Journal of Biomedical Informatics 45 (2012) 101–106

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

VisualDecisionLinc: A visual analytics approach for comparative effectiveness-based clinical decision support in psychiatry

Ketan K. Mane [a,*], Chris Bizon [a], Charles Schmitt [a], Phillips Owen [a], Bruce Burchett [b], Ricardo Pietrobon [b], Kenneth Gersing [b]

[a] Renaissance Computing Institute (RENCI), University of North Carolina, Chapel Hill, NC, USA
[b] Duke Medical Center, Duke University, Durham, NC, USA

ARTICLE INFO

Article history:
Received 7 February 2011
Accepted 11 September 2011
Available online 20 September 2011

Keywords:
Visual analytics
Decision support
Comparative Effectiveness Research
Electronic health records
Visualization

ABSTRACT

Comparative Effectiveness Research (CER) is designed to provide research evidence on the effectiveness and risks of different therapeutic options on the basis of data compiled from subpopulations of patients with similar medical conditions. Electronic Health Record (EHR) system contain large volumes of patient data that could be used for CER, but the data contained in EHR system are typically accessible only in formats that are not conducive to rapid synthesis and interpretation of therapeutic outcomes. In the time-pressured clinical setting, clinicians faced with large amounts of patient data in formats that are not readily interpretable often feel 'information overload'. Decision support tools that enable rapid access at the point of care to aggregate data on the most effective therapeutic outcomes derived from CER would greatly aid the clinical decision-making process and individualize patient care.

In this manuscript, we highlight the role that visual analytics can play in CER-based clinical decision support. We developed a 'VisualDecisionLinc' (VDL) tool prototype that uses visual analytics to provide summarized CER-derived data views to facilitate rapid interpretation of large amounts of data. We highlight the flexibility that visual analytics offers to gain an overview of therapeutic options and outcomes and if needed, to instantly customize the evidence to the needs of the patient or clinician. The VDL tool uses visual analytics to help the clinician evaluate and understand the effectiveness and risk of different therapeutic options for different subpopulations of patients.

1. Introduction

Evidence-based medicine, or evidence-based practice, refers to the incorporation of sound research evidence, clinical experience, and patient values into the decision-making process of the clinician, in terms of therapeutic choices and overall patient care [1]. Comparative Effectiveness Research (CER) is designed to improve the clinical decision-making process by providing research evidence on the effectiveness and risk–benefit profile of different therapeutic options for specific patient subpopulations [2]. CER often involves the use of existing patient data to identify subpopulations of patients who share similar clinical characteristics and to determine which therapeutic outcomes produce improvements in medical status with minimal side-effects [3]. CER aims to improve and simplify the clinical decision-making process by reducing the number of available treatment options to those that have proven to be most effective with little risk. Clinicians and health policy experts agree that decision support tools that aid clinical decision-making hold enormous potential to improve overall patient care [4,5].

The application of CER-based, Clinical Decision Support (CDS) tools at the point of care should lead to improvements in health care quality and reductions in medical costs. Electronic Health Record (EHR) system provides access to large volumes of patient data that can be used for CER, but the data are typically accessible only in formats that are not conducive to rapid synthesis and interpretation of therapeutic outcomes. Clinicians often face information overload when presented with large amounts of patient data in the time-pressured clinical setting, which limits the use of CER-based decision support at the point of care. In addition, the information-processing abilities of humans are limited when dealing with large amounts of data in real time [6]. Further, in the clinical setting, the clinician has a limited amount of time to spend on each patient and cannot spend that time filtering through large amounts of data to evaluate the effectiveness of different treatment options.

Computers can augment the information-processing abilities of humans. Computational approaches can be designed to rapidly process large datasets, identify the underlying characteristics of the data, and present the results in a user-friendly format. External aids of this type can enhance human decision-making. For

* Corresponding author. Address: Renaissance Computing Institute (RENCI), 100 Europa Drive, Suite 540, Chapel Hill, NC 27517, USA. Fax: +1 919 445 9669.
E-mail address: kmane@renci.org (K.K. Mane).

