

VizADR - Visualizing Adverse Drug Reactions

Alexander Pfundner*
0828326

Jasmin Kuric†
0703684

ABSTRACT

VizADR has the purpose to help medical staff to get an overview over adverse drug reactions on patients. The user is able to identify drug reactions in combination with a certain disease for patients in regard to their gender and age. To achieve this goal we have studied and designed several visualization methodologies with the ambition to find the concept that fits best for this task. Finally we have developed VizADR, with which we think is a proper solution.

1 INTRODUCTION

This document was written during the course "050065 VU Visualisierung" at the University of Vienna during summer semester 2014. It describes the project "VizADR - Visual Adverse Drug Reactions" we have developed, in order to achieve our course grade.

VizADR is available online under following resource [8].

2 MOTIVATION

During the "Visualisierung" course this semester we have acquired a wide range of knowledge that helped us develop a software tool that is able to identify adverse drug reactions of specific drugs. In practice this tool should help medical staff and medical scientists to explore type and frequency of adverse drug reactions regarding a specific drug.

2.1 Problem

A general task that a physician has to do when he gives medication to a patient is to calculate the risk of adverse drug reaction based on different circumstances regarding the patient (e.g. age, gender, ...). To help the physician with his choice and to identify common adverse reactions with possible prescribed drugs, we developed VizADR.

2.2 Users and their Tasks

There are mainly two user groups with specific use cases that we want to cover with VizADR.

The first user-group we target are physicians. VizADR should help a doctor making the choice of picking the most appropriate drug for his/her patient easier. By choosing the applicable disease class the user can explore which kind of drugs have been dispensed and which type (gender, age) of patients had adverse reactions and how likely it is that those reactions will occur on his/her patient.

The second user-group we are targeting are medical scientists from different fields of research. They should be able to identify which kind of drugs lead to what kind of adverse drug reactions. With this information, medical scientists can identify proper research fields that need further development, because the existing drugs cause too much side effects.

*e-mail: alepfu@gmail.com

†e-mail: jasmin.kuric@gmail.com

2.3 Data

The data set used for this project was provided by "Health Canada", which is the Canadian Federal department responsible for health affairs. [5]

Unfortunately Health Canada does not support an access via an interface to read the data simultaneously. But does allow a full extract of the database into another database, provided by the project members. An instruction manual is provided as well, by Health Canada. The database is not maintained up-to-date on a regularly basis, but there are major bulk-updates on irregular intervals.

3 RELATED WORK

3.1 Previous designs

In order to get to our final design choice, we tried several other designs with paper prototypes to find a solution for the visualization problem.

3.1.1 Heat maps

One of our first design choices included heat maps. Figure 1 shows a prototype of an overview and a detailed view of drug reactions which can be filtered by demographic items. A comparison between two demographic groups is shown in Figure 2.

This design choice would have fulfilled our needs, but we declined it because of aesthetic values.

