Vision, Modeling, and Visualization (2024)
L. Linsen and J. Thies (Eds.)

Raccoon: Supporting Risk Communicators in Visualizing Health Data for the Public

Anna Kleinau, Bernhard Preim, Monique Meuschke

University of Magdeburg, Germany

Abstract

The urgent need to improve health communication is highlighted by the millions of premature deaths worldwide each year due to lifestyle choices and behavioral risks. These losses reveal that researching and understanding these risks is not sufficient; we must also communicate them effectively to the public. In this paper, we discuss how we can assist experts in creating data-based risk visualizations for the general public. Our tool, RACCOON, *is able to identify and suggest the most important risk factors in a data set, visualizing them in a way that allows seamless exploration of the data set. Then, we use the latest research in risk communication, narrative visualization, and affective visualization to generate engaging visualizations for the general public. Extensive customization options allow the expert to integrate their domain knowledge, and tailor the visualizations to their data story and communicative intent. We evaluated* RACCOON *with domain experts, as well as our visualizations with the general public. The findings highlight* RACCOON*'s effectiveness in providing intuitive and engaging visualizations that appeal to a broad audience. They also provide first insights into the interplay of visualization design and communicative intent. By fusing the research fields of risk communication, narrative visualization, and affective visualization in one visualization generation tool, we provide a novel approach to support domain experts in communicating risks and risk factors to the general public.*

CCS Concepts

• ***Human-centered computing*** → ***Visualization systems and tools; Information visualization; Visual analytics;***

1. Motivation

In today's data-driven world, effective communication of insights is crucial, especially in healthcare. In 2019, smoking caused over 8.7 million deaths, dietary risks 7.9 million, and alcohol use 2.4 million [MAZ*20]. These preventable risks highlight the importance of communicating health risks to the public.

Epidemiological studies analyze the links between risk factors and diseases [RBK10]. However, results need to be simplified to be presented to a general audience. Data visualizations can improve the effectiveness of risk communication [ASKS06, TZFE*13], but can also cause misunderstandings or be too complex [TZFE*13].

With our tool, RACCOON (**R**isk f**AC**tor **CO**mmunicati**ON**), we aim to assist experts in crafting impactful visualizations for communicating risks right from the initial data. This process begins by automatically preprocessing the data and identifying the most crucial risk factors. Subsequently, RACCOON interactively creates visualizations that are not only effective but also easily comprehensible to the lay audience. In addition to incorporating current research in risk visualization, RACCOON incorporates narrative visualization to captivate [SH10], and affective visualization to resonate emotionally [LWC24] with the audience.

Recognizing that the optimal visualization often varies with the expert's objectives, we delve into how our visualizations can be tailored to meet diverse intents. Extensive customization options allow both the integration of domain-specific knowledge and personalization. Our evaluation of RACCOON includes two key components: an assessment of the tool's usability for experts and an evaluation of the public's reception of the visualizations produced. This approach ensures that RACCOON not only meets the needs of professionals but also effectively communicates risks to a broader public. In summary, we make the following contributions:

- We provide a visualization generation tool supporting experts in determining the most important health risk factors, as well as in health risk communication.
- We incorporate techniques from risk communication, as well as narrative and affective visualization, to generate engaging and easily comprehensible data-driven health risk visualizations.
- We conducted studies with experts and the general public to assess the tool's usefulness and the visualizations it generated, as well as the effect of different communicative intents.

RACCOON is publicly available at https://akleinau.github.io/raccoon/.

DOI: 10.2312/vmv.20241200

2. Related Work

In this section, we will explore the research fields of risk communication, narrative visualization, affective visualization, and visualization generation.

2.1. Risk Communication

Communicating risks requires making complex, probabilistic information easily understandable for a general public [SWW99]. Effective risk communication can reduce fears by setting them into context as well as increase awareness [Lei04].

Guidelines by Trevena et al. [TZFE*13] show how health risks should be communicated. In general, pictographs are recommended as a very effective method for risk communication [ASKS06, HZFU*08]. However, the best format varies depending on intent, for example, behavior change or likeability [ASKS06]. Each of these intents requires different, sometimes contradictory design choices. For accuracy, visualizations should provide extensive and detailed information, but omitting information may be effective for encouraging behavior change. When participants used their preferred format, it did not lead to an increase in performance [ASKS06]. Additionally, visualizations will be perceived differently depending on the characteristics of the viewer. For example, participants with higher numeracy generally performed well across all formats [HZFU*08]. Therefore, visualization design should focus on low numeracy users who are most impacted by the choice of visualization format [HZFU*08].

A study on the needs of health communicators [SAZFB23] revealed that communicators preferred graphics for communication but often had problems clearly defining their communication goals and applying guidelines accordingly. RACCOON addresses this need by providing tailored risk communication graphics.

2.2. Narrative Visualization

Narrative visualization merges methodologies from visualization with storytelling, crafting visualizations and data narratives that are engaging, easily comprehensible, and memorable [SH10].

