

Interactive Clinical Pedigree Visualization Using an Open Source Pedigree Drawing Engine

João Miguel Santos¹(✉), Beatriz Sousa Santos^{1,2},
and Leonor Teixeira^{2,3}

¹ DETI/Department Electrónica, Telecomunicações e Informática,
Universidade de Aveiro, Aveiro, Portugal
{miguel santos, bss}@ua.pt

² IEETA/Instituto de Engenharia Electrónica e Telemática,
Universidade de Aveiro, Aveiro, Portugal

³ DEGEI/Department Economia, Gestão e Engenharia Industrial,
Universidade de Aveiro, Aveiro, Portugal
lteixeira@ua.pt

Abstract. Advances in Genetics have revealed that many diseases are related to genetic factors. In this context, family health histories play an increasingly important role in healthcare, aiding practitioners in the diagnosis, risk assessment and treatment of various conditions. The clinical pedigree, a graphic representation combining family structure and clinical information, is a well-accepted tool to represent family health histories. At present this tool remains underused, possibly due to the lack of pedigree management tools in health information systems. OntoFam addresses this problem by offering a clinical pedigree information system that can be integrated with current health information systems. This paper presents the method used to create OntoFam's interactive pedigree visualization by wrapping an existing open source pedigree drawing engine. The resulting environment allows practitioners to interactively view, create and manipulate pedigrees. This paper also describes the evaluation strategy that was developed to assess the system and includes its preliminary results.

Keywords: Family health history · Clinical family history · Pedigree · Health information system · Electronic health record · Ontofam · Hemophilia care

1 Introduction

The utility of Family Health Histories (FHHs) has been widely recognized, and is increasing as more and more diseases are linked to genetic factors [1, 2]. FHHs, also known as Clinical Family Histories, help practitioners in the diagnosis and risk assessment of patients and their family members in a cost-effective manner; provide insight as to the most adequate course of treatment for patients [2–4]; and enable the identification of at-risk individuals before diseases manifest, opening the possibility for counselling and preventive measures to delay, diminish or completely avoid diseases or symptoms [5].

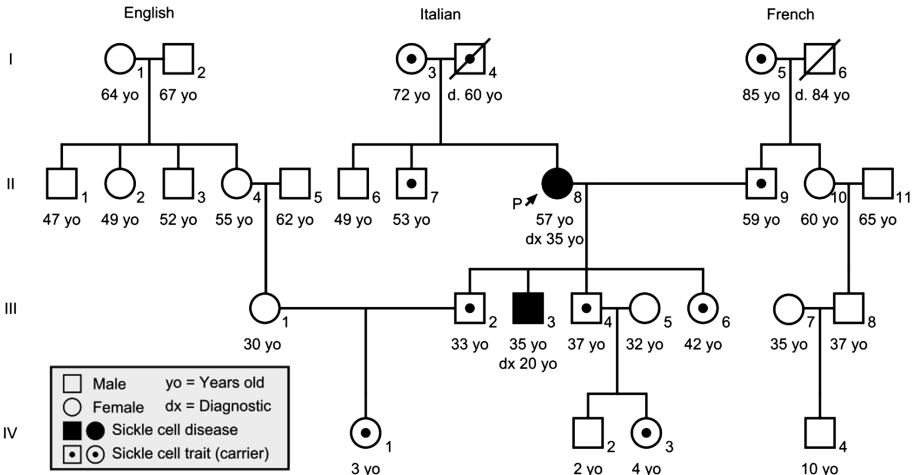


Fig. 1. Example pedigree in PSWG notation

There are several formats available to represent FHHs, both in text and graphical form [6–8]. The clinical pedigree, henceforward referred to as pedigree, is a broadly used graphical representation that combines genealogical and clinical information and offers a good balance of simplicity, expressivity and detail [4]. It is subject to standardization by the Pedigree Standardization Work Group (PSWG) that regularly issues recommendations regarding the use of symbols, annotations and layout of pedigrees [9, 10]. Figure 1 presents an example pedigree in PSWG notation, regarding a family with a history of sickle cell disease.

Several software packages allow the creation of such pedigrees [11–24], yet few integrate with existing Health Information Systems (HIS) or Electronic Health Records (EHR). Using different, disconnected tools for recording patient data and FHHs not only makes information less readily available, but also potentially results in out-of-date information and incoherent data. This problem is addressed by OntoFam [25], an ontology-based clinical pedigree information system that can be integrated with existing HIS.

