A Study on Semantic Diabetes Diagnosis Based on Age/Gender and Drug Suggestion

Poongothai P1, Dr. T. Devi2

1Ph.D. Research Scholar, Department of Computer Applications, Bharathiar University, Coimbatore, Tamilnadu, India
poongothaimca39@gmail.com
2Professor and Head, Department of Computer Applications, Bharathiar University, Coimbatore, Tamilnadu, India
tdevi@buc.edu.in

ABSTRACT

Diabetes is one of the metabolic disorders that causes the threat for healthcare community and medically doctors call diabetes as diabetes mellitus. Semantic analytics is an ontology usage for identifying the content in web resources. Ontologies are essential for knowledge administration, data integration, semantic interoperability, decision support and reasoning. Ontology protects the semantic link between their concepts and increases intelligence of decision support systems. The existing fuzzy ontology technique relates the patient’s historical data with the web resources in order to diagnose the diabetes. In addition, the diagnosis of diabetes is carried out with age/gender form and drugs are suggested based on the severity of diabetes. Diabetes awareness can be created among the people in order to reduce the risk among patients due to diabetes. Existing literature shows that the diabetes diagnosing efficiency and diabetes awareness rate are not improved much. In this work, different diabetes diagnosing techniques are studied and their drawbacks are listed. This research work aims to identify the diabetics at earlier stage and suggest drugs as well as create diabetic awareness among people.

Keywords: Diabetes, Ontology, Semantic Analytics, Diabetic Awareness, Decision Support Systems, Age/Gender.

I. INTRODUCTION

Diabetes is a public health problem due to the high prevalence, induced difficulties and increasing health care cost [14]. Diabetes is a chronic illness in requirement of continuous medical care with multifactorial risk-reduction plan beyond glycemic control. Diabetes is divided into three types, namely type 1 diabetes, type 2 diabetes and gestational diabetes mellitus. Type 1 diabetes is otherwise known as Insulin dependent Diabetes Mellitus (IDDM) as the patient’s require injecting or taking the insulin directly because of not sufficient production of insulin from body [15]. Type 2 diabetes is otherwise termed as Non-insulin dependent diabetes mellitus (NIDDM) in which the patient’s cells failed to react for the insulin produced in body. Gestational diabetes takes place in pregnant women because of rise in glucose level [15]. Diabetes mellitus causes Retinopathy (eye disease), Nephropathy (kidney disease), and Neuropathy (nerve disease) [16]. The main contribution of the work is to study the existing issues of diabetes diagnosis. In addition, this study helps in future work by identifying the diabetics at earlier stage and suggesting drugs as well as creating diabetic awareness among people.

This paper is ordered as follows: Section II discusses reviews on diabetes diagnosis with awareness, Section III describes the existing diabetes diagnosis, Section IV explains diagnosis of diabetes based on age/gender with drug suggestion and possible comparison, Section V describes the diabetes awareness with performance analysis comparison, Section IV explains the limitation of as well as future work.

II. LITERATURE REVIEW

Diabetes mellitus (DM) is described as collection of metabolic disorders exerting essential pressure on human health worldwide [17]. A new development to diabetes diagnosis CBR was designed with new case-base fuzzy OWL2 ontology (CBRDiabOnto) [1]. The
ontology has been taken as first fuzzy case-base ontology in medical domain with case-base fuzzy Extended Entity Relation (EER) data model. However, the diabetic diagnosing time was not reduced. The ontology with more cases failed to increase the case retrieval accuracy. A fuzzy ontology-based CBR framework was introduced with fuzzy case-base OWL2 ontology and fuzzy semantic retrieval algorithm by researchers [2]. The designed framework manages many feature types. The combination of CBR with EHR environment increases the automation of decision support process. But, the remaining CBR steps were not used in case adaptation process.

For guaranteeing the annotation consistency in entire corpus, harmonization tools were used to recognize the annotation differences addressed by improving the overall corpus quality [3]. The features of diabetes were resulted in generation of large amount of data. The associations between age, diabetes duration and main macrovascular events, death cause and microvascular events were observed with type 2 diabetes assigned to standard glucose control in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation (ADVANCE) Trial [4]. Logistic regression analysis has been introduced to find out the factors with diabetes diagnosis and late diabetes diagnosis [5]. The study explained in T2D population showed novel oral GLDs, dapagliflozin or DPP-4 inhibitors with lower risk of all-cause mortality, CVD and hypoglycaemia [6].

