Visualization of Medicine Prescription Behavior

Paul van der Corput¹, Johan Arends²,³, and Jarke J. van Wijk¹

¹Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands
²Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands
³Kempenhaeghe Institute for Epileptology, Sleep Medicine, and Neurocognition, The Netherlands

Abstract
Medicine prescriptions play an important role in medical treatments. More insight in medicine prescription behavior can lead to more efficient and effective treatments, as well as reflection on prescription behavior for specific physicians, types of medicines, or classes of patients. Most current medical visualization systems show health data only from the perspective of patients, whereas to understand prescription behavior multiple perspectives are relevant. We present a new approach to visualize prescription data from four different perspectives: physician, patient, medicine, and prescription. Information about physicians, patients, and medicines is shown in three tables; relations between selected items in these tables are shown using custom glyphs and histograms. These tables can also be used to define selections of prescriptions which can be compared to each other by showing a variety of metrics. This enables physicians and possibly other stakeholders to perform a wide variety of queries and inspections, while the use of familiar metaphors, such as tables and histograms, enables them to use the system in short time. This was confirmed by an evaluation session with six neurologists from an institute of epileptology. Our system is tailored to medicine prescription data, but we argue that the underlying pattern in the data is ubiquitous, and that hence our approach can be useful for many other cases where A provides B to C.

Categories and Subject Descriptors (according to ACM CCS): H5.2 [Information Systems]: User Interfaces—Graphical user interfaces H5.2 [Information Systems]: User Interfaces—Interaction styles

1. Introduction
Electronic patient data and medical prescription histories are a valuable source of information. These are primarily stored to help physicians to find suitable medications for their patients. Based on the patient’s history and current circumstances, some medicine and dose can be selected. This prescription history is often a simple log file that describes which medicine is prescribed to which patient by which physician, and can nowadays easily contain hundreds of physicians, thousands of patients, and millions of prescriptions. Our aim is to enable physicians or other medical experts to discover patterns in this data and use this information to support physicians, patients, and other stakeholders.

As long as there exist electronic patient records, tools have been developed to visualize them. In the nineties, systems like the Time Line Browser [CK91] and LifeLines [PMS*98] were developed to show per-patient information. More recently, studies were done to visualize the interaction between patients and medicines, which resulted in solutions like LifeLines 2 [WPS*09], the Chronograph [NHB*10], and MatrixFlow [PS12]. Those visualizations enable the discovery of correlations between events (diagnoses, medicine side effects, etc.) and medicine prescriptions. Systems that focus on the role of the physician, however, do not exist yet.

Insight in prescriptions from a physician’s perspective is highly relevant, because physicians are the link between patients and medicines. Physicians have often freedom which medication to prescribe. The way they deal with this freedom can be described as their prescription behavior. In order to discover the best possible treatments, we need a way to reflect on this. By visualizing the prescription behavior, we are able to discover differences between physicians. Such discoveries can give rise to discussions, which will eventually lead to better treatments.

This project is done in cooperation with the Kempenhaeghe Institute for Epileptology, which has logged all medical prescriptions of its neurologists for over three decades. Epilepsy is a neurological disorder that can in most cases be controlled, but not cured with medication. Prescribing the
right medication is therefore crucial to achieve an acceptable quality of life for patients with epilepsy.

In this paper, we present an approach to give physicians insight into their prescription behavior. Some existing visualizations also take additional events and measurements into account, but a problem is that these are often not available. It is for instance impractical to let patients log all their epileptic seizures at home. Prescription data is already interesting on its own, because it can reveal many properties of physicians, patients, and medicines. Although this logged data may look simple, it is however not trivial to visualize its structure.

Prescription behavior such as described above is not covered by related work, as we discuss in Section 2. Therefore, we analyze the problem domain in Section 3, and present our approach in Sections 4 and 5. An evaluation with a prototype confirms that experts are able to answer complex questions using our approach. In Sections 7 and 8 we discuss a broader use of our methods, future work, and give our conclusions.

