Effective Visualization of Tuberculosis Three-Drug Assays: A Design Study

Sunganani Silubonde
Department of Computer Science
University of Cape Town,
Rondebosch, 7701, South Africa
ssilubonde@gmail.com

Digby Warner
Mycobacteriology Research Unit
University of Cape Town,
Rondebosch, 7701, South Africa
digby.warner@uct.ac.za

Michelle Kuttel
Department of Computer Science,
University of Cape Town,
Rondebosch, 7701, South Africa
+27 21 6505107
mkuttel@cs.uct.ac.za

ABSTRACT

The rise in multidrug-resistant tuberculosis means that new drugs and new drug combinations are needed to address the problems associated with current treatments. Drug screening facilities aim to identify new high quality drug compounds, or novel drug combinations, for treatment of tuberculosis. The experimental drug assay procedure produces multivariate data that is difficult to analyse and onerous to process in order to determine which drug combinations should be pursued for further development.

In this design study, we have developed a visualization tool to assist with analysis and processing of this multidimensional data. The tool was developed with an iterative user-centred design process, beginning with a low fidelity paper mock-up through to the deployment of a fully functional, computer-based prototype that expert users judge to be both usable and effective.

CCS Concepts

• **Human-centered computing→Scientific visualization** • *Human-centered computing→User centered design*

Keywords

visualization; Tuberculosis; drug screening assay; user-centred design; design study.

1. INTRODUCTION

A design study has been defined as "*project in which visualization researchers analyze a specific real-world problem faced by domain experts, design a visualization system that supports solving this problem, validate the design, and reflect about lessons learned in order to refine visualization design guidelines*" [8]. Our design study described here aims to develop a data visualization that will aid drug discovery teams in identifying the most promising active drug combinations against drug-resistant tuberculosis.

SAICSIT '16, September 26-28, 2016, Johannesburg, South Africa.
 ACM 978-1-4503-4805-8/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2987491.2987501

Tuberculosis (TB) is caused by *Mycobacterium tuberculosis* and kills approximately 1.5 million people annually [1]. Further, about a third of the world's population is infected with latent TB [2]. The rise in multidrug-resistant TB means that new drugs and new drug combinations are needed to address the problems associated with current treatments [2].

TB drug screening facilities, such as the Molecular Mycobacteriology Research Unit (MMRU) at the University of Cape Town, aim to identify new high quality drug compounds, or novel drug combinations, which inhibit the growth of *M. Tuberculosis*. The screening process uses 96-well microplate assays to measure drug potency and efficacy. Three-drug combinations are tested using five assay plates per experiment, a procedure which produces a three-dimensional array of drug interactions.

Each assay experiment places candidate drugs at varying concentrations across wells in the microtitre plate. A constant volume of stained bacteria is added to each well. After a defined period, the microplate reader measures the absorbance/luminosity for each well. High absorbance values indicate thriving bacteria and hence an ineffective combination. Conversely, a low absorbance value indicates few bacteria remaining, suggesting successful growth inhibition and an effective drug combination. Drug concentration values and luminosity/absorbance values are used to determine drug potency, efficacy and drug-drug interactions by calculation of the Minimum Inhibitory Concentration (MIC) and Fractional Inhibitory Concentration Index (FIC index) values for each well. The MIC is the minimum concentration of a drug, when acting alone, required to inhibit the growth of bacteria. A concentration lower than this value results in unsuccessful inhibition, one higher prevents bacteria growth. Consideration of the drug concentration is important for avoiding toxicity in human treatment: ideal drugs inhibit bacterial growth at low concentrations. The FIC index is a linear sum of the drug concentrations weighted according to the MIC, when two or more drugs are in combination. The FIC index value represents the type of drug interaction: synergistic, antagonistic, no effect or additive. A drug screen thus ultimately produces a 5-tuple data set comprising: drug concentration, luminosity/absorbance, Minimum Inhibitory Concentration (MIC), Fractional Inhibitory Concentration Index (FIC index) and percentage inhibition.

For three-drug combination assays, the resultant 5-tuple 3D data set is difficult and time consuming to analyse and explore. To our knowledge, there are no existing visualization tools that focus on the display of microplate data, let alone three-drug combination assays: TB researchers currently plot and compare the 2D data from individual microplate assays using Microsoft Excel. An effective visualization of microplate data that leverages the ability of the human visual system to rapidly detect and identify patterns has the potential to accelerate analysis [3]. Developing a scientific visualization that is both effective and usable to the users is a form of problem-driven research that entails interaction with domain experts to design a satisfactory solution. The design study described here comprises a contextual enquiry followed by a cyclic user-centred design process: working in consultation with a researcher, we built from a paper-prototype, and iterated to a fully functional implementation of our final visualization design.

2. METHODOLOGY

2.1 Visualization design guidelines.

An effective data visualization acts as a cognitive tool, facilitating understanding of, and reasoning about, data. The design phases in this study were informed by the visual analytics mantra adapted from Schneiderman [10]: analyse first, show the important; zoom, filter and analyse further; details on demand [11]. This mantra establishes that the user initially has an overview of the data, then identifies regions or patterns of interest and only, if they wish to, then analyses further and accesses the details of the data. This ensures that there is no information overload at any time and exemplifies the need for user–system interaction: an overview of the data is presented and additional data can be requested through interaction with the system. This is particularly important for multivariate data, which gives rise to unique challenges in designing an effective visualization. For example, care must be taken when mapping data attributes to graphics to avoid overwhelming the observer's viewing ability. Too much detail at once increases the difficulty of distinguishing relevant and irrelevant information. Good designs maximize the use of space and intelligently structure space. Further, interaction techniques play a key role in determining the effectiveness of a visualization [19]: the use of consistent usable interaction design models with clear visual feedback is essential. Data representation without interaction is merely a static image: in order for visualizations to be a means of cognitive support, they must allow for manipulation and data exploration [14].