The patient’s reactions to diagnosis of type 2 diabetes mellitus (T2DM) were monitored with health related quality of life [7]. A qualitative exploratory was carried out with thematic analysis. However, the diabetes awareness rate was not improved. In order to improve the diabetes awareness rate, the evidence from randomized controlled trials (RCTs) on CHITs are examined by meta-analysis or narrative synthesis approach [8]. A systematic search of seven electronic databases recognized the relevant reports of RCTs for analysis. A systematic review of machine learning and data mining techniques were carried out with Prediction and Diagnosis, Diabetic Complications, Genetic Background and Environment, Health Care and Management [9]. However, the diabetes was not efficiently diagnosed. The analysis of multi-stage natural language processing system was carried out with entity recognition, Bayesian statistics and rule logic for recognizing the heart disease risk factor events in diabetic patients [10].

Many factors with diabetes and late diabetes diagnosis from males and females were analyzed [11]. But, the risk factors to impact males and female was not computed in effective manner. K-Nearest neighbor’s algorithm with fuzzy K-nearest neighbors was designed to improve the diabetes diseases diagnosis accuracy [12]. However, the algorithm failed to increase the disease diagnosing efficiency. An intelligent naïve Bayes approach based system was introduced for diabetes diagnosis [13]. However, the approach failed to reduce the diabetes diagnosing time.

### III. Diabetic Diagnosis

Diabetes mellitus is a metabolic disorder classified by presence of hyperglycemia because of defective insulin production and defective insulin action [18]. The chronic hyperglycemia of diabetes is linked with long-term microvascular difficulties affecting eyes, kidneys and nerves for cardiovascular disease (CVD) [19]. The diagnostic principles for diabetes are used with the glycemia threshold and microvascular disease. A diabetes diagnosis follows the person for medical advice after experiencing symptoms of diabetes like feeling thirsty, frequent urination, fatigue and unexpected weight loss. Type 2 is diagnosed by diabetes symptoms like polyuria and polydipsia.

#### A. Case base knowledge in diabetes mellitus domain in a fuzzy ontology model

Diabetic mellitus is a collection of metabolic disease where person has high blood sugar because of not sufficient insulin production by pancreas or cells and failed to react for insulin that produced by body. For addressing the experience-based issues, case based reasoning (CBR) is an essential AI technique for decision support. CBR imitates human reasoning when it failed to create the generalized rules. CBR model is used in different medical fields from lung disease through eating disorders to diabetes and Alzheimer’s disease [20]. Ontologies increase the capabilities of CBR through creating knowledge intensive-CBR (KI-CBR) systems. It is mainly used in CBR like background domain ontology, case-base ontology, semantic similarity measurement etc.
Knowledge-Intensive Case-Based Reasoning Systems (KI-CBR) are based on ontologies [1]. Ontology is an essential task for case-base knowledge. The mixture of ontology and fuzzy logic reasoning is an important one in medical domain. Case-base representation is carried out depending on fuzzy ontology to increase the semantic and storage of CBR knowledge-base. The diabetes diagnosis of CBR is carried out using case-base fuzzy OWL2 ontology (CBRDiaabOnto). By means of case-base fuzzy Extended Entity Relation (EER) data model, it is considered as the fuzzy case-base ontology.

Knowledge Intensive Case-Based Reasoning (KI-CBR) is carried out with Case Base (CB) structure and content [1]. Ontology collects the domain knowledge in machine-readable format where humans can also understand. It is employed in CBR systems as knowledge representation formalism for representing the domain background knowledge and CB knowledge. KI-CBR allocates the automatic reasoning with semantic knowledge and syntactic property. Ontology maintains the creation of semantic retrieval algorithms to improve intelligence of CBR systems. It presents the conceptualization of domain with concepts, properties and axioms. Each case has a description and solution.

Case structure = \{P, S(P)\} \hspace{1cm} (1)

From (1), ‘P’ is the problem description including patient symptoms and lab test results.

