Comprehensive Visualization of Temporal Patient Data for the Dermatological Oncological Tumor Board

- Detailed Results of the Qualitative Evaluation Study -

To ensure the quality of our prototype, we conducted a qualitative questionnaire to collect conclusions about its clinical relevance for the tumor board, the effectiveness of the applied visualizations, and its usability, with regard to an understandable design and minimal interactions. The questionnaire consists of 5 inquiring participant details, 8 tasks with respective solutions (correct, incorrect) and 27 closed-ended questions on a five-point Likert scale to evaluate different aspects of relevance, effectiveness and usability (--, -, o, +, ++) using one run based on one patient case. The questionnaire is based on the usability principles proposed by Forsell and Johansson [1] and the seven scenarios for the evaluation of information visualizations according to Lam et al. [2]. Five domain experts in dermatology (three female and two male) filled out the questionnaire, whose level of experience is between 2 and 10 years. One of the participants is co-author of this work. In the following, we provide their individual accuracy for solving the tasks and the answered Likert scores together with the number of experts, e.g. S(correct) = 5, means that all experts solved the tasks correctly and S(++) = 5 means that all experts gave the score ++ for the accuracy of a question.

Correctness of Fulfilling Tasks. The eight tasks serve to assess the comprehensibility of the visual elements and their functions (Figure 1). The determination of the time of change in the clinical staging (S(correct) = 5) or the first appearance of metastases (S(correct) = 5) was performed by all participants correctly. However, the identification of the periods of critical parameter values (LDH or S100) was answered differently (S(correct) = 2, S(false) = 3), which could be due to the different interpretation of the line graphs, as the critical red line starts just above the standard range, while the observed values occure later. However, it cannot be excluded that elevated values have already occurred between the two observations. The naming of certain therapies applied to the patient (S(correct) = 3, S(false) = 2), the tumor stage at the time of the diagnosis (S(correct) = 5) and the presence of affected lymph nodes (S(correct) = 5) was mostly determined correctly. Furthermore, the investigation of the first and last therapy applied (S(correct) = 5) and the one in which the ECOG stage increased for the first time (S(correct) = 5) resulted in agreement among all respondents.

Evaluation of the Usability Principles. Forsell and Johansson [1] have established ten usability principles serving to evaluate the design and functions of visualizations under predefined guidelines. For each principle, we have considered specific questions that represent its core message (Figure 2). The first principle of *information coding* concerns the truthful mapping of data to visual elements. Most of the questions representing this principle regarding truthfulness (S(++) = 3, S(+) = 2) and learning effort (S(++) = 3, S(o) = 2) were found to be accurate. The principle of *minimal actions* determines the minimum number of actions required to achieve a task. The majority tended to agree with this regarding the prototype (S(++) = 2, S(+) = 2, S(o) = 1). Furthermore, the principle of *flexibility* indicates the availability of different ways possible to achieve a certain goal. Thus, individual adaptation with regard to certain habits and tasks can be ensured. This was stated controversially as two experts doubted the compliance of this principle (S(++) = 2, S(+) = 1, S(-) = 2), but partly because they confirmed that “they are no visual communication types”. For the majority the

tool, moreover, it provides sufficient supporting functions regarding *orientation and help* (S(++) = 2, S(+) = 2, S(o) = 1). The principle of *spatial organization* refers to the legibility and distribution of visual elements. Only one of the respondents found the localization of the elements difficult (S(++) = 4, S(-) = 1). In addition, *consistency* in operation and design was confirmed by all experts (S(++) = 4, S(+) = 1). The principle *recognition rather than recall* is based on the fact that the user has to remember as little information as possible in order to perform tasks. Only one respondent rated this as neutrally (S(++) = 3, S(+) = 1, S(o) = 1). For the purpose of good user guidance, showing all permissible actions was rated good (principle of *prompting*) (S(++) = 3, S(+) = 2). One respondent criticized the presence of superfluous information in respect of the principle *remove the extraneous* (S(++) = 2, S(+) = 2, S(--) = 1). Following the last principle of *data set reduction*, subsets of the data sets can be easily focused in the prototype (S(++) = 3, S(+) = 2).

