DESIGN OF A REFERENCE HANDLING SYSTEM FOR CLINICAL DNA SEQUENCING ANALYSIS

Jorun Børsting
Department of Informatics, University of Oslo
P. Boks 1080, 0316 Blindern, Oslo, Norway

Alma L. Culén
Department of Informatics, University of Oslo
P. Boks 1080, 0316 Blindern, Oslo, Norway

Morten C. Eike
Department of Medical Genetics, Oslo University Hospital
P. Boks 4950, 0424 Oslo, Norway

ABSTRACT

This paper is concerned with the design of a system that handles references related to clinical DNA sequencing and analysis of genetic variants. The reference handling system is part of a large system, currently under development at the Department of Medical Genetics at Oslo University Hospital, inquiring into procedures and supportive bioinformatics tools for analysis of genetic data. Finding and evaluating relevant literature that reports on classifications of genetic variants is of high importance. In many cases, it is a requirement to compare local findings with those published in high-quality external references. This process further ensures that the correct decision on the nature of the variant is reached. The implications of the decision are relevant for both patients and knowledge production and its transferability to new patients. A user-centered design was chosen as a research approach, with user participation in both qualitative (walk-troughs, interviews and talk-aloud evaluations) and quantitative (questionnaire) inquiries. Users involvement in design and evaluation of the reference handling prototype was important for identifying diverse usability problems and design issues, which could then be improved in later iterations of the prototype. Users have also contributed to defining more general guidelines for the re-design of later versions, e.g., a need for customization, as users often have different strategies for working with references. Another problem brought forth by users had to do with handling of options when assessments were uncertain, including ways to convey that uncertainty to the next analyst working with the same references. Users involvement in design and evaluation processes, such as described in this paper, leads to system design that is more in tune with users' needs, also making the adoption of the system easier. Furthermore, it improves efficiency and quality of the analysis.

KEYWORDS

Usability, complex systems design, genetic sequencing, user-centered design, user participation, reference handling.

1. INTRODUCTION

Genetic testing based on DNA sequencing is used in clinical practices for both diagnostic and prognostic purposes. Recent advances in the underlying technology termed high-throughput sequencing (HTS), have resulted in vastly greater sequencing output for a fraction of the cost when compared to older techniques. This change has opened up for a massive increase in the clinical application of DNA sequencing, reaching more, and larger, patient groups. HTS is therefore considered a crucial factor in making personalized medicine feasible. However, the enormous quantity of data generated by HTS and issues around knowledge extraction from that data are deeply connected with an increasingly important issue in bioinformatics, the handling of so-called "big data" (Schadt et al., 2010). At present, the analysis of DNA variants found in sequencing data involves a large and fragmented set of bioinformatics tools and informational resources,

placing a high cognitive load on the individual analyst (Mardis, 2010). Although HTS technologies are effective in generating data, there is still a large developmental gap between sequencing output and final analysis results tailored to answer specific questions related to the genetic material (McKenna et al., 2010). Moreover, the lack of sophisticated and flexible applications that enable downstream analysts to access and manipulate massive sequencing data has been a hindrance to further development of tools and methods for sequencing (ibid.). Thus, it seems timely to look into ways of designing new systems based on research methods, tools and empirical findings from research fields such as human-computer interaction (HCI) and computer-supported collaborative work (CSCW). The approaches from these fields may help in design of systems that aid analysts in managing and interpreting data, focusing on their needs and their workflow, with an aim to reduce cognitive load and increase accuracy of the analysis (Eike et al., 2014).

