DE GRUYTER i-com 2025; 24(1): 259–281

Research Article

Sarah Mittenentzwei*, Laura A. Garrison, Beatrice Budich, Kai Lawonn, Alexander Dockhorn, Bernhard Preim and Monique Meuschke

AI-based character generation for disease stories: a case study using epidemiological data to highlight preventable risk factors

https://doi.org/10.1515/icom-2024-0041
Received August 5, 2024; accepted January 31, 2025; published online February 18, 2025

Abstract: Data-driven storytelling has grown significantly, becoming prevalent in various fields, including healthcare. In medical narratives, characters are crucial for engaging audiences, making complex medical information accessible, and potentially influencing positive behavioral and lifestyle changes. However, designing characters that are both educational and relatable to effectively engage audiences is challenging. We propose a GenAI-assisted pipeline for character design in data-driven medical stories, utilizing Stable Diffusion, a deep learning text-to-image model, to transform data into visual character representations. This approach reduces the time and artistic skills required to create characters that reflect the underlying data. As a proof-of-concept, we generated and evaluated two characters in a crowd-sourced case study, assessing their authenticity to the underlying data and consistency over time. In a qualitative evaluation with four experts with knowledge in design and health communication, the characters were discussed regarding their quality and refinement opportunities. The characters effectively conveyed various aspects of the data, such as emotions, age, and body weight. However, generating multiple consistent images of the same character proved to be a significant challenge. This underscores a key issue in using generative AI for character creation: the limited control designers have over the output.

Keywords: data-storytelling; GenAI; medicine; healthcare; character generation

***Corresponding author: Sarah Mittenentzwei**, Department of Simulation and Graphics, Otto-von-Guericke University of Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany,
E-mail: sarah.mittenentzwei@ovgu.de. https://orcid.org/0000-0003-4124-3132
Laura A. Garrison, Department of Informatics, University of Bergen, Thormøhlens Gate 55, 5020 Bergen, Norway,
E-mail: Laura.Garrison@uib.no. https://orcid.org/0000-0001-7134-2006
Beatrice Budich, **Bernhard Preim and Monique Meuschke**, Department of Simulation and Graphics, Otto-von-Guericke University of Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany,
E-mail: beatrice.budich@ovgu.de (B. Budich),
bernhard@isg.cs.uni-magdeburg.de (B. Preim),
meuschke@isg.cs.uni-magdeburg.de (M. Meuschke).
https://orcid.org/0000-0001-9826-9478 (B. Preim).
https://orcid.org/0000-0001-9098-504X (M. Meuschke)
Kai Lawonn, Faculty of Mathematics and Computer Science, University of Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany,
E-mail: kai.lawonn@uni-jena.de
Alexander Dockhorn, Institute of Information Processing, Leibniz University Hannover, Welfengarten 1, 30167 Hannover, Germany, E-mail: dockhorn@tnt.uni-hannover.de. https://orcid.org/0000-0001-8711-7428

1 Introduction

Data-driven storytelling, particularly in healthcare, is gaining traction as a means to communicate complex medical information to patients and the general public. Characters, which we define as embodiment of specific conditions derived from patient data to make health information more relatable, can strengthen the audience's emotional engagement with both the narrative and the underlying data.[1,2] Creating convincing characters is still challenging and requires either creativity, research, and technical skills in the creation of illustrations or the consent of patients to have their photos taken. Deep learning text-to-image generative AI (GenAI) tools like Midjourney[3] and Stable Diffusion[4] have transformed this process by generating images from natural language prompts. In character design, *prompt engineering* involves creating prompts to develop traits and variations in the generated output images using specific queries and style specifications.[5] In our previous work,[6] we developed a pipeline for AI-enhanced character creation in data-driven medical narratives, addressing data protection challenges by not using real patient photos. We have extended this approach to generate virtual patients from epidemiological datasets such as the *Study of Health in Pomerania* (SHIP),[7] leveraging richer data on education, occupation, and

Open Access. © 2025 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

marital status to create more detailed characters. Therefore, some of the text passages closely follow our previous work.

We contribute a list of design criteria for characters in medical storytelling, developed by a domain expert with extensive industry experience. By integrating Stable Diffusion,[4] we enhance reproducibility in the design process, which is often impacted by frequent model updates on external platforms such as Midjourney.[3] Our case study evaluates the authenticity and consistency of characters, addressing the challenges and opportunities in AI-assisted character design. Our findings advance character design in healthcare and encourage further research in this area. Our contributions compared to the work by Budich et al.[6] are:

- Extending the GenAI character creation pipeline by using detailed epidemiological data, enriching narratives with attributes like education, occupation, and marital status.
- Enhancing character generation reproducibility and increasing data privacy by using a self-hosted Stable Diffusion model.
- Introducing a two-step evaluation of our proof-of-concept to validate AI-enhanced creation of educational visuals and address AI-assisted design challenges, providing insights to refine medical storytelling.

2 Related work

Our semi-automated character design pipeline for medical stories builds on principles from traditional storytelling, narrative medical visualization, and the creative potential of Generative AI (GenAI). This section reviews related approaches, highlighting gaps and positioning our contribution.

2.1 Narrative medical visualization

Narrative visualization blends storytelling with data visualization to craft data-driven stories.[8] In medicine, storytelling adapts to healthcare's unique challenges, employing diverse data while respecting patient privacy and sensitivities.[9] Medical data stories frequently appear in the news and health media. The methods for crafting these stories are an ongoing topic in narrative medical visualization research,[10,11] which explores how visualization can facilitate physician-patient interaction and motivate care.[12] McCurdy et al.[13] employed visual storytelling for patients to express health issues, using a timeline-based template to show symptom extent and duration. Meuschke et al.[9] established a template for crafting data-driven disease stories for patients and the general public based on analyzed web blogs. Mittenentzwei et al.[14] provide a design process for narrative medical visualizations where a character is one of the four main ingredients for a narrative. The book chapter by Garrison et al.[15] aligns closely with our work, highlighting the importance of narrative in health literacy. Their focus on narrative techniques, personalization, and data integration offers a theoretical and practical framework that complements our character-driven storytelling approach.

Kleinau et al.[16] applied a template in a case study to enhance engagement and knowledge retention through structure, personalization, characters, and metaphors. Building on this, Mittenentzwei et al.[17] studied the effects of narrative genres (slideshow and scrollytelling) in disease communication for usability and aesthetics. Later, they examined the impact of different story protagonists on engagement and memorability[1] and found that a human protagonist exerts various influences on story perception, especially improving engagement. While these stories are data-driven, the characters they feature are hand-crafted and designed based on discussions with clinicians. In our work, we propose a way to derive characters directly from epidemiological data.

2.2 Towards data-driven storytelling with characters

Characters enhance a story's narrative, particularly in the media and entertainment industry, with the gaming sector leading character-driven storytelling research and innovation. Cavazza et al.[18] focus on the dynamic interaction of autonomous characters. More recently, Mariani et al.[19] and Sheldon[20] investigated character-driven storytelling. Mariani et al. focus on developing characters for interactive storytelling, exploring how richly crafted characters with complex traits and backstories can drive narratives, engage audiences, and deepen immersion in story worlds. Katherine Isbister[21] illustrates the complexity of the design space for characters by exploring game character design from a psychological perspective. Notably, the concept of nonplayer characters (characters not controlled by the user) shares similarities with characters in data-driven storytelling. In game design, decisions primarily revolve around social roles and integrating characters into gameplay, emphasizing that characters with a greater presence should be developed with more depth. However, unlike typical nonplayer characters in games (e.g., traders, guides, or enemies), the characters in data-driven storytelling lack interactivity with the user. Both data-driven stories and games share the common goal of fostering empathy for their characters. Katherine Isbister[21] specifically discusses the concept of *Social Syncing*, highlighting how emotions can be "contagious." This underscores the significance of emotions in character design, as users can internalize and adopt the emotions conveyed through a character's expressions. Faces play an important role in conveying the emotions of a single character, while body language is especially useful for communicating connections with others. Chen et al.[22] investigates the user's identification with the protagonist, concluding that a high degree of identification can lead to self-referencing (meaning that users connect the story's content to their own lives).

In game development, narrative design (story and character roles) and visual realization by concept artists are typically handled by two distinct individuals.[23] We utilize GenAI to assist the concept artist's role, while the narrative design is driven by data. We have also included additional context, such as emotions, in our GenAI prompts, as they also play an important role in character design. In data-driven storytelling, characters play a more nascent role. Dasu et al.[24] introduced a framework to improve data storytelling by integrating characters, making visualizations engaging and accessible. They analyzed 160 data stories to outline character roles – main, supporting, and antagonistic – and their interactions within story plots, demonstrating how characters can effectively bridge complex data and audience understanding through visual metaphors and narrative structure. While this work demonstrates how characters can clarify and humanize complex data, they remain limited to manually defined character roles and lack automation or design aspects like emotional expression or visual coherence.

2.3 Generative AI art

The recent popularity of openly available (latent) diffusion models like Stable Diffusion,[4] along with user-friendly applications such as Midjourney[3] and Leonardo AI,[25] has brought GenAI art to the forefront. These models and tools enable image creation from text across various domains.