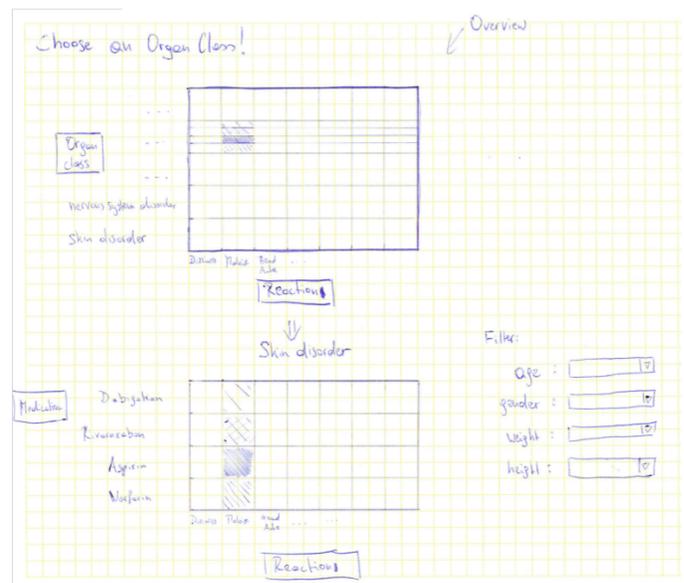


Figure 1: Paper prototype - Heat map: Overview and detail of VizADR

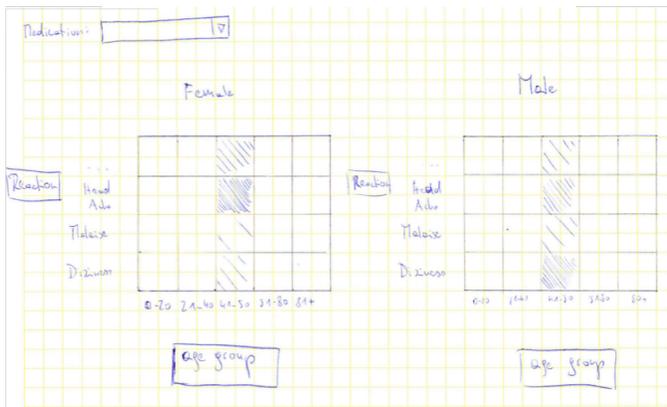


Figure 2: Paper prototype - Heat map: Detailed view with gender comparison.

3.1.2 Sunburst

Figure 3 shows another design idea we have came up with. At beginning we were kind of enthusiastic about this idea, but in further discussions we realized it cannot fulfil our requirements. With a high number of drugs or diseases, this design choice becomes quite confusing for the user. So we rejected this design.

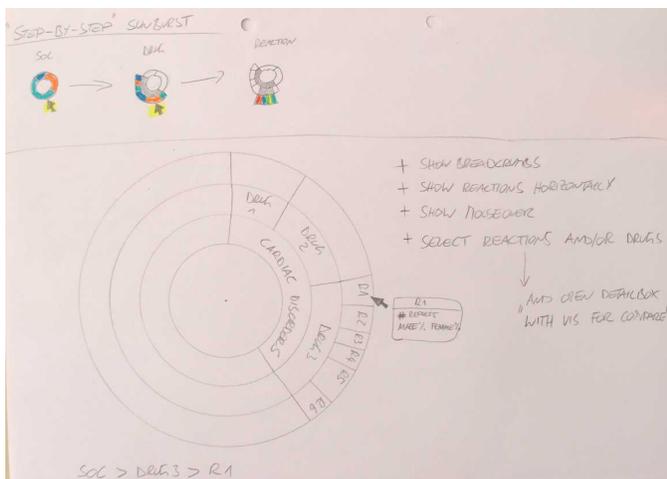


Figure 3: Paper prototype - Sunburst: Interactive view of adverse drug reactions.

3.1.3 Bar charts

Our third alternative design was including bar charts. While we took elements shown in Figure 4 in our final design, it has become apparent that the other ideas in Figure 4 and Figure 5 were not very user oriented and also maybe too "ordinary" in our opinion.

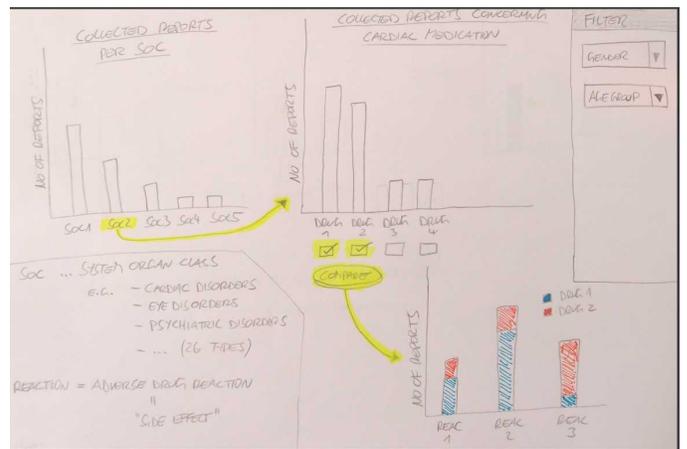


Figure 4: Paper prototype - Bar chart: Interactive overview of adverse drug reactions.

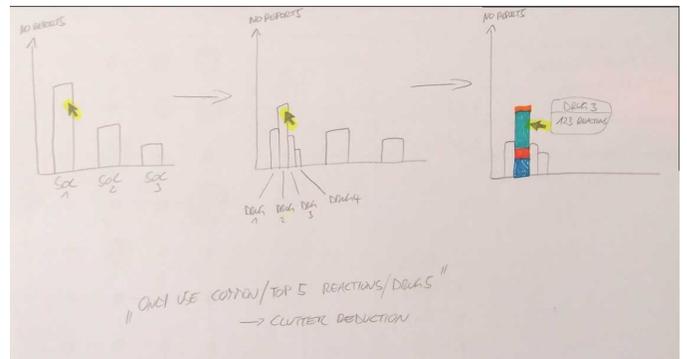


Figure 5: Paper prototype - Bar chart: Interactive detailed view of adverse drug reactions.