P=LiF, LiP, GIS, Age, BMI, Residence, GenderOccupation, KiF, Lab, UrS, HaP \hspace{1cm} (2)

Where, ‘LiF’ denotes liver function tests, ‘LiP’ represents lipid profile, ‘GIS’ symbolizes global symptoms, ‘KiF’ represents kidney function tests, ‘Lab’ denotes lab tests, ‘UrS’ represents urination symptoms, ‘HaP’ denotes haematological profile. The solution S(P) is described by five features as shown in (3).

\[
S(P) = \{D, L, N, C, H\} \hspace{1cm} (3)
\]

Where, ‘D’ represents DM diagnosis, ‘L’ represents liver problem, ‘N’ denotes nephropathy problem, ‘C’ represents cancer type and ‘H’ denotes hypercholesterolemia issues [1]. A standard CB data model is introduced with HL7 Reference Information Model (RIM) and SNOMED CT (SCT) ontology. Each patient is represented by row or vector in data model. The patient is illustrated with features and CBR system suggests diagnosis of patient conditions. The features are string data types like sex and residence. The most of the features are numerical called fuzzy attributes as described in figure 1.

![Fuzzy Ontology Construction Framework](image)

**Figure 1. Fuzzy Ontology Construction Framework**

A case base crisp EER model generated is extended to fuzzy EER model (Fig. 1). Each crisp value is fuzzified into all of the attribute’s linguistic values. For supporting the similarity calculation by means of similarity matrices, a single fuzzy label is used to explain feature crisp value. Then, the generated fuzzy EER model is mapped to FO structure (TBOX) by means of formal methodology. Finally, generated ontology is populated with requests from fuzzy database (ABOX) same as fuzzy EER model [1].

**B. Case-based reasoning framework for semantic diabetes diagnosis**

Case-based reasoning (CBR) is a problem-solving technique that employs past knowledge to interpret or address new issues. It is appropriate for experience-based and theory-less issues. Knowledge-intensive CBR (KI-CBR) by formal ontologies is the development of new model. The case representation and storage are utilized by Ontologies. By means of standard medical ontologies like SNOMED CT, the interoperability is improved in health care systems. With help of vague or imprecise knowledge, the efficiency of CBR semantic is improved. A fuzzy ontology-based CBR framework is designed with fuzzy case-baseOWL2 ontology and fuzzy semantic retrieval algorithm for managing different feature types.

A fuzzy KI-CBR framework uses an imprecise knowledge through the efficient integration of fuzzy logic in ontology-based CBR model [2]. Fuzzy case-base ontology and fuzzy semantic retrieval algorithm are designed with intelligent CBR for diabetes diagnosis. A fuzzy semantics of CBR are employed in two places.
Initially, the algorithm is used in demonstration of inaccurate knowledge and secondly in case retrieval. A fuzzy ontology controls the demonstration of inaccurate case-specific knowledge while retrieval of cases is enabled with fuzzy semantic similarity framework. The designed system is carried out with six modules, namely Case source preparation, case base ontology engineering, terminology server, fuzzy case-base ontology population, case retrieval engine and case query parser. The designed system comprises user-friendly interface for selection of standard medical concepts from SCT dialog and employs the clinical distance in case retrieval process.

An efficient way to develop the case-base fuzzy ontology is discussed in fuzzy KI-CBR framework [2]. The ontology is designed with the crisp ontology and the top-level CBR crisp ontology namely CBROnt. The designed fuzzy ontology is mainly used in medical domain. A fuzzy semantic retrieval algorithm is employed for retrieving the cases from fuzzy ontology consistent with the issues. The hybrid algorithm (i.e., Fuzzy case-base ontology and fuzzy semantic retrieval algorithm) is an exact and considers patient’s features with numerical, fuzzy, lexical, and semantic types. The fuzzy types are denoted in fuzzy ontology and semantic types are depending on standard diabetes diagnosis SCT ontology [2].

C. Medical Entity and Relation LIMSI annOtated Text corpus (MERLOT)

Medical Entity and Relation LIMSI annOtated Text corpus (MERLOT) is employed in large high-quality corpus of clinical documents with entities, attributes and relations [3]. A simple rule-based algorithm has been designed to reconstruct the split lines inside a paragraph or sentence. A typology in clinical notes is used to describe the contents of documents. Four high-level types are discussed namely, generic header with contact information for health care unit, specific header with information, core medical content of note and footer with physician’s signature. The annotation scheme was introduced to present broad coverage of clinical domain for the annotation of medical events in the clinical documents. Semantic annotations in the scheme include the entities, attributes, relations between entities, and temporal annotations.

D. Performance Analysis of Diabetic Diagnosis

In order to compare different techniques for diabetic disease diagnosis, number of test data (i.e., patient data) is taken to perform the experiment. The performance analysis of different diabetic diagnosis techniques is carried out using common dataset called Diabetes 130-US Hospitals for 1999-2008 Dataset from UCI Machine learning repository. Diabetes 130-US Hospitals for 1999-2008 Dataset contains 55 attributes sets and 100000 instances. The results are obtained by using above mentioned dataset. Various parameters used for diabetic diagnosis are true positive rate and diabetic diagnosis time.