While the training of diffusion models may require large amounts of training data and compute resources, a technique called Low-Rank Adaption[26] allows for efficiently fine-tuning existing models for specific domains and applications, e.g., stylized or near photo-realistic images. Community-driven platforms like Civitai[27] allow requesting and sharing such models, while tools like the stable-diffusion-web UI by Automatic1111[28] allow users to easily use them. In addition to creating images from text, diffusion models can also be used for inpainting (replacing content in a masked image), sketch-to-image (conditioning the creation of images based on a simple sketch), and tasks such as up- or down-sampling. These advancements in GenAI art offer various opportunities to support design processes. So et al.[29] introduced a GenAI tool for generating narrative medical visuals from social media posts on conditions like diabetes, based on the 'bio-psycho-social model' encompassing medical, psychological, and social factors.[30]

Existing research in interactive storytelling and game design reveals a gap in leveraging GenAI to synthesize narrative roles, emotions, and data-driven contexts. Our work directly addresses this void by exploring how GenAI tools can assist in creating emotionally resonant, visually impactful characters tailored for medical storytelling – a domain where characters can bridge technical data and audience empathy but remain underexplored.

3 From traditional to AI-enhanced character design

Character design is a complex, time-intensive process requiring creativity, research, and technical skills. We describe the typical design pipeline for patient characters in medical stories based on one of the coauthors' industry experiences in such a production pipeline and their adaptation of modern GenAI tools. Traditional character design is divided into four broad phases: research, concept, refine, and launch.[6]

Research. Character design begins with the artist receiving a project brief that outlines the disease and desired learning outcomes for the audience. The artist researches the disease's signs and symptoms and identifies a typical patient profile, including age, gender, race, weight, and psycho-social factors like depression. Access to medical literature databases and online resources like UpToDate,[31] Healthline,[32] Cleveland Clinic's Health Library,[33] or JAMA's Patient Pages[34] provide epidemiological data or patient narratives to inform character design. In addition, the artist collects photos and stylistic references to ensure the depiction of patients' physical features, proportions, and behaviors. In our pipeline, we extract patient data from epidemiological studies, reducing the online research workload. However, the representativeness of selected patients for the condition should still be validated with external sources.

Concept. The artist creates thumbnail sketches of potential character designs, either analog or digital, using industry-standard software like Adobe Creative Cloud[35] or Procreate.[36] The artist uses references from the research phase in an iterative concept phase, refining the character design before collaborating with the team or client for feedback. The phase concludes when one character design is chosen for further refinement. We use GenAI to rapidly create initial character designs, generating multiple images in seconds to minutes by deriving prompts from previous research results.

Refine. The artist refines the character design iteratively, using references from the research phase and feedback from collaborators and experts. For clinical training, disease stories may enhance diagnostic skills by having users observe a virtual patient's respiration rate or other behaviors. Like earlier, GenAI speeds up image modifications.

Launch. In the final phase of character design, the completed character and related assets such as textures, models, and animations are integrated with other story elements like text and multimedia, validated, and launched to end users. This process is similar for both traditional and AI-enhanced designs.

4 Criteria for character design

When designing characters for medical data stories, considering key factors is crucial for effective communication of medical information. We outline ten criteria for character design in medical stories, again based on one of the coauthors' industry experiences and building on the methods described in Section 3.

C1. Relevance to Medical Context: The visual appearance of characters is essential for effectively conveying medical conditions, as it can represent symptoms, evoke emotional responses, and provide a human context to make data more relatable. Characters should clearly connect to the narrative's healthcare focus, with backgrounds and professions that are relevant to the story. Designers can use evolving visual cues to depict changes in health status and the condition's impact on well-being.[37]

C2. Informative and Educational: The character should serve as a vehicle for conveying complex medical information and data. The character's appearance should provide informative insights into medical concepts, conditions, or procedures.[37,38]

C3. Memorable: To capture the audience's attention and leave a lasting impression, characters must have unique traits, qualities, or quirks that let them stand out from other characters. For example, *Naming* characters significantly contribute to their identification, relatability, and memorability.[39]

C4. Empathy and Emotional Connection: The character should evoke empathy and foster an emotional connection, enabling deeper audience engagement. Integrating authentic visual elements that represent symptoms, effects, or challenges of the medical condition enhances understanding and connection.[39]

C5. Credibility: To make a character credible, its visual design should align with the narrative to immerse readers fully. Clothing choices and authentic expressions enhance the story's atmosphere and engagement while maintaining a consistent style among characters is crucial for credibility.[39]

C6. Personalization and Audience Relevance: The character's attributes and traits should resonate with the target audience, representing individuals from risk groups defined by factors like age or preventable risks. Prioritizing diversity in design – such as various ages, genders, ethnicities, and body types – ensures medical conditions are inclusively portrayed and not confined to a single demographic.[39]

C7. Behavioral Change and Motivation: The primary role of a patient character is to inspire positive behavioral changes, motivating the audience to adopt healthier habits and make informed medical decisions. As role models, characters can demonstrate overcoming challenges and the benefits of changes.[38]

C8. Data-Informed Design: Character design should be grounded in data analysis and corroborated by credible sources like literature and reports. In longitudinal studies, attributes such as demographics, background, disease progression, and risk factors can be directly extracted from the dataset.[40]

C9. Ethical Considerations: Character portrayal should adhere to ethical guidelines, respecting patient privacy, confidentiality, and cultural sensitivities. Designers should avoid identifiable features, real-life resemblances, and any stereotypes or stigmatizing portrayals of medical conditions or communities. Thus, they should uphold ethical standards and prioritize the well-being of those with medical conditions.[38]

C10. Iterative Design and Evaluation: The character's design should be subject to iterative refinement based on empirical evidence and audience feedback. User testing and evaluation should be conducted to assess the character's impact on audience comprehension, engagement, and behavioral change.[37]

5 Proof of concept for a semi-automated character pipeline

We present our semi-automated character design pipeline to support data-driven medical stories, illustrated in Figure 1. Our pipeline includes *requirement definition* for extracting data, followed by designing physical appearance through *text-to-image generation* and *image-to-image translation*, and concluding with *refinement* where results are evaluated, reworked, and finally launched.

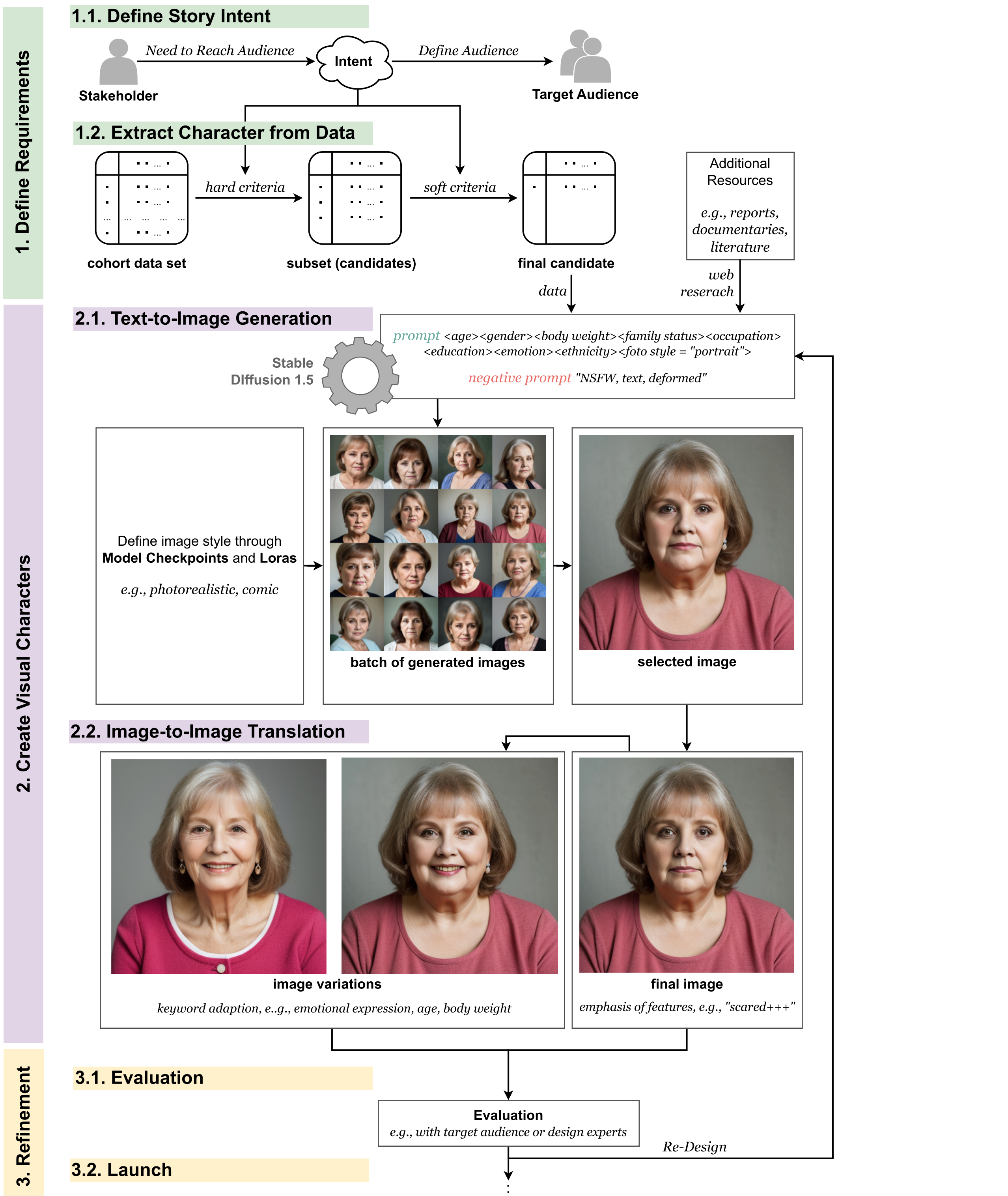

Figure 1: Semi-automated character design pipeline. Based on the needs of the stakeholder (typically a clinician), a story intent and target audience are defined. Hard and soft criteria are defined to filter data according to the intent. Based on the final candidate and optionally additional resources (e.g., internet sources on the disease and research on character design, such as suitable names), a prompt is formulated to generate a batch of character images using text-to-image generation. Specific styles (e.g., photorealistic or comic) can be achieved by using model checkpoints or LoRAs. The most appropriate image is selected from the batch and can be modified using image-to-image translation. The resulting character should be evaluated to ensure that it conveys the intent of the story, and the feedback can be used to modify the prompt to generate better characters. When a suitable character is generated, it can be finally launched.

