BMI 506: Clinical Decision Support

Assignment 1: Motivate the development of the proposed clinical decision support system

PLM: Eric Holden & Tara Salehpour

February 16, 2015

Description of the search strategy

1. Research question: What previous clinical decision support tools exist for comparing population health data and cohort similarity?
2. An initial literature search was conducted to obtain list of Medical Subject Headings (MeSH) and keywords to construct a search strategy.

Terms from Initial Literature Search

MeSH		Keywords
Algorithms	Medical Informatics/organization & administration*	Treatment Outcome
Artificial Intelligence	Medical Informatics/trends*	real time
Artificial Intelligence*/economics	**Medical Records Systems, Computerized**	**prediction patient specific**
Artificial Intelligence*/trends	Medical Records Systems, Computerized/standards	**personalized risk**
Biomedical Research/economics	Models, Statistical	**personalized**
Biomedical Research/standards	Multivariate Analysis	Pattern Recognition
Biomedical Technology	**Natural Language Processing**	Patient-Specific
Biostatistics	Neural Networks (Computer)	**patient-centered**
Cluster Analysis	Observation/methods*	**patient similarity analytics**
Comparative Effectiveness Research	Observer Variation	**patient similarity**
Comparative Effectiveness Research*/statistics & numerical data	Outcome and Process Assessment (Health Care)	**NOT pharmacogenomics**
Comparative Effectiveness Research/methods	Outcome and Process Assessment (Health Care)/statistics & numerical data*	**NOT genetic**
Comparative Effectiveness Research/organization & administration*	Outcome Assessment (Health Care)	**Natural Language Processing**
Comparative Study	Patient Compliance/statistics & numerical data	**Multi-Patient Queries**
Computer Simulation*/economics	Patient Dropouts/statistics & numerical data	**informatics**
Confidence Intervals	Patient Selection	**individualized**
Cost-Benefit Analysis	**Patient-Centered Care**	**individualized**
Data Collection*/statistics & numerical data	**Patient-Specific Modeling**	**Humans**
Data Collection/standards	**Pattern Recognition, Automated**	**Electronic Health Record**
Data Interpretation, Statistical	**Physician's Practice Patterns/statistics and numerical data**	cluster analysis
Data Mining	Postoperative Care/standards	**clinical decision support**
Databases, Factual	Postoperative Complications/blood	
Databases, Factual/statistics & numerical data	Postoperative Complications/diagnosis*	
Decision Making	Postoperative Complications/drug therapy	
Decision Making, Computer-Assisted	Prognosis*	
Decision Support Systems, Clinical	Program Development	
Decision Support Systems, Clinical*/economics	Program Evaluation	
Decision Support Systems, Clinical/instrumentation*	Quality Control	
Decision Support Techniques*	Quality Improvement	
Decision Trees	Quality Improvement/organization & administration*	
Delivery of Health Care	Quality Improvement/statistics & numerical data	
Delivery of Health Care, Integrated	Randomized Controlled Trials as Topic*/methods	
Delivery of Health Care/trends*	Randomized Controlled Trials as Topic*/statistics & numerical data	
Diffusion of Innovation	Registries*	
Electronic Health Records	Registries/standards*	
Electronic Health Records/organization & administration*	Registries/statistics & numerical data	
Electronic Health Records/standards*	Reoperation	
Electronic Health Records/statistics & numerical data	Research Design	
Electronic Health Records/utilization	Residence Characteristics*/statistics & numerical data	
Evaluation Studies	Retrospective Studies	
Evidence-Based Medicine*	Risk	
Expert Systems	Risk Assessment	
Health Services Research	Sample Size	
Humans	Semantics	
Individualized Medicine	Social Determinants of Health*/statistics & numerical data	
Individualized Medicine/methods*	Software*	
Information Storage and Retrieval/standards	Specialization/statistics & numerical data*	
Knowledge Bases	**Support Vector Machines**	
Linear Models	Systems Integration	
Logistic Models	**Time**	
Longitudinal Studies	Time Factors	
Markov Chains*	Treatment Outcome	
Medical Informatics	United States	
Medical Informatics/methods	United States Agency for Healthcare Research and Quality	

