

Research Article

DBA's Perspective on Use of Information Visualization in Electronic Health Records

Muhammad Sheraz Arshad Malik and Suziah Sulaiman

Department of Computer and Information Sciences, Universiti Teknologi PETRONAS,
Bandar Seri Iskandar, 31750, Tronoh, Perak, Malaysia

Abstract: DBA's perspective for Information Visualization (IV) assists in providing a detailed insight about patient history in Electronic Health Records (EHR). These data visualizations are used by different stake holders in public hospitals such as Doctors, Database Administrators (DBA) and visual designers in order to improve health services for patients. This is not possible without the contribution of DBAs, who are directly involved in launching efficient query system and required data retrieval for resulting visualization. EHR database is comprising of different forms of data i.e., numbers, texts, pictures and graphs resulting in a challenge for creating a meaningful visualization for single or multiple patients. DBAs have EHR knowledge gaps, lesser exposure to the transformation process of data into effective visualization and understanding about IV limitations result in poor and less efficient IV tools. This study highlights the challenging factors that affect DBAs in understanding, developing and controlling EHR visualization. These factors are selected after a literature review and further analysis by a questionnaire based study on DBAs in two different government hospitals. This will help in highlighting the limitations of DBAs skills and experience and assist in determining the contribution of their role in the development of future IV applications. This work is a second phase of our ongoing research for designing a model, CARE 1.0, as a future IV solution in multiple EHR.

Keywords: CARE 1.0, EHR, information visualization, IV knowledge areas, patient data visualization, visual interfaces

INTRODUCTION

Visual Analytic Systems (VAS) are used to represent different important patient history facts or events in Electronic Health Records (EHR), in a very intricate and user friendly manner. Analysis of such Information Visualization (IV) systems always revealed the fact that design of such graphic applications should consider human sensory capabilities in such a fashion so that important data elements and data patterns can be perceived efficiently by Ware (2013) and Malik and Sulaiman (2013a). The primary objective of IV in EHR is to facilitate the end users that can be mostly doctors and physicians and hospital management professionals for understanding the patient history and data trends in an easier format using different colors and pictorial shapes by Mane *et al.* (2012). Perception of such knowledge is considered majorly on the requirements of doctors while ignoring the important factors that are coming as other constituent components and stake holders e.g., EHR database, Data Base Administrators (DBA) and Visual designers by Wilcox *et al.* (2010a).

DBA plays a vital role as IT professional in the whole system and considered as a potential stake holder by Krist *et al.* (2012). Usually, in an IV

analytical application for EHR, three stake holders are involved:

- Doctors
- DBA
- Visual designer

Most of the data mining role and information retrieving is carried out by this category of stake holders. Understanding of proper data flow procedures, code identification in International Code for Disease such as ICD9 or ICD10 and existing of any other information either for medical professionals or for insurance companies is quite significant with their involvement by Hirsch (2012) (<http://www.microsoft.com/Amalga/>).

Most of the existing IV applications highly focus on the role of easy understanding of pictures, more knowledge sharing and user friendly interfaces but lesser attention to the conclusion of these results from the basic EHR data sets that are involve in query process initiation by Wilcox *et al.* (2010a) and Krist *et al.* (2012) (<http://www.microsoft.com/Amalga/>). Still, unavailability of unified standard interfaces due to different format of databases in operation and

Corresponding Author: Muhammad Sheraz Arshad Malik, Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750, Tronoh, Perak, Malaysia

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

architecture results in transforming the true EHR details on screens in the form of visualizations by Hirsch (2012). Most hospitals in various states and countries still have no similar EHR database set that result in many interpretations due to multiple representations although the disease and symptoms of the patients are the same. This results in failure or non adaptation of EHR systems and their frequent use in health care units due to variance in data base entities, difference in data formats from one data base of one hospital to other and data retrieval from one application to another application by Hirsch (2012) (<http://www.microsoft.com/Amalga/>.) The situation gets worse when doctors do not have enough time to search, explore and mine the information from complicated interfaces that are intended to provide much data.

Even though there is a pattern in getting more applications in EHR developed with a focus on analysis, lesser attention is attributed towards understanding the analyzed information about patients by Malik and Sulaiman (2013b), Mane *et al.* (2012), Wilcox *et al.* (2010a) and Krist *et al.* (2012). In public health care units, due to maximum number of patients and lesser number of available doctors in emergency, outdoor patients ward and normal wards, it is very challenging to use the computer information to understand a single and comparison of multiple patients history based on database features by Wilcox *et al.* (2010a), Wang *et al.* (2011a) and Malik and Sulaiman (2012).

Each government hospital has its own database system to input, store, retrieve and process patient data depending upon their financial, geographical and management decisions by Wang *et al.* (2011b). It also employs different IV technologies to accumulate it from different locations such as from Intensive Care Unit (ICU), operation theatre, laboratories and other facilities to represent it in a single location. However, most of them fail when it needed to be shared with different other formats or location by Schaeffbauer and Siek (2011), Nair *et al.* (2011), Malik and Sulaiman (2013a), Wilcox *et al.* (2010b) and Wang *et al.* (2011a). One of the reasons is due to non understanding of database professionals needs and dataset demands. As majority of IV applications are designed primarily concerning the primary stake holders requirements irrespective to the architecture of database and understanding of professionals demands working on them by Malik and Sulaiman (2012), Zhang *et al.* (2012), Bui *et al.* (2007) and Wang *et al.* (2013). EHR database for most of the hospitals are directly under the control of updation, backup and full provision facilities by IT team with varied set of experience experts. This team may include IT manager, DBA team and Media team who are mostly web developers and IT technicians. This works as support team to other departments for instance Human Resources, Accounts

and Finance and Doctors officers as well as Ward Units where computers are provided on intra-network by Malik and Sulaiman (2013a), Wilcox *et al.* (2010b), Wang *et al.* (2011b), Malik and Sulaiman (2012) and Zhang *et al.* (2012). It may vary from single location to multiple location to populate data. This data format is further used for developing visualization by Ware (2013), Wilcox *et al.* (2010b) and Wang *et al.* (2011a).