1) True Positive Rate (TPR)

True positive rate is defined as the ratio of number of data that are correctly identified as diabetic disease affected data to the total number of test data. It is measured in terms of percentage (%).

\[
TPR = \frac{\text{number of data correctly identified as diabetic disease affected data}}{\text{total number of test data}} \times 100
\]

When the true positive rate is higher, the method is said to be more efficient.

For Example:

Case-base Fuzzy EER Data Model: \( TPR = \frac{8}{10} \times 100 = 80\% \)

Fuzzy KI-CBR Framework: \( TPR = \frac{7}{10} \times 100 = 70\% \)

MERLOT: \( TPR = \frac{6}{10} \times 100 = 60\% \)

<table>
<thead>
<tr>
<th>Number of test data (Number)</th>
<th>True Positive Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-base Fuzzy EER Data Model</td>
<td>Fuzzy KI-CBR Framework</td>
</tr>
<tr>
<td>10</td>
<td>79</td>
</tr>
<tr>
<td>20</td>
<td>81</td>
</tr>
<tr>
<td>30</td>
<td>83</td>
</tr>
<tr>
<td>40</td>
<td>86</td>
</tr>
<tr>
<td>50</td>
<td>87</td>
</tr>
<tr>
<td>60</td>
<td>89</td>
</tr>
<tr>
<td>70</td>
<td>91</td>
</tr>
<tr>
<td>80</td>
<td>93</td>
</tr>
<tr>
<td>90</td>
<td>94</td>
</tr>
<tr>
<td>100</td>
<td>95</td>
</tr>
</tbody>
</table>

Table 1. Tabulation of True Positive Rate
Table 1 shows the true positive rate with respect to number of test data ranging from 10 to 100. True Positive Rate comparison has been carried on existing case-base fuzzy Extended Entity Relation (EER) data model [1], fuzzy Knowledge-intensive Case-based reasoning (KI-CBR) [2] and Medical Entity and Relation LIMSI annotated Text corpus (MERLOT) [3].

![Figure 2. Measurement of True Positive Rate](image)

From figure 2, true positive rate of existing techniques are compared. When the number of test data is increased, the true positive rate gets increases correspondingly. True positive rate of case-base fuzzy Extended Entity Relation (EER) data model [1] is comparatively higher than that of fuzzy Knowledge-intensive Case-based reasoning (KI-CBR) [2] and Medical Entity and Relation LIMSI annotated Text corpus (MERLOT) [3]. Research in case-base fuzzy Extended Entity Relation (EER) data model has 10.8% higher true positive rate than fuzzy Knowledge-intensive Case-based reasoning and 18.8% higher true positive rate than Medical Entity and Relation LIMSI annotated Text corpus.

2) Diabetic Diagnosis Time (DDT)

DDT is defined as the amount of time taken for diagnosing the diabetic disease from the given number of test data and it is measured in terms of milliseconds (ms).

\[ DDT = \text{number of test data} \times \text{Time (diagnosing time for one test data)} \]

When the disease diagnosing time is lesser, the efficiency of the method is more.

For Example:

**Case-base Fuzzy EER Data Model:** \[ DDT = 10 \times 5.6ms = 56ms \]

**Fuzzy KI-CBR Framework:** \[ DDT = 10 \times 4ms = 40ms \]

**MERLOT:** \[ DDT = 10 \times 6.7ms = 67ms \]

Table 2 shows the disease diagnosing time with respect to number of test data ranging from 10 to 100. Disease diagnosing time comparison takes place on existing case-base fuzzy Extended Entity Relation (EER) data model [1], fuzzy Knowledge-intensive Case-based reasoning (KI-CBR) [2] and Medical Entity and Relation LIMSI annotated Text corpus (MERLOT) [3].

