The high level goal of this project to develop state of the art clinical decision support (CDS) tools capable of advancing clinical care towards the ideal of personalized medicine. The idea of individualized or personalized medicine represents an ideal healthcare diagnosis and treatment model. Personalized medicine is the promise of patient specific, customized care; it is the perfect care at the perfect time based on exactly what the patient needs. Patients Like Mine (PLM) key concepts parallel the ideal of individualized medicine; including real-time, relevant, patient specific clinical insight precisely when it is needed.

Clinical databases contain a wealth of valuable information that is used for retrospective studies1. Utilizing clinical database information at the point of care has the potential to revolutionize clinical care by disambiguating disease heterogeneity and providing personalized treatment. This project aims to utilize that knowledge at the point of care to achieve improved, personalized clinical outcomes through near-time cohort similarity analysis and prediction. The project objective is to implement a tool within the framework of the PLM system that allows a domain expert to author a template for use in the generation of a database query. The templates are composed of an input template, an output template, relevant patient data concepts and query functions.

The project focus is to implement a tool within the framework of the PLM system that allows a domain expert to author a template for use in the generation of a database query. The templates are composed of an input template, which defines the variables, their ranges, and their relative timeframes in order to determine which patients are considered similar to the patient of interest, and an output template, which defines the variables and their relative timeframe to be extracted from the cohort of similar patients. The authoring tool will allow the domain expert to select the variables to be included in both the input and output templates from an internal knowledge base. The template will then be used in conjunction with a patient of interest's information in order to generate a database query which will then be run on the past patients database(s) to extract information from patients that are determined to be suitably similar.

The project domain expert, Dr. Peter Li, has defined the following patient data concepts relevant to risk of post-op colorectal bleed:

* Date of Birth
* Gender
* Demographics
* Colorectal Surgery
* Blood Transfusion
* Blood Pressure
* Hematocrit
* Hemoglobin
* White Blood Cell (WBC) Count
* Temperature
* Medications: Blood Thinners
* Date of observations/events/lab values or change in values (date test administered)

These relevant patient data concepts are available in the input template for selection or exclusion and contain adjustable similarity range relative to the current patient’s values or change in values and their relative timeframes. The domain user can also choose from a set of available functions within the template. These functions increase the expressivity of the query criteria. The functions apply to a specified time interval defined by the domain user. Some examples include:

• First, last, mean, greater than, less than, equal to

• (Similarity) distance functions

Data extraction methodologies must be cognizant of existing data representation, location and the effect of a temporal dimension. These aspects will influence how relevant data is accessed and retrieved.

Patient data concepts will be mapped to ontologies. Implementation requires a template to map patient data to biomedical ontological concept codes. Representing the data concepts of interest with ontologies, formalizes the requisite data concepts into an atomic set. A query for the concept “blood pressure”, can reference the template and all the instances for data values of the concept blood pressure can be identified. Ontology based queries permit controlled, precise, reproducible access to data.

Once the domain expert has finished authoring a template the template parameters need to be transferred into a query. The parameters used in the query must be a relative representation of initial values from the current patient and similar patients. This can be achieved with prepared statements3. Prepared statements exist in a pre-designed query framework. They are reusable query statements that can accept parameters. The input template supplies the query parameters to the prepared statements during runtime.

At this point, the query contains the necessary knowledge about the data concepts and ranges. There isn’t any information directing the query where to find the data values defined by the query. The database schema contains the information necessary to define the location of the data values. The database schema is defined by the HL7 Fast Healthcare Interoperability Resources (FHIR) Clinical Information Model. The FHIR schema is object-oriented. The query operates on a relational database. It is necessary to convert the FHIR schema into a relational schema. Hibernate is a tool that will execute this conversion.

Criteria Application Program Interface (API) will be used to store knowledge about the database schema. This knowledge will define the location of the data values in a query. The combined information and knowledge needed to execute the query are: patient data, the template defined data search criteria, the prepared statements that accept the parameters from the input template, ontology template mapping data to coded concepts, ontological concepts used to represent query elements, location of data values derived form database schema, FHIR schema translated in to a relational schema, and Criteria API to supply database schema knowledge.