ELSEVIER

Special Section on EG VCBM 2024

Interactive data comics for communicating medical data to the general public: A study of engagement and ease of understanding

Melissa Fogwill, Areti Manataki *

School of Computer Science, Jack Cole Building, North Haugh, St Andrews, KY16 9SX, UK

ARTICLE INFO

ABSTRACT

We are experiencing a health literacy crisis worldwide, which has alarming effects on individuals' medical outcomes. This poses the challenge of communicating key information about health conditions and their management in a way that is easily understood by a general audience. In this paper, we propose the use of data-driven storytelling to address this challenge, in particular through interactive data comics. We developed an interactive data comic that communicates cancer data. A between-group study with 98 participants was carried out to evaluate the data comic's ease of understanding and engagement, compared to a text medium that captures the same information. The study reveals that the data comic is perceived to be more engaging, and participants have greater recall and understanding of the data within the story, compared with the text medium.

1. Introduction

We are experiencing a health literacy crisis worldwide, with nearly half of all Europeans having inadequate or problematic health literacy skills [1]. This is a major issue, as poor health literacy adversely affects people's health [2,3]. With medical information prominent in every individual's life, it is essential to think about how to address this issue to reduce the alarming outcomes it has upon a large proportion of the worldwide population. Our paper pioneers a method to enhance recall and understanding of medical information in an engaging way through interactive data comics.

Data comics are built on the foundations of traditional comics' concept of a sequence of panels where text and images convey a series of messages [4]. Data comics focus on delivering information that we may normally see within the data visualisation sphere, incorporated into the narrative of the comic.

Even though data comics have been developed for immigration [5], geography [6], and social justice [7], there is a scarcity of data comics communicating medical information. In this paper, we address this research gap by developing a data comic that communicates cancer data to the general public, i.e. non-expert adults.

Our work heightens the benefits of data comics through the additional element of interactivity by allowing readers more control over the story that we, the authors, have generated. Therefore, the interactivity can seek to improve their understanding with a more engaging

and fun format to 'personalise' their own story, spending more time on aspects that interest them and quickly passing elements less so.

We consider engagement and ease of understanding to be key success factors of a data comic that is motivated by the health literacy crisis and which can serve as a public health intervention. Considering these two aspects, we carry out a study to compare the data comic to a text version that replicates how current public-facing cancer information is given, ensuring the data visualisation and information remain the same within both formats. We define the following two research questions:

RQ1 *How does engagement of the medical data comic compare to that of the text form?*

RQ2 *How does ease of understanding of the medical data comic compare to that of the text form?*

Through a between-group study with 98 participants, we compared these formats and collected quantitative and qualitative data from participants through an anonymous online survey. Feedback suggests that the data comic is more engaging, resulting from quantitative measures surrounding their opinion of the medium's visual appeal, ability to focus, enjoyability, engagement, and the desire to spend additional time exploring. The data comic participants also recall and answer factual questions regarding the medical information more accurately than the text group. Interestingly, those who interacted with the data comic also stated to be more likely to attend a cancer screening appointment than their counterparts who were exposed to the text format.

---

* Corresponding author.
E-mail addresses: mf246@st-andrews.ac.uk (M. Fogwill), A.Manataki@st-andrews.ac.uk (A. Manataki).

2. Background

2.1. Health literacy

Health literacy is described as the basic and personal skills to read, write, and understand health-related information, allowing patients to make suitable health decisions [8]. Despite the growing understanding of the importance of health literacy [8,9], a study conducted by the World Health Organisation (WHO) in 2013 found that health literacy among 48% of adults was problematic or inadequate [1]. This is concerning, as poor health literacy is associated with increased rates of hospital admissions, ineffective use of preventive services, and adverse mental and physical health, which can also lead to death [2,3]. It is, thus, imperative to develop accessible tools that enable people to make informed health decisions. We believe that data comics are a way to achieve this.

2.2. Graphic medicine

Graphic medicine is a related field, which involves the use of comics in medicine and health [10,11]. Comics are recognised as an intuitive, engaging and accessible method of communicating medical narratives [10,11], allowing readers to empathise with characters [12]. Comics have been successfully used in a variety of healthcare settings and for a range of purposes, including patient education in the context of informed consent [13], public heath communication during the COVID-19 pandemic [14], to promote cancer screening [15], towards improving medication compliance [16], to communicate patient experiences of illness [17], as well as in undergraduate and postgraduate medical education [18].

2.3. Memorability and persuasion in data visualisation

Research in data visualisation has shown that visual embellishments and unique types of visualisations can improve memorability [19,20], while titles and supporting text help with message recall [21]. It has also been found that data and message redundancy in visualisations help with understanding, and improve recognition and recall [21]. These findings are important to consider when developing visualisations to be included in data comics, too.

Presenting data through visualisations is more persuasive compared to a table-based presentation, especially if people's initial attitude is not strongly polarised [22]. Data visualisation interventions have been found to have a positive impact on cognitive and behaviour change, including decision-making, attitude, motivation, understanding, engagement and perception [23]. Recent studies on affective visualisation design [24,25] are particularly relevant here.

Recent research in the context of public health has shown that effectively designed visual aids can improve risk assessments, helping people make decisions that they feel more confident about [26]. The fact that these benefits are strongest for people with low numeracy or low risk literacy [26] is particularly relevant for our work.

2.4. Narrative medical visualisation

Narrative medical visualisation is a relatively new but growing field of study that combines traditional storytelling techniques with data visualisation in medicine and health [27]. It helps to tell data-driven medical stories in an engaging and memorable fashion, with great potential for educating patients and the general public about disease occurrence, treatment and prevention [27].

Key considerations when creating data-driven medical stories for a broad audience include narrative genres and structures. Segel and Heer [28] identify seven genres of narrative visualisation: magazine style, annotated chart, partitioned poster, flow chart, comic strip, slideshow, and film/video/animation. Several of these have been employed for creating data-driven medical stories, including slideshows [29–32], partitioned posters [31] and videos [33]. However, to the best of our knowledge, the genre of data comics has been underexplored in the context of medicine.

In terms of narrative structures, Meuschke et al. [31] propose a seven-stage template for telling a disease story to a broad audience, while in their slideshow-based narrative visualisation about cerebral small vessel disease, Mittenentzwei et al. [32] follow the character-driven structure "The Hero's Journey", which describes typical stages that a hero goes through during a story. Freytag's pyramid, which divides a story into five acts (exposition, rising action, climax, falling action, and denouement), is followed in a data-driven story that communicates how blood flow data is used to diagnose and treat a cardiovascular disease [29].

Personalising data-driven medical stories can encourage the audience to empathise and relate to the story [27], and has great potential for promoting behavioural change [33]. Personalisation is often achieved through the choice of narrator, which typically is a patient [29–32] or a clinician [32] in the story.

