Proceedings of the 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), Larnaca, Cyprus, 11-13 November 2012

A Method to Summarize Disease Based Temporal State of Human Organ using Laboratory Test Data and UMLS Knowledge

Soubhik Paul, Raj Roy,
Swarupananda Bissoyi, Sandip Bhaumik,
Meru A Patil, Vivekeshwar Sai Bhardwaj
Samsung India Software Operations
Bangalore, India
Email: soubhik.p@samsung.com

***Abstract*—A novel concept of Disease Based Temporal Score (DT-Score) is introduced to efficiently represent periodic laboratory test data. Through this score, temporal state of an organ can be represented by summarizing periodic laboratory test data. The score can be used to indicate early trend for chronic abnormalities and thus results in an effective wellness measure. Many of these chronic abnormalities have a late manifestation and are major contributors for healthcare cost and mortality. Resources of Unified Medical Language System (UMLS) are introduced for automatic generation of relational tree between laboratory test, disease and organs with relative rank. Doctor's annotations are used to create reference score and data mining techniques are employed in deriving a mathematical model for estimating the DT-Score. A novel human body based summarization is employed for an intuitive view of the DT-Score and resultant temporal state of the associated organ. The proposed method enables an efficient temporal summarization of high volume of laboratory data and eventually reduces the cognitive load on physician. This method has potential to impact larger population as this can be effectively built over low cost regular laboratory test.**

***Index Terms*—Disease Based Temporal Score, Periodic Laboratory Test, UMLS, Temporal State of Human Organ, Regression Mechanism**

I. Introduction

Efficient visual representation of patient's lifelong clinical data plays an important role in delivering quality patient care. With advent of sophisticated diagnosis procedures, increased affordability, higher health consciousness and friendly corporate policies, periodic health checkups are becoming norms for urban populations. As a result of this, series of laboratory test data is captured throughout the life of a person. Most of the present day's Health Information Tools visualize these data as a time-value pair by means of graphs or charts [1] [2] [3]. An intuitive efficient summarization of this temporal data is an open challenge.

Many abnormalities in human organ are chronic in nature with late manifestation [4] [5]. Medical data collected across years can be a good early indication for these abnormalities. The challenge is to analyze the huge temporal data and identify the organs state, considering large effort involved. In the medical domain, organ state representation is limited to critical clinical conditions, like tracking of ICU patients. MODS (Multiple Organ Dysfunction Score) [6] and LODS (Logisitc Organ Dysfunction Score) [7] are popular example of organ state representation used in tracking critical clinical conditions, like post-operative state.

For the last few decades researcher worldwide are working on effective visualization of patient health data and a number of such visualization tools are already in place. One common practice of visualizing temporal laboratory test data is through data table, trend chart and timeline, such as Time Line Browser [1], LifeLines [2], visualization interface proposed in [3]. Along with this, some of the tools introduce color coding scheme to indicate the states of the laboratory test results [8] [9]. Lam H. et al [10] has introduced visualization of laboratory tests through human anatomical system. Most of the known visualizing tools, techniques only focus on representation of data over the course of time and they lack in summarization of temporal laboratory data. The existing medical analysis of laboratory data is confined to snapshot data alone [11]. This results in discarding important temporal information of the laboratory data and thus it prevents doctors to correlate clinical findings with chronicity of the condition. Takabayashi et. al. [12] shows how temporal abstraction together with data mining can be used to derive rules to distinguish between hepatitis B and Hepatitis C from time series laboratory data.

Doctor's annotation is one of the important diagnostic information in patient's clinical records. The temporal state of an organ can be effectively derived using these annotations. But, regular annotations by Doctor in patient clinical record like Annual Checkup are not guaranteed. Regular laboratory test data are mostly the only information present in data set like Annual Checkup. In this paper, an effort is made to capture the temporal state of the organ using regular laboratory test data in patient record like Annual Checkup program.