2. Related work

Several surveys give an overview of electronic health record (EHR) visualization. Rind et al. [RWA’13] give a very comprehensive overview, where they make a distinction between visualizations of single patient histories and visualizations that aggregate over multiple patients. Another survey, by Kosara and Miksch [KM02], is about medical data analysis and planning, and distinguishes between past and future events. Past events contain measurements and incidents, future events have some level of uncertainty, which has to be taken into account. Other overview papers are by Chit-taro [Chi01] and Roque et al. [RST10]. The most relevant visualizations for prescription visualization are discussed below.

2.1. Electronic health record visualizations

The Time Line Browser by Cousins and Kahn [CK91] and LifeLines by Plaisant et al. [PMS’98] are examples of early systems that visualize medical patient histories. In these systems, the time axis plays a central role, which works excellent for single patient’s EHR visualization.

Wang et al. [WPQ’08] presented the concept of a sentinel event, which is the first event of a certain type in one record, for example the prescription of medicine m. Each patient record that contains a prescription of medicine m, can then be aligned by the first event of this type, such that a relative time scale appears. This idea is adopted in the Chronograph by Norén et al. [NHB’10] and LifeLines 2 by Wang et al. [WPS’09]. These visualizations show histograms with the temporal distribution of all events relative to the sentinel event. EventFlow by Monroe et al. [MLL’13] also aligns patient records in order to summarize temporal events that come from noisy datasets. Instead of looking at the temporal spacing between an event and a prescription, we can also look at the co-occurrence between them in several time segments. This idea is used in MatrixFlow by Perer et al. [PS12], which constructs an adjacency matrix with co-occurring events for predefined time segments.

2.2. Generic solutions

Many physicians can work at the same hospital, and in addition, a patient can be treated by more than one physician. The systems mentioned above do not take into account who makes important decisions such as prescriptions. We can therefore only make statements about prescriptions in a hospital as a whole, but are unable to compare individual physicians, which is crucial when analyzing prescription behavior.

For this reason we investigate techniques other than existing EHR visualizations. Since we want to focus on exploratory browsing through individual (or groups of) physicians based on certain properties, faceted search [HEE’02] and NASA’s GCMD [GTP’99] are a source of inspiration. In our solution, some ideas are adopted such as enabling the user to view the intermediate result at any time. In case of EHRs, we have to deal with a large amount of multivariate data. Rao and Card show with the Table Lens [RC94] that tables are well suited for this purpose, and how focus + context can be applied on them for convenience. In this paper we again show, in this case by using extendable rows, that focus on table rows is a powerful tool.

We argue that EHRs can be seen as network data. ManyNets by Freire et al. [FPSG10] uses dynamic tables to present various properties of such networks. One thing that has to be taken into consideration is the existence of related entities with multivariate data, such as physicians and patients. Examples of systems that can deal with such related entities are PaperLens by Lee et al. [LCRB05] and Jigsaw by Stasko et al. [SGL08]. They do this by using interconnected tables and other views. With interconnected tables, we take a similar approach, but we use different techniques to show relations between entities.

3. Problem analysis

Kempenhaeghe has logged their physician’s medicine prescriptions for more than three decades, which resulted in a database containing hundreds of thousands of entries. This data can be used to discover certain trends in the use of medicines, but only if visualized properly.

3.1. Use cases

The overall aim was identified as obtaining insight into prescription behavior. To make this more specific, we performed several brainstorm sessions with medical experts to obtain four typical use cases that cover most of the problem domain:
U1 How does the usage of a medicine develop over time? In most cases, physicians can choose from several types of medicine to treat their patients. We want to know how their preference changes over time.

U2 Has the average medication load for patients increased in the last 20 years? Patients who receive too many (high doses of) anticonvulsants usually feel spiritless. Can we notice a trend in this medication load?

