EG UK Computer Graphics & Visual Computing (2024)
A. Slingsby and D. Hunter (Editors)

Exploring Data Analysts' Uncertainty Reasoning Strategies for Effective Uncertainty Visualization Design

Christoph Kinkeldey[1], Mirela Reljan-Delaney[2], Georgia Panagiotidou[3] and Jason Dykes[2]

[1]Hamburg University of Applied Sciences, Germany
[2]City University London, United Kingdom
[3]King's College London, United Kingdom

Abstract

Despite its proven positive effects, visual data analysis rarely includes information about data uncertainty. Building on past research, we explore the hypothesis that effective uncertainty visualizations must support reasoning strategies that enable data analysts to utilize uncertainty information ('uncertainty reasoning strategies', UnReSt). Through this work, we seek to gain insights into the reasoning strategies employed by domain experts for incorporating uncertainty into their visual analysis. Additionally, we aim to explore effective ways of designing uncertainty visualizations that support these strategies. For this purpose, we developed a methodology involving online meetings that included think-aloud protocols and interviews. We applied the methodology in a user study with five domain experts from the field of epidemiology. Our findings identify, describe, and discuss the UnReSt employed by our participants, allowing for initial recommendations as a foundation for future design guidelines: uncertainty visualization should (i) visually support data analysts in adapting or developing UnReSt, (ii) not facilitate ignoring the uncertainty, (iii) aid in the definition of acceptable levels of uncertainty, and (iv) not hide uncertain parts of the data by default. We reflect on the methodology we developed and applied in our study, addressing challenges related to the recruiting process, the examination of an existing tool along with familiar tasks and data, the design of bespoke prototypes in collaboration with visualization experts, and the timing of the meetings. We encourage visualization researchers to adapt this methodology to gain deeper insights into the UnReSt of data analysts and how uncertainty visualization can effectively support them. The supplemental materials can be found at `https://osf.io/s2nwf/`.

CCS Concepts

• ***Human-centered computing*** **→** ***Empirical studies in visualization; Visualization design and evaluation methods;***

1. Introduction

As previous research suggests, visualizing uncertainty has the potential to support people in conducting informed data analysis, interpretation and decision making [KMRS15]. Yet, visual data analysis rarely includes representations of uncertainty, and research has shown that visualization designers have difficulties including uncertainty because of "challenges calculating, visualizing, and explaining uncertainty to viewers" [Hul20]. A major cause for this situation is the common prioritization in uncertainty visualization of accurately communicating specific quantities of uncertainty, rather than supporting individuals' reasoning with uncertainty [KMS14].

We build upon past research, arguing that understanding the reasoning strategies of data analysts can contribute to addressing these challenges [KKH21, RPJ19]. While it is important for effective uncertainty visualization designs to account for analysis tasks and the nature of the data, their ultimate goal should be to support *reasoning strategies* that incorporate uncertainty information – 'uncertainty reasoning strategies' ('UnReSt'). This highlights the importance of understanding these strategies to provide enhanced guidelines for uncertainty visualizations that enable data analysts to effectively utilize information on uncertainty.

This work is based on the overarching research question: **how can we design uncertainty visualizations that effectively support data analysts' uncertainty reasoning strategies?** We report on a user study exploring this question by involving domain experts and a visual data analysis tool from epidemiology (the study of distribution and determinants of health-related states among specified populations and the application of that study to the control of health problems [Las86]). This field is particularly suitable for our study because epidemiological research inherently involves uncertain and complex data, and domain experts encounter uncertainty in many aspects of their work. Our strategy was to learn from domain experts how they deal with uncertainty and to find clues for what might be helpful for data analysis in other domains that have different backgrounds. The *Australian Cancer Atlas*, the visual analysis tool on which we base our study, integrates several representations of uncertainty, making it an interesting subject for our research. Our

DOI: 10.2312/cgvc.20241232

methodology follows a four step approach. We first capture participants' reasoning strategies for a specific task from their work through think aloud and interviews; second, we develop bespoke visual prototypes through co-design with visualization experts; we then evaluate whether the prototypes can help integrate uncertainty information into participants' analyses (again through think aloud and interviews); and finally, in the fourth step, we derive recommendations for effective uncertainty visualization design through thematic analysis.

Our findings identify three strategies for reasoning with uncertainty: ignoring the uncertainty, filtering by uncertainty, and qualifying hypotheses based on it. These insights lead to initial recommendations for designing visualizations that not only make uncertainty visible but also integrate it meaningfully into the analytical process. Our recommendations emphasize the importance of helping analysts develop robust uncertainty reasoning strategies, ensuring that uncertainty is neither ignored nor misrepresented, and supporting the definition of acceptable uncertainty thresholds.

2. Background

2.1. Visual Reasoning Strategies and Uncertainty

In the field of visual data analysis, reasoning can be defined as encapsulating "all tasks that result in the generation of thoughts, insights or decisions" [BPHE17]. Data analysts use *reasoning strategies*, "self-contained set[s] of processes that need to be applied for solving a task" [LF05] that can be simple or complex, conscious or unconscious, and optimized for speed or accuracy. A popular example is *heuristics* (such as rules of thumb) which are quick, reasonably effective reasoning strategies designed to avoid overly exhaustive search and optimization.