In this paper, we have introduced an unique disease based temporal score (DT-Score) to effectively represent temporal state of an organ. A mathematical model is proposed to detect temporal abnormality through DT-Score calculation. Semantic

978-1-4673-4358-9/12/$31.00 ©2012 IEEE

network and co-occurrence score of Unified Medical Language System (UMLS) is used to discover the ranked relations between the medical concepts used in our methodology. Relation rank is used to specify the confidence of the DT-Score. A novel Human Anatomy based visual interface is introduced to visualize this score.

The Section II highlights the proposed approach in detail, where Subsection II-A discusses about the concept relation discovery from UMLS, concept ranking and deviation of ranking confidence, followed by the methodology to build model for DT-Score prediction in Subsection II-B and visualization technique of DT-Score using Human Body Model in Subsection II-C. A brief discussion on future work in Section III draws the conclusion.

II. Proposed Approach

In our proposed approach a concept called "Temporal State of Human Organ" is introduced to efficiently represent periodic laboratory test data. A temporal abnormality detection mechanism, through DT-Score, is formulated. The mechanism is based on UMLS's semantic network and co-occurrence table. The temporal summarization of an organ is represented through a list of relevant DT-Scores. A confidence score is also introduced for each disease score (DT-Score) for better interpretation. The confidence score is derived using rank of the laboratory procedures, used in calculating the DT-Score, in UMLS co-occurrence table against the particular disease. Real patient data collected from a hospital record is used to derive the mathematical model for the DT-Score. In this data, the doctors annotation in "decision report" is used as the classification reference. In the preparation of training data, temporal score of a disease (DT-Score) is derived using a sequence of historical annotations, employing separate weightage for each severity level (e.g., mild, moderate etc). Temporal variation of top ranked laboratory test parameters from the biological mean along with the frequency of abnormality of the particular parameter is considered as feature for classification. In the training phase, Support Vector Machine (SVM) based Regression mechanism is used to build the mathematical model (temporal state analyzer model). In the execution phase, feature vectors are derived for each disease (of an organ) and is passed through the temporal state analyzer model to calculate the temporal score (DT-Score). Fig. 1 depicts the overall flow of the proposed system.

The estimated DT-Score signifies the probability of a doctor annotating a particular patient with a particular condition and severity in the time window considered for analysis. Suppose, the estimated DT-Score for "Fatty Liver" is 25% for a particular patient, that means in case a doctor has done regular annotation for all the laboratory tests within the time window, there would have been "Mild Fatty Liver" annotation ("mild" has an assigned weight of 0.5 in this study) in 50% instances (or "Severe Fatty Liver" in 25% instance ("severe" has an assigned weight of 1.0 in this study) or other possible permutation of severity of Fatty Liver resulting in 25% overall score).

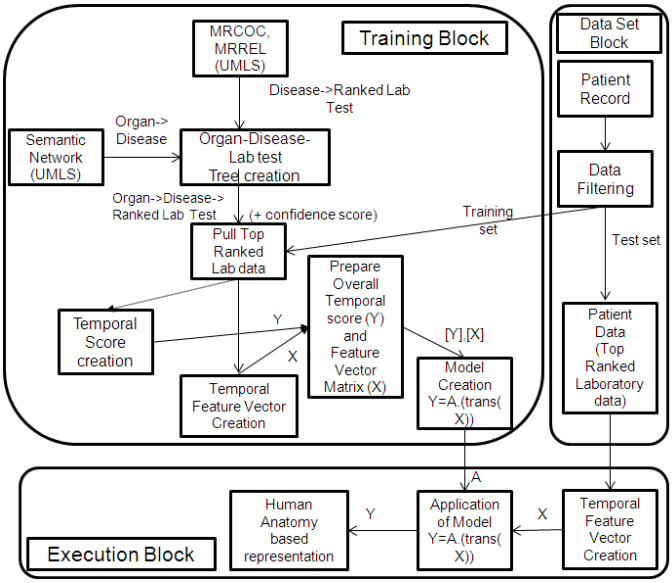

Fig. 1. Overall System Flow

The proposed methodology is described using Liver organ of human body, its related diseases and corresponding laboratory tests. The same concepts can be extended to other organs.