Active Notes by Wilcox *et al.* (2010b), SOAP by Schaeffbauer and Siek (2011), Problem Oriented Medical Record (POMR) by Nair *et al.* (2011), Event Flow by Krist *et al.* (2012) and by Wang *et al.* (2011a) and many other existing and previously developed techniques are used to address the issues of EHR database influence on IV. Although these techniques are used in different visualization applications to address the factors that are less highlighted by designers still they did not encompass the full requirements of DBAs. Most of these applications fail to provide the desired result due to not understanding the needs and query structure development with reference to database architecture itself by Schaeffbauer and Siek (2011), Nair *et al.* (2011) and Malik and Sulaiman (2013a). This can be addressed if requirements of DBAs related to EHRs are properly understood and given a significant role to better demonstrate within existing systems. Much time is spent on understanding the requirements of doctors but lesser attention is paid to join the gap of knowledge sharing between doctors and DBAs by Hirsch (2012), Schaeffbauer and Siek (2011), Nair *et al.* (2011) and Malik and Sulaiman (2013a) (<http://www.microsoft.com/Amalga/>). This results in representing different formats of the same data as well as incomplete information retrieval; making the process of understanding EHR more cumbersome. More time consumption and difficult application interface are also the outcomes of complicated query results representation in such cases. The biggest challenge is unavailability of single or similar EHR database at different hospitals or health care units in even a whole one country, except a few states in USA and CANADA by Malik and Sulaiman (2013a) and Wang *et al.* (2011b). In such a hospital, for instance in British Columbia, normally there are different people designated with different operational roles in EHR visualization systems by Wang *et al.* (2011a), Malik and Sulaiman (2012), Zhang *et al.* (2012) and Bui *et al.* (2007). Usually notes taken by doctors are key punched by secretary of the health care unit or customer care unit while laboratory results are entered into the computer systems by general nurses or senior nurses at various locations. But database handling is either done remotely or onsite by government centralized units for insurance beneficiaries as well as managing health ministry operational procedures for future by Krist *et al.* (2012), Wilcox *et al.* (2010b) and Hirsch (2012).

The research work describes in this study involves identification of the core factors or areas of improvement for an EHR visualization system at database part with DBAs' perception in mind. This is to address the issue on previous applications whose focus are on data presentation centric concepts but not data formatting and understanding theme with reference to other contributing stake holders by A survey study on a group of DBAs was conducted for this purpose. The study findings are envisaged to be able to assist in future development of a visual analytic system in a more systematic manner by way of getting input from each participant user directly or indirectly involved to increase the efficiency, understandability and adaptability of such tools.

LITERATURE REVIEW USING EHR DATABASE

DBA normally administrate EHR database and its queries based on the health care professionals demands and requirements. Each hospital usually uses different database systems, design and architecture based on their budget availability, number and amount of resources and complexity of data by Bui *et al.* (2007), Wang *et al.* (2011b) and UK NHS (2013). Nearly all general purpose EHR queries have the same data format for results such as patient ID, personal details including name, age, location, sex, weight, symptoms, test reports, doctor's recommendations, medicines and visit details similar as patient paper file format by Malik and Sulaiman (2013b), Mane *et al.* (2012) and Wilcox *et al.* (2010a). Based on stake holder requirements, query structure changes the resulting information format in the form of simple to complex data representations.

Most of such scenarios are related to retrieving information from a single source of database or multiple data sources. This is actually a result of doctors and stake holder's particular interest to some specific event or areas of special attention by Schaeffbauer and Siek (2011), Nair *et al.* (2011) and Malik and Sulaiman (2013a). Just as in some situations, doctors are interested to find the past heart beat patterns of a diabetic patient with smoke habit and compare it with non smoking person suffering through same disease in a particular region. This results in complex visualization based on mapping of such events by Nair *et al.* (2011), Malik and Sulaiman (2013a) and Wang *et al.* (2011a).

Representation of information could be in any format but the results are based on the structure of query developed, sometimes merely on the basis of records mismatching information such as blood sample and pulse rate comparison of patients of diabetic and artheritis by Wang *et al.* (2013b). This results in visualization that is not at times completely fulfilling the demand of doctors. This is due to many factors as DBAs are from IT field and doctors are from medicine with different set of requirements, lesser attention towards the limitations of database, on a sparsed entities distribution in complex dataset by Bui *et al.* (2007), Wang *et al.* (2011a) and UK NHS (2013). To address these issues, more work is done towards the visualization side but lesser on standardizing DBA's perception in the past. Several approaches Active Notes by Nair *et al.* (2011), temporal querying by Krist *et al.* (2012), data mapping by Bui *et al.* (2007) and ontological based semantics by Wang *et al.* (2013) are

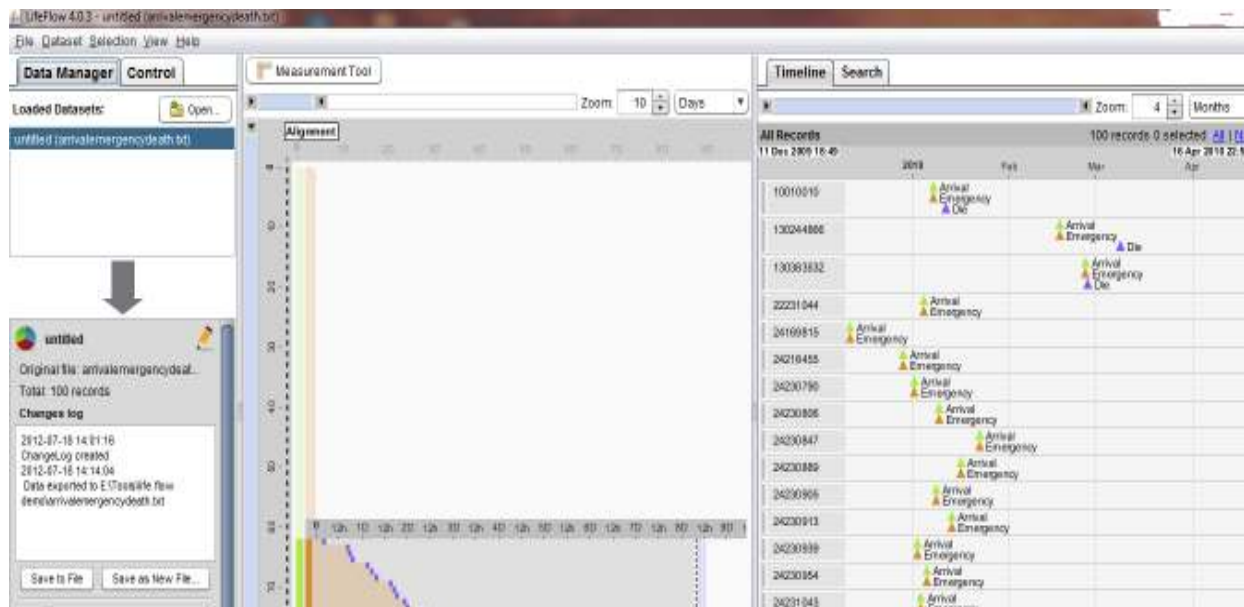


Fig. 1: Snapshot of life flow (Wong *et al.*, 2009)

solution oriented indirectly to address these issues. But these tools addressed more on the data sorting, filtering and query development rather than highlighting the shortcomings and deficiency areas for DBA and their interaction abilities with visualizations.