The goal of incorporating uncertainty into data analysis is to achieve positive effects such as more in-depth analyses, recognizing problems with data quality, greater confidence in the results, and others [KMRS15]. We agree with authors arguing that "[u]nderstanding how human[s] reason about uncertainty is [...] fundamental to designing useful visualizations" [RPJ19] and that "characterizing possible strategies may lead to design recommendations based on how users reason in practice" [KKH21].

Along these lines we hypothesize that in order to be effective, uncertainty visualization designs must support analysts' *uncertainty reasoning strategies (UnReSt)*. This is a related concept to "uncertainty coping strategies" [BPHE17] but with a different focus—while the latter aims to mitigate the negative effects of uncertainty on decision-making, well-being, and organizational performance, UnReSt aim to enhance the accuracy and reliability of data analysis by explicitly accounting for uncertainty.

2.2. Design Guidelines for Uncertainty Visualization

Visualizing uncertainty is a challenging endeavour because it draws from a large design space [PKH21], including design choices about explicitly symbolizing the uncertainty information or not (explicit/implicit), using separate symbolization or not (extrinsic/intrinsic), integrating data and uncertainty into one visualization or not (coincident/adjacent), and the decision whether to use animation and/or interactive elements to communicate the uncertainty [KMS14]. That is why guidelines have been proposed to support these design choices, e.g., classifications or taxonomies of uncertainty visualization techniques mostly based on data characteristics [PRJ12, BAL12, PWL97]. However, they are often of limited use because the effective choice of uncertainty visualization techniques depends not only on data characteristics but also on users, their tasks, and goals [SWM21, Mac15]. Thus, we aim to establish a foundation for more advanced design guidelines for uncertainty visualizations that support UnReSt.

3. User Study with Domain Experts from Epidemiology

Our user study draws from the domain of epidemiology and involves domain experts who use the Australian Cancer Atlas for their work. In the following, we present the methodology of our user study.

3.1. Australian Cancer Atlas (ACA)

Our plan involved conducting a user study in a real-world scenario with actual data analysts and analysis tools to ensure ecological validity (which is often lacking in sandbox studies).

The ACA is a freely accessible online atlas providing interactive visual access to data about various cancer types across Australia (while this paper refers to version 1.0 of the ACA, a new advanced version 2.0 is currently online at `https://atlas.cancer.org.au/`). It seemed particularly suitable because of the high-quality visualization design achieved through co-design involving experts in visualization design, and experts in epidemiology. It provides a map interface with various menus and options to filter the data and, crucially for our study, several expressive visual and verbal representations of uncertainty.

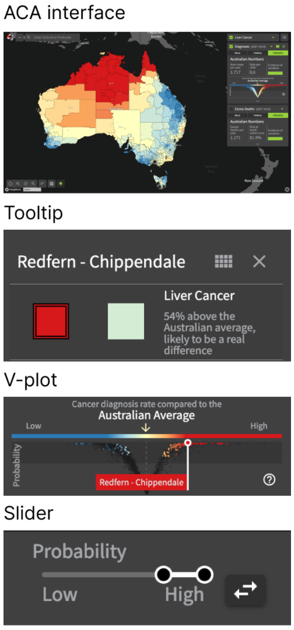
The cancer data represented in the atlas comprises SIR (*standard incidence ratio*) and EHR (*excess hazard ratio*) of various cancer types for Statistical Area Level 2 (SA2) regions in Australia. Both are related to the Australian average, so an SIR of 1 indicates that the incidence corresponds exactly to the expected rate, i.e. the Australian average. An uncertainty measure expresses the probability that a value (SIR or EHR) actually deviates from the average (values with low uncertainty are likely to be different from the average, while values with high uncertainty are not). The uncertainty is represented in the atlas in different ways (see figure on the next page):

In the interactive map, the uncertainty is encoded by manipulation of **color transparency**. This results in a fading effect, visually highlighting regions with low uncertainty.

In a **tooltip** for each region, a verbal expression appears, either "likely to be a real difference" (probability$\geq$60%) or "not likely to be a real difference"(probability$<$60%), thus providing binary information on the uncertainty of the value in this region.

A chart called the **v-plot** shows the (SIR or EHR) value per region in the x-axis and the level of uncertainty on the y-axis. It serves as an overview of the distribution of values and their uncertainty across Australia.

A **slider** allows regions on the map to be filtered (displayed or grayed out) according to their uncertainties. For instance, we can display just those regions with values that are likely to be real differences from the average (low uncertainty).



3.2. Recruitment

We were assisted in recruiting participants by one of the principal investigators of the ACA project, who sent out our call for participation to members of a mailing list for ACA users (*convenience sampling*). Those who responded by email received more detailed information about the process. We recruited five participants, who were experienced users from different organizations and with different emphases of use, so we were likely to capture a range of UnReSts. They agreed to participate in the two half-hour online sessions (see Sections 3.3 and 3.5) and to send us a description of a data analysis task from their work beforehand. Before the first meeting, all participants signed a consent form which informed them in detail about the purpose of the study, the procedure and the protection of their personal data.

The participant breakdown for the five participants was as follows: P1) senior epidemiologist at a major research institute for viral hepatitis, P2) cancer epidemiologist at a major research institute, P3) manager at a cancer registry and PhD student in non-melanoma skin cancers, P4) epidemiologist and health services researcher at a major research centre on cancer control and policy, and P5) senior research fellow in health data science at a university.