![Figure 3. Measurement of Diabetic Diagnosis Time](image)

From figure 3, disease diagnosing time of existing techniques is compared. When the number of test data gets increased, the disease diagnosing time gets increased correspondingly. Disease diagnosing time consumption of fuzzy Knowledge-intensive Case-based reasoning (KI-CBR) [2] is comparatively lesser than that of case-base fuzzy Extended Entity Relation (EER) data model [1] and Medical Entity and Relation LIMSI annotated Text corpus (MERLOT) [3]. Research in fuzzy Knowledge-intensive Case-based reasoning (KI-CBR) has 18% lesser disease diagnosing time consumption than case-base fuzzy Extended Entity Relation (EER) data model and 32% lesser disease diagnosing time consumption than Medical Entity and Relation LIMSI annotated Text corpus.
IV. Diabetes Diagnosis Based on Age/Gender and Drug Suggestions

Diabetes is a disease where the body failed to utilize the energy it gets from the food [21]. The food is broken down or digested into sugar or glucose. Glucose presents the body's cells with the energy they need. Insulin is the hormone created in pancreas that helps the glucose gets inside cells where glucose is burned for energy. Symptoms of diabetes comprises the frequent urination and thirsty than usual. Diabetes gradually increases with age. Many diabetes drugs are prescribed for people with type 2 diabetes who are not able to control blood sugar levels by following the strict diet and exercise. Metformin is the drug used in insulin treatment for people with the type 1 diabetes [22].

A. Diabetes Diagnosis and Late Diabetes Diagnosis for Males and Females

Type 2 diabetes is used for diagnosing the disease to prevent the progression. Type 2 diabetes presents 9-12 years before it gets diagnosed and many problems are faced during the diagnosis process. Canadian Chronic Disease Surveillance System (CCDSS) is a collaborative network of simple and defensive surveillance systems and employs the validated case definition to identify the individuals with diabetes. Medical Care Plan (MCP) system includes the information with the services presented by fee-for-service physicians under provincial MCP[4]. Many factors are connected with diabetes and these factors change for males and females with respect to early and late detection.

For categorization of individuals diagnosed at earlier stage and later stage, records with diabetes are connected to the MCP and CDMS data. These are mainly used to recognize when the physician visits for diabetes related comorbidities. Incident diabetes patients devoid of diabetes are linked to comorbidities. Characteristics of population are denoted as weighted percentages and compared the individuals with and without diabetes. It is diagnosed at earlier and later stage with diabetes using chi-square tests and t-tests. For identifying the factors with diabetes diagnosis and late diabetes diagnosis, logistic regression analysis computes the odds ratios (OR) [4].

B. Age at Diagnosis and Duration of Diabetes on Macrovascular and Microvascular Complications Risks

Data are not consistent about links with age, age at diagnosis of diabetes, diabetes period and vascular complications. The associations between age diabetes duration and macrovascular results, microvascular events were identified in patients with type 2 diabetes that are assigned to standard glucose control in Diabetes and Vascular Disease. A spearman correlation identifies the correlation between key variables. The recorded time for all participants is calculated from registration date to date of event. Cox proportional hazard models are computed with age or age at diagnosis and diabetes duration by two-step approach [5]. The first event is used for every outcome that comprised in analysis. The crude rates and risk models are computed for randomized treatments. Intensive glycaemic control of people diagnosed with type 2 diabetes is guaranteed to reduce the microvascular complications problems. The designed analysis observes the individual outcomes of age or age at diagnosis and duration of diabetes on macrovascular andmicrovascular events that causes mortality with type 2 diabetes.

C. Novel Oral Glucose-Lowering Drugs for Patients with Type 2 Diabetes

Metformin is the first-line drug treatment for many patients with type 2 diabetes (T2D). After changing the metformin time, patients with T2D require an intensified treatment due to the disease progression and inadequate glycaemic control [6]. Glucose-lowering drug (GLD) is used for ongoing glucose-lowering therapy with pharmacological treatments like insulin, dipeptidyl peptidase-4 (DPP-4) inhibitors, sulphonylureas, sodium-glucose co-transporter-2 (SGLT2) inhibitors, thiazolidinediones, acarbose or glucagon-like peptide-1 (GLP-1) receptor agonists.

The main aim is to study novel oral GLDs, namely SGLT-2 inhibitors and DPP-4 inhibitors with variation in mortality problems, CVD events or hypoglycaemia [6]. All patients with filled prescription are carried out for DPP-4 inhibitors or SGLT2 inhibitors termed novel GLD group of treatments or insulin. The index date describes the date of first filled prescription. Patients with diagnosis of gestational diabetes with one year duration of index date and patients with type 1 diabetes are rejected. Patients with type 1 diabetes are described
with type 1 diabetes diagnosis and treated with insulin in first year of GLD. The patients aged lesser than 30 years are given insulin treatment and patients who aged lesser than 15 years are given diabetes medication. Patients with index date for drug classes in novel group are joined in SGLT2 inhibitor group and in DPP-4 inhibitor group.