Garrison et al. [GMPB23] provide an introduction to how narrative visualization can be used for medical visualizations. Meuschke et al. [MGS*22] present an initial proof-of-concept, illustrating how narrative visualization can be used for disease education. Their goal is to help non-designer scientists use narrative visualization in science communication, closely aligning with our goals.

Lee et al. [LRIC15] describe the narrative process, starting with data exploration over making a story to telling a story. They highlight the need for more computer support in this process.

2.3. Affective Visualization

Affective visualization investigates how emotions can be used in visualization [LWC24]. Lan et al. [LSZC21] found 12 primary affective responses associated with infographics. Research in affective visualization includes the effect of color on data visualization [BPS17]. We will use colors in RACCOON to support specific communicative intents. Lan et al. [LWS*22] show that even inducing negative emotions might have advantages, like increasing long-term memory and user engagement. In RACCOON, we will aim at using negative emotions to increase the perceived risk and urgency in risk communication.

2.4. Visualization Generation and Recommendation

Visualization recommendation tools have great commercial success, with a popular tool being *Tableau*[†] based on Stolte et al. [STH02]. Mackinlay [Mac86] presented the foundational work in automatic visualization generation. Further tools also automate data exploration. For example, Wongsuphasawat et al. [WMA*15] introduced *Voyager*, a tool that generates visualization recommendations based on a tabular data set. Unlike our approach, these recommendations are general and do not focus on a specific domain.

Wang et al. [WSZ*19] presented a tool to automatically generate a fact sheet from a given data set. Moreover, there are special-purpose systems designed to generate and annotate visualizations for specific fields like stock data [HDA13]. RACCOON is such a special-purpose system, targeting risk communication. Differing from Cullen et al.'s tool [CHB*24] which also designs public health visualizations, our tool starts from the data set, aiding experts until the final visualization.

3. Presenting RACCOON

We consider a scenario in which an expert wants to create a data story about diabetes risk factors. After importing their data set and selecting diabetes as the variable of interest, the expert will receive visualization recommendations generated from the data set, e.g. of potential risk factors or the size of the data set. The expert can collect relevant visualizations by adding them to their dashboard. As default, all visualizations are designed in a way best suited for an expert exploring the data set. With two clicks, the expert can change all visualizations to one of the designs more suited for the general public, as described in Section 6.1. Global settings and a detailed fact view enable the modification of bins, visualization aspects, and annotations. Visualizations can be exported as PNG or PDF files or directly copied to the clipboard for easy import into other applications, like PowerPoint.

RACCOON is usable with epidemiological data sets of cross-sectional studies in CSV format. An example data set containing diabetes risk factors is available from the Behavioral Risk Factor Surveillance System (BRFSS) [Cen15], a telephone survey conducted by the CDC in the United States.

RACCOON is implemented as a static website, developed using the javascript framework Vue[‡]. Data protection is ensured with all data processing being performed locally on the client side. For visualizations, the javascript library d3 [BOH11] was used. RACCOON is optimized for the screen sizes of laptops and desktops. The browsers Google Chrome and Microsoft Edge are recommended. Figure 1 shows a part of the user interface of RACCOON.

[†] https://www.tableau.com/

[‡] https://vuejs.org/

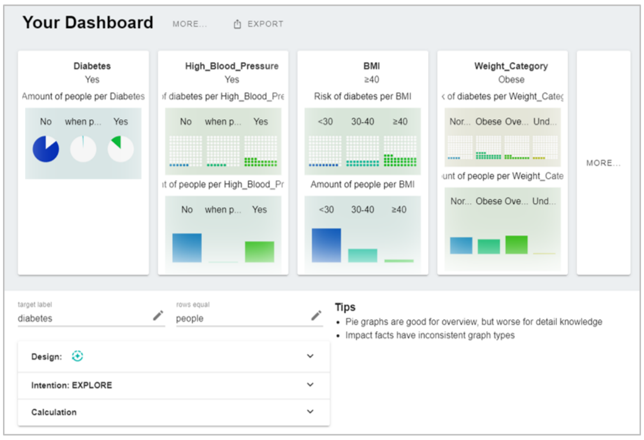

Figure 1: *The user interface of* RACCOON *with the dashboard containing previews of selected risk factors, settings, and tips.*

In the following sections, we will describe in detail how an expert is supported in data exploration (Section 4), how risks can be communicated to the general public (Section 5), and how the expert can be supported in this communication process (Section 6).

4. Data Exploration

How can RACCOON support experts in exploring their data set? We adopted the requirements defined by Preim and Lawonn [PL20], aimed at visual analytics tools for public health, to risk factors:

- **(R1) Provide an overview of the data** and potential risk factors.
- **(R2) Enable analysts to integrate expert knowledge** to choose the risk factors they know are most important.
- **(R3) Provide familiar visualizations** that experts are accustomed to or can easily interpret.
- **(R4) Provide integrated information,** for example, relevant statistical information when displaying a risk factor.
- **(R5) Provide visual support for association analysis** between the disease and risk factors.
- **(R6) Provide visual support for comparisons** of risk factors to find the most relevant ones.