3.3. Step 1: Explore Current Practice

Before the first meeting, we asked each participant to send us a description of a data analysis task from their work involving the ACA. The participants chose the following tasks:

P1- Compare liver cancer incidence to hepatitis B and C prevalence for larger areas ('primary health networks') across Australia,

P2- Correlate melanoma incidence and mortality with ambient UV and temperature across Australia,

P3- Compare the distribution of incidence of various cancers between Tasmania and mainland Australia,

P4- Collect information for fact sheets on prostate cancer (incidence, mortality, survival and others) for different regions in Australia, and,

P5- Compare above-average liver cancer diagnoses and excess deaths of the population in two regions (Sydney and Northern Territory) in Australia.

Each participant demonstrated their task in a recorded *think aloud* session as that enabled us to follow their reasoning and the interaction with the tool. The demonstration was accompanied by a *semi-structured interview* to facilitate in-depth discussions about the role of uncertainty in their tasks and work practice.

3.4. Step 2: Develop Visual Prototype

We developed individual visual prototypes tailored to each participant's specific tasks. These prototypes were designed to prompt reflection on strategies for reasoning with uncertainty in step 3 of the study and were based on the data from the ACA. They included similar uncertainty visualizations and GUI elements, such as maps with uncertainty tooltips, dot plots with credible intervals, and uncertainty filter sliders (see Subsection 3.1), depending on each participant's task. We used Observable (`https://observablehq.com`) notebooks, combining (and reusing) different GUI elements. For instance, all prototypes provided a map, as geography played an important role for all tasks. But the prototypes for tasks that compared specific regions (P3, P5) or geographical patterns of SIR and EHR (P2) each contained a pair of linked maps (as shown in Figure 1).

We recruited visualization experts from the university context, specifically seeking individuals with professional experience in building visual interfaces. Their expertise was instrumental in designing the prototypes, reducing the risk of bias from individual design decisions, and improving the overall quality of the visual prototypes. We chose to recruit one visualization expert per prototype (rather than one visualization expert for all designs) to limit the workload of each individual and to accommodate a variety of perspectives on the designs that we could synthesize for development. One member of the research team documented each task and sent it to one of the visualization experts. In a meeting with each visualization expert, the requirements and possible designs of the prototype were discussed and documented. Based on this design, one member of the team implemented each prototype.

3.5. Step 3: Evaluate Visual Prototype

In the second online meeting, after a thorough introduction to the respective visual prototype, we asked the participants to repeat their task from step 1 using the prototype while thinking aloud. This was followed by an interview about the prototype and its potential usefulness for their task. The goal was to observe whether they developed new strategies to solve their task with the new interface. This sparked discussions about the uncertain nature of parts of the data and about possible (new) strategies to reason with the uncertainty.

3.6. Step 4: Derive Recommendations

From the five participants (P1 to P5) and two rounds of think aloud / interviews (steps 1 and 3) we obtained ten videos with 442 minutes of screen capture footage and ten automatically generated transcripts with with 5255 lines (the material in anonymized form