1532-0464/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2011.09.003

clinicians faced with large amounts of patient data at the point of care, external aids can improve the decision-making process if they can rapidly summarize large datasets to provide an overview of the data, instantly update the data to determine the effects of different parameters or to incorporate new data, and aptly present the results in a format that is easy to interpret [7]. Visual Analytics (VA) offers an integrated computational approach that combines visualization, human factors, and data analysis. VA can be applied as an external aid to filter, analyze, and visualize data [8,9].

The aim of this paper is to discuss the use of VA for CER-based CDS using patient data from an EHR system. We developed a prototype external tool, termed 'VisualDecisionLinc' (VDL), to identify subpopulations of patients who share a similar medical profile and to succinctly present results from large datasets using visual cues to guide the clinician in their decision-making process. We show how a VA approach can be used for CER at the point of care and how it can be customized to the needs of the patient or the clinician. We demonstrate how our VA-based approach permits the use of EHR data as a supplemental CDS tool to improve decision-making at the point of care.

This manuscript is organized as follows: Section 2 provides an overview of related work using visual approaches for clinical decision support; Section 3 provides a brief overview of EHR system (MindLinc) and the psychiatric dataset that was used to develop and test VDL; Section 4 provides an overview of the VDL user interface (UI), and the application of VDL for CDS, Section 5 provides details about the limitations of the current work, and Section 6 concludes with a discussion of the current results and future work.

2. Related work using visual approaches

Current CDS systems use EHR data, but often constrain displays of the data to tabular views or text formats. Large datasets displayed in tabular or text format fail to rapidly communicate underlying characteristics of the data and in turn, limit data interpretation. Other visualization approaches display critical temporal events for a particular patient [10,11] or temporal trends across patients [12]. The temporal visualization efforts have been used primarily for data exploration and retrieval; these tools have not been refined for use in CDS at the point of care. Thus, a need exists for new visual approaches to process large volumes of patient data for CER and to present the results in a form that can be used for CDS at the point of care.

The ideal point of care CDS tool should be flexible and able to accept input from the clinician, including clinician-approved changes to the data, and it should provide clinicians with quick feedback on any changes in the displayed information [13–15]. Furthermore, the ideal point of care CDS tool should be able to quickly retrieve and synthesize large amounts of patient data and outcomes for use at the point of care.

We developed a prototype VA-based tool – VDL, that we believe overcomes the problems associated with previous visualization approaches and achieves the characteristics of an ideal point of care CDS tool, as described above. Before discussing the methodological details involved in the development of VDL, we highlight in the next section, the EHR system and psychiatric dataset that we used to develop and test VDL.

3. MindLinc EHR system and dataset details

The MindLinc EHR system is the largest de-identified psychiatry outcome data warehouse in the United States, and it is a clinically representative sample of data collected in psychiatric practice [16]. MindLinc data warehouse represents 110,000 patients or 2,400,000 clinical encounters collected over a 10 year span. The MindLinc data are drawn from various types of mental health facilities—academic medical centers (25%), community mental health centers (50%) and other practices (25%), from geographically different areas of the country (North, East, South and West). Patient data from each psychiatric practice site is periodically pooled (~every 6 months) into a de-identified, HIPAA compliant data warehouse. The MindLinc EHR system stores data on patient demographics, current and past medications, side-effects, comorbidities, and other related clinical data, including psychiatric diagnoses and therapeutic outcomes. The data used for the development of VDL was extracted from the MindLinc EHR system.

We limited our focus to psychiatric patients with Major Depressive Disorder (MDD) as their primary diagnosis. Within MindLinc, patient records are tagged with codes from the Diagnostic and Statistical Manual of Mental Disorder (commonly referred to as DSM codes). We use this DSM codes to identify patients with a primary diagnosis of MDD, and extract them for use in the development of VDL. From the available MindLinc data, the resulting MDD dataset had 33,536 patient encounters from 3016 unique patients. The final, compiled MDD dataset included data on prescribed medications, comorbid conditions, demographics (race, gender, and age), visit type (inpatient, outpatient, or emergency), and treatment outcomes.

