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.



1. 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

1. 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**

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

**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.