The primary objective of the visualization tool is to aid in identifying the drug combinations that are most promising, a process which cannot be done by merely looking at the assay plate images. The researchers would like to directly elucidate the dose (drug concentrations)-response (inhibition/potency) relationship and identify regions of significant synergy/antagonism. The specific design goals for the visualization module were identified in collaboration with by Charles Omollo – a PHD student in the MMRU specializing in TB Drug Research using three drug combinations and thus an expert user of the system. The high-level analysis process requirements for the tool were identified as:

- an accurate and clear representation of the well plates, which displays the number of wells in which all three drugs work together, including both synergistic and antagonistic interactions;
- the ability to select a plate of interest to view in more detail;
- the ability to view interaction type, FIC and % inhibition for a well.

Key requirements were that visualization must provide a faithful representation of the raw data to prevent inaccurate observations and that the output must be able to be mapped onto the visualization of the plates.

2.2 Design Study Methodology

As a standard procedure for the analysis of three-drug experiments could not be found in literature, experts in the field of drug research were included in an iterative user-centred design process. User-centred design involves the user in the design process to ensure that the resulting product is understandable, usable and accomplishes the desired tasks [12]. Sedlmair et al. have proposed a methodological framework for design of a visualization that requires collaboration between the designer and user expert [8]. Three categories exist in the model: a precondition phase that describes what must be done before starting a design cycle; a core phase presenting the main steps of design; and an analysis phase depicting the analytical reasoning at the end. This work aims to produce a validated visualization design and is thus is primarily concerned with the core phase of the design study comprising four stages: *discover*, *design*, *implement*, and *deploy* [8].

The first step in the core phase is *discover:* problem characterization and abstraction. For this phase, we performed a detailed task analysis as part of a contextual enquiry [13], where an expert user was observed working in the laboratory and interrogated as necessary.

The next three stages in the core phase are *design, implement* and *deploy.* When the nature of the problem does not allow the designer to move with certainty from problem to solution, an iterative design process employing successive prototypes that are evaluated by the user is appropriate [9]. Such prototypes allow users to see and experience the system long before the final version is deployed. Our methodology incorporated four *design-implement-deploy* iterations of increasing complexity, each of which involved user evaluation. Details on each of the processes follow below.

2.2.1 Discover

The aim of the *discover* phase of data gathering and requirement analysis is to determine and clearly define the user requirements for the system being designed. The typical user profile for the system under design is a researcher in a drug screening facility investigating the efficacy and potency of a specific three-drug combination for treatment of TB. We used contextual inquiry – a semi-structured interview method that combines observation and interview – to gather information about the context of use [13]. This combination of structured queries with observation of the user reduces the possibility of omitting pertinent information [14].

The contextual enquiry took place in an MMRU research laboratory. The aim of this process was to understand the users requirements, the tasks they carry out and the environment in which they work. To gain further understanding of the tasks, the designer took the role of an experimenter: under the supervision of the expert user we conducted a three-drug combination assay experiment with a microplate reader and analysed the results, taking notes and asking questions during the process. A non-

harmful bacterium, with similar behaviour to *M. Tuberculosis,* was used and known effective drugs were used as reagents.

2.2.2 Design-Implement-Deploy Iterations

Four iterations of the design process were conducted, with prototype implementations increasing in fidelity with each iteration. The design was thus refined progressively, which lessens the risk of over-committing to a design too early.

The first two iterations were the simplest prototypes, the first a low-fidelity paper prototype and the second a simple interactive prototype. These prototypes were evaluated qualitatively, using the Wizard of Oz [15] and Think-out-loud [18] protocols respectively. User opinions were collected for both. Both of these prototypes used a synthetic data set.

The third prototype was a high-fidelity implementation that used a real data set and served to test the accuracy of the system and its integration into the experimental process. This prototype was evaluated with a questionnaire addressing usability and usefulness.

The final iteration was used to address any outstanding issues with the system and was assessed with an expert evaluation and subsequent interviews.

2.2.2.1 Iteration 1

This iteration involved a low-fidelity paper prototype which was used to verify the tasks identified in the Discover phase and to reveal the methods of interaction the user employs to achieve the goals. The iteration developed multiple visualizations for each task and identified the methods of interaction that are most natural to the user. This is a rapid iteration intermediary between idea and implementation, allowing the system to be evaluated before too much time is consumed in development.

The paper prototypes were hand-drawn mock-ups of the system design, which are inexpensive, easy to change and quick to complete. Rough sketches of the system were drawn on A4 paper, which represented the application window. The Wizard of Oz method was to evaluate these prototypes [15]. In this method, the user is presented with the mock-ups and allowed to "use" the prototype application as if it really existed by pressing on parts of the "screen" to mimic actions they would have done on a functioning machine [16].

This iteration was tested with a combination of expert users who had knowledge of TB drug assays and non-expert users with knowledge of the TB drug discovery experimentation process. Expert user testing using Wizard of Oz is a way to clarify tasks and user requirements; while non-expert user testing validates that the tasks can be completed. Both sets of users reveal interaction methods that can be used in the development of the system.

2.2.2.2 Iteration 2

This iteration developed a simple, interactive prototype to ensure that the core functionality of the system could be implemented successfully. The Think-Out-Aloud evaluation method was used to evaluate the prototype, where the users are encouraged to express their thoughts verbally and an observer records them [18]. The purpose of this qualitative evaluation was to identify the aspects that experts liked and disliked and what they found useful about the system. A brief interview session was held after the user interacted with the system and the observer questioned the user on the actions they performed. This was done because most people find it difficult to act and speak at the same time. As interactions were implemented, the system was also assessed for ease of interaction: the number of times the user asked a question was noted, and we also documented when they could not manipulate the system or were not able to perform a task.