Recently, Meuschke et al. [31] investigated the impact of data-driven interactive medical stories on a general non-expert audience. A between-subject study was conducted, comparing an interactive story containing 3D liver visualisation to a traditional web blog design about liver cancer. The study found that the interactive medical story was associated with higher memorability and enjoyment, and it contributed to a shift in personal perception of individual risk. This study has been a source of inspiration for the research presented in this paper.

2.5. Data comics

Data comics are a relatively novel approach to information-driven storytelling within the data visualisation realm [4]. They combine comic strip elements (e.g. sequence of panels) with data visualisation techniques to tell a story and convey information in data. According to Bach et al. [4], there are four essential components of data comics: visualisation (ranging from realistic to abstract), flow (ranging from undirected to directed), narration (ranging from factual to narrated), and words and pictures (ranging from text-heavy to image-heavy). Data comics have been found to improve understanding and to be more engaging and more enjoyable compared to infographics [34] and PowerPoint-based slideshows [35].

A growing number of data comics have been created in the last few years for a range of domains, including history [36], geography [6], immigration [5] and cybersecurity [37]. However, there is a lack of data comics for medicine and health. Encouraged by the positive results regarding data-driven interactive medical stories by [31], we set out in this paper to address the research gap surrounding data comics for communicating medical data to the general public.

3. Design approach

We decided to communicate cancer data through data comics, as compared to other narrative genres, data comics have a unique combination of characteristics that work particularly well for communicating medical information to low health literate audiences: (1) They are effective in helping users to understand data, and their sequential nature allows to split information into smaller, more manageable chunks [34]. (2) They provide quick access to content across a page, allowing users to skim the content first to get an overall view and then move on to more detailed reading [38]. This is not the case with videos or slideshows, for instance. (3) According to a recent study comparing data comics to infographics, data comics are more engaging and more enjoyable [34]. (4) The familiar and non-threatening visual language of comics [39] makes data comics accessible for diverse audiences. (5) It is an appealing medium, especially for people that are intimidated

by media associated with academic and corporate worlds, such as annotated charts and slideshows.

For the creation of the data comic, we adopted ideas from user-centred design [40] and we invited a cancer care nurse, with experience in communicating medical information to the general public, to join as a design participant for the data comic. This is in line with recommendations regarding medical expert involvement in the process of developing data-driven medical stories [27]. Given that the target audience is the general public, working with an expert in this area was an effective solution to the practical and ethical challenges associated with working with members of the public. The nurse provided feedback throughout the main design phases (i.e. requirements elicitation, low-fidelity paper-based prototype design, and high-fidelity prototype development), paying particular attention to the narrative, the clarity of the message and the data illustrated. Consultations took place in the form of unstructured discussion and lasted for approximately an hour each.

Upon eliciting requirements, we designed the data comic following the Five Design Sheet methodology [41], which structures the design method for sketching visualisations. Even though this methodology is not widely used for data comics, we found it to be valuable for developing sketches from ideation to realisation. Three distinct design directions for the data comic were generated through this process. Inspired by participatory design methods [42], low-fidelity paper prototypes were generated to allow for feedback quickly and early in the design process, therefore making it less challenging to change at later stages.

The final, high-fidelity data comic was created using Adobe Illustrator and Tableau. The latter offers functionalities regarding layout, animation, and interactivity, which are useful for creating a more exploratory data comic that is publicly available to all.

4. The data comic

4.1. Data

The primary source of information was synthetic cancer data from the Simulacrum v1.2.0 datamart [43]. Simulacrum data imitates the statistical distributions of Public Health England National Cancer Registration and Analysis Service's (NCRAS) real information, without being constrained by patient confidentiality restrictions.

Our analysis focused on all cancer-related disease codes from the International Classification Disease, tenth revision (ICD-10), in particular C00-C97 and D00-D48. Exploratory analysis helped us identify key cancer statistics worth communicating through the data comic, such as incidence and mortality. We included additional datasets, such as the Office of National Statistics (ONS) English age population data, which allowed us to calculate age-standardised cancer incidence rates. We also included death-related data, as death is an unavoidable topic when discussing cancer.

4.2. Story

The next stage of creating the data comic surrounded story creation. The narrative intent of the data comic is to inform a broad audience on cancer incidence and mortality, with a key idea being that not only old people are at risk. A second intent is to raise awareness about cancer prevention, and the related take-home message is that it is possible to reduce one's risk of developing cancer. The initial step was to create a persona that identifies the target audience to ensure that all design decisions such as language tone, characters, colour palette, and style of the data comic were all considered choices [44]. Based upon our analysis, the persona of the data comic was identified as UK females. Females were chosen, as the data identified strong trends within age groups and the story naturally honed in on these; and the UK, as the data is English in origin, and the more relatable and comprehensible
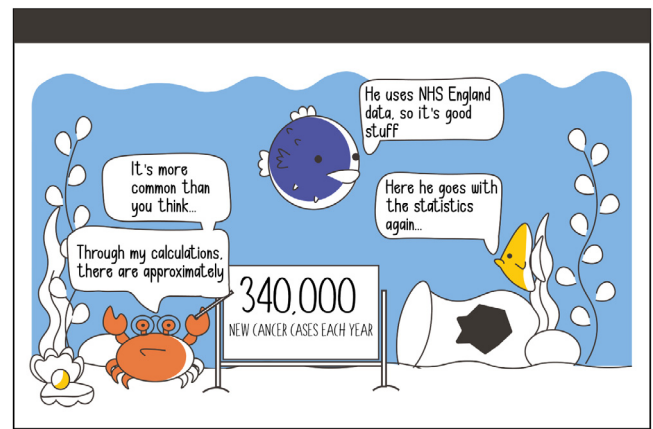

Fig. 1. The first scene that introduces the protagonist that will guide readers through the data.

the information and messages are to the audience, the more it resonates with them.

The user-centred design approach proved particularly useful for improving and finalising the story. For example, the conclusion included a call to action for the audience, for which we trialled differing tones and discussed them with the cancer nurse. These included a soft, emotional plea to check yourself for cancer; a warning/telling off; and a more light-hearted reference to the characters within the data comic.

The final story begins with a doctor diagnosing one of his patients with cancer. Within his office there is a fish tank, in which a crab and two fish live. Over the years, the characters within the tank have watched patients come and go, picking up medical terminology and the statistics surrounding cancer diagnoses. The crab (McCrab) has a keen interest in statistics and has kept track of the facts and figures surrounding patients' journeys. McCrab uses objects he can see within the tank, doctor's office and outside to help visualise the data he has collected and debunks misconceptions surrounding cancer, such as 'it just affects old people'. McCrab then highlights how cancer impacts gender-age groups differently and reinforces the importance of attending cancer screening appointments and looking after your body. The take-home message is that by doing this, the reader can avoid being another cancer statistic for McCrab to use.