A. Concept relation discovery, concept ranking, derivation of ranking confidence

In the proposed approach, medical concepts like organ, disease and laboratory test, and their relationship are used to derive the temporal state of a human organ. Unified Medical Language System (UMLS) [13] is consulted to achieve this. UMLS is a comprehensive thesaurus and ontology of biomedical concepts [14]. UMLS has a Metathesaurus which contains 1 million biomedical concepts and 5 million concept names collected from 150 source vocabularies. UMLS has another knowledge called Semantic Network which categorizes all the Metathesaurus concepts into 133 semantic types and provides useful relationships between these concepts as well [13] [14]. UMLS also has a very useful concept ranked relations in the form of co-occurrence score, stored in MRCOC table. The co-occurrence score is derived using three sources, namely MEDLINE, AI/RHEUM, CCPS [13]. Qing et. al. [15] has used co-occurrence score information from MRCOC for extraction of diseases-drug relation as well as disease-lab chemical relationship. They observed 93 percent sensitivity for disease-drug relation and 68 percent sensitivity for disease-lab chemical relationship. In development of Personalized PageRank system [16] the subset of knowledge base graph was created using strength of co-occurrence information available in MRCOC table. Information available in MRCOC table also helped development and evaluation of KnowledgeMap computation tool [17].

In our methodology we are interested in three medical concepts organ, disease and laboratory procedure. Using UMLS Semantic Network and MRCOC table we have identified the relationships [refer Table I] that suffice our requirement. Information from MRCOC table is used to identify the ranked

relation between disease and laboratory procedure under the relation "diagnose". This resulted in a Organ-Disease-Labtest tree (refer Fig. 2). The leaf node of the tree contains the top ranked laboratory tests (including diagnostic tests) sorted by co-occurrence score. This ranking is also used to derive the confidence score. Higher the rank of the laboratory test, used for calculating DT-Score, higher is the associated value of confidence score.

TABLE I
CONCEPT RELATIONS

Concept1	Concept2	Relation
Organ	Disease	Location of
Laboratory Procedure	Disease	Diagnose
Laboratory Procedure	Laboratory Procedure	Is a

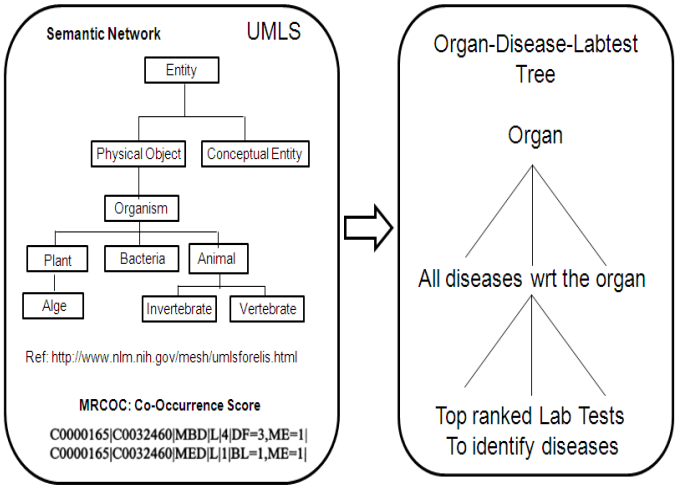

Fig. 2. Organ-Disease-Labtest tree

1) Methodology to Extract Related Concept: As already discussed, a organ-disease-laboratory procedure tree is required for temporal organ state. To achieve this we developed an algorithm (refer Algorithm 1) that extracts the list of related concepts for a given concept (parent concept) and corresponding relationship. Parent concept, semantic type of required concepts and relation are passed to the method as input parameter. First the relation ID is fetched using Semantic Network in UMLS database and then all the concepts of input semantic type is searched in UMLS Metathesaurus. If relationship not found, the MRCOC data is searched for co-occurred concepts. Finally, the related/co-occurred concepts along with the co-occurrence score is returned as the output.