2.2.2.3 Iteration 3

Iteration 3 refined the design and developed a full implementation of the visualization system: challenges and problems identified in Iteration 2 were addressed. This iteration was the first to use a real data set from a three-drug assay previously conducted. This iteration was evaluated both quantitatively and qualitatively with six users who all work in TB drug research. The users were given the system with the task of identifying the drug concentration combination that best inhibits the growth of the bacteria. Each user was observed as they used the system, the tasks performed were compared to the original tasks identified in the *discover* phase. In addition, the prototype was judged on its usability (effectiveness, efficiency, learnability and error prevention) and user experience (satisfaction and value). The usability of the system was evaluated both through observation of the users, who were recorded on video as they interacted with the system, as well as with the widely used System Usability Scale (SUS) questionnaire [6]. SUS comprises ten statements, equally distributed between positive and negative, that are rated on a five-point scale. The statements cover measures of effectiveness (the ability of the user to complete a task and the quality of the output of the task performed), efficiency (how long the user takes to complete the task and usefulness of the system) and satisfaction (subjective reactions of the user to using the system). The user (respondent) indicates the degree of agreement or disagreement with each statement. SUS has the advantage in that the final score is a single number between 0 to 100, which represents the composite measure of the usability of the system. The SUS questionnaire has been found to be a robust measure of system usability even with a small sample size (as few as two users) [18]. To interpret the SUS score, we adopted the approach proposed by Bangor et al. [21]: the acceptability score is mapped to five descriptions, "worst imaginable", "poor", "OK", "good", "best imaginable".

Table 1: Statements on usefulness, users indicate degree of agreement on a 5-point scale.

1. This system enables me to accomplish tasks more quickly.
2. This system supports critical aspects and answers critical questions.
3. This allows me to accomplish more work than would otherwise be possible.
4. Overall, I find this system useful.
5. The system displays information in a logical order.
6. The system highlights potentially effective drug combinations.
7. I would recommend integration of this system with current TB drug discovery experimental methods.

After assessing usability with the SUS, the usefulness of the system was also probed with seven statements aimed at

determining whether they found the system to be useful and if they would recommend its use, as listed in Table 1.

2.2.2.4 Iteration 4

The final iteration of the project incorporated no new design ideas, but tested the system with real data and two expert users. The system was evaluated with a new data set obtained from the expert user and not previously analysed – the data set comprised results of the most recent experiment he had conducted. In the expert evaluation process [18], one expert user was asked to analyse the data set with the visualization system and to draw conclusions, while another expert user analysed the same data set using the current manual methods. The conclusions were compared for similarity, and key values such as the FIC and MIC compared for accuracy.

2.2.3 Evaluation Data

The first two iterations validated design and feasibility respectively and were therefore tested with synthetic data resembling real data sets. These data sets were generated using a real data set as a template. The microplate reader measures absorbance between 20000 and 65000, and assigns values to a 12 x 8 grid for each well. We used Microsoft Excel to generate random absorbances and concentration values in this grid.

The final two iterations were tested with real data sets, representative of a single experiment and including drug names and concentrations and absorbance values (CSV) for each well on all five plates. The FIC, MIC and well-specific drug concentrations were collected and used to ensure that the data processing module produced accurate results.

3. IMPLEMENTATION

The visualization module takes as input (in CSV format) luminosity/absorbance values and initial concentrations of the drugs which the user inputs. The application has two aspects: data processing, which takes the user input and produces data in a suitable format for the visualisation, and a graphic representation aspect which produces images the user can interact with.

The visualization tool in Iterations 3 and 4 was implemented in JavaScript to create a rapid prototype that could be hosted on a web page. The algorithms focused on ease of implementation, not efficiency: as the size of the input data set is relatively small, performance is not an issue. D3-js, a library for data driven applications, was used to visualize the plate data. Snap.svg, a JavaScript library for scalable vector graphics (SVG), was to show images. Interaction was implemented with jQuery. A Python server hosts a static web page on which the visualization is run.

4. RESULTS

4.1 Discover

The pre-condition phase of the design study (not discussed here) established the primary objective of the visualization tool as to aid in identifying promising drug combinations for further pharmacological development as a TB treatment. The contextual enquiry in the *Discover* phase of this design study was used to elucidate the exact steps in the experimental procedure from which we abstracted the key sequence of events, as described below.

4.1.1 Experimental procedure

The assay experimental procedure places a combination of drugs and bacteria in each well of a 96 well plate at various concentrations. The bacteria are stained - interaction with the drugs results in a blue, pink or mixed colour change in the well. The colour intensity is measured using a microplate reader. A maximum reading signifies uninhibited growth.

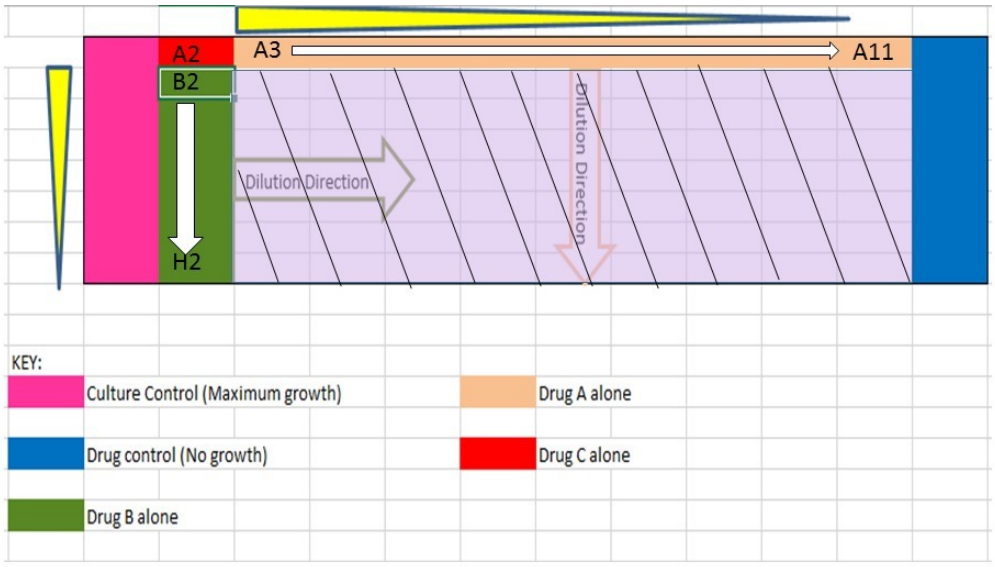


Figure 1: Schematic of the layout of a single microplate for a 3-drug TB assay.