4.3. Interactive data comic

The final data comic is accessible via https://tinyurl.com/2pe46vrm . Fig. 1 shows the story's protagonist, 'McCrab', discuss cancer incidence rate. The data source is mentioned too, which is an important aspect of credibility in data-driven storytelling in medicine [27]. Combining fact and fiction to accentuate and elucidate the data visualisations is where the medium breaks down barriers for low-health literate audiences. The different characters support the narrative and interpretation of cancer data, as they provide explanations and point out things that the users should not miss. For example, one of the fish say "look at how big the difference is" in a pane that utilises visual juxtaposition [45] to highlight the difference in cancer incidence between genders of the same age demographic.

We draw on the storytelling power of data comics to work around the fact that very large or incomprehensible numbers do not resonate with the audience. Breaking such numbers into digestible and relatable terms enhances the reasoning of why they should care, making the information more memorable [46]. We map the '340,000 new cancer cases' statistic in Fig. 1 to a night and day background that shows '932 cases a day', over which there is a clock that shows '40 new cases an hour'.
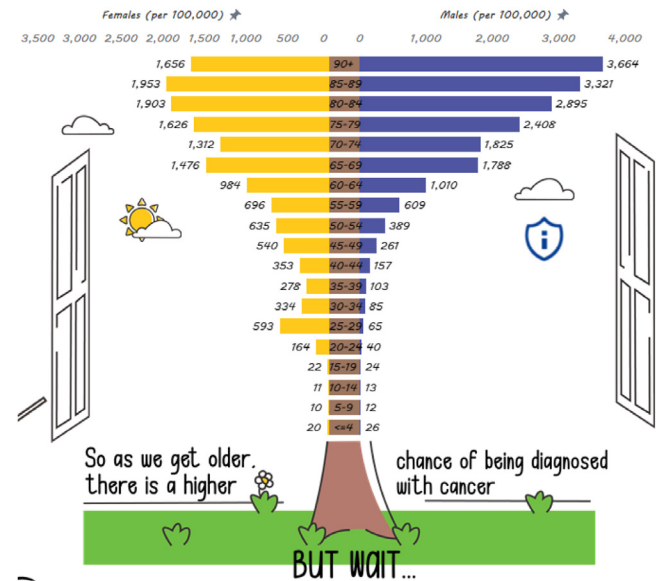
Fig. 2. The first interactive data comic pane illustrating cancer incidence by gender and age, incorporated into a tree.
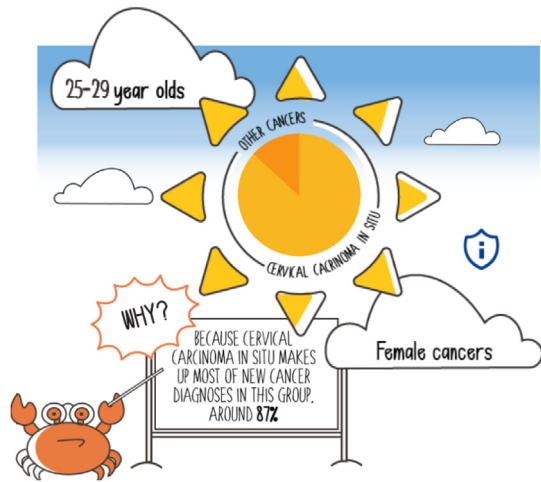

Fig. 3. The interactive sun data comic pane, seamlessly integrated into the data comic theme.

Fig. 4. An extract from the medical text, which communicates the same information as the data comic. This illustrates that the same data visualisations are included as in the data comic.

The visual theme of the data comic, in terms of colour palette and font, flows through all visualisations; however, ensuring that this consistency and flow was kept in the interactive visualisations was more challenging. Innovative solutions were required to overcome these challenges. An example here involves the 'stat tree', whereby we have integrated the commonly developed population age pyramid into the data comic (Fig. 2). Back-to-back histograms represent the tree's crown showing the incidence rate for all cancers of the two genders separately, by age group. The interactivity of this pane encourages the reader to explore the data in their own time whilst also guided by the narrative. Another example of seamless integration of a fully interactive data visualisation into the data comic is the interactive sun (Fig. 3). This combines an interactive pie chart created in Tableau, which represents the centre of the sun, with the sun perimeter created using Adobe Illustrator.

An important message that the data comic aims to get across is that 25–29 year-old females have a higher risk of cancer diagnosis compared to some older groups (i.e. 593 females per 100,000). This statistic is explicitly highlighted, utilising concretisation of the cancer
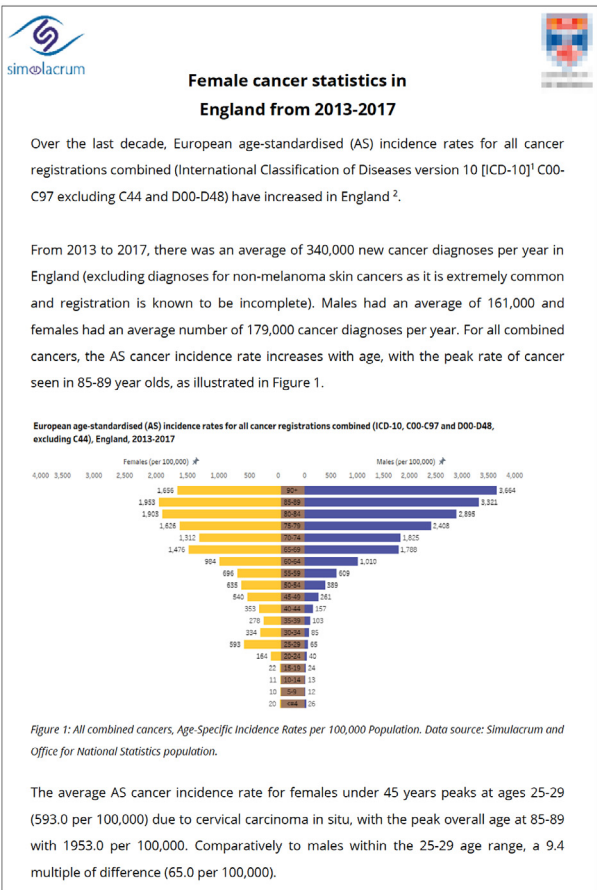
incidence rate to illustrate an abstract number into concrete objects, in this case as female interactive shapes [45]. To emphasise this message, we compromised the accuracy of the visualisation to create a more simplified and memorable panel, a trade-off that is worthwhile when the audience is of low-health literacy [46].

Simplified and concise language is used throughout the data comic to ensure it is inclusive of the low-health literacy readers. For example, 'colorectal cancer' is exchanged for 'bowel cancer', and 'likelihood of survival' is reworded to 'chance of death'. Where deemed appropriate, concepts are visually explained. For instance, in order to explain stage zero cancer, we visually depict cancer stages through cancer cells' position on a common scale, increasing in area and angry facial expressions to describe cancer stages in an interactive and simplified way.