3.2 Existing solutions

During our project we looked up already existing solutions for this problem and found "Data Visualization Speeds Review of Potential Adverse Drug Events in Patients on Multiple Medications" [10], "Advanced Visual Analytics Interfaces for Adverse Drug Event Detection" [11], "A Novel Visualization Tool for Evaluating Medication Side-Effects in Multi-drug Regimens" [9].

Especially the pixel-based visualization [11] expanded our thoughts regarding potential visualization possibilities.

4 APPROACH

In order to achieve our visualization we had to identify our goal. As mentioned in Section 2.2 we first defined the target users and their tasks to make the use case clear.

Next step was to decide which kind of design we want to implement, which is discussed in the subsection below.

And the last step was the decision which technology respectively framework we want to use. This decision will be elucidated in section 5.

4.1 Visualization design

We implemented different levels of visualization for VizADR.

First of all there is an overview (as seen in Figure 6) of all the disease classes. The size of the bar chart accords to the number

of reports there are available for this particular disease class. By hovering over one stack with the mouse there pops up a window with information about the disease class, as seen in Figure 7. This information contains the drug reaction distribution and a list of the top five drugs, which are used in this disease class.

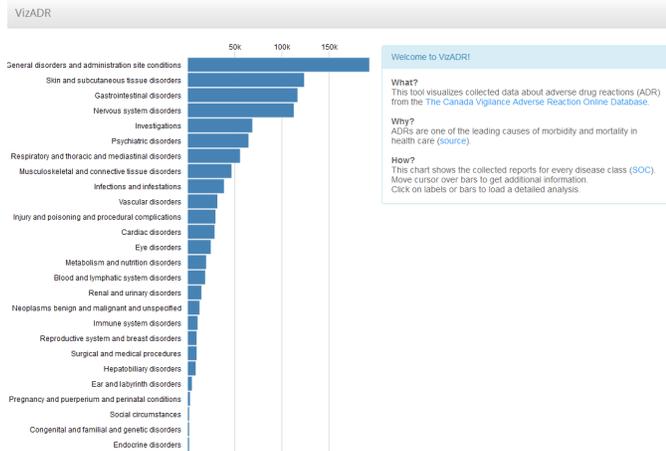


Figure 6: VizADR Level 1 - Overview of the disease classes with amount of reports.

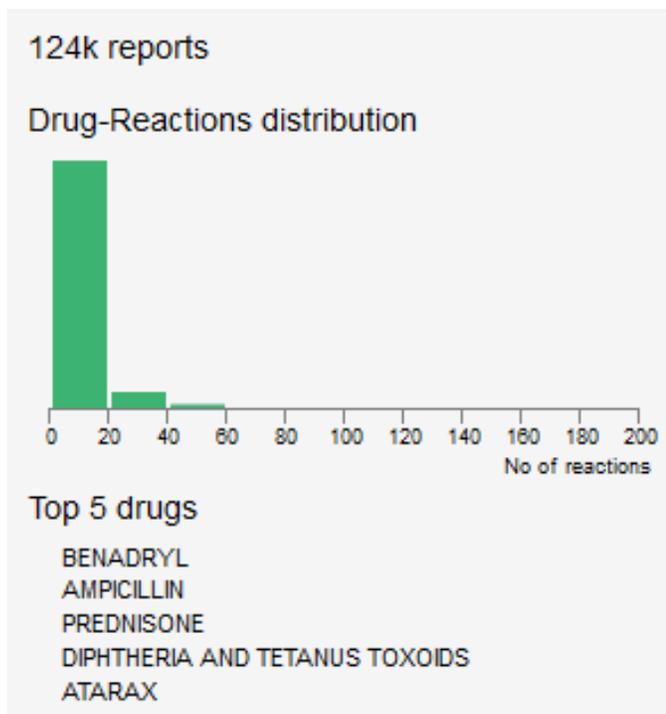


Figure 7: VizADR Level 1 - Popup window with information regarding a disease class.

By clicking on one bar stack (as seen in Figure 6), you arrive in the second level which is a scatter plot with informations about the possible drugs for the chosen disease class and the drug reactions this drug can cause (as seen in Figure 8). The size of one point tells us how likely it is possible that this drug causes this specific adverse

reaction. The larger the point, the bigger the chance that this drug will create this side effect.

