, Frequency: 0.2685)
        weight (Tree:6.40.136.168, Frequency: 0.1064)
        weight_percentile (Tree:6.40.136.169, Frequency: 0.0000)
        height (Tree:6.40.136.170, Frequency: 0.0160)
        height_percentile (Tree:6.40.136.171, Frequency: 0.0000)
        head_circumference (Tree:6.40.136.172, Frequency: 0.0000)
        head_percentile (Tree:6.40.136.173, Frequency: 0.0000)
        length (Tree:6.40.136.174, Frequency: 0.0014)
        occipitofrontal_diameter (Tree:6.40.136.175, Frequency: 0.0000)
        blood_pressure (Tree:6.40.136.176, Frequency: 0.3375)
        oxygen_saturation (Tree:6.40.136.177, Frequency: 0.0598)

pain (Tree:6.40.136.178, Frequency: 0.0316)
body_mass_index (Tree:6.40.136.179, Frequency: 0.0008)
neck_circumference (Tree:6.40.136.180, Frequency: 0.0000)
body_surface_area (Tree:6.40.136.181, Frequency: 0.0000)
fluid_balance (Tree:6.40.136.182, Frequency: 0.0010)
asa_class (Tree:6.40.136.183, Frequency: 0.0000)
ballard_exam (Tree:6.40.137, Frequency: 0.0000)
apgar_score (Tree:6.40.138, Frequency: 0.0002)
1_minute_apgar (Tree:6.40.138.186, Frequency: 0.0000)
5_minute_apgar (Tree:6.40.138.187, Frequency: 0.0000)
10_minute_apgar (Tree:6.40.138.188, Frequency: 0.0000)
anatomic_exam (Tree:6.40.139, Frequency: 0.0126)
derm_exam (Tree:6.40.139.190, Frequency: 0.3954)
skin_and_breast_exam (Tree:6.40.139.190-31, Frequency: 0.0007) , composite concept
head_neck_exam (Tree:6.40.139.191, Frequency: 0.0068)
endocrine_exam (Tree:6.40.139.192, Frequency: 0.0025)
hematologic_lymphatic_oncologic_exam (Tree:6.40.139.193, Frequency: 0.0000)
cardiovascular_exam (Tree:6.40.139.194, Frequency: 0.7641)
chest_exam (Tree:6.40.139.195, Frequency: 0.1988)
gastrointestinal_exam (Tree:6.40.139.196, Frequency: 0.4877)
gastrointestinal_genitourinary_exam (Tree:6.40.139.196-15, Frequency: 0.0007) , composite concept
anorectal_exam (Tree:6.40.139.197, Frequency: 0.0000)
genitourinary_exam (Tree:6.40.139.198, Frequency: 0.1186)
genital_rectal_exam (Tree:6.40.139.198-7, Frequency: 0.0029) , composite concept
neuro_psych_exam (Tree:6.40.139.199, Frequency: 0.0000)
musculoskeletal_rheumatological_exam (Tree:6.40.139.200, Frequency: 0.1105)
musculoskeletal_extremity_exam (Tree:6.40.139.200-26, Frequency: 0.3772) , composite concept
vascular_exam (Tree:6.40.139.201, Frequency: 0.0360)
extremity_exam (Tree:6.40.139.202, Frequency: 0.4182)
extremities_hip_exam (Tree:6.40.139.202-22, Frequency: 0.0000) , composite concept
extremity_and_joint_exam (Tree:6.40.139.202-0, Frequency: 0.0000) , composite concept
non_anatomic_examinations (Tree:6.40.140, Frequency: 0.0000)
wound_exam (Tree:6.40.140.203, Frequency: 0.0051)
ulcer_exam (Tree:6.40.140.204, Frequency: 0.0051)
burn_exam (Tree:6.40.140.205, Frequency: 0.0019)
laboratory_and_radiology_data (Tree:6.41, Frequency: 0.5331)
telemetry (Tree:6.41.142, Frequency: 0.0044)
laboratory_data (Tree:6.41.144, Frequency: 0.2420)
common_chemistries (Tree:6.41.144.206, Frequency: 0.0003)
csf_analysis (Tree:6.41.144.207, Frequency: 0.0000)
urinalysis (Tree:6.41.144.208, Frequency: 0.0921)
ascitic_fluid_or_peritoneal_aspirate_analysis (Tree:6.41.144.209, Frequency: 0.0000)
joint_fluid_analysis (Tree:6.41.144.210, Frequency: 0.0000)
pericardial_fluid_analysis (Tree:6.41.144.211, Frequency: 0.0000)
pleural_fluid_analysis (Tree:6.41.144.212, Frequency: 0.0000)
skin_testing (Tree:6.41.144.213, Frequency: 0.0002)
toxicological_studies (Tree:6.41.144.214, Frequency: 0.0000)
microbiology (Tree:6.41.144.215, Frequency: 0.0005)
biopsies_consistent_with_malignant_neoplastic_diso (Tree:6.41.144.216, Frequency: 0.0000)
electron_microscopy (Tree:6.41.144.217, Frequency: 0.0000)
cardiovascular_studies (Tree:6.41.145, Frequency: 0.0008)
radiological_cardiological_studies (Tree:6.41.145-25, Frequency: 0.0000) , composite concept
electrophysiology_results (Tree:6.41.145.218, Frequency: 0.0000)
echocardiogram (Tree:6.41.145.219, Frequency: 0.0379)
valvular_data (Tree:6.41.145.220, Frequency: 0.0019)