A data set entered by the expert will be automatically preprocessed, as described in Section 4.1. RACCOON displays general information (**R1**), like the number of data entries in the dataset and the prevalence (frequency) of the disease. All risk factors are shown as simplified previews to allow an overview of the data set (**R1**) as well as an easy comparison (**R6**). Risk factors are ranked by their computed importance, as described in Section 4.2. Additionally, we wanted to enable experts to manually compare risk factors by showing them adequate visualizations. Here, we were inspired by the fact-scoring method of Wang et al. [WSZ*19] separating each fact's relevance into its *significance* and *impact*. Adapting these concepts to risk factors, *significance* describes the strength of the correlation between a risk factor and the outcome, whilst *impact* describes how many persons are affected by an increased risk. Selected risk factors are visualized using commonly used visualization types, like pictographs and bar charts (**R3**). When a risk factor is selected, additional information is shown (**R4**), including potential biases and correlations (**R5**). Extensive customization options allow the integration of domain knowledge (**R2**).

4.1. Automatic Data Processing

Data sets of epidemiological studies are typically tabular, multivariate data. We assume the data set is cross-sectional, meaning that prevalences can be inferred, and time will not be considered. The values of each continuous factor in the dataset are binned using a custom algorithm. The algorithm aims to create a small number of bins of the same range, merging outer bins if necessary to maintain reasonable bin sizes. Furthermore, the algorithm ensures that the boundaries of the bins are multiples of five, which facilitates human interpretation. The expert can also adapt the bins manually. Missing values are excluded, if not otherwise specified, as they are assumed to be unimportant to the general public.

4.2. Automatic Risk Factor Recommendation

Potential risk factors are identified by using logistic regression [YTS96] to measure their potential in predicting the disease. Categorical factors are encoded using one-hot encoding. Previously selected factors can be used to improve subsequent recommendations, facilitating the choice of a diverse set of risk factors. Factor rankings are then based on the enhancement of the model's goodness of fit when trained with the factor alongside all previously selected ones, relative to a model trained exclusively with the previously selected factors.

For the final scoring, to capture even small improvements in a model's goodness of fit caused by a risk factor, we considered the predicted likelihoods, instead of final predictions. With our classes being "having the disease" or "not having the disease", we calculate the average error of the likelihoods for each data set item within each class: $e = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i|$). Then, to ensure equal weighting of both classes, we compute the mean of these two errors: $f = 0.5 * (e_1 + e_2)$ Alternative scoring metrics include the epidemiological measures relative risk and odds ratio [ZCB97].

5. Data Communication

To communicate risks to the general public, we first decided what information is important. To introduce the topic, RACCOON provides general information about the disease, like prevalence information, and basic information on the data set, like the number of participants. Then, individual risk factors are displayed. Lastly, RACCOON can contextualize the risk factors. This includes comparing risk factors by their absolute risks, the relative risk increase associated with them, or their importance when assessing the disease likelihood.

Risk visualizations typically aim for the following outcomes [ASKS06]:

- **(R7) Likeable:** People will rather accept visualizations that they like.
- **(R8) Accurate:** Visualizations should display information in a format optimized for accurate human perception.

- **(R9) Behavior change:** The visualizations should motivate the audience to change their risky behavior.

In the context of our work, we identified two more requirements:

- **(R10) Engaging:** We will leverage narrative and affective visualization to support engagement [SH10, LWC24].
- **(R11) Consistency:** Our tool must ensure consistency across multiple visualizations. [SH10].

We will now describe the design of the visualizations and their annotation. Our visualizations were iteratively developed, involving regular presentations to various laypersons and experts.

5.1. Designing Risk Factor Visualizations

We start by describing some general design considerations. For easy export, all visualizations are static. We opted for creating multiple simple visualizations, as opposed to a single convoluted visualization. Simple visualizations increase likeability and are better understood, especially by persons with low visualization literacy [ASKS06]. To keep those visualizations consistent (**R11**), we adapt the proposed concepts by Wang et al. [WSZ*19] of *Inter-Consistency* and *Intra-Diversity*. Across all risk factors, the same designs are used for the same facts, making it easier to switch between factors (*Inter-Consistency*). However, when looking at one risk factor, there is a diversity in visualization formats created for its facts (*Intra-Diversity*).

5.1.1. Visualization Format

The influence of the visualization format on risk perception is well researched [ASKS06, HZFU*08]. Bar charts have a high likeability (**R7**) due to their simplicity [ASKS06] as well as utilizing the accurate human perception (**R8**) of lengths. Pictographs are one of the best formats for risk communication [ASKS06, HZFU*08] because ordered icons still let the human perception use area information for accurate assessment, but the icons add intuitivity and highlight the part-to-whole relationship [HZFU*08]. Pie charts have a bad reputation in the visualization community. Their use of degrees significantly decreases accuracy. However, they are simple and liked by the public (**R7**) as well as great for visualizing the part-to-whole relationship [HZFU*08]. RACCOON enables experts to choose between pictographs, bar charts, pie charts, and a textual summary.