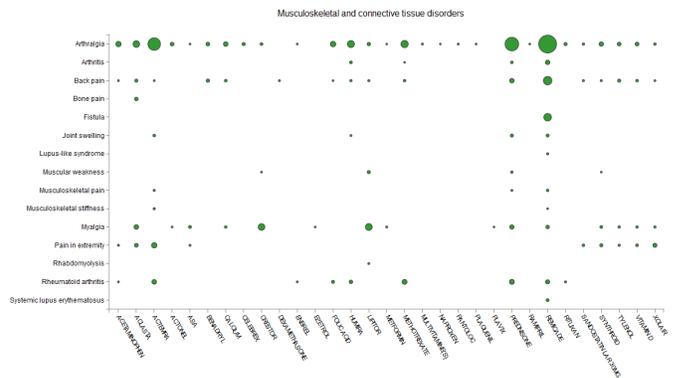


Figure 8: VizADR Level 2 - Scatterplot diagram of the relation between drug and adverse reaction.

When you then click on a point in the scatter plot, you get to the third level (as seen in Figure 9). This third level contains detailed information about the drug combined with the drug reaction. First there are text based informations like number of reports or seriousness of the adverse reaction. By hovering over a coloured bar stack you get the percent-information of the coloured field in a tool-tip. We describe this diagram as a sort of mix between a mosaic plot and a stacked bar chart.

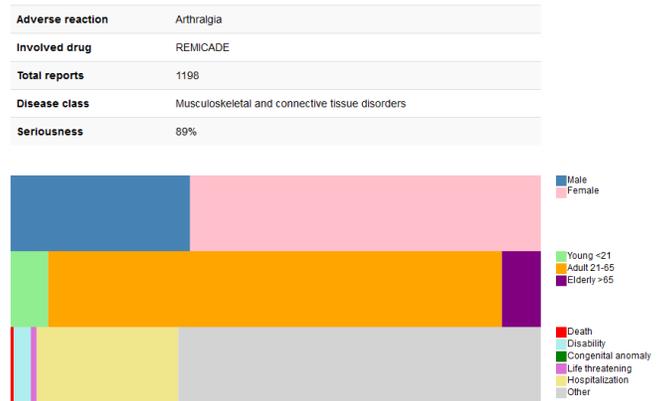


Figure 9: VizADR Level 3 - Detailed diagram of the impact of the adverse reaction on the patients, divided by gender, age and outcome

4.2 Reason for design choices

We have discussed about which design we should use for quite a time. This section explains our design choices.

4.2.1 First level

Figure 6 shows our first view or rather first level of the application. So the question is, why does it look that way? First of all, we want to identify how it is actually looking: It's a diagram with vertical bar charts, without showing the exact value of one bar chart. The bar charts are sorted by their size.

So why did we design it this way? We decided to use bar charts [12] because they are an easy way to visualize a quantity values. The bar value represents the given amount of drug reports that the database provides us with. We sorted the bar charts by size, because the bigger items are more useful for the user. The more reports there

are, the more meaningful are the further statistics. This implementation is based on the concept of "reduce"[13] and "overview" [14] by Tamara Munzner. By hovering over a bar chart, a small window with a histogram will pop-up (as seen in Figure 7). We used this diagram to make a easier crossing from level 1 to level 2 (as seen in Figure 8). This diagram shows us the amount drugs with similar number of drug-reactions.

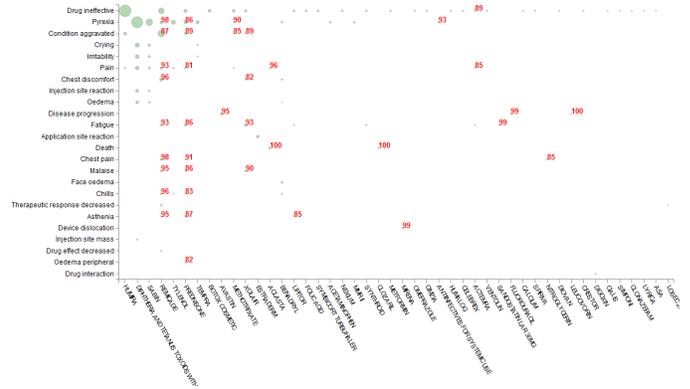


Figure 11: VizADR Level 2 - Scatter plot with seriousness

4.2.2 Second level

Our second level of visualization shows us a scatter plot [15] (as seen in Figure 8) with the information of drug, drug-reaction and the amount of reports which have been reported in combination of a certain drug with a certain drug-reaction. So it is possible to visualize three dimensions with this type of diagram.