U3 Is a specific physician more cautious than his colleagues? Can we find the physicians who are (not) cautious with prescribing high doses (of a certain medicine, or to a specific group of patients)?

U4 Which medicine is preferred as first medicine? When a patient is diagnosed with epilepsy, no information about the patient’s medicine response is known. We want to know the first medicine of choice by each physician.

In these four use cases, we can identify four perspectives: the medicine’s (U1), patient’s (U2), and physician’s (U3, U4). The fourth perspective is that of the prescription, which appears in all use cases.

3.2. Prescriptions
To understand how prescriptions are established, we look at a basic procedure that is used by neurologists to manage epileptic seizures. A neurologist first chooses a medicine and an initial dose. Then, this dose is increased stepwise until the seizures are under control. If side-effects occur, the dose can (temporarily) be decreased or the medicine is replaced by another one. We define a single prescription as this dose/time function such as illustrated by Figure 1. For some patients it can be a real challenge to tune the doses of all medicines such that there is a good balance between desired effects and side effects. Therefore, there is a wild variety of prescription patterns, which is difficult to generalize.

Figure 1: A prescription can be seen as the dose as a function of the time. We can derive several metrics: the initial dose ($d_{init}$), the maximum dose ($d_{max}$), and the time it takes to reach the maximum dose ($t_{stm}$).

We observe three main entities in this procedure: physician, patient, and medicine. These are related to each other by means of prescriptions: a physician prescribes a medicine to a patient. Figure 2 shows that the data can hence be seen as a hypergraph where prescriptions are hyperedges that connect physicians, patients, and medicines. Since medicines are only prescribed for a time interval, the hyperedges exist only for these intervals. Besides relations, these entities can have many attributes. Patients have for instance a date of birth, gender, and syndromes. Hence we can look at these entities as being multivariate data.

Figure 2: A prescription can be seen as a hyperedge connecting a physician, a patient, and a medicine.

To be able to compare doses of different medicines, we use a medicine’s Defined Daily Dose (DDD), which is “the assumed average maintenance dose per day for a drug used for its main indication in adults” [WHO03]. By dividing all doses by the DDD of the prescribed medicine, we get normalized and comparable doses.

3.3. Requirements
From the problem analysis, we can derive the following major requirements:

R1 The system must show relations between entities, such as physician (or doctor) $d$ prescribed medicine $m$ to patient $p$, in a clear and understandable way.

R2 The system must show multivariate data that comes along with these entities such as the patient’s date of birth or a medicine’s DDD.

R3 The system must show the data from multiple perspectives, for one user may be interested in the prescriptions of a physician, while another is interested in the usage of a specific medicine.

R4 The system must enable the user to select entities or a time interval of interest and get relevant statistics about this selection.

R5 The system must facilitate comparisons between these selections.

R6 The system must be easy to use by neurologists and other medical experts.

4. The Three Table View
In the previous section we discussed that the data can be seen as a hypergraph, so the obvious solution to visualize this data would be to just draw a graph like in Figure 2. Such a graph will however quickly become unreadable once the number of prescriptions increases. To resolve this violation of R1, we can present the data in a more structured way by abstracting from individual prescriptions. One solution is to
use three lists: physicians, patients, and medicines; and to connect pairs of items with edges if there are prescriptions that match these items. This results in a 3-partite graph as shown in Figure 3.

While now having a structured representation of the relations, the graph is not very scalable and does not provide a nice way to show the multivariate data that comes along with the vertices. There is however a generic, simple, and well-accepted solution for displaying large amounts of multivariate data (R2), namely tables. This idea is for instance used in Jigsaw’s List View [SGL08], where relations between items in tables are shown as lines connecting rows in the tables. In order not to clutter the picture, these lines are only shown for selected rows of the tables. Figure 3 shows how this can be applied for our prescription data.