One of the main problems with the usability of highly specialized systems, such as those used in DNA sequencing, is that highly qualified users are often not engaged in the design processes directly, resulting in systems that are not optimal for their use. The bioengineers, molecular biologists and physicians working with interpretation of results from DNA sequencing are users with high levels of expertise; they possess both tacit and complex domain knowledge, which are crucial for the analysis process. For the design of a clinical genetic analysis software to be successful, these users should be involved in the software development process, as has been argued by (Bolchini et al., 2009; Neri et al., 2012) among others. Accordingly, this paper focuses on user-centered design (Javahery et al., 2004), with user participation in both qualitative (walk-troughs, interviews, and talk-aloud evaluations) and quantitative (questionnaire) inquiries. The aim is to discover how genetic analysts work, what they do and how the new system could better support them in their work. A large number of possibilities for system improvement was identified and described in detail in (Børsting, 2014). A central tenet, crucial for the design of new systems for clinical genetic analysis, is to engage analysts in the design process and to include designers who, at least to some degree, understand the analysis processes.

In this paper, we chose to focus on a small, but important, part of the new system interface that enables analysts to handle relevant literature references to genetic variants. The main research questions addressed are: 1) How important are references for the analysis process? 2) How are references to variants handled today? 3) Are there some implications for the design of a new reference handling system that can ease the work or reduce the cognitive load for the analysts?

The results show that, while indeed very important for a decision process, references are handled in different ways by different analysts. Thus, rather than forcing the analysts to comply to the system, the new solution needed to provide some customization possibilities. Further, clear options for conveying uncertainty in assessments is necessary so that the next person looking at the same references may reach the same understandings. The identified issues, in conjunction with deeper understanding of existing practices around literature and its use in decision processes, provided guidelines for the design of a more successful handling of references in the re-designed system. Consequently, our third research question was answered in positive.

The paper is structured as follows: in the next section we provide more information on methods used during design and evaluation of the reference-handling system. In addition, we show why participants consider reference handling to be an important problem, and how it is done today. Section 3 then addresses implications for the design that our data-gathering methods have yielded. Section 4 shows how the prototype integrates users' suggestions, followed by a short discussion of results and, finally, conclusions.

2. DESIGNING A REFERENCE HANDLING SYSTEM

In order to understand the context of the problem, we have studied the literature on the general workflow of genetic analysts, and usability problems that they experience with new sequencing interfaces. Several studies were found, such as that of Shyr et al. where the authors state that *"a consensus opinion about a causal gene candidate may arise from a series of email exchanges, face-to-face meetings and sharing of references such as hyperlinks to scientific abstracts"* (Shyr et al., 2014, p. 134). The authors point out that most software does not provide suitable functionalities for facilitating multiple users to collaborate on the same data, but that such software would be highly desirable and would accelerate the clinical analysis process. In our work, one of the first things we learned was that handling the literature references was one of the hardest things for

analysts. It was the collaborative, multi-user nature of the process that was central here, and the need to convey assessments clearly, also pin-pointing any levels of uncertainty in the assessment.

2.1 Method

User-centered design, with user participation, was the approach we chose in order to find out how users with high professional and domain knowledge actually work with literature related to genetic variants. The methods chosen for the inquiry were both qualitative and quantitative; see Table 1 for a detailed summary. The qualitative methods, such as observations and interviews, were used following the basic guidelines for user-centered design: 1) Focus on the user and tasks from the start, 2) Involve users in the process of trying to find the solutions to right problem, rather than solving a pre-defined problem in a better way, 3) When the right problem is identified, a solution, with users, may be iteratively improved (Gould and Lewis, 1985).

Table 1. Methods used in order to identify usability problems and guide design of the reference handling system

Name	Focus	Method	Data gathered	Time	Participant(s)
Identifying issues - User 1	To identify the most challenging tasks and identify usability issues.	Talk aloud. Usability testing of prototype. Semi-structured interview.	Audio, video, pictures, notes. Document containing literature evaluation.	1 hour 40min	1 laboratory engineer
Identifying issues - User 2	The same as for "Identifying issues user 1" In addition, to further explore and validate identified issues.	Talk aloud. Usability testing of prototype. Semi-structured interview.	Audio, video, notes and pictures.	1 hour	1 laboratory engineer
Observation 1	To gather data about the reference evaluation functionalities.	Observation followed by a semi- structured interview.	Audio, notes and pictures. Variant classification documents.	2 hours	2 laboratory engineers
Observation 2	The same as for the first observation	Observation followed by a semi- structured interview.	Audio, notes and pictures. Variant classification documents.	1 hour 12min	1 laboratory engineer
Interview	The same as for the observations	Semi-structured interview.	Audio recording and notes.	1 hour 45min	1 lab physician
Survey	The same as for the interview and observations. In addition, validate and further explore findings.	Survey sent out by e-mail to the future users of the system.	The Microsoft Word documents containing the survey answers.		11 participants