Temporal queries are developed by extracting different approaches either using time stamp sorting of data such as Life Flow (Fig. 1) by Wongsuphasawat and Shneiderman (2009), event based division i.e., Event Flow by Wang *et al.* (2011a), in some applications on the bases of case context approach by UK NHS (2013), HSCIC (2012) and Maldonado *et al.* (2012) and a few are addressed using grouping of same feature object classes such as mental disorders or other associated medical division terms by Maldonado *et al.* (2012). Event based or temporal queries bring the results that are formulated in visualizations such as 3D forms that are complex enough to understand or explore much from other stake holder's point of view. DBAs lack in the areas of knowledge about IV tools facilities, their complete use, limitations of EHR dataset with the visualization; thus, leads to an absence of improvement in future applications.

In another reported work, the approach suffers from mismatch and match option in categorical events with the same issues and also with less time difference by Wongsuphasawat and Shneiderman (2009). This approach tries to segregate similar events groups into similar categories based on mismatch and match events using Hungarian algorithm that involves applying distance function by McGuinness *et al.* (2012), Jiabin (2010), Huser *et al.* (2010) and Wongsuphasawat and Shneiderman (2009). The same principle has been used to focus on smaller databases but for the case of larger databases it misses a lot of records with different events with similar patterns by Huser *et al.* (2010). Query adjustment for complicated databases is really a challenge to visualize complete information representation.

Transformation of available datasets into single or multiple patients in brief visualization creates a gap of information for doctors when it comes into useful data presentation by Ware (2013) and Bui *et al.* (2007). One of the main reasons behind such complexities are non understanding of DBA's requirements and their lesser involvement within the final data selection and aggregation processes for resulting visualization. Although data is present, retrieval techniques and exploration of data from complex to simpler form is another challenge. For instance, to visualize medication results used by similar group of patients against intestinal infection in hospital may be easier till determining patient chart. But in the case where patients are diabetic or hyper tension associated then to measure affect with reference to side effects monitoring on other body parts involves more complex query

within given data set. This scenario could be solved if a unified and understandable format is available from DBAs point of view for query development as well as to understand their IV to database by Hirsch (2012), Schaeffbauer and Siek (2011), Nair *et al.* (2011), Malik and Sulaiman (2013a), Wilcox *et al.* (2010b) and Wang *et al.* (2011a).

Several applications have also tried to solve this challenge by segregating the expertise, level of navigation and information in addition to perception of stake holders about queries in different groups by Ware (2013), Mane *et al.* (2012), Wilcox *et al.* (2010b) and Krist *et al.* (2012). Aligning, filtering, regrouping, temporal based queries, formulation of single patient and later on multiple patient, information granularity, command on retrieval procedures and comparing the results are different research interest areas in similar data set or heterogeneous data set by Maldonado *et al.* (2012) and Viitanen *et al.* (2011). Data from multiple places come in different formats and to design and operate a data query is very difficult to bring all data fields in a single format on a single dialogue screen. Most of the problems reported were also incomplete information because of ignoring important patient fields within temporal events by Wang *et al.* (2013). Also it results in holding of system for few moments as no particular reason came in front of it yet by UK NHS (2013).

Even the most efficient and very well formulated Database Management Systems (DBMS) cannot bring results itself till they are properly operated and mechanized by DBAs based on end user requirements. IV majorly depends upon the results based on the queries and level of complexity in the queries addressed. But the best results can only be obtained related to the information required by Malik and Sulaiman (2013a) and Zhang *et al.* (2012). DBAs requirements as well as influencing factors pertaining to IV knowledge and its understanding should be identified and considered a valid importance for designing and utilizing such applications by Krist *et al.* (2012), Hirsch (2012) and Schaeffbauer and Siek (2011). As for doctors, information is the primary goal but for DBAs the goal is to understand the demand based on the system and then to present the results. If there is no alignment and proper flow of information between these stake holders, it will result in only complex or incomplete visualization to waste resources by Kwon *et al.* (2011), Zhang *et al.* (2013), Farri *et al.* (2012) and Wang *et al.* (2011a). To solve this problem, it is utmost important to group the requirements of DBAs and their expectations from the DBMS as well as future exploration procedures clarification by Zhang *et al.* (2013), Farri *et al.* (2012) and Wang *et al.* (2011b).

In another approach, a framework based on sequence of questions that depend upon the where, what and why and how 5 w's concept by Zhang *et al.* (2013) is introduced. The approach involves a combination of time and event based techniques on a particular sequential interfacing approach to a specific segment of database depending upon the number of events and interested events against a query based approach by Farri *et al.* (2012). The drawback within this approach is lesser in addressing multiple issues in multiple patients for chronic or complicated EHR data for multiple resources by Farri *et al.* (2012), Bowman *et al.* (2012) and Faisal *et al.* (2013). For extensive and complicated databases such as the NHS United Kingdom, data sections are distributed in different groups to better address the query demands by doctors and query holders by Zhang and Waliji (2011) and Farri *et al.* (2012).

Our research work describes in this study has used a similar approach as reported in previous tools and development process that include survey and close interview based study with DBAs of different government hospitals using different available databases by Mane *et al.* (2012), Nair *et al.* (2011), Viitanen *et al.* (2011) and Kwon *et al.* (2011). DBAs are considered as the secondary stake holder dealing with EHR visualization development yet their feedback is very important for IV tools efficient use by Bowman *et al.* (2012), Faisal *et al.* (2013) and McGuinness *et al.* (2012). Identification of their limitations and shortcomings will help in addressing the areas of pitfalls in the forthcoming IV applications for single or multiple patient EHR solutions.

Motivation for the study: The main purpose of this research work is to identify the gap areas within use and knowledge set for DBAs and their interaction with others i.e., Doctors and Visual designers. This will help to improve the application use as well as strengthen the flow of only required information upon demand process. DBAs from different domains and from different databases handling experience usually encounter with the situation of less understanding about deriving the real time information from EHR data set due to lesser information provisions and poor coordination. Returning results from different forms of queries such as temporal, event, phase oriented or entity based does not merely fulfill the demands of a desired visualization. Although from a database, results are deducted but they are not completely related or not fulfilling the requirements of doctors due to lesser IV knowledge of its stake holders by Malik and Sulaiman (2012), Zhang *et al.* (2012) and Bui *et al.* (2007).