5.1.2. Labels

Our titles are inspired by Wang et al. [WSZ*19] and are kept simple, but clearly state what the visualization is about. Ratios, for example, how many persons have a disease, can be displayed as percentages, using the “1 in X” format, and using natural frequencies with fixed denominators (e.g., 5 out of 10 and 8 out of 10). There is clear evidence against the use of “1 in X”, as it resulted in very low accuracy when comparing numbers [ASKS06, TZFE*13]. Natural frequencies are well suited for pictographs, as they can directly align with the numbers of icons. Percentages may be more familiar to users. RACCOON supports the use of percentages and natural frequencies. For data analysis, it also supports displaying the absolute values of persons per bin.

5.1.3. Icons

Imagery or metaphors are one design method to create an affective reaction [LWC24]. RACCOON supports custom icons for pictographs by using the Material Design Icon libary[§]. Preselected icons include humans and emoticons displaying different emotions. Additionally, pictographs can be designed with or without showing the denominator. For example, when 4 out of 100 persons have a disease, only the four affected persons can be shown, or all 100 persons. Displaying the denominator prevents the phenomenon of denominator neglect bias [TZFE*13], where numbers appear more significant than they actually are. However, hiding the denominator might be a valuable option for very small risks or when aiming to increase the perceived risk (**R9**). Therefore, RACCOON gives experts both the option to show as well as hide the denominator.

5.1.4. Styling

A well-designed style improves likeability (**R7**) and engagement (**R10**). In RACCOON, we integrated multiple predefined color schemes. However, such schemes are often designed (only) for maximum differentiability of colors. Therefore, we decided to additionally provide the option to create a unique color palette based on analogous color schemes [PC20] that provide a more coherent look. For storytelling, we incorporated a color scheme emphasizing a particular bin of each factor, coloring it and graying out the rest.

5.2. Annotations for Risk Factor Visualizations

RACCOON automatically annotates the generated visualizations with two types of annotations. A *summary annotation* is created for each factor, separate from their visualizations, to aid in understanding how the visualizations work together. It merges data about high-risk bins and their size, compared to the majority's risk. The annotation is created using a template that the expert can modify. *Visualization annotations* are created as an addition to each visualization, intended to provide a clear starting point on how to read and interpret it. For more flexibility, we provide a list of further generated annotations as well as the ability to add custom ones.

6. Bridging the Gap

How can RACCOON integrate data exploration and communication? Here, the last two requirements arise.

- **(R12) Supporting effective risk communication and visualization** to support experts in adhering to current research in risk visualization design [Lei04, SWW99].
- **(R13) Extensive customizability** to allow the addition of domain knowledge, custom narratives, or individualization.

The last two sections have already described the broad range of customizability options available (**R13**). Guidance (**R12**) is provided through hints when, for example, a data set is too small or there is inconsistency in visualization styles. Additionally, to tackle the process from data exploration to communication, RACCOON features multiple visualization designs for different communicative intents.