Evaluation Based on the Seven Scenarios. Since two of the seven scenarios by Lam et al. [2] focus more on the performance of either the visualization algorithms or the user, only five of the seven scenarios were taking into account for evaluating the visual elements and functions of this tool (Figure 3). For each scenario, we considered questions representing their core message and evaluated their quality content. *The Understanding Environments and Work Practices* (UWP) scenario considers the adaptation to individual expert requirements. The adaptation of the tool to the requirements of the tumor board was approved by the majority of the respondents (S(++) = 2, S(+) = 2, S(-) = 1). The *Evaluating Visual Data Analysis and Reasoning* (VDAR) scenario was mostly assessed correctly with regard to the exploration of details (S(++) = 2, S(+) = 2, S(o) = 1) and the generation of conclusions for the clinical prognosis of the patient (S(++) = 3, S(+) = 1, S(o) = 1). Regarding the *Evaluating Communication Through Visualization* (CTV) scenario, the faster and better acquisition of the data was mostly confirmed by the experts in comparison to listening only to the physician introducing the patient (S(++) = 2, S(+) = 2, S(o) = 1). The *Evaluating Collaborative Data Analysis* (CDA) scenario supports the generation of group knowledge (e.g. tumor board recommendations) (S(++) = 3, S(o) = 2) and the communication about the data in groups of people (S(++) = 3 , S(+) = 1, S(o) = 1). The majority agreed to this requirements. The last scenario of *User Experience* (UE) refers to the subjective opinion regarding the visualization. In this sense, the score spreads over a wide range and there were some suggestions regarding the adjustment (S(++) = 1, S(+) = 2, S(-) = 2) and the addition (S(++) = 1 , S(+) = 2, S(o) = 1, S(-) = 1) of important interactions and visualizations in the prototype.

Clinical Relevance. The majority of the experts confirmed the clinical benefit of the application for tumor board meetings (S(++) = 2, S(+) = 1, S(o) = 2) (Figure 4). Especially in comparison to the conventional preparation of data (verbal presentation of important patient data), the advantages of this application were considered to be valuable (S(++) = 2, S(+) = 2, S(o) = 1). The use of the application for the preparation for the tumor board was rated to be very supportive for the dermatologists concerning the amount of work required for the preparation of patient data, but also for all other experts to get to know the patients in advance (S(+) = 4, S(o) = 1).

[1] FORSELL C., JOHANSSON J.: An heuristic set for evaluation in Information Visualization. pp. 199-206. doi:10.1145/1842993.1843029.

[2] LAM H., BERTINI E., ISENBERG P., PLAISANT C., CARPEN-DALE S.: Empirical Studies in Information Visualization: Seven Scenarios. IEEE Transactions on Visualization and Computer Graphics 18, 9 (Sept. 2012), 1520–1536. doi:10.1109/TVCG.2011.279.

Figure 1: Correctness of Fulfilling Tasks.