Complex dataset and most intricated database solutions by different vendors can handle the data efficiently but techniques and way to handle the data by

and for users is the real key solution. Currently there are different EHR applications like Amalga by Microsoft (<http://www.microsoft.com/Amalga/>) and a few others at different hospitals are being used within different states of USA and CANADA that try to fulfill the query based requirements of its stake holders but still they are not widely adopted due to complex understanding procedures and differentiation in data set structure by Wilcox *et al.* (2010b), Wang *et al.* (2011a) and Malik and Sulaiman (2012). DBAs may possess excellent knowledge and expertise in EHR data retrieval queries but if the time requirements and proper goal for that query is not understood, it is not viable to extract a complete insight of results. Data set sometimes does not contain the required information and if it does sometimes even complete previous notes are not able to be retrieved based on multiple patients records due to various factors such as lesser understanding of requirement explanation, limitations of onsite data availability, increasing demand of doctors with complexity of past data analysis.

This work is a second phase of our ongoing project as CARE 1.0 model for designing of an EHR visualization solution for doctors as primary stake holders and with incorporation of DBAs and Visual designers as secondary stake holders indirectly involved in this process by Malik and Sulaiman (2012).

METHODOLOGY

The methodology used involves a questionnaire based survey on a selected strata of participants by Wilcox *et al.* (2010a), Schaeffbauer and Siek (2011) and Wang *et al.* (2011a). As DBAs are the main stakeholders of this study, sampling was done on this group of target audience. Based upon the availability of resources, data has been collected from two different government hospitals in Pakistan; one is a teaching hospital while the other is a non teaching hospital. Selection is made on the basis of measuring effects within sample taken from both teaching and non teaching hospitals. As in teaching hospitals normally number of patients and EHR data is more refined as compared to non teaching hospital. Also there are more number of man power resources available in teaching hospital. But combining from both kinds of hospitals will help to get better form of unbiased results. Twenty DBAs participated voluntarily for this study. As number of participants are always less with relation to this strata so we have taken maximum sample size in available population to avoid any proximity. They are on job DBA with average experience of 2 years or more in government hospitals. This is justifiable as the number of DBAs are always far lesser than the number of doctors working on HER in hospitals. In addition,

Table 1: Cumulative factors related to DBA

EHR applications/tools knowledge	Infra structure support
Single patient data retrieval	Information extraction
Data set creation	Simplification in operation
Temporal data knowledge and identification	Data set information
Development of visualization	Entities and field applicability
Ease in understanding	Errors identification
Expertise level in existing EHR tool	Complex solution development
Multiple patients data comparison	Results interpretation
Dataset updation	Information sharing
IV and dataset	IV development
Strength and weakness of DBMS	Database architecture limitations
Level of extensive Info	Data set resizing
Tools features knowledge	End user requirement handling
Data set management	Complete solution
Query development	User friendly GUI
	Multiple data flow

with reference to past work in a similar domain, the number of participants suggested are 6-10 varying upon the available resources and type of facilities by Krist *et al.* (2012), Hirsch (2012) and Viitanen *et al.* (2011).

Thirty one selected factors analyzed from past reported work in relation to DBAs were used as a basis to formulate the survey questions. Table 1 summarizes these factors. The factors were further divided into four main categories based on similar functions. This is to facilitate the DBA limitations based on the same problem set although discussed individually using previous work by Krist *et al.* (2012), Hirsch (2012), Nair *et al.* (2011) and Malik and Sulaiman (2013a). The categories are termed as DBA's perspective, EHR data knowledge set, Data set assessment and individual future expectation for data base areas related to EHR visualizations for its potential benefit to doctors and other stake holders. In order to facilitate the participants in understanding the questionnaire, the questions were phrased using the keywords found in each factor.

Each participant was introduced to the study questionnaire and was given a briefing about the questions for 15 min. They were invited to provide the feedback against the questions based on their EHR user experience and operational skills in IV. Each participant took a maximum of 20 min to complete the study. They could seek clarification from the researcher regarding the questions asked at anytime during the study.

To measure the expertise and level of skills against each factor, Likert scale has been used for choice of expertise from 1-5 whereby 1 denotes for less or no knowledge and 5 for expert rank. Feedback were taken from this sample of DBA population and then analyzed in a 2D chart form representing five ability levels as used in likert scale using color coding as expressed in Fig. 2 to 5 in the results section. The area of each color represents the strength of expertise level of a DBA

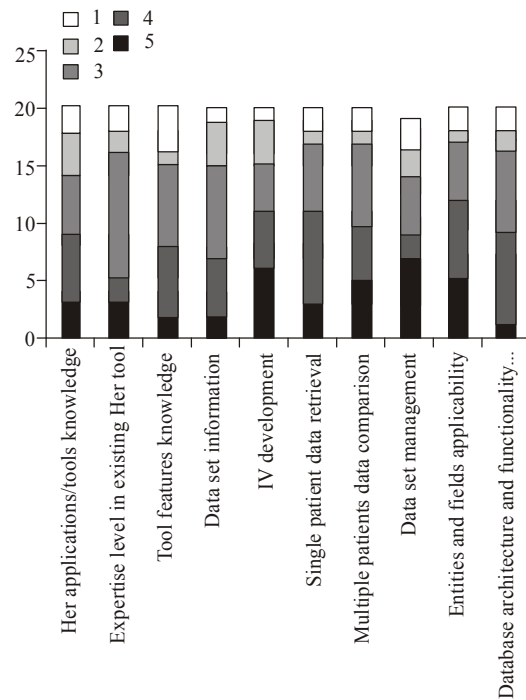


Fig. 2: Comparison of factors in DBA's belief for IV

strata while height represents the number of sample in that strata based on similar technique used by previous researchers by Wilcox *et al.* (2010a), Wang *et al.* (2011a), Malik and Sulaiman (2012) and Zhang *et al.* (2012).

Data is analyzed in the form of graphs as in Fig. 2 to 5, where X-axis is showing the enlisted factors in each category while Y-axis is representing the feedback of participants. Color coding is mentioned from 1-5 representing different categories of likert scale to facilitate feedback evaluation.