2) Ranking of Laboratory Procedures co-occurred with Disease: For a given disease as parent concept and diagnose as relationship a list of laboratory tests are derived from the above mentioned algorithm along with co-occurrence score. In the derived tree diagnostic tests like diagnostic imaging are also considered as part of laboratory test. Next, the list is sorted with co-occurrence score for defining the confidence score and low ranked concepts are removed. The algorithmic steps of ranking Laboratory Procedures for a particular disease is described in Algorithm 2.

Algorithm 1 Extract Related Concept

procedure GETRELCONCEPT($concept, semTyp, reqRel$)
  $relID \leftarrow getrelationID(reqRelation)$
  if $searchRelatedConcepts(semTyp, relID)$ **then**
    $relConcepts \leftarrow getRelatedConcepts(semTyp, relID)$
    $coocScore \leftarrow null$
  else
    $relConcepts \leftarrow getCoocConcepts(concept, semTyp, relID)$
    $coocScore \leftarrow getCoocScores(relConcepts)$
  end if
  return $relConcepts, coocScore$
end procedure

Algorithm 2 Rank Laboratory Procedures

procedure RANKLABPROC($DisName, semTypLab, reqRel$)
  $relLabProc \leftarrow getRelConcept(DisName, semTypLab, reqRel)$
  while $relLabProc.count \neq null$ **do**
    if $relLabProc.coocScore \geq thresholdVal$ **then**
      $sortedLabProc \leftarrow sort(relLabProc)$
    end if
  end while
  return $sortedLabProc$
end procedure

These two algorithms resulted in the Organ-Disease-Laboratory procedure tree.

The top ranked laboratory procedure (including diagnostic procedures) in this tree is given the highest confidence score. The score goes down along the rank. Fig. 3, shows an example of a sub-set of this tree for organ "Liver" and disease "Fatty Liver". The leaf nodes list the top ranked tests for "Fatty Liver". The tree shows that the specificity of diagnosis of "Fatty Liver" is high with MRI (Magnetic Resonance Imaging) and MRS (Magnetic Resonance Spectroscopy) while the specificity is low with LFT. Hence, a Low (L) confidence score is assigned with DT-Score generated with regular lab test like Liver Function Test (LFT). This tree is further used to derive the mathematical model for DT-Score.

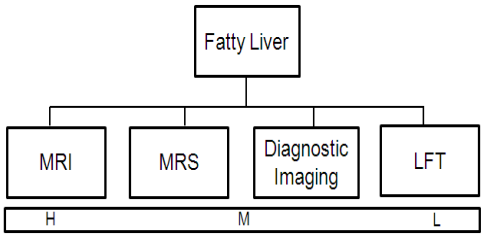

Fig. 3. A sub-set of Organ-Disease-Labtest tree for Fatty Liver with associated confidence values (H-High, M-Medium, L-Low)

B. Model Building for DT-Score Prediction

The popular abnormality classification rules [11] are snapshot based, and not directly suitable for "temporal state"

detection. To address this issue, we have formulated a mathematical model to detect such "temporal state" using DT-Score. We have used real patient data collected during Annual Checkup program of the selected hospital for this purpose. The data also contains followup test results and associated doctor's diagnostic annotations. A disease based temporal score (DT-Score) is derived from doctor's annotations that are available as part of decision report and this score is used as the classification reference for the model. Patients' temporal laboratory test data is used to generate feature vectors for classification.

1) Classification of diseases states and weight selection for each class: DT-Score is represented by different phases of the disease. An empirical weight is assigned to each phase of disease to integrate it with a mathematical model. In the experiment conducted by us, Fatty Liver was taken as the sample disease. Following seven types of phases of Fatty Liver are identified from the doctor's annotation in patient record.

1) Minimal
2) Minimal to mild
3) Mild
4) Mild to moderate
5) Moderate
6) Moderate to severe
7) Severe

A normalization of disease phases are done (refer Table II) to reduce number of transitive phases (like minimal-to-mild, mild-to-moderate etc).