2.2 Case: a reference handling system for genetic sequencing

Today, the process of analyzing gene variants and references is done consecutively by up to a minimum of three users. Typically, the procedure is as follows: a lab engineer performs the initial analysis of observed gene variants, using different supporting software tools and external databases, as well as checking the existing literature for references. In this process, judgments are made based on general knowledge from molecular biology (such as the effect of a given variant on protein function) and genetics, but also based on literature references they have found. The latter implies finding out whether conclusions about the clinical significance of a variant in question already exist, and are published through trustworthy scientific channels. The results of this work are then checked by another lab engineer and, finally, by a lab physician. Usually only the first two users, but sometimes all three, comment on individual findings and articles and collaborate to determine the classification of the gene variant, which describes the clinical significance of the variant in standardized terms. Although the analysts always strive to reach classifications that conclusively describe the variant as either pathogenic or neutral, in many cases, current knowledge is too limited. This then results in the classification category named Variants of Unknown clinical Significance (VUS) (Plon et al., 2008).

2.2.1 How important are references for the analysis process?

Despite the existence of local and external databases with large collections of previously classified variants and their references, the literature is still largely hand-curated and serves the purpose of eliminating as much uncertainty in classification as possible. Currently, references are also the only way of definitely classifying a genetic variant as pathogenic. In addition, this is a field with high research activity, with new findings

published at a fast pace. It is therefore important to ensure that the latest research is taken into account and that the local database is properly updated (Dienstmann et al., 2014).

In our investigations, the first step was to identify main issues encountered in testing the new analysis system prototype (Table 1). We used a walk-through with two users, using the talk-aloud technique, followed by a semi-structured interview. The main finding from these user sessions was that the most difficult issue for analysts had to do with handling of literature references.

2.2.2 How are references about variants handled today?

Scientific research articles are reviewed and evaluated by the genetic analysts in order to determine if an observed gene variant is associated with the development of a hereditary disease. The analysts starts with a list of references to evaluate, which are usually obtained from various external sources such as the universal Human Gene Mutation Database (HGMD) and gene-specific databases such as the Breast Cancer Information Core (BIC) and others that use the Leiden Open Variant Database (LOVD) system, as well as from manual Google and PubMed searches. When at least two independent articles are evaluated to be of high quality and with the same, high-confidence conclusion regarding the clinical significance of a gene variant, additional research references are often not further evaluated.

Local workflow documents on reference handling were available and served as the basis for the prototype tested with users, pre-defining the solution to the problem. The workflow data was obtained from interviews with the users, as well as from documents on Standard Operating Procedures. The purpose of including users in the "Identifying Issues" phase from Table 1, was to see if there is a difference between what users say and what they actually do (say-do problem, see (Kensing et al., 1998)), and to make sure that the right problem was identified. Findings from the user studies show that the analysts deploy different user strategies for handling challenges encountered during the variant classification process, some of which are described below. These differences were not addressed in the local workflow documents.

- **Selecting the first article to evaluate**

Based on the answers from the survey, users base their choice of which article to evaluate on different elements, some of which are shown in Figure 1. Since the analyst can stop evaluating new references when at least two articles that meet the requirements are found, supporting user strategies that shorten the time spent on finding the right articles would increase the efficiency of the evaluation process.