During the conception of OntoFam, the team was confronted with the decision of whether to build or reuse a pedigree drawing engine (PDE) for the system’s user interface. An investigation revealed that several open-source PDEs were available for reuse, though none inherently offered the desired degree of interactivity. In fact, most existing engines follow a “static” approach: they accept an input in a structured text format and output a graphic pedigree representation; any editing involves constructing a new input to obtain an updated pedigree. However, by wrapping an existing static PDE with additional functionality, it is possible to create an interactive visualization. This was the approach chosen for OntoFam.

This paper presents OntoFam’s interactive pedigree visualization, which allows the dynamic creation and management of pedigrees. It reuses an existing static pedigree drawing engine and extends it to allow end users to directly manipulate patient

information and family structure. Section 2 presents the method used to construct the visualization and Sect. 3 describes the evaluation strategies and presents the corresponding results. Section 4 concludes this paper.

2 Method

Madeline 2.0 Pedigree Drawing Engine [18] is the chosen PDE for the basis of OntoFam's pedigree visualization tool due to a unique combination of favorable characteristics:

- It generates pedigrees that conforms to most PSWG recommendations;
- It is able to correctly represent complex family structures that some other engines do not handle [26];
- It generates pedigrees in Scalable Vector Graphics (SVG), an XML (eXtensible Markup Language) based format that is rendered as vector graphics in modern browsers and is editable in external software, such as Inkscape, CorelDRAW or Adobe Illustrator;
- It is an open-source project, allowing aspects of the representation to be extended or fine-tuned, should the need arise;
- It is actively maintained (the most recent version is from mid-August, 2014).

In its current form, Madeline is a “*non-interactive program that is executed from a shell environment*” [27], that is, it is invoked through the command-line and accepts parameters to customize the output. Input data must be formatted in one of the several table-based formats allowed. Upon invocation, Madeline renders the corresponding pedigree onto a file in SVG format. Any changes to the rendered pedigree involve either editing the SVG in a compatible image editor or altering the input data and repeat the pedigree generation process.

A separate project that shares some of Madeline's authors, dubbed Madeline 2.0 Web Service, does allow pedigrees to be built interactively [28]. However, at present this service is closed-sourced, does not fully exploit Madeline's capabilities and requires direct manipulation of the inherent input data-table for relatively trivial tasks (such as identifying the proband, setting deceased states of specifying clinical conditions, among others).

OntoFam uses a different approach, in which the data-table input is completely hidden from the end user. The intent is to provide a user interface that is closer to practitioners' mindset, allowing data to be manipulated in a way that conforms to their needs and view-points, rather than those of the pedigree drawing engine.

The fact that Madeline uses SVGs was particularly relevant in the construction of OntoFam's visualization method. SVG is an XML language that uses elements to represent lines and shapes that, once rendered on a web browser, allows client-side script to interact with the elements and corresponding shapes. This enables end users to interact with pedigree elements and provides immediate solutions to the needs of selecting or activating represented individuals – client scripts react to the user action of clicking an individual by changing its visual aspect (i.e., changing the CSS style of the

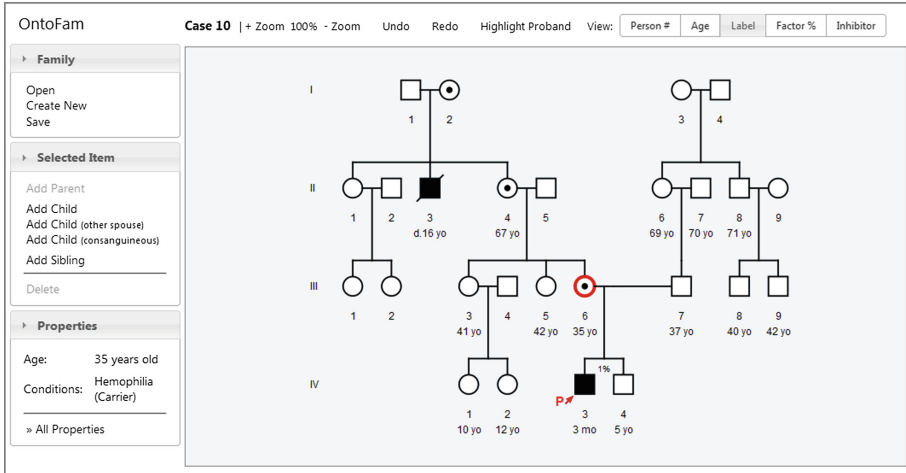


Fig. 2. OntoFam’s user interface editing a 4-generation family pedigree

corresponding SVG element) and display a pop-up window with detailed patient information when an individual is double-clicked.