Within MindLinc, the patient treatment outcome is recorded in the form of a Clinical Global Impression (CGI) score. The CGI score was developed and validated for use by the National Institute of Mental Health [17] and is used in virtually all FDA-regulated and most CNS trials. The CGI provides a brief, stand-alone score of the clinician's assessment of the patient's global functioning prior to and after initiating a treatment. The CGI score range is from 1 through 7, with 1 indicating 'no illness/maximum improvement' and 7 indicating 'severe illness/maximum worsening'. Typically, good medications outcome are indicated with CGI score less than two ($CGI \leqslant 2$). The CGI score is applied to two dimensions: CGI-improvement (CGI-I) and CGI-severity (CGI-S). While stability is the goal, our goal is more to aid the doctor with transitions, e.g., if the doctor believes a patient's medication should be switched then what should it be switched to. In that case, CGI-I offers insight into patient's response to the new treatment. Therefore, we focused on the CGI-I scale (referred to as 'CGI' henceforth) for VDL.

4. VisualDecisionLinc: Overview and details

4.1. MDD comparative population dataset derived from EHR data

A patient data-driven approach was used to define the MDD comparative population. With this approach, the medical profile of the patient was used as 'seed' data to identify comparable patients with similar medical profiles. A similar medical profile was defined initially as 'patients with the same primary diagnosis'. The similar medical profile can be defined using multiple data fields, including patient demographics, comorbidities, medication switching, and others. These definitions of patient similarity can be applied dynamically at the UI level (Section 4.2). In the current study, because MDD patients often switch medications, patients with a primary diagnosis of MDD and their last prescribed medication were treated as 'seed' data identifiers to identify the MDD comparative patient population. This initial dataset is real world de-identified data; hence we have to deal with aspects which include patients with wide gaps between visits, as well as patients with wide gaps in prescription refills/renewals. To filter out the data variations (time gaps, and inconsistent refill visits), we only include patient visits where: (a) the same medication was prescribed for at least 120 days; (b) same medication for at least 120 days, an upper bound on the time; and (c) to exclude patients

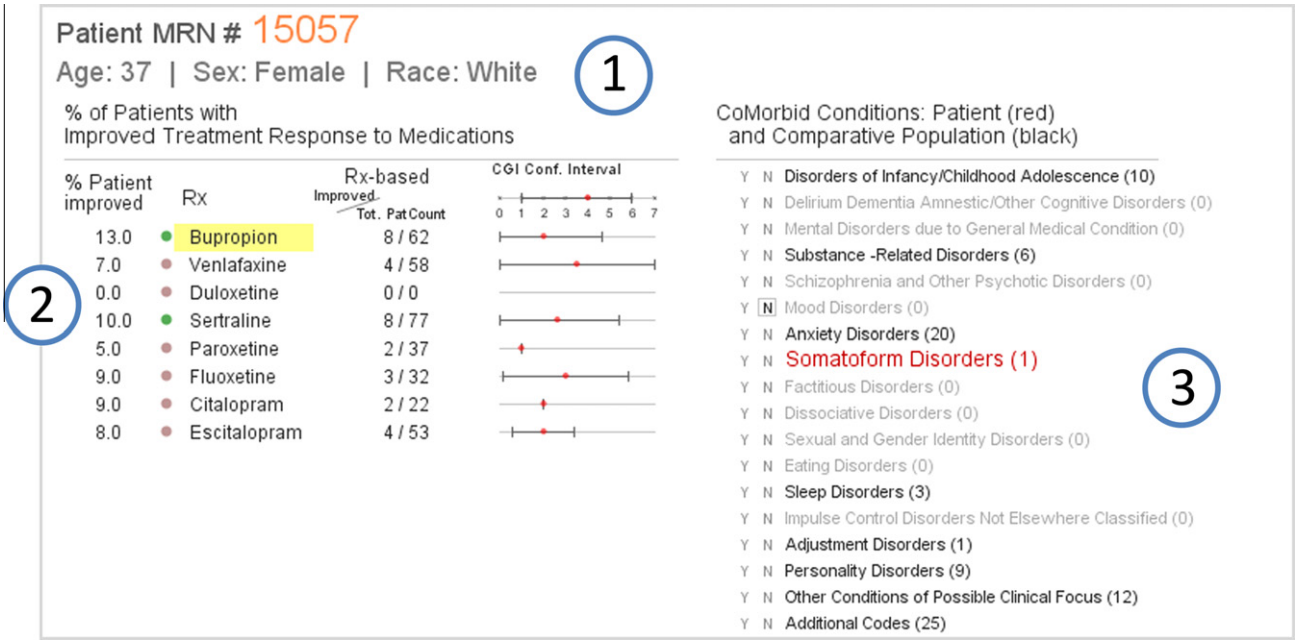

Fig. 1. Data view for patient demographics (label 1), summarized medication response (label 2), and comorbidities (label 3). A pseudo patient MRN 15057 is used for demonstration.