Question: If you find more references for the current variant, how do you choose which one to evaluate first?	
Categories	Details
Publication.	Newest, tittle, author, journal, abstract.
Databases.	Google, HGMD, LOVD, BIC.
Variant related.	Relevance for variant and Variant mentioned in article.
Annotated articles (by others).	Descriptions from others.
Random choice of article.	
Study type.	Functional study, segregation analysis, family information.

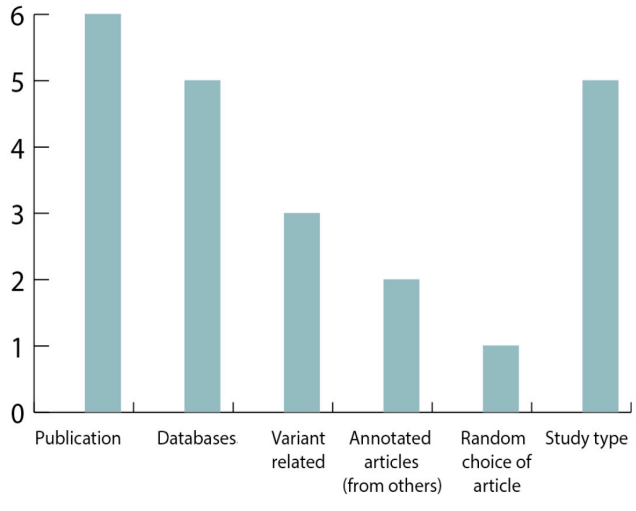
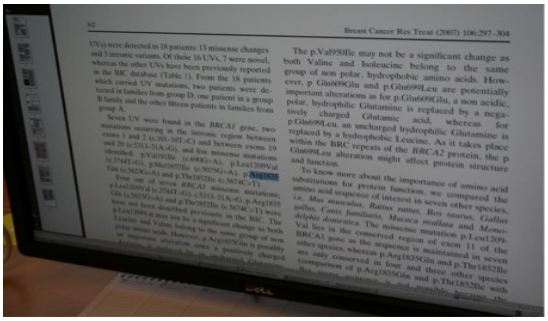


Figure 1. Results from the survey question regarding which articles that are evaluated first.

Two analysts were observed searching the PDFs of each article from the list to see how many times the variant in question was mentioned before they started the actual evaluation. This, they stated, was a way to insure that the variant was actually mentioned, but was also done to get an initial gut feeling of the article, which influenced whether it was considered a good candidate for being evaluated first. A common

observation was that the name used for the variant in the references was not always consistent, so the analyst had to search manually using many different names. Based on this observation, a suggestion for automatically providing the number of occurrences of a variant in an article was included in the survey. The question referred to a graph displaying variant name occurrences in addition to the occurrences of the words "Segregation", "protein" and "RNA", asking "Would this word count be useful for you?" The answers are displayed in the graph in Figure 2.



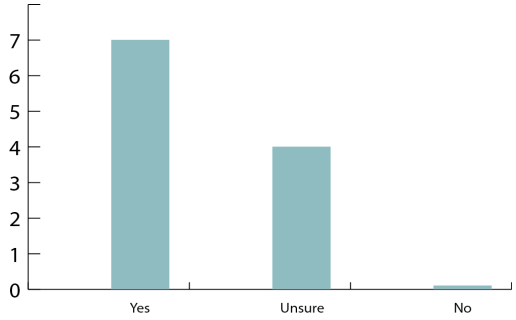


Figure 2. One of the genetic analysts searches for exact references to a variant in the PDF file of a potentially relevant article. The graph shows the answers to the question related to variant occurrences.

Although this suggestion received positive feedback from all of the users, some were unsure if this should be connected to the usefulness of the article in the variant classification.

- **How are issues of uncertainty handled today**

The biggest difference in findings from the two first user sessions were related to how these users handled being unsure of something with regards to an assessment of a reference. One of the users discussed all such matters directly with a locally available colleague. In contrast, the other user preferred to make an assessment on her own, but make a note also that the assessment was done with some level of uncertainty. The next person doing the evaluation could easily see this note and add a new assessment or a comment. This strategy is also reflected in a paper-based system developed by some analysts at the Hospital (Figure 2), where uncertain statements are color-coded in gray.