It is also possible to sort the axis alphabetically or by amount of reports as shown in Figure 10. There is also the option to differ the data based on gender or age groups. Furthermore you can enable the view for seriousness, which is indicator for how strong the drug reaction took place (as seen in Figure 11). The higher the number, the more serious the drug reaction was. It is possible to filter the seriousness from a range between 1 to 100.

We chose a scatter plot because we can easily implement three dimensions of the data: drug, drug-reaction and report amount. The difference to an usual scatter plot is, that our two axes are categorical and not based on ratio data. That's why the several scatter plots cannot overlap each other. [16] With the option to sort the axes by Report count (as seen in Figure 10) the user can group drug/drug-reaction combinations on the screen and identify them faster. Furthermore we choose to make the option available to differ between gender or age groups, because we wanted to give the user the ability to identify his target group faster. And last we made the seriousness visualization possible, to make the user see the overview of all drug/drug-reaction combinations with the seriousness, and without the need of clicking into a single combination to see the seriousness only there.



Figure 10: VizADR Level 2 - Adjustments for scatter plot

4.2.3 Third level

The third level shows text details regarding the drug/drug-reaction combination. And also a diagram (as seen in Figure 9) with the disposition of the reports regarding gender, age group and outcome.

This is the first view, where we actually see some details. So the user is forced to click himself from level one (which represents the big overview) to level two (which shows filtered data) to level three where he actually gets the details. We choose this procedure to make sure the user can see aggregated parts of all of the data, and not only his desired detail. So he eventually can spot something he actually didn't even search for. The number of interactions stays still the same, with the overview in level one and two or without it.

We call the diagram we used in the third level a mix between stacked bars and a mosaic plot. We wanted to represent visually the distribution of the reports at first sight, without need to read and interpret the numbers. When you move with the mouse over a mosaic, you can see the exact number that the mosaic represents in a tooltip.

5 IMPLEMENTATION

5.1 System requirements

VizADR is a Tool which is based on a http-Server (e.g. Apache Web server). You can access VizADR by an Internet browser of your choice which is able to deal with Java Script [3], CSS and HTML.

On the server side the D3.js (Version 3.4.8) framework [2], the jQuery framework (Version 2.1.1)[4] and the Bootstrap framework (Version 3.1.1) [1] are required.

5.2 Implementation challenges

There have been several challenges we have dealt with. Some functions we wished to implement in first place, but weren't be able to do so. Those are documented in the subsequent subsections below this paragraph.

We had also a few challenges during the coding, but those have been solved pretty fast due to the big D3 community and the existing documentation.

5.2.1 Omitted feature: BMI differences in the scatter plot

Since the dataset includes columns for the patient's body height and weight, an additional filter displaying labels with the average Body-Mass-Index on the dots (like the seriousness labels) would be possible.

Unfortunately, these columns have a very low data quality, in terms of consistency and completeness, so a visualization brings none or little benefit.

5.2.2 Omitted feature: Parallel sets in detail view

For the detailed view (as seen in Figure 9) it would have been well suited to implement a parallel sets [7] representation between the gender, age group and outcome bars.

We analyzed this feature by giving a closer look to an example implementation (as seen in Figure 12). In the end we decided to omit this feature for different reasons which would have led to a very complex custom implementation. The example uses a custom built D3 library in combination with another JavaScript library we don't have experience with.

Furthermore the example's source code is very complex and has little documentation. Additionally the example uses a specific input file format, applied to our dataset and implementation this would lead to a complex extraction process, resulting in huge files.

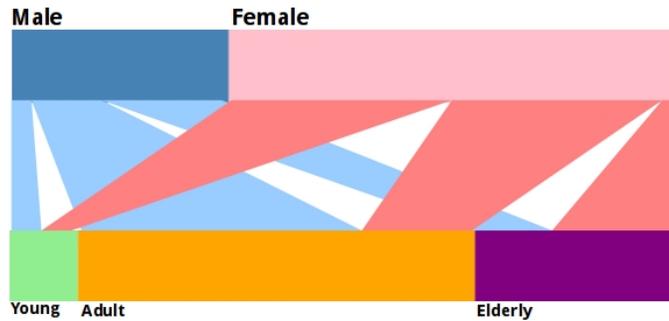


Figure 12: Possible example of parallel sets in VizADR

6 RESULTS

6.1 Use case scenario

6.1.1 Use case one: Find medication for a patient

Fictional user: Dr. Angelika Schmidt (physician, 39 years, female, family doctor) Angelika Schmidt has a male patient who suffers from gastrointestinal disorder. She has different treatment drugs she could prescribe him. Her patient is 60 years old. She wants to know which kind of drug she should give her patient. She wants to minimize the possible adverse drug reactions.