To add new individuals to the pedigree, menu options are provided to allow the insertion of relatives to a selected individual. For example, upon selecting an individual whose ancestry is not yet filled in, the interface will provide options to add the corresponding mother and father. Spouses, siblings and offspring can be added to individuals at any time. This “chained” insertion of individuals ensures that the family structure is always valid. Figure 2 presents OntoFam’s user interface while editing a pedigree with 4 generations, and Fig. 3 exemplifies an individual’s popup form, allowing its information to be viewed and edited.

OntoFam consists of a web application that uses single-page application (SPA) techniques to wrap Madeline’s invocations and provide an interactive experience. Whenever the user requests the pedigree information to be changed, the following process occurs:

1. The user requests the addition, edition or deletion of an individual;
2. Client-side script validates the user action and sends the corresponding information to the web server;
3. OntoFam’s family database is updated with the new information;
4. The current FHH is formatted according to Madeline’s requirements;
5. Madeline is invoked and an SVG pedigree is generated;
6. The SVG is post-processed on the server;
7. The post-processed SVG is returned to the client and rendered on the browser, replacing the previous pedigree representation.

Step 6 is used to fine-tune visual aspects and overcome Madeline’s limitations, such as the lack of individual and generation sequential numbering. The same effect could have been achieved by extending Madeline’s source code, but the team found preferable to

The screenshot shows the 'Person Properties' window in the OntoFam application. The window is titled 'Person Properties' and has a close button (X) in the top right corner. The main content area contains the following fields and controls:

- Process Number:** 1356634 (text input) with a link 'Import data from Hemo@Care'.
- Name:** John Smith (text input).
- Label:** (empty text input).
- Gender:** Male (dropdown menu).
- Birthdate:** 1946-01-01 (text input) with a checked **Deceased** checkbox and a **Death Date:** 1962-01-01 (text input).
- Study:** Studied Patient Proband (radio buttons).
- Diagnosis:** Affected (dropdown menu).
- Condition:** Hemophilia (dropdown menu).
- Factor %:** (empty text input) with an **Inhibitor Present** checkbox.
- Notes:** (empty text area).

At the bottom right of the window are 'Save' and 'Close' buttons. On the left side of the main window, there is a sidebar with the following sections:

- Family:** Open, Create New, Save.
- Selected Item:** Add Parent, Add Child, Add Child (other sp), Add Child (consang), Add Sibling, Delete.
- Properties:** Name: John, Age: Dece years, Conditions: Hem. Below this is a link '> All Properties'.

Fig. 3. Example popup form displaying the properties of an individual

use Madeline as-is and treat it as a “black-box”. This illustrates a useful characteristic of SVGs: because they are XML documents, it is fairly easy to process and manipulate an existing graphic, unlike raster-based image formats.

Wrapping Madeline in this manner also allows OntoFam to replace its PDE with relative ease, should the need arise. Assuming the replacement engine produces comparable SVGs, the system changes will be limited to the translation of FHH information into the format accepted by the new engine.

Even though the process involves a number of steps, response times are adequate and users are not aware of delays caused by server roundtrips. The XMLHttpRequest (XHR) Application Programming Interface (API) is used in the background to exchange data with the server and, on step 7, user selections are preserved and scroll positions are maintained to ensure that, from the end user’s point of view, the pedigree representation was “instantly” complemented with the new information.

3 Evaluation and Results

OntoFam is currently being integrated with an existing HIS for Hemophilia care – Hemo@Care [29]. Hemophilia is an X-linked congenital bleeding disorder, with an estimated frequency of 1 in 10,000 births [30]. Because Hemo@Care deals with a hereditary condition and is also developed in the University of Aveiro, it is an excellent candidate for integration with OntoFam. However, the fact that Hemophilia is a rare disease poses some challenges when it comes to evaluating the resulting system, namely the limited number of practitioners knowledgeable in Hemophilia care. Thus, it proved necessary to use evaluation strategies that could be applied with a limited

number of participants. The chosen approach involves three distinct evaluation techniques, executed at different development stages:

- Use of an iterative development process, resorting to expert reviews to validate the most significant iterations and guide the forthcoming development;
- Conduction of user experience tests with (necessarily few) real users (practitioners knowledgeable in Hemophilia care) at project milestones, to assess overall system acceptance in real-world scenarios;
- Conduction of usability tests with non-expert users in greater numbers (students of Human-Computer Interaction [HCI] courses), to fine-tune the user interface and possibly detect generic (non-domain) usability problems.