5. The study

A between-group user study was conducted, focusing on the differences in a general population's experience with either the interactive data comic or a text version, both including the same medical information and data visualisations. The text version uses the same language and terminology as Cancer Research UK's public facing website (see Fig. 4).

5.1. Hypotheses

The hypotheses of our project were developed upon current literature while considering a universal, non-specialist audience:
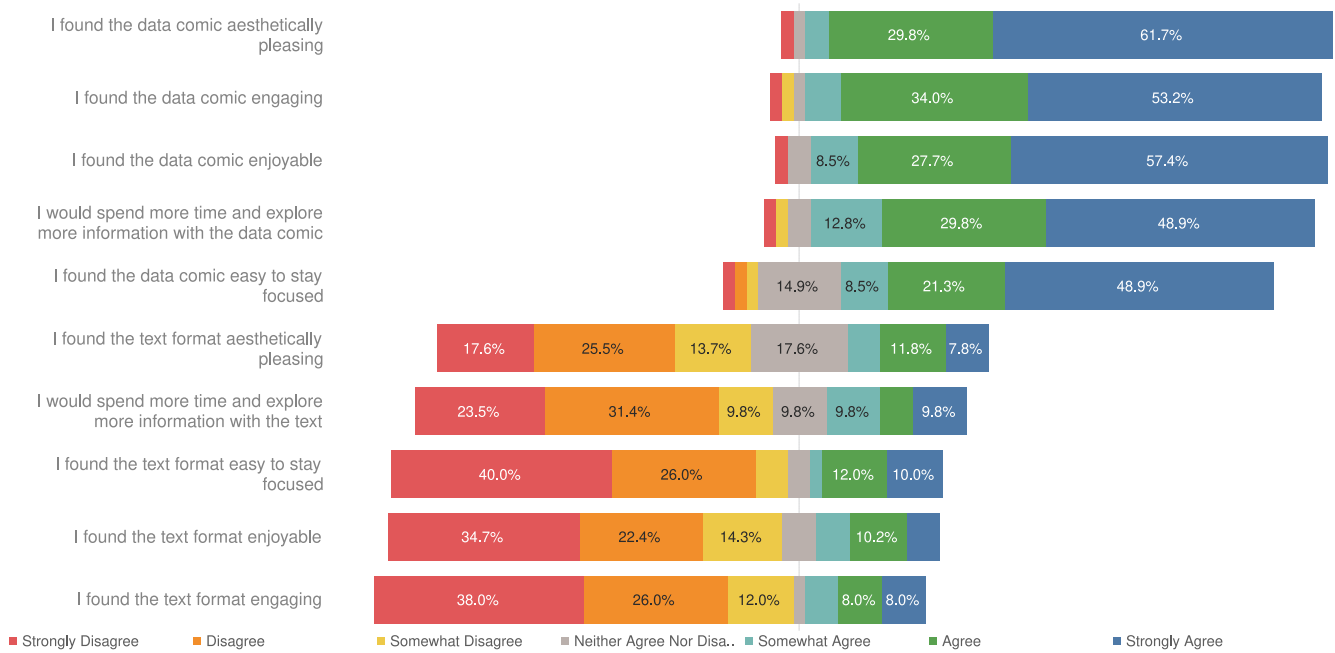
Fig. 5. Feedback received from participants in each group regarding engagement.

- **Engagement** is expected to be rated higher for the data comic than the text. Therefore, we believe that readers will find the data comic a more enjoyable medium that they will be more willing to spend time reading than the text.
- **Ease of Understanding** is expected to have participants recall the data and messages more accurately from reading the data comic than the text. Therefore, we believe there will be a lower error rate for the data comic compared to text.
- **Self-Assessed Likelihood of Attending Screening** is expected to be more positively influenced than the text. The characterisation and narrative of a data comic can impact readers emotionally, encouraging them to feel more empowered to act upon the take-home message.

5.2. Participants

Our study focuses on evaluating how a universal, non-specialist audience interprets two mediums. To maximise accessibility to the general public, recruitment for participants was via social media (Facebook, Twitter, Reddit) and through the University of St Andrew's internal mailing lists for Computer Science students and staff.

In total, 98 participants completed the study, 47 were assigned to the data comic group (Group1) and 51 to the text group (Group2). Group1 includes 29 women and 18 men, while Group2 includes 28 women, 21 men and 1 non-binary. Ages range between 18 and 63 years old, with the following distribution per group (Group1/Group2): 1/0 18–19 yr., 25/26 20–29 yr., 8/12 30–39 yr., 3/4 40–49 yr., 3/7 50–59 yr., 3/1 60–69 yr. and 4/1 not stated. The industries that participants work within range vastly, from healthcare, construction, engineering, retail, to retired, which gives the study a more accurate representation of the general public than approaching specific demographics separately. No clear differences were observed in terms of industries between the two groups.

5.3. Procedure

Our study employed a between-group design, each group was either given the data comic or the text. The online anonymous survey was created using Qualtrics and to ensure that the baseline participant characteristics are comparable across groups, participants were randomly assigned to one of the medical information mediums. The between-group design benefits from removing any potential 'learning effect' that could happen if a within-group design was employed; exposure to both mediums would affect the ease of understanding section of the survey.

Prior to consenting to participate in the study, participants were informed about the main aims of the study. Upon reading and interacting with the assigned medium (a link was provided), participants were invited to answer questions related to the following aspects: (1) Engagement, which was inspired by [34]. We measured engagement by assessing the reader's enjoyment, willingness to spend time exploring, attention, aesthetic feeling, and perceived engagement with the medium that they were presented with by using 7-point Likert scale questions (see Fig. 5 for list of questions). (2) Qualitative comments on the medium, which were gathered via a free-text box within the survey (see Supplemental Material for study question). (3) Participants were asked seven multiple-choice questions covering single facts to evaluate the ease of understanding of the data and messages within the chosen medium. All questions had five options, including 'I'm not sure' and both groups received the same questions. As shown in Table 1, the questions were a mix of visualisation-supported and narrative-only facts. (4) The perceived likelihood of attending screening after reading the corresponding medium was assessed through a 5-point Likert scale question.

6. Results

6.1. Engagement

Overall, participants found the data comic considerably more engaging than a text article with the same information. Fig. 5 visualises the survey results, sorted in descending order from the most positive feedback to the poorest responses at the bottom. The most positive feedback emerges the data comic group, with participants finding it very **aesthetically pleasing**. 91.5% of participants who were exposed to the data comic agreed or strongly agreed that it is aesthetically pleasing (median = 7), while this was the case for only 19% of participants exposed to the text medium (median = 3). The difference

Table 1

Ease of understanding: list of survey questions and answer source within each medium.