- **What is included in the comment field**

Assessments are communicated to the next analyst via comment fields, and are important for article evaluation. As one analyst puts it: *"We copy and paste from the articles to convey to the next person that this was what we found. Then, the next person can find the places we copied from in the article and read it on their own* [note: assess and verify the content themselves]*"*. In the Survey the users where asked "What do you find important to include in the comment field?" The results are displayed in the table in Figure 3.

Categories	Users	Total
The article's conclusion regarding the variant	#1, #2, #3, #4, #5, #6, #10, #11	8
Frequency data	#1, #3	2
Information about patients	#1, #3, #4	3
Splicing analyses	#1	1
Description of study method	#1, #4 #9	3
Functional study	#1, #3, #6, #9	4
Family information	#1, #4	2
Whether the article is well written	#1	1
A summary of the article	#1	1
Personal opinions about the article	#1, #3, #7, #10	4
Date and name of prior evaluators	#2, #4	2
Findings from articles referred to in the article	#5	1

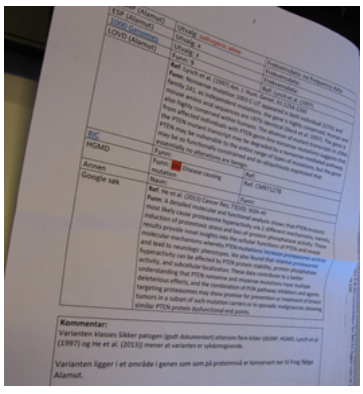


Figure 3. The table displays the Survey result for the question related to information in the comment field.

3.2 Issues of uncertainty

The analysts stated that they needed a clear set of options for how to handle assessments that involve uncertainties.

3.2.1 Color-coding the uncertainty

Some analysts suggested a color-coding scheme for handling uncertainty and conveying the fact that there is uncertainty and how much of it. This information should indicate the level of uncertainty to the next analyst performing the evaluation. This functionality was included in a new prototype, see Figure 5. Since assessments are based on a variety of different inputs, uncertainty is frequent, and handled through collaboration.

3.2.2 Handling the study type

Perhaps the most time-consuming issue related to reference articles is the assessment of the study's quality. Some studies declare, for example, that they are functional studies, but it is not enough to rely on authors' declaration of the study type. The analyst must check whether all procedures were done properly and assess if conclusions can be trusted. Also, the information they need on the specific variant under consideration is not always easy to extract from the article. For example, finding out what kind of study material (patient data, family history etc.) was used on the specific variant is often difficult. Therefore, analysts need to have a study type and study material used on a specific variant specified. A comment field is suggested for this purpose.

3.3 Prototyping a new solution for reference handling

Based on the analysis of all the data collected, two paper prototypes were developed to investigate the research questions further through the survey that all involved participants filled out. The prototypes are Photoshop redesign of a high fidelity prototype that was being developed. Based on the survey results, final prototypes were developed and are shown in Figure 4 and Figure 5.