electrocardiogram (Tree:6.41.145.221, Frequency: 0.2508)
stress_test (Tree:6.41.145.222, Frequency: 0.0036)
pulmonary_tests (Tree:6.41.146, Frequency: 0.0000)
sleep_studies (Tree:6.41.146.224, Frequency: 0.0003)
abg (Tree:6.41.146.225, Frequency: 0.0146)
pfts (Tree:6.41.146.226, Frequency: 0.0054)
neurological_studies (Tree:6.41.147, Frequency: 0.0000)
eeg (Tree:6.41.147.227, Frequency: 0.0029)
visual_and_auditory_evoked_potentials (Tree:6.41.147.228, Frequency: 0.0000)
electronystagmogram_and_vestibular_function_tests (Tree:6.41.147.229, Frequency: 0.0000)
electromyography (Tree:6.41.147.230, Frequency: 0.0000)
nerve_conduction_velocities (Tree:6.41.147.231, Frequency: 0.0000)
eyes_formal_visual_field_testing (Tree:6.41.148, Frequency: 0.0000)
radiographic_studies (Tree:6.41.149, Frequency: 0.0177)
imaging_findings (Tree:6.41.149.232, Frequency: 0.0037)
imaging_impression (Tree:6.41.149.233, Frequency: 0.0000)
types_of_radiographic_studies (Tree:6.41.149.234, Frequency: 0.0000)
endoscopic_procedure (Tree:6.41.150, Frequency: 0.0000)
endoscopic_findings (Tree:6.41.150.235, Frequency: 0.0000)
endoscopic_impressions (Tree:6.41.150.236, Frequency: 0.0000)
respiratory_tract_endoscopy (Tree:6.41.150.237, Frequency: 0.0000)
gastrointestinal_endoscopy (Tree:6.41.150.238, Frequency: 0.0002)
peritoneoscopy (Tree:6.41.150.239, Frequency: 0.0000)
cystoscopy_and_cystometrogram (Tree:6.41.150.240, Frequency: 0.0000)
culdoscopy_and_culdocentesis (Tree:6.41.150.241, Frequency: 0.0000)
culposcopy (Tree:6.41.150.242, Frequency: 0.0000)
arthroscopy (Tree:6.41.150.243, Frequency: 0.0000)
psychological_data (Tree:6.41.151, Frequency: 0.0000)
educational_testing (Tree:6.41.151.244, Frequency: 0.0000)
developmental_testing (Tree:6.41.151.245, Frequency: 0.0000)
functional_assessment (Tree:6.41.152, Frequency: 0.0000)
pathologic_data (Tree:6.41.153, Frequency: 0.0078)
biopsy_nervous_system (Tree:6.41.153.246, Frequency: 0.0000)
respiratory_tract_histopathological_studies (Tree:6.41.153.247, Frequency: 0.0000)
breast_biopsy_aspirate (Tree:6.41.153.248, Frequency: 0.0000)
pituitary_biopsy (Tree:6.41.153.249, Frequency: 0.0000)
thyroid_biopsy (Tree:6.41.153.250, Frequency: 0.0000)
parathyroid_biospy (Tree:6.41.153.251, Frequency: 0.0000)
adrenal_biopsy (Tree:6.41.153.252, Frequency: 0.0000)
gonadal_biopsy (Tree:6.41.153.253, Frequency: 0.0000)
myocardial_biopsy (Tree:6.41.153.254, Frequency: 0.0000)
pericardial_biopsy (Tree:6.41.153.255, Frequency: 0.0000)
blood_vessel_biopsy (Tree:6.41.153.256, Frequency: 0.0000)
esophagus_biopsy (Tree:6.41.153.257, Frequency: 0.0000)
stomach_biopsy (Tree:6.41.153.258, Frequency: 0.0003)
si_biopsy (Tree:6.41.153.259, Frequency: 0.0087)
colon_biopsy (Tree:6.41.153.260, Frequency: 0.0000)
rectal_biopsy (Tree:6.41.153.261, Frequency: 0.0000)
liver_biopsy (Tree:6.41.153.262, Frequency: 0.0007)
pancreas_biospy (Tree:6.41.153.263, Frequency: 0.0000)
peritoneum_biopsy (Tree:6.41.153.264, Frequency: 0.0000)
kidney_biopsy (Tree:6.41.153.265, Frequency: 0.0003)
ureter_or_urinary_bladder_biopsy (Tree:6.41.153.266, Frequency: 0.0000)
uterine_biopsy (Tree:6.41.153.267, Frequency: 0.0000)
prostate_biopsy (Tree:6.41.153.268, Frequency: 0.0002)
hematopoetic_and_reticuloendothelial_histopatholog (Tree:6.41.153.269, Frequency: 0.0000)