TABLE II
CLASSIFICATION OF DISEASE TYPES

Disease	Specialized Phases	Normalized Phases	Assigned Weight
Fatty Liver	Minimal Minimal to Mild	Minimal	0.25
	Mild Mild to Moderate	Mild	0.5
	Moderate Moderate to Severe	Moderate	0.75
	Severe	Severe	1

2) Preparation of patient data set for training and validation: The selection of the patient set started with 9154 real patients data from the annual checkup program of the selected hospital. Last 10 years data for these patients are taken for analysis. The patient database contained multiple annotations in the form of natural language text (Originally bi-lingual data was present, but in the current experiment only English is used) which describe the snapshot state. Patients are categorized as Fatty and Non Fatty through text mining of the corresponding annotations made by doctors. All possible combinational string for fatty liver, like "moderate fatty liver", "fatty liver moderate" are considered in the text mining. As a result of this process, 5815 patients were found to have diagnosed as "Fatty Liver" at least once in last 10 years. The accuracy of the DT-Score model is highly dependent on the number of available data points, like number of annotations and laboratory instances. To ensure a better model the fatty patient set selected for model training, was further pruned through following two conditions:

1) Patient with minimal 5 annotations in past medical records identifying different Fatty Liver stages.
2) Patients with at least 10 laboratory test (related to Fatty Liver) instances.

Through this condition, the final data set came down to 146 fatty patients. The patient set was further divided into 98 patients as training set and 48 patients as test set. These two sets have near-identical distribution of severity for fatty patients.

3) Calculation of Disease based temporal score: Based on the past diagnosis done across years for a patient, a disease based temporal score (DT-Score) is introduced which indicates disease chronicity. The DT-Score is calculated in the following way:

Step 1 For each patient find how many times any of the four type of Fatty Liver has been diagnosed and use their weights to calculate the instance score (SI) for each type.

$$SI = I * W \tag{1}$$

where I is the number of instance of a particular type/phase.
W is the weight of that type/phase

Step 2 Calculate the temporal score using all the SI and the number of laboratory test instances (M) when all the related laboratory tests are done together.

$$ST = (\sum_{i=1}^{N} SI_i)/M \tag{2}$$

where N is the number of diagnosis done (and having annotation).
M is laboratory test instances when all the related laboratory tests are done together.
This DT-Score will be used for identifying disease chronicity. The patient who never diagnose as Fatty Liver, the score will be zero for them.

4) Selection of Feature Vectors: Selection of feature vectors is important for mathematical model creation. For the proposed model, we need feature vectors which can represent the disease temporal state most efficiently. Methodology discussed in [Subsection II-A] is used to identify the related laboratory tests (with ranks) to represent the temporal state of disease "Fatty Liver". Table III shows the laboratory tests that are selected as feature vectors for Fatty Liver model from associated ranked Organ-Disease-Labtest tree.

5) Representation of feature vectors: Each of the selected laboratory test results were collected as a series of data over time. Feature vectors are derived from these time series data of laboratory test results. The primary aim of the feature vectors

TABLE III
FEATURE VECTORS FOR FATTY LIVER MODEL

Disease	Top Ranked Laboratory Test	Feature Vector
Fatty Liver	Liver Function Test (LFT)	$GGT_{Temporal}$ $AST_{Temporal}$ $ALP_{Temporal}$ $ALT_{Temporal}$

is to highlight the chronicity of the disorder or temporal abnormality. To achieve this, we have used a variation of Standard Deviation called "Standard Deviation using Biological Mean" to represent the time series laboratory test data. In this representation Standard Deviation of the laboratory test values (temporal) are calculated using biological mean of that test, and termed as SD(temporal). Then the calculated SD(temporal) value is multiplied with the number of times the corresponding laboratory test went abnormal. This resultant value is used as the temporal feature for one candidate laboratory parameter.

The data set is, then, broken into two sets, training and test. In the training set, the DT-Score are calculated using the methods described earlier. Fig. 4 depicts the overall process of deriving temporal disease score and feature vectors for training.