![Figure 3: A comparison between (a) the 3-partite graph and (b) Jigsaw’s List View, visualizing the same sample data. In the List View, physician 1, patient 1, medicine 1, and medicine 2 are selected simultaneously.](image_url)

4.1. The Row Relation Glyph

The example shows that lines are not very practical for our case, since they may cross the patients table. We present a novel method that solves this problem and results in a cleaner interface. The idea is to show the connections between items by marking the endpoints of relations with glyphs. Suppose that we select a physician who treats patient p1, then we only draw a glyph in front of p1 indicating that there is a relation between the selected physician and p1. In the case of a patient, this Row Relation Glyph (RRG) consists of two boxes: one showing the relation to the selected physician(s) and one showing the relation to the selected medicine(s). In the other two tables, this is done in a similar way as illustrated in Figure 4.

To emphasize the relations between various components, each table has its own color that is also used to highlight selected items to draw and their corresponding RRGs. In case multiple items in the same table are selected, medicines m1...ma for instance, there is a chance that there exist physicians or patients who prescribe/receive only a subset of these. To indicate such a weaker relation, the height of the box is proportional to the fraction between the number of relations and the selection size. To ensure that such a box remains visible at low proportions, there is a fixed minimal height. There is also a stronger type of relation, denoted by a thick border around the RRG. We speak of a strong relation for row x when each combination of selected rows in the other tables appears in a prescription together with x. This is for example the case for a patient who gets all selected medicines prescribed by all selected physicians.

![Figure 4: The same tables as in Figure 3b, but the relations are denoted with RRGs. Selected rows are highlighted in the table’s color (blue, green, and orange), and the RRGs are shown in corresponding colors. (d1) A physician who prescribed both selected medicines to the selected patient, indicated by the green and orange boxes in front; (p1) A patient who received m1 and m2 from d1; (m1, m2) Medicines that d1 prescribed to p1; (d2) A physician who prescribed only one of the selected medicines, indicated by a small orange box; (p2) A patient who received m1 and m2, but not both from d1, indicated by a missing thick border around the two boxes.](image_url)

4.2. Configuration

The three tables together with the RRGs form the basis of the interactive visualization. The tables support basic functionality like scroll, select, sort, and show the RRGs and other multivariate data like the patient’s age and gender. If one sorts on the RRG column, rows with the highest connectivity (largest boxes) traverse to the top. A timeline is used to select only prescriptions that overlap with the selected time interval. The content of the tables is automatically updated when the interval is changed. The tables together with the timeline are called the Three Table View (3TV). Figure 5 shows the basic configuration of the 3TV, how selections are made, and how RRGs are used. Figure 6 depicts a more advanced configuration. The accompanying video shows an interactive session with the system.

4.3. Augmenting the 3TV

So far, we have a basic visualization that at least supports Requirement 1 and 2. In order to view the data from multiple perspectives (R3) and present (R4) and compare (R5)
detailed information about a selection, we have augmented the 3TV as follows.

**Special cells.** To obtain an overview in distribution characteristics, such as the number of patients per physician, we added distribution columns, as depicted in Figure 5b. A cell in such a column depicts the number of relations of an item with items in another table, or in graph terminology, the out-degree of nodes to nodes with another color. Bars are used to visually emphasize the magnitude of the numbers in the columns. Besides showing the distribution, it also enables quick querying of for instance the number of patients that were given a certain medicine.

In a similar way, we defined other special cell types like **Sparkline Cells**, that show how the number of prescriptions associated with a row evolves over time. Since space is limited in a table cell, we use a Sparkline [Tuf06] of which the time axis corresponds to the selected interval in the timeline.