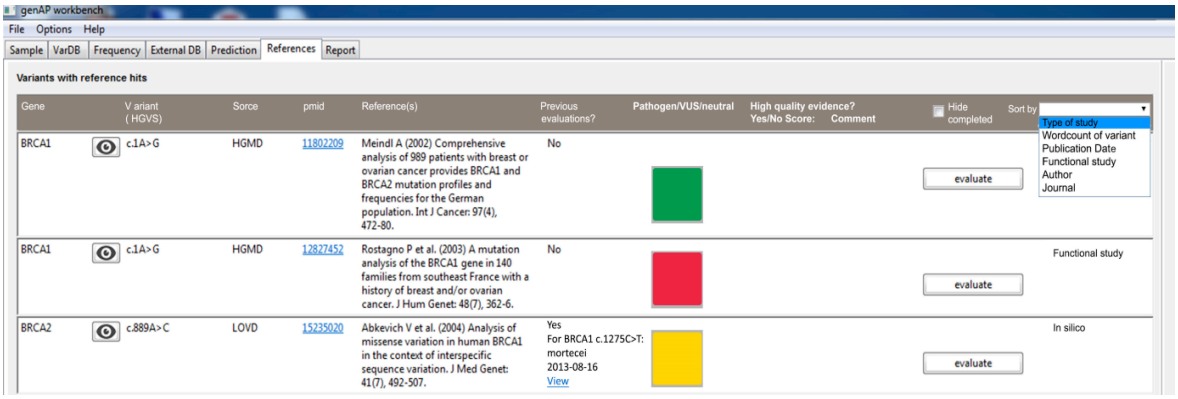


Figure 4. Prototype showing the list of references presented for analysts to evaluate.

In Figure 4, the evaluated articles are color coded with red meaning pathogenic variants, green neutral variants and yellow for VUS variants. This was based on the user suggestions. In addition, the top field is colored in a dark color with white text to highlight this area, again, as a user suggested in the survey. The right-hand corner contains a drop-down menu for selecting the type of the study, word-count, date of publication and other criteria users practice.

On the top right side, the variant word-count is displayed together with the variant name used in the article. The button labeled "find variant" can be used to find all occurrences of the variant in the text. To handle issues of uncertainties, it is a possibility to click the button "mark with some level of uncertainty" then the selected text is colored gray, like in the user developed strategy to evaluate research references. The type of study can be selected in a drop down menu. The registered study type is later displayed in the list of

articles to be evaluated. It is possible to register several studies related to one article. The comment field is refined into study method, study material and a comment itself.

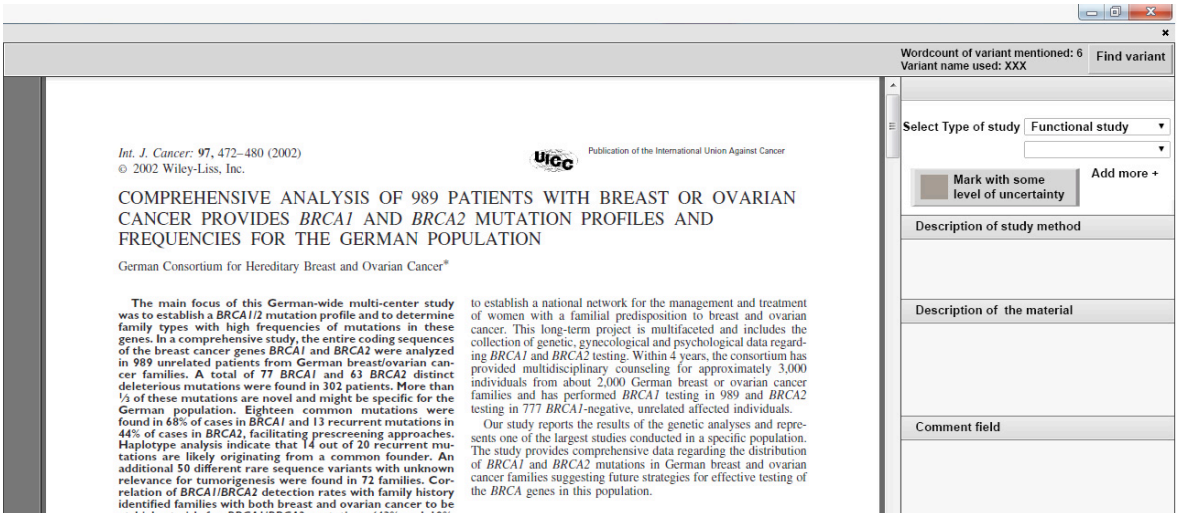


Figure 5. A prototype that displays the selected article and the comment field used for evaluation of the article.