RESULTS

This section depicts the under pinning results about DBAs factors based on the participants feedback in four categories. Each category grouped on the basis of closely related factors with IV and DBMS relation. These categories are:

- DBA's belief
- Dataset Knowledge for Visualization
- Assessment of EHR dataset
- Individual and future perspective

DBA's belief: DBA's belief is the primary component within this study showing the basic expertise level of database administrators about understanding and operating existing EHR database, IV tools and medical dataset of patients. The study was carried out in

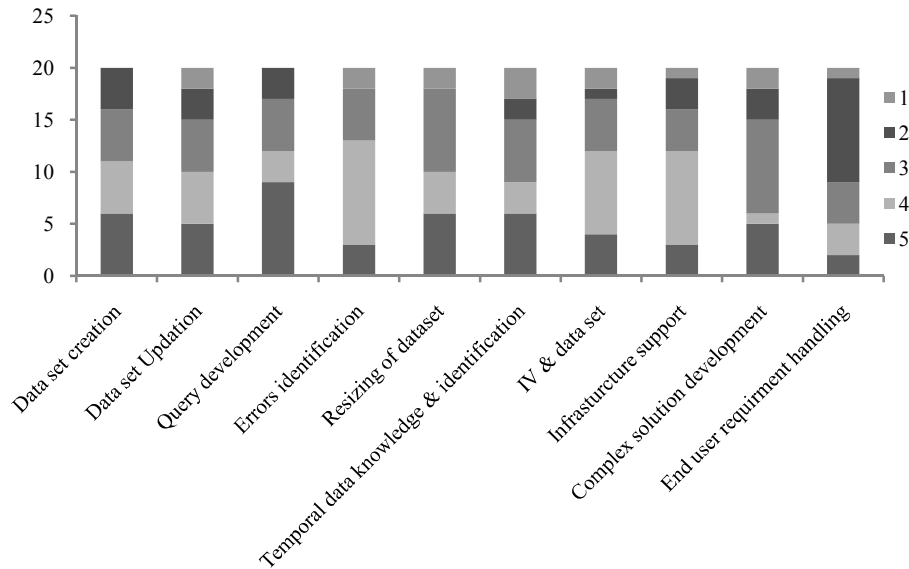


Fig. 3: Comparison of factors in DBA knowledge set for IV

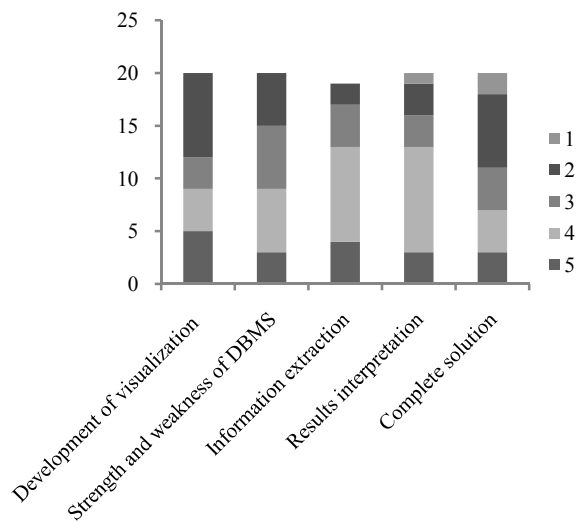


Fig. 4: Comparison of data set assessment factors for IV

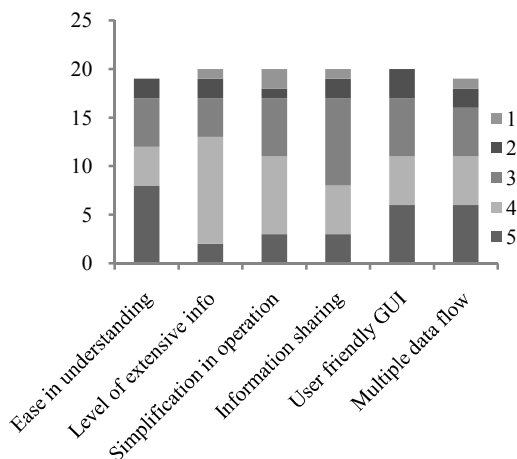


Fig. 5: Analysis of future perspective areas w.r.t DBAs

government hospitals where DBA performs the role of administration and development of hospital database with smaller or bigger IT team.

From Fig. 2, about 50% of DBA professionals have EHR application knowledge, Dataset Information, Multiple Patients data comparison and EHR tools expertise level at scale 3 and 4. About 25% of DBAs gave feedback at scale 1-2 for IV development information and its entities field applicability. But there are also some facts lacking their attention and requiring more expertise from them mentioned later in the same section.

The results for factors within this category represents that even DBAs have prior knowledge but still majority of them feels difficulty in understanding the complete information about features in tools and database, lesser number are marked to higher level of expertise in existing EHR tools. Most participants fall at level 3 and level 4 for multiple patients data retrieval, single patient history retrieval and database architecture information. But still the graph shows lack of dataset management techniques that are highly demanded for query development and complex IV formulation. These techniques are quite highly desired from DBAs particularly when an unorganized dataset from various sources is updating on daily basis.

Data set knowledge for visualization: Data set provides a base line for developing a visualization that results in an easier understanding of single of multiple patients history. As multiple patients history always contain difficult data entities constructing complicated dataset, so complete information to handle and understand such data set is the key for a useful visualization. Ten different factors within this category were selected based on previous work in IV tools and studies. They are also quite much related with their

titles for use of different stake holders i.e., doctors and visual designers but considered important at similar grounds for DBA as well by Wilcox *et al.* (2010b), Nair *et al.* (2011), Maldonado *et al.* (2012), Viitanen *et al.* (2011) and Wang *et al.* (2011b). Figure 3 represents the analytical comparison of these factors with reference to DBA expertise level. These factors are nominated based on the past work criteria by different researchers Data set creation, Data set updation, Query development, Errors identification, Resizing of dataset, Temporal dataset knowledge, IV and dataset, Infrastructure support, Complex solution development and End user requirement handling.

Dataset creation and End user requirements handling are significant areas where DBA needs more expertise with reference to Fig. 3 as most of them are falling at level 2. The results show that 60% database professionals working in different hospitals are very less in number who feel confident in handling dataset updation, query development process, IV and dataset relationship and Infra structure link with database support. The graph is also showing that 40% of database professionals are falling at level 2 to 3 for complex solution development from complex dataset that is one of the leading requirement for IV in various scenarios. This is mentioning the deficiency provoking areas that still need to be addressed. Most of EHR database related professionals lack to provide complex solution development due to two major factors:

- Dataset patterns that are not always friendly with each database query versus infra structure support
- Missing a complete solution support availability based of diversified information requirement of doctor's demand in complex EHR visualization

Assessment of EHR data set: Development of Visualization, Strength and weakness of database and data set, Information extraction on demand, Results interpretation easiness and complete solution to end user requirements are key factors within the category of assessment of EHR dataset (s). These factors attempt to determine the capabilities of a database professional about assessing the data set and database system that is going to be operated. It also helps to understand the short comings and future enhancement milestones for oncoming challenges. Very few number of DBAs lack the ability of driving a complete solution for other stake holders such as doctors and visual designers based on their requirements as mentioned in Fig. 4. But still there are about 25% of population in database professionals who still lack the ability of results interpretation and solution derivation. Sixty five percent professionals are comfortable at average level of scale 3 with information extraction features and techniques based on requirements in addition to results interpretation.