D. Performance Analysis of Diabetic Diagnosis based on Age/ Gender and Drug Suggestion

In order to compare different techniques for diabetic disease diagnosis based on age/gender with drug suggestion, number of test data (i.e., patient data) is taken to perform the experiment. The performance analysis of different diabetic diagnosis based on age/gender and drug suggestion is carried out Diabetes 130-US Hospitals for 1999-2008 Dataset from UCI Machine learning repository. Various parameters used for diabetic diagnosis are memory consumption and diabetic diagnosis efficiency.

1) Memory Consumption (MC)

Memory consumption is defined as the amount of memory space used to store the patient’s data for diabetic diagnosis. It is measured in terms of megabytes (MB). When the memory consumption is lesser, the method is said to be more efficient.

\[ Memory\ Consumption = Number\ of\ test\ data \times Memory\ consumed\ by\ one\ data \]

For Example:
- Logistic Regression Analysis: \( MC = 10 \times 3.9MB = 39MB \)
- Cox Proportional Hazard Models: \( MC = 10 \times 5.6MB = 56MB \)
- Novel oral GLDs: \( MC = 10 \times 6.5MB = 65MB \)

**Table 3. Tabulation of Memory Consumption**

<table>
<thead>
<tr>
<th>Number of test data (Number)</th>
<th>MC (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>39</td>
</tr>
<tr>
<td>20</td>
<td>43</td>
</tr>
<tr>
<td>30</td>
<td>46</td>
</tr>
<tr>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>50</td>
<td>53</td>
</tr>
<tr>
<td>60</td>
<td>56</td>
</tr>
<tr>
<td>70</td>
<td>59</td>
</tr>
</tbody>
</table>

Table 3 shows the memory consumption with respect to number of test data ranging from 10 to 100. Memory consumption comparison for Logistic Regression Analysis [4], Cox Proportional Hazard Models [5] and Novel oral Glucose-Lowering Drugs (GLD) has been carried out [6].

\[ DDE = \frac{number\ of\ diabetic\ data\ detected\ and\ suggested\ drugs}{Total\ number\ of\ test\ data} \times 100 \]

For Example:
- Logistic Regression Analysis: \( DDE = \frac{6}{10} \times 100 = 60\% \)
Cox Proportional Hazard Models: \[ \text{DDE} = \frac{7}{10} \times 100 = 70\% \]

Novel oral GLDs: \[ \text{DDE} = \frac{8}{10} \times 100 = 80\% \]

Table 4. Tabulation of Diabetic Diagnosing Efficiency

<table>
<thead>
<tr>
<th>Number of test data (Number)</th>
<th>Diabetic Diagnosing Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Logistic Regression Analysis</td>
</tr>
<tr>
<td>10</td>
<td>65</td>
</tr>
<tr>
<td>20</td>
<td>67</td>
</tr>
<tr>
<td>30</td>
<td>69</td>
</tr>
<tr>
<td>40</td>
<td>72</td>
</tr>
<tr>
<td>50</td>
<td>74</td>
</tr>
<tr>
<td>60</td>
<td>76</td>
</tr>
<tr>
<td>70</td>
<td>78</td>
</tr>
<tr>
<td>80</td>
<td>81</td>
</tr>
<tr>
<td>90</td>
<td>83</td>
</tr>
<tr>
<td>100</td>
<td>84</td>
</tr>
</tbody>
</table>

Table 4 shows the diabetic diagnosing efficiency with respect to number of test data ranging from 10 to 100. Diabetic diagnosing efficiency comparison has been done for methods such as Logistic Regression Analysis [4], Cox Proportional Hazard Models [5] and Novel oral Glucose-Lowering Drugs (GLD) [6].

From figure 5, diabetic diagnosing efficiency of existing techniques are compared. When the number of test data gets increased, the diabetic diagnosing efficiency gets increased correspondingly. Diabetic diagnosing efficiency of Novel oral Glucose-Lowering Drugs (GLD) [6] is comparatively higher than that of Cox Proportional Hazard Models [5] and Logistic Regression Analysis [4]. Research in Novel oral Glucose-Lowering Drugs has 17% higher diabetic diagnosing efficiency than Logistic Regression Analysis and 10% higher diabetic diagnosing efficiency than Cox Proportional Hazard Models.

