Designing a Visual Analytics System for Medication Error Screening & Detection

Tabassum Kakar[1], Xiao Qin[1], Cory M. Tapply[1], Oliver Spring[1], Derek Murphy[1], Daniel Yun[1], Elke A. Rundensteiner[1], Lane Harrison[1], Thang La[2], Sanjay K. Sahoo[2], and Suranjan De[2]

[1] Worcester Polytechnic Institute, Worcester, MA 01609, USA
{tkakar, xqin, cmtapply, obspring, dmurphy2, dyun, rundenst, ltharrison}@wpi.edu
[2] U.S. Food and Drug Administration, Silverspring, MD 20993, USA
{thang.la, sanjay.sahoo, suranjan.de}@fda.hhs.gov

Abstract. Drug safety analysts at the the U.S. Food & Drug Administration analyze medication error reports submitted to the Adverse Event Reporting System (FAERS) to detect and prevent detrimental errors from happening in the future. Currently this review process is time-consuming, involving manual extraction and sense-making of the key information from each report narrative. There is a need for a visual analytics approach that leverages both computational techniques and interactive visualizations to empower analysts to quickly gain insights from reports. To assist analysts responsible for identifying medication errors in these reports, we design an interactive Medication Error Visual analytics (MEV) system. In this paper, we describe the detailed study of the Pharmacovigilance at the FDA and the iterative design process that lead to the final design of MEV technology. MEV a multi-layer treemap based visualization system, guides analysts towards the most critical medication errors by displaying interactive reports distributions over multiple data attributes such as stages, causes and types of errors. A user study with ten drug safety analysts at the FDA confirms that screening and review tasks performed with MEV are perceived as being more efficient as well as easier than when using their existing tools. Expert subjective interviews highlight opportunities for improving MEV and the utilization of visual analytics techniques in general for analyzing critical FAERS reports at scale.

Keywords: Treemaps · Visual Analytics · Pharmacovigilance · Medication Errors

1 Introduction

Medication errors represent major threats to public health safety [19]. A medication error involves preventable mistake that can lead to either inappropriate use of the medication or patient harm while the medication is in the control of the patient, or health care professional [33]. Examples of such medication errors include but are not limited to the wrong administration or handling of drugs due

to ambiguous drug names or drug labels. It is estimated that medication errors cause 1 of 854 inpatient and 1 of 131 outpatient deaths [25] and an annual cost burden of $20 billion [1]. Medication errors may also cause adverse drug events [31], which can have detrimental consequences such as patient harm, unnecessary hospitalization, time away from work and additional resource utilization [43]. Such adverse events not only impact a patient's quality of life but can even lead to death [7]. Therefore, detection and prevention of medication errors is a high priority task for health care systems worldwide [2].

To detect and prevent medication errors, developed countries have established medication vigilance systems that maintain post marketing drug surveillance programs to capture reports about potential medication errors and adverse reactions, once the drug is released to the market [35]. In the U.S., the Food and Drug Administration (FDA) maintains the Adverse Event Reporting System (*FAERS*), to which consumers, health care professionals, and drug manufacturers submit reports related to medication errors and adverse events. At the FDA, drug safety analysts in the Division of Medication Errors Prevention and Analysis (DMEPA) analyze the FAERS reports to detect any errors in the administration, prescription or distribution of drugs. If a medication error is identified then it is addressed via regulatory actions such as, revising the usage instructions, or container labels, or changing a drugname.

A drug safety analyst reviews a report based on various factors including the severity, type and cause of the error and determines if the incident corresponds to a more general problem that may potentially warrant an action. Therefore, this analysis tends to require the evaluation of many reports over a long period of time. While these reports contain useful information such as patient demographics and administered drugs in a structured format, actual details about the error tend to be discussed in more depth in the text narrative associated with the report.

Currently, the tools used by the FDA's drug safety evaluators mainly support retrieval of reports from the FAERS database associated with their assigned set of products [8]. The analyst gathers the information regarding a specific error type manually by reading through each report narrative one by one. SQL is used to collect basic statistics about a report's collection, e.g., getting a count of the total number of reports for a particular product within a specified time, or retrieving reports based on other structured information such as demographics or reporter location.

These manual approaches are tedious and time consuming particularly as the volume of reports grows. Computational techniques alone are not sufficient in supporting the drug review process, as they lack the power to enhance the human perception and cognition to interactively manipulate data [15]. Visual analytics, on the other hand, is useful in allowing analysts interactively gain insights from data by analyzing data summaries and distributions [15]. There are a few widely accepted works on leveraging automation and combining it with a human-in-the-loop approach [32] to support analysts workflows. Our goal here

is to design interactive visualization and analytics techniques to address these limitations.

For this, we design MEV, a visual analytics system that supports the screening and analysis of medication error reports (Fig. 8). Based on the limitations of the FDA's current workflow, we first leverage recently developed natural language processing techniques [41, 4] to extract key information from the reports' narratives crucial for the detection of the medication errors. MEV uses a multilayer treemap visualization to display this extracted key information along with other structured information such as demographics of the patients affected by the errors. MEV's visual interactions guide analysts towards the most critical errors by empowering them to identify the associated data attributes. A timeline view allows analysts to see the overall distribution of the reports over a period of time. Finally, analysts can interactively study the reports associated with the screened errors.

This paper is an extension of our preliminary work [24] on leveraging visual analytics for the detection of medication errors. We now introduce a detailed analysis of the requirements that led to the design for MEV. In particular, we provide a description of the design process of MEV including the formative interviews with domain experts and the insights gained that helped in the design of MEV (Section 3). We elaborate on the importance of medication error detection via motivating use-case examples. We also present a detailed analysis of the current tools and workflow of the analysts at the FDA that lead us to gain an understanding of the analysts' pain points and limitations of the current tools. Based on these insights we solicit a set of design requirements for MEV (Section 3). We also include a detailed design process discussing various design alternatives and the decisions that led to the final design of MEV (Section 4). We enhance the evaluation of the system by including a detailed analysis of the system usability study (Section 6.2).

The results of our user study with ten analysts at the FDA suggest that with MEV, participants were able to perform review related tasks more quickly and perceived the tasks to be easier as compared with using their existing tools. Further, post-study qualitative interviews illustrate participants' interest regarding the use of visual analytics for medication error detection and suggestions for the improvement of the technology.

2 Related Work

In this section we focus on the existing work that study similar data types and analysis goals. The majority of our data including the key elements extracted using NLP such as drugnames, types and causes of errors are categorical, also called facets. A wide range of works have used facets as interactive filters to search and browse data [13, 27, 38]. FacetMap [38] focuses on using interactive visualizations for the exploration of facets in a data set. The system supports filtering, however, it does not support the analysis of relationships among facets. FacetLens [27], an extension of FacetMap, supports users to observe trends and

explore relationships within faceted datasets. Although these faceted systems [27, 38] are useful in data exploration, these systems are designed in a way that the interface is divided into a main view and a facet area. With such a design, only one data item (facet) can be browsed at a time. In the case of medication error screening, however, it is crucial to see the effect of selection of one item on others to support the quick identification of the data points representing more evidence towards the concerning identified errors.

Treemaps [5] have been widely used in visualization-based systems [28, 18]. For example, NV [18] allows analysts and system administrators to manage and analyze vulnerabilities on their networks via histograms and treemaps. SellTrend [28] utilizes treemaps to support the analysis and exploration of transaction failures in airline travel requests. These tools do not support the extraction of name-entities from textual data. JigSaw [39], on the other hand, is a comprehensive tool for investigating textual documents by visualizing name entities and their relationships to reveal hidden criminal plots. There is a need for the support of temporal data analysis for reviewing and screening reports that is not dealt with in JigSaw.

In the medical domain, few systems have been designed to help in preventing medication errors, such as clinical information systems [21] and medication-reconciliation tools [34]. Ozturk et al. [34] designed a system to help clinicians and emergency room staff review a patient's one year long prescription history to support their medication-reconciliation efforts. Varkey et al. [40] studied the effect of interventions on decreasing drug administration related medication errors. Other tools have focused on designing user-friendly and efficient interfaces to assist in error reporting [37]. On the other hand, clinical decision support systems focus on reducing medication errors during prescription [21]. Others [43] have attempted to prevent medication errors in clinical settings by using clustering techniques to group reports based on similarity scores. However, these tools are designed with the goal of medication error prevention during the administration or prescription of the drugs.