**Row grouping.** In many cases the experts want information about a group of physicians, patients, or medicines, that are in some way similar. The 3TV enables the users to group table rows based on their column values, which is illustrated in the following example. Suppose that (1) we want to find the most commonly prescribed medicine for patients with a symptomatic localization-related syndrome, and (2) we have information about whether or not patients have this syndrome stored in a column. We can simply group based on this column, after which we get two groups: patients with this syndrome, and patients for which this syndrome is not (yet) determined, see Figure 6. These groups behave the same in the table as if they were single patients. We can still select them, but the columns now contain aggregated information. Additionally, it is possible to group on multiple columns, for example to select all males with this syndrome.

Because data in each column has its own meaning, aggregations are column-specific. The gender of patients in a group is for instance displayed as a histogram. Their ages can also be shown as a histogram, but also as a box plot. The columns for physician, patient, and medicine names display the number of items in a group, see Figure 6a. For RRGs we take per box the average height so that the connectivity with respect to a selected group remains visible, see Figure 6.

**Row extensions.** In many cases we have more information about table items than can be displayed in a row. The 3TV provides an extend/collapse mechanism for each table row to show and hide additional entity-specific charts, which can be alternated by buttons in the top-right corner. Figure 5d shows how an extended row is used to inspect a patient’s individual prescriptions. Medicine extensions (Figure 6d) provide charts about its prevalence (number of patients using this medicine), incidence (number of patients who start using this medicine), and retention (average time that patients use this medicine). Medicine extensions show basic statistics over time. Charts can relatively easily be added to each of the three extensions to suit specific requirements.

**Table context menu.** The user can customize the table by means of a context menu that is attached to the table’s header. It provides the user the option to group rows and toggle columns to show only relevant information.

---

**Figure 5:** Screenshot of the 3TV showing its tables with physicians, patients, and medicines: (a) A selected physician is highlighted with a blue color; (b) Additional columns show distribution characteristics such as the number of patients treated by a physician; (c) A legend explains in natural language how the RRGs have to be interpreted, and shows between brackets how many of these exist in the corresponding table; (d) Row extensions can be used to obtain detailed information about a specific item, in this case the patient’s medication history; (e) The selection on the timeline indicates the interval from which the data is taken; (f) Grayed out rows show physicians, patient, or medicines that have no prescriptions in the selected time window.
Figure 6: Screenshot of the 3TV with a different configuration compared to Figure 5: (a) Patients are grouped by the column that indicates whether they have a symptomatic localization-related syndrome (classified as ILAE 12), this results in two groups; (b) By sorting on the RRG column, we find out that Carbamazepine is the most commonly prescribed medicine for patients with this syndrome; (c) Sparklines give a global indication of how the number of prescriptions associated with each row develops over time; (d) An extension showing a medicine’s prevalence (the number of patients using it) over time, the green graph represents the selected patients. In case of Vigabatrin we spot a sudden decrease of patients just before the year 2000, due to a discovery that this medicine can cause blindness. Other views on medicines can be selected by pressing the buttons in the top right corner of the extension.

5. The Prescription View

The 3TV provides a detailed view on physicians, patients, and medicines. An important aspect of the data is that these three entities are connected by prescriptions. The 3TV does not provide that much information about prescriptions themselves. In order to do so, we added an additional view that is specifically designed to display prescription statistics.

5.1. Prescription metrics

We defined prescriptions as the dose as a function of time, the next step is to identify interesting metrics on these. Since we want to reveal the prescription behavior of physicians, we need to extract and show information about important decisions. This can for example be the initially prescribed dose as illustrated in Figure 1. For this project we derived 10 metrics $f_1(r) \ldots f_{10}(r)$ for a prescription $r$ divided into 3 categories:

- **Dose metrics.** Regarding the dose, we are interested in three metrics. To begin with, a physician determines an initial dose to find out how a patient responds to the medicine. This dose is increased until the desired or side effects occur, this is the maximum dose. We can also compute the average dose over the time interval it was prescribed.

- **Temporal metrics.** Besides the dose, also the duration of the prescription plays an important role. A measure for how quickly the physician increases the dose, we measure the time it takes to reach the maximum dose $t_{\text{start-to-max}}$. Similarly, one may be interested how quickly a physician wants to decrease at the end of the prescription interval $t_{\text{max-to-end}}$.