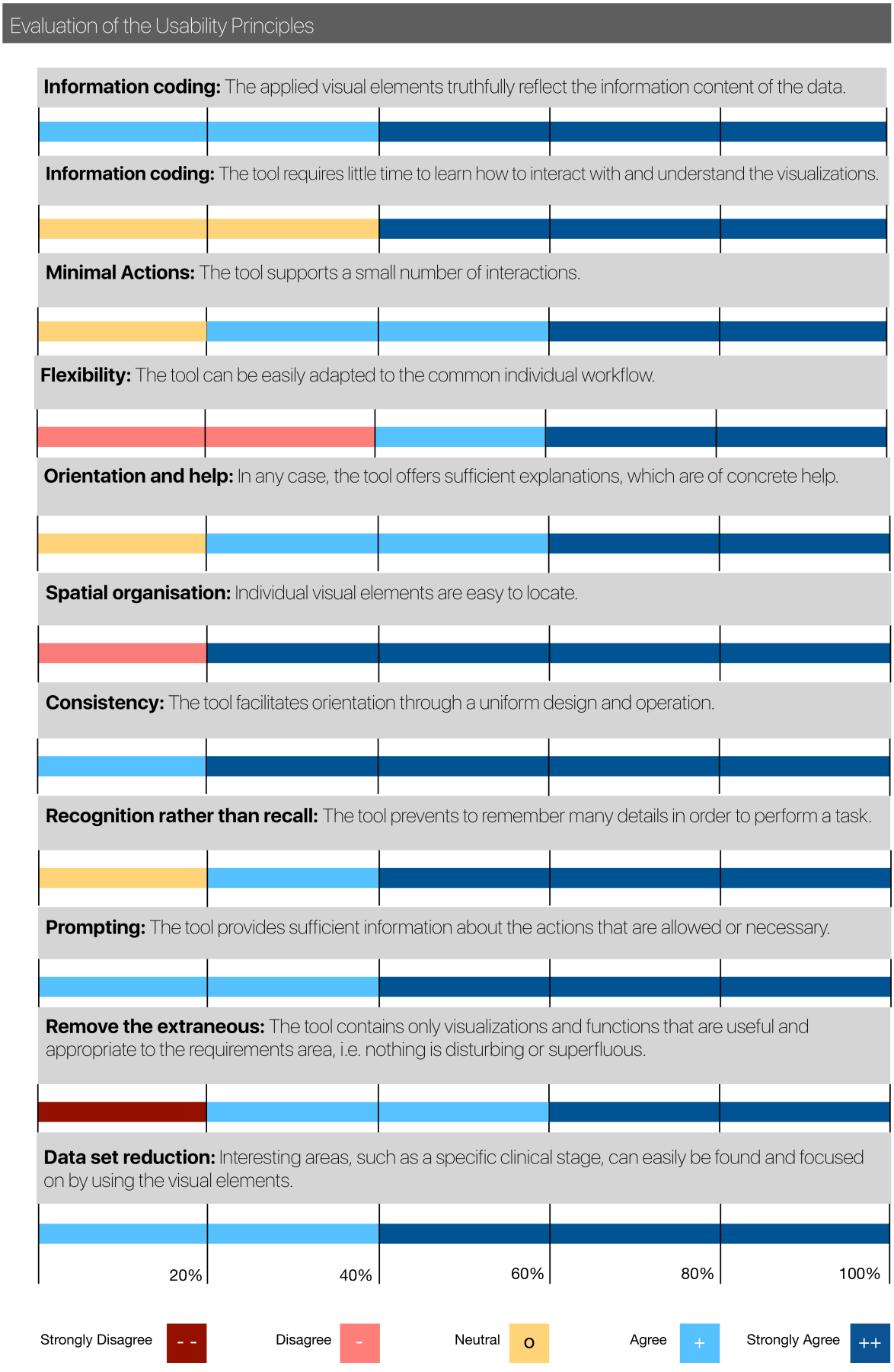

Figure 2: Evaluation of the Usability Principles.