Number	Question content	Answer source within the data comic	Answer source within the text medium
Q1	How many people in England are diagnosed with cancer each year?	Narrative	Text
Q2	For females under age 49, what age range do we see a peak in cancer diagnoses?	Visual aids and narrative	Visualisation and text
Q3	How many more females in this peak age range are diagnosed compared to males in the same age group?	Visual aids and narrative	Text
Q4	What cancer type is mostly diagnosed for the younger females in the peak age group?	Visual aids and Narrative	Visualisation and Text
Q5	When we consider females of all ages, what is the most common cancer?	Visual aids and narrative	Visualisation and Text
Q6	What is the average chance of death after a lung cancer diagnosis?	Narrative	Text
Q7	How many female cancer diagnoses could be prevented?	Narrative	Text

between the two groups was found to be statistically significant based on a Mann–Whitney test (U = 243, p <.001).

Perceived engagement is considerably different in the two groups, with the median for the data comic being 7, while for the text medium it is 2. A Mann–Whitney test showed that this difference was significant (U = 241.5, p <.001). This agrees with Wang et al.'s [34] research on static data comic's engagement compared with infographics. Engagement for the data comic was received very positively, with 93.6% of participants rating it as engaging. Conversely, it was also the worst-rated for the text format, with 76.0% of responses disagreeing with the statement. It is concerning that the text was considered so disengaging as this is how most medical information is currently communicated.

Participants' **enjoyment** of the formats varied significantly, again having a much a more positive experience with the data comic (median = 7) compared to the medical text article (median = 2). A Mann–Whitney test showed that this was significant (U = 215.5, p <.001), suggesting that a real difference in enjoyment between the mediums was experienced.

91.5% of participants who received the data comic responded positively to the proposal of spending **more time exploring** further information with this type of format. A considerable difference is illustrated when compared to the text format, where only 25.5% agreed with this proposition. A Mann–Whitney test showed that the difference between the two groups was significant (U = 308, p <.001). To have readers interested or enjoying the data comic format substantially more provides a basis that data and often complex medical terminology can be enjoyed by many if developed with the user in mind.

The final engagement question asked the participants if they found it easy to **stay focused** on the medium. Similar to the outcome for the other questions, the data comic was received very positively (median = 6) compared to the medical text (median = 2), with a Mann–Whitney test indicating that this difference was statistically significant (U = 353.5, p <.001). This is likely due to the panelling within the data comic, creating small sections of information and the narrative sequencing engaging the reader through the storytelling ability of the medium.

6.2. Ease of understanding

Ease of understanding was evaluated as the participants' accuracy for the seven survey questions in Table 1, calculated for each participant as the mean accuracy score across all questions (ranging from 0 to 1). As participants either selected the correct or incorrect answer, the error for each question was binary. Fig. 6 shows the distribution of accuracy for each group.

Overall, the data comic was found to have considerably higher accuracy in understanding (median = 0.7143) than the comparable medical text containing the same information (median = 0.3571). The difference is statistically significant based on a Mann–Whitney test (U = 544.5, p <.001). A Mann–Whitney test was chosen, as upon a Shapiro–Wilk test we found the accuracy scores for the two groups not to be normally distributed.
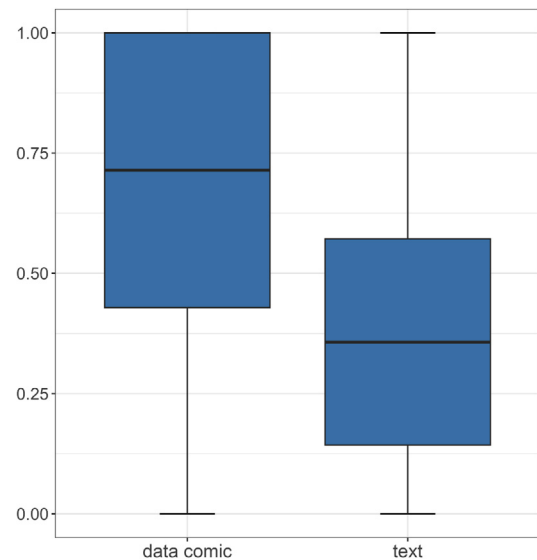

Fig. 6. The distribution of overall ease of understanding for each group.

We used the same visualisations in both mediums. For example, a pie chart illustrating cancer incidence in 25–29 year old females by site and a bar-in-bar chart including the most common female cancers with the respective number of deaths. As we controlled the number of variables, an accurate comparison between the mediums can be made.

The most accurately recalled fact occurred for Q4. This question was supported by a visualisation in both mediums: a pie chart including only two categories. This simplicity has aided in the memorability of the fact. 89.4% of data comic participants answered this correctly, which we consider to be substantial. The scene is interactive and further explained by McCrab, which appears to have made it easier for participants to understand the medical information.

The participants presented with the text medium scored most accurately on the questions containing visual aids, exemplifying that a general audience can understand information more easily with visual aids. The least accurate question answered by the participants within the text group was Q3, with only 16.7% answering it correctly (see Fig. 7). The answer is a multiple and only included within the text. In contrast, within the data comic, the addition of the fish communicating directly with the reader to "look at how big the difference is!" has undoubtedly impacted participants' understanding and recall of this fact, as Q3 has the largest difference (45.0%) between accuracy scores.

Interestingly, the question answered least accurately by the participants within the data comic group was Q6, with 55.3% answering it correctly. The answer was only communicated in a text format within a double speech bubble that included interactivity for the lower section of text explaining lung cancer symptoms. It is posited that the
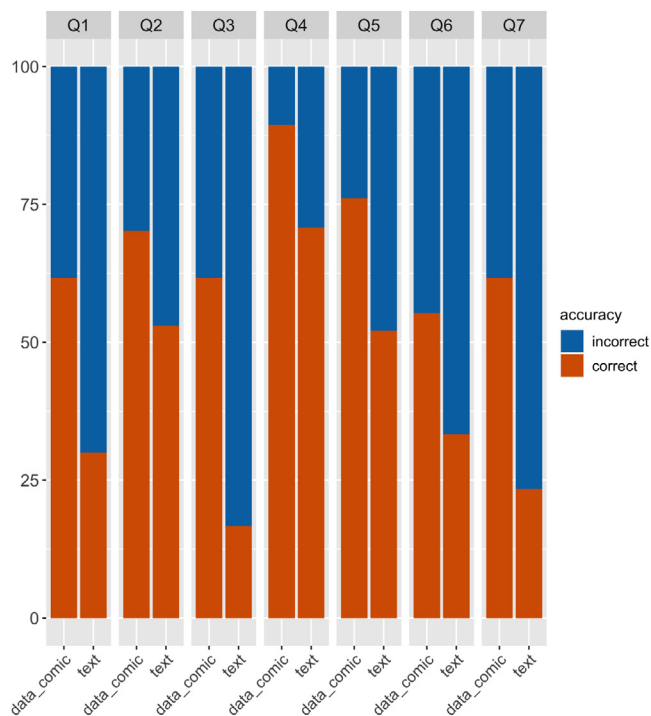
Fig. 7. The accuracy per question for each group.

greater engagement that the interactive element brought to this section circumvented participants from the statistic, reducing recall and understanding of the data. Therefore, the requirement for a visualisation to aid ease of understanding for readers of the medical information reaches into the interactivity aspect.