Bold Words: Used in final search

3. Base search queries were created for each database.

   Google Scholar:

   Use Filter: Full text; Articles

   "clinical decision support" AND (personalized OR individualized) AND "electronic health

record"

PubMed:
Use Filter: Full text
((((("Decision Support Systems, Clinical"[Mesh]) AND "Individualized Medicine"[MeSH Terms]) OR "Patient-Specific Modeling"[Mesh] AND "Electronic Health Records"[Mesh])) NOT pharmacogenomics NOT genetic

4. Design search strategy.

Objective: Return a search query with 10 results
If number of results returned:
+/- 4 of objective, apply eligibility criteria
< 10 remove search terms from base search
>20 add search terms to base search
15 - 20 restrict publication date to previous 4 years (2011)
If number of results returned is <20
Remove duplicate articles found from previous searches
If titles and/or abstracts are similar, choose most cited or most recent
>0 keep results

Description of the eligibility criteria and the rationale
Eligibility Criteria
- NOT genetic[Title/Abstract]
- NOT pharmacogenomics[Title/Abstract]
- NOT genomic[Title/Abstract]
- Most significant (most cited or most recent)

The eligibility criteria was created after conducting the initial searches without excluding: genetic, genomic or pharmacogenomics. The MeSH search terms: Individualized Medicine and Patient-Centered Care returned a large number of results about genomics and personalized care. Using the 'NOT' search operator, eliminated the erroneous results from the search.

Highly specific search queries tended to return similar results. In this case, it is useful to determine which is the most significant article. If one article was cited at least three times more than any of the other articles, it was deemed the most significant. Else, the most recent article with relatively average or above average citations was selected as the most significant article.