To demonstrate our approach as proof of concept, we extracted characters from the *Study of Health in Pomerania* (SHIP), an epidemiological database in West Pomerania that identifies disease risk factors.[41] Data subsets from *SHIP-0* (1997–2001), *SHIP-1* (2002–2006), and *SHIP-2* (2008–2012) were used. We created two characters to illustrate the link between body weight and Non-alcoholic fatty liver (NAFL) to a lay audience. The first character, a woman, recovered from NAFL, while the second, a man, developed NAFL due to lifestyle factors like weight gain and insufficient exercise. The main objectives of communicating SHIP data on NAFL disease are (1) to inform the public about the risks and preventive measures for NAFL and (2) to promote the importance of the SHIP study to generate acceptance among potential participants. Therefore, the target audience is not NAFL patients but the general public. This is described in Figure 1 as Step 1.1. Define Story Intent.

Similar epidemiological studies are widely available, e.g., the UK Biobank[42] and the Rotterdam Study.[43] These are even more comprehensive in terms of the people and data collected, e.g., genetics. Risks for a disease are often influenced by both genetic predisposition and lifestyle-related parameters. Thus, while we use a concrete data set as an example, our pipeline is generalizable to other epidemiological data.

5.1 Extract character from data

Our character design process is built upon high-quality longitudinal data. We propose a query-filtering approach to extract a targeted subset of data for subsequent image generation using the SHIP dataset. This filtering applies two sets of criteria, aligned with design principles C1, C2, C7, and C8, see Figure 1 (Step 1.2. Extract Character from Data). **Hard criteria** select candidates embodying the core medical message and story intent, such as advocating for positive lifestyle changes. These include individuals with profiles aligning with disease risk factors, who have either overcome health issues or developed diseases due to lifestyle choices. **Soft criteria** consider candidates with notable lifestyle improvements or deteriorations, requiring background knowledge of disease risk factors for identification.

For the narrative context, we chose additional variables that can help describe the personal background and personality of the protagonist, e.g., gender, family, education, and occupation. This shows many parallels to the creation of personas, where user data is gathered and analyzed for demographic or behavioral patterns.[44] The resulting personas are often enhanced with, e.g., pictures or names to create a comprehensive profile. We added emotional context based on **additional resources** (e.g., documentaries about patients with non-alcoholic fatty liver disease[45,46]). Other additional resources may include information provided by medical experts or internet sources about the disease and research into character design, e.g., appropriate names that fit the age and background of the characters.

Algorithm 1 shows our hard and soft criteria for extracting the female character recovered from NAFL, while Algorithm 2 depicts the extraction of the male patient that develops NAFL. The **hard criteria** fall into three groups, focusing on variables strongly correlated with NAFL identified from SHIP studies:[7]

1. Diagnose-related variables: For the diagnosis of NAFL, only subjects with *fatty liver* combined with no *alcohol problems* and either recovered or developed the disease were considered.
2. Demographic variables, i.e., *gender*, *family*, *education*, and *occupation* to give context and history to the character.
3. Risk factors or lifestyle-related variables for NAFL, e.g., age, *BMI*, *alcohol use*, *physical activity*.[7]

Soft criteria identified those participants with positive lifestyle changes: *weight loss*, *exercise*, *smoking cessation*, and *reduced alcohol consumption*.

Algorithm 1. Extract character (case study example female).

Candidates_hardcriteria ← subjects S where
(stea_s0 = 1 **and** stea_s2 = 0) **and** (alcproblem_s0 = 0)
and (overweight = True) **and** (weight-loss = True) **and** (physact_s2 = 1)
and(smoking_s2 = 0 | 1)
where:
overweight: **if** (whtr_s0 > 0.5) **then** True
weight-loss: **if** (som_tail_s0 > som_tail_s2)
then True
Candidate_softcriteria ← Candidates_hardcriteria where s
maximizes the following conditions:
(whtr_s0 > 0.5 **and** whtr_s2 < 0.5) **or**
(som_bmi_s0 > 25 **and** som_bmi_s2 <25) **or**
(som_bmi_s0 > som_bmi_s2) **or** (physact_s0 = 0 **and** physact_s2 = 1) **or**
(alkligt_s0 >alkligt_s2) **or** (LFV_s0 > LFV_s2)

Legend: *s0*, *s1*, and *s2* refer to data subsets SHIP-0/-1/-2 recorded at five-year intervals. ***Variables:*** *stea* = steatosis hepatis (fatty liver), *physact* = physical activity, *smoking* = 0 (non-smoker); 1 (ex-smoker), *whtr* = waist-to-height ratio, *som_bmi* = body mass index, *alkligt* = alcohol in g per day, *LFV* = liver function values: Triglycerides, ALAT, GGT, ASAT (indicators for liver health, e.g., high dietary fat intake, liver inflammation or damage).

Algorithm 2. Extract character (case study example male).

Candidates_hardcriteria ← *subjects S where*
(stea_s0 = 0 **and** stea_s2 = 1) **and** (alcproblem_s0 = 0) **and**
(overweight = False) **and** (weight-loss = False) **and** (physact_s2 = 0)
where:
overweight: **if** (whtr_s0 > 0.5) **then** True
weight-loss: **if** (som_tail_s0 > som_tail_s2)
then True
Candidate_softcriteria ← *Candidates_hardcriteria where s maximizes the following conditions*:
(whtr_s0 > 0.5 **and** whtr_s2 < 0.5) **or**
(som_bmi_s0 > 25 **and** som_bmi_s2 <25) **or**
(som_bmi_s0 > som_bmi_s2) **or**(physact_s0 = 0 **and** physact_s2 = 1) **or**
(alkligt_s0 >alkligt_s2) **or** (LFV_s0 > LFV_s2)

5.2 Text-to-image generation

We employ a text-to-image generation method to create characters based on hard and soft criteria, see Figure 1 (Step 2.1. Text-to-Image Generation). Budich et al.[6] utilized Leonardo AI,[25] a web-based GenAI tool with pre-trained models for various image styles. However, users have limited control over the models, as regular updates may alter the functionality of older prompts. Given this significant drawback, both for the long-term application of the tool in supporting character design for data-driven medical stories and for research studies, we adapted our image generation pipeline to locally run Stable Diffusion version 1.5. This decision was based on the following criteria:

- **Reproducability:** Stable Diffusion is widely recognized as a benchmark for evaluating generative AI techniques, including diffusion models, and is frequently cited in academic papers and studies as a comparative baseline. Its fixed releases ensure that researchers can utilize the exact same version of the GenAI model that we employed.
- **Open source:** Stable Diffusion's open-source nature allows users to thoroughly examine its model architecture and algorithms, providing insights into its inner workings, an essential aspect of ensuring reproducibility. The absence of licensing fees for non-commercial use makes it a cost-effective choice for institutions, researchers, and developers, especially in academic or non-commercial settings.
- **Data privacy:** Stable Diffusion can be executed locally, ensuring that both the input and output data remain exclusively accessible to the creator. This is particularly critical for applications in sensitive areas such as medical topics, where data privacy is paramount.
- **Multimodal capabilities:** Stable Diffusion offers versatile functionality essential for character creation, including text-to-image generation, image-to-image translation, and tools for refining outputs, such as specifying areas of an image to be altered or preserved. This adaptability makes it suitable for a wide range of GenAI tasks. Its all-in-one functionality streamlines the pipeline, eliminating the need for multiple platforms and enhancing accessibility for character creation workflows.
- **Available resources:** As an open-source platform, Stable Diffusion provides access to a diverse range of models and a flexible toolchain developed by its active community. This supports fine-tuning the model and storing optimization results, including network weights and learning rates, e.g., as model checkpoints tailored to various image styles and user-friendly GUIs. Its high degree of customizability makes Stable Diffusion a preferred choice for creators seeking adaptability and control.
- **Relevance:** Stable Diffusion is a leading diffusion-based generative model that creates high-quality, photorealistic, and stylized images from text prompts. Competing with proprietary models such as OpenAI's DALL-E and MidJourney, it leverages advanced latent diffusion models (LDMs) to deliver high-quality outputs with optimized computational efficiency and resource utilization.