skin_biopsy (Tree:6.41.153.270, Frequency: 0.0000)
musculoskeletal_histopathological_studies (Tree:6.41.153.271, Frequency: 0.0000)
orders (Tree:7, Frequency: 0.0048)
instructions (Tree:8, Frequency: 0.0002)
discharge_instructions (Tree:8.42, Frequency: 0.0000)
activity (Tree:8.43, Frequency: 0.0048)
discharge_activity (Tree:8.44, Frequency: 0.0000)
follow_up (Tree:9, Frequency: 0.0025)
discharge_followup (Tree:9.45, Frequency: 0.0007)
functional_status (Tree:10, Frequency: 0.0071)
problem_list (Tree:11, Frequency: 0.2245)
discharge_problem_list (Tree:11.46, Frequency: 0.0000)
ed_problem (Tree:11.47, Frequency: 0.0000)
health_status (Tree:12, Frequency: 0.0000)
assessment_and_plan (Tree:13, Frequency: 0.3128)
summary (Tree:13.48, Frequency: 0.0401)
assessment (Tree:13.49, Frequency: 0.4041)
analysis (Tree:13.50, Frequency: 0.2398)
plan (Tree:13.51, Frequency: 0.4442)
treatment_goals (Tree:13.51.154, Frequency: 0.0049)
patient_education (Tree:13.51.155, Frequency: 0.0003)
plan_by_type (Tree:13.51.156, Frequency: 0.0000)
diagnostic_plan,_workup (Tree:13.51.156.272, Frequency: 0.0000)
therapeutic_plan (Tree:13.51.156.273, Frequency: 0.0017)
anesthesia_plan (Tree:13.51.156.274, Frequency: 0.0012)
wound_care_plan (Tree:13.51.156.275, Frequency: 0.0080)
prophylaxis_plan (Tree:13.51.156.276, Frequency: 0.0427)
health_maintenance_plan (Tree:13.51.156.277, Frequency: 0.0005)
disposition_plan (Tree:13.51.156.278, Frequency: 0.0197)
transplant_plan (Tree:13.51.156.279, Frequency: 0.0019)
plan_by_system (Tree:13.51.157, Frequency: 0.0034)
general_plan (Tree:13.51.157.281, Frequency: 0.0027)
skin_plan (Tree:13.51.157.282, Frequency: 0.0031)
heent_plan (Tree:13.51.157.283, Frequency: 0.0003)
lypmhatic_plan (Tree:13.51.157.284, Frequency: 0.0000)
hematology_plan (Tree:13.51.157.285, Frequency: 0.0292)
hematologic_lymphatic_plan (Tree:13.51.157.285-9, Frequency: 0.0000) , composite concept
oncology_plan (Tree:13.51.157.286, Frequency: 0.0037)
hematologic_oncologic_plan (Tree:13.51.157.286-8, Frequency: 0.0000) , composite concept
cardiovascular_plan (Tree:13.51.157.287, Frequency: 0.0790)
thorax_plan (Tree:13.51.157.288, Frequency: 0.0000)
gastrointestinal_plan (Tree:13.51.157.289, Frequency: 0.0702)
genitourinary_plan (Tree:13.51.157.290, Frequency: 0.0034)
neuro_psych_plan (Tree:13.51.157.291, Frequency: 0.0003)
musculoskeletal_rheumatological_plan (Tree:13.51.157.292, Frequency: 0.0000)
endocrine_metabolic_plan (Tree:13.51.157.293, Frequency: 0.0000)
infectious_disease_plan (Tree:13.51.157.294, Frequency: 0.0627)
congenital_plan (Tree:13.51.157.295, Frequency: 0.0000)
pulmonary_plan (Tree:13.51.157.296, Frequency: 0.0639)
immunologic_plan (Tree:13.51.157.297, Frequency: 0.0000)
extremity_plan (Tree:13.51.157.298, Frequency: 0.0008)
renal_plan (Tree:13.51.157.299, Frequency: 0.0477)
fluid_electrolyte_nutrition_plan (Tree:13.51.157.300, Frequency: 0.0554)
fluid_electrolyte_nutrition_gastrointestinal_plan (Tree:13.51.157.300-32, Frequency: 0.0007) , composite concept
findings (Tree:13.53, Frequency: 0.0015)

        operative_findings (Tree:13.53.158, Frequency: 0.0000)
closing (Tree:14, Frequency: 0.0003)
    carbon_copy (Tree:14.54, Frequency: 0.0015)
    report_status (Tree:14.55, Frequency: 0.0020)
    total_time (Tree:14.56, Frequency: 0.0000)
        critical_care_time (Tree:14.56.159, Frequency: 0.0002)
        counseling_time (Tree:14.56.160, Frequency: 0.0000)
operative_procedural (Tree:15, Frequency: 0.0000)
    estimated_blood_loss (Tree:15.57, Frequency: 0.0002)
    counts (Tree:15.58, Frequency: 0.0070)
    type_of_procedure (Tree:15.59, Frequency: 0.0005)
    technique (Tree:15.60, Frequency: 0.0000)
    anesthesia (Tree:15.61, Frequency: 0.0005)
    complications (Tree:15.62, Frequency: 0.0029)
description (Tree:16, Frequency: 0.0012)
technical_quality (Tree:17, Frequency: 0.0000)
addendum (Tree:18, Frequency: 0.0034)
    attending_addendum (Tree:18.63, Frequency: 0.0000)
report (Tree:19, Frequency: 0.0248)
references (Tree:20, Frequency: 0.0005)
    recommended_reading (Tree:20.64, Frequency: 0.0000)
attestation (Tree:21, Frequency: 0.0015)
    house_staff_attestation (Tree:21.65, Frequency: 0.0000)
    attending_attestation (Tree:21.66, Frequency: 0.0022)
author (Tree:22, Frequency: 0.0984)
data_base (Tree:23, Frequency: 0.0002)
consent (Tree:24, Frequency: 0.0019)
comment (Tree:25, Frequency: 0.0754)
resuscitation (Tree:26, Frequency: 0.0034)