6.3. Self-assessed likelihood of screening attendance

Both mediums highlighted the risk of, and how to prevent cancer. The results confirm our hypothesis that the data comic would influence readers more positively than the comparable medical text. In particular, the perceived likelihood of screening attendance was higher for the data comic group (median = 4) compared to the text group (median = 3), and this difference was found to be statistically significant based on a Mann–Whitney test (U = 707, p=.001).

Fig. 8 shows that an overwhelming majority of participants (76.6%) stated that they would be more likely to attend a screening appointment following their interaction with the data comic. This indicates that the data comic had a considerable impact on their attitudes and willingness to change their health behaviour. On the other hand, most participants who were presented with the medical text medium (45.8%) stated that they felt unaffected by the data and information presented and would not be more likely or unlikely to attend a screening appointment after their interaction. A smaller proportion (41.7%) was still positively influenced by the messages within the text and felt more likely to attend a screening appointment after reading the article.

6.4. Qualitative feedback

Twenty text participants (39.2%) and 25 data comic participants (53.2%) provided qualitative feedback regarding the medium they were presented with. Sentiment analysis was carried out on this text using Qualtrics' Text iQ functionality. Generating the topics was an iterative procedure. A top-down process was initially followed to generate topics that we had expected to see. Upon the identification of trends, a

bottom-up method was adopted to encapsulate the appropriate comments.

As shown in Fig. 9, a much more positive sentiment was attached, overall, to the data comic compared to the text. In particular, 19 of the 25 data comic participants to leave feedback gave positive comments, 5 negative, and 1 neutral. The two main topics involve the data viz and understanding. This is a comment left with both topics included: *"Found the information very interesting and more easy to read on comic format as it didn't just look like boring text"*. New topics, such as 'novel', formed as participants commented on how it is an *"unusual way to show data but it works!"* and how the interactivity aided in their understanding of the data. The negative sentiment within the device topic was that participants who had used a tablet to load the data comic found it more challenging to follow due to the side-to-side scrolling required.

Regarding the text medium, only four of the 20 comments were positive, 13 were negative, and 3 were neutral. The positive feedback surrounded the data viz aspect of the text. Given that both mediums have the same number of data visualisations, a lack of visualisations was not the driving factor for the negativity. One participant stated that the formality and all-around conciseness of the text made it *"...quite obvious why people generally do not like to read this type of stuff in their free time even if it's supposed to be something that was done to raise awareness and educate people to have a better self-protection for their health.."*..

7. Discussion

7.1. Main findings

Higher engagement and understanding for the medical data comic - The data comic was rated as considerably more engaging, enjoyable and aesthetically pleasing than the comparable medical text, with a much higher proportion of participants finding it easy to stay focused on the medium and stating a desire to spend additional time exploring. The data comic also led to more correct answers regarding medical information. We were not surprised that the data comic was more positively received, especially in terms of enjoyability and visual appeal. However, some of the differences were larger than we were expecting. For example, perceived engagement for the text format scored surprisingly low, which is a concern for current practices regarding communicating medical information. With 91.5% of participants in the data comic group stating that they were willing to spend more time exploring further information, there is great potential for data comics to include additional information to people interested, for example through interactivity for details on demand. The fact that participants recalled the data and messages more accurately from reading the data comic than the text can potentially be explained by the fact that "panels help to divide information into easily memorable chunks" [34], as well as by the fact that data comic characters can comment on the data visualised and draw attention where needed.

Higher self-assessed likelihood to attend screening for the data comic - Even though we were not surprised to find that the data comic was rated more highly than the text, the proportion of data comic participants who stated that they were much more likely to attend a screening appointment (46.8%) surpassed our expectations by far. We view this as an important finding for the field of data visualisation in medicine, that health interventions could benefit from. A likely reason for these results is that data comics are a relatable and accessible aesthetic form [47], which allows users to create their own reality through the characterisation and narrative of the story, while at the same time providing new ways of understanding complex medical data. The interactivity in our data comic also allows users to explore the information and potentially personalise the narrative.

Much more positive sentiment attached to the data comic - In line with our findings regarding enjoyability, engagement and visual appeal, the subjective free text provided by participants had, overall, an encouraging and positive tone, which was not the case with the text medium. The data comic's interactivity was well received, and it was interesting to see participants comment on the novelty of the medium.
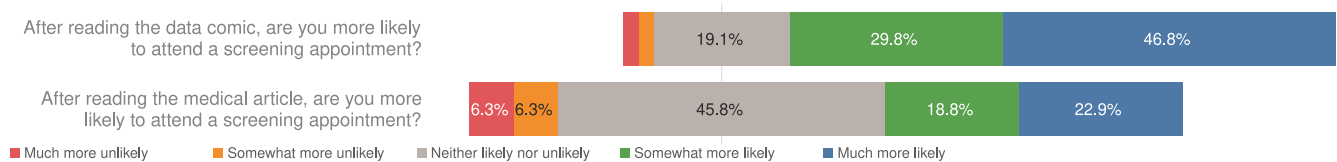
Fig. 8. Feedback received from participants regarding their likelihood of attending screening upon engaging with each medium.
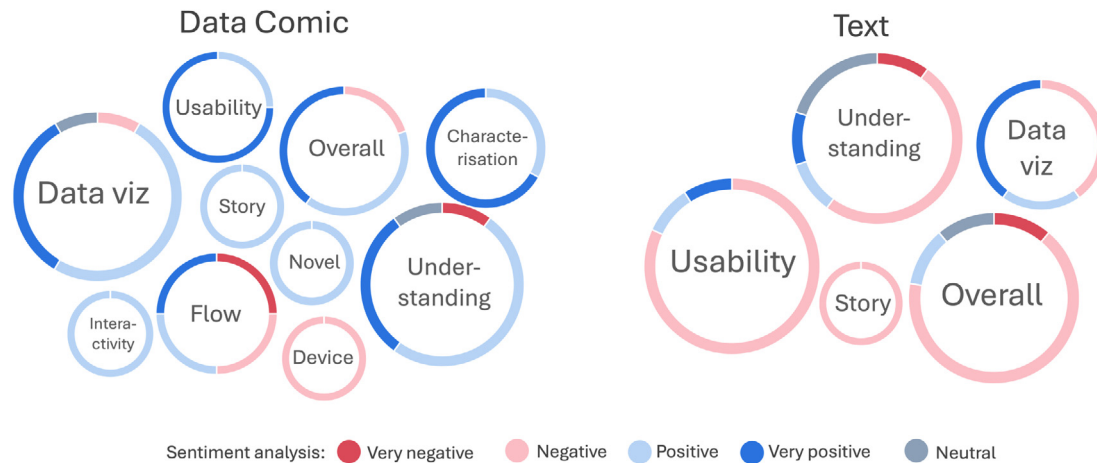

Fig. 9. Sentiment analysis results of the feedback gathered from participants in both groups. The sentiment is encoded through the colour on the outer ring. The size of the bubble indicates the number of comments on that topic.