Our work instead supports the workflow that starts after the medication errors have already occurred and have been reported to the concerned regulatory authorities such as the FDA. For example, if two drugs have similar names and FDA receives error reports about these drugs being prescribed interchangeably. Then the FDA analysts carefully examine such reports to recommend actions so that different products can be differentiated easily, in this case, changing the drugname to prevent such error from happening in the future. To the best of our knowledge, no visual analytics tool exists that directly support the exploration and analysis of medication error reports.

3 Requirement Analysis

This section provides background information on the medication error reports and the importance of the detection of these errors. We conducted formative

interviews with drug safety analysts at the FDA to understand the data, tools and tasks involved in the analysis of these medication error reports.

3.1 Interviews with Domain Experts

We organized a series of one-on-one expert observations and semi-formal interview sessions with five drug safety analysts at the Division of Medication Error Prevention and Analysis (DMEPA) at the FDA. A similar approach had been taken for the design of DIVA [23], a visual analytics tool to explore adverse reactions related to drug-drug interactions. Our main goal was to gain an understanding of the current workflow of reviewing the stream of incoming error reports and in particular, to identify the challenges faced by the analysts while reviewing these reports.

Our first session was to become *familiar with the domain* and included discussions about the overall medication error detection process, the reports and the challenges of analyzing these reports. As a second step, to get precise facts and figures, we designed a *questionnaire* that included detailed questions about the workflow based on the initial discussions. Questions included, for instance, how many drugs a safety analyst gets assigned to? or, how many reports are received on a weekly basis? or, what is the most challenging aspect involved in the review of these reports?

Our second session with the same five drug safety analysts after few weeks involved field observation [26] followed by administering our carefully prepared questionnaire. We observed these analysts at the FDA in a think-aloud manner while performing their routine tasks of reports analysis. This helped us derive information on their existing practices and tools used for reviewing reports. These interviews revealed that certain information was critical to the workflow and also helped us identify the limitations of their current tools. These insights then were instrumental in being able to derive use-cases and elicit detailed requirements related to the tasks that could potentially leverage visual analytics to make the review process efficient.

The requirements were refined iteratively in a sequence of follow-up sessions through brainstorming and discussions with the same five analysts. Based on the collected requirements, we selected existing visualizations that could best support these processes and thus meet the identified requirements. That is, we showed these analysts sketches of multiple visual designs, such as parallel coordinates and node-link diagrams.The discussions about design alternatives helped us collect further design requirements, such as scalability and readability of the visualization. Based on these discussions, some of the design alternatives were rejected, while others were considered as potential candidates for the design of system. Subsequent discussions with the analysts lead to further refinements of the design and eventually one design namely, treemaps, was selected for implementation to realize the requirements which resulted in MEV.

In the later interviews, a working prototype of MEV was presented to these analysts to assess if MEV meets their needs. This then led to further suggestions to improve our visual and interaction design.

Lastly, we invited a larger group composed of ten FDA analysts, who were not involved in the design process of MEV, to participate in the user study to evaluate MEV (Section 6.2). This activity led to additional insights on the utility of MEV, a visual analytics tool in supporting the drug safety review process. We repeat on these final study results.

We next discuss the medication error reports, the review workflow and the challenges of current tools derived from the above interviews.

3.2 FAERS Data Description

Drug regulatory agencies worldwide maintain reporting systems as a part post-marketing drug surveillance programs to ensure public health safety [35]. In the U.S. the FDA maintains an Adverse Event Reporting Systems (FAERS) to monitor the safety of medications and therapeutic biologic products. FAERS contains voluntary reports submitted by health care professionals and consumers, and mandatory reports submitted by the drug manufacturers. These reports are composed of structured information about patient demographics, drugs being taken, therapies, and adverse reactions or medication errors. Each report also contains an unstructured textual narrative that provides a detailed description of the incidents such as medication errors or adverse reactions. These details about the incident contain richer information that can help the analysts decide if the incident is worthy of investigation and warrants an action. This information is critical for the analysis and includes the drugs or products, type and cause of error. Most of this information is categorical in nature.

Drug safety analysts review FAERS reports based on their assigned therapeutic classes of products. Each analyst receives an average number of approximately 200 reports on a weekly basis and monitors an average number of 50-100 products. These numbers may vary for different teams based on the assigned products. For instance, an analyst may maintain more than 100 products, as majority of her assigned products may rarely be causing an issue. Or, patients being familiar with an old product, might be making less mistakes than a newly approved drug which may be unfamiliar to use properly, hence generating more error reports. Additionally, for the investigation and detailed analysis of a particular error or product, thousands of reports from the past several months may be analyzed, if available.

3.3 Motivating Example of Medication Error Detection

To understand medication errors and the importance of their detection, we provide two motivating real examples. The first example includes a report received by the Institute for Safe Medication Practices (ISMP) about the medication Belbuca (Buprenorphine) [16]. Belbuca is prescribed to manage pain severe enough to require regular and long-term opioid treatment and for which alternative treatment options are inadequate. The medication is available in multiple strengths including 10 fold strengths, i.e., 75mcg and 750mcg (Fig. 1).

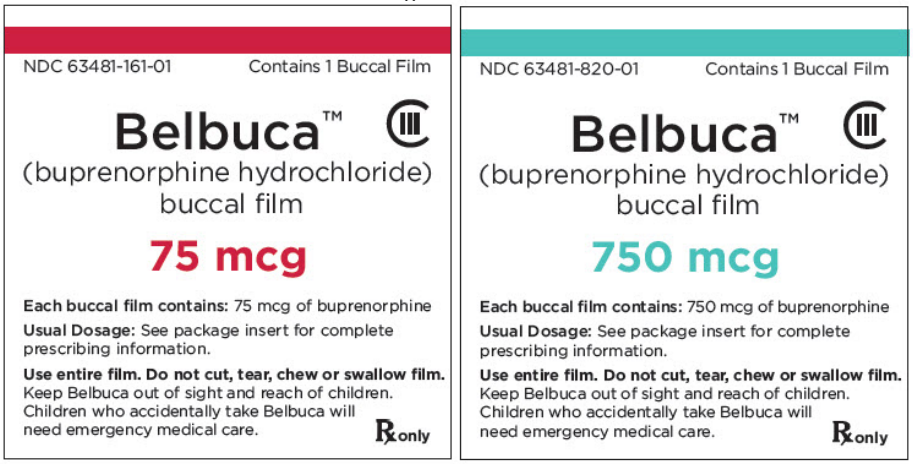

Fig. 1. Example of a potential 10-fold wrong strength error. A trailing zero (e.g., 75.0 mcg) can cause a mix-up of the two doses in prescription or dispensing of the drug.

The report describes a consumer who received a 750mcg of Belbuca instead of 75mcg, which is the recommended starting dose, from her pharmacy. After taking 5 doses of erroneous 750mcg strength, the patient experienced dizziness, lightheadedness, and vomiting. Upon follow up with her doctor, she learned that he had prescribed 75mcg, not 750mcg. One possible reason for such error could be misreading the hand written or typed prescription with a trailing zero (i.e., 75.0 mcg). Or, selecting the wrong strength from the drug drop-down list if 75mcg and 750mcg strengths appear next to each other in the computer system.

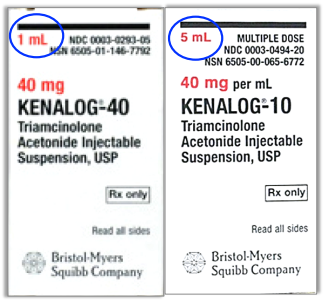

Fig. 2. Example of a look-alike carton label for the drug Kenalog resulting in an error.

Multiple instances of similar errors of mixing up 10 fold strengths of Belbuca as well as other medications such as Abilify (2mg and 20mg) are reported to drug safety authorities including ISMP and the FDA. A possible solution to these errors could be encouraging drug manufacturers to use strengths above or below an exact 10-fold difference, such as 749mcg instead of 750mcg [16].

Another example includes wrong strength error caused by a look-alike carton label. For instance, Fig. 2 represents two different strengths of Kenalog (i.e., 1ml and 5ml). The similarity between the two carton labels has caused errors of mistakenly using a higher concentration (5ml) instead of 1ml [6]. In this case, if the FDA receives multiple reports of similar errors with a product, then FDA may recommend to use a different color for the carton labels of each strength to easily distinguish between them. Seemingly benign errors like these can have adverse consequences such as death, if this were to cause patients to use the wrong drugs for treatment of critical conditions or diseases.