APPENDIX C

IMPORTANT SECTION LIST FOR EVALUATION

Chief complaint
History present illness
Past medical history
Family medical history
    Medical history of 1st degree relatives: mother, father, siblings, and children
Health maintenance, if it exists as a grouped section
Personal and social history
    Substance use or a more granular section (ethanol, tobacco, or drug use)
Medications
Allergies/Adverse reactions
Review of Systems
    Any identified sections here. Only add tags for sections tagged by the authors. Common sections would parallel physical exam subsections (where appropriate).
Physical exam
    Vital signs
    General
    Dermatologic
    Lymph nodes
    HEENT (or any of Head, Eye, Ear, Nose, or Throat)
    Cardiovascular
    Any of gastrointestinal, rectal, or abdominal
    Any of pulmonary/thorax/chest
    Any of genitourinary, pelvic, or genital
    Neurological
    Psychological
    Any of musculoskeletal/rheumatological, or subcomponents back, costrovertebral angle, or spine
    Extremities
Labs and radiology
    Score subsections identified by SecTag, but do not need to labeled unidentified sections
Assessment and plan
    Of note, in the terminology, “analysis”, “assessment”, and “plan” and their subcategories are components of “Assessment and plan.” Score identified subsections.

REFERENCES

1. Denny J, Peterson J. Identifying QT prolongation from ECG impressions using Natural Language Processing and Negation Detection. Medinfo 2007:(accepted).

2. Denny JC, Irani PR, Wehbe FH, Smithers JD, Spickard A, 3rd. The KnowledgeMap project: development of a concept-based medical school curriculum database. AMIA Annual Symposium proceedings / AMIA Symposium 2003:195-9.

3. Denny JC, Smithers JD, Armstrong B, Spickard A, 3rd. "Where do we teach what?" Finding broad concepts in the medical school curriculum. J Gen Intern Med 2005;20(10):943-6.

4. Denny JC, Smithers JD, Miller RA, Spickard A, 3rd. "Understanding" medical school curriculum content using KnowledgeMap. J Am Med Inform Assoc 2003;10(4):351-62.

5. Denny JC, Spickard A, Miller RA, et al. Identifying UMLS concepts from ECG Impressions using KnowledgeMap. AMIA Annual Symposium proceedings / AMIA Symposium 2005:196-200.

6. Brown SH, Lincoln M, Hardenbrook S, et al. Derivation and evaluation of a document-naming nomenclature. J Am Med Inform Assoc 2001;8(4):379-90.

7. Johnson S. Preface. In: Dictionary. London; 1755.

8. Davis LS, Collen MF, Rubin L, Van Brunt EE. Computer-stored medical record. Computers and biomedical research, an international journal 1968;1(5):452-69.

9. Garrett LE, Jr., Hammond WE, Stead WW. The effects of computerized medical records on provider efficiency and quality of care. Methods of information in medicine 1986;25(3):151-7.

10. Leslie SJ, Hartswood M, Meurig C, et al. Clinical decision support software for management of chronic heart failure: development and evaluation. Comput Biol Med 2006;36(5):495-506.

11. Kuperman GJ, Bobb A, Payne TH, et al. Medication-related clinical decision support in computerized provider order entry systems: a review. J Am Med Inform Assoc 2007;14(1):29-40.

12. Meystre S, Haug PJ. Automation of a problem list using natural language processing. BMC Med Inform Decis Mak 2005;5:30.

13. Shultz E, Rosenbloom T, Kiepek W, et al. Quill: a novel approach to structured reporting. AMIA Annual Symposium proceedings / AMIA Symposium 2003:1074.

14. Rosenbloom ST, Grande J, Geissbuhler A, Miller RA. Experience in implementing inpatient clinical note capture via a provider order entry system. J Am Med Inform Assoc 2004;11(4):310-5.

15. Johnson KB, Cowan J. Clictate: a computer-based documentation tool for guideline-based care. J Med Syst 2002;26(1):47-60.

16. Lieberman MI, Ricciardi TN. Using NLP to extract concepts from chief complaints. AMIA Annual Symposium proceedings / AMIA Symposium 2005:1029.

17. Heinze DT, Morsch ML, Holbrook J. Mining free-text medical records. Proceedings / AMIA Annual Symposium 2001:254-8.

18. Garrison FH. Contributions to the history of medicine : from the Bulletin of the New York Academy of Medicine 1925-1935. New York and London: Hafner Publishing Company; 1966.

19. Von Staden H. Herophilus : the art of medicine in early Alexandria : edition, translation and essays. Cambridge ; New York: Cambridge University Press; 1989.

20. Vesalius A. De humani corporis fabrica. Lugduni,: Apud Ioan. Tornaesium; 1552.

21. Thomas Sydenham. (Accessed at http://en.wikipedia.org/wiki/Thomas_Sydenham.)

22. Carling PC. Sir Thomas Sydenham revisited. Clin Infect Dis 2001;32(10):1519.

23. Dewhurst K. Thomas Sydenham (1624-1689) Reformer of Clinical Medicine. Med Hist 1962;6(2):101-18.

24. Libby W. History of medicine in its salient features. Boston: Houghton Mifflin; 1922.

25. René Laennec. (Accessed July 20, 2007, at http://en.wikipedia.org/wiki/Laennec.)

26. William Osler. (Accessed July 20, 2007, at http://en.wikipedia.org/wiki/William_Osler.)

27. Norris GW, Landis HRM, Krumbhaar EB, Montgomery CM. Diseases of the chest and the principles of physical diagnosis. 5th ed. Philadelphia, London,: W.B. Saunders company; 1933.