1. Model Configuration. We utilized the Stable Diffusion web UI,[28] offering a graphical, browser-based interface to specify parameters like model selection, number of output images, and prompts. Customized model checkpoints can be developed by retraining neural network weights using a small set of images with consistent styles. Pretrained model checkpoints (files containing the weights of a Stable Diffusion model specialized in a particular style, subject, or aesthetic) are available on the Civitai website,[27] with our selection being "HT Photorealism v4.1.7",[47] tailored for portrait photography. Table 1 outlines our parameter setup, predominantly using default settings from the Stable Diffusion web UI v1.6.1.

2. Positive and Negative Prompts. Positive prompts indicate the desired image outcome, while negative prompts specify undesired features. Image generation models may

Table 1: Parameters of our text-to-image generation pipeline.

Parameter name	Parameter value
Sampling method	DPM++ 2M Karras
Sampling steps	20
CFG scale	7
Image width	512
Image height	512

lack regulation of their generated output, potentially generating not-safe-for-work (NSFW) content, including nudity. To encourage safe results, we included "NSFW" as a negative prompt for all images. Additionally, "text" and "deformed" were added as negative prompts, as models often struggle with text and complex structures like hands.

3. Prompt Adaptation due to Model Specifics. Using Stable Diffusion instead of Leonardo AI requires adjusting previous prompts used by Budich et al.[6] for optimal outcomes. Style-specific keywords (e.g., semi-realistic or photorealistic) were removed from the prompts as they are now defined by the model checkpoint. Additionally, character names were removed from the prompts, as they no longer significantly affect their appearance.

4. Effects of Different Keywords. The keyword "kindergarten teacher" introduced occupation-specific backgrounds like classrooms and blackboards in some images, often altering clothing to blouses or plain long-sleeved shirts, occasionally with floral patterns. However, the occupation's influence on output varies, and it may perpetuate stereotypes and clichés for certain jobs. Keywords like "educated," "mother," and "married" had minimal impact on output, mainly altering characters' clothing and jewelry in some images. For instance, "educated" occasionally produced more elegant looks with prominent jewelry and blazers. Despite their minor influence, these keywords remain in the prompt scheme as they are derived directly from data. Additionally, even if keywords have minimal effect, the model can still generate believable characters consistent with patient data.

Certain keywords, like "platinum blonde" or "wearing glasses" have a more significant impact than others, overriding descriptors such as "mother," "kindergarten teacher," and "educated" from the original prompt. However, terms like "open-minded" have negligible influence on results. The influence of keywords on generated images varies based on the model's training data. Reducing keywords in a prompt enhances control over the output by decreasing the weight of each keyword. Consequently, we eliminated keywords not derived from patient data and precise appearance descriptions, as they override data-derived keywords.

5. Keyword Weights. Prompt weights can be modified to adjust specific character features. The default weight is one, which can be increased by adding + or a number between 1.1 and 2 or decreased by − or a number between 0 and 0.9. + stands for a weight of 1.1, ++ adds a weight of 1.1^2, and so on. The same scheme applies for reducing the weights where – equals 0.9 and −− equals 0.9^2. For example, *"A beautiful, friendly, 65 years old, (overweight)+, educated woman"* assigns a weight of 1.1 to "overweight".

6. Base Prompt Input. We modified the prompt scheme from Budich et al.[6] to suit model specifics, resulting in the following base prompt input: `<age> <gender> <body weight> <family> <occupation> <education> <emotion> <ethnicity> <picture section>`.

All keywords, except for *emotion* and *picture section*, were derived from data. The base prompt scheme yielded satisfactory results, though additional keywords could further modify the images, such as altering hair color or style. We generated images for each character at time point *s0*, selecting emotions based on their health status. The prompts used were:

`65 years old women, overweight, mother, kindergarten teacher, educated, scared, European, portrait`

`32 years old man, normal weight, married, salesman, strong and focused, European, moderate drinking, portrait`

7. Refinement. In the initial phase, a range of output images is generated, providing the designer with options. Illustrated in Figure 1, even with the same prompt, Stable Diffusion produces diverse characters. From here, designers can adjust prompt component weights to experiment further.

5.3 Image-to-image translation

The character is finalized using Image-to-Image translation,[48] refining the design by generating similar images. This process preserves key features like pose, hairstyle, and clothing, allowing the selection of the best character variant. Refer to Figure 1 (Step 2.2. Image-to-Image Translation) for the finalized female base character. Our pipeline generates character variants from this final image, covering diverse attributes such as emotions, body weights, and ages crucial for depicting long-term medical narratives.

Image-to-image translation alters an image based on a text prompt. The *inpaint* approach allows users to specify regions for alteration by drawing on the image, as illustrated in Figure 2. To modify emotions, we painted within the facial region. Adjusting weight and age involved including the outer boundaries of the head, considering potential changes in head size. Using inpaint helped to preserve features affected by unstable diffusion. We maintained consistency by employing the same negative prompt as in text-to-image generation and adjusted the positive prompt by modifying emotion, age, and weight keywords.

We used image-to-image translation to show the character's evolution from time step *s0* to *s2* in the SHIP study. Over ten years, the female character lost weight

Figure 2: The original image (a) and an inpaint area (b) are combined with a text prompt for image-to-image translation. Different emotions, like scared (c) and happy (d), are generated within the painted area.

and recovered from fatty liver, while the male character gained weight and developed liver inflammation. Their facial expressions were adjusted, with healthy characters showing positive emotions and diseased characters showing negative ones. See Figure 1 for the female character's changing expression, body weight, and age.

5.4 Generalizability

There are multiple epidemiological studies that offer their data for research purposes, such as the UK Biobank,[42] Rotterdam Study,[43] and German National Cohort (NAKO).[49] Our pipeline can be applied to any epidemiological study, but hard and soft criteria need to be adjusted based on the story intent. Cohort data, collected through studies such as the ones mentioned above, usually consists of sociodemographic data (employment, education, ethnicity, etc.), lifestyle information (preventive and risk factors, e.g., smoking, diet, activity, etc.), and health status and physical measures/biological samples (body weight, blood pressure, blood samples, etc.). These are the most important aspects of developing a narrative medical visualization focusing on one disease and featuring an individual health journey.

In our proof of concept, we created characters that have either a poor health status that improved due to lifestyle improvements, or a good health status that worsened due to poor lifestyle. Based on diseases tracked in the epidemiological study, concrete hard and soft criteria can be defined to find fitting patients. Sociodemographic data helps shape the AI-generated character's appearance, such as age in the prompt. Lifestyle information can be used to visualize lifestyle aspects relevant to a certain condition, e.g., a character can be depicted as smoking. Lifestyle information visualizes relevant aspects, like depicting a character smoking. Health data, physical measurements, and biological samples are used to represent the disease, its causes, symptoms, or consequences, such as body weight.

It is important to emphasize that since individual patients are used as models for the GenAI characters, consultation with a clinician – e.g., someone who works with the cohort data – is essential. This will ensure that the patients selected are *representative* and that the final character matches the physician's approval, especially with regard to risk factors for certain diseases.

6 Evaluation of the proof of concept

We evaluated both characters concerning the three AI-based steps in our pipeline: (1) Text-to-Image Generation, (2) Image-to-Image Translation, and (3) Refinement. Therefore, we addressed the following questions:

Q1 – Authenticity: Is the character authentic to the underlying data?

Q2 – Resemblance: Do the altered images resemble the same person?

Q3 – Refinement: How to refine the characters to meet design standards?

With **Q1**, we assess if the generated images represent the input information. Given that these characters convey health details to the public, **authenticity** is crucial. As the underlying data spans a lengthy epidemiological study, the character must reflect changes over time, primarily in age and body weight. **Q2** examines our ability to modify the original images using GenAI while ensuring users recognize both versions as the same individual. **Q3** investigates how the generated characters can be refined to produce results that align more closely with standards upheld by designers and medical illustrators.

Q1 and **Q2** are evaluated with a quantitative evaluation targeted at an audience without medical or design expertise in Section 6. **Q3** is evaluated with a qualitative evaluation with design experts familiar with medical application cases in Section 7.

6.1 Participants

To answer **Q1** and **Q2**, we conducted a quantitative evaluation using an online questionnaire over six weeks. 29 participants completed the questionnaire, while 14 participants dropped out before completion (their responses were not included). The questionnaire and the responses are available as Supplementary Material. While we covered a range of different backgrounds, most participants are under 30, have a university degree, and live in Germany, see Table 2. Gender distribution is nearly equal between males and females.

6.2 Results of Q1 – authenticity

Select keywords for an image. We explore the authenticity of the images in representing the underlying data. Participants were tasked with assigning keywords to the images, covering six categories corresponding to each keyword used in the image generation prompt:

Age: young adult (18–29), middle-aged (30–59), mature (60+).

Body Weight: overweight, normal weight, underweight.

Family: married, widowed, separated, divorced, single, parent.

Occupation: Military, Agriculture, Manufacturing, Construction, Natural Sciences, Traffic/Logistics, Sales/Trading, Business Organization, Health Care, Social Sciences (short forms based on the occupational sectors defined in the German Classification of Occupations 2010 – Revised Version 2020[50]).