3.4 Workflow of Reports Review by Domain Experts

At the FDA, the drug safety analysts receive reports related to the drugs assigned to them. Every week, they receive a new set of reports which they analyze throughout the week along with other Pharmacovigilance activities such as reviewing drug labels or investigating a particular error. The goal of the analysis is to find potential severe medication errors that need action. To detect potential signals, the analysts first screen these reports based on criticality factors such as severe outcomes including hospitalization or death, particular error types or products. Then the text narrative of the screened reports are analyzed to get in-depth analysis of the errors. Once any alarming error is found, then FAERS is searched for other reports about similar errors searching backwards in time for at least last six months or a year. In some cases, the search can extend to the date when the product or error was last reviewed. Other sources are also analyzed for the investigation of a particular error.

3.5 Current Report Analysis Tools at the FDA

Currently, FDA uses FAERS Business Intelligence System (FBIS) [8] as one of the major tools to retrieve reports from FAERS database and analyze them. FBIS is designed with the goal of allowing analysts to review reports based on their assigned set of drugs, using queries to efficiently retrieve reports based on multiple dimensions. FBIS displays the list of reports in a tabular format, and provides controls to filter reports based on the structured information such as the age group or a product. For each report, only by reading the report narrative, the analysts can examine the details of the incident, such as, the type, cause or stage at which error has occurred. Examples of types of errors include taking a wrong dosage or using a wrong technique, the root causes of the errors include confusing instructions or container labels, and the stage in which the error has occurred include preparation versus dispensing.

For a detailed analysis, the analysts export the filtered reports for a given drug from FBIS into the Microsoft Excel and populate columns for the manually extracted information such as type, cause, and stage of error along with the any additional remarks added by the analysts about the report.

3.6 MEV Design Rationale

Our interviews with the analysts and their workflow observation (see Section 3.1) helped us identify the limitations of the current tools. These interviews also revealed that there is a need for an automatic system that first extracts the key information (type, cause and stage of error) from text narratives and then allows analysts to interactively explore and analyze this information, hence, making the report review process efficient.

This resulted in the identification of the below design requirements of a system that can support the exploration and analysis of medication error reports.

R1: ***Provide an overview of the core data elements for data screening.*** Analysts expressed a need for an overview of the key data attributes important for the analysis of medication errors. These attributes include the drugnames or products, types, causes and stages of errors. Each of these attributes consist of multiple categorical values. Such an overview should help an analyst see the distribution of reports per each attribute value to help them screen the critical errors.

R2: ***Support the exploration of reports over time.*** Drug safety analysts review reports on a weekly basis, however, they often analyze reports accumulated over a longer period of time. Hence, a way to interactively explore reports for a specified time is needed.

R3: ***Facilitate the analysis of demographics.*** The demographics of the patient also play an important role in the analysis of medication errors. The analysts should be able to see the age, location and occupation of the reporters. This would allow them to see, for instance, if an error is more prevalent in certain groups.

R4: ***Allow the analysts to quickly see the related attributes of a selected data point.*** Multiple errors may be reported for a given product. An error can be happening due to multiple reasons at different stages from the distribution of the drug to the administration. Therefore, analysts should be able to quickly get the gist of the associated data elements once a particular data value is selected.

R5: ***Facilitate identification of critical reports.*** The outcome of reports is an important factor in screening and prioritization of error reports. Analysts expressed a desire of being able to quickly identify any critical data points related to serious report outcomes. These outcomes indicate if the medication errors resulted in a serious outcome such as death or hospitalization of the patient, or were non-serious.

R6: ***Ready access to the reports narratives associated with selection.*** Analysts also indicated the importance of having direct access to the actual reports once a set of errors are screened. As these reports provide the details essential for decision making about the critical errors.

Besides the above mentioned requirements, the following have been found to be also equally important:

R7: ***Support smooth and interactive exploration of a large number of reports.*** Drug safety analysts review hundreds of drugs and their associated errors. Therefore, the system should be scalable to manage a large volume of reports.

R8: ***Using simple and familiar visual metaphors.*** Since the majority of the analysts had experience with basic visual analytics applications, it was emphasized to keep the visual designs intuitive and easy to understand. Clearly, careful design choices need to be made to consider these aspects.

4 MEV Design Process

We now discuss the design process of an interactive visual analytics system to help drug safety analysts in the screening reports by fulfilling the requirements solicited in Section 3. For (**R1**), a visualization was needed that display the distributions of categorical data representing our core data attributes (drugs, error types, causes and stages).

Also, visual interactions that help in assessing attribute values and their relationships with the rest of the data (**R4**) were needed. Therefore, we sketched several visual designs that meet these requirements (**R1, R4**) based on studying the state-of-the-art visualizations for multiple categorical attributes. Fig. 3 through 6 depict our alternate designs for (**R1**) that were repeatedly discussed with the domain experts (safety analysts) to get their feedback. We used color hue to encode the seriousness of reports (**R5**) with brown depicting highest count of serious reports while blue depicting the lowest count. This way the brown color can help analysts quickly identify critical data points. These color encodings are kept consistent throughout all designs.

We now discuss these designs and their pros and cons highlighted by the experts.

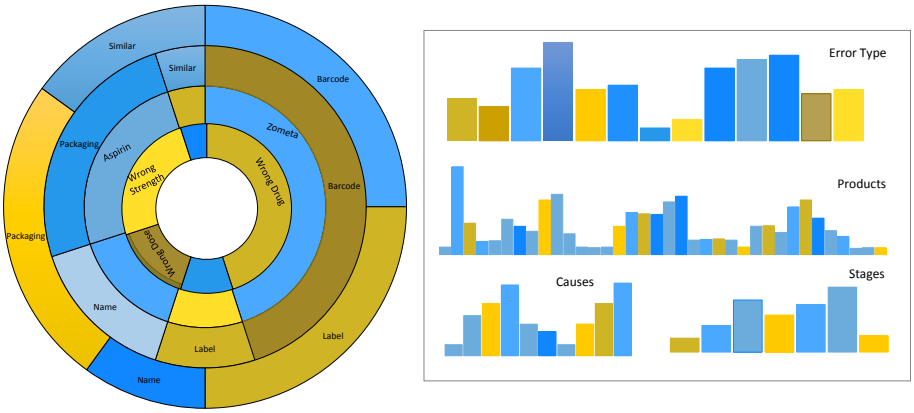

Fig. 3. Multi-layer radial chart (left) and bar-charts (right) for representing products, types, stages and causes of the errors **R1**. Color depicts the outcome of the reports.

4.1 Radial and Bar Charts

We started the design process with basic visualizations such as radial and bar charts to visualize the core data elements (**R1**) of the reports including the types, medications, stages, causes and outcomes of the errors (Fig. 3). Each layer in the radial chart (Fig. 3-left) represented the core attributes with the size of the arc mapped to the frequency of reports for each attribute value. Selecting any

Fig. 4. A node-link diagram sketch for one medication error "Wrong drug". The second layer depicts the products associated with the error. The third layer represents the distribution of causes and stages in the form of barcharts.

value on any layer updates the rest of the layers to reflect the data associated with the selection (**R4**). Similar mappings were used for the bar charts as well.

After discussions with the analysts, the radial chart was discarded due to its limited readability and scalability (**R7, R8**). The analysts expressed some concern about the complexity of the chart. On the other hand, the analysts seemed to like bar charts better than radial charts. However, their concern was that for a larger number of products (50 or more) the bars may be too tiny to read (**R7**).

4.2 Node-link Diagrams

Inspired by the existing works [9, 22] that has used node-link diagrams to visualize drug-related adverse reactions, we also considered a similar design to visualize the core attributes related to medication errors. As depicted in Fig. 4, the first layer of nodes represents the medication error type, while the second layer represents the products associated with each error type. The size of the nodes is mapped to the frequency of reports for the corresponding attribute value. The third layer uses a bar-chart to represent the distribution of stages and causes of the error.

Even though the domain experts considered the design sophisticated, they expressed some concern regarding the clutter and readability of the graph (**R7, R8**) when multiple error types and their associated products (50 or more) and other attributes are visualized at once.