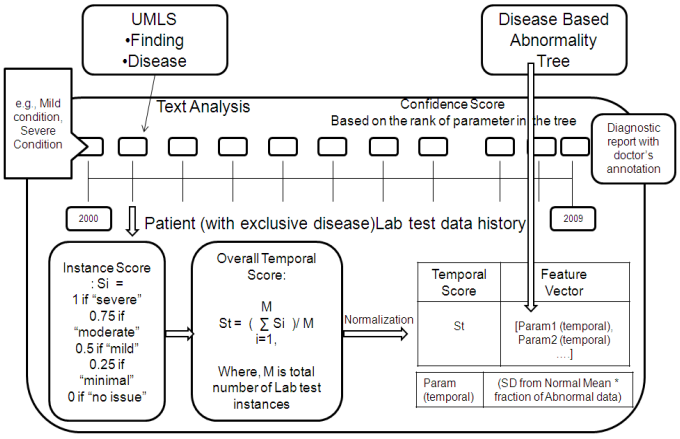

Fig. 4. DT-Score and Feature

6) Model creation using training data: A mathematical model called temporal state analyzer is formulated for a specific disease, based on the training data. The model is trained using historical annotations and feature vectors [refer Fig. 5]. This model can derive the disease based temporal score (DT-Score) using the laboratory test result as input. The final mathematical model, derived through Regression using Support Vector Machine (SVM), is as follows:

$$\begin{aligned} FL\_SCORE \quad = \quad & 0.0772 * (normalized)(ALP_{Temporal}) \\ & -0.0509 * (normalized)(AST_{Trmpotal}) \\ & +0.3212 * (normalized)(ALT_{Temporal}) \\ & -0.059 * (normalized)(GGT_{Temporal}) \\ & +0.3407 \end{aligned} \tag{3}$$

where FL_SCORE represents the estimated Disease Based Temporal Score (DT-Score).

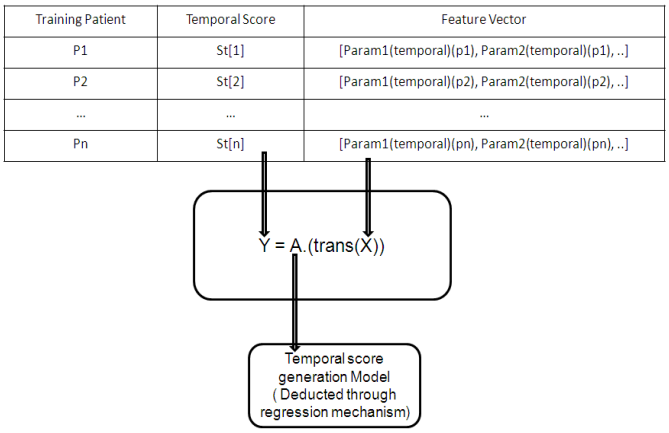

Fig. 5. Disease Model Creation

7) Application of the model: In the execution phase, feature vectors are derived for each disease (of an organ) and is passed through the temporal state analyzer model to calculate the DT-Score. The calculated score is then interpreted to decide the chronicity of the disease. A confidence score [refer Subsection II-A] is also added to the final output using rank of the laboratory tests used in the Organ-Disease-Labtest tree. Fig. 6 depicts the overall execution phase of our proposed model.