Education: No qualifications, Secondary school-leaving certificate, High school diploma, Completed apprenticeship, University degree.

Emotion: angry, afraid, sad, calm, strong/focused, happy (Based on the "Feelings Wheel" developed by Willcox[51] for aiding individuals in identifying and articulating emotions. It is a structured framework that categorizes and subcategorizes different emotions. To keep the identification of emotions simple, we chose to only use the six top categories.)

Age, body weight, and education were single-choice questions because each participant could only select one correct answer per category. In contrast, family, occupation, and emotion were multiple-choice questions, allowing participants to select all plausible options, as multiple options could apply to a person simultaneously.

In Figure 3, participants' keyword selections reveal that most ($n = 21$) perceived both characters as *Middle Aged (30–59)*, aligning with our prompt for the male character's age. Although the female character is intended to be 65 years old, most participants selected an age range only slightly above this. Overall, participants tended to view the male character as younger and the female character as older, highlighting a noticeable age gap between them. Regarding body weight, participants categorized the male character as normal weight ($n = 28$) and the female character as overweight ($n = 23$), consistent with our original prompt. While participants mainly interpreted the male character's emotion as *calm* ($n = 27$), which was not in our original prompt, they consistently chose positive or neutral emotions, reflecting our intention for the healthy character to be depicted positively. *Calm* was also the most frequently selected emotion for the female character ($n = 26$). Few participants opted for *afraid* ($n = 3$), the emotion we used in our prompt, or another negative emotion like *sad* ($n = 3$). We often had to increase the keyword weights for emotions in our prompts to elicit more pronounced facial expressions. However, our findings suggest that the model struggled to produce strong negative emotions, as these were rarely identified by participants.

Regarding the characters' occupations, the most common responses for the male character were *Business organization* ($n = 24$) and *Sales/Trading* ($n = 24$), the latter aligning with our original prompt input. For the female character, the predominant response was *Occupations in health care, the social sector, teaching, and education* ($n = 27$), which corresponds to our original prompt featuring the keyword *kindergarten teacher*. While participants may have based

Table 2: Metadata of study participants. Only selected options are displayed.

Age		Gender		Education		Residency	
18–24	7	Male	14	School pupil	1	Germany	26
25–29	14	Female	12	Junior high diploma	1	Norway	2
30–34	4	Other	2	High school diploma	7	Canada	1
50–54	1	No answer	1	Vocational training	2		
60–64	2			University degree	17		
>65	1			Other	1		

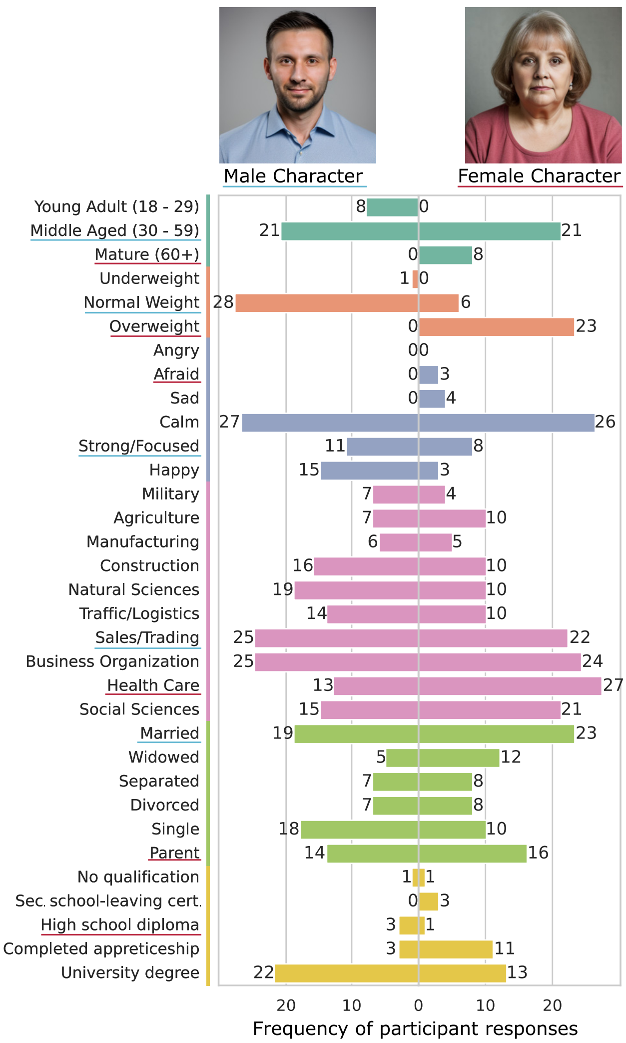

Figure 3: The bar chart displays participant responses for male and female characters. Colors represent different categories: age (single choice), body weight (single choice), emotion (multiple choice), occupation (multiple choice), family (multiple choice), and highest level of education (single choice). Keywords from the original prompt used to generate the images are underlined for the male (blue) and female (red) characters.

their choices on visual cues like clothing, these occupations are also heavily gender-typical, suggesting that gender alone could strongly influence the results.

The most frequently selected family term was *Married* for both the male ($n = 19$) and female characters ($n = 23$). Many participants also thought of the male character as being single ($n = 18$). While the family was not specified for the female character, we used the keyword *married* in our original prompt for the male character. About half of the participants also thought that the male ($n = 14$) and female characters ($n = 16$) were parents, which is the case for the latter according to the underlying SHIP data. The male ($n = 22$) as well as the female character ($n = 13$) were assumed to have a high educational level, i.e., a University degree, by many participants. While the education of the male character was not specified in the data, the female character has a *High school diploma*, which was only selected by one participant. When asked whether the characters in the static images felt like real or computer-generated people, participants frequently chose both options without a noticeable tendency.

Select images for keyword combinations. We showed the participants a subset of keywords used to generate the image variations for *s0* and *s2* for the male (A and B) as well as female (C and D) characters. The participants had to select all the images matching the keywords. The keyword combinations are:

A) 32 years old, Strong/Focused, Normal Weight
B) 65 years old, Scared, Overweight
C) 42 years old, Sad, Overweight
D) 75 years old, Happy, Normal Weight

The results are shown in Figure 4. Nearly all users chose character images with the correct age and weight. For A, the image generated with the keyword “angry” was selected most frequently ($n = 24$), with the image generated with “strong/focused” as the second most selected ($n = 18$). The angry expression, characterized by a furrowed brow, may also be interpreted as concentration or focus, indicating the ambiguity of this facial expression.

For B, C, and D, participants most commonly selected the correct images. However, for character C, the top image received only 17 votes (the top images for A, B, and D received at least 24 votes) indicating more difficulty in recognizing this emotion. Although correct images generally dominated the votes, participants often interpreted other images as depicting the same emotions. Furthermore, the model’s subtle alterations in the images made distinguishing emotions challenging, as seen in the comparison of strong and calm emotions in Figure 4(A).

Select images for character descriptions. This approach assessed whether participants correctly identified emotions, but we also explored whether the selected emotions resonated with the audience or if they would choose images depicting other emotions. We provided the participants with these short stories:

a) “Thomas Schmidt is a 32-year-old salesman. He is married and has a secondary school level. Although he has a poor diet, he has a normal body weight and moderate alcohol consumption but no alcohol problem.”

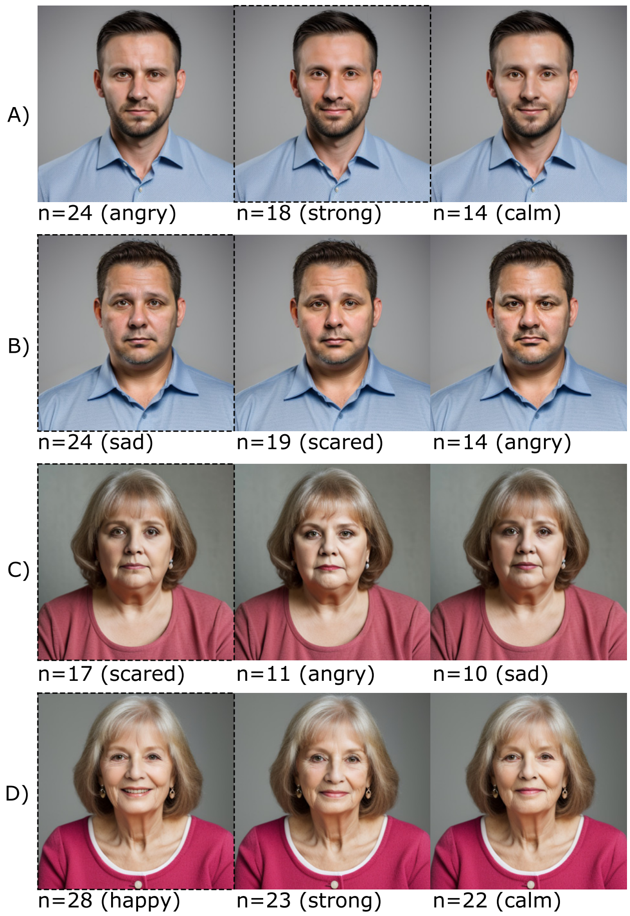

Figure 4: Top three images selected for each keyword combination. Below each image is the number of responses for that image and the emotion keyword we used in the prompt. The task was multiple-choice. Each participant could select all the images they found that matched the keywords. Images matching our prompt are highlighted with a dashed frame.

b) "Thomas Schmidt is a 43-year-old salesman. He is married and has a secondary school level. Due to his poor diet, he became overweight and developed critical stomach fat and liver inflammation."

c) "Emma Winter is a 65-year-old woman. She is married and a mother. After graduating from high school, she became a kindergarten teacher. She is overweight and has been diagnosed with a fatty liver. However, she has no alcohol problem and is an ex-smoker."

d) "Emma Winter is a 75-year-old woman. She is married and a mother. After graduating from high school, she became a kindergarten teacher. She started exercising and lost weight during the last years and successfully recovered from her fatty liver."