3. DISCUSSION AND RESULTS

The research in this study highlights the importance of providing system support for multiple user strategies and drawing upon expert knowledge through user involvement. From user sessions, it was evident that users valued that the system supported their current workflow. This is in line with findings from (Shyr et al., 2014).

The importance of user involvement in the development of clinical decision support tools is also evident in how local work practices are often unique. Lindgren argues that *"the organization of clinical practice differs between clinics and countries. Local routines, work division, amount and characteristics of teamwork, etc., affect who may benefit from the support provided by a clinical decision support. Such factors need also to be taken into account when the user environment is assessed, and requirements for a CDSS* (Clinical Decision Support System) *are formulated"*, (Lindgren, 2011, p. 129). Our research also reflects numerous characteristics of local work practices and how the system can benefit from understanding them. The local collaboration that consists of verbal discussions is not currently supported by the system. Many users did not see the benefit in supporting online discussions since they had a close physical proximity. One user's strategy for limiting the need for verbal discussions was to mark uncertainties in the text. This user's solution based on highlighting in gray statements that had some level of uncertainty could be classified as an awareness-making mechanism and could be viewed and implemented into the new system as a support for collaboration.

The users were positive towards a support for collaborative analyzes. This is consistent with Shyr et al. findings that *"users expressed that an ideal system would allow users to attach notes, links to scholarly articles, as well as comments on individual genes or genetic variations, and that such information be available to multiple users in the same clinical setting. Software that empowers collaborative analysis would be well received."* (Shyr et al., 2014, p. 134)

Users handle different gene variant classification cases by deploying diverse user strategies. These strategies should be addressed in the design of the system to accommodate the need for presenting the right information to the user at the time of decision-making. Some of the strategies are built upon individual preferences, while others are brought on by the different demands from variant classification cases. As the sequencing technology and its uses is advancing and increasing the workload, most likely the user strategies will alter. Strategies are especially important to handle special variant classification cases. One of the users addresses this by stating that the "*system has to be flexible*." This is consistent with Shyr et al.'s warning that "*there are unique cases, which require unusual analysis approaches. Therefore while the software should be structured around specific standard analysis models, it needs to remain flexible*" (Shyr et al., 2014, p. 134).

During the course of this study, we focused systematically on applying the user-centered design approach and methods. These were an aid in maneuvering this complex research domain and helped handle the large amount of usability issues uncovered in user sessions.

The focus of this research and the identification of recurring design issues and themes were not about making generalizations, but exploring how user studies can be applied in the complex domain of bioinformatics. The focus has been on understanding the design context through the eyes of the users. In this research, the document containing the user-developed strategy for evaluating research references was a good source for understanding how the users make sense of the scientific research articles. This was used, in combination with all the other research data collected, to discover and understand how users make meaning related to different elements from research articles contributing to local genetic knowledge production.

Through users involvement in the system development process, we identified a total of 24 design issues and themes related to handling of literature references. Our research re-confirmed the importance of triangulation, i.e. the use of different methods in the study. We could be more confident in our findings when two or more methods gave the same results. In addition, the use of diverse methods helped uncover interesting data that would otherwise remain hidden. For example, without observing users while they performed the genetic classification, some data, as well as findings, would have been missed. Users could not articulate precisely what it is that they do when searching for variants in the literature. What they said that they did, and what they actually did, differed (a say/do problem, see (Simonsen and Kensing, 1998)). For instance, during interviews they said that they highly value Google search by the variant's name. However, a single variant frequently has several different names. Observations provided a possibility to see how analysts searched repeatedly through a set of articles using diverse names for the variant in question. Furthermore, when an article containing the correct variant was found, its PDF was searched in order to check the number of variant's occurrences. The perceived relevance of the article was proportional to the number of times the variant occurred in the article. The need for these two manual processes of getting information (variant name used and the number of its occurrences) from articles were perceived as ineffective, but since users valued the functionality highly for other reasons, the ineffectiveness was not mentioned. This provided a design opportunity, where a modified Google search displaying all names of the variant and the number of their occurrences was made. Thus, the time spent on article evaluation was reduced. The new system added value to an already popular functionality and contributed to increased efficiency of the article evaluation process for the analysts.