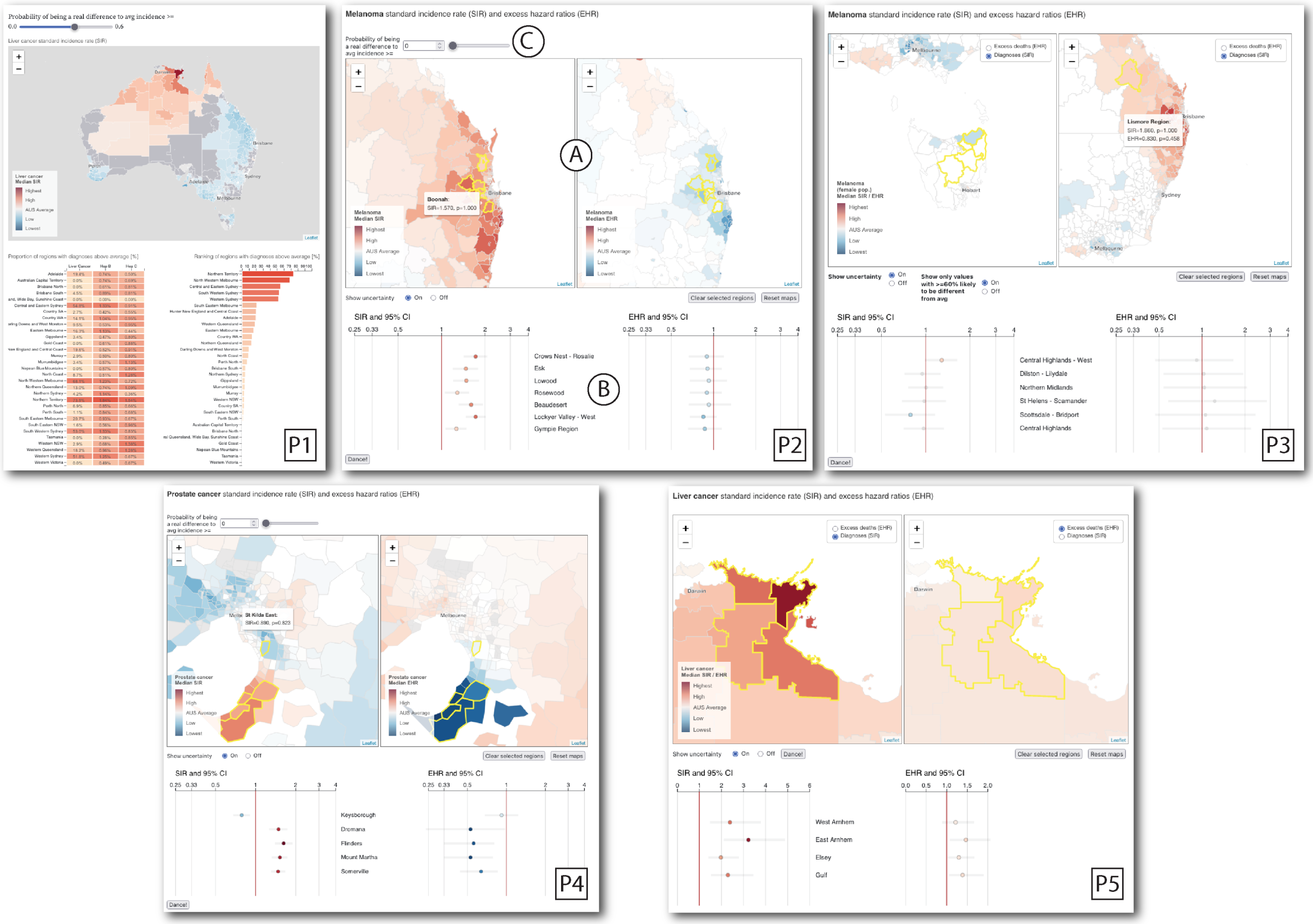


Figure 1: *Visual prototypes from our user study with domain experts from epidemiology. We created five prototypes tailored to each participant's task to explore their reasoning strategies when analyzing uncertain cancer data, for participants P1 (upper left) to P5 (bottom right). Example: The interactive visual prototype for P2 (upper line, center) was developed for the task of correlating incidence (SIR) and excess death (EHR) rates of melanoma in specific regions of Australia. The GUI provides a pair of linked maps (A) showing the geographical distribution of SIR and EHR but without the color transparency manipulation as in the ACA, which proved difficult for our participants to read. A dot plot contains SIR and EHR median values and credible intervals for selected regions (B), and a probability slider facilitates filtering of regions by uncertainty (C). Links to the prototypes can be found at* `https://osf.io/s2nwf/`

can be found at `https://osf.io/s2nwf/`). From this material, one author corrected the generated transcripts with the help of the videos and added annotations about how participants interacted with the GUIs. We used thematic analysis [BC21] to analyze the annotated transcripts regarding the overarching research question. The coding approach was abductive: some codes were derived from existing theory and prior research, while others emerged from patterns observed directly in the data. This led to the five main codes `Task [T]`, `Data [D]`, `Visualization [V]`, `Reasoning [R]`, and `Communication [C]`, refined by subcodes such as `Reasoning / Strategy [RST]` or `Reasoning / Uncertainty [RUN]`. Two authors independently conducted the coding process to enhance intersubjective agreement by ensuring that multiple perspectives were considered and discrepancies were discussed and resolved [DW14]. One of the authors merged the codes into one final solution per participant and round, following a consensus reached through discussion and resolution of any conflicting coding. Overall, we extracted 752 coded segments from the transcripts.

For the findings we present here, we compared and clustered coded segments with subcodes `[RST]` and `[RUN]` (see above). Through discussion of the findings, we identified themes for the initial recommendations in Section 5.

4. Findings about Uncertainty Reasoning Strategies (UnReSt)

The first important finding was that although all participants were convinced of the ACA's usefulness, *none of the participants used the uncertainty visualizations in the ACA* (see Section 3.1). This was unexpected, so we demonstrated the uncertainty representations in the ACA to facilitate discussion. After both rounds of meetings (step 1 and 3), we identified UnReSt in the following categories, including the strategies that our participants developed during the meetings (F2 and F3):

F1) Ignoring the Uncertainty

Information about data uncertainty is not (consciously) used for solving a task. The major strategy (or non-strategy) identified in round 1 of our study is to ignore uncertainty information during data analysis. Despite the general consensus among participants that uncertainty information can be valuable, four of our five participants did not formally take the uncertainty into account when demonstrating their analysis tasks based on the cancer data (with the exception of P1). We identified the following reasons for this:

- **Absence of Formal Strategies:** Despite being experienced domain experts accustomed to working with uncertain data, participants often lacked formal strategies or institutional guidelines to integrate uncertainty into their analyses.
  - While P1 did consider uncertainty in their analysis, the other participants did not have formal strategies in place for reasoning with uncertainty, indicating the reason for ignoring the uncertainty.
- **Lack of Awareness or Familiarity:** Some participants were simply not aware of the available uncertainty information in the ACA, or they were not familiar enough with the features to incorporate them into their analysis.
  - After our demonstration, P2 found the uncertainty information useful but stated, "I didn't [use uncertainty information because I didn't] know it was there."
- **Caution in Communication:** Another reason for ignoring uncertainty was caution in communicating uncertain results. There are concerns that including uncertainty might confuse the audience or undermine the perceived reliability of the data.
  - P4, who creates fact sheets with regional information about prostate cancer for the general public, chose not to include uncertainty information because of the "difficulty [...] of getting that across to an audience in a very concise, quick manner."