**Context metrics.** The context in which a prescription is done is also important information. It often matters which combination of medicines is prescribed because they may interfere. We are therefore interested in co-medicines: the set of medicines prescribed simultaneously to this patient. Besides co-occurrence, there are successive relations: the set of medicines prescribed strictly before and after this prescription $\text{preceding/succeeding medicines}$. Finally, we can observe whether this was the first medicine prescribed to the patient, i.e., there are no prescriptions for this patient that started earlier.

5.2. Prescription set visualization

There are thousands of prescriptions, which we do not want to investigate all individually. Instead we are interested in all prescriptions that satisfy a certain query, for example: all prescriptions by physician $d$ with medicine Carbamazepine. To this end we calculate aggregated metrics $f(R)$, for prescriptions $r$ in the set $R$ of prescriptions that match the query. We use histograms to show the frequency distribution for each metric, where the metric is on the $x$-axis and the number of prescriptions on the $y$-axis.
Figure 7: Screenshot of two tabs of the Prescription View, comparing Physician 9 with his colleagues. The left tab shows the dose distribution for Carbamazepine, while the right tab shows how other medicines are combined with Lamotrigine. (a) The size of the primary and secondary selection; (b) Metrics are grouped by tabs; (c) Histogram showing that Physician 9 starts with relatively low doses (either because of a different patient population, or a real difference in individual behavior); (d) A legend indicates the queries for both selections; (e) Histogram showing that Pipamperon is usually not combined with Lamotrigine, but Physician 9 does this for 30% of the prescriptions. (f) Radio buttons to select whether the primary and/or secondary selection is currently adjusted by the 3TV.

5.3. Configuration

The user creates the query for selecting relevant prescriptions via selection of items in the 3TV. If both physician d and Carbamazepine are selected in the 3TV, then the system automatically selects the prescriptions by physician d with Carbamazepine. More precisely, the system selects all prescriptions that connect one of the selected physicians, one of the selected patients, and one of the selected medicines that also overlap with the selected time interval. For convenience, tables without selected items do however not constrain a prescription’s selection: a table with the empty selection in this case equals the selection of all rows.

Per metric, one histogram is shown in a separate Prescription View (PV). The PV has one tab for each of the three metric categories, see Figure 7. For each selection, the histograms give the characteristics of the corresponding set of prescriptions. According to Requirement 5, we need to facilitate comparisons between selected items. Therefore we introduce a primary and secondary selection. To make comparisons between these sets more convenient, we draw them in the same histogram instead of separate histograms. The reference set is drawn in the background as gray bars, the primary set as the foreground in red. Since the two selections may contain a different number of prescriptions, the user has the option to normalize the height of the histogram bars by dividing by the set size. As a result, the y-axis then displays percentages.

6. User evaluation

To test the usability (R6), we built a prototype and organized a user evaluation session of about 80 minutes in which 6 neurologists participated. The participants, three female and three male, had ages between 30 and 60 years.

6.1. Setup

The evaluation consisted of three parts: (1) a short 15 minute demo to present how the system works, (2) a set of pre-defined assignments to find out how well the participants are able to use the system, and (3) a questionnaire to ask
them for their opinion on the usefulness of the system, and to evaluate the difficulty of the various components. During the demo, the participants were encouraged to repeat all the actions done by the instructor. The assignments were 5 realistically formulated questions with some sub-questions, where the overall complexity was gradually increased:

1. How many patients have been treated with Lacosamide, and how many neurologists prescribed it?
2. Name the most prescribed medicines between the years 2000 and 2010.
3. Are neurologists more cautious with Lamotrigine than with Clobazam?
4. Does the prescribed dose for children with an intellectual disability differ from the general population?
5. What is the most popular first choice medicine for patients with a symptomatic localization-related syndrome?