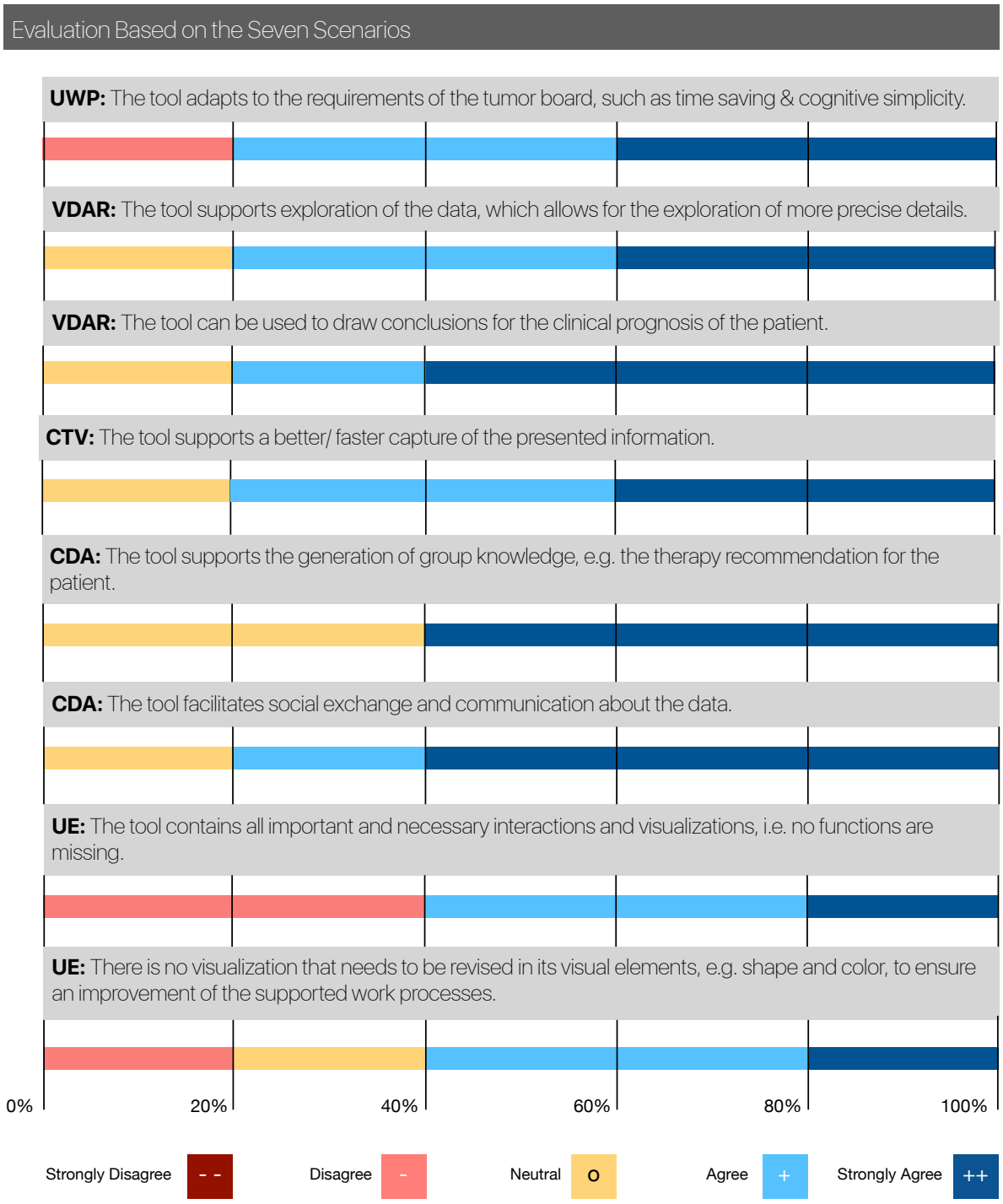

Figure 3: Evaluation Based on the Seven Scenarios.

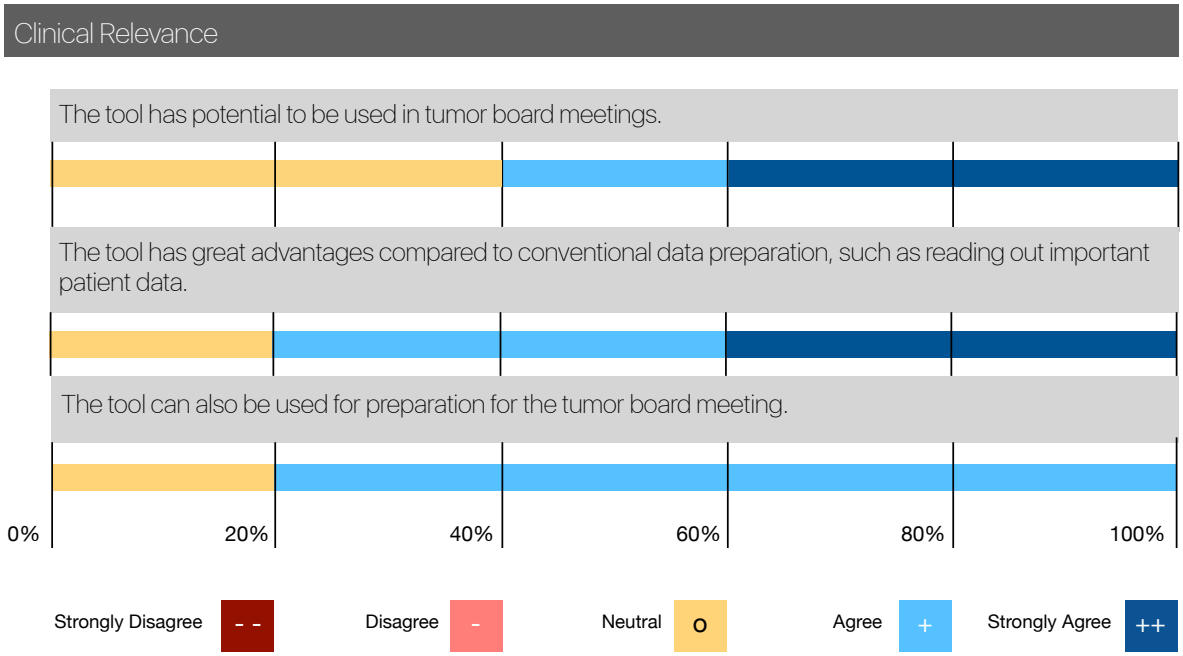

Figure 4: Clinical Relevance.