F2) Filtering by Uncertainty

Only keeping parts of the data with a specific range of uncertainty. During the discussion in step 1, we have made the uncertainty slider (see Section 3.1) a topic of discussion. All participants expressed the idea to use filtering to focus on data points that were more likely to represent "real" values, thus reducing the noise in their analysis. This approach helped them concentrate on what they perceived as the most reliable information. In order to deepen this discussion in step 3 of the study, we integrated an uncertainty slider into the visual prototypes to enable this strategy (see Figure 1-C). Related to this strategy, the following challenges were discussed:

- **Setting Uncertainty Filter Thresholds:** A significant challenge associated with filtering by uncertainty was determining the appropriate threshold for what constitutes acceptable uncertainty. Participants highlighted the trade-off between excluding too much useful data and retaining only what they deemed reliable. This balance is crucial because filtering out too much data could lead to missing important insights, while including too much uncertain data might dilute the clarity of the analysis.
  - Some participants saw the danger "to exclude too much useful data" (P1). While P1 and P3 found it reasonable to rely on a recommended threshold, such as the 60% probability used in the ACA, P5 described the selection of this threshold as "completely arbitrary". P4 emphasized that the choice of threshold might depend heavily on the context of the analysis and the audience for whom the results are intended. P4 pointed out that people needed statistical knowledge for meaningful filtering and that "it might be very much dependent upon [...] who I'm talking to and what I'm showing, which cancers out of which periods". This suggests that selecting a meaningful threshold may require domain knowledge, statistical knowledge, and an understanding of the target audience.
    As an alternative strategy for filtering by uncertainty, P3, P4, and P5 suggested a heuristic to use the credible intervals shown in the dot plot of the prototype. When the credible interval does not overlap with SIR=1 (or EHR=1) it is likely that the value is different from the average. This was proposed as a filtering strategy without a fixed, general uncertainty threshold, as defined by the uncertainty slider.
- **Visibility of Uncertain Values:** Another related aspect was whether parts of the data deemed too uncertain should be hidden from the data analyst or remain visible but visually deemphasized. There were arguments in favor of both visualization strategies.
  - On the one hand, participants wanted to hide uncertain data, to "basically blank out all of the other[...] areas" (P5) in order "to sort of ignore the noise and just focus on what I'm trying to actually get out of it" (P3). On the other hand, P3, although preferring to consider only values with low uncertainty, further suggested keeping an eye on the uncertain values: "it will definitely be something that perhaps I would monitor over time, [...] see if it changes". P4 supports this by stating that it makes sense to gray out the filtered values and "they're still visible if you put [...] your cursor on them [...] and that's [...] fine."

F3) Qualifying Hypotheses Using Uncertainty

Using uncertainty information to strengthen, question, or refine a hypothesis. All participants developed strategies to use uncertainty information to qualify their hypotheses based on the data from the ACA in the following ways:

- **Strengthening Hypotheses:** Participants indicated that when uncertainty was low (i.e., data points were more certain), they felt more confident in the validity of their hypotheses.
  - For example, if the data showed a significant difference in cancer incidence across regions with low uncertainty, participants were more likely to conclude that the observed pattern was meaningful and not due to random variation. P5, for instance, suggested that filtering by uncertainty may help "suggest or reinforce particular hypotheses."
- **Questioning Hypotheses:** Conversely, when uncertainty was high, participants were more cautious about their conclusions. High uncertainty prompted them to re-evaluate their initial assumptions, considering whether the observed patterns could be attributed to chance rather than a true underlying effect.
  - When confronted with hypotheses relying on uncertain values, participants would "interpret with caution" (P2) or "tend to discount or downrate the importance of that information" (P5), especially if that makes the data appear questionable: "I think that [the areas with above-average melanoma rates in Hobart are][...] a little bit suspicious but then again I would want to know whether the differences I'm seeing are due to chance or not" (P3).
- **Refining Hypotheses:** Participants also used uncertainty to refine their hypotheses. For instance, if initial data suggested a particular trend but included regions with high uncertainty, participants would consider adjusting their hypothesis to account for this uncertainty. This might involve exploring additional data sources, focusing on more reliable subsets of the data, or redefining the scope of the hypothesis.
  - P3 stated that in case of a spatial cluster of a high number of cases, the presence of highly uncertain values "might tell me that ascertainment [the process of identifying and documenting cases of a particular disease or health condition within a population] might be a problem", meaning that differences in case reporting could lead to the appearance of a hotspot. This implies that for this participant, uncertain values within a spatial cluster may indicate the need for further examination of the underlying reasons. This, in turn, could lead to the refinement of the hypothesis.
- **Avoiding Misleading Conclusions:** High uncertainty often led participants to downplay or dismiss certain data points, preventing them from forming potentially incorrect hypotheses based on unreliable information. This cautious approach underscores the importance of integrating uncertainty into the analytical process to avoid overconfidence in the results.
  - P5 explained that by filtering out highly uncertain data, they could "stop them grabbing your attention," and prevent "the formulation of [...] potentially incorrect hypotheses."