with wide gaps between visits, we only include patients with at least three visits of same medication during the 120-day time period. These criteria ensured that we include only those patients who are consistent in taking medications and schedule visits for refill. We use these criteria to filter initial dataset (3036 MDD patients) and identify patients who closely match the presenting patient profile. Thus, a large dataset is analyzed before we narrow down to a refined set of comparative patients who were prescribed the same medication.

The filtered dataset for the MDD comparative population is large and complex due to inclusion of multiple data variables such as demographics, comorbidities, medication, and treatment outcomes. The next challenge is to present the large dataset in a meaningful format that would enable the clinician to easily interpret the presented data. In addition, we wanted to provide an on-demand mechanism to further filter the data to allow the data view to be customized by the clinician at the point of care. The subsequent sub-sections below cover different data views that present the data characteristics, and on-demand filter mechanism.

4.2. Data views to display comparative population evidence from EHR data

4.2.1. Data view of patient demographics

Each patient within the EHR system had a unique medical record number (MRN) assigned to them. We used the MRN to identify patients and display demographic information. Fig. 1, label 1 is the patient demographics data view that shows patient's MRN, age, gender, and race.

4.2.2. Data view of summarized medication response

The prescribed medication and the associated CGI scores were used to build a data view on the MDD comparative population. At the computational level, a bin was created for every medication. From the MDD dataset (in Section 4.1), patients were added to bins of their prescribed medication, and at the same time were tagged based on their treatment outcome response (good response as CGI score $\leqslant 2$). Post binning process, additional computational approach was used to quantify the collective comparative MDD patient response into '% Patient Improved' score. This percentage score reflects a ratio of the number of patients who showed improvement (CGI score $\leqslant 2$) to the total number of patients on a given medication. Label 2 in Fig. 1 shows the data view of the comparative medication response for patients in the MDD dataset. The data view shows '% Patient Improved' score (on the left) and the absolute number of patients used to calculate the percentage (on the right). Visual encoded dots were added next to the medication names to distinguish between medications that reflect% patient improved score $\geqslant 10$ (green[1] dot) and less (red dot). Yellow shading was used to highlight the selected medication (in the middle). The CGI variance bound of the improved population is indicated by the confidence interval displayed for each medication (on the far right). At the UI level, single medication selection is allowed and the selected medication name is highlighted with a yellow background for visual feedback. A medication selection triggers updates to other data views (Sections 4.2.3 and 4.2.5) of the comparative population of patients on that medication.

4.2.3. Data view of comorbidities

A comorbid condition was defined as 'the presence of one or more disorders in addition to the primary diagnosed disorder'. Within the MindLinc data, every patient visit is tagged with the comorbid condition data at the DSM code level. We use this DSM code data to identify comorbid conditions for patients that are part of the comparative population. Label 3 in Fig. 1 shows the data view of comorbid conditions among patients on 'Bupropion'. At the UI level, we quantify the comorbid conditions shared by the comparative population. Visual encoding was applied to distinguish between the comorbidities of: the patient (larger font in red), the comparative population (in black), and not present comorbidities (in gray). By default for the selected medication, we show the entire comparative population irrespective of the presence or absence of comorbidities. The letters 'Y' and 'N' to

[1] For interpretation of color in Figs. 1–4, the reader is referred to the web version of this article.

the left of each comorbidity can be used as filter to indicate the inclusion or removal of patients that are in the comparative patients data. At the UI level, clicking the letter (Y or N) toggles its state. A small square around the selected letter was used to indicate the selected state of the comorbid condition. The selection of a comorbidity results in an instant update to other data views (Sections 4.2.2 and 4.2.5). Multiple selections of comorbid conditions are allowed at the UI level. This *ad hoc* ability to filter data and instant updates to the data view is useful for decision support at the point of care.