28. Wartenberg R. The examination of reflexes, a simplification. Chicago,: The Year Book Publishers; 1945.

29. Burch GE. A primer of venous pressure. Philadelphia,: Lea & Febiger; 1950.

30. Walker H. Physical diagnosis. St. Louis,: Mosby; 1952.

31. Fowler NO. Physical diagnosis of heart disease. New York,: Macmillan; 1962.

32. Martini P. Principles and practice of physical diagnosis. 3d ed. Philadelphia,: Lippincott; 1962.

33. Perloff JK. Physical examination of the heart and circulation. Philadelphia: Saunders; 1982.

34. Swartz MH. Textbook of physical diagnosis : history and examination. 5th ed. Philadelphia: Saunders; 2006.

35. Evans LR, Bybee JR. Evaluation of Student Skills in Physical Diagnosis. J Med Educ 1965;40:199-204.

36. Wasson J, Sox HC, Jr., Tompkins RK, Walsh BT, Petersen J, Rund DA. Teaching physical diagnosis: the effect of a structured course taught by medical students. J Med Educ 1976;51(12):1014-5.

37. McGlynn TJ, Sayre A, Kennedy D. Physical diagnosis courses--a question of emphasis. J Fam Pract 1978;6(3):565-71.

38. Hunt DK, Badgett RG, Woodling AE, Pugh JA. Medical student career choice: do physical diagnosis preceptors influence decisions? Am J Med Sci 1995;310(1):19-23.

39. Hamann C, Volkan K, Fishman MB, Silvestri RC, Simon SR, Fletcher SW. How well do second-year students learn physical diagnosis? Observational study of an Objective Structured Clinical Examination (OSCE). BMC Med Educ 2002;2:1.

40. Fagan MJ, Griffith RA, Obbard L, O'Connor CJ. Improving the physical diagnosis skills of third-year medical students: a controlled trial of a literature-based curriculum. J Gen Intern Med 2003;18(8):652-5.

41. Nieman LZ, Cheng L, Hormann M, Farnie MA, Molony DA, Butler P. The impact of preclinical preceptorships on learning the fundamentals of clinical medicine and physical diagnosis skills. Acad Med 2006;81(4):342-6.

42. Bell DS, Greenes RA. Evaluation of UltraSTAR: performance of a collaborative structured data entry system. Proceedings / the Annual Symposium on Computer Application [sic] in Medical Care 1994:216-22.

43. Kahn CE, Jr., Wang K, Bell DS. Structured entry of radiology reports using World Wide Web technology. Radiographics 1996;16(3):683-91.

44. Institute of Medicine (U.S.). Committee on Quality of Health Care in America. Crossing the quality chasm : a new health system for the 21st century. Washington, D.C.: National Academy Press; 2001.

45. E/M History Criteria. (Accessed July 13, 2007, at http://www.fpnotebook.com/MAN3.htm.)

46. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. The New England journal of medicine 1991;324(6):370-6.

47. Leape LL, Brennan TA, Laird N, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. The New England journal of medicine 1991;324(6):377-84.

48. Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. Bmj 2000;320(7237):741-4.

49. Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. Jama 2001;286(4):415-20.

50. Holder AR. The importance of medical records. Jama 1974;228(1):118-9.

51. Niklas M. A flexible dictation system speeds reporting and cuts clerical time. Hospital management 1963;95:118-9.

52. Toussaint JL. Central dictation system promotes complete and legible medical records. Hospitals 1966;40(1):44-5.

53. Schwartz LH, Kijewski P, Hertogen H, Roossin PS, Castellino RA. Voice recognition in radiology reporting. AJR Am J Roentgenol 1997;169(1):27-9.

54. Zick RG, Olsen J. Voice recognition software versus a traditional transcription service for physician charting in the ED. Am J Emerg Med 2001;19(4):295-8.

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

56. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. J Am Med Inform Assoc 2005;12(4):448-57.

57. Friedman C, Shagina L, Lussier Y, Hripcsak G. Automated encoding of clinical documents based on natural language processing. J Am Med Inform Assoc 2004;11(5):392-402.

58. Jain NL, Friedman C. Identification of findings suspicious for breast cancer based on natural language processing of mammogram reports. Proc AMIA Annu Fall Symp 1997:829-33.

59. Sager N, Lyman M, Bucknall C, Nhan N, Tick LJ. Natural language processing and the representation of clinical data. J Am Med Inform Assoc 1994;1(2):142-60.

60. Huske-Kraus D. Text generation in clinical medicine--a review. Methods of information in medicine 2003;42(1):51-60.

61. Friedman C. Towards a comprehensive medical language processing system: methods and issues. Proc AMIA Annu Fall Symp 1997:595-9.

62. Fiszman M, Rindflesch TC, Kilicoglu H. Integrating a hypernymic proposition interpreter into a semantic processor for biomedical texts. AMIA Annual Symposium proceedings / AMIA Symposium 2003:239-43.

63. Wulfman CE, Rua M, Lane CD, Shortliffe EH, Fagan LM. Graphical access to medical expert systems: V. Integration with continuous-speech recognition. Methods of information in medicine 1993;32(1):33-46.

64. Aronsky D, Kasworm E, Jacobson JA, Haug PJ, Dean NC. Electronic screening of dictated reports to identify patients with do-not-resuscitate status. J Am Med Inform Assoc 2004;11(5):403-9.