The wells in each plate are filled as shown in Figure 1. The first and last columns are culture (pink) and drug control (blue) respectively. In each plate, the cell A2 (red) is reserved for drug C alone, rows A3-A11 are reserved for Drug A alone and Drug B is

placed in column B2-H2 (green). Drug A and B are two-fold diluted in each well, moving right across the columns (A) and down across the rows (B), with the maximum concentration of the drugs occurring in wells A3 and B2. Drug C increases in concentration from Plate 1 to Plate 5. The rows containing drugs alone enable MIC values for each drug to be calculated. Each column has a constant concentration of Drug A and each row has a constant concentration of Drug B. Drug C is added to the interacting region shaded lilac.

On completion of the experiment, the researcher follows the procedure below for each plate.

1. Determine the MIC of the drugs alone (MICAalone, MICBalone and MICalone). This is usually the concentration of the drug in the last well that has luminosity lower than the maximum. This is done by visual inspection of the microplate: the researcher judges where there is a colour change and concludes that the well before the colour change contains the MIC.

2. Determine the MIC of the three drugs in combination. This is found in the shaded region (Figure 1), and is signified by the drug concentrations in the well that is not pink on the plate and has the lowest concentrations of drugs A and B.

3. For each well in the interacting region, calculate the FIC values of the three drugs using the formula
$FIC = (concetration\ of\ drug\ in\ well)/(MIC\ of\ drug)$.

4. For each well in the interacting region, determine the FIC index (FICI) using the formula: $FICI = FIC_A + FIC_B + FIC_C$.

5. Determine drug efficacy, measured by the absorbance/luminosity values produced by the microplate reader. Absorbance signifies the amount of bacteria growth in the well and can thus be manipulated into % inhibition. High values mean a large amount of bacteria and low values indicate high bacteria inhibition. The percentage inhibition is calculated using:

$$Luminosity = well - ave_{control}$$

$$\%inhibition = \frac{luminosity}{ave_{culture} - ave_{control}} -$$

6. Determine the lowest FICI that gives % Inhibition >90% and determine the synergy for that well.
Synergy determination:

FICI <= 0.75	*Synergistic*
0.75>FICI<=4	*no effect/additive effect*
FICI >4	*Antagonistic*

A drug combination that is of interest has typically has the lowest FICI. However, for human treatment, another consideration is that the drug concentrations should also be relatively low, as, high concentrations can be harmful.

4.1.2 Task analysis

From the contextual enquiry, we identified the following sequence of events that an expert completes to determine if a candidate drug combination should be pursued further:

1. *Run the assay experiment and obtain absorbance values obtained for each well in the plate. Visually compare all the plates.*

In this task the user determines, by visual inspection of the plates, which plate appears to have promising results.

2. *Select a single plate for futher analysis.*

This process allows for analysis and data processing and has to be done for all the plates.

3. *Analyse the plate to determine the region of interaction, type of interaction, MIC, FIC and % Inhibition.*

This data is important because it helps the user determine whether the drug combination is to be investigated further. It is also a means for the user to gain a deeper understanding on the efficacy and potency of the drugs.

4. *Compare multiple plates to identify data points with synergistic interactions and high % inhibition.*

This task is the most important, as it determines which drug combination should be pursued for further development. With this information the user identifies the concentrations at which the drugs are most effective and how effective they are.
This analysis revealed that a single graphical representation is unable to address each task because dependencies exist between them. As a result, each task is associated with its own design and the following interactions are employed to link the representations:

- *Select:* mark something as interesting.
- *Explore:* show something else.
- *Reconfigure:* show a different arrangement.
- *Encode:* show a different representation
- *Abstract/Elaborate:* show more or less detail.
- *Filter:* show something conditionally.

During the task analysis, we identified only two misuse cases: an error with the experiment which requires the user to redo the experiment and the microplate producing faulty results which calls for the absorbance values to be re-measured or the experiment to do redone.