To find the best treatment drug she first selects the disease class in the overview (as seen in Figure 13 in the red rectangle).

After that she gets to the level two view (as seen in Figure 14, where she has an overview of all drug reactions and possible treatment drugs. She sorts the scatter plot axes by report count, so she can identify the drugs with the most and the least adverse reactions (as seen in Figure 14). She also wants to see how serious the adverse reaction were, so she clicks on "Show seriousness" (as seen in Figure 14). Now she can identify which drugs have the least drug reactions and can choose one of those on the right side, which are applicable to her patient.

After she clicked on the drug/drug-reaction combination that is possible, she sees how likely this adverse reaction will happen on her patient by seeing the gender, age and outcome distribution of this adverse reaction (as seen in Figure 15).

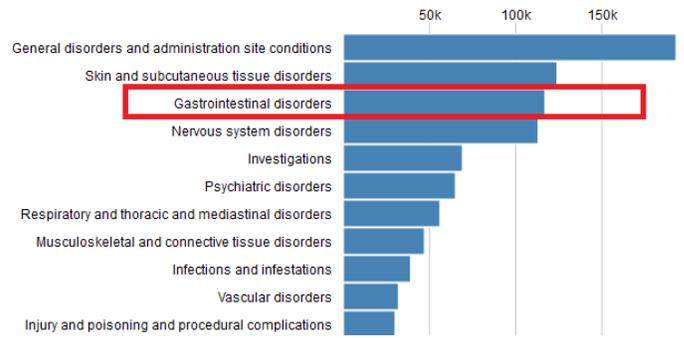


Figure 13: Use case one in VizADR - Level one view

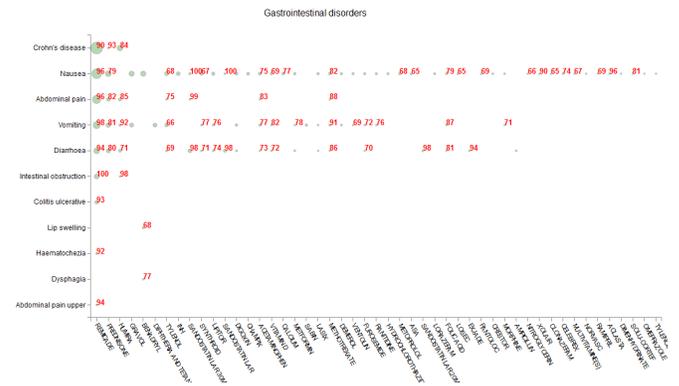


Figure 14: Use case one in VizADR - Level two view

Adverse reaction	Nausea
Involved drug	OMEPRAZOLE
Total reports	268
Disease class	Gastrointestinal disorders
Seriousness	50%

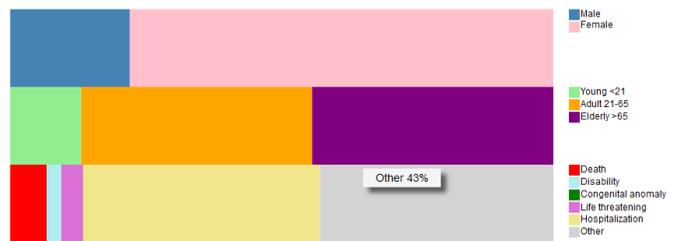


Figure 15: Use case one in VizADR - Level three view

6.1.2 Use case two: Discover new research fields regarding drugs

Fictional user: Univ.-Prof. Dr. Herbert Berger (pharmacologist, 36 years, male, medical scientist) Herbert Berger is doing research in the field of respiratory and thoracic and mediastinal disorders, his goal is to optimize medication plans for different groups of patients and to single out patient groups that might profit from new forms of medication. In the field of respiratory and thoracic and mediastinal disorders the available drugs often trigger adverse reactions, so for Herbert Berger's research it is important to identify those adverse reactions.