Before and After Eligibility Criteria

Eligibility Criteria: Compare Before & After

Database	Search Terms	Before Eligibility Criteria: Article Titles	Number of Articles	Eligibility Criteria	After Eligibility Criteria: Article Titles	Number of Articles
PubMed	((("Electronic Health Records"[MAJR]) AND "Natural Language Processing"[MAJR]) AND "Humans"[MeSH Terms]) AND "Support Vector Machines"[MAJR]	A flexible framework for deriving assertions from electronic medical records. A supervised framework for resolving coreference in clinical records. Assessing the similarity of surface linguistic features related to epilepsy across pediatric hospitals. Automatic extraction of relations between medical concepts in clinical texts. Hybrid methods for improving information access in clinical documents: concept, assertion, and relation identification.	5	selected most recent	Assessing the similarity of surface linguistic features related to epilepsy across pediatric hospitals	1
PubMed	((((("Artificial Intelligence"[MAJR]) AND "Electronic Health Records"[MAJR]) AND "Humans"[MeSH Terms]) AND "Natural Language Processing"[MeSH Terms]) AND "Time"[MeSH Terms]]	A flexible framework for recognizing events, temporal expressions, and temporal relations in clinical text. A la Recherche du Temps Perdu: extracting temporal relations from medical text in the 2012 i2b2 NLP challenge. An end-to-end system to identify temporal relation in discharge summaries: 2012 i2b2 challenge. Applying a natural language processing tool to electronic health records to assess performance on colonoscopy quality measures. Classifying temporal relations in clinical data: a hybrid, knowledge-rich approach. Combining rules and machine learning for extraction of temporal expressions and events from clinical narratives. Comprehensive temporal information detection from clinical text: medical events, time, and TLINK identification. Development of a natural language processing system to identify timing and status of colonoscopy testing in electronic medical records. Evaluating temporal relations in clinical text: 2012 i2b2 Challenge. Eventual situations for timeline extraction from clinical reports. Extracting temporal information from electronic patient records. Temporal reasoning over clinical text: the state of the art. Temporal relation discovery between events and temporal expressions identified in clinical narrative. TEMPTING system: a hybrid method of rule and machine learning for temporal relation extraction in patient discharge summaries. Towards generating a patient's timeline: extracting temporal relationships from clinical notes.	15	selected most cited	Evaluating temporal relations in clinical text 2012 i2b2 Challenge	1
PubMed	Arthritis[Title] AND Rheumatism[Title] AND Aging[Title] AND Medical[Title] AND Information[Title] AND System[Title]	ARAMIS and toxicity measurement. (Arthritis Rheumatism and Aging Medical Information System). Arthritis, Rheumatism and Aging Medical Information System Post-Marketing Surveillance Program. Gastrointestinal complications of prescription and over-the-counter nonsteroidal anti-inflammatory drugs: a view from the ARAMIS database. Arthritis, Rheumatism, and Aging Medical Information System. NSAID induced gastrointestinal complications: the ARAMIS perspective--1997. Arthritis, Rheumatism, and Aging Medical Information System. The Arthritis, Rheumatism and Aging Medical Information System (ARAMIS): still young at 30 years.	5	selected most recent	The Arthritis, Rheumatism and Aging Medical Information System (ARAMIS)/ Still young at 30 years	1
PubMed	"The Asgaard Project"	An intention-based language for representing clinical guidelines. The Asgaard Project/ A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines	2	selected most recent	The Asgaard Project/ A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines	1
PubMed	((((("Individualized Medicine"[MAJR]) AND "Electronic Health Records"[MAJR]) AND "Decision Support Systems, Clinical"[MAJR]))	Adaptive semi-supervised recursive tree partitioning The ART towards large scale patient indexing in personalized healthcare Electronic medical records and personalized medicine Improving Health Care Outcomes Based on Electronic Health Records Usability of a novel clinician interface for genetic results.	4	ADDED: NOT genetic[Title/Abstract]	Adaptive semi-supervised recursive tree partitioning The ART towards large scale patient indexing in personalized healthcare Electronic medical records and personalized medicine Improving Health Care Outcomes Based on Electronic Health Records	3
PubMed	(((((("Decision Support Systems, Clinical"[Mesh]) AND "Individualized Medicine"[MeSH Terms]) OR "Patient-Specific Modeling"[Mesh] AND "Algorithms"[MeSH Terms]) AND "Electronic Health Records"[Mesh]))	Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. Evicase: an evidence-based case structuring approach for personalized healthcare. Fine-grained clinical outcome extraction and polarity classification. Integrating pharmacogenetic information and clinical decision support into the electronic health record.	4	ADDED: NOT pharmacogenomics	Artificial intelligence framework for simulating clinical decision-making/ a Markov decision process approach. Artificial intelligence framework for simulating clinical decision-making/ a Markov decision process approach Electronic medical records and personalized medicine Evicase: an evidence-based case structuring approach for personalized healthcare.	3
Subtotal			35			10
Duplicates Subrated		**Total Before**	**34**		**Total After**	**9**

Duplicate Article Found in Search

The results of the search (publications identified) and the application of the criteria (publications considered eligible after applying the criteria)