65. Aphinyanaphongs Y, Tsamardinos I, Statnikov A, Hardin D, Aliferis CF. Text categorization models for high-quality article retrieval in internal medicine. J Am Med Inform Assoc 2005;12(2):207-16.

66. Duda S, Aliferis C, Miller R, Statnikov A, Johnson K. Extracting drug-drug interaction articles from MEDLINE to improve the content of drug databases. AMIA Annual Symposium proceedings / AMIA Symposium 2005:216-20.

67. Jain NL, Knirsch CA, Friedman C, Hripcsak G. Identification of suspected tuberculosis patients based on natural language processing of chest radiograph reports. Proc AMIA Annu Fall Symp 1996:542-6.

68. Cooper GF, Miller RA. An experiment comparing lexical and statistical methods for extracting MeSH terms from clinical free text. J Am Med Inform Assoc 1998;5(1):62-75.

69. Aronson AR. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. Proceedings / AMIA Annual Symposium 2001:17-21.

70. Hersh WR, Donohoe LC. SAPHIRE International: a tool for cross-language information retrieval. Proceedings / AMIA Annual Symposium 1998:673-7.

71. Nadkarni P, Chen R, Brandt C. UMLS concept indexing for production databases: a feasibility study. J Am Med Inform Assoc 2001;8(1):80-91.

72. Zou Q, Chu WW, Morioka C, Leazer GH, Kangarloo H. IndexFinder: a method of extracting key concepts from clinical texts for indexing. AMIA Annual Symposium proceedings / AMIA Symposium 2003:763-7.

73. Fan JW, Friedman C. Semantic Classification of Biomedical Concepts Using Distributional Similarity. J Am Med Inform Assoc 2007;14(4):467-77.

74. Hsieh Y, Hardardottir GA, Brennan PF. Linguistic analysis: terms and phrases used by patients in e-mail messages to nurses. Medinfo 2004;11(Pt 1):511-5.

75. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 2001;34(5):301-10.

76. Elkin PL, Brown SH, Bauer BA, et al. A controlled trial of automated classification of negation from clinical notes. BMC Med Inform Decis Mak 2005;5(1):13.

77. Elkin PL, Brown SH, Husser CS, et al. Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists. Mayo Clin Proc 2006;81(6):741-8.

78. Sneiderman CA, Rindflesch TC, Aronson AR. Finding the findings: identification of findings in medical literature using restricted natural language processing. Proc AMIA Annu Fall Symp 1996:239-43.

79. Srinivasan P, Rindflesch T. Exploring text mining from MEDLINE. Proceedings / AMIA Annual Symposium 2002:722-6.

80. Xu H, Anderson K, Grann VR, Friedman C. Facilitating cancer research using natural language processing of pathology reports. Medinfo 2004;11(Pt 1):565-72.

81. Chen L, Friedman C. Extracting phenotypic information from the literature via natural language processing. Medinfo 2004;11(Pt 2):758-62.

82. Hazlehurst B, Frost HR, Sittig DF, Stevens VJ. MediClass: A system for detecting and classifying encounter-based clinical events in any electronic medical record. J Am Med Inform Assoc 2005;12(5):517-29.

83. de Bruijn B, Cranney A, O'Donnell S, Martin JD, Forster AJ. Identifying wrist fracture patients with high accuracy by automatic categorization of X-ray reports. J Am Med Inform Assoc 2006;13(6):696-8.

84. Drucker H, Wu D, Vapnik V. Support vector machines for spam categorization. IEEE Trans Neural Network 1999;10(5):1048–54.

85. Joachims T. A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization In: Proceedings of the Fourteenth International Conference on Machine Learning; 1997; 1997. p. 143-51.

86. Joachims T. Transductive Inference for Text Classification using Support Vector Machines. In: Proceedings of the Sixteenth International Conference on Machine Learning; 1999; 1999. p. 200-9.

87. Meyer T, Whateley B. SpamBayes: Effective open-source, Bayesian based, email classification system. In: Conference on Email and Anti-Spam 2004; 2004; 2004.

88. Mitchell T. Bayesian Learning. In: Machine Learning. Boston: McGraw-Hill; 1997:154-200.

89. Sahami M, Dumais S, Heckerman D, Horvitz E. A Bayesian approach to filtering junk email. In: AAAI Workshop on Learning for Text Categorization; 1998; Madison, Wisconsin; 1998.

90. Sakkis G, Androutsopoulos I, Paliouras G, Karkaletsis V, Spyropoulos CD, Stamatopoulos P. A Memory-Based Approach to Anti-Spam Filtering for Mailing Lists. Information Retrieval 2003;6(1):49-73.

91. Zorkadis V, Karras DA, Panayotou M. Efficient information theoretic strategies for classifier combination, feature extraction and performance evaluation in improving false positives and false negatives for spam e-mail filtering. Neural Netw 2005;18(5-6):799-807.

92. Berrios DC, Cucina RJ, Fagan LM. Methods for semi-automated indexing for high precision information retrieval. J Am Med Inform Assoc 2002;9(6):637-52.

93. PubMed. (Accessed at http://www.ncbi.nlm.nih.gov/sites/entrez.)

94. Kolcz A, Alspector J. SVM-based filtering of e-mail spam with content-specific misclassification costs. In: Proc of TextDM'01 Workshop on Text Mining; 2001; 2001.

95. Pakhomov SV, Buntrock JD, Chute CG. Automating the assignment of diagnosis codes to patient encounters using example-based and machine learning techniques. J Am Med Inform Assoc 2006;13(5):516-25.