[§] https://pictogrammers.com/library/mdi/

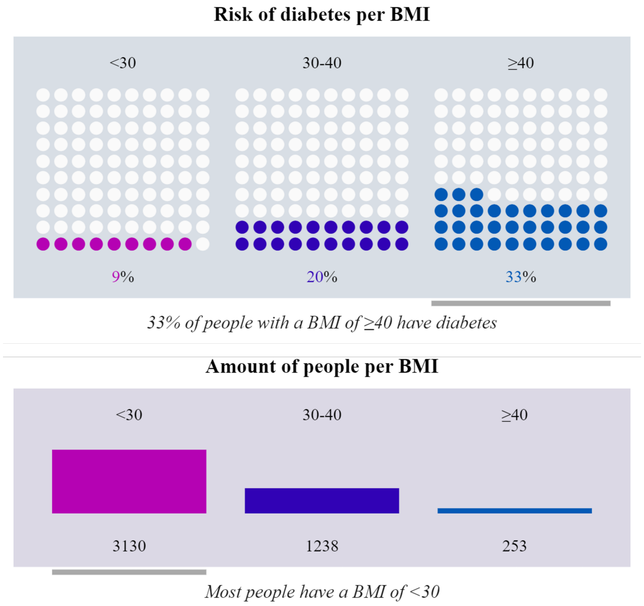

Figure 2: *The **Explore** intent aims to support data exploration.*

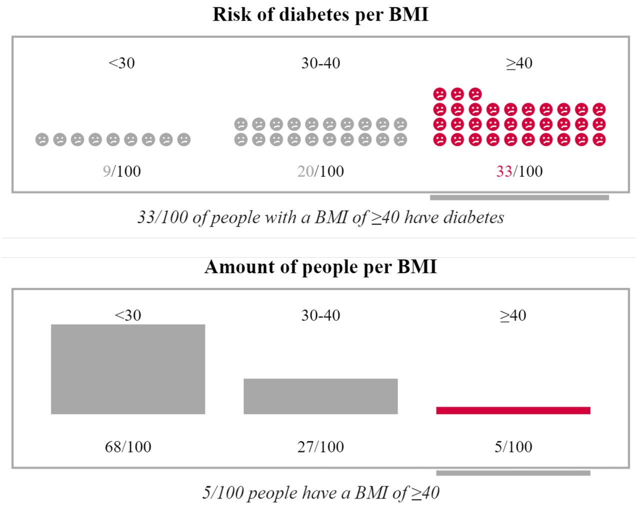

Figure 3: *The **Convince** intent aims to motivate behavior change.*

Figure 4: *The **Educate** intent aims to increase knowledge.*

6.1. Communicative Intents

Except for the intent directed towards the expert, our intents align with the risk communication goals as described by Rohrmann [Roh92]. We excluded the goal of facilitating cooperative conflict resolution, as it is not applicable to our use case.

The *Explore* intent (Figure 2) is targeted at the needs of the expert when exploring the data set. They require high accuracy to make educated decisions. Intuitivity is not as essential as the expert has more time to get used to the visualizations. Experts commonly use percentages. The impact visualization reveals the exact count per bin, aiding in determining whether the population size warrants further analysis. The color scheme has a high saturation and diverse hues to easily visually separate different bins. Icons are kept as neutral circles to reduce visual clutter. Annotations are used for hints when observations are not statistically relevant.

The *Convince* intent (Figure 3) is targeted at a general audience with the goal of increasing the perception of risk to influence behavior. We hide denominators in order to make differences between bins appear bigger. The color scheme is dramatic, with heavy use of read to elicit a stronger emotional reaction [LWS*22]. The highest risk is highlighted in red, whilst the rest is kept grey. We chose sad emoticons as icons to create a strong affective reaction [LWS*22]. Annotations draw attention to the bin with the highest risk.

The *Educate* intent (Figure 4) is targeted at a general audience to increase their knowledge about a disease. The focus is to keep the audience engaged and convey the main messages. We use pie charts as they are well-liked by the public and are efficient in providing a good overview of the data [HZFU*08]. We trade-off accuracy with this choice, but gain an intuitive understanding of the part-to-whole relationship. Colors are chosen to support trust, by being lighter, less saturated greens and blues [BPS17]. The icon design is kept relatable by choosing human silhouettes without facial expressions. The annotations provide a good overview of each risk factor.

7. Evaluation

We performed multiple studies to evaluate RACCOON both by domain experts and the general audience. The authors of this paper did not participate in the evaluations. The expert user study and the first public user study were evaluated with a preliminary version of our visualizations. We then finalized our visualizations and performed a second public user study.

7.1. Expert User Study

We started by evaluating with the experts' perspective, using preliminary visualization designs. We interviewed clinical, visualization, and risk communication experts in a qualitative study. We used the think-aloud protocol [PRI18] for direct feedback during the exploration of the tool. Guiding tasks included asking the experts to choose a data set, select risk factors, and customize visualizations. Afterwards, they were requested to complete a questionnaire that initially gathered demographic data and information about their experience. Then, they were asked to rate the tool using various questions detailed in Figure 5.

7.2. Public User Studies

We then evaluated the perception of the visualizations by the general public. Therefore, we focused on the "Convince" and "Educate" intent. As the general public is very diverse, we used online surveys, which allowed us to reach out further. The survey was shared with friends and colleagues, as well as multiple student and online groups.

To evaluate the visualizations, we created short data stories in which all visualizations are kept as close to our templates as possible. The only adaptations done were aligning bins with domain conventions, and fixing grammatical errors generated by the automatic annotation. We implemented the study in German, as it took place in Germany. This prevents bias based on the participants' English competency. The data stories are based on the diseases diabetes, arthritis and heart attack. They were chosen as diseases in which behavior change could be fruitful in reducing an individual's risk. Two versions of each data story, differing in the communicative intent, were used. As an example, one of the data stories is provided as supplementary material. The study starts by asking about demographic data. Then, it will randomly present one of the data stories to the user. Afterwards, various questions are asked, as detailed in Figure 6.

8. Results

We have structured the results of our user studies by first describing the expert evaluation, and then each of the two public user studies.

8.1. Expert User Study

We conducted online interviews with one clinician and four visualization experts, including a risk communication expert (three females, two males), ages 28 to 42.

The experts found RACCOON beneficial for its easy visualization generation. The clinician expressed a strong liking for the tool, as it enables quick creation of visualizations without requiring programming knowledge. However, the usage of the tool was not directly clear and required initial guidance. Then, the experts appreciated how easily they could use RACCOON to customize the visualizations. The experts also appreciated the connectivity, which automatically propagates changes across visualizations (**R11**). However, two experts requested a more direct manipulation of the visualizations. The experts also provided valuable feedback regarding the design of our preliminary visualizations. For example, they recommended simplifying the visualizations and visually relating all visualizations of a risk factor together (**R4**). Three experts commented that they liked the usage of pictographs (**R3**), because, as one stated, they are "reflective of current research". All experts liked the annotations, as they helped in interpreting the visualizations and served as notifications when there were insufficient data points (**R5**). The intents were explored with different preferences as to which design changes they liked or disliked. Three experts expressed concerns about misleading the audience when hiding the denominators. They provided various suggestions for improvement, such as exploring the use of color saliency and icons.