4.3 Parallel Coordinates

We also sketched parallel coordinates [20], another well-known visualization method to represent multiple data attributes. Each coordinate represents one

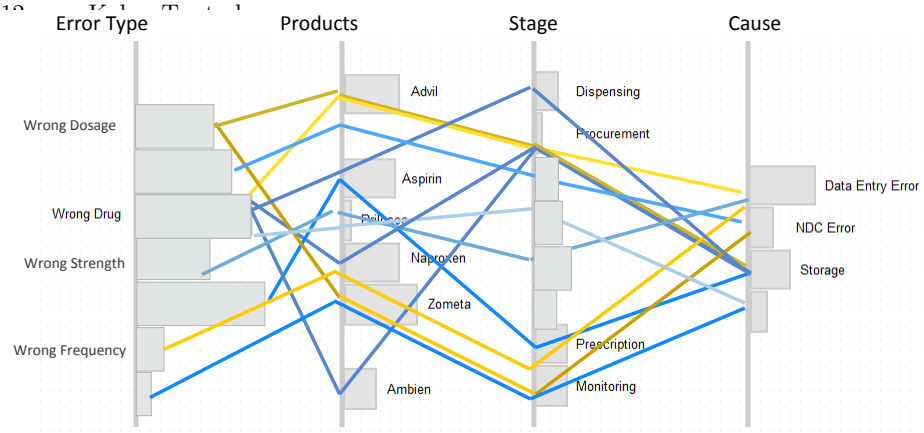

Fig. 5. A parallel coordinates sketch with each coordinate representing the products, medication errors, stage, causes of the errors.

of the core attributes, with bars depicting the frequency of the corresponding attribute values. The domain experts had similar concerns related to the readability of parallel coordinates due to visual clutter (**R7**).

4.4 Treemaps

Finally inspired by Selltrend [28], a tool that uses treemaps to visualize failures in transactions related to airline travel requests, we sketched a treemap design for our medication error screening problem (Fig. 6). The multiple layers of treemaps depict the different core attributes with each layer displaying the distribution of reports related to the corresponding attribute. A rectangle within a treemap layer represents one category for the attribute represented by that particular layer. For example, a rectangle in the product treemap represents one drugname. The size of the rectangle is mapped to the number of reports related to that specific category and the color encodes the count of the severe outcomes (**R5**) which is obtained from the structured field.

We received positive responses from the domain experts about the treemap design. The space-filling characteristic of treemaps makes them more readable. We then iteratively refined this design through discussions with domain experts, until the final design which was developed into a prototype system (Fig. 8). Some of these refinements included the change of color, as after implementation, the shades of brown and blue colors were getting mixed and were forming new colors (green). Therefore, color saturation was identified as a better choice to encode the seriousness of reports (**R5**) instead of color hue [32].

To fulfill (**R2, R3**), we sketched barcharts to display the distribution of reports over time (Fig. 7) as well as demographics (Fig. 8a). The analysts liked the barcharts for demographics (**R2**) and timeline (**R3**) due to their simplicity. Which is why, we do not show design alternatives for them in this paper. However, when the system was implemented, the timeline with barcharts (Fig. 7) did not look visually appealing. Hence, upon discussion with the experts, we

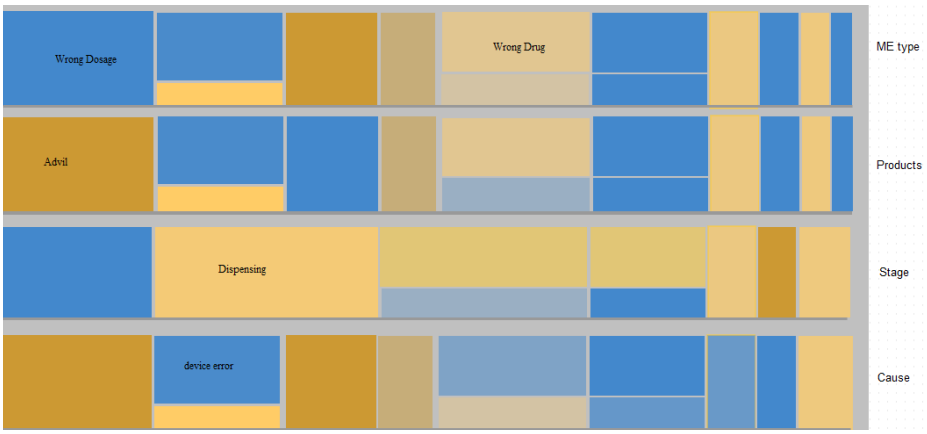

Fig. 6. A treemap based sketch that was realized into MEV.

decided to use area-curves to display the weekly reports distributions (Fig. 8-c). For examining the reports associated with the screened data (**R6**), the analysts suggested to have a view with both a line-listing of the reports as well as the text narratives themselves. This resulted in the design of the reports view as depicted in Fig. 10. Further suggestions on the reports view included adding search features for searching through both the narratives as well as the line-listings.

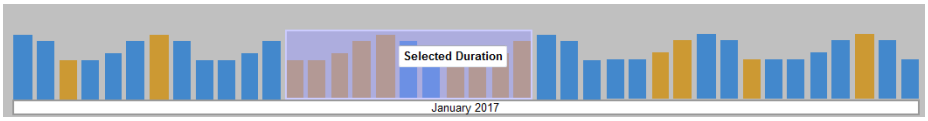

Fig. 7. Reports count over time with color representing serious outcomes.

5 The MEV System

In this section, we describe MEV's framework and interface design in detail.

5.1 The MEV Framework

The MEV framework, depicted in Fig. 9, is designed for analysts to efficiently perform the tasks of reports review. As we introduced earlier, FAERS reports consist of structured fields as well as text narrative with details about the medical events. In the case of medication error reports, the critical information about the error types and causes are often described in the text narrative rather than explicitly being stated in the structured field. To quickly extract these valuable information from the text narrative for the analysts, we adopt rule-based name-entity recognition techniques [41] in MEV.

The extractor is equipped with domain specific lexicons [33, 11] to identify key data attributes including types, causes and stages of medication errors. In MEV, the *Natural Language Processor* (Fig. 9) also performs preprocessing such

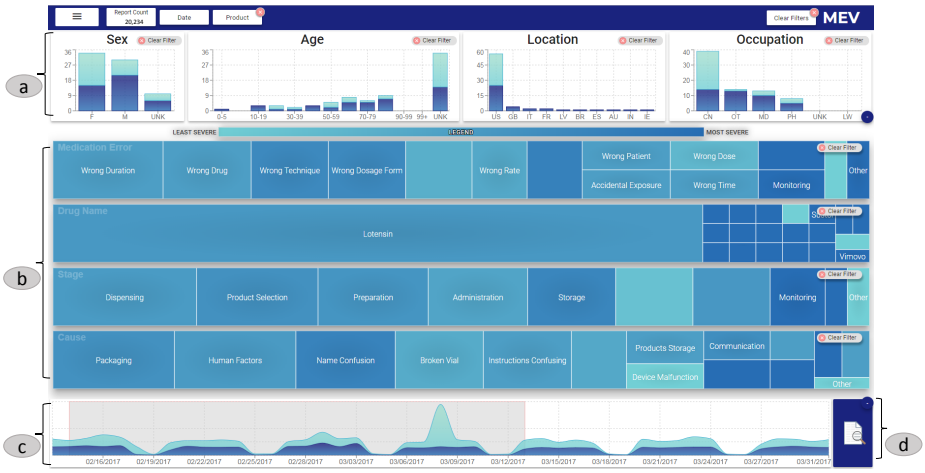

Fig. 8. The user interface of MEV (a) The demographics panel. (b) The treemap panel. (c) The timeline panel. (d) Reports icon to access the reports view to analyze the associated report narratives [24].

as word tokenization, stop-words removal and stemming before the core extraction process. The extracted results are then standardized by mapping them to the NCC-MERP terminology [33]. The mapping strategy is based on an *edit distance* based string matching algorithm [14].

On average, each of the standardized entity type contains approximately 15-20 categories. As illustrated in Fig. 9, both extraction results and the structured information about the demographics are stored in the *Data Store*. The *Query Executor* is responsible for handling data retrieval and transformation requests specified via online MEV visual interface. For better user experience, this module also has a caching mechanism to quickly respond to the most frequently issued requests. By using coordinated interactive visualizations, MEV assists analysts in the efficient exploration of reports as described below.

As depicted in Fig. 8, MEV consists of four core interactive displays – the treemap view, the demographics panel, the timeline panel and the reports view.