5. Recommendations for Uncertainty Visualization Design Supporting UnReSt

Based on our findings and the current state of research in this field, we present four recommendations as a first step towards guidelines for designing uncertainty visualizations that support UnReSt:

R1) Support Adapting or Developing UnReSt

As summarized in F1, one of the reasons why participants tended to ignore uncertainty was a lack of formal strategies for utilizing this information. This is consistent with past research suggesting that visualization authors often acknowledge the importance of visualizing uncertainty but end up omitting it [Hul20], Additionally, data analysts often do not apply the optimal strategy and even switch reasoning strategies during the analysis [KKH21]. However, when facing the bespoke visual prototype in step 3 of our study, our participants *developed* ideas for possible UnReSt. This is in line with research claiming that "different types of [uncertainty] visualizations prompt people to reason in different ways" [RPJ19].

Therefore, rather than designing uncertainty visualizations solely to support *existing* UnReSt of data analysts, visualizations should therefore integrate uncertainty in a way that prompts analysts to engage with it, helping them to recognize its importance and *adapt* or *develop* new, alternative, or improved UnReSt. While this puts new requirements to uncertainty visualization design it should nevertheless be part of future research.

R2) Do Not Facilitate Ignoring the Uncertainty

In F1 we stated that most of our participants had not used the uncertainty representations in the ACA simply because they did not know about them or did not see the need to engage with them. This reflects a key challenge in uncertainty research: understanding why people avoid uncertainty information and enabling them to change this. To address this, we recommend creating visualizations that *encourage* data analysts to incorporate uncertainty information. One approach could be to not only integrate uncertainty visually into the graphical representation of the data but adding graphical hints that make it obvious that uncertainty is displayed. Another potential solution could involve interface design that integrates uncertainty information as an integral part of the functionality, rather than treating it as additional information accessible only in an 'expert mode'.

To support these goals, visualization designs could include features such as mandatory uncertainty overlays, prompts or alerts when uncertainty is being ignored or filtered, and interactive elements that require users to acknowledge and address uncertainty before proceeding with their analysis.

R3) Support Defining Acceptable Levels of Uncertainty

Based on the findings in F2, our participants generally considered filtering by uncertainty to be useful for their tasks. However, the question of how they define a meaningful filter threshold remained largely unresolved. Some participants preferred using a fixed threshold for the filtering process to avoid this decision. This issue is related to the general question about the usefulness of thresholds (such as p-values) in statistical reporting in visualization [Dra16].

As a step forward, a feasible approach could be to extend the concept of the uncertainty slider to enable analysts to understand the distribution of uncertainty better. For example, this could involve animating the effect of different thresholds using animated transitions [HR07]. Another approach would be to extend the concept of filtering so that not a single binary threshold defines the filter but rather some kind of fuzzy filter. Further research is necessary to understand what the requirements are for defining acceptable levels of uncertainty by data analysts for typical tasks from various domains.

R4) Do Not Hide Uncertain Parts of the Data By Default

Findings in F2 highlighted the challenge of filtering by uncertainty, where analysts might exclude uncertain data altogether, and findings in F3 suggested that some participants recognized the value of removing uncertain values from the data to avoid incorrect hypotheses. In F2, however, two participants also acknowledged that *even highly uncertain data can be valuable for interpretation*. This aspect is often overlooked, as people tend to avoid uncertainty and try to minimize or ignore it [BPHE17]. To address this, visualizations should avoid designs that allow for the easy exclusion of uncertain data without careful consideration. Instead, visual tools should guide users in understanding the implications of filtering and ensure that uncertain data remains accessible or is visually marked, even if de-emphasized.

6. Reflecting on the Methodology

In order to explore our research question we developed a novel methodology for our study as described in Section 3. Our user study applied this methodology, including a small number of participants and focusing on a specific domain, region, analysis tool and dataset. This allowed for a detailed, in-depth analysis of reasoning strategies used by data analysts, without claiming generalizability. There are a number of lessons we learned during the course of the study. In the following we describe aspects that we see as novel and worth reflecting. For each aspect we describe the challenge and our learning, as well as a recommendation based on our experience.

6.1. Finding Expert Participants

A core requirement of our approach was to recruit participants with a specific expertise in the domain (here: epidemiology) and experience with a specific visual analysis tool (here: the ACA). The rationale was to allow us to examine reasoning strategies in context.

Challenge: Finding experts with the described background who are willing to volunteer for a study, particularly for full-day workshops, is challenging. Additionally, identifying experts who use a specific tool, especially when they are geographically dispersed and affiliated with various institutions, presents another significant hurdle.

Learning: In our case, identifying experts with the required background was not difficult, as we had a facilitator (although the number remained small). To lower the threshold for participation, we arranged two half-hour meetings per participant. Combined with the fact that the sessions were held online (with dates and times requested by participants), this lowered the cost of time for participants and greatly facilitated the search.

Recommendation: We believe that the strategy of scheduling multiple shorter meetings in an online environment was a successful choice and recommend it for recruiting domain experts.

6.2. Defining Participants' Expertise

To learn about how domain experts handle uncertainty during data analysis in the field of epidemiology, our goal was to recruit experts from the field who use the ACA (and the dataset it is based on) for their work.

Challenge: While the experience with the ACA varied, to our surprise, none of the participants were familiar with the uncertainty visualizations it provides (although there are accessible tutorials introducing them).

Learning: This raises a point that is rarely addressed: just because a tool provides representations of uncertainty does not guarantee that they are used. In our case, however, this did not limit the value of the discussions, as participants were aware of the uncertain nature of the data.