A. Consequence of Adjusting to Health Related Quality of Life

The main aim is to search the patient’s reactions for diagnosis of type 2 diabetes mellitus (T2DM) and health related quality of life (HRQOL) model [7]. A qualitative exploratory study design is introduced for the thematic analysis. Twelve patients with T2DM are interviewed through semi-structured interview guide. The in-depth interviews are audio-taped and recorded verbatim tracked through line-by-line coding to recognize the themes. The mixed feelings regarding the diagnosis of T2DM are shared between the patients. Six areas where quality of life emerged from interviews, namely physical and social functioning, work function and social obligations, dietary freedom and conforming to treatment standard as described in figure 6.

Diabetes management needs the drug therapy and addresses the social determining factors with change in patient’s lifestyle [7]. It aims on clinical results and patient’s perceived results that reflect person’s quality of life and psychological problem on her/his daily life. Health Related Quality Of Life (HRQOL) model is explained or determined by three techniques, namely standard disease-specific instruments, generic instruments and utility instruments [7]. HRQOL is a subjective assessment in cultural, social and an environmental context. Many disease-specific instruments are used to calculate the HRQOL for many chronic illnesses or diabetes HRQOL studies.

B. Meta-Analysis and Narrative Review of Randomized Controlled Trials

The evidence from randomized controlled trials (RCTs) on results of consumer health information technologies (CHITs) on patient outcomes was examined by Calvin
et al., through meta-analysis or narrative synthesis approach [8]. A systematic search of seven electronic databases was carried out to identify the relevant details of RCTs for analysis. In meta-analysis, standardized mean differences in patient outcomes are calculated and random-effects models are used in many areas. The heterogeneity of results was moderate or high and fixed-effects models are employed. An article screening and selection processes were systematic and comprehensive to detect the relevant RCTs. A large number of eligible trials with high quality increase the review performances. The narrative synthesis revealed the evidence about the results of CHITs on behavioral and knowledge outcomes. CHITs failed to increase the patient outcomes such as fasting blood glucose level, body weight, body mass index, high-density lipoprotein (HDL) cholesterol level, low-density lipoprotein (LDL) cholesterol level, quality of life, depression and it was unclear as well as efficiency was not observed.

C. Performance Analysis of Diabetic Disease Awareness

In order to compare different techniques for diabetic disease awareness, number of test data (i.e., patient data) is taken to perform the experiment. The performance analysis of different diabetic disease awareness technique is carried out with the help of common dataset called Diabetes 130-US Hospitals for 1999-2008 Dataset from UCI Machine learning repository. Various parameters used for diabetic awareness are false positive rate and diabetic awareness rate.

1) False Positive Rate (FPR)

False positive rate is defined as the ratio of number of data that are incorrectly identified as diabetic disease affected data to the total number of test data. It is measured in terms of percentage (%).

\[
FPR = \frac{\text{number of data incorrectly identified as diabetic disease affected data}}{\text{affected data}} \div \text{total number of test data}
\]

When the false positive rate is lesser, the method is said to be more efficient.

For Example:

HRQOL Model : \( FPR = \frac{6}{10} \times 100 = 60\% \)

Meta-analysis Approach : \( FPR = \frac{7}{10} \times 100 = 70\% \)

Table 5. Tabulation of False Positive Rate

<table>
<thead>
<tr>
<th>Number of test data (Number)</th>
<th>False Positive Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HRQOL Model</td>
</tr>
<tr>
<td>10</td>
<td>65</td>
</tr>
<tr>
<td>20</td>
<td>68</td>
</tr>
<tr>
<td>30</td>
<td>71</td>
</tr>
<tr>
<td>40</td>
<td>73</td>
</tr>
<tr>
<td>50</td>
<td>76</td>
</tr>
<tr>
<td>60</td>
<td>79</td>
</tr>
<tr>
<td>70</td>
<td>81</td>
</tr>
<tr>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>90</td>
<td>87</td>
</tr>
<tr>
<td>100</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 5 shows the false positive rate with respect to number of test data ranging from 10 to 100. False positive rate comparison takes place on existing Health Related Quality Of Life (HRQOL) Model [7] and Meta-analysis Approach [8].

Figure 7. Measurement of False Positive Rate

From figure 7, false positive rates of existing techniques are compared. When the number of test data gets increased, the false positive rate gets increased correspondingly. False positive rate of Health Related Quality Of Life (HRQOL) Model [7] is comparatively lesser than that of Meta-analysis Approach [8]. Research in Health Related Quality Of Life (HRQOL) Model has 6.7% lesser false positive rate than Meta-analysis Approach.