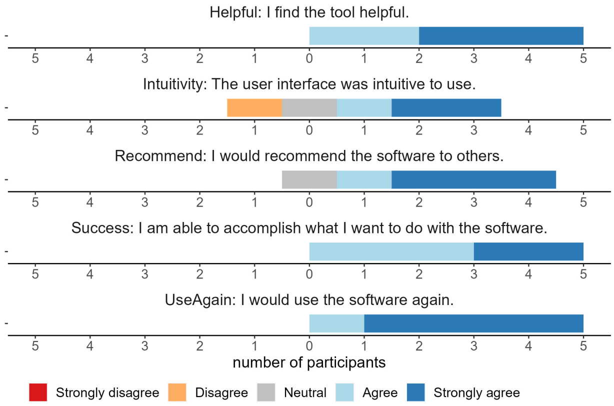

Figure 5: *Results of the expert questionnaire. For each question, colored bars display the number of people per answer, centered around the neutral answer with disagreeing answers to the left and agreeing answers to the right. The answers are separated by intent.*

All five experts agreed that the tool is helpful (see Figure 5). However, opinions on intuitivity varied. Four experts would recommend the software to others. All felt that they were able to accomplish what they wanted to do with it and would use it again.

8.2. First Public User Study

The first public user study was performed with 62 persons and a story about diabetes. Our preliminary visualization designs already received good feedback on detail, trust, likeability and understandability. However, they were often perceived as overloaded. Additionally, people had a hard time connecting the different visualizations together. Both intents did not achieve their intended effect.

8.3. Second Public User Study

The second user study was performed using our final visualization design, which features simplified visualizations and an enhanced connectedness of visualizations. Intents now also integrate research from affective visualization and customized narratives.

Proceedings published by Eurographics - The European Association for Computer Graphics.

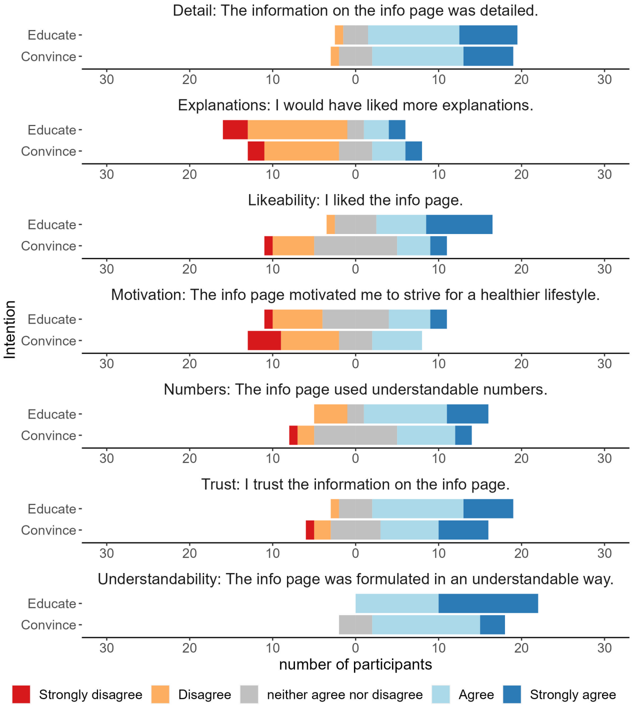

Figure 6: *Results of the second public study. For each question, colored bars display the number of people per answer, centered around the neutral answer with disagreeing answers to the left and agreeing answers to the right. The answers are separated by intent.*

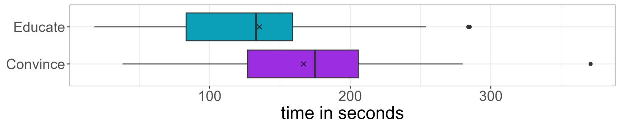

Figure 7: *For each intent, the distribution of watch time is shown as a boxplot of the time in seconds.*

We created info pages on the topics of heart attack and arthritis. In total, 44 people (17 females, 22 males, 5 diverse) completed the study. The mean age was 33, ranging from 19 to 88. Of those, 11 persons were assigned to each combination of intention and topic, resulting in four groups. Most persons had *a little bit* of prior knowledge about their disease (26/44) and were *familiar* with visualizations (20/44). Most persons paid *neutral* (21/44) or strong (16/44) attention to a healthy lifestyle.

Participants generally agreed that the data stories were detailed (35/44), trustworthy (30/44), understandable (38/44), and had understandable numbers (24/44), see Figure 6. Likeability was rated more diversely, but still most persons liked the info page (20 out of 44). Most persons did not want more explanations (26/44). Subsequent motivation to strive for a healthier lifestyle differed, with 18/44 persons disagreeing and only 13/44 agreeing. The *educate* intention was rated slightly more positive in understandability, trust, and likeability. Surprisingly, the *convince* intention is slightly worse in motivating persons to strive for a healthier lifestyle. However, as shown in Figure 7, persons generally spent more time looking at the *convince* version than the *educate* version.