4.2.4. Data view of contextual patient treatment outcome

The label 1 in Fig. 2 shows the actual patient treatment outcome (CGI score) over time. At the UI level, years of treatment (2002–2007 in Fig. 2) are indicated by vertical axes (in gray) with patient CGI scores represented by a horizontal line (in orange). Summary of different visit types is available in the view (right bottom). Orange dots reflect outpatient visits, and an option is available show other visit types in the view. Prescribed medications and their time span is shown using horizontal bars, right below the CGI temporal view. Blue and red lines reflect the median response trend to the selected medication (details in next section). At the UI level, an available zoom feature was available to magnify a temporal portion of the data view.

4.2.5. Data view of median-based historical response to medication

The comparative population identified for the selected medication (Section 4.2.2) is used to build this data view. Historic CGI response of each patient in the comparative population with visits in proximity to 30, 60, 90, and 120 day time points were aggregated, and a median value was computed for each of the time points. Median value reflects the historic outcome response at each time point. At the UI level, the median historic response trend line (in blue) and the 95% confidence interval plot are shown at different time points (30, 60, 90, 120 days), see label 2 in both Figs. 2 and 3. Any change in the comparative population of the selected medication triggers automatic updates to this view.

4.3. VisualDecisionLinc: A dashboard-style coordinated display UI

The VDL UI shown in Fig. 4 represents a dashboard-style interface with a collection of different data views (Section 4.2). These

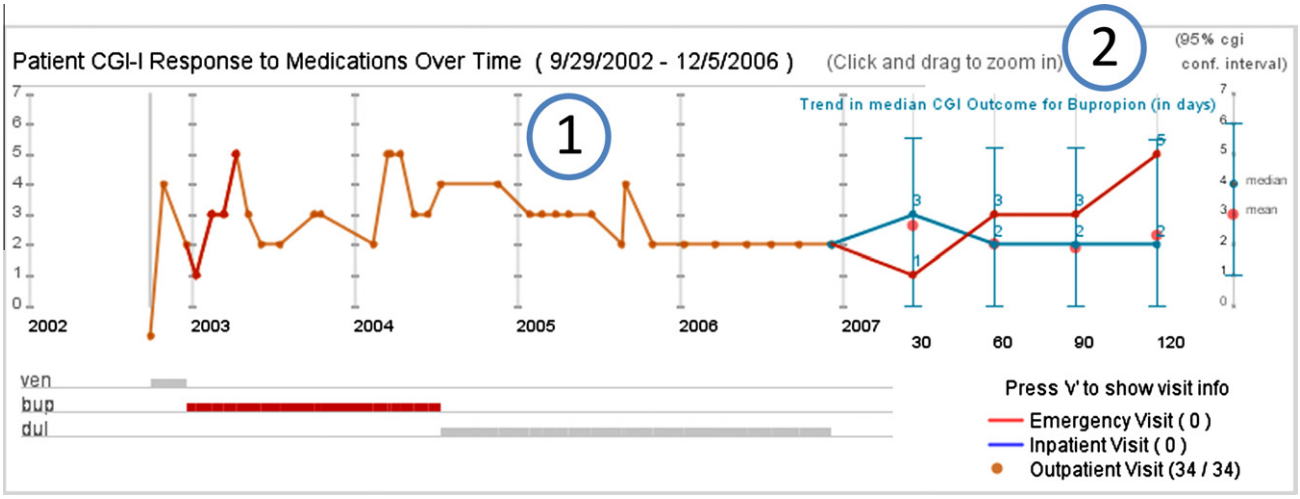

Fig. 2. Temporal overview of patient (pseudo MRN 15057) treatment outcome (in orange, label 1), and median based historical trend in medication outcome from a comparative population (in blue, label 2), along with patient's past 120 day response (in red, label 2) to the selected medication, if it was prescribed to the patient in the past.