The participants worked simultaneously and in the same room on the predefined assignments. After completion, they were encouraged to think of other questions and answer them using the system. During the experiment, questions to the instructor were allowed to prevent the participants from getting stuck.

Some features were not tested during the evaluation session. The amount of time that neurologists are available for such an experiment was simply not enough for a thorough evaluation of all aspects of the system. We therefore decided to present the most generic features such as the RRGs, the selection mechanisms of the tables, row grouping, the timeline, and the PV. We did not test the Row Extensions, Sparkline cells, and the secondary selection mechanism of the PV, since we considered these as useful for more advanced users.

6.2. During the session

During the short demo, we had the impression that the audience almost instantly understood how the 3TV and PV function. This feeling was strengthened by requests for features that had not yet been demonstrated. An example of such a request that popped up was the secondary selection mechanism of the PV, just after we showed that it is possible to visualize the prescriptions of each individual physician: “Is it also possible to compare a single physician with the average of all physicians?” Indeed, this is possible using a secondary selection as described in Section 5. We were positively surprised by these questions, especially because we considered these as a strong signal that the neurologists understood the principles of the visualization and could imagine how to use such a tool to answer their questions.

6.3. Results

The results of the predefined assignments were clear: All participants were able to complete them correctly without notable trouble. Completion times per task were not recorded, but they were all comfortably finished within one hour. Quotes from the neurologists support the usefulness of this project: “When can we start using this tool?”, “The project surprised me in a pleasant way. We certainly have to go further with this.”, and “A beautiful project, especially for audits.”

The neurologists commented that they needed more quantitative, textual information, for instance on the exact number of children that use a certain medicine. In Section 7.4, we describe how this can be solved by using hoverboxes. Furthermore, one participant wanted to filter out adults. This appears to be difficult, since age develops over time, i.e., if you make a range selection on the timeline, the patient’s age becomes undefined. Finally, another participant wanted to investigate side-effects, but this type of data was not taken into account in this project. In Section 7.3, we explain how such data can be incorporated. Besides these issues, the 3TV and PV are generally well accepted by the target audience. The results of the questionnaire [vdC13] show that all participants found it easy to use the tables and the timeline, but they are more divided on using row grouping and the PV. Some found it easy, some did not. They also agreed (with at least 4 points on a scale from 1 to 5) that they would like to use the system in the future, as well as that the used methods enabled them to obtain new insights into their prescription behavior.

7. Discussion

The 3TV and PV were developed to provide insight into medicine prescription data. However, they appear to be generic, scalable, and extensible such that they can be used for many other applications.

7.1. Other applications

In case data contains three entities A, B, and C that are in some way related by a concept D, then the 3TV and PV are likely to be applicable. We illustrate by means of two examples that the pattern A gives B to C, or in general, subject acts on direct object to affect an indirect object, is ubiquitous.

Suppose that we have to visualize insurance policy data. We could for example replace our three tables by a company, customer, and insurances table. Each company sells a set of predefined insurances, such as health or home insurances. When customers sign up for an insurance, they receive an insurance policy. This policy then connects the company, customer, and insurance, and can hence be seen as an equivalent of the prescription. An insurance policy has similar properties as a prescription, like temporal existence and a varying price. Using the 3TV and PV (Policy View), we can answer questions like: “What are popular insurances for people younger than 25 years?”.

Suppose that we have to visualize airline data. We could
then use three tables: airlines, passengers, and flights. A flight is a link between two airports, which is flown by several airlines from time to time. A passenger can make a journey by reserving a seat on a flight that is flown by an airliner. Hence the journey is the prescription equivalent. The 3TV can for instance help us to find the passengers who travelled between Amsterdam and New York in the year 2013.