For each story, they should select one image that best fits the character described in the text.

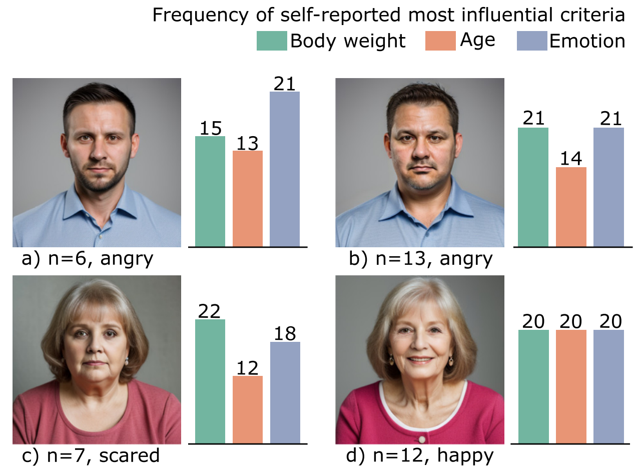

Figure 5: Most voted images for every story including the frequency of votes it received and the emotional expression shown. On the right side of every image, the frequency of self-reported criteria that were most influential for the participants' choice is depicted.

The most chosen image for each story is shown in Figure 5. Participants generally selected images with accurate age and weight. For the male character, the top images for both the (a) healthy and (b) diseased versions depicted anger. While anger in the diseased character could reflect frustration about his condition, the reason for its prevalence in the healthy character is less clear. It might be that participants interpreted the expression not as anger but as focus, similar to observations in a previous task. For the female character, participants predominantly chose scared for the diseased version (c) and happy for the healthy version (d), aligning with the images selected for our stories later in the evaluation. This suggests participants often use negative emotions to depict disease. The task required a single image selection per character. Notably, options (a) and (c) received fewer votes, 6 and 7, respectively, compared to (b) and (d), which garnered 13 and 12 votes, respectively.

We asked the participants which criteria they considered for their answers. Body weight, age, and emotion were frequently selected. We offered participants to enter their own criteria. Five participants used this option and stated that the following criteria played a role in their decision:

- Alcohol consumption, Dietary, Occupation
- Signs of health from poor diet such as eye bags, wrinkles, unhealthy skin color
- Bad teeth

The image generation model sometimes incorrectly rendered teeth, suggesting that the normal-weight male character had poor dental health when smiling. Although unintended, this coincidentally aligned with the character's backstory of having a poor diet that might affect his teeth. Participant responses indicate they also look for subtle health indicators like eye bags.

6.3 Results of Q2 – resemblance

We assessed whether participants viewed images from image-to-image translation as depicting the same person as the original. They rated the resemblance of each altered image on a 5-point Likert scale; results are shown in Figure 6.

Participants rated images where only the emotion was altered, but age and body weight were constant, as most similar. However, in cases where age and body weight changed, images of the female character transitioning from overweight to normal weight were seen as more similar than those of the male character going from normal to overweight. Emotional changes did not significantly affect the similarity ratings. In the image-to-image translation process, the model frequently altered the noses of characters, especially giving overweight characters more upturned noses. This suggests that the model often applies learned features associated with certain body weights instead of preserving the original features, leading participants to see them as different individuals.

7 Refinement to resonate with target audience

To address **Q3: How to refine the characters to meet design standards?**, we created new character images based on the quantitative user feedback and discussed these with four different design experts regarding the character design criteria from Section 4.

7.1 Refinement

The main feedback of the participants regarding which cues are helpful to identify a disease was (1) showing more than a portrait section to enable the audience to assess body weight better and (2) showing more subtle signs of disease such as eye bags, unhealthy skin color, or bad teeth.

While the former was easy to implement by generating torso shots instead of portraits, the latter could not be achieved as the generative model was not able to realize these subtle changes in a sufficient way. For example, even when using *inpaint* to define that only the area below the eyes should be changed to make eye bags more prominent, it generally led to deformed eyes. Therefore, instead of depicting the male character's poor diet using bad teeth, we chose to place him in a fast-food restaurant. This framing also allows us to discuss the choice of backgrounds with the design experts in the next step by comparing the female character with a neutral background and the male character with a non-neutral background. Based on the torso shots, we generated comic variants of the characters. Therefore, we used the model checkpoint revAnimated V1.2.2.[52] As there is a huge body of medical comics used for health education,[53] our intention is to investigate whether photorealism or comic style is preferred by design experts in certain aspects. We created six different pairs of images (three for each character, including the ones used in the previous user study). Each pair shows a character at time point *s0* and *s2* in the SHIP study, see Figure 7.

We used the following prompts for the generation of photorealistic torso shots:

```
65 years old women, overweight, mother, kindergarten teacher, educated, scared, European, waist and torso shot, standing, solid color background
```

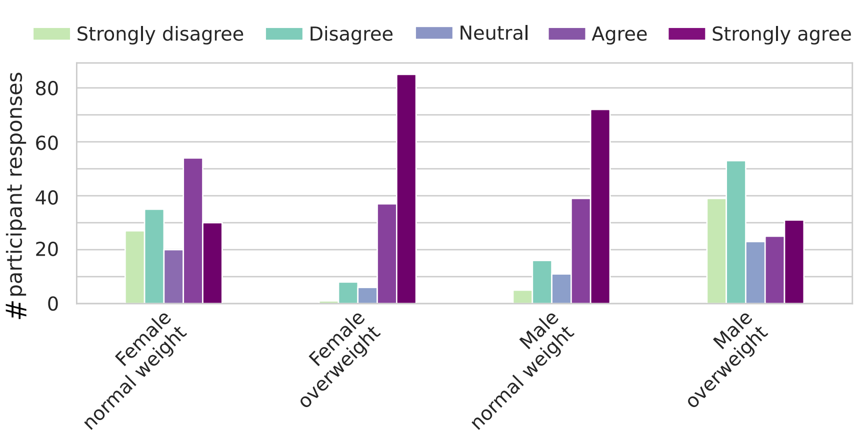

Figure 6: Frequency of participants' responses on the similarity of two images, aggregated by gender and weight to the statement "both images above show the same person, probably at different ages, weights, and moods."

Figure 7: The refined images.

`75 years old women, normal weight, mother, kindergarten teacher, educated, happy, European, waist and torso shot, standing, solid color background, sports`

`32 years old man, normal weight, married, salesman, focused, European, waist and torso shot, standing in a fast-food restaurant`

`42 years old man, overweight, married, salesman, sad, European, waist and torso shot, standing in a fast-food restaurant`

Using image-to-image translation, we generated the comic adaptions by adding the keyword "comic character" and including the LoRA (Low-Rank Adaptation) inkSketch V1.5.[54] LoRA is a fine-tuning technique that modifies a pre-trained model without requiring full retraining. The LoRA was activated using the command "lora:inkSketch_V2:1" in the prompt window.

7.2 Qualitative evaluation

We gathered feedback from four design experts *(D1 – D4)*, one of them being a co-author of this paper.

D1 holds bachelor's degrees in design and computer science. They are familiar with medical visualizations and narrative medical visualization, and have practical experience in designing them.

D2 has a bachelor's degree in pharmacology and a master's degree in medical illustration. They have practical expertise in illustration, especially portraits, and working experience in a public health agency.

D3 has a master's degree in interaction design and holds a design professorship. They are familiar with illustration, character design, and storytelling and have practical experience in designing them.

D4 has a graduate degree in medical illustration and has experience as a healthcare communications designer.

We used Miro[55] as a virtual whiteboard to show background information about SHIP, NAFL disease, patient backstories, and data. We prepared a tier list on the Miro board for each of the ten design criteria from Section 4. The design experts were individually asked to rank all the image pairs (one pair being the character at *s0* and *s2* in the same style) from one (very good) to six (very bad) based on how well the image pairs fit each individual criterion. This was achieved by the design experts dragging the generated character images into the tier list, see Figure 8. Furthermore, they were asked to express their thoughts and reasons for their decisions, which were noted down according to the think-aloud method.

7.3 Results of Q3 – refinement

While all design experts have a background in illustration or design and narrative medical visualization, they showed different preferences when discussing the generated images. Their rankings can be seen in Figure 9. It became apparent, that despite their professional background, personal experiences and preferences played an important role in their assessment. In the following, a detailed overview of the design experts' feedback regarding each criterion is presented. A concluding overview can be found in Section 9.1.