Database	Search Terms	Article Title	Number of Articles
Google Scholar	"Multi-Patient Queries" AND "Electronic Health Records" AND "individualized OR personalized OR patient-centered"	Utilizing IHE-based Electronic Health Record Systems for Secondary Use	1
PubMed	(("Biomedical Technology"[MeSH]) AND "Electronic Health Records"[MeSH]) AND "Natural Language Processing"[MAJR]	Natural language processing in biomedicine a unified system architecture overview	1
PubMed	(("Electronic Health Records"[MAJR]) AND "Expert Systems"[MAJR]) AND "Knowledge Bases"[MeSH Terms]	Modeling and Executing Electronic Health Records Driven Phenotyping Algorithms using the NQF Quality Data Model and JBoss® Drools Engine	1
PubMed	(("Pattern Recognition, Automated"[Mesh]) AND "Patient-Centered Care"[Mesh]) AND "Decision Support Systems, Clinical"[Mesh]	Case-based medical informatics.	1
PubMed	((("Electronic Health Records"[MAJR]) AND "Natural Language Processing"[MAJR]) AND "Humans"[MeSH Terms]) AND "Support Vector Machines"[MAJR]; selected most recent	Assessing the similarity of surface linguistic features related to epilepsy across pediatric hospitals	1
PubMed	((("Medical Records Systems, Computerized"[MeSH Terms]) AND "Medical Informatics"[MAJR]) AND "Comparative Effectiveness Research"[MAJR]) AND "Review"[pt]	Building the informatics infrastructure for comparative effectiveness research (CER): a review of the literature.	1
PubMed	((("Patient-Centered Care"[MeSH Terms] AND "Electronic Health Records"[MAJR]) AND "Decision Support Systems, Clinical"[MAJR]) AND "Data Mining"[MeSH Terms])	Patient-tailored prioritization for a pediatric care decision support system through machine learning.	1
PubMed	((("Artificial Intelligence"[MAJR]) AND "Electronic Health Records"[MAJR]) AND "Humans"[MeSH Terms]) AND "Natural Language Processing"[MeSH Terms]) AND "Time"[MeSH Terms]]; selected most cited	Evaluating temporal relations in clinical text 2012 i2b2 Challenge	1
PubMed	((("Databases, Factual"[MAJR]) AND "Health Services Research"[MAJR]) AND "Medical Records Systems, Computerized"[MAJR]) AND "Physician's Practice Patterns/statistics and numerical data"[MAJR])	Use of the mediplus patient database in healthcare research	1
PubMed	Arthritis[Title] AND Rheumatism[Title] AND Aging[Title] AND Medical[Title] AND Information[Title] AND System[Title]]; selected most recent	The Arthritis, Rheumatism and Aging Medical Information System (ARAMIS)/ Still young at 30 years	1
PubMed	The Asgaard Project]; selected most recent	The Asgaard Project/ A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines	1
PubMed	"Cluster Analysis"[MAJR] AND "Electronic Health Records"[MAJR]	Modeling temporal relationships in large scale clinical associations Visual cluster analysis in support of clinical decision intelligence	2
PubMed	"patient similarity" AND "clinical decision support" AND (personalized OR individualized) AND "electronic health record"	Continual development of a personalized decision support system Visual Cluster Analysis in Support of Clinical Decision Intelligence	2
Google Scholar	allintitle: prediction "patient specific" OR "personalized risk" "clinical decision support"	An Intelligent Clinical Decision Support System for Patient-Specific Predictions to Improve Cervical Intraepithelial Neoplasia Detection A patient-driven adaptive prediction technique to improve personalized risk estimation for clinical decision support	2
PubMed	(((((("Decision Support Systems, Clinical"[Mesh]) AND "Individualized Medicine"[MeSH Terms]) OR "Patient-Specific Modeling"[Mesh] AND "Algorithms"[MeSH Terms]) AND "Electronic Health Records"[Mesh])) NOT pharmacogenomics	Artificial intelligence framework for simulating clinical decision-making/ a Markov decision process approach. Artificial intelligence framework for simulating clinical decision-making/ a Markov decision process approach Electronic medical records and personalized medicine Evicase: an evidence-based case structuring approach for personalized healthcare.	3
PubMed	((("Individualized Medicine"[MAJR]) AND "Electronic Health Records"[MAJR]) AND "Decision Support Systems, Clinical"[MAJR])) NOT genetic[Title/Abstract]	Adaptive semi-supervised recursive tree partitioning The ART towards large scale patient indexing in personalized healthcare Electronic medical records and personalized medicine Improving Health Care Outcomes Based on Electronic Health Records	3
Google Scholar	"patient similarity analytics" AND (personalized OR individualized)	Information technology 2 3 for healthcare transformation Iterative Cohort Analysis and Exploration Towards Personalized Medicine/ Leveraging Patient Similarity and Drug Similarity Analytics Visual Cluster Analysis in Support of Clinical Decision Intelligence	4
Google Scholar	"patient similarity" AND "clinical decision support" AND (personalized OR individualized) AND "electronic health record"	From Micro to Macro/ Data Driven Phenotyping by Densification of Longitudinal Electronic Medical Records A 'Green Button' For Using Aggregate Patient Data At The Point Of Care Challenges in Designing an Online Healthcare Platform for Personalised Patient Analytics Information technology 2 3 for healthcare transformation Machine learning of patient similarity/ a case study on predicting survival in cancer patient after locoregional chemotherapy Mining Diabetes Complication and Treatment Patterns for Clinical Decision Support Open Issues in Intelligent Personal Health Record – An Updated Status Report for 2012 PARAMO/ A PARAllel predictive MOdeling platform for healthcare analytic research using electronic health records	8
Google Scholar	"patient similarity" AND "clinical decision support" AND "electronic health records" AND "informatics" since 2011	A 'Green Button' For Using Aggregate Patient Data At The Point Of Care A method for inferring medical diagnoses from patient similarities A Search Engine for Structured Health Data A SNOMED supported ontological vector model for subclinical disorder detection using EHR similarity Discrimination and stratification tests of cardiovascular disease risk assessment models against ultrasound detection of carotid plaques in type 2 diabetics Information technology 2 3 for healthcare transformation Mining Diabetes Complication and Treatment Patterns for Clinical Decision Support Nonconfidential Patient Types in Emergency Clinical Decision Support PARAMO/ A PARAllel predictive MOdeling platform for healthcare analytic research using electronic health records Patient Clustering with Uncoded Text in Electronic Medical Records Predicting ICU Death with Summarized Data: The Emerging Health Data Search Engine Sim•TwentyFive: An Interactive Visualization System for Data-Driven Decision Support Supervised Patient Similarity Measure of Heterogeneous Patient Records VisualDecisionLinc/ A visual analytics approach for comparative effectiveness-based clinical decision support in psychiatry	14
Subtotal			49
Duplicates Subrated			**40**