The first step is to choose his disease class (as seen in Figure 16 in the red rectangle). Then he orders the scatter plot by the report count and also shows the gender differences. Now he sees that the adverse reaction "Dyspnoea" is occurring on almost every possible drug. He also sees that almost every adverse reaction occur mostly on female patients (as seen in Figure 17). So Mr. Berger can compile the information that there are no adequate drugs for female patients, and "Dyspnoea" is a occurring adverse reaction in almost every drug treatment. So he can interpret the visualization in that way, that there is more research necessary for adequate drugs for female patients and there should be a development initiated to prevent "Dyspnoea". He also sees that "XOLAIR" causes almost every possible adverse reaction. So there should be research done to invent a adequate replacement-drug for "XOLAIR".

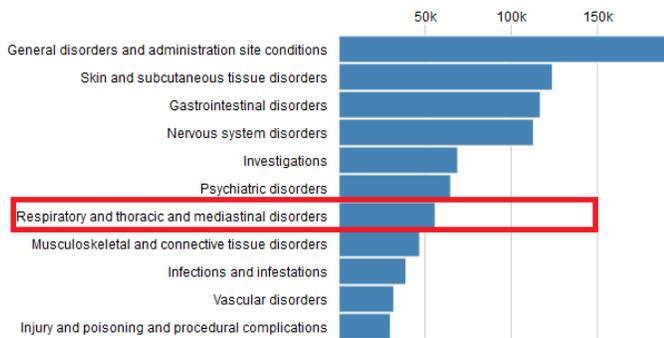


Figure 16: Use case two in VizADR - Level one view

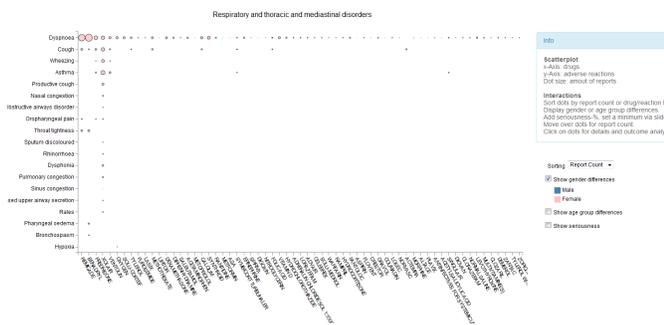


Figure 17: Use case two in VizADR - Level two view

6.2 Performance

Since we do not have any bottle necks on which we depend on (e.g. external database), VizADR performs pretty well. There are no delays during query's, because the data is already generated and aggregated previously into CSV files, which are stored on the same server as the actual application. Interaction with VizADR runs very fluid.

Of course this depends on the machine which the application is running on, as well as the client machine and the clients ability to perform Java script in a proper way, which usually shouldn't be a major issue.

One thing, we haven't tested during our project is the access of the application with a high amount of different clients, as it is usual for a frequently used web page. We considered it not the primary goal of this project to ensure the applicability in the world wide web.

6.3 Evaluation feedback

Due to the missing opportunity of evaluating VizADR by medical scientists or physicians we have asked other teams of the "Visualisierung" course to test our application and give as a feedback. The reactions were mainly positive. The break down in three levels has been mentioned as a pleasant way to view the data.

Too small font sizes, cropped labels in some diagrams have been criticized by the evaluation group.

During the development we got response from our course instructor. In order to that input we changed from pie charts to stacked bars, added the possibility to sort items on the axes, added the histogram on level one as a pop-up(as seen in Figure 7)

Although we would still need a feedback from a person of the medical profession, to have a sight from someone which could actually use the application on a regularly basis.

7 DISCUSSION

7.1 Strengths and weaknesses

7.1.1 Strengths

We consider our application as well fitted to our user groups (but still need a feedback from a medical person as mentioned in Section 6.3). It isn't really a training necessary to be able to use and understand the outcome of the application.

There is no big click depth, the user is forced to enter. With only two clicks you can get to the most detailed view of the application. Still there are some adjustments you can do in the second level (as seen in Figure 8), but you are not obliged to do them.

One of our most likely features is that VizADR has an immediate overview of the data. It is not required to choose e.g. between disease classes before you even see the data. The entering page is already a overview and a possibility to select data.

Thanks to D3, VizADR has some nice looking effects and appealing looks, which makes the whole tool very user-friendly.

7.1.2 Naming of drugs and reaction terms

In the medical field it is common to use taxonomies to classify terms, their relations and the hierarchy between them. In case of adverse reaction terms this leads to shallow definitions.

E.g. "Pain" or "Drug interaction" which, when used too often, lower the quality of the dataset and the resulting visualization. On the other side, drug names are set by the companies producing them, the resulting spectrum of names is very wide and leads to cryptic letter number combinations or names consisting of five ore more words.