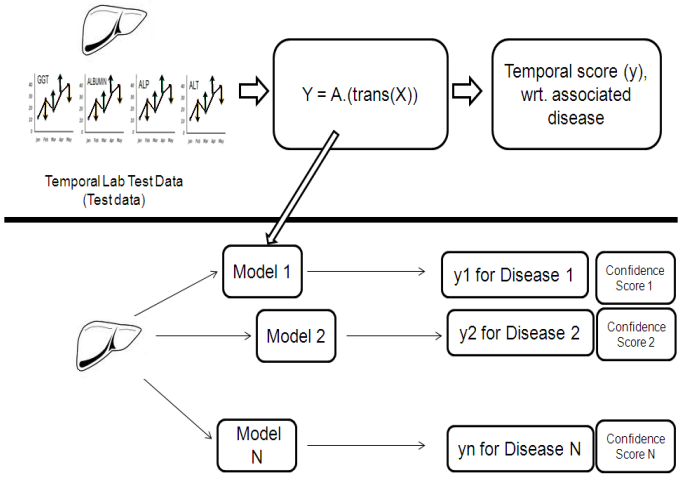

Fig. 6. Disease Model Execution

8) Validation of the model: In the validation phase, we feed the model with the temporal laboratory test data of test patient set and derived the DT-Score for fatty liver. Next, we compare the estimated DT-Score with the calculated score generated from the available annotation for these patients. Table IV shows the summarized results of this validation process. The result shows a better estimation at DT-Score ranges of (20%-30%, 30%-40%), where higher density of training data is available. The error in estimated DT-Score goes up in DT-Score ranges (e.g., 50%-60%) with lower density of training data. On an overall basis, the average absolute error

in estimation stands at 11%, which means a difference of 11% in DT-Score between actual (calculated using doctor's annotation) and derived (calculated using our model).

TABLE IV
VALIDATION RESULT FOR FATTY LIVER MODEL

Actual DT-Score Range(%)	No. of Test Patient	Average Absolute Error(%)	Cumulative Average Absolute Error(%)
10-20	10	14.5	
20-30	13	3.8	
30-40	13	5.2	11
40-50	8	14.6	
50-60	4	25	

C. Visualization of DT-Score using Human Body Model

We have introduced a visualization interface using Human Body Model to represent the DT-Score. The output of the mathematical model is feed to the visualization interface and it shows Disease Based Temporal Score (DT-Score), derived from our model. The interface also represents trend chart of actual result, configurable time window for better reference and control. Fig. 7 depicts the visualization interface for the organ Liver.

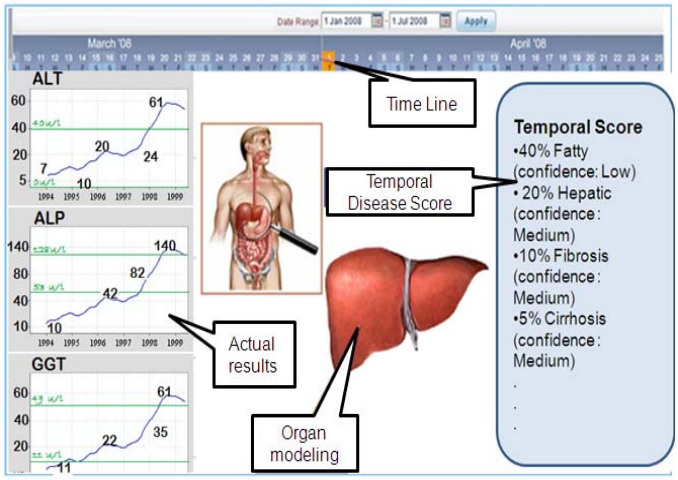

Fig. 7. Human Body Model

III. CONCLUSION & FUTURE WORK

In this paper we have discussed a novel technique to visualize temporal organ state for quick summarization of large periodic laboratory test data, which are otherwise unused. An unique Disease Based Temporal Score (DT-Score) is introduced to effectively represent temporal state of an organ. Each DT-Score is attributed with a confidence value, derived from the rank of the associated lab procedures in UMLSs co-occurrence table (co-occurrence between disease and lab procedure). This score will help in early indication of chronic condition without doing costly diagnostic process, like MRI, Biopsy. A mathematical model is designed to calculate DT-Score of an organ using medical relations from Semantic Network (UMLS) and rank from co-occurrence score (MRCOC table of UMLS). As a future scope, the proposed method can be applied to summarize diverse periodic clinical test data, including and not limited to, tracking and monitoring of patient under specific treatment/clinical procedure. The proposed method can also be extended to self learn the latest medical knowledge through periodic training with incremental version of UMLS.