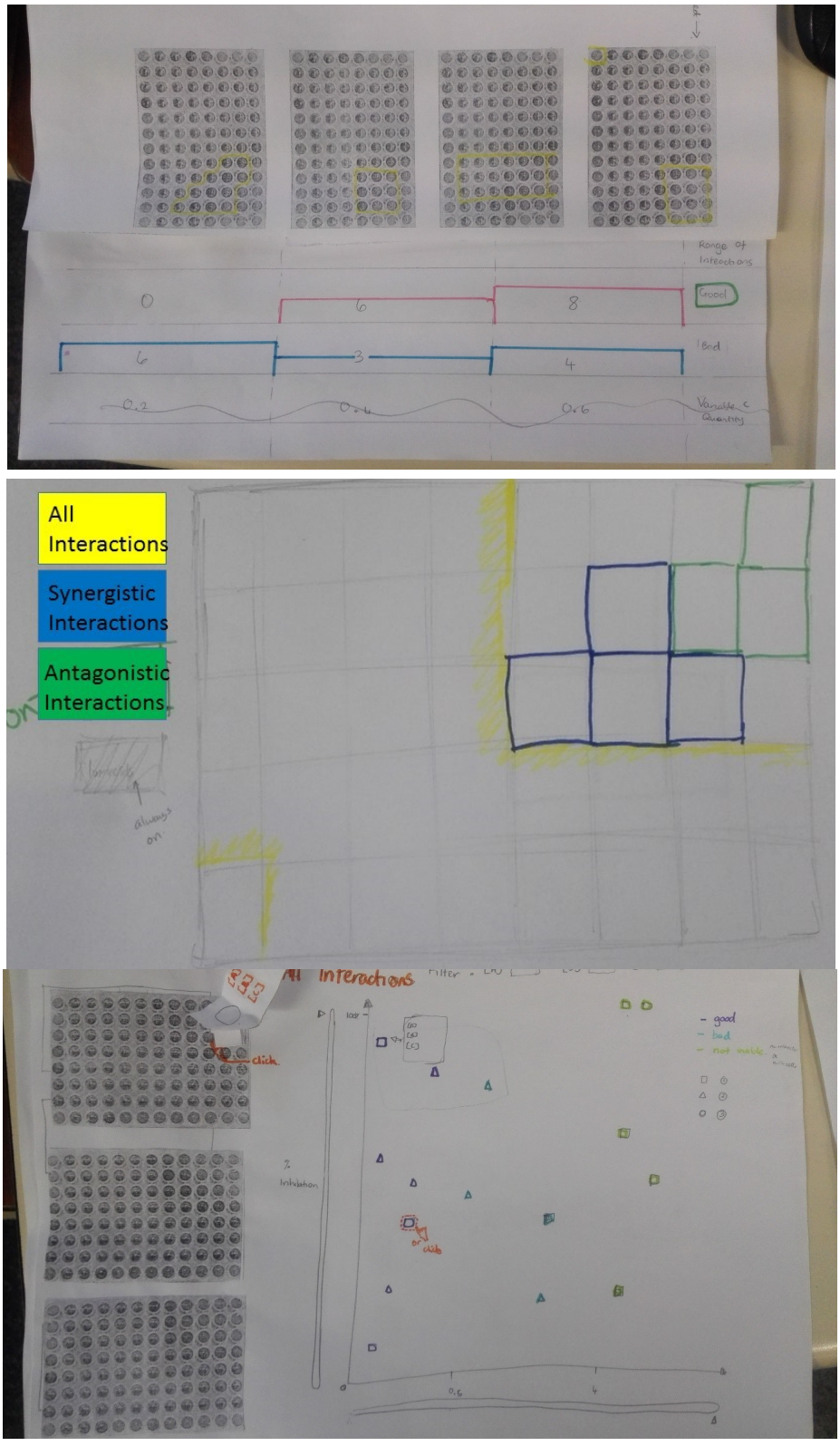
4.1.3 Prototype architecture

Task analysis enabled us to determine a suitable architecture for the software prototype, with the assumption that the user has setup identical to that used currently in the MMRU. The architecture of the visualization module has three components: user input, data processing pipeline and data representation.

First, the user inputs the initial concentration values for drugs A and B and the concentrations of drug C in each plate. This process aims to reduce the amount of work the user does before a graphical output is displayed. The user enters only the initial concentrations of the drugs and a twofold dilution for Drug A and B is assumed. Drug C concentrations are entered manually, as these values differ in dilution.

(a)

(b)



(c)

Figure 2: Paper Prototype design. (a) (Left) Start page for the Visual application. (Right) Main page of the visualisation, which serves as the home page where the user views an overview of the experiment. (b) A zoomed in plate view which appears when the user selects to view a single plate. (c) Graphic representation produced when multiple plates are selected simultaneously.

Then the user then uploads a CSV file from the microplate reader containing luminosity/absorbance values. This raw data is processed in the processing module. Here, each well is assigned drug concentration values for all three drugs based on a twofold

dilution, absorbance values from the CSV file, as well calculated FICI, MIC and % inhibition values (calculation details in 4.1). In the step, the data is converted into a form suitable for the visualization. The % inhibition value for each well is calculated.

The data is then passed to the visualisation system without the absorbance values, as they are not used in the graphic display. The visualisation data is an aggregate of five values (drug concentration, FICI, MIC and % inhibition) per well. The visualization system produces a graphic display of the data and allows the user to interact with it.

4.2 Design Iteration 1

In this first stage of our user-centred design process, we designed a rough paper prototype (Figure 2). In this mock-up, the main page view was chosen to be a flat structure instead of a 3D stack of the plates, because of the difficulty in viewing all plates in 3D. The microplates are each represented by an image glued onto the paper (Figure 2a). When a user selects a single plate (Figure 2b), the total number of synergistic and antagonistic interactions for each plate is plotted under each plate. On the left, buttons allow the user to view synergistic interactions, antagonistic interactions or both. These regions can also be viewed on the plate, with different border highlights indicating different types of interaction. Luminosity/absorbance values are represented by the shading in each cell: a fully shaded cell represents 100% inhibition whereas a cell that is not shaded shows 0% inhibition.

When multiple places are selected simultaneously (Figure 3c), the plates are arranged on the left hand side with a scatterplot of percentage inhibition versus FICI on the right. The scatterplot uses different shapes and colours for the data points of each plate, so that visual queries can be answered quickly: the user is able to link the shape and colour of a plot point to a given plate. Sliders on the x- and y-axis allow for filtering of the data points to reduce the amount of data shown at any moment. The pop-up shows the result of a user selecting a well. Users were asked to complete the following tasks on the paper prototype:

1. Obtain an overview of the experimental results.
2. View a single plate. Analyse the plate to determine the region of interaction and the type of interaction.
3. Compare two or more plates.
4. Determine the drug concentrations, MIC, FIC and % Inhibition values for any well in any plate.
5. View data points that fall within the region of low FICI and high % inhibition.
6. Establish which plate has the data point with the highest % inhibition and lowest FICI.

Evaluation of the prototype revealed both good and bad design decisions and confirmed the sequence of the tasks. The first task required users to obtain an overview of the experiment; this was completed by submitting a user name, initial concentration values for each drug and uploading a CSV file with absorbance values on the start page. All participants were able to complete this activity easily. Participants appreciated the fact that they could easily compare the number of good and bad interactions across all the plates, and that it allowed them make quick decisions about which plates are potential candidates.