4. CONCLUSION

The findings of this study indicate that user-centered design can be a good way of overcoming some usability challenges when working in complex domains. By including users, issues related to human-to-human interactions and collaborations also become visible. Thus, the chances of designing a system that provides wider and better support to analysts increases. The application of user-centered methods revealed how users contributed with valuable input for the design of the future system. Such rich input could hardly be gathered in other ways, e.g., studying workflow charts. Understanding the ecology of the system, all relations between technology and people needed to be considered and understood. Placing the analysts in the center, however, helped to adjust the focus on human productivity regarding the support for accuracy and speed of assessment. The case of reference management hopefully illustrates well these points.

ACKNOWLEDGEMENT

This research was part of the genAP project (Norwegian genetic Analysis Platform). Our thanks are due to all lab engineers and physicians at the Department of Medical Genetics, Oslo University Hospital who participated in this study.

REFERENCES

Bolchini, D. et al., 2009. Better bioinformatics through usability analysis. *Bioinformatics*, 25(3), pp. 406–412.

Børsting, J., 2014. Design of Genetic Classification Software: The Case of Representation of Research References. Available at: https://www.mn.uio.no/ifi/forskning/grupper/design/undervisning/master-projects/borsting-master.pdf.

Boyd, S.D., 2013. Diagnostic Applications of High-Throughput DNA Sequencing. *Annual Review of Pathology: Mechanisms of Disease*, 8(1), pp. 381–410.

Chilana, P.K., Wobbrock, J.O. and Ko, A.J., 2010. Understanding Usability Practices in Complex Domains. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. New York, NY, USA: ACM, pp. 2337–2346.

Dienstmann, R. et al., 2014. Standardized decision support in next generation sequencing reports of somatic cancer variants. *Molecular Oncology*, 8(5), pp. 859–873.

Eike, M.C. et al., 2014. GenAP workbench: aiding variant classification in clinical diagnostic settings. In *American Society of Human Genetics Annual Meeting*. American Society of Human Genetics. San Diego.

Gould, J.D. and Lewis, C., 1985. Designing for Usability: Key Principles and What Designers Think. *Commun. ACM*, 28(3), pp. 300–311.

Javahery, H., Seffah, A. and Radhakrishnan, T., 2004. Beyond Power: Making Bioinformatics Tools User-centered. *Commun. ACM*, 47(11), pp. 58–63.

Kensing, F., Simonsen, J. and Bodker, K., 1998. MUST: A Method for Participatory Design. *Human–Computer Interaction*, 13(2), pp. 167–198.

Lindgren, H., 2011. Towards personalized decision support in the dementia domain based on clinical practice guidelines. *User Modeling and User-Adapted Interaction*, 21(4-5), pp. 377–406.

Mardis, E.R., 2010. The $1,000 genome, the $100,000 analysis? *Genome Medicine*, 2(11), p.84.

McKenna, A. et al., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), pp. 1297–1303.

Neri, P.M. et al., 2012. Usability of a novel clinician interface for genetic results. *Journal of Biomedical Informatics*, 45(5), pp. 950–957.

Plon, S.E. et al., 2008. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human Mutation*, 29(11), pp. 1282–1291.

Schadt, E.E. et al., 2010. Computational solutions to large-scale data management and analysis. *Nature Reviews Genetics*, 11(9), pp. 647–657.

Shyr, C., Kushniruk, A. and Wasserman, W.W., 2014. Usability study of clinical exome analysis software: Top lessons learned and recommendations. *Journal of Biomedical Informatics*, 51, pp. 129–136.

Simonsen, J., and Kensing, F., 1998. Make Room for Ethnography in Design!: Overlooked Collaborative and Educational Prospects. SIGDOC Asterisk J Comput Doc 22, pp. 20–30.