Recommendation: We recommend extending our approach by conducting a preliminary survey among potential participants to assess their familiarity with the visual analysis tool. This will ensure that participants possess the necessary expertise in handling uncertain data within the specific domain and have practical experience using the tool.

6.3. Studying a Familiar Software Tool

We based our case study on the ACA (and the data it provides) that all participants had already used. This allowed us to capture participants' strategies for dealing with data uncertainty in a context they were familiar with.

Challenge: This strategy limits the pool of potential participants, making recruitment more difficult.

Learning: Although most participants had not been using the ACA as a central part of their tasks, their familiarity with the data represented in the tool was crucial for the study's objectives. Overall, this enabled us to delve deeper into discussions about the role of uncertainty during the analysis.

Recommendation: We recommend our strategy to involve tools and data that are familiar to the participants, despite the additional effort required for recruitment.

6.4. Studying Familiar Tasks

Our approach asked participants to prepare tasks from their work in advance. This ensured the evaluation was based on realistic tasks they have experience with instead of predefined tasks from the authors of the study that might not reflect their practice. Tasks ranged from simple retrieval of values from the tool to analyzing geographic patterns in incidence rates or examining the correlation between incidence and excess deaths (see Subsection 3.3).

Challenge: This approach, while overall successful, meant that tasks were usually not comparable between participants. This also makes the comparison between participants' UnReSt difficult.

Learning: Despite the limited comparability, asking participants to demonstrate a task from their work during the think aloud session in step 1 of the study was a successful way to learn about their workflows and visualization needs.

Recommendation: We recommend working with participants' tasks instead of generating artificial tasks. Asking participants for a specific *type* of task may make comparison more feasible, but limits scope, which is a delicate trade-off.

6.5. Creating Individual Prototypes for Each Domain Expert

In step 2 of our study, we created visual prototypes using Observable notebooks in the fashion of data sketches [LD11]. Their purpose is to provide an interface providing the data and uncertainty visualization tailored to each participant's task to stimulate discussions about UnReSt (see Figure 1).

Challenge: The scalability of the methodology is limited because a separate prototype must be designed for each participant.
Learning: Our approach facilitated rapid prototyping and easy deployment since no software installation was required on the part of the participants prior to the meetings. Reusing parts of the visualizations (e.g., the map or the dot plot) reduced the effort of designing bespoke visualizations.
Recommendation: We recommend our approach of using Observable notebooks or similar platforms for visual prototype development. In addition to the aforementioned reuse of parts of the visualizations, a single design session (instead of individual ones) can reduce the effort.

6.6. Including Visualization Experts in the Collaborative Design Phase

During the prototyping phase in step 2 of the study, we brought in visualization experts to support us with the design of the prototypes. This helped to both limit the risks of design bias associated with individual design decisions and increase the quality of the visual prototype design. We chose to recruit one visualization expert per prototype, rather than one expert for all designs, to limit individual workloads and include multiple perspectives.

Challenge: Finding five visualization experts to volunteer was a challenging task.
Learning: It proved advantageous that in step 2—as in steps 1 and 3 with the domain experts—we limited the duration of the meetings to half an hour. Overall, our approach yielded justified designs sampled from a variety of design approaches and perspectives. This helped mitigate bias, although it is important to note that all designers used similar tools and had comparable backgrounds.
Recommendation: We recommend our approach because we found it helped minimize biases resulting from a single designer's perspective and led to further insights about the nature of uncertainty visualization during the design sessions. If possible, it may be worthwhile to hold co-creation workshops with multiple visualization experts simultaneously so as to leverage the interaction between the experts.

6.7. Timing of the Meetings

We scheduled one month between steps 1 and 3 to develop the visual prototypes and plan the second round of meetings with the domain experts, based on insights gained from the first round.

Challenge: The time between meetings was relatively short for developing prototypes, particularly because each prototype required discussion with a visualization expert. Consequently, some prototypes were not as advanced as desired. However, rescheduling the second meeting to allow more time was impractical due to the risk of potential scheduling conflicts.
Learning: It is challenging to determine an optimal interval between meetings in advance. Furthermore, estimating the necessary time for prototype development is difficult as it depends on the outcomes of the first round of meetings and the availability of visualization experts. Nevertheless, we argue that the longer the break between meetings, the more likely crucial details from the first meeting may be forgotten, potentially disrupting the continuity of the process.
Recommendation: For the timing of the study, we recommend taking into account the capacity for analysis of the first round of meetings (step 1), as well as the time needed for design and coding, and the availability of visualization experts during the prototyping phase (step 2). Our initial choice of one month proved to be a reasonable guideline.

7. Conclusion

This work does not yet answer how we can design uncertainty visualizations that effectively support the UnReSt of data analysts. However, it confirms previous research suggesting that a better understanding of reasoning strategies is key to effective uncertainty visualization and provides new insights on how we can take a step towards supporting UnReSt. Since our findings did not focus on aspects specific to epidemiology, we are confident that they are also applicable to other domains.

The initial recommendations serve as a basis for future guidelines on the design of uncertainty visualizations to support UnReSt. To advance toward developing these guidelines, we recommend adapting and refining our study methodology to gain deeper insights into data analysts' UnReSt and how uncertainty visualization can effectively support them. The next step will be to conceptualize how these design guidelines can be developed, given the recommendation that uncertainty visualizations should support the adaptation or development of UnReSt for data analysts with different tasks and backgrounds.