Task two, which required the participants to view a single plate for further analysis, revealed confusions around interaction with the system. Some users clicked the plate, some clicked the bar highlighting the number of interactions, while others were confused and asked the facilitator for help.

Viewing different types of interactions was a challenging subtask for some of the participants, mainly because the paper prototype did not provide sufficient affordances. A number of participants did not see certain widgets as buttons, which made the execution of the given task difficult: they thought that they were labels (Figure 2b).

For task 3, comparison of multiple plates (Figure 2c), we expected the user to go back to the home page and from there select more than one plate by clicking. However, this was not how the users attempted the task: the majority of them were unsuccessful at this activity.

Paper prototypes, although cheap and quick, do not offer the same imagery that a computer based prototype does. Many times the users were unclear on whether certain widgets represented buttons or legend labels. We also identified affordances to add to the design, such as a "select all" button to allow the user to compare all the plates. The pop-up that appears when a well is selected was found to be restrictive, as only the one well's data could be viewed at a given time. The placement and sizes of certain widgets, such as buttons, affected the user's ability to perform certain tasks intuitively. The slider bars were found to be an effective filtering tool: expert users recommended the addition of slider bars for the drug concentrations. The expert users highlighted the importance of having a constant view of the plates throughout the analysis process.

Testing with non-experts required the facilitator to explain the process more thoroughly and at times tasks could not be completed because the user did not have enough knowledge of the experimental procedure. Expert users were new to the concept of paper prototypes and often got caught up in details that did not pertain to the tasks they had to complete. This was both challenging and useful, it helped to get a better understanding of expectations, but at times the additions suggested were either beyond the scope of the project or too tailored to their specific experimental procedure. Overall testing with both experts and non-experts is beneficial to ensuring the correct level of usability and determining the accuracy and logic of the system before committing to a design.

4.3 Design Iteration 2

The paper prototype evaluation provided insights into strengths and weaknesses of the initial design. Design changes included implementing a static view of all the plates at the top of each page. Check boxes added to each plate enable selection of multiple plates. The option to select all or none was also added.

Instead of the pop-up a table containing drug concentration, FIC index and luminosity values were placed below the plate representation when viewing a single plate and below the scatterplot when viewing multiple plates. The table was placed at the bottom of the screen to provide more horizontal room for the plate, allowing each well to be larger and consequently easier to click. Also, the scroll wheel on mice make vertical scrolling easier than horizontal scrolling, which requires use of the browser's scrollbars. Sliders were added for the concentration values when multiple plates are selected (Figure 9) to allow the user to filter

the data by concentration as well as luminosity and FIC index. The sliders are double sided so that the user can select a maximum and minimum value as they see fit.

After making the necessary adjustments, a new prototype was tested with six users who had not seen the system before, so as to eliminate bias, as previous users may recall tasks and know what to do. The users were able to complete all the tasks that were asked of them and did so with much more ease than before. The users liked that they could easily identify interesting data points. The detailed view of a single plate was found to be useful for exploring each plate individually and changing the type of interaction displayed at will. This indicates regions on the plate which revealed a trend to the user. The check boxes made it easier to select multiple plates and the users appreciated the ability to filter the scatter plot points using the sliders. A majority of the users requested that percentage inhibition be used instead of absorbance values, as this shows the potency of the drug combination more effectively without the user having to process the data any further. To aid the user visually it was recommended that the colours used for the data points on the scatter plot also be used to form a border around the plates they correspond to. This helps the user to make the relation between data point and plate. An additional feature that was added to the visualization enabled a data value on the table to be highlighted when a well is selected in the detailed view of the plate or a graph point in the scatter plot.

This iteration helped to determine feasibility of the system design as well as improve the graphic representations. Testing using the Think-Out-Aloud protocol is effective if done well, the observer must be attentive so as not to miss any detail and the user sometimes requires encouragement to express their thoughts. Recording the evaluation procedure is beneficial and discussing the system with the user immediately after their interaction with it provides valuable information that could have been missed.

4.4 Design Iteration 3

In this iteration, the usability of a full prototype was tested. Users were tasked to find the lowest drug concentration combination that interacts synergistically and inhibits the growth of bacteria, after which they completed the SUS questionnaire and questions on usefulness. Interestingly, although all the users managed to complete this single task successfully, they used a variety of different steps to do so. This illustrates the inherent exploratory nature of drug candidate identification. The SUS scores listed in Table 2 indicates that the system is acceptable in terms of usability: the score of 72.5 corresponds to a rating of 'good' and the highest of 95.5 is “excellent”. All six users strongly agreed on the ability of the system to help them conclude if a candidate drug combination should be pursued for further study. The users also indicated that the system would allow them to accomplish the analysis task more quickly and that they would recommend the use of the system. The users comments highlighted the positive aspects of the design as well as possible improvements. The mapping of the microplates was a primary concern for users. A good mapping affords clarity and ease of use allowing the user to rely on recognition rather than recall. Users suggested changing the ID values of the data points in the table to indicate row and column so that the user could easily pick out the corresponding well; this is the conventional way of labelling a microplate. In addition, the random colours used for the well concentrations confused many users. It was recommended that the wells take on the conventional blue, purple or pink colours, which correspond to growth, some growth and no growth of bacteria. Both of these suggestions were implemented in the final iteration.