Eight persons gave additional textual feedback. One person found the separate display of bins confusing, One wanted more labels and three suggested spelling improvements. Two persons found the context visualization unclear. Four commented on the content of the data stories, wishing for more information or calling it informative. For the *educate* version, one person wrote that it is good for education. However, changes in the text and colors would be necessary to convince people. For the *convince* version, one person commented that the info page is overloaded. Another person suggested improving the color coding of bins.

9. Discussion

The expert evaluation of RACCOON indicates good usability and perceived helpfulness. Four experts strongly agreed that they would use the software again, suggesting that the tool fills a gap. RACCOON was valued for its ability to create visualizations without requiring programming skills (**R12**). However, the evaluation revealed a need for better user guidance and an improvement in the tool's design. The experts also valued the customizability provided by RACCOON to adapt the visualizations to their needs (**R2, R13**).

In our public user studies, the visualizations were generally well-received, especially in terms of detail, trust, and understandability (**R8**). Exploring various communicative intents, we observed mixed outcomes. The *educate* version performed slightly better in terms of likeability (**R7**), trust, and understandability (**R8**). Interestingly, whilst neither version had significant success in persuading participants to consider adopting a healthier lifestyle (**R9**), the *educate* version was marginally more successful. This suggests that increasing risk perception in the *convince* version may have to be better balanced with other requirements like viewer engagement.

Unexpectedly, the *convince* version led to longer reading times (**R10**). This pattern was also observed in individuals affected by at least one risk factor, whom we anticipated would be more motivated to change their behavior, yet we only saw a slight correlation. This similar behavior prompts speculation. It suggests that the *convince* version may have heightened the perceived risk but did not effectively translate it into motivation for behavioral change.

Limitations of the expert evaluation include a small sample size. Results are based solely on one data set and might differ for others. Participants followed specific tasks and may use the tool differently in a real-world scenario. Depending on the device and data set, longer runtimes may occur. To avoid this, calculations can be performed on a randomly selected subset of observations instead.

Limitations of the public studies include a small sample size when generalizing the results to a general audience. Recruitment of participants through mainly university groups biased the sample towards younger individuals with a higher education level. The story's German translation could have altered the meaning of certain words, impacting the results.

10. Conclusion

Preventable health risks are a major cause of disease, underscoring the crucial importance of risk communication. Our tool RACCOON assists experts in exploring and communicating risk factors. It generates data-driven, customizable risk visualizations based on research in risk communication, affective visualization, and narrative visualization. Its utility has been confirmed across several diseases. RACCOON contributes to the growing research on how specialized tools can empower experts to convey insights through impactful, tailored visualizations. Future work could support the exploration of specific subgroups, temporal or geospatial data, and narrative creation. Finally, more research is needed on the interplay between communicative intent and visualization design.

References

[ASKS06] ANCKER J. S., SENATHIRAJAH Y., KUKAFKA R., STARREN J. B.: Design features of graphs in health risk communication: a systematic review. *Journal of the American Medical Informatics Association 13*, 6 (2006), 608–618. 1, 2, 3, 4

[BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D3 data-driven documents. *IEEE Transactions on visualization and computer graphics 17*, 12 (2011), 2301–2309. 2

[BPS17] BARTRAM L., PATRA A., STONE M.: Affective color in visualization. In *Proc. of SIGCHI conference on human factors in computing systems* (2017), pp. 1364–1374. 2, 5

[Cen15] CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC): Behavioral risk factor surveillance system survey data. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2015. 2

[CHB*24] CULLEN R., HEITKEMPER E., BACKONJA U., BEKEMEIER B., KONG H.-K.: Designing an infographic webtool for public health. *Journal of the American Medical Informatics Association 31*, 2 (2024), 342–353. 2

[GMPB23] GARRISON L. A., MEUSCHKE M., PREIM B., BRUCKNER S.: Current approaches in narrative medical visualization. In *Approaches for Science Illustration and Communication*. Springer, 2023, ch. 4. in publication. 2

[HDA13] HULLMAN J., DIAKOPOULOS N., ADAR E.: Contextifier: automatic generation of annotated stock visualizations. In *Proc. of SIGCHI Conference on human factors in computing systems* (2013), pp. 2707–2716. 2

[HZFU*08] HAWLEY S. T., ZIKMUND-FISHER B., UBEL P., JANCOVIC A., LUCAS T., FAGERLIN A.: The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient education and counseling 73*, 3 (2008), 448–455. 2, 4, 5

[Lei04] LEISS W.: Effective risk communication practice. *Toxicology letters 149*, 1-3 (2004), 399–404. 2, 4

[LRIC15] LEE B., RICHE N. H., ISENBERG P., CARPENDALE S.: More than telling a story: Transforming data into visually shared stories. *IEEE Computer graphics and applications 35*, 5 (2015), 84–90. 2