C1. Relevance to Medical Context. *D1* favored the torso shots as they show the stomach fat, which is an important indicator for NAFL. *D2* highlighted positively that the change in posture in the torso photo of the woman indicates a change in lifestyle. The portraits of the woman also show that she glows more in the second image, but because the rest of the body is not shown, the stomach fat is not visible. The comic version of the male character was ranked low by

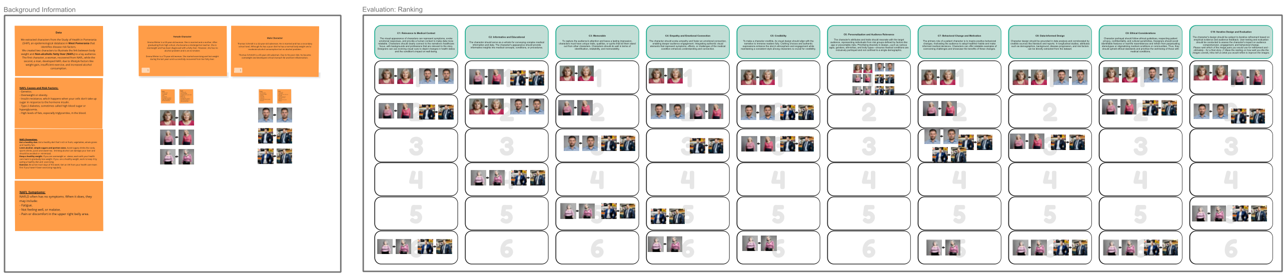

Figure 8: This screenshot illustrates the setup of one of the Miro boards used during the qualitative evaluation with design experts (specifically D4). On the left, orange boxes provide information about the data and the disease. On the right, a ranking list is presented for each criterion (outlined in blue boxes, similar to Section 4). Here, design experts were tasked with ranking the character images for each criterion by arranging them based on their assessments. Images placed higher in the column were deemed to better fulfill the respective criterion. The detailed results can be seen in Figure 9.

D1 because the style felt inappropriate. *D2* said that using comics would give the creator the opportunity to exaggerate more on, e.g., symptoms. However, especially the female comic character does not look like the same person in the two images. *D3* thinks that the comic version of the male character looks too strong and confident like he is in control of his life, which does not fit the story. *D1* appreciated the male torso shots' background for reflecting the character's poor diet, while *D3* found it too bold and stereotypical, preferring the less intrusive grey background.

C2. Informative and Educational. *D1* argued that having portraits that focus on the face leads the focus of the story toward the emotions of the characters. However, for NAFL the body shows important indicators, in particular stomach fat. *D2* and *D4* agree, adding that it is also possible to guess the body weight based on the portraits but it is much harder. While *D1* sees a lot of potential in the comic style, but especially the images of the female character do not fit the story as she looks too young in the second image. Additionally, *D2* noticed that the facial expression for the comic version of the male character stays the same, and that shifts the focus of the viewer to the changes in the body. However, emotion can show the social implications of the disease, so it depends on the intent of the story if it should be shown or not. *D3* thinks that the emotions of the comic versions and the photorealistic body shot do not match the story's intent. *D1* liked that the background of the male torso shots shows context relevant to the story. *D2* liked that the female character glows and shows positive emotions after recovering from NAFL. However, the deterioration is more stark in their personal opinion.

C3. Memorable. *D1* argued that memorability, in general, depends on the ability of the individual viewer to memorize faces, which differs per person. For the photorealistic images, the portraits might be more memorable as they show the face in more detail. *D3* thinks that the expression of the photorealistic torso shot of the female character is especially memorable. *D4* thinks that the portrait of the female character conveys a warm and motherly personality while the portrait of the male character shows no traits that stand out in particular. Furthermore, the male torso shots look too cliche. *D1* said that the comic version of the female character has a lot of expression, making it more memorable. *D3* agrees, but thinks that the increase of memorability might come at a loss of conveying the relevant content. The male comic character is described as being too sleek which reduces memorability. *D2* thought that comics are, in general, more memorable. The change of the posture and clothing of the female character is done nicely in both comic and photorealistic torso shots. The stark shading and general art style of the male comic character make him a distinguishable character and not just a random person. In general, *D2* said that changing clothing and backgrounds for the individual time steps would be better. The portrait photo of the female character has a very good contrast between the clothes and the background.

C4. Empathy and Emotional Connection. *D4* highlighted that a general issue might be that all kinds of characters might act as projection areas, where viewers, despite recognizing that they are close to their own demographics, will think, “This happens to other persons; this will not happen to me.” According to *D1* and *D2*, the portraits are better at conveying emotions as the face is visible in more detail. *D2* added that while the male character's portrait evokes empathy, it does not really look like the same person and the character ages too much. The torso shots of the male character do not evoke much emotion, according to *D2*. The facial expressions are neutral and the overweight level does not look unusual. *D3* mentioned that the photorealistic torso shot of the female character feels more distant to them than the portraits and the photorealistic torso shot of the male character. *D1* said that the photorealistic torso shot of the female character shows a large emotional difference between the time steps. In the second time step, she

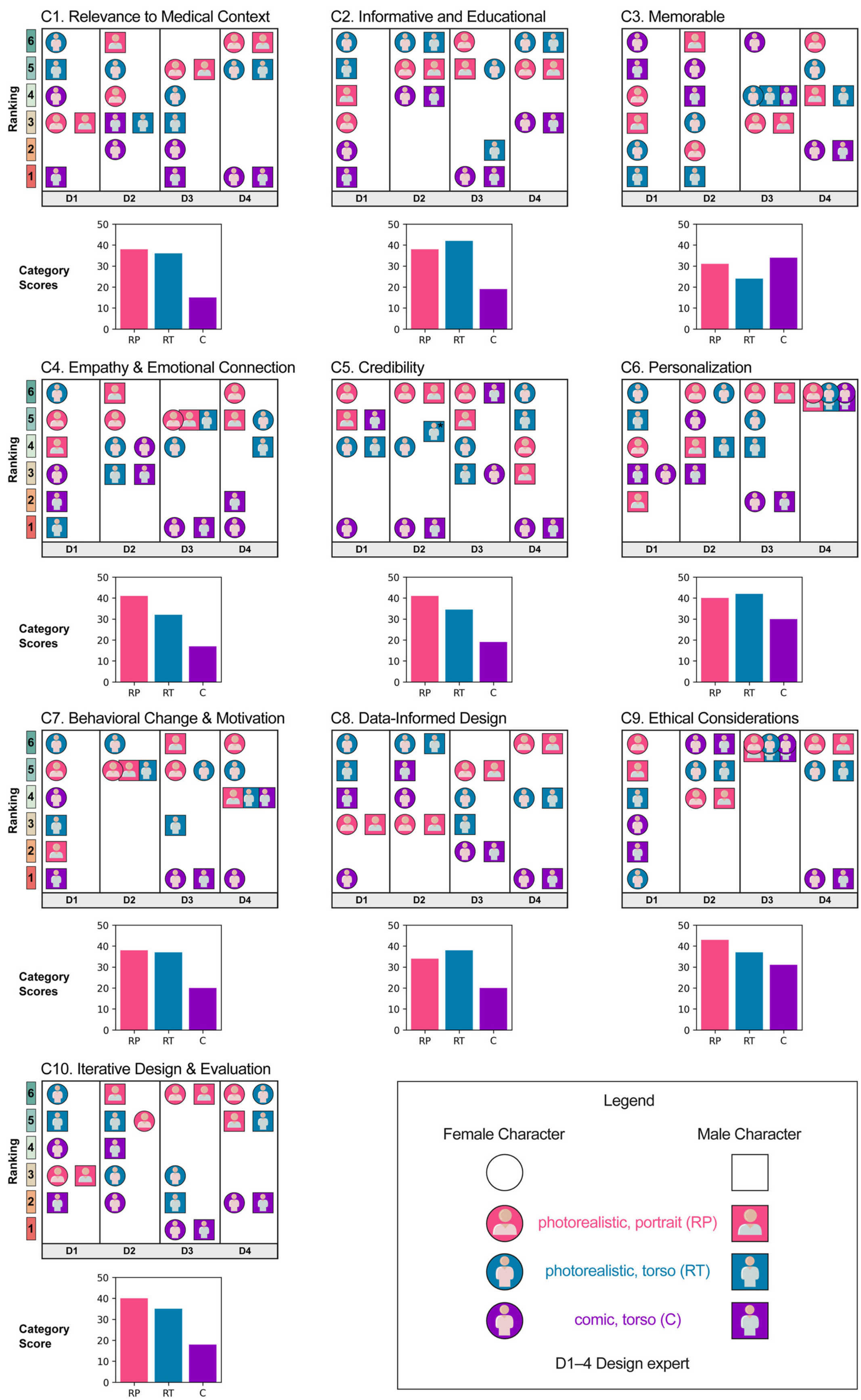

Figure 9: The design experts (D1 – D4) ranked the generated images regarding each individual design criteria (C1 – C10).

appears proud of overcoming NAFL. On the other hand, the comic version of the male character looks unemotional. His aggressive glance establishes a certain distance from the viewer, and the story is not conveyed that well. However, the background is praised again. *D2* said that the style of the comic version of the female character is too different to feel like this is the same person which prevents the viewer from feeling empathy. Furthermore, the comic style looks more abstract, like taken from a textbook, which builds distance between the characters. On the other hand, the photorealistic images feel less distant and give the impression of being real people. *D4* favored the female portrait as it looks motherly and warm, but showing the character in a fitting environment, e.g., at home, would increase the emotional connection. While the male torso shots successfully convey the character's lifestyle choice by placing him in a well-known setting, such as a fast food restaurant, the character itself does not evoke empathy.