96. Yang Y, Chute CG. An application of least squares fit mapping to clinical classification. Proceedings / the Annual Symposium on Computer Application [sic] in Medical Care 1992:460-4.

97. Wilcox AB, Hripcsak G. The role of domain knowledge in automating medical text report classification. J Am Med Inform Assoc 2003;10(4):330-8.

98. Tsumoto S, Tanaka H. Induction of expert system rules based on rough sets and resampling methods. Medinfo 1995;8 Pt 1:861-5.

99. Statnikov A, Aliferis CF, Tsamardinos I, Hardin D, Levy S. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. Bioinformatics (Oxford, England) 2005;21(5):631-43.

100. Lee M, Wang W, Yu H. Exploring supervised and unsupervised methods to detect topics in biomedical text. BMC Bioinformatics 2006;7:140.

101. Muscatello DJ, Churches T, Kaldor J, et al. An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia. BMC Public Health 2005;5:141.

102. Cheng BY, Carbonell JG, Klein-Seetharaman J. Protein classification based on text document classification techniques. Proteins 2005;58(4):955-70.

103. Berrios DC. Automated indexing for full text information retrieval. Proceedings / AMIA Annual Symposium 2000:71-5.

104. Berrios DC, Kehler A, Fagan LM. Knowledge requirements for automated inference of medical textbook markup. Proceedings / AMIA Annual Symposium 1999:676-80.

105. Cucina RJ, Shah MK, Berrios DC, Fagan LM. Empirical formulation of a generic query set for clinical information retrieval systems. Medinfo 2001;10(Pt 1):181-5.

106. Kim DK, Fagan LM, Jones KT, Berrios DC, Yu VL. MYCIN II: design and implementation of a therapy reference with complex content-based indexing. Proceedings / AMIA Annual Symposium 1998:175-9.

107. The Leapfrog Group. (Accessed July 27, 2007, at http://www.leapfroggroup.org/.)

108. Evans DA, Cimino JJ, Hersh WR, Huff SM, Bell DS. Toward a medical-concept representation language. The Canon Group. J Am Med Inform Assoc 1994;1(3):207-17.

109. Masys DR. Of codes and keywords: standards for biomedical nomenclature. Acad Med 1990;65(10):627-9.

110. Ozbolt JG, Fruchtnicht JN, Hayden JR. Toward data standards for clinical nursing information. J Am Med Inform Assoc 1994;1(2):175-85.

111. Humphreys BL, Lindberg DA, Schoolman HM, Barnett GO. The Unified Medical Language System: an informatics research collaboration. J Am Med Inform Assoc 1998;5(1):1-11.

112. Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Interface terminologies: facilitating direct entry of clinical data into electronic health record systems. J Am Med Inform Assoc 2006;13(3):277-88.

113. McCray AT, Bodenreider O, Malley JD, Browne AC. Evaluating UMLS strings for natural language processing. Proceedings / AMIA Annual Symposium 2001:448-52.

114. Johnson SB, Aguirre A, Peng P, Cimino J. Interpreting natural language queries using the UMLS. Proceedings / the Annual Symposium on Computer Application [sic] in Medical Care 1993:294-8.

115. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods of information in medicine 1998;37(4-5):394-403.

116. Bodenreider O, Smith B, Kumar A, Burgun A. Investigating subsumption in SNOMED CT: an exploration into large description logic-based biomedical terminologies. Artif Intell Med 2007;39(3):183-95.

117. Gene Ontology. 2007. (Accessed July 2, 2007, at http://www.geneontology.org/GO.doc.shtml.)

118. Chute CG, Cohn SP, Campbell JR. A framework for comprehensive health terminology systems in the United States: development guidelines, criteria for selection, and public policy implications. ANSI Healthcare Informatics Standards Board Vocabulary Working Group and the Computer-Based Patient Records Institute Working Group on Codes and Structures. J Am Med Inform Assoc 1998;5(6):503-10.

119. ISO 1087-1: Terminology work-vocabulary, part 1: theory and application: technical committee TC 37/SC 1. . In: ISO Standards—terminology (principles and coordination); 1996.

120. ISO/TS 17117:2002(E): Health informatics - controlled health terminology - structure and high-level indicators: technical committee ISO/TC 215. In: Health informatics; 2002.

121. ISO 1087-2: Terminology work-vocabulary, part 2: computer applications: technical committee TC 37/SC 3. In: ISO Standards—computer applications for terminology; 1996.

122. Spackman KA, Campbell KE, Cote RA. SNOMED RT: a reference terminology for health care. Proc AMIA Annu Fall Symp 1997:640-4.

123. Chute CG, Elkin PL, Sherertz DD, Tuttle MS. Desiderata for a clinical terminology server. Proceedings / AMIA Annual Symposium 1999:42-6.

124. Dr. Donald A. B. Lindberg DotNLoM. Hearings Before the Subcommittee on the Departments of Labor, Health and Human Services, Education, and Related Agencies of the House Committee on Appropriations. In: Departments of Labor HaHS, Education, and Related Agencies Appropriations for 1986, ed.; 1985.

125. UMLS Source Vocabularies. National Library of Medicine. (Accessed June 27, 2007, at http://www.nlm.nih.gov/research/umls/metab4.html.)

126. Logical Observation Identifiers Names and Codes. (Accessed 6/19, 2007, at http://www.regenstrief.org/medinformatics/loinc/.)