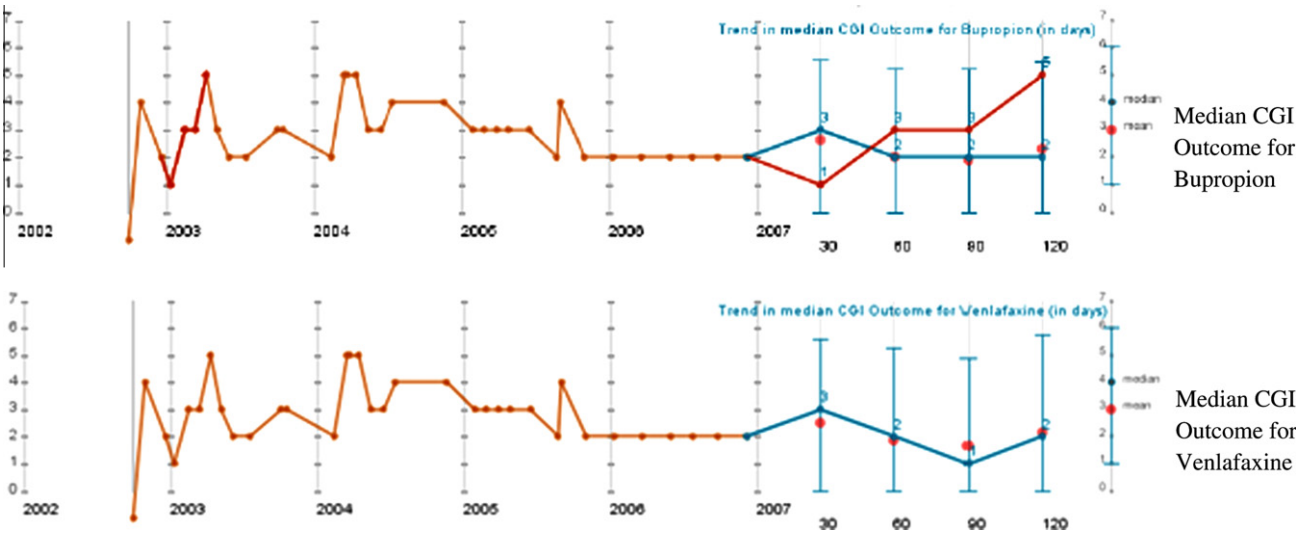

Fig. 3. For the same patient, the comparative patient based median CGI response is shown in blue for two different medications: bupropion (top) and venlafaxine (bottom).