The number of participants who are expert in identifying the strength and weakness of Database

management systems, driving complete solution is also very less. So these are more areas of interest for junior and fresh database professionals in EHR visualization domain. Mostly other stake holders in IV are very much interested in finding insights of information and additional details by Hirsch (2012), Malik and Sulaiman (2013), Maldonado *et al.* (2012) and Bowman *et al.* (2012). These features are actually the supply line for providing a useful visualization within EHR and help to under-fold different facts. Most of the researchers are quite much interested to follow the past information and alignment of information in a way so that it can help to improvise the functionality and adaptability of IV systems by Faisal *et al.* (2013) and McGuinness *et al.* (2012). This is also highlighting that the underlying reason behind the colored pattern is development of a smart query system to better solve the user problem with the help of database professional that could only be achieved by improving expertise in these factors.

Individual and future perspective: The fourth category particularly constitutes the factors that apprehend the DBA's personal perspective as well future expectations from a dataset, DBMS and data base operations for interaction with other stake holders. Ease in understanding of information retrieved, Level of depth of information, Operational simplification, Information sharing with other users, contribution towards friendly GUI and multiple data flow from different sources or sets at single place are different factors present within this area. These factors are being narrow down here based on already existing applications proposed future work and areas of research, short comings and other related areas that are not being attended or put in the future work section by Hirsch (2012), Mane *et al.* (2012) and Krist *et al.* (2012). Database professionals work for information digging, retrieval, management and updation processes. The information retrieved will be the base source for developing a useful and easily understandable IV. So their input has a vital role for helping in boosting efficiency and improvement of IV applications as future work. DBA's feedback is as mentioned in Fig. 5 about such factors.

Majority of database professionals are highly interested in increasing the information sharing and expect the systems should help in under-folding more level of information with other stake holders. The rest are highly interested in level of information by going in more depth of information about patient history. So this feature in database systems will not only help database professionals but also facilitate others using IV. Up to 45% of DBAs are also in favor of simplification of processes in data retrieval and data mining steps as it facilitate in better understanding of the information. They also tried to highlight the importance of user friendly GUI. The interface of such applications and

tools are more tailored towards the requirements of end users based on their previous interaction with similar tools and visualization knowledge.

DISCUSSION

EHR database and dataset play a vital role in developing a useful Information visualization for medical professionals and associated stake holders. Single to multiple patients record retrieval, patient history on demand and different formats of results reports and development of queries for such functions are routine job of database professionals. Most of database professionals are interested in operating and deriving results based on IT query set but sometimes these queries are not as same as desired by other stake holders of visual analytic systems of EHR by Huser *et al.* (2010). Although it looks a simple task to say, but complicated in multiple patients scenarios suffering from same disease accompanying different physical background, utilizing different medicines and have complex medical history.

Different formats of EHR dataset and different operational procedures and set up of EHR database management systems result in retrieving different type and quality of results. This reflects a tremendous change in resulting visualizations that are end product for any visual analytic systems for efficient analysis and problem understanding. Different visualizations with similar results will result in different interpretations and in complete information not only to end users but also to database professionals. This increases more time consumption for results finding, facts gathering and poor perception of inferences from such systems resulting in failure or their lesser adaptability in public hospitals. In government hospitals, lesser number of doctors are available per patient time in emergency as well as in normal wards and operation theatres by Wilcox *et al.* (2010a). More number of patients, lesser available doctors and limited IT facilities affects the use and adaption of EHR visualization for complicated tools. Specified training for such tools and also comparison of facts in multi facet formats often create confusion in recording doctors notes.

The primary objective of this study is to enlist the number of challenges faced by working DBA on existing EHRs in public hospitals and categorize them accordingly. These include their skills, knowledge and understanding to develop a simple query is very much directly linked for an effective and efficient user friendly visualization. Although finding information about a single patient or multiple patient is just to compare the patient daily chart or hospital file from cabinet or record branch. But when comparing the results in hundreds on single screen or sheet then it becomes difficult and complex. Different vendors such

as Microsoft, ORACLE, IBM provide different DBMS solutions for information retrieval, recording, query development and reports generation by Mane *et al.* (2012) and Krist *et al.* (2012). The records are kept in different hospitals in different formats like MS word, Excel, Fox Pro or Fortran based programs or SQL based modern data storage tools. The real job is to bring the records to front end in a useful format using a transparent and understandable fashion to visualize that are directly related to DBAs. Conventional databases are good to record and retrieve data as per end user requirements but real problem is complex medical data fields such as representation of multiple patients past six months injection frequency measurement for particular pulmonary infection by McGuinness *et al.* (2012). For example to represent the numerical data of diabetic patients like pulse rate, blood sugar with or without fasting is easier. But to represent the smoking habit affect on the heart arteries based on ECG heart condition of multiple patients is difficult to represent in time gaps versus number of heartbeats. Although this problem can be noted as general field record like numbers or text in the form of record field but sometimes condition is pertaining to such formats, it could be picture and visualize such differentiation and representation is not possible in simple visualization. This is easier to require for a query but its directly related to limitation of the dataset and database capacity to handle such multi variated information.

Patient dataset format is although quite much similar with reference to name, address, basic disease and doctor's notes structure in various databases widely used in the public hospitals around the world. But with latest diseases, complicated lab reports, extensive and exhaustive and temporal events in patient health results in driving more information granularity that was not possible couple of years before. This leads to theoretically a better and efficient health care system but based on scarce resources, minimum technology change patterns and latest training trends unavailability at public hospitals also result in poor knowledge provision results, in complete query oriented schemas and complex visualizations for doctors and stake holders. This results in overall lower efficiency of DBMS system supporting to visualize results as well as also lesser orientation towards the systems updation. As different DBMS varies from hospital to hospital that also effect on bringing a uniformity among the visualization that can only be achieved by updating the skills of DBAs as mentioned with the above results.