ACKNOWLEDGMENT

The authors would like to thank Dr. Sanghyun Yoo and Mr. Seungwoo Ryu of SAIT-Korea for valuable review comments and guidance.

REFERENCES

[1] S. B. Cousins, M. G. Kahn, "The Visual Display of Temporal Information", *Artif. Intell. Med.*, vol.3, no.6, pp. 341-357, 1991.
[2] C. Plaisant, B. Milash, A. Rose, S. Widoff, B. Shneiderman, "LifeLines: Visualizing Personal Histories", *In Proc. SIGCHI Conf. on Human factors in computing systems: common ground*, Vancouver, pp. 221-227, 1996.
[3] J. An, X. Lu, H. Duan, "Integrated Visualization of Multi-Modal Electronic Health Record Data", *In Proc. Bioinformatics and Biomedical Engineering*, pp. 640-643, May 2008.
[4] *Polycystic kidney disease* [Online]. Available: http://www.nlm.nih.gov/medlineplus/ency/article/000502.htm
[5] J. Rakela, M. Goldschmiedt, J. Ludwig, "Late manifestation of chronic liver disease in adults with alpha-1-antitrypsin deficiency", *Digest. Dis. Sci.*, vol. 32, no. 12, pp. 1358-1362, 1987.
[6] J. C. Marshall, D. J. Cook, N. V. Christou, G. R. Bernard, C. L. Sprung, W. J. Sibbald, "Multiple Organ Dysfunction score: a reliable descriptor of a complex clinical outcome", *Crit. Care. Med.*, vol.23, no.10, pp. 1638-1652, Oct. 1995.
[7] M. B. Heldwein, A. Badreldin, F. Doerr, T. Lehmann, O. Bayer, T. Doenst, K. Hekmat, "Logisitic Organ Dysfunction score (LODS): A reliable postoperatie risk management score also in cardiac surgical patients ?", *J. Cardiothorac. Surg.*, vol.6, no.1, pp. 1-6, 2011.
[8] A. Bodnar, D. Nekrasovski, "LabVis: Simple and Intuitive Visualization of Medical Laboratory Results", *Infovis 2004*.
[9] A. Lowe, R. W. Jones, M. J. Harrison, "The graphical presentation of decision support information in an intelligent anaesthesia monitor", *Artif. Intell. Med.*, vol. 22, no.2, pp. 171-183, 2001.
[10] H. Lam, "Displaying Medical Laboratory Reports on small screens", M.S. thesis, School of Engineering Science, Simon Fraser University, Apr. 2004.
[11] S. Karthik, A. Priyadarishini, J. Anuradha, B. K. Tripathy, "Classification and Rule Extraction using Rough Set for Diagnosis of Liver Disease and its Types", *Adv. Appl. Sci. Res.*, vol. 2 no. 3, pp. 334-345, 2011.
[12] K. Takabayashi, T. B. Ho, H. Yokoi, T. D. Nguyen, S. Kawasaki, S. Q. Le, T. Suzuki, O. Yokosuka, "Temporal abstraction and data mining with visualization of laboratory data", *Stud. Health. Technol. Inform.*, vol.129, pp. 1304-1308, 2007.
[13] *Fact Sheet Unified Medical Language System* [Online]. Available: http://www.nlm.nih.gov/pubs/factsheets/umls.html
[14] *Unified Medical Language System* [Online]. Available: http://en.wikipedia.org/wiki/Unified_Medical_Language_System
[15] Q. Zeng, J. J. Cimino, "Automated Knowledge Extraction from the UMLS", *In Proc. AMIA Symposium*, pp. 568-572, 1998.
[16] E. Agirre, A. Soroa, M. Stevenson, "Graph-based Word Sense Disambiguation of biomedical documents", *Bioinformatics*, vol.26, no.22, pp. 2889-2896, 2010.
[17] J. C. Denny, J. D. Smithers, R. A. Miller, A. Spickard, "Understanding Medical School Curriculum Content Using KnowledgeMap", *J. Am. Med. Inform. Assn.*, vol.10, pp. 351-362, 2003.