Details and results of each technique are described in the following subsections.

3.1 Expert Reviews

A number of reasons determined the use of expert reviews to guide the development of the visualization. As previously mentioned, the number of available practitioners is very limited, due to the disease's rarity. Also, at early exploratory stages of development the user tasks are not always completely understood, hence controlled experiments with concrete tasks do not yet apply [31]. Finally, the team was interested in understanding what high-level cognitive tasks, such as decision-making and insight generation, are involved when using a pedigree management system in the context of Hemophilia care.

Two experts were involved in the periodic reviews: an HCI expert and a medical doctor with ample experience in Hemophilia care. The HCI expert relied on heuristic evaluation and personal experience to detect problems and validate design choices, while the medical doctor acted as a test user and, most importantly, enlightened the team regarding what low and high-level tasks were deemed necessary in her field. For example, it was revealed that women in reproductive age within a family with a history of Hemophilia should be educated about the possible implications in offspring, thus gender and age should be clearly represented in the visualization. The presence of inhibitors is also of great importance; therefore it should be visible on the pedigree. It was also learned that due to the X-linked nature of the disease, a pedigree can be used as a "map" to differentiate individuals that should be studied from those that can do without study (and thereby spare human and financial resources).

3.2 User Experience Tests

When expert reviews determined that the system was useable and feature-complete, the team proceeded to conduct user experience tests. The objectives of such tests include obtaining "*subjective feedback and opinions on the visualization tool*" [32] and assess the system adequacy in real-world scenarios. This is not a "final test", but rather an iterative process, where the results of a test session influence the next iterations of the development process and the design of the following test session.

Thus far, the team has conducted the first round of these tests. Two medical doctors knowledgeable in Hemophilia care were invited to participate, guided by three observers. After an open exploratory conversation, where the utility of pedigrees within the Hemophilia field was discussed, they were given a 10 min presentation on the system goals and functionality. They were then allowed to freely explore the system for another 10 min, where they created fictional test pedigrees. The team had prepared a set of introductory tasks to ensure familiarity with the system functionality; however, the participants found their way around the user interface with great ease and the observers were unanimous in deeming these tasks unnecessary.

The participants proceeded to the second set of tasks, which involved constructing a 4-generation pedigree with 26 individuals, similar to that represented on **Fig. 2**; analyzing existing pedigrees to detect individuals that should be studied and individuals that were unnecessarily studied; and analyzing an incorrect pedigree to detect errors, taking into account the known hereditary patterns of Hemophilia. Upon completing each task, participants were asked to rate its difficulty using a Likert-like scale. After performing all tasks, the participants were asked to answer a questionnaire to evaluate system quality and adequacy.

All tasks were correctly performed by both participants, with the exception of the pedigree construction. On this task, one of the participants confused the 4th generation branches and constructed an incorrect family structure. Nearing the end of the task the participant realized this problem but was told not to correct it, seeing as it was the system that was being evaluated, and not the user. This error suggests that functionality to indicate an individual's relation to the proband may be a useful addition to the user interface. Both users rated all tasks as easy or very easy, and all tasks were completed within the estimated completion time.

The answers to the questionnaire were very positive, with both participants agreeing that the system is adequate to represent FHHs related to Hemophilia care, and that they would both use the system and recommend it to colleagues (one participant mentioned that she would also like to use the system applied to the Thrombophilia field). When asked to rate the overall system on a scale of 1-10 (greater is better), one participant awarded an 8 and the other awarded a 9-10. No major problems with the system were found by the participants, though both provided suggestions for further development: the inclusion of molecular study information (on the patient data but not necessarily on the pedigree) and the possibility to represent adoptions (not completely necessary when dealing with hereditary diseases, but relevant nonetheless to get an holistic view of families with adopted members).

A second round of user experience tests is scheduled to occur shortly following the time of this writing, with participants distinct from those of the first round.