Acknowledgments

The authors wish to thank Peter Baade (Cancer Council Queensland) and Sarah Goodwin (Monash University Melbourne) for their support with the ACA, the data behind it, and the recruiting process. This work was supported by a research fellowship from the German Academic Exchange Service (DAAD).

References

[BAL12] BRODLIE K., ALLENDES OSORIO R., LOPES A.: A Review of Uncertainty in Data Visualization. In *Expanding the Frontiers of Visual Analytics and Visualization*, Dill J., Earnshaw R., Kasik D., Vince J., Wong P. C., (Eds.). Springer, London, 2012, pp. 81–109. doi:10.1007/978-1-4471-2804-5_6.

[BC21] BRAUN V., CLARKE V.: *Thematic Analysis: A Practical Guide*. SAGE Publications, 2021. URL: https://us.sagepub.com/en-us/nam/thematic-analysis/book248481.

[BPHE17] BOUKHELIFA N., PERRIN M.-E., HURON S., EAGAN J.: How Data Workers Cope with Uncertainty: A Task Characterisation Study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17* (Denver, Colorado, USA, 2017), ACM Press, pp. 3645–3656. doi:10.1145/3025453.3025738.

[Dra16] DRAGICEVIC P.: Fair Statistical Communication in HCI. In *Modern Statistical Methods for HCI*, Robertson J., Kaptein M., (Eds.). Springer International Publishing, Cham, 2016, pp. 291–330. Series Title: Human–Computer Interaction Series. doi:10.1007/978-3-319-26633-6_13.

[DW14] DÖÖS M., WILHELMSON L.: Proximity and distance: phases of intersubjective qualitative data analysis in a research team. *Quality & Quantity 48*, 2 (Mar. 2014), 1089–1106. doi:10.1007/s11135-012-9816-y.

[HR07] HEER J., ROBERTSON G.: Animated Transitions in Statistical Data Graphics. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (Nov. 2007), 1240–1247. doi:10.1109/TVCG.2007.70539.

[Hul20] HULLMAN J.: Why authors don't visualize uncertainty. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (2020), 130–139. doi:10.1109/TVCG.2019.2934287.

[KKH21] KALE A., KAY M., HULLMAN J.: Visual reasoning strategies for effect size judgments and decisions. *IEEE Transactions on Visualization & Computer Graphics 27*, 02 (feb 2021), 272–282. doi:10.1109/TVCG.2020.3030335.

[KMRS15] KINKELDEY C., MACEACHREN A. M., RIVEIRO M., SCHIEWE J.: Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science 44*, 1 (Jan. 2015), 1–21. doi:10.1080/15230406.2015.1089792.

[KMS14] KINKELDEY C., MACEACHREN A. M., SCHIEWE J.: How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies. *The Cartographic Journal 51*, 4 (Nov. 2014), 372–386. doi:10.1179/1743277414Y.0000000099.

[Las86] LAST J. M.: A dictionary of epidemiology. *International Journal of Epidemiology 15*, 2 (1986), 277–277. doi:10.1136/jech.47.5.430.

[LD11] LLOYD D., DYKES J.: Human-Centered Approaches in Geovisualization Design: Investigating Multiple Methods Through a Long-Term Case Study. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec. 2011), 2498–2507. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi:10.1109/TVCG.2011.209.

[LF05] LEMAIRE P., FABRE L.: Strategic aspects of human cognition: Implications for understanding human reasoning. In *Methods of thought: Individual differences in reasoning strategies*, Current issues in thinking and reasoning. Psychology Press, New York, NY, US, 2005, pp. 11–29. doi:10.4324/9780203503881.

[Mac15] MACEACHREN A. M.: Visual analytics and uncertainty: Its not about the data. In *EuroVis Workshop on Visual Analytics EuroVA* (2015), The Eurographics Association;, pp. 55–59. doi:10.2312/eurova.20151104.

[PKH21] PADILLA L., KAY M., HULLMAN J.: *Uncertainty Visualization*. John Wiley & Sons, Ltd, 2021, pp. 1–18. doi:10.1002/9781118445112.stat08296.

[PRJ12] POTTER K., ROSEN P., JOHNSON C. R.: From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches. In *Uncertainty Quantification in Scientific Computing* (Berlin, Heidelberg, 2012), Dienstfrey A. M., Boisvert R. F., (Eds.), IFIP Advances in Information and Communication Technology, Springer, pp. 226–249. doi:10.1007/978-3-642-32677-6_15.

[PWL97] PANG A. T., WITTENBRINK C. M., LODHA S. K.: Approaches to uncertainty visualization. *The Visual Computer 13*, 8 (Nov. 1997), 370–390. doi:10.1007/s003710050111.

[RPJ19] REANI M., PEEK N., JAY C.: How different visualizations affect human reasoning about uncertainty: An analysis of visual behaviour. *Comput. Hum. Behav.* (2019). doi:10.1016/j.chb.2018.10.033.

[SWM21] SABBURG L., WOODLEY A., MENGERSEN K.: A data-and task-oriented design framework for bivariate communication of uncertainty. *arXiv preprint arXiv:2112.06921* (2021). doi:https://doi.org/10.48550/arXiv.2112.06921.