127. Miller RA, Pople HE, Jr., Myers JD. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. The New England journal of medicine 1982;307(8):468-76.

128. Lemaire JB, Schaefer JP, Martin LA, Faris P, Ainslie MD, Hull RD. Effectiveness of the Quick Medical Reference as a diagnostic tool. Cmaj 1999;161(6):725-8.

129. Miller RA, McNeil MA, Challinor SM, Masarie FE, Jr., Myers JD. The INTERNIST-1/QUICK MEDICAL REFERENCE project--status report. The Western journal of medicine 1986;145(6):816-22.

130. Berner ES, Maisiak RS, Cobbs CG, Taunton OD. Effects of a decision support system on physicians' diagnostic performance. J Am Med Inform Assoc 1999;6(5):420-7.

131. Friedman CP, Elstein AS, Wolf FM, et al. Enhancement of clinicians' diagnostic reasoning by computer-based consultation: a multisite study of 2 systems. Jama 1999;282(19):1851-6.

132. Srinivasan A, Kunapareddy N, Mirhaji P, Casscells SW. Semantic web representation of LOINC: an ontological perspective. AMIA Annual Symposium proceedings / AMIA Symposium 2006:1107.

133. Bakken S, Cimino JJ, Haskell R, et al. Evaluation of the clinical LOINC (Logical Observation Identifiers, Names, and Codes) semantic structure as a terminology model for standardized assessment measures. J Am Med Inform Assoc 2000;7(6):529-38.

134. Huff SM, Rocha RA, McDonald CJ, et al. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. J Am Med Inform Assoc 1998;5(3):276-92.

135. McDonald CJ, Huff SM, Suico JG, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. Clinical chemistry 2003;49(4):624-33.

136. UMLS Knowledge Source Server. (Accessed July 3, 2007, at http://umlsks.nlm.nih.gov/kss/.)

137. SNOMED® CT Browser. (Accessed July 3, 2007, at http://terminology.vetmed.vt.edu/SCT/menu.cfm.)

138. Dolin RH, Alschuler L, Beebe C, et al. The HL7 Clinical Document Architecture. J Am Med Inform Assoc 2001;8(6):552-69.

139. Dolin RH, Alschuler L, Boyer S, et al. HL7 Clinical Document Architecture, Release 2. J Am Med Inform Assoc 2006;13(1):30-9.

140. Rishel W. A new approach to claims attachments. Use of the proposed XML-based messages might hasten selective use of autoadjudication. Healthc Inform 2003;20(9):56-8.

141. Klein A, Ganslandt T, Brinkmann L, Spitzer M, Ueckert F, Prokosch HU. Experiences with an Interoperable Data Acquisition Platform for Multi-Centric Research Networks Based on HL7 CDA. AMIA Annual Symposium proceedings / AMIA Symposium 2006:986.

142. Muller M, Frankewitsch T, Ganslandt T, Burkle T, Prokosch HU. The Clinical Document Architecture (CDA) enables electronic medical records to wireless mobile computing. Medinfo 2004;11(Pt 2):1448-52.

143. 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.

144. Hunscher DA, Boyd AD, Green LA, Clauw DJ. Representing Natural-Language Case Report Form Terminology Using Health Level 7 Common Document Architecture, LOINC, and SNOMED-CT: Lessons Learned. AMIA Annual Symposium proceedings / AMIA Symposium 2006:961.

145. Health Language 7. (Accessed July 2, 2007, at http://www.hl7.org/index.cfm.)

146. Meystre S. In; 2006.

147. Gupta D, Saul M, Gilbertson J. Evaluation of a deidentification (De-Id) software engine to share pathology reports and clinical documents for research. Am J Clin Pathol 2004;121(2):176-86.

148. Medical Words. (Accessed June 21, 2007, at http://sourceforge.net/projects/medicalwords/.)

149. GNU Aspell. (Accessed June 21, 2007, at http://aspell.net/.)

150. Crowell J, Zeng Q, Ngo L, Lacroix EM. A frequency-based technique to improve the spelling suggestion rank in medical queries. J Am Med Inform Assoc 2004;11(3):179-85.

151. Crowell JB, Zeng QT, Kogan S. A technique to improve the spelling suggestion rank in medical queries. AMIA Annual Symposium proceedings / AMIA Symposium 2003:823.

152. QuickTag (Accessed at http://www.cogilex.com/.)

153. Porter MF. An algorithm for suffix stripping. Program 1980;14(3):130-7.

154. Domingos P, Pazzani M. On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning 1997;29(2-3):103-30.

155. Cestnik B. Estimating probabilities: A crucial task in machine learning. . In: Proceedings of the Ninth European Conference on Artificial Intelligence; 1990; 1990. p. 147-9.

156. Winawer S, Fletcher R, Rex D, et al. Colorectal cancer screening and surveillance: clinical guidelines and rationale-Update based on new evidence. Gastroenterology 2003;124(2):544-60.

157. Recommended Adult Immunization Scheduled United States, October 2006-September 2007. Center for Disease Control. (Accessed July 20, 2007, at http://cdc.gov/vaccines/recs/schedules/downloads/adult/06-07/adult-schedule.pdf.)

158. Goroll AH, Sirio C, Duffy FD, et al. A new model for accreditation of residency programs in internal medicine. Annals of internal medicine 2004;140(11):902-9.

159. Corbett E, Whitcomb M. The AAMC Project on the clinical education of medical students: Clinical Skills education. Washington, DC: Association of American Medical Colleges; 2004.