2) Diabetic Awareness Rate

Diabetic awareness rate is defined as the rate at which the diabetes awareness was created among people. When the awareness is created, the people affected by
Table 6 shows the diabetic awareness rate comparison takes place on existing Health Related Quality Of Life (HRQOL) Model [7] and Meta-analysis Approach [8]. Table 6. Tabulation of Diabetic Awareness Rate

<table>
<thead>
<tr>
<th>Approach/Model</th>
<th>Diabetic Awareness Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRQOL Model</td>
<td>85</td>
</tr>
<tr>
<td>Meta-analysis Approach</td>
<td>92</td>
</tr>
</tbody>
</table>

Figure 8. Measurement of Diabetic Awareness Rate

Figure 8, shows the graphical representation of diabetic awareness rate of existing techniques. When the number of test data gets increased, the diabetic awareness rate gets increased correspondingly. Diabetic awareness rate of Meta-analysis Approach [8] is comparatively lesser than that of Health Related Quality Of Life (HRQOL) Model [7]. Research in Meta-analysis Approach has 7% higher diabetic awareness rate than Health Related Quality Of Life (HRQOL) Model.

VI. Discussion on Limitation of Diabetic Disease Diagnosis using Different Techniques

A case-base fuzzy OWL2 ontology (CBRDiabOnto) was designed for diabetes diagnosis of CBR. Fuzzy case-base ontology is taken with case-base fuzzy Extended Entity Relation (EER) data model [1]. However, the diabetic diagnosing time was not reduced and case retrieval accuracy is not increased. A fuzzy ontology-based CBR framework with fuzzy case-base OWL2 ontology and fuzzy semantic retrieval algorithm controls many feature types [2]. The mixture of CBR with EHR improves automation of decision support process. But, remaining CBR steps are not used in case adaptation process.

Cox proportional hazard models evaluated with age or age at diagnosis and diabetes period with help of two-step approach [5]. The first one is used for every outcome in analysis. The crude rates and risk models are computed for randomized treatments in second one. However, the diabetic diagnosing efficiency was not improved. For recognizing the factors with diabetes diagnosis and late diabetes diagnosis, logistic regression analysis is designed with odds ratios (OR) [4]. But, the false positive rate is not reduced below certain level.

The patient’s reactions are identified for the diagnosis of type 2 diabetes mellitus (T2DM) and health related quality of life (HRQOL) model [7]. A qualitative exploratory design was introduced for thematic analysis. Randomized controlled trials (RCTs) of consumer health information technologies (CHITs) on patient results were examined by means of meta-analysis or narrative synthesis approach [8]. The narrative synthesis demonstrated the evidence concerning the results of CHITs on behavioral and knowledge results. However, CHITs failed to increase the patient diabetic diagnosing efficiency and it was unclear as well as efficiency was not observed.

A. Future Direction

The future direction of diabetic diagnosis is to increase the diabetic diagnosing efficiency and diabetic awareness rate through finding the semantic data similarity with patient history. In addition, the drugs can be suggested to the patients based on age/gender in presence of diabetics.

VII. Conclusion

A comparison of different techniques for semantic diabetic diagnosis process is carried out. From the survival study, case-base fuzzy OWL2 ontology (CBRDiabOnto) failed to reduce the diagnosing time. Oral GLDs with variation in mortality problems are addressed. The existing randomized controlled trials (RCTs) of consumer health information technologies (CHITs) on patient results with meta-analysis or narrative synthesis approach failed to improve the diabetic diagnosing efficiency. The wide range of experiments on existing techniques and the comparative performance of the various semantic diabetic diagnosis techniques and its drawbacks are discussed. Finally, from the result, the research work can be carried out in the semantic diabetic diagnosis with higher diabetic diagnosis to improve the diabetic diagnosis efficiency.
VIII. ACKNOWLEDGEMENT

The authors thank the University Grant Commission for providing Rajiv Gandhi National Fellowship to support this research by research fellowship.

IX. REFERENCES


[7]. Lee Lan Low, Seng Fah Tong and Wah Yun Low, “Mixed Feelings about the Diagnosis of Type 2 Diabetes Mellitus: A Consequence of Adjusting to Health Related Quality”, Collegium Antropologicum, NCBI Journals, 38(1), Mar 2014, pp. 11-20.