7.2. Scalability
In case the data has a slightly different structure or size, we expect the 3TV and PV to remain applicable. If we for example have additional entities, then we could add more tables. As a result, the RRG will contain one extra box per added table. Screen space and cognitive workload are however limiting factors in the number of tables that can be added.

For some applications including medicine prescriptions, the number of attributes per entity can become very large and result in a large number of columns. Since users probably do not have to view all columns at once, the column toggle mechanism should be sufficient to keep the display clean.

In our project, we used a database containing about 60,000 prescriptions having an average of 5 dose changes (300,000 changes in total). Under the assumption that a prescription has a bounded number of dose changes, i.e., the dose is not a continuous function of the time, we can derive [vdC13] that the running time is linear in the number of prescriptions. So the 3TV and PV are well scalable in the size of the data.

7.3. Additional data
Although prescription data on its own can provide new insights compared to other medical data sets, the visualization of additional data can be combined with the presented methods. We already showed how multivariate data of different types can be used in the 3TV. When looking at the literature, the visualization of electronic health records mostly involves event based data. In case of epilepsy, one could for example think of logged epileptic seizures, which is a list of tuples (patient, time stamp). Such data can be inserted in several ways: (1) By displaying aggregated data in the 3TV, e.g., the number of seizures per patient in the selected time interval; (2) By showing events per (group of) patient(s) in a Row Extension; Or (3) by using seizure frequency as a metric in the PV, for instance to show how physicians, patients, and medicines are correlated with seizures. The same can be done with other events such as headache, insomnia, heart attack, and stroke.

7.4. Future work
Hoverboxes. The physicians in the user evaluation noted that they were unable to read exact numbers from RRGs. Since there is not enough space in the interface to show all numbers textually, we intend to add hoverboxes to show additional and accurate information. Hoverboxes can also make the system more intuitive by providing information in full sentences. A hoverbox for a RRG could then for example say: “75% of the patients in this group use both Valproic Acid and Carbamazepine, and 25% of them are treated by doctor Walden.”

Outlier detection. While it is possible to compare physicians with each other, the system currently does not enable the user to quickly find “interesting” physicians for comparison, i.e., outliers. A possible solution is to compute prescription behavior metrics per physician such as those in the PV, and present them as a new column of the physicians table. No matter how this is visualized, it is important to present the statistical significance of the differences to confirm that someone who looks like an outlier is indeed one.

Visualizing treatment effectiveness. Due to the absence of detailed epileptic seizure and side effect reports, we are unable to visualize the results of medicine prescription behavior, and therefore its effectiveness. Techniques for seizure registration are becoming more accessible these days, making the options presented in Section 7.3 relevant.

Sentinel events. In Section 2 we discussed that the use of sentinel events can reveal correlations between events. The ability to select the prescription of a certain medicine as a sentinel event would be an interesting feature. As a result, all prescriptions per patient will be aligned by this event, turning the timeline into a relative time scale.

8. Conclusions
We presented a system to interactively visualize medicine prescriptions in order to give physicians insight into their prescription behavior. An evaluation session with a prototype demonstrated that our techniques are well accepted by physicians, our target audience. They were able to quickly understand the concepts of the Three Table View and Prescription View. After a short introduction, physicians are able to answer quite complex questions about their prescription data. The evaluation session also revealed some potential improvements, such as the use of hoverboxes to present more detailed information.

We discovered that the (seemingly simple) structure of prescription log data is difficult to visualize using standard techniques. We approached this structure as being a hypergraph, which allowed us to maintain a high level of abstraction. This way, we can also use the presented techniques to visualize data from other domains but with a similar structure. The Three Table View can be generalized further by using a dynamic configuration with an arbitrary number of tables. We see two major directions for future work in this area: (1) generalization of hypergraph visualization methods, and (2) integration of domain specific requirements in the approach presented here.
References

[Ch01] Chittaro L.: Information visualization and its application to medicine. Artificial Intelligence in Medicine 22 (May 2001), 81–88. 2


© 2014 The Author(s)