Both issues have disturbing influences on the design and implementation of the visualization.

7.1.3 Visualizing age group differences as glyphs

In contrast to the two-dimensional gender differences visualization, the three-dimensional age group differences visualization would benefit from a different form of encoding.

Since gender is binary, we can simply calculate the difference between male and female report amounts and use the dot's size and colour for encoding. For three values, calculating a meaningful difference becomes impossible. Our course adviser gave us the tip to use glyph-based encoding for this scenario, he mentioned slides from the lecture concerning the visualization of tensor data (method of isotropic/anisotropic glyphs by Gordon Kindleman) where 3D-glyphs are used. Since we have no three dimensional rendering available in SVG, this matter evolved into the research question:

”How to encode multiple dimensions in a 2D glyph”. To give an outlook to this question, we worked out an idea where we encode the three age group dimensions by using shifts of irregular hexagons, to get a first impression see this [6] JSFiddle we created.

7.2 Lessons learned

First of all the research for a proper visualization gave us some kind of enlightenment. Through the investigation we discovered a whole new range of methodologies which could have been used for this case, which we didn’t really expect at the beginning. But how could we, there hasn’t ever been really a course or topic at high school or university which had engaged us to look deeper into this matter.

Another lesson we certainly learned, was to give a little more time for the design choices. Don’t be too quick on deciding which design should be implemented. Give yourself one or two days time to think about the design in your free time, it isn’t exceptional that you come up with an even better idea after some time of mental processing.

In addition to our whole project, we experienced D3. And we must say it is pretty impressive. The possibilities for design implementation are sheer never ending. The already developed visualizations in the D3 community gave us a lot of ideas and influenced our design process in a good way.

We have also discovered that the research potential for this kind of visualization is still existent, though we couldn’t find many scientific papers or articles for this visualization problem.

Knowing your dataset is a key issue in this specific project, but probably for every visualization problem. It’s necessary on one side to have a good knowledge of the actual topic or subject you want to visualize, and also to know visualization well. Every new visualization implies a new partial study of the topic field.

REFERENCES

- [1] Bootstrap framework. Available online at <http://getbootstrap.com/>; visited on June 1st 2014.
- [2] D3 framework. Available online at <http://d3js.org/>; visited on June 1st 2014.
- [3] Javascript programming language. Available online at <https://developer.mozilla.org/en-US/docs/Web/JavaScript>; visited on June 1st 2014.
- [4] jquery framework. Available online at <http://jquery.com/>; visited on June 1st 2014.
- [5] Medeffect canada - canada vigilance adverse reaction online database - health canada. Available online at <http://www.hc-sc.gc.ca/dhp-mps/medeff/databasdon/index-eng.php>; visited on April 4th 2014.
- [6] Multiple dimensions in a 2d glyph. Available online at <http://jsfiddle.net/alepfu/B8D5r/>; visited on June 20th 2014.
- [7] Parallel sets framework. Available online at <http://www.jasondavies.com/parallel-sets/>; visited on June 10th 2014.
- [8] Vizadr webapplication. Available online at <http://www.unet.univie.ac.at/a0828326/Impl/overview.html>; visited on June 22nd 2014.
- [9] J. Duke, A. Faiola, and H. Kharrazi. A novel visualization tool for evaluating medication side-effects in multi-drug regimens. In *Human-Computer Interaction. Interacting in Various Application Domains*, pages 478–487, 2009.
- [10] J. Duke, X. Li, and S. J. Grannis. Data visualization speeds review of potential adverse drug events in patients on multiple medications. 2009.
- [11] S. Mittelsttdt, M. C. Hao, U. Dayal, M. Hsu, J. Terdiman, and D. A. Keim. Advanced Visual Analytics Interfaces for Adverse Drug Event Detection. In *In Proceedings of the Working Conference on Advanced Visual Interfaces (AVI 2014)*, 2014.
- [12] T. Munzner. *Visualization Analysis and Design - DRAFT - 19th March, 2014*. 2014. pages 136-137.

- [13] T. Munzner. *Visualization Analysis and Design - DRAFT - 19th March, 2014*. 2014. pages 50-51, 237-267.
- [14] T. Munzner. *Visualization Analysis and Design - DRAFT - 19th March, 2014*. 2014. page 53.
- [15] T. Munzner. *Visualization Analysis and Design - DRAFT - 19th March, 2014*. 2014. pages 133-134.
- [16] T. Munzner. *Visualization Analysis and Design - DRAFT - 19th March, 2014*. 2014. page 180.