3.3 Usability Tests

Usability tests with a much larger number of participants are planned for the 1st semester of 2015. These will consist of controlled experiments performed by students of Human-Computer Interaction (HCI) courses in the University of Aveiro. Although these users are not domain experts, their familiarity with user interfaces and usability

guidelines is likely to allow the detection of generic usability problems. User performance metrics, such as task completion time, error ratio and number of clicks, among others, may also help improve interface design.

4 Conclusions and Future Work

The method used to create OntoFam's interactive pedigree visualization by wrapping an existing pedigree drawing engine has proved successful. It hides the process of constantly feeding the PDE with new input to generate new output, and instead presents users with an environment where pedigrees can be directly manipulated. Even though the process involves server roundtrips and constant pedigree regenerations, the fact that these occur in the background and that the visual context is maintained (such as selection and scroll position) keeps users unaware of the actual complexity. In their view, the pedigree is instantly updated on the client browser whenever changes are made.

The results obtained in the concluded evaluation stages have been favorable, with comments and suggestions mostly pertaining to additional functionality rather than faulting the existing system. Results also suggest that practitioners already familiar with clinical pedigrees will have little or no trouble adapting to the visualization. Even though these are promising results, it must be noted that, thus far, the number of participants has been very limited and that broader evaluations should be performed. This will be the focus of forthcoming evaluation stages.

Acknowledgments. The authors would like to thank Dr.^a Natália Martins, Dr.^a Lúcia Borges and Dr.^a Marina Costa for their collaboration and valuable suggestions and observations.

References

1. Kmiecik, T., Sanders, D.: Integration of Genetic and Familial Data into Electronic Medical Records and Healthcare Processes (2009). www.surgery.northwestern.edu/docs/KmiecikSandersArticle.pdf. Accessed 13 November 2012
2. Rich, E.C., Burke, W., Heaton, C.J., Haga, S., Pinsky, L., Short, M.P., Acheson, L.: Reconsidering the family history in primary care. *J. Gen. Intern. Med.* **19**, 273–280 (2004)
3. Morales, A., Cowan, J., Dagua, J., Hershberger, R.E.: Family history: an essential tool for cardiovascular genetic medicine. *Congestive Heart Fail.* (Greenwich, Conn.) **14**, 37–45 (2008)
4. Bennett, R.L.: *The Practical Guide to the Genetic Family History*, 2nd edn. Wiley, New York (2010)
5. Frezzo, T.M., Rubinstein, W.S., Dunham, D., Ormond, K.E.: The genetic family history as a risk assessment tool in internal medicine. *Genet. Med.* **5**, 84–91 (2003)
6. Butler, J.F.: The Family Diagram and Genogram: Comparisons and Contrasts. *The American Journal of Family Therapy* **36**, 169–180 (2008). 07 May 2008
7. Rempel, G.R., Neufeld, A., Kushner, K.E.: Interactive use of genograms and ecomaps in family caregiving research. *J. Fam. Nurs.* **13**, 403–419 (2007)