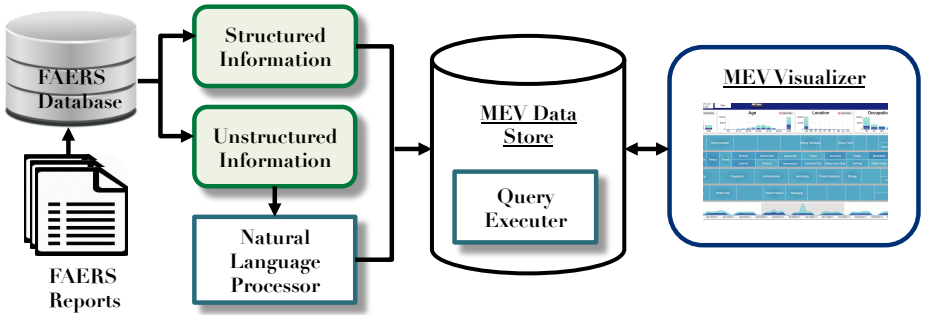

Fig. 9. The MEV framework [24].

5.2 MEV Interface Overview

The Treemap Panel. As discussed earlier, the treemap design (Fig. 8b) displays the distribution of each attribute over its possible categories.

This treemap visualization support analysts to interactively filter attribute such as drugname with even large number of categorical values (**R7**). MEV allows analysts to select multiple data values on each treemap at the same time and the other treemaps will be refreshed accordingly with the updated information corresponding to the selection. Such filtering functionality can facilitate analysts to quickly identify interesting information based on the presented data distribution (**R4**) which would take many tidy steps to achieve using their existing tools.

The Timeline Panel. The timeline panel, depicted in Fig. 8c, displays the report distribution based on volume and seriousness over a period of time using an area chart. Such design allows analyst to find spikes in the severity levels associated with the events of certain products (**R2**). It also supports interactive brushing and selection through zooming allowing analysts to quickly drill down to a particular time frame and explore specific reports. Similar to the interactivity on treemap, once a time frame is specified, other displays are then updated to reflect data corresponding to the selection (**R4**).

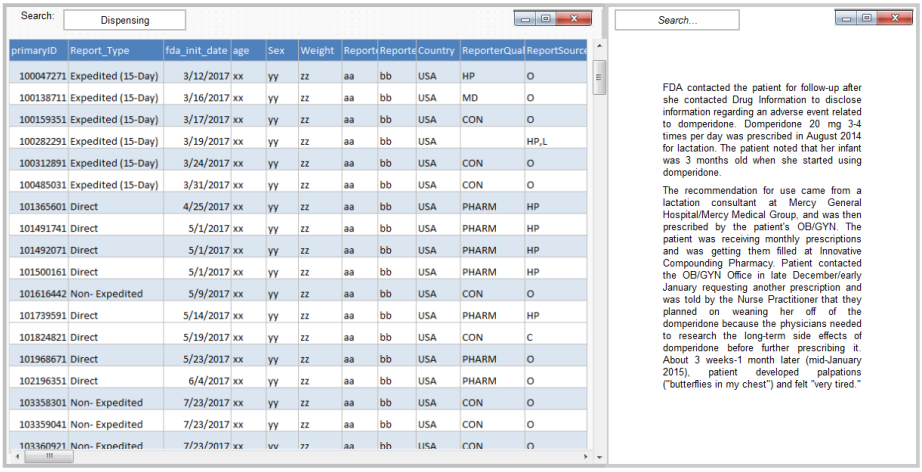

Fig. 10. Reports view with no personal information with line-listing (Left). Example of a de-identified text narrative (Right) [24]

The Demographic Panel. As pointed out by the domain analysts, the demographics information of the patients are critical for their analysis. For example, a severe outcome triggered by the administration of a certain medical product

Fig. 11. MEV with selection of date range, demographics, drugname and medication error type [24]

may only happen to pediatric patients as opposed to every age group. The demographics panel (Fig. 8a) can assist the analysts in quick identification and selection of problematic population defined by attributes such as age, gender or location (**R3**). This way the analysts can not only screen and prioritize reports based on selected attributes but also can immediately view the distributions of other associated attributes via linked displays (**R4**).

The Reports View. The analysts are able to browse the respective reports after they select the interesting medical products or error types as depicted in Fig. 10. This allows them to further investigate if these reports are indicative of errors with serious consequences for patient health warranting regulatory action (**R6**). The selected reports are accessible by clicking on the reports icon shown in Fig. 8d.

6 Evaluation

To validate the system, we first present a use case to highlight the usage and capabilities of MEV in data analysis tasks.

6.1 Use Case

To understand the use of MEV and how it can help in screening of reports, we describe an overview of actual series of steps performed by Alex, a drug safety analyst. Alex uses MEV to review reports related to her assigned set of products. At a glance on the timeline panel, she sees an overall weekly distribution of reports count and their severity for the last month (**R2**). She instantly can see that reports are only received during the weekdays, with no report visible during the weekend. She further observes a spike in the number of reports for the week

of march 3rd - 7th (2017), of which around 40% of reports are severe while the rest are non-severe (Fig. 11-bottom).

Using the brush tool on the timeline panel, she selects this week's reports to analyze them (**R2**). The treemap and demographics panels are updated for the selected week (**R4**). She sees that total number of reports for this selected week are 28,123. On the demographics panel, she sees that majority of patients are female and the most prevalent age group is above 30 (**R3**). Her assigned set of products are mostly for elderly women so the distribution of the age and gender is expected. She selects females from the gender barchart and U.S. from the location barchart to investigate the reported drugs and errors. This selection reduced the number of reports to 11,174.

She notices (on the treemap) that the medication error "wrong-technique" is reported with the most severe outcomes (**R1, R5**). She is now curious about the products associated with this "wrong-technique" error. She selects this error by clicking on the rectangle labeled as "wrong-technique" in the first treemap. Now the report counts is reduced to 2,786. On the second treemap for products, she observes that the drug Lotensin has the highest count of severe outcomes. She thus selects Lotensin in the second treemap. She is now curious that at which stage are these errors happening and what is causing them.

From the third and fourth treemaps which corresponds to the cause and stage of errors respectively, she observes that the major causes of this error are "packaging" and "name confusion". She says "The errors seem to be happening in the preparation of the drug". She also notices that the number of reports that she needs to analyze in detail, has reduced to 49 (Fig. 11). She opens the reports view by clicking on the reports icon (Fig. 8d) to read the details of each narrative to see if the reports indeed have compelling evidence about these errors (**R6**). Thus, MEV supports the exploration and screening of reports and interactively guides the analysts towards concerning errors.

6.2 User Study

We conducted a user study with drug safety analysts at the Division of Medication Error and Prevention Analysis (DMEPA) at the FDA to evaluate the effectiveness of MEV in various review tasks.

Study Design. We conducted a one hour in-person study session with ten drug safety analysts at the DMEPA. These participants included nine females and one male and were within the age range of 30-50 years. These participants were pharmacists with PharmD degrees and had experience with basic visualizations such as barcharts. These analysts are responsible to conduct regular reports review to identify any concerning errors that may need regulatory action.

Tasks and Assessment Measures. Based on our initial interviews with the analysts and the study of their current practices and workflow, we designed as set of nine commonly performed tasks to evaluate the usefulness of MEV (Table 1). These tasks contained simple one-step tasks to complex multi-step tasks.

Table 1. List of 9 Tasks designed to evaluate the effectiveness of MEV [24]

Task #	Description
T1	How many total reports have been reported during a time period?
T2	Which medication error is reported the most for a time period?
T3	Which drug has most severe outcomes for a selected medication error?
T4	Which gender and age have most severe outcomes?
T5	Which age group is most prevalent in reports related to a selected product?
T6	What are the two most frequent medication errors reported with a select product, age group, and gender?
T7	Given the report distribution of a drug for female patients with a specified age group, what are the critical medication errors that need to be analyzed?
T8	What are the two most frequent root causes of error for a selected drug and medication error?
T9	What are the two most common reported stages of errors for a drug and a medication error?

Example of a one-step task (T1-T2) included finding a particular attribute value for a selected time (**R2**). While complex tasks (T3-T6) included exploring reports associated with the analysis of the distribution of multiple data attributes (**R3-R5**). On the other hand, T8 & T9 included the analysis of reports based on the extracted data entities i.e., stage and cause of error (**R1**). The composite tasks (T3-T9) involved filtering based on the analysis of relationships between and among data attributes (**R4**). The main goal of the tasks was to find interesting data points to prioritize reports associated with critical errors (**R5**).