[LSZC21] LAN X., SHI Y., ZHANG Y., CAO N.: Smile or scowl? looking at infographic design through the affective lens. *IEEE Transactions on visualization and computer graphics 27*, 6 (2021), 2796–2807. 2

[LWC24] LAN X., WU Y., CAO N.: Affective Visualization Design: Leveraging the Emotional Impact of Data. *IEEE Transactions on visualization and computer graphics 30*, 1 (2024), 1–11. 1, 2, 4

[LWS*22] LAN X., WU Y., SHI Y., CHEN Q., CAO N.: Negative emotions, positive outcomes? exploring the communication of negativity in serious data stories. In *Proc. of SIGCHI Conference on Human Factors in Computing Systems* (2022), pp. 1–14. 2, 5

[Mac86] MACKINLAY J.: Automating the design of graphical presentations of relational information. *ACM Transactions On Graphics (Tog) 5*, 2 (1986), 110–141. 2

[MAZ*20] MURRAY C. J., ARAVKIN A. Y., ZHENG P., ABBAFATI C., ABBAS K. M., ABBASI-KANGEVARI M., ABD-ALLAH F., ABDELALIM A., ABDOLLAHI M., ABDOLLAHPOUR I., ET AL.: Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet 396*, 10258 (2020), 1223–1249. 1

[MGS*22] MEUSCHKE M., GARRISON L. A., SMIT N. N., BACH B., MITTENENTZWEI S., WEISS V., BRUCKNER S., LAWONN K., PREIM B.: Narrative medical visualization to communicate disease data. *Computers & Graphics 107* (2022), 144–157. 2

[PC20] PLANTE T. B., CUSHMAN M.: Choosing color palettes for scientific figures. *Research and Practice in Thrombosis and Haemostasis 4*, 2 (2020), 176. 4

[PL20] PREIM B., LAWONN K.: A survey of visual analytics for public health. *Computer Graphics Forum 39*, 1 (2020), 543–580. 3

[PRI18] PREIM B., ROPINSKI T., ISENBERG P.: A Critical Analysis of the Evaluation Practice in Medical Visualization. In *Eurographics Workshop on Visual Computing for Biology and Medicine* (2018), pp. 45–56. 6

[RBK10] RESSING M., BLETTNER M., KLUG S. J.: Data analysis of epidemiological studies: part 11 of a series on evaluation of scientific publications. *Deutsches Arzteblatt International 107*, 11 (2010), 187. 1

[Roh92] ROHRMANN B.: The evaluation of risk communication effectiveness. *Acta psychologica 81*, 2 (1992), 169–192. 5

[SAZFB23] SURESH U., ANCKER J. S., ZIKMUND-FISHER B. J., BENDA N. C.: Designing support to help health communication professionals convey numbers clearly to the public–a needs assessment and formative usability evaluation. In *AMIA Annual Symposium Proceedings* (2023), vol. 2023, pp. 1277–1286. 2

[SH10] SEGEL E., HEER J.: Narrative visualization: Telling stories with data. *IEEE Transactions on visualization and computer graphics 16*, 6 (2010), 1139–1148. 1, 2, 4

[STH02] STOLTE C., TANG D., HANRAHAN P.: Polaris: A system for query, analysis, and visualization of multidimensional relational databases. *IEEE Transactions on visualization and computer graphics 8*, 1 (2002), 52–65. 2

[SWW99] SCHWARTZ L. M., WOLOSHIN S., WELCH H. G.: Risk communication in clinical practice: putting cancer in context. *JNCI Monographs 1999*, 25 (1999), 124–133. 2, 4

[TZFE*13] TREVENA L. J., ZIKMUND-FISHER B. J., EDWARDS A., GAISSMAIER W., GALESIC M., HAN P. K., KING J., LAWSON M. L., LINDER S. K., LIPKUS I., ET AL.: Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. *BMC Medical informatics and decision making 13*, 2 (2013), 1–15. 1, 2, 4

[WMA*15] WONGSUPHASAWAT K., MORITZ D., ANAND A., MACKINLAY J., HOWE B., HEER J.: Voyager: Exploratory analysis via faceted browsing of visualization recommendations. *IEEE Transactions on visualization and computer graphics 22*, 1 (2015), 649–658. 2

[WSZ*19] WANG Y., SUN Z., ZHANG H., CUI W., XU K., MA X., ZHANG D.: Datashot: Automatic generation of fact sheets from tabular data. *IEEE Transactions on visualization and computer graphics 26*, 1 (2019), 895–905. 2, 3, 4

[YTS96] YUEN J., TWENGSTRÖM E., SIGVALD R.: Calibration and verification of risk algorithms using logistic regression. *European Journal of Plant Pathology 102* (1996), 847–854. 3

[ZCB97] ZOCCHETTI C., CONSONNI D., BERTAZZI P. A.: Relationship between prevalence rate ratios and odds ratios in cross-sectional studies. *International journal of epidemiology 26*, 1 (1997), 220–223. 3