Table 2: Converted user responses for each statement and final SUS Score for each user calculated as follows:
Odd statement= (User Response -1) and Even number statement = (5-User Response)

	User	User	User	User	User	User
Statement 1	4	4	4	4	4	3
Statement 2	3	4	3	3	4	3
Statement 3	4	3	3	3	4	3
Statement 4	3	3	3	3	3	3
Statement 5	3	3	3	3	4	3
Statement 6	3	3	3	3	4	3
Statement 7	3	4	3	2	3	3
Statement 8	4	3	3	3	4	3
Statement 9	3	3	3	3	4	3
Statement 10	3	3	3	2	4	3
Sum of Converted	33	33	31	29	38	30
Overall Score = Sum *	**82.5**	**82.5**	**77.5**	**72.5**	**95**	**75**

4.5 Design Iteration 4

This final iteration ensured precision (a key concern for researchers) and robustness. Minor changes also were made to the design (Figure 3), principally involving alterations to make the representation of the microplates as close as possible to the real microplates. A single plate is represented as a 2D array which akin to the grid structure of a plate (Figure 3a). Each element represents a well in a plate consisting of percentage inhibition, drug concentrations and FICI values. The well colours were changed to blue, purple and pink, according to the percentage inhibition value for each well. Blue represents inhibition greater than 90%, purple inhibition between 90%-50% and pink inhibition less than 50%. Comparison of two or more plates creates a scatter plot of percentage inhibition vs FICI (Figure 3b).

The evaluation of this final system with two experts confirmed that our visualisation tool was able to produce the same results as those obtained using manual methods. The only clear difference was that manual methods give a range for MIC values. Manual methods rely on human vision to differentiate a colour change from blue to pink; if the researcher cannot distinguish when exactly the colour change took place they give a maximum and minimum concentration value. The true MIC lies within that given range. Our system uses percentage inhibition values and can therefore pick out the exact MIC based on the last well in the row that gave a percentage greater than zero.

(a)

(b)

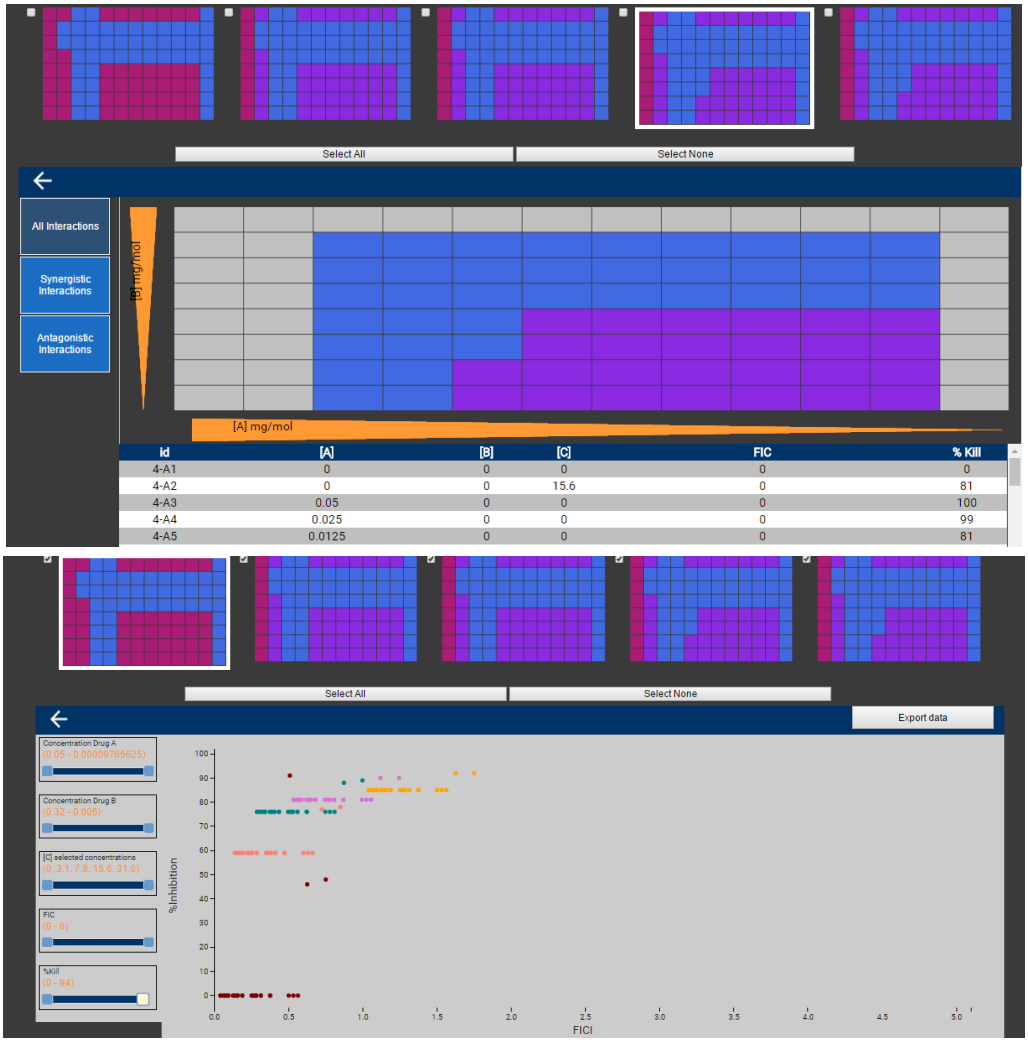


Figure 3: Final Visualisation Design. (a) Display when a single microplate is selected. The overview of all plates in the experiment is shown at the top. (b) Graphic representation produced when multiple plates are selected simultaneously, showing the scatterplot of % Inhibition versus FIC index values for the plates selected. A set of sliders on the left allows users to filter the scatterplot.

The final questionnaire evaluation with two experts had extremely positive ratings for both usability (extremely and usefulness. Most importantly, comments indicated that the tool would both reduce time and be helpful for identification.

5. DISCUSSION AND REFLECTION

This design study benefitted to a large degree from a user-centred methodology we employed involving domain experts in the design: because of the complexity of the drug screening data set and the very specific requirements of the researchers, it would have been very difficult to come up with an effective design without their direct involvement. However, difficulties arose in involving active researchers in the design cycle, as this required a time commitment, which they initially underestimated. To prevent delays in the design cycle, we had to work around the users' very full and rather inflexible time schedules.