7.2. Limitations and future work

Audience and sampling bias - The target audience for the interactive data comic is the general public. However, given the recruitment methods, there are no guarantees that participants are representative of a general audience. There is a risk of self-selection bias, for example with people that are interested in medical information or data visualisation. This is a limitation for a data comic that is motivated by the health literacy crisis, and future studies need to pay special attention to it.

Beyond a research setting and towards health behaviour change - Future studies need to evaluate how such a medical data comic can be used in a real-world setting, and what might affect users' engagement and understanding. Considering the great potential of data comics as part of public health interventions, it is important to understand where, when and how to make them available in order to maximise accessibility, engagement and actual health behaviour change. Building on the data comic presented in this paper, we see great value in exploring topics around promoting prevention (e.g. making healthy lifestyle choices) and supporting informed decision-making about screening.

In-depth qualitative feedback The survey-based data collected in this study is an important first step to understanding and ultimately improving the communication of medical data through interactive data comics. Future studies can extend our work by collecting in-depth qualitative feedback from participants (e.g. through interviews), allowing for a deeper understanding of the user experience. This could include an investigation of the reasons behind a change of screening likelihood and its relationship to users' affect, similarly to [33], especially given the fact that some of the wording or messages in the data comic could be perceived as threatening (e.g. "chance of death" and "40 new cases an hour").

8. Conclusion

Recognising the challenge of communicating medical information to a general audience, we propose interactive data comics as an accessible and engaging medium. We presented an interactive data comic about cancer and we discussed key design considerations. Findings from a survey-based study indicate that data comics ease the readers' understanding of complex medical information through an enjoyable and engaging format, while improving their perceived likelihood of attending screening. These are encouraging results for data visualisation in medicine, with valuable implications for public health campaigns and initiatives.

CRediT authorship contribution statement

Melissa Fogwill: Writing – original draft, Visualization, Validation, Investigation, Formal analysis. **Areti Manataki:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

A link to the study data is available within the Supplementary Materials.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cag.2024.104055.

M. Fogwill and A. Manataki Computers & Graphics 124 (2024) 104055

References

[1] World Health Organisation. Health literacy the solid facts. 2013, https://apps.who.int/iris/bitstream/handle/10665/128703/e96854.pdf.

[2] Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155(2):97–107.

[3] Bostock S, Steptoe A. Association between low functional health literacy and mortality in older adults: longitudinal cohort study. BMJ 2012;344.

[4] Bach B, Riche NH, Carpendale S, Pfister H. The emerging genre of data comics. IEEE Comput Graph Appl 2017;37(3):6–13. http://dx.doi.org/10.1109/MCG.2017.33.

[5] Caplan B. Open borders: the science and ethics of immigration. 2019.

[6] Moore AB, Nowostawski M, Frantz C, Hulbe C. Comic strip narratives in time geography. ISPRS Int J Geo-Inf 2018;7(7):245.

[7] Scottish Government. Scottish household survey 2019: Twenty years of Scotland's people. 2019, URL https://shs.theapsgroup.scot/2019/data-comic/.

[8] Sørensen K, den Broucke SV, Fullam J, Doyle G, Pelikan J, Slonska Z, Brand H. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12(1):1–13. http://dx.doi.org/10.1186/1471-2458-12-80.

[9] Nutbeam D, Lloyd JE. Understanding and responding to health literacy as a social determinant of health. Annu Rev Public Health 2021;42:159—173. http://dx.doi.org/10.1146/annurev-publhealth-090419-102529.

[10] Green MJ, Myers KR. Graphic medicine: use of comics in medical education and patient care. BMJ 2010;340:863. http://dx.doi.org/10.1136/bmj.c863.

[11] Williams ICM. Graphic medicine: comics as medical narrative. Med Hum 2012;38(1):21–7. http://dx.doi.org/10.1136/medhum-2011-010093.

[12] McNicol S. The potential of educational comics as a health information medium. Health Inf Libr J 2017;34(1):20–31.

[13] Brand A, Gao L, Hamann A, Crayen C, Brand H, Squier SM, Stangl K, Kendel F, Stangl V. Medical graphic narratives to improve patient comprehension and periprocedural anxiety before coronary angiography and percutaneous coronary intervention: A randomized trial. Ann Intern Med 2019;170(8):579–81. http://dx.doi.org/10.7326/M18-2976.

[14] Kearns C, Kearns N. The role of comics in public health communication during the COVID-19 pandemic. J Vis Commun Med 2020;43(3):139–49. http://dx.doi.org/10.1080/17453054.2020.1761248.

[15] Krakow M. Graphic narratives and cancer prevention: A case study of an American cancer society comic book. Health Commun 2017;32(5):525–8. http://dx.doi.org/10.1080/10410236.2016.1211075.

[16] Leung AY, Leung IS, Liu JY, Ting S, Lo S. Improving health literacy and medication compliance through comic books: a quasi-experimental study of Chinese community-dwelling older adults. Glob Health Promot 2018;25(4):67–78. http://dx.doi.org/10.1177/1757975918798364.

[17] Rhode M, Connor J. Graphic tales of cancer. Int J Comic Art 2012;14:112–56.

[18] Li O, Gray NA. Drawn together: Merging the worlds of health and comics through graphic medicine. In: Shapiro L, editor. Graphic medicine, humanizing healthcare and novel approaches in anatomical education. Cham: Springer Nature Switzerland; 2023, p. 3–21. http://dx.doi.org/10.1007/978-3-031-39035-7_1.

[19] Bateman S, Mandryk RL, Gutwin C, Genest A, McDine D, Brooks C. Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In: Proceedings of the CHI conference on human factors in computing systems. 2010, p. 2573–82. http://dx.doi.org/10.1145/1753326.1753716.

[20] Borkin MA, Vo AA, Bylinskii Z, Isola P, Sunkavalli S, Oliva A, Pfister H. What makes a visualization memorable? IEEE Trans Vis Comput Graphics 2013;19(12):2306–15. http://dx.doi.org/10.1109/TVCG.2013.234.

[21] Borkin MA, Bylinskii Z, Kim NW, Bainbridge CM, Yeh CS, Borkin D, Pfister H, Oliva A. Beyond memorability: Visualization recognition and recall. IEEE Trans Vis Comput Graphics 2016;22(1):519–28. http://dx.doi.org/10.1109/TVCG.2015.2467732.