Temporal patient data representation is a main focus in a visualization for presenting data by DBAs as different events and health status reports of single or multiple patients becomes difficult for various database professionals. The main reasons behind this fact are incomplete information understanding about

requirements of other stake holders due to poor information sharing, complicated data set with different DBMS formats as well as malfunctioning of one scenario query set in other scenarios. This piece of work also depicts the weakness areas of DBAs particularly in organizing, managing and replicating the results in an understandable format. Normally formats are available or developed based on end user requirements and also data is received on the same requirements. But representation format are not always same for same data. In different tools previously used, patient name, ID or doctor's name are placed while in comparison of multiple patients is obviously clear that at many instances this information is not that much important if area of interest for end user is other than these fields. Like doctors if are interested in comparing multiple patients entry and exit dates in emergency related to road side accident or against an outbreak dengue virus then patient IDs and patient details are not that much important more than just bringing the information. But representation of such information will make the visualization more populated and lesser addressing to the fact finding details.

A general comparison of all four categories of factors bring forward facts related to data retrieval, dataset basic knowledge information, management of database and comparison of future trends within different patients records. Most of DBAs are expert in developing and running basic queries to bring out the single patient data and multiple patient general records but when it comes to retrieve complex data in easy form it becomes difficult. The reasons behind these facts are less training on multiple database management systems for health care units in public hospital and lesser exposure to the real understanding about objectives of other stake holders. Different vendors and IT solution providers in different countries are providing different solutions about EHR management and its organization. But the format and structure of database nearly in each medical and teaching hospital in developing countries is varying from to another as due to non standardization in data keeping practice, availability of limited resources and other social influences. This result in sparse and disintegrated form of EHR data, that is replicated many times and complex in segregation for query development at multiple sources.

Results also reveal on the less interaction between DBA and other stake holders. Normally doctors expect more from IV applications without considering the limitation of the available DBMS. While DBAs develop queries that bring maximum amount of results and information with a poor and not well defined filter. This results in bringing the information but that information is not to the point resulting in extra information with non interesting facts or unrelated insights of patient history. Thus results are representing a continuous upgrade in technical knowledge and information

retrieval techniques update in DBAs as well as a standardized information flow cycle by aligning the requirements of all EHR stake holders at single junction to avoid confusion and complications in data retrieval. As data retrieval and formulation into a visualization is not as much time taking process as to analytically understand that visualization and upon not up to the mark its rejection rates are high. Thus causing frustration and time wasting for different stake holders.

Requirements and future needs of medical domain professionals will change in a particular span of time that is not identified or clarified yet. There could also be change in requirements based on some events like some disease outbreak or some natural disaster activity that will result the data retrieval of multiple patients at single or multiple places. So both the DBMS and DBA should be well acquainted with these facts to have a flexibility mode and ease in data set to facilitate such changes that are both timely intimated or un-intimated. Although DBMS for EHR from different solution providers try to encompass the future needs but still the requirements of medical professionals at public hospitals as due to vitality as well as for more research mode, thus also change the requirements for DBAs.

A brief analysis of all results represent a four key areas improvement note for DBAs in which the first one is making more knowledge improvement with temporal dynamic query set structure i.e., live multiple patients data from multiple sources at different or similar time zones have similarity such as diabetic B patients heart beat rate at morning time before eating any food in different age groups etc. The second area is understanding the limitation of EHR DBMS and visualization format and with different data set of previous years and current years in an agreeable format. The third is availability of level of information depth based on doctor's demand in public hospitals both for research and day to day operational needs. The fourth is understanding and standardization in dataset information not only for junior level DBAs but also for other IT and non IT professionals associated with EHE visual analytic systems by trainings, workshops as well as update in bringing the implementation of novel ideas in front.

This piece of work is an integral part of our ongoing research work on EHR visualization model based on integration of studies carried out on needs, gaps and requirements all three stake holders associated with it i.e., doctors, DBAs and visual designers. This is a step forward to integrate our under process work as CARE 1.0 a future EHR model as an analytic base for patient history records. The importance of such factors highly recognizable by other stake holders to properly integrate the flow of requirements adjunctly. This will result in formulation and aligning the solution part based on input from three stake holders and having

more grip on the visualization of desired results in a well organized pattern.

RECOMMENDATIONS EXPLORATION

IV in EHR is direct proportional in context of its means to the results of patient database management system and personals working on it. More granular results would be more simplified and understandable visualization format GUI in the EHR tool will be obtained. Although within this study, the best efforts were focused on the needs, gaps and working requirements influencing factors of DBAs but still there is more extensive and in depth work is still pending to be done specifically on the EHR database standardization and temporal query format uniformity sides. A lot of applications and tools will be emulating the results in visible and colored format from different EHR database systems but still there is no uniformity in obtaining even at the structure of basic patient record and history setup. Each vendor is producing different format supporting but at the user end, due to unavailability and lesser attention to the features identification very less implementation is observed. So more knowledge exploration about EHR visualizations should be disseminated not only in developed countries hospitals but also in under developed countries.

Currently most of public hospitals having DBAs with minimized qualification and lesser expertise at future trends and innovation lines in future EHR visualization tools that is an alarming situation. Information filtration and development of efficient temporal queries containing numbers, events, status or any other divisional category based are the areas of future exploration. Although previous related work covers particularly needs of physicians in public hospitals as in our ongoing research project of CARE 1.0 and this work encompassed the interaction of same needs in DBAs. But as authors, we feel more input is required from different geographical regions and from different hospitals to improve the IV applications in patients records both from DBAs as well as other stake holders. Aligning the needs of all stake holders with proper information filter and perceptual design patterns are future research goals in this area of domain. Data flow with query development is also not quite much defined from user and administration prespective as well.

ACKNOWLEDGMENT

The authors would like to thank Universiti Teknologi PETRONAS for providing technical and financial supports for this research work. Special thanks to The MediaSol Pakistan for providing feedback and active participation in conducting this study with the help of Database administrators working in EHR

domain at Government Civil Hospital and Allied Teaching Hospital at Faisalabad Pakistan.