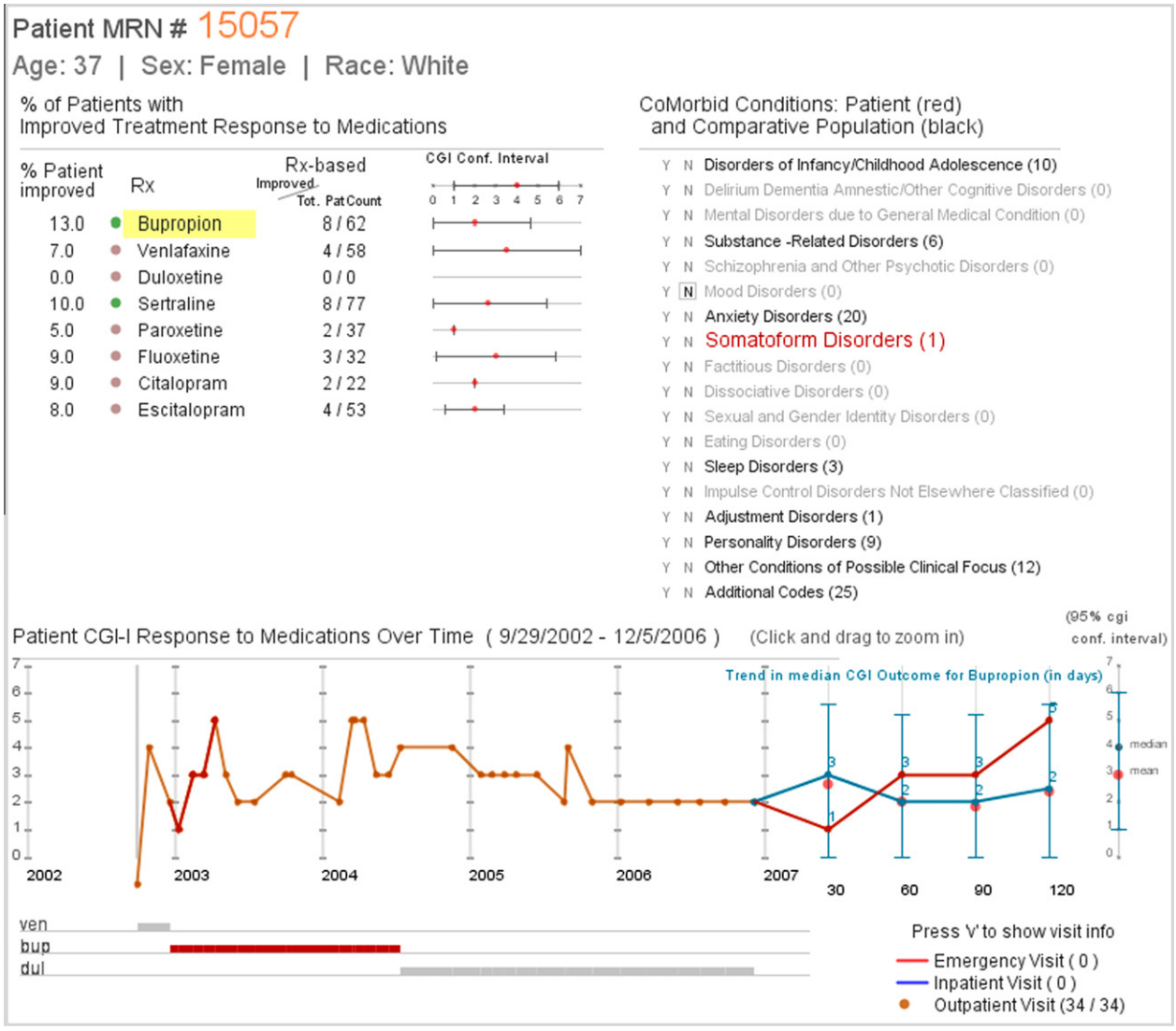

Fig. 4. VDL dashboard-style UI with all different data views combined in a single UI.

data views are linked together to form a coordinated display setup – filter updates to one view triggers relevant updates to data in other views. Coordinated display setup helps the user understand the relation between different data elements. *Ad hoc* filtering capability with instant updates to the data views allows for the evaluation of treatment options to better aid the decision-making process.

5. Limitations of the work

As part of our VDL design, we present real-time and non-peer reviewed comparative results drawn from EHR data. From an epidemiology perspective, we acknowledge EHR data related issues that need to be addressed, especially on – data inconsistencies, missing data, and biased data. Hence, as part of the VDL UI, we offer the clinicians an option to customize the data view to draw their own conclusions, but avoid drawing any direct conclusions. This is one of the reasons why we have avoided using modeled or machine-learning based results in our displays – instead we show summarizations of the data that can be readily consumed by the clinician. We accounted for the data inconsistencies, and filtered out the missing data elements from inclusion in the comparative patient population set by using data dictionaries. But comprehensive investigation of the inconsistencies in EHR data and mechanisms to account for missing data does demand more research. Particular bias in the presented data can be normalized to a certain extent by aggregating data from multiple locations. When presenting EHR data, we do realize that some data issues do indeed exist and need more thorough investigation.

The presented work focus is directed at VA use in the VDL UI to offer the clinicians an opportunity to better understand the EHR data, which is the source of their evidence. Inclusion of UI level indicators (like confidence intervals) offers insight into the underlying characteristics of the presented data. Temporal view provides general awareness of the trend in a patient's treatment outcome profile. The UI also offers an insight into the overlap between comorbid conditions of the patient and the comparative population. The scope of the presented work is to demonstrate the use of VA based UI to facilitate the clinician's understanding of the EHR data in the context of the presenting patient, and help evaluate viable treatment options based on a comparative population.

6. Discussion and future work

This paper highlights the role that VA can play in CDS. We used VA to bridge the complementary skill sets of humans and computers (as an external aid) to rapidly derive useful information from a large dataset. The external aid was used to augment the informa-

tion-processing abilities of the user to better guide the decision-making process. The VDL UI uses VA to facilitate the clinician's understanding of the data available in EHR system, and to help the clinician evaluate treatment options for a given patient. The VDL data views present aggregated observational data for clinicians to explore which treatment option is most likely to result in clinical improvements for patients with different medical conditions and comorbidities. Built-in interactions between the VDL data views and embedded data filters enable the customization of views to better suit the needs of the patient and the clinician. VDL offers a data-centric approach for leveraging the use of VA in CDS by displaying the relative effectiveness of different treatments options, showing the interactions between different data elements in a large dataset, reducing information overload, data integration at the point of care, and facilitating rapid clinical decision-making.