Duplicate Article Found in Search

Search Limitations

There were quite a few limiting factors in our literature search. This was due in large part to the novel nature of the PLM system. As such, there was no literature directly pertaining to the issue at hand. Instead, we had to focus on literature that contained fragments of information that could, in theory, be applied to the development of PLM. Since these fragments only appear useful at this early stage of development, it is unclear if they will be truly valuable at a later stage. At the same time, this means our literature search was limited by our expectations of the requirements of PLM, rather than the potential reality. As a result, it is entirely possible that our search missed some crucial aspect that will only become apparent at a later date. Because this eventuality is uncertain, there is no sense in belaboring the point, but we must be prepared to perform additional literature searches, should the need arise.

Conclusions

The focus of this project will be the development of a clinical decision support tool within the framework of the Patients Like Mine (PLM) system. The overall goal of the PLM system is to compare the medical values of an individual patient against a database of prior patients in order to help predict possible outcomes. The values used in the patient comparison, as well as the predictive values extracted from the cohort are both predefined by constructs called input templates and output templates respectively. For this project, we will be focusing on the input templates. In order for two patients to be considered similar, their medical values must both fall within the same range of similarity. This range of similarity is informed both by an internal knowledge base, as well as the patient of interest's actual values. The reason for this is because the range of similarity for an approximately normal value is different from the range of similarity for an extreme value. Determining the range of similarity must, therefore, be done on-the-fly, requiring a clinical decision support tool. Once the ranges of similarity have been determined, a suitable query will be constructed to search the database and compile a cohort of similar patients. Due to the scope of the PLM system, our project for this class will stop at the construction of the query, since access to actual patient databases is restricted.

The ultimate goal of the PLM system is to allow clinicians to perform a real-time search for patient similarity in order to predict treatment outcomes. Since guidelines are already in place to inform clinical decisions, the PLM system's main target is outlier patients that fall outside of established guidelines. The current work on the system has so far been limited to colorectal surgery patients, but the ultimate goal is to make the system generalizable to any medical situation, so long as the templates exist. To that end, the internal knowledge base of the PLM system must be robust enough to allow for the comparison of most potential medical values.