For each task, we measured the time taken to successfully complete the task as well as the easiness perceived by the participants. The perceived ease was recorded on a 5-point Likert scale (1 extremely difficult and 5 extremely easy). As data loading in the existing tool (FBIS) takes a long time to accomplish, so we recorded the time for completing a task **after** FAERS data was loaded into both tools for one week (from 2017). To compare the existing tools (*Control*) and MEV, the participants performed the same set of tasks with both tools.

Study Procedure. We conducted one hour in-person interview with the participants to observe them closely interacting with MEV and get their detailed feedback. The study consisted of three sessions. The first session (20 minutes) was dedicated to the demonstration and training. After the participants felt comfortable interacting with the system, in the second session the participants were asked to perform the set of previously designed tasks (Table. 1) using both MEV and their existing tool. In the third session, a post-study questionnaire was provided to the participants at which was not timed. The first part of the questionnaire included questions related to the demographics of the participants such as age and gender. The second section contained questions about the usability [10] of MEV on a 5-point Likert scale (1 strongly disagree & 5 strongly agree). The last part included an open-ended questionnaire to solicit qualitative feedback about MEV.

Analysis Method. We used the non-parametric Mann-Whitney U Test (Wilcoxon Rank Sum Test) [30] to compare the two conditions, as the recorded

time and perceived easiness for some tasks were not normally distributed (Table 2). Moreover, we calculated the 95% confidence intervals (Figure 12) for the task completion time and perceived easiness score.

Study Results. We now discuss the participants' performance on the tasks and their response regarding the overall system usability.

Quantitative Analysis. Table 2 shows that, for the majority of the tasks except Task T1, there are significant differences between the recorded time and perceived ease score when tasks were performed using MEV and the existing tool *control*.

Table 2. U-Test with significant results for both time and perceived ease for all tasks except T1 [24].

Tasks	Medviz - Avg Time (sec) [95% CI]	Current Tools Avg Time (sec) [95% CI]	Siginificance Test (α = 0.05)	MedViz Avg Easiness [95% CI]	Current Tools Avg Easiness [95% CI]	Siginificance Test (α = 0.05)
T1	5.11 [3.47,6.76]	3.62 [1.80, 5.44]	**(U = 29.5,p = 0.13104)**	4.8 [4.54, 5.06]	4.9 [4.70, 5.10]	**(U = 45, p = 0.72786)**
T2	7.54 [3.57, 11.52]	31.84 [15.78, 47.91]	(U = 14, p = 0.00736)	4.9 [4.70, 5.10]	4.0 [3.59, 4.41]	(U = 14, p = 0.00736)
T3	13.74 [9.83, 17.65]	78.31 [57.80, 98.83]	(U = 0, p = 0.00018)	4.7 [4.40, 5.00]	2.5 [2.06, 2.94]	(U = 1.5, p = 0.00028)
T4	10.04 [7.39, 12.69]	82.75 [46.51, 119.00]	(U = 0, p = 0.00018)	4.4 [3.97, 4.83]	2.6 [2.00, 3.20]	(U = 7.5, p = 0.00152)
T5	12.45 [9.12, 15.77]	67.05 [46.26, 87.84]	(U = 4, p = 0.00058)	4.5 [4.06, 4.94]	1.7 [1.19, 2.21]	(U = 1 , p = 0.00024)
T6	18.11 [13.75, 22.47]	83.58 [55.19, 111.96]	(U = 4, p = 0.00058)	4.4 [3.97, 4.83]	2.4 [1.73, 3.07]	(U = 7, p = 0.00132)
T7	18.52 [15.60 21.45]	64.84 [40.38 89.31]	(U = 11, p = 0.00362)	4.6 [4.28, 4.92]	2.2 [1.71, 2.69]	(U = 0, p = 0.00018)
T8	12.51 [8.45, 16.56]	Not Supported	Not Applicable	4.7 [4.40, 5.00]	Not Supported	Not Applicable
T9	13.07 [10.32, 15.83]	Not Supported	Not Applicable	4.6 [4.28, 4.92]	Not Supported	Not Applicable

The differences for T1 were not significant. T1 was a simple task that involved the finding of the total number of reports for a given duration of time. One possible explanation for this outcome could be that participants knew exactly where they will find this information in their current tool as they were used to it. On the other hand, being new to MEV, they took little longer (M=5.11 seconds [3.47, 6.76]) as compared to using their current tool (M=3.62 [1.80, 5.44]). Participants also rated this task easier under the *control* condition than using MEV.

For Task T2, there were significant differences between the performance using MEV (M =31.84 seconds [15.78, 47.91]) and existing tool (M = 7.54 seconds [3.57, 11.52]). T2 involved finding the most reported medication errors for a selected time period. In addition to time, we also observe that participants found it easier to perform the task using MEV (M = 4.9 [4.70, 5.10]) as compared to the control condition (M=4.0 [3.59, 4.41]).
Similarly, significant differences can be seen between the time and perceived ease scores for the multi-step tasks (T3 - T7) which involved the analysis of the distribution and severity of reports across multiple data attributes (Table 2). The composite tasks T8 and T9 involved retrieval of reports based on the stages and causes of errors with severe outcomes. The comparison for these two tasks was not possible as the causes and stages of errors were extracted using NLP for MEV and their current tools do not provide them.

Additionally, Fig. 12 (Top) shows that participants have a relatively consistent performance for all tasks. That is, all participants were able to successfully

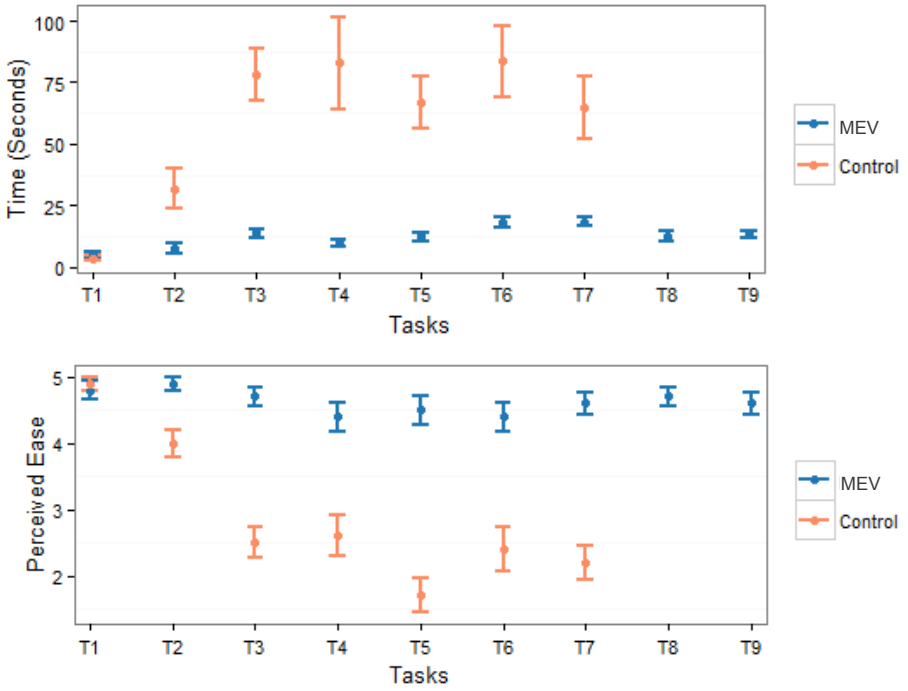

Fig. 12. 95% Confidence Interval for performing tasks using both MEV and existing tool (*Control*). (Top): Time (in sec). (Bottom): Perceived Ease Score. Task 8 & 9 are not supported by *Control* [24].

complete the tasks quickly. On the other hand, under the *control* condition participants had a highly varied performance. Similarly, from Fig. 12 (Bottom) we see that, the participants perceived it easier to perform tasks (T2-T9) using MEV than the existing tool. Participants rated the Task T5 which involved the analysis of the distribution of age for a selected product, as the most difficult under *control* condition. The reason is that with their existing tool it is tedious to explore the distribution of data attributes as it requires to filter each attribute value individually and then analyze the outcome of each value.

Qualitative Analysis & Overall Impression of MEV. The goal of qualitative questionnaires was to collect participants subjective feedback about the tool and their experience using it. The analysis of the questionnaire answers suggests that the participants experiences with the tool differed depending on their prior experience with similar interactive systems. For instance, some participants found it difficult to interact with the timeline panel to select dates, while others liked it.