Also, the VDL presents an opportunity to take a similar approach where data are temporal in nature and associated with different events at different time points. Data views can be adapted to show different data types and their characteristics. The different views can be associated with each other to form a linked visualization. Underlying algorithms can be changed to build aggregate bins and perform trend analysis.

We plan to expand our work from one drug view to a combination medication view to account for the complex ways multiple medications are prescribed. Clearly, further research is needed to effectively view medication combinations. Also, where prescription overlap exists, techniques need to be explored to incorporate the presenting patient's past prescription history in the visualization of the median trend. Now that a prototype exists, it presents an opportunity to use a working foundation to use to identify the requirements of an EHR system to make it compatible for real-time CER use with reliable and valid data available. We plan to perform a formal evaluation of the clinical utility of our developed VDL CDS tool. Additionally, further efforts will include studies on how we can expand our VA approach and leverage other types of patient data (sequencing, metabolomics, etc.) as supplemental resources for CDS.

Acknowledgments

The authors declare no conflict of interest. We would like to thank Ricardo Pietrobon for his suggestions during the early stages of the VDL development. Also, we would like to thank Kara Fecho and Kirk Wilhelmsen for their input and for proof reading the document. This work was supported through funds from RENCI, and Agency for Healthcare Research and Quality (AHRQ) Grant R21HS019023.

References

[1] Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996;312(7023):71–2.
[2] Sox HC, Greenfield S. Comparative effectiveness research: a report from the institute of medicine. Ann Intern Med 2009;151(3):203–5.
[3] Iglehart JK. Prioritizing comparative-effectiveness research IOM recommendations. New Engl J Med 2009;361(4):325–8.
[4] Chaudhry B et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006;144(10):12–22.
[5] Johnston D, Pan E, Walker J, Bates DW, Middleton B. Assessing the value of computerized provider order entry in ambulatory settings. Boston (MA): Center for Information Technology; 2003.
[6] Ware C. Information visualization: perception for design. 2nd ed. San Francisco (CA): Morgan Kaufmann; 2004.
[7] Sittig DF et al. Grand challenges in clinical decision support. J Biomed Inform 2008;41(2):387–92.
[8] Thomas JJ, Cook K. Illuminating the path: the research and development agenda for visual analytics. National Visualization and Analytics Center; 2005.
[9] Keim DA. Visual exploration of large data sets. Commun ACM 2001;44:38–44.
[10] Plaisant C et al. LifeLines: using visualization to enhance navigation and analysis of patient records. In: Proceedings of the 1999 American medical informatics association symposium; 1998. p. 76–80.
[11] Shabtai A, Klimov D, Shahar Y, Elovici Y. An intelligent, interactive tool for exploration and visualization of time-oriented security data. In: Proceedings of the 3rd international workshop on visualization for computer security; 2006. p. 15–22.
[12] Wang TD, Plaisant C, Quinn AJ, Stanchak R, Murphy S, Shneiderman B. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In: Proceeding of the twenty-sixth annual SIGCHI conference on human factors in computing systems; 2008. p. 457–66.
[13] Moran M, Raju B, Saunders J, Meagher D. Achieving evidence-based prescribing practice in an adult community mental health service. Psychiatr Bull 2006;30(2):51–5.
[14] Haynes RB, Hayward RS, Lomas J. Bridges between health care research evidence and clinical practice. J Am Med Inform Assoc 1995;2(6):342–50.
[15] Bates DW et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc: JAMIA 2003;10(6):523–30.
[16] Gersing K, Krishnan R, Computing Clinical. Clinical management research information system (CRIS). Psychiatr Serv 2003;54(9):1199–200.
[17] Guy W. Clinical global impression (CGI): ECDEU assessment manual for psychopharmacology. Rockville (MD): US Department of Health, Education, and Welfare; 1976.