8. American College of Obstetricians and Gynecologists: Committee opinion no. 478: family history as a risk assessment tool. *Obstet. Gynecol.* **117**, 747–750 (2011)
9. Bennett, R.L., Steinhaus, K.A., Uhrich, S.B., O’Sullivan, C.K., Resta, R.G., Lochner-Doyle, D., Markel, D.S., Vincent, V., Hamanishi, J.: Recommendations for standardized human pedigree nomenclature. *J. Genet. Couns.* **4**, 267–279 (1995)
10. Bennett, R.L., French, K.S., Resta, R.G., Doyle, D.L.: Standardized human pedigree nomenclature: update and assessment of the recommendations of the national society of genetic counselors. *J. Genet. Couns.* **17**, 424–433 (2008)
11. Agarwala, R., Biesecker, L.G., Hopkins, K.A., Francomano, C.A., Schaffer, A.A.: Software for constructing and verifying pedigrees within large genealogies and an application to the old order Amish of Lancaster County. *Genome Res.* **8**, 211–221 (1998). 1 March 1998
12. Dudbridge, F., Carver, T., Williams, G.W.: Pelican: pedigree editor for linkage computer analysis. *Bioinformatics* **20**, 2327–2328 (2004). 22 September 2004
13. HaploPainter Team. HaploPainter Manual 1.0 (2012). http://haplopainter.sourceforge.net/HaploPainterManual_V1.0.pdf, 08 December 2012
14. He, M., Li, W.: PediDraw: a web-based tool for drawing a pedigree in genetic counseling. *BMC Med. Genet.* **8**, 31 (2007)
15. Hughes Risk Apps. Hughes Risk Apps User Guide (2009). <http://www.hughesriskapps.com/Files/UserGuide-26x.aspx>, 08 December 2012
16. Lee, W.J., Pollin, T.I., O’Connell, J.R., Agarwala, R., Schaffer, A.A.: PedHunter 2.0 and its usage to characterize the founder structure of the Old Order Amish of Lancaster County. *BMC Med. Genet.* **11**, 68 (2010)
17. Makinen, V.-P., Parkkonen, M., Wessman, M., Groop, P.-H., Kanninen, T., Kaski, K.: High-throughput pedigree drawing. *Eur. J. Hum. Genet.* **13**, 987–989 (2005)
18. Trager, E.H., Khanna, R., Marrs, A., Siden, L., Branham, K.E.H., Swaroop, A., Richards, J.E.: Madeline 2.0 PDE: a new program for local and web-based pedigree drawing. *Bioinformatics* **23**, 1854–1856 (2007). 15 July 2007
19. Brun-Samarçq, L., Gallina, S., Philippi, A., Demenais, F., Vaysseix, G., Barillot, E.: CoPE: a collaborative pedigree drawing environment. *Bioinformatics* **15**, 345–346 (1999). 1 April 1999
20. Cohn, W.F., Ropka, M.E., Pelletier, S.L., Barrett, J.R., Kinzie, M.B., Harrison, M.B., Liu, Z., Miesfeldt, S., Tucker, A.L., Worrall, B.B., Gibson, J., Mullins, I.M., Elward, K.S., Franko, J., Guterbock, T.M., Knaus, W.A.: Health Heritage ©, a web-based tool for the collection and assessment of family health history: initial user experience and analytic validity. *Public Health Genomics* **13**, 477–491 (2010)
21. Wong, L.: Visualization and manipulation of pedigree diagrams. *Genome Inform Ser Workshop Genome Inform* **11**, 63–72 (2000)
22. U.S. Department of Health & Human Services. My Family Health Portrait (2009). <http://familyhistory.hhs.gov/>. Accessed 12 September 2014
23. Zhao, J.H.: Pedigree-drawing with R and graphviz. *Bioinformatics* **22**, 1013–1014 (2006). Accessed 15 April 2006
24. Progeny Software, Progeny Clinical, version 8.0, (2011). <http://www.progenygenetics.com/clinical/>
25. Santos, J.M., Sousa Santos, B., Teixeira, L.: Using ontologies and semantic web technology on a clinical pedigree information system. In: Duffy, V.G. (ed.) DHM 2014. LNCS, vol. 8529, pp. 448–459. Springer, Heidelberg (2014)
26. Trager, E.H., Khanna, R., Marrs, A.: Comparison of Pedigree Drawing Programs (2011). <http://eyegene.ophthy.med.umich.edu/madeline/comparisons/>. Accessed 20 December 2014

27. Trager, E.H., Khanna, R., Marrs, A.: Madeline 2.0 Pedigree Drawing Engine Documentation (2011). <http://eyegene.ophthy.med.umich.edu/madeline/documentation.php>. Accessed 20 December 2014
28. Khanna, R., Trager, E.H.: Madeline 2.0 Public Web Service (2007). <http://eyegene.ophthy.med.umich.edu/madeline/webservice.php>. Accessed 1 October 2014
29. Saavedra, V., Teixeira, L., Ferreira, C., Sousa Santos, B.: A preliminary usability evaluation of Hemo@Care: a web-based application for managing clinical information in hemophilia care. In: Kurosu, M. (ed.) HCD 2009. LNCS, vol. 5619, pp. 785–794. Springer, Heidelberg (2009)
30. World Federation of Hemophilia. Guidelines for the Management of Hemophilia (2nd ed.) (2012). <http://www.wfh.org/en/resources/wfh-treatment-guidelines>
31. Tory, M., Moller, T.: Evaluating visualizations: do expert reviews work? *IEEE Comput. Graph. Appl.* **25**, 8–11 (2005)
32. Isenberg, T., Isenberg, P., Jian, C., Sedlmair, M., Moller, T.: A systematic review on the practice of evaluating visualization. *IEEE Trans. Vis. Comput. Graph.* **19**, 2818–2827 (2013)