Overall, the majority of participants found the goal of the MEV system in exploring and analyzing medication errors to be promising and potentially useful. Around 60% of the participants explicitly mentioned the usefulness of integrating the extracted data elements into the visualization and the intuitiveness of

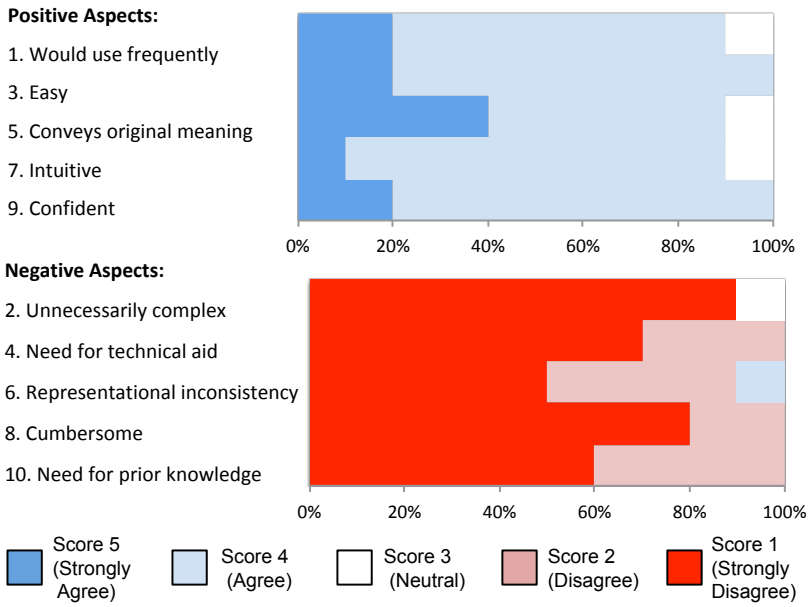

Fig. 13. System Usability Scale (SUS) Questionnaire Results

the tool itself. According to the participant P2: "Though the text-extraction is not perfect but it gives us a big sense of what kind of errors are being reported". Participant P5 said: "It takes some time to get used to the tool, then it is very easy and intuitive to use". Participant P10 mentioned: "Well, I think this tool makes it very easy to see what the reports are describing without going into much detail".

Participants constructive feedback about the potential improvement of the design of the system were also solicited using an open response option. For instance, four participants suggested to include an individual search option on each treemap to look up a particular error or drug. Others suggested to include a quick preview of the narratives on the treemaps view to get an instant gist of the text depicting an error.

Usability Questionnaire. At the end of the study, the participants were asked to complete the system usability scale (SUS) questionnaire [10]. Figure 13 depicts the distribution of the results of the questionnaire grouped as positive and negative questions. The majority of the participants agreed or strongly agreed with the questions related to the positive aspects of MEV, while only three participants had a neutral response. None of the participants had a negative response (scores of 1 or 2). On the other hand, for the negative questions, except for one participant who needed technical assistance with the tool, 80% of the participants disagreed or strongly disagreed (score of 1 or 2) with the questions. The overall score for the usability of the system was 85 out of 100.

7 Discussion

MEV aims to support the pharmacovigilance workflow using a visual analytics approach. The results of our user study suggest that analysts can perform review and screening tasks using MEV by exploring the relationships among data attributes. More broadly, our discussions with the drug safety analysts have highlighted additional challenges and opportunities for the medical professionals in the space of human-in-the-loop systems.

Scale is a key issue in modern visualization systems. Although the goal of MEV is to be used weekly by individual analyst for reports screening based on their assigned set of products which constitutes a count of thousands of reports. We tested MEV by loading data for the year 2017 which represented over 1.82 million reports, where the system took few seconds to load the data and transformed it to the initial overview. Other challenges of scale are associated with the visualizations themselves. If the analyst were to steer to a view with hundreds or more drugs, the rectangles on the treemap will become tiny and can cause visual clutter [36]. One solution to avoid the clutter is to only display a subset of drugs on the treemap and provide a search option to access a desired drugname. Another option to address the scale issue is to add a layer of treemap to represent drug classes. Analyst can first select a drug class and the second layer of treemap representing the drugs will be updated for the selected class. Alternatively, incorporating the domain practices into the system can also be a solution. For example, in typical cases clutter may not be a problem for MEV as the maximum number of distinct products in the reports for an analyst does not exceed 100.

During our qualitative interviews although the majority of the analysts acknowledged the usefulness of MEV in reports screening, few analysts expressed their preference for reading each and every report narrative when the number of reports was small (e.g., 10 or 20), rather than using MEV for screening. For such analysts, a feature to highlight the key information within the text narratives can be added. During our user study, we also observed that some of the extracted information was incorrect, when the participants retrieved the reports to analyze their narratives. For this system, we leveraged the MEFA [41] framework for name-entity extraction such as the cause and stage of the error. For improved accuracy, more advanced deep learning-based extraction techniques [42] could be plugged into MEV. However, name-entity recognition techniques for biomedical text itself is known to be a challenging problem and research efforts continue to improve their accuracy [3].

Our user study has a number of limitations. First, participants' familiarity with their existing tools allowed them to complete some complicated tasks in a short time. Also, few participants found some tasks irrelevant. For instance, participants who usually investigate one particular product found it irrelevant to find the reports associated with multiple products based on severity. Study participants, although a small number, are real domain experts who would be the ultimate users of MEV in daily analysis. Long term studies with these experts are needed to further assess the capabilities of MEV in their workflow.

In the future, we plan to provide direct access to external sources such as DailyMed and PubMed from within MEV. This will help the analysts to investigate these sources to reject or confirm a hypothesis about a possible medication error formed using the current views. We also plan to extend the capability of the current system by providing support for the report text analysis. Finally, we will add the capability of visual provenance [17] to allow analysts to share their thought-processes and findings with their team members during an investigation.

8 Conclusion

In this paper, we described the detailed design process of MEV - a prototype visual analytics tool for medication error detection using spontaneous reporting system. MEV uses multiple coordinated views including interactive treemaps, bar charts and timeline visualization to support the exploration and screening of spontaneous reports. MEV guides the analysts to visually identify critical errors by allowing them to evaluate the distributions of multiple facets of data across several weeks. A task-based user study with drug safety analysts at the FDA suggests that MEV is effective in supporting the analysts in their review tasks, allowing them to complete their tasks more efficiently, in comparison to their existing tool. Moreover, qualitative feedback from the participants highlights opportunities to improve the system. This is valuable for future developments of MEV specifically as well as of for the design of visual analytics tools for critical incident reports analysis in general – a ubiquitous task in domains such as aviation [29] and finance [12].

References

1. *A $21 Billion Opportunity National Priorities Partnership convened by the National Quality Forum file* (2010), `http://www.nehi.net/bendthecurve/sup/documents/Medication_Errors_Brief.pdf` [Accessed: 2018-01-07]
2. Agrawal, A.: Medication errors: prevention using information technology systems. British journal of clinical pharmacology **67**(6), 681–686 (2009). https://doi.org/10.1111/j.1365-2125.2009.03427.x
3. Alshaikhdeeb, B., Ahmad, K.: Biomedical named entity recognition: A review. International Journal on Advanced Science, Engineering and Information Technology **6**(6), 889–895 (2016). https://doi.org/10.18517/ijaseit.6.6.1367
4. Aronson, A.R., Lang, ç.M.: An overview of metamap: historical perspective and recent advances. Journal of the American Medical Informatics Association **17**(3), 229–236 (2010). https://doi.org/10.1136/jamia.2009.002733
5. Asahi, T., Turo, D., Shneiderman, B.: Visual decision-making: using treemaps for the analytic hierarchy process. In: The Craft of Information Visualization, pp. 235–236. Elsevier (2003). https://doi.org/10.1016/b978-155860915-0/50030-5
6. Authority, P.P.S.: *Drug Labeling and Packaging Looking Beyond What Meets the Eye* (2017), `http://patientsafety.pa.gov/ADVISORIES/documents/200709_69b.pdf` [Accessed: 2019-06-25]