In this study, our initial contextual inquiry involved both interviews and careful observation of an actual drug screening experiment. This detailed process was very useful for acquiring insight into the complex tasks and processes researchers perform when analysing a drug screening experiment. It also helped to clarify the specific data analysis challenges faced by the researchers.

The initial low-fidelity prototypes allowed for evaluating speculative techniques and mechanisms without the time and cost of implementing them. This method was effective because the ideas were not influenced by the designers' capabilities or programming knowledge. However, the Wizard of Oz evaluation of the non-functional prototype was unexpectedly challenging and time-consuming with users very familiar with paper-prototypes.
We found that the combination of expert and non-expert testing to be beneficial. While non-experts may find a complex system confusing, experts may be too invested in the outcome to be sufficiently critical of the initial designs. We found that our expert users appreciated many of the changes that were made based on evaluations with non-experts.

Overall, we found that the incremental build process involving four design iterations was very useful in that large changes to the design suggested by the user evaluations were easily addressed in

the early stages before actual implementation. This was particularly useful in the process of refining what was to be implemented, as we avoided addressing technical details of the non-functional requirements. In addition, we discovered that users found a close mapping of the data images to the actual laboratory microplates to be helpful. This suggests a general design heuristic for scientific visualization: where possible, to facilitate user recognition rather than recall, images or representations that map to familiar laboratory equipment or processes should be used.

The ultimate goal of our study was design of a data set visualization that aids in identifying promising drug candidates. Our qualitative final evaluation suggests that our design will assist in more rapidly identifying promising drug combinations. However, as there is no other visualization system to which expert users' could compare our design, the results of the evaluation may not be an accurate reflection of the effectiveness of our design, but rather enthusiasm for automation. Testing the system with more users and comparing it to other visualization tools would be beneficial to revealing the true usefulness of the system. Further, because of the exploratory nature of drug discovery research, it is difficult to measure success on this relatively short time scale and to what extent it really meets the requirements of drug researchers.

6. CONCLUSIONS AND FUTURE WORK

This study produced an effective design of a visualization tool to facilitate evaluation of the data from assays of three-drug combinations against tuberculosis in order to identify promising drug combinations. Our methodology combining an initial detailed contextual enquiry with an iterative user-centred design process ensured that not only was the project completed on time with all features and functions as initially specified but also that the system was usable.

Future work will involve a full implementation of the design with subsequent longitudinal testing.

7. ACKNOWLEDGMENTS

We thank Charles Omollo for training, expert testing and providing the data sets used. The South African Medical Research Council funding agency provided financial support for this work.

8. REFERENCES

[1] World Health Organization. 2015. WHO Report 2015: Global Tuberculosis Report. WHO Press, Geneva, Switzerland.

[2] Ma, Z., Lienhardt, C., McIlleron, H., Nunn, A. J. and Wang, X. 2010. Global tuberculosis drug development pipeline: the need and the reality. *The Lancet,* 375,2100–2109.

[3] Ware, C. 2008. *Visual Thinking for Design*. Morgan Kaufman.

[4] Card, S. K. and Mackinlay, J. 1997. The structure of the information visualization design space. *Proceedings of the 1997 IEEE Symposium on Information Visualization,*: 92–99.

[5] Kehrer, J. and Hauser, H. 2013. Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE Transactions on Visualization and Computer Graphics* 19, 3, 495–513

[6] Nesbitt, K. V. 2005. Using guidelines to assist in the visualisation design process. *Proceedings of the 2005 Asia-Pacific symposium on Information visualisation - Volume 45*. 115–123.

[7] Munzner, T. 2009. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics* 15, 6, 921–928.

[8] Sedlmair, M., Meyer, M. and Munzner, T.. 2012. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics* 18, 12, 2431–2440.

[9] Jones, M. and Marsden, G. 2006. *Mobile interaction design*. John Wiley & Sons, Chichester, UK.

[10] Schneiderman, B. 1996. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *Proceedings of IEEE Symposium on Visual Languages,* 336-343.

[11] Keim, D. A., Mansmann, F., Schneidewind, J. and Ziegler, H. 2006. Challenges in Visual Data Analysis. *Tenth International Conference on Information Visualization (IV'06)*: 2–7.

[12] Gabbard, J., Hix, D., and Swan, J. E. II,. 1999. User-Centered Design and Evaluation of Virtual Environments. *Computer Graphics and Applications* 19, 6,51–59.

[13] Holtzblatt, H. and Jones, S. 1993. Contextual inquiry: A participatory technique for system design. In D. Schuler, A. Namioka (eds.): Participatory Design: Principles and Practices, pages 177–210. Lawrence Erlbaum Associates.

[14] Tory, M. and Möller, T. 2004. Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics* 10, 1, 72–84.

[15] Dow, S., MacIntyre, B. and Lee, J. 2005. Wizard of Oz support throughout an iterative design process. IEEE Pervasive Computing, 4, 4,18–26.

[16] Dahlbäck, N., Jönsson, A. and Ahrenberg, L. 1993. Wizard of Oz studies — why and how. *Knowledge-Based Systems* 6, 4: 258–266.

[17] Brooke, J. 1996. SUS - A quick and dirty usability scale. *Usability evaluation in industry* 189, 194, 4–7.

[18] Carpendale, S. 2008. Evaluating information visualizations. Lecture Notes in Computer Science 4950, 19–45.

[19] Bangor, A., Kortum, P. T., and Miller, J. T. 2008. An empirical evaluation of the System Usability Scale. International Journal of Human-Computer Interaction, 6, 574-594.

[20] Saraiya, P., North, C. and Duca, K. 2004. An Evaluation of Microarray Visualization Tools for Biological Insight. *IEEE Symposium on Information Visualization*: 1–8.

[21] Bangor, A., Kortum, P., and Miller, J. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3, 114–123