REFERENCES

- Bowman, B., N. Elmqvist and T.J. Kelly, 2012. Toward visualization for games: Theory, design space and patterns. *IEEE T. Vis. Comput. Gr.*, 99: 1-1.
- Bui, A.A.T., D.R. Aberle and H. Kangaroo, 2007. TimeLine: Visualizing integrated patient records. *IEEE T. Inf. Technol. B.*, 11: 462-473.
- Faisal, S., A. Blandford and H.W.W. Potts, 2013. Making sense of personal health information: Challenges for information visualization. *Health Inform. J.*, 19(198): 21.
- Farri, O., D.S. Pieckiewicz, A.S. Rahman, T.J. Adam, S.V. Pakhom and G.B. Melton, 2012. A qualitative analysis of EHR clinical document synthesis by clinicians. *Proceeding of the AMIA Annual Symposium*, pp: 1211-1220.
- Hirsch, M.D., 2012. Lack of Standardized HER Interface Delaying Interoperability. Retrieved from: <http://www.fierceemr.com/story/lack-standardized-ehr-interface-delaying-interoperability/2012-02-14>.
- HSCIC (Health and Social Care Information Center), 2012. Quality of Dataset Report in Various Groups. Retrieved from: <http://www.hscic.gov.uk/catalogue/PUB08687>.
- Huser, V., S.P. Narus and R.A. Rocha, 2010. Evaluation of a flowchart-based EHR query system: A case study of RetroGuide. *J. Biomed. Inform.*, 43: 41-50.
- Jiaxin, W., 2010. WIVF: Web information visualization framework based on information architecture 2.0. *Proceeding of the 2nd International Conference in Computer and Automation Engineering (ICCAE)*, pp: 734-738.
- Krist, W., M. Monroe, C. Plaisant, B. Shneiderman, J. Millstein and S. Gold, 2012. Exploring point and interval event patterns: Display methods and interactive visual query. Retrieved from: <http://www.cs.umd.edu/hcil/eventflow/>.
- Kwon, B.C., B. Fisher and J.S. Yi, 2011. Visual analytic roadblocks for novice investigators. *Proceeding of the IEEE Conference in Visual Analytics Science and Technology (VAST)*. RI, USA, pp: 3-11.
- Maldonado, J.A., C.M. Costa, D. Moner, M.M. Tortosa, D. Bosca, J.A.M. Gimenez, J.T.F. Breis and M. Robles, 2012. Using the research EHR platform to facilitate the practical application of the EHR standards. *J. Biomed. Inform.*, 45: 746-762.
- Malik, M.S.A. and S. Sulaiman, 2012. An integrated modular approach for visual analytic systems in electronic health records. *Int. J. Adv. Comput. Sci. Appl.*, 3: 246-250.

- Malik, M.S.A. and S. Sulaiman, 2013a. Towards the development of an interface model for information visualization in multiple electronic health records. *Proceeding of the IEEE International Conference on Computer Medical Applications (ICCMA)*, IEEE, pp: 1-5.
- Malik, M. and S. Sulaiman, 2013b. Analytical Comparison of Factors Affecting EHR Visualization for Physicians in Public Health Care Units. In: Zaman, H., P. Robinson, P. Olivier, T. Shih and S. Velastin (Eds.), *Advances in Visual Informatics*. Springer International Publishing, Switzerland, 8237: 605-614.
- Mane, K.K., C. Bizon, C. Schmitt, P. Owen, B. Burchett, R. Pietrobon and K. Gersing, 2012. Visual decision line: A visual analytics approach for comparative effectiveness-based clinical decision support in psychiatry. *J. Biomed. Inform.*, 45: 101-106.
- McGuinness, D.L., T. Lebo, A.R. Shaikh, R.P. Moser, L. Ding, J.P. McCusker, G.D. Morgan, Z. Tatalovich, G. Willis, B.W. Hesse, N. Contractor and P. Courtney. 2012. Towards semantically enabled next generation community health information portals: The PopSciGrid pilot. *Proceeding of the 45th Hawaii International Conference in System Science (HICSS)*, pp: 2752-2760.
- Nair, V., M. Kaduskar, P. Bhaskaran, S. Bhaumik and L. Hodong, 2011. Preserving narratives in electronic health records. *Proceeding of the IEEE International Conference in Bioinformatics and Biomedicine (BIBM)*, pp: 418-421.
- Schaeffbauer, C.L. and K.A. Siek, 2011. Cautious, but optimistic: An ethnographic study on location and content of primary care providers using electronic medical records. *Proceeding of the 5th International Conference in Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pp: 63-70.
- UK NHS (UK National Health Service), 2013. NHS Datamodel and Dictionary Service. Retrieved from: <http://www.connectingforhealth.nhs.uk/systemsandservices/data/nhsdmds>.
- Viitanen, J., H. Hypponen, T. Laaveri, J. Vanska, J. Reponen and I. Winblad, 2011. National questionnaire study on clinical ICT systems proofs: Physicians suffer from poor usability. *Int. J. Med. Inform.*, 80: 708-725.
- Wang, H.Q., J.S. Li, Y.F. Zhang, M. Suzuki and K. Araki, 2013. Creating personalised clinical pathways by semantic interoperability with electronic health records. *J. Artif. Intell. Med.*, 58(2): 81-89.
- Wang, T.D., K. Wongsuphasawat, C. Plaisant and B. Shneiderman, 2011a. Extracting insights from electronic health records: Case studies, a visual analytics process model and design recommendations. *J. Med. Syst.*, 35(5): 1135-1152.
- Wang, X., D. Wenwen, T. Butkiewicz, E.A. Bier and W. Ribarsky, 2011b. A two-stage framework for designing visual analytics system in organizational environments. *Proceeding of the IEEE Conference on Visual Analytics and Technology (VAST)*. Providence, RI, pp: 251-260.
- Ware, C., 2013. *Information Visualization Perception for Design*. 3rd Edn., Morgan Kaufmann Publishers, USA.
- Wilcox, L., D. Morris, D. Tan and J. Gatewood, 2010b. Designing patient-centric information displays for hospitals. *Proceedings of the 28th International Conference on Human Factors in Computing Systems*. Atlanta, Georgia, USA.
- Wilcox, L., J. Lu, J. Lai, S. Feiner and D. Jordan, 2010a. Physician-driven management of patient progress notes in an intensive care unit. *Proceedings of the 28th International Conference on Human Factors in Computing Systems*. Atlanta, Georgia, USA.
- Wongsuphasawat, K. and B. Shneiderman, 2009. Finding comparable temporal categorical records: A similarity measure with an interactive visualization. *Proceeding of the IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pp: 27-34.
- Zhang, J. and M.F. Waliji, 2011. TURF: Toward a unified framework of EHR usability. *J. Biomed. Inform.*, 44(6): 1056-1067.
- Zhang, R., S. Pakhomov and G.B. Melton, 2012. Automated identification of relevant new information in clinical narrative. *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*. Miami, Florida, USA.
- Zhang, Z., B. Wang, F. Ahmed, I.V. Ramakrishnan, R. Zhao, A. Viccellio and K. Mueller, 2013. The five Ws for information visualization with application to healthcare informatics. *IEEE T. Vis. Comput. Gr.*, 19(11): 1895-1910.