7. Bates, D.W., Spell, N., Cullen, D.J., Burdick, E., Laird, N., Petersen, L.A., Small, S.D., Sweitzer, B.J., Leape, L.L.: The costs of adverse drug events in hospitalized patients. Jama **277**(4), 307–311 (1997). https://doi.org/10.1001/jama.1997.03540280045032
8. BIFACT: *FAERS Business Intelligence System (FBIS)*, `http://www.bifact.com/faers-bifact.html` [Accessed: 2019-06-19]
9. Botsis, T., Jankosky, C., Arya, D., Kreimeyer, K., Foster, M., Pandey, A., Wang, W., Zhang, G., Forshee, R., et al.: Decision support environment for medical product safety surveillance. Journal of biomedical informatics **64**, 354–362 (2016). https://doi.org/10.1016/j.jbi.2016.07.023
10. Brooke, J., et al.: Sus-a quick and dirty usability scale. Usability evaluation in industry **189**(194), 4–7 (1996)
11. Brown, E.G., Wood, L., Wood, S.: The medical dictionary for regulatory activities (meddra). Drug safety **20**(2), 109–117 (1999). https://doi.org/10.1002/9780470059210.ch13
12. CFPB: *Consumer Financial Protection Bureau*, `www.consumerfinance.gov/` [Accessed: 2019-06-03]
13. Dachselt, R., Frisch, M., Weiland, M.: Facetzoom: a continuous multi-scale widget for navigating hierarchical metadata. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1353–1356. ACM (2008). https://doi.org/10.1145/1357054.1357265
14. Du, M.: Approximate name matching. NADA, Numerisk Analys och Datalogi, KTH, Kungliga Tekniska Högskolan. Stockholm: un pp. 3–15 (2005)
15. Fekete, J.D., Van Wijk, J.J., Stasko, J.T., North, C.: The value of information visualization. In: Information Visualization, pp. 1–18. Springer (2008). https://doi.org/10.1007/978-3-540-70956-5$_1$
16. Gaunt, M.J.: *Preventing 10-Fold Dosage Errors*, `www.pharmacytimes.com/publications/issue/2017/july2017/preventing-10fold-dosage-errors` [Accessed: 2019-06-24]
17. Groth, D.P., Streefkerk, K.: Provenance and annotation for visual exploration systems. IEEE Transactions on visualization and Computer Graphics **12**(6), 1500–1510 (2006). https://doi.org/10.1109/tvcg.2006.101
18. Harrison, L., Spahn, R., Iannacone, M., Downing, E., Goodall, J.R.: Nv: Nessus vulnerability visualization for the web. In: Proceedings of the ninth International Symposium on Visualization for Cyber Security. pp. 25–32. ACM (2012). https://doi.org/10.1145/2379690.2379694
19. Huckels-Baumgart, S., Manser, T.: Identifying medication error chains from critical incident reports: a new analytic approach. The Journal of Clinical Pharmacology **54**(10), 1188–1197 (2014). https://doi.org/10.1002/jcph.319
20. Inselberg, A.: The plane with parallel coordinates. The visual computer **1**(2), 69–91 (1985). https://doi.org/10.1007/BF0189835
21. Jia, P., Zhang, L., Chen, J., Zhao, P., Zhang, M.: The effects of clinical decision support systems on medication safety: an overview. Public Library of Science One **11**(12), e0167683 (2016). https://doi.org/10.1371/journal.pone.0167683
22. Kakar, T., Qin, X., McCarthy, B., Schade, A., Tran, H.Q., Zylich, B., Rundensteiner, E.A., Harrison, L., Sahoo, S.K., De, S.: DEVES: Interactive signal analytics for drug safety. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management. pp. 1891–1894 (2018). https://doi.org/10.1145/3269206.3269211

23. Kakar, T., Qin, X., Rundensteiner, E.A., Harrison, L., Sahoo, S.K., De, S.: DIVA: Towards Validation of Hypothesized Drug-Drug Interactions via Visual Analysis. In: Eurographics (2019). https://doi.org/10.1111/cgf.13674
24. Kakar, T., Qin, X., Tapply, C.M., Spring, O., Murphy, D., Yun, D., Rundensteiner, E.A., Harrison, L., La, T., Sahoo, S.K., De, S.: MEV: Visual Analytics for Medication Error Detection. In: International Conference on Information Visualization Theory and Applications (IVAPP), Prague, 2019. SciTePress (2019). https://doi.org/10.5220/0007366200720082
25. Kohn, L.T., Corrigan, J., Donaldson, M.S., et al.: To err is human: building a safer health system, vol. 6. National academy press Washington, DC (2000). https://doi.org/10.1016/s1051-0443(01)70072-3
26. Lam, H., Bertini, E., Isenberg, P., Plaisant, C., Carpendale, S.: Empirical studies in information visualization: Seven scenarios. IEEE transactions on visualization and computer graphics **18**(9), 1520–1536 (2011). https://doi.org/10.1109/TVCG.2011.279
27. Lee, B., Smith, G., Robertson, G.G., et al.: Facetlens: exposing trends and relationships to support sensemaking within faceted datasets. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1293–1302. ACM (2009). https://doi.org/10.1145/1518701.1518896
28. Liu, Z., Stasko, J., Sullivan, T.: Selltrend: Inter-attribute visual analysis of temporal transaction data. IEEE Transactions on Visualization and Computer Graphics **15**(6), 1025–1032 (2009). https://doi.org/10.1109/TVCG.2009.180
29. Marais, K.B., Robichaud, M.R.: Analysis of trends in aviation maintenance risk: An empirical approach. Reliability Engineering & System Safety **106**, 104–118 (2012). https://doi.org/10.1016/j.ress.2012.06.003
30. McKnight, P.E., Najab, J.: Mann-whitney u test. The Corsini encyclopedia of psychology pp. 1–1 (2010). https://doi.org/10.1002/9780470479216.corpsy0524
31. Morimoto, T., Gandhi, T., Seger, A., Hsieh, T., Bates, D.: Adverse drug events and medication errors: detection and classification methods. BMJ Quality & Safety **13**(4), 306–314 (2004)
32. Munzner, T.: Visualization analysis and design. AK Peters/CRC Press (2014). https://doi.org/10.1201/b17511
33. NCC-MERP: *National Coordinating Council for Medication Error Reporting and Prevention* (1995), `http://www.nccmerp.org/` [Accessed: 2018-02-28]
34. Ozturk, S., Kayaalp, M., McDonald, C.J.: Visualization of patient prescription history data in emergency care. In: AMIA Annual Symposium Proceedings. vol. 2014, p. 963. American Medical Informatics Association (2014)
35. Patel, I., Balkrishnan, R.: Medication error management around the globe: an overview. Indian journal of pharmaceutical sciences **72**(5), 539 (2010). https://doi.org/10.4103/0250-474x.78518
36. Peng, W., Ward, M.O., Rundensteiner, E.A.: Clutter reduction in multi-dimensional data visualization using dimension reordering. In: IEEE Symposium on Information Visualization. pp. 89–96. IEEE (2004). https://doi.org/10.1109/INFVIS.2004.15
37. Singh, R., Pace, W., Singh, A., Fox, C., Singh, G.: A visual computer interface concept for making error reporting useful at the point of care. In: Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 1: Assessment). Agency for Healthcare Research and Quality (2008)
38. Smith, G., Czerwinski, M., Meyers, B.R., et al.: Facetmap: A scalable search and browse visualization. IEEE Transactions on Visualization and Computer Graphics **12**(5), 797–804 (2006). https://doi.org/10.1109/TVCG.2006.142

39. Stasko, J., Gørg, C., Liu, Z.: Jigsaw: supporting investigative analysis through interactive visualization. Information Visualization **7**(2), 118–132 (2008). https://doi.org/10.1057/palgrave.ivs.9500180
40. Varkey, P., Cunningham, J., Bisping, S.: Improving medication reconciliation in the outpatient setting. Joint Commission Journal on Quality and Patient Safety **33**(5), 286–292 (2007)
41. Wunnava, S., Qin, X., Kakar, T., Socrates, V., Wallace, A., Rundensteiner, E.A.: Towards transforming fda adverse event narratives into actionable structured data for improved pharmacovigilance. In: Proceedings of the Symposium on Applied Computing. pp. 777–782. ACM (2017). https://doi.org/10.1145/3019612.3022875
42. Yadav, V., Bethard, S.: A survey on recent advances in named entity recognition from deep learning models. In: Proceedings of the 27th International Conference on Computational Linguistics. pp. 2145–2158 (2018)
43. Zhou, S., Kang, H., Yao, B., Gong, Y.: Analyzing medication error reports in clinical settings: An automated pipeline approach. In: AMIA Annual Symposium Proceedings. vol. 2018, p. 1611. American Medical Informatics Association (2018)