[22] Pandey AV, Manivannan A, Nov O, Satterthwaite M, Bertini E. The persuasive power of data visualization. IEEE Trans Vis Comput Graphics 2014;20(12):2211–20. http://dx.doi.org/10.1109/TVCG.2014.2346419.

[23] Park S, Bekemeier B, Flaxman A, Schultz M. Impact of data visualization on decision-making and its implications for public health practice: a systematic literature review. Inform Health Soc Care 2022;47(2):175–93. http://dx.doi.org/10.1080/17538157.2021.1982949.

[24] Lee-Robbins E, Adar E. Affective learning objectives for communicative visualizations. IEEE Trans Vis Comput Graphics 2023;29(1):1–11. http://dx.doi.org/10.1109/TVCG.2022.3209500.

[25] Lan X, Wu Y, Cao N. Affective visualization design: Leveraging the emotional impact of data. IEEE Trans Vis Comput Graphics 2024;30(1):1–11. http://dx.doi.org/10.1109/TVCG.2023.3327385.

[26] Franconeri SL, Padilla LM, Shah P, Zacks JM, Hullman J. The science of visual data communication: What works. Psychol Sci Public Interest 2021;22(3):110–61. http://dx.doi.org/10.1177/15291006211051956, PMID: 34907835.

[27] Garrison LA, Meuschke M, Preim B, Bruckner S. Current approaches in narrative medical visualization. In: Roughley M, editor. Approaches for science illustration and communication. Cham: Springer Nature Switzerland; 2023, p. 95–116. http://dx.doi.org/10.1007/978-3-031-41652-1_4.

[28] Segel E, Heer J. Narrative visualization: Telling stories with data. IEEE Trans Vis Comput Graphics 2010;16(6):1139–48. http://dx.doi.org/10.1109/TVCG.2010.179.

[29] Kleinau A, Stupak E, Meuschke M, Mörth E, Garrison LA, Mittenentzwei S, Smit NN, Lawonn K, Bruckner S, Gutberlet M, Preim B. Is there a Tornado in Alex's Blood Flow? A Case Study for Narrative Medical Visualization. In: Raidou RG, Sommer B, Kuhlen TW, Krone M, Schultz T, Wu H-Y, editors. Eurographics workshop on visual computing for biology and medicine. The Eurographics Association; 2022, http://dx.doi.org/10.2312/vcbm.20221183.

[30] Mittenentzwei S, Garrison LA, Mörth E, Lawonn K, Bruckner S, Preim B, Meuschke M. Investigating user behavior in slideshows and scrollytelling as narrative genres in medical visualization. Comput Graph 2023;114:229–38. http://dx.doi.org/10.1016/j.cag.2023.06.011.

[31] Meuschke M, Garrison LA, Smit NN, Bach B, Mittenentzwei S, Weiß V, Bruckner S, Lawonn K, Preim B. Narrative medical visualization to communicate disease data. Comput Graph 2022;107:144–57. http://dx.doi.org/10.1016/j.cag.2022.07.017.

[32] Mittenentzwei S, Weiß V, Schreiber S, Garrison LA, Bruckner S, Pfister M, Preim B, Meuschke M. Do disease stories need a hero? Effects of human protagonists on a narrative visualization about cerebral small vessel disease. Comput Graph Forum 2023;42(3):123–35. http://dx.doi.org/10.1111/cgf.14817.

[33] Sallam S, Sakamoto Y, Leboe-McGowan J, Latulipe C, Irani P. Towards design guidelines for effective health-related data videos: An empirical investigation of affect, personality, and video content. In: Proceedings of the CHI conference on human factors in computing systems. 2022, http://dx.doi.org/10.1145/3491102.3517727.

[34] Wang Z, Wang S, Farinella M, Murray-Rust D, Riche NH, Bach B. Comparing effectiveness and engagement of data comics and infographics. In: Proceedings of the CHI conference on human factors in computing systems. 2019, p. 1–12. http://dx.doi.org/10.1145/3290605.3300483.

[35] Zhao Z, Marr R, Elmqvist N. Data comics: Sequential art for data-driven storytelling. tech. report, Univ. of Maryland; 2015.

[36] Bach B, Kerracher N, Hall KW, Carpendale S, Kennedy J, Henry Riche N. Telling stories about dynamic networks with graph comics. In: Proceedings of the CHI conference on human factors in computing systems. 2016, p. 3670–82. http://dx.doi.org/10.1145/2858036.2858387.

[37] Bach B, Wang Z, Farinella M, Murray-Rust D, Riche NH. Design patterns for data comics. In: Proceedings of the CHI conference on human factors in computing systems. 2018, p. 1–12. http://dx.doi.org/10.1145/3173574.3173612.

[38] Boucher M, Bach B, Stoiber C, Wang Z, Aigner W. Educational data comics: What can comics do for education in visualization? In: 2023 IEEE VIS workshop on visualization education, literacy, and activities (eduVis). 2023, p. 34–40. http://dx.doi.org/10.1109/EduVis60792.2023.00012.

[39] McAllister MP. Comic books and AIDS. J Pop Cult 1992;26(2):1–24. http://dx.doi.org/10.1111/j.0022-3840.1992.26021.x.

[40] Lowdermilk T. User-centered design: a developer's guide to building user-friendly applications. O'Reilly Media, Inc.; 2013.

[41] Roberts JC, Headleand C, Ritsos PD. Sketching designs using the five design-sheet methodology. IEEE Trans Vis Comput Graphics 2016;22(1):419–28. http://dx.doi.org/10.1109/TVCG.2015.2467271.

[42] Robertson T, Simonsen J. Participatory design: An introduction. In: Routledge international handbook of participatory design. Routledge; 2012, p. 1–18.

[43] Health Data Insight CIC. Simulacrum: Artificial patient-like cancer data to help researchers gain insights. 2022, https://simulacrum.healthdatainsight.org.uk/.

[44] Wang Z, Dingwall H, Bach B. Teaching data visualization and storytelling with data comic workshops. In: Extended abstracts of the CHI conference on human factors in computing systems. 2019, p. 1–9. http://dx.doi.org/10.1145/3290607.3299043.

[45] Bach B, Stefaner M, Boy J, Drucker S, Bartram L, Wood J, Ciuccarelli P, Engelhardt Y, Koeppen U, Tversky B. Narrative design patterns for data-driven storytelling. In: Data-driven storytelling. AK Peters/CRC Press; 2018, p. 107–33.

[46] Ma K-L, Liao I, Frazier J, Hauser H, Kostis H-N. Scientific storytelling using visualization. IEEE Comput Graph Appl 2011;32(1):12–9. http://dx.doi.org/10.1109/MCG.2012.24.

[47] Alamalhodaei A, Alberda AP, Feigenbaum A. Humanizing data through 'data comics': An introduction to graphic medicine and graphic social science. Data Vis Soc 2020;347.