As a consequence of the PLM system focusing on outlier cases, there is no specific guideline that can be followed to help inform this project. This issue is further compounded by the fact that the system should be generalizable to multiple disciplines. Instead, our main knowledge resources will be standards databases such as LOINC and RxNorm, as well as subject matter experts. Because work has already been done on PLM in regards to the prediction of bleeds in post-operative colorectal surgery patients, it makes sense that testing of our proposed tool should focus on that particular template set.

According to the papers described above, the current state of the field surrounding PLM is as follows. For the sake of this discussion, the information gathered from papers have been divided into general topic categories so their value to this project and PLM in general can be made clear. The categories will be sorted by their general relevancy to this project.

The first topic category is distance measures used in the comparison of patients. This topic is clearly at the crux of this project, since the query generated by our tool will contain these similarity measures. A large number of the papers that touched on this topic only did so as part of comparative effectiveness studies whereby cohorts of patients could be compared against each other to determine which drugs or treatments are most useful. While the algorithms described in these papers might be useful, it is important to note that they are fundamentally different from the one proposed by our project. Instead of entire groups of patients being compared, our tool would only compare patients on an individual basis in order to build a cohort. Comparative effectiveness studies compare already established cohorts, so the comparison algorithms are understandably more complex. Therefore these algorithms' usefulness to our project depends on whether or not they can be scaled down effectively. Due to the relatively simple nature of our comparison, however, the use of these large-scale algorithms seems excessive. In addition to the discussion of the algorithms used in comparative effectiveness studies,

visualization methods were discussed as well. While the visualization of patient similarities is outside the scope of this project, it is still an important aspect of the PLM system as a whole. Therefore, it is an important aspect to be aware of for future work on PLM.

The next major topic covered by these papers is the construction of knowledge bases. This is an important aspect of this project, as all the clinical decision support must be informed by the tool's internal knowledge base. A number of papers describe various structures for knowledge base rules, as well as their usage in larger clinical decision support systems. Some of the structures described are MDA, IDAN, KNAVE-II, and VISITORS. At this early stage it is unclear which of these structures, if any, could be applied to our project, but having multiple options allows us to be flexible in our approach moving forward. It should be noted that the structures in these papers are, again, applied to more complex clinical decision support problems, so they would likely need to be modified to suit our needs.

Another important aspect identified within these papers is natural language processing. NLP serves an important role for our project, since information needs to be extracted from a patient's records before it can be interpreted by the knowledge base. In particular, a considerable amount of information resides within medical notes that could be useful for determining patient similarity. To that end, we found several papers which discussed extracting information from medical notes. Most papers focused on extracting information under a specific medical context, which might limit the generalizability, but the techniques employed can still be utilized, provided we have access to the requisite medical knowledge. Another aspect mentioned by these papers is the extraction of temporal information, which would be particularly useful for PLM, since comparisons are done by relative time (pre-surgery, post-surgery day one, etc).

Finally, we found several papers pertaining to databases of patient information. This aspect is highly relevant to PLM as a whole, but not our project in particular, since we are creating a tool that assumes the existence of an already-built patient database. Still, the structures of the databases described by these papers may provide some insights that allow our project to produce more effective and useful queries. It should be noted that none of the databases described match the exact type of information that would be stored in PLM's patient database, potentially resulting in different indexing and search methodologies.

The Self and Peer Appraisal

Tara:

- Conducted initial search and determined MeSH Terms and keywords to be used
- Designed search strategy and eligibility criteria
- Amassed research articles through search
- Compiled final paper

  Authored:
  - Tables/ figures
  - Description of the search strategy
  - Description of the eligibility criteria and the rationale

- Before and After Eligibility Criteria
- The results of the search (publications identified) and the application of the criteria (publications considered eligible after applying the criteria)

Peer Appraisal: I am satisfied with the division of work; it seemed to compliment our strengths and preferences.