C5. Credibility. *D1* appreciated that in the photorealistic torso shots of the female character, similarities but also changes are visualized in a credible manner. The character in the second image (S2) looks a bit taller than in the first (S0), which should be improved. *D1* said that the portraits of the male character do look like the same person, however, the healthy version (S0) should look happy instead of neutral. The healthy version of the photorealistic torso shots of the male character looks good, but the diseased version does not seem to show the same person. In general, it is important to show the visual and emotional changes of the characters. Changing the clothing style in different images of the same character would add to this. *D2* said that in the case of the male character, the portraits, and the comic torso shots are identifiable as the same person. *D4* agreed in the case of the comic versions. There, the nose and jawline look similar. The background conveys the eating habits of this person. In the portraits of the female characters, most of the facial features are similar, however, the eyes appear to be too different. *D3* and *D4* said that the comic versions of the female character and the photorealistic torso shot of the male character do not look like the same person at all as their facial features are very different. *D3* only recognized a consistency in the portrait photos. While the poses in the torso shots were praised, the comics did not give *D3* a sense of credibility. *D4* said that the portrait of the female character looks most consistent.

C6. Personalization and Audience Relevance. *D1* and *D2* thought that the female character matches the main risk group very well, as women tend to be affected more often than men. Additionally, the demographics in western Pomerania, where SHIP takes place, show that more than every third person is 65 years or older. Furthermore, *D2* and *D3* said that the male character is too much of a businessman which might not resonate well with a large part of the audience. *D3* says that the male character looks too helpless in the portrait photo, which does not invite the viewer to resonate with him. *D2* assumed that a face might be more relatable than, e.g., body or pose. The portraits leave more room for imagination, potentially addressing a wider variety of body types. *D1* and *D2* assumed that comics might not resonate well with the older target group, but this would need further investigation. In the case of the torso shots for the male character, the context in the background increases personalization, according to *D1*. *D3* suggested that another strategy could be to not visualize detailed individuals but instead use archetypes. This way, a larger part of the audience might be able to resonate with one archetype as it is more general. However, it is difficult to create archetypes that are not too stereotypical. Additionally, it would be great for the audience to show both progressions for the same character. Starting with a healthy character, the story can show how an unhealthy lifestyle leads to NAFL and how to recover from the disease by making lifestyle changes. *D4* does not see the need to apply this criterion to the characters as they are based on real persons and thus are relevant real-world examples of the disease.

C7. Behavioral Change and Motivation. *D1* was unsure if a positive or negative example is more motivating for the audience but has a personal preference for the positive example, comparing it to the psychological concept of positive reinforcement. *D2* said that both examples can motivate, and it might also depend on the disease and what works best. For example, there might be a difference in the case of diseases where risk factors are hard to quit due to an addictive factor such as smoking, where an example of a character that quickly progresses towards a healthier lifestyle might not be appropriate. *D4* assumed that what works best for an individual viewer might be the character

Similar to a tier list, the design experts were asked to assign the character images on a scale ranging from 6 (very good) to 1 (very bad), while it was possible to skip ranks and assign the same rank to multiple images. The original ranking from the Miro boards is shown along with an aggregated score for each image style (RP, RT, C) per category (C1 – C10). The scores are calculated by adding up the ranking results, where each image could receive 1–6 points from each design expert (D1 – D4), and allow for a comparison of how the image styles performed per category.

that is closer to the viewers on demographics and lifestyle. *D3* argued that the portraits have a different level of meaning, focusing on the emotions instead of on the physique. *D1* argues that the comic version of the male character looks cool, not diseased. Thus, the gravity of the disease is not conveyed. *D2* agrees, adding that the character does not look like he wants to change which will not be motivating for the audience. *D3* sees similar issues in the photorealistic torso shot of the male character.

C8. Data-Informed Design. *D1* said that the torso shots are crucial to show the critical stomach fat of the male character as this was one important aspect in the data. For both characters, the change of age and weight is visible in all the torso shots, except the comic version of the female character where the age does not seem to match the data. Concerning the portraits, only age is conveyed prominently, neglecting body fat. *D2* agrees, highlighting that the poor diet of the male character is also visible in the torso shots due to the background. For *D3* the age was not correctly conveyed in the comic versions. The female character looks too young in the second image, while the male character looks too young in the first image. *D4* thought as well that the photographs are generally better at conveying the data.

C9. Ethical Considerations. *D1* argues that the SHIP participants' privacy is guaranteed in all cases. As the portraits do look like real people, *D2* assumed that it is hard to get away from any resemblance with real people. *D2* added that in the comic versions, an accidental resemblance of a real person is unlikely. The comics are furthermore portraying the characters in a sensitive way as the proportions are still good looking, e.g., going too much in the direction of a caricature might ridicule certain conditions. Portraits are more likely to avoid fatphobia, as it is mostly concentrated on the body. But since weight is an important aspect of the disease, *D2* thought it reasonable to show obese bodies in the story. However, designers should think about what visual cues are crucial to show and focus on these. *D3* argued that it is not possible to avoid stereotypes completely, as there will always be some aspects that match certain stereotypes. *D4* said that they think the comic versions do not get the necessary sensitivity of the topic across. They questioned if the comic characters were respectful to the people from which the data was taken. *D1* assumed that it might be possible that images showing an expression that is too bitter might convey the impression of the character being isolated or stigmatized potentially triggering viewers with similar experiences.

C10. Iterative Design and Evaluation. For this criterion, design experts were asked to consider using one of the characters in an NAFL story. They were tasked with selecting images for the iterative design process and suggesting improvements for the first step. *D1* preferred the photorealistic torso shot of the female character but suggested reworking the facial expression on the left side, such as reducing the shadow contrast. While portraits alone would not fully represent NAFL disease, a mix of portraits for emotions and torso shots for physique or lifestyle would be ideal. *D2* preferred the male character's portrait, finding it most relatable. A stressful job with long hours leading to poor diet and NAFL was seen as realistic. However, D2 suggested adding more intermediate images to show the characters' gradual journey and depict how they could recover by adopting a healthier lifestyle. *D3* thought the portraits are looking the most natural and credible. It is also a good start for the story to first create empathy so the audience cares for the character and only in the second step show them in the context of the disease. Similar to *D2*, *D3* also liked the idea of showing how a person develops and recovers from NAFL. *D4* liked the story of the female character better. However, they disliked the concrete style of the comic. For a medical topic, they would opt for a softer style, e.g., inspired by watercolor.

8 Reflections and limitations of AI-enhanced character design

We selected a specific GenAI tool for image generation. Although our study aimed to establish a generalizable AI-enhanced character pipeline rather than evaluating specific tools, the choice of tool influences the results. Comparing various tools is challenging because outcomes depend on the underlying model (e.g., Stable Diffusion or Midjourney) and the specific checkpoints used, with many available on the Civitai website.[27] Due to this and the rapid development of GenAI models, which make prior versions irrelevant, we do not believe a comparison of different models is a valuable contribution. Investigating broader aspects like styles (photorealistic, artistic, etc.) and general strategies to create characters provides more meaningful insights for future character design.

During our evaluation with the design experts, it became clear that there are many different illustrative styles for visualizing characters. As we only feature a small sample of styles in our proof-of-concept, it is not representative of the whole body of illustrative styles. Keeping an illustrative style consistent adds another challenge additionally to the requirement to show images of similar persons with different body weights and ages. To guarantee a consistent style, more specialized models are necessary.

Overall, the issue of consistency and resemblance, as well as the inability to add certain visual cues (e.g., eye bags), highlights the limited control creators have over the generative process. As generative models can only reproduce what they learned from a set of training data, the way how the model was trained significantly influences the results a creator can achieve. However, we are only able to discuss these issues regarding the results of our proof-of-concept approach for which we utilized a specific software, namely Stable Diffusion 1.5. It is important to note that other existing (commercial) software, such as SnapChat[56] or Adobe[35] might produce better results regarding altering facial images. Furthermore, regardless of the training data, the models still struggle with complex shapes such as hands. We also tried Stable Diffusion 3 (as it is currently the latest release), but could not see any difference in the quality of the resulting characters. Since Stable Diffusion 3 was promoted with the promise of improving text rendering in particular, this was to be expected. In addition, Stable Diffusion 3 can provide improved resolution, but this is only noticeable with larger output images or small details, neither of which are relevant problems in our application example. Also, the set of available checkpoints and Loras is significantly smaller for Stable Diffusion 3, which is why we decided to stay with Stable Diffusion 1.5.

The majority of participants in our study are under 30 years old and hold a university degree, resulting in limited representativity of the broader and more diverse lay audience targeted by the characters. Additionally, the chosen application example of NAFL and the associated characters may resonate more with older individuals, as NAFL predominantly affects older adults. These demographic and contextual factors should be considered when interpreting our results, as they may introduce potential biases.

It is essential for character design to avoid stigmatization and stereotypes, ensuring sensitivity, accuracy, and authenticity to promote inclusivity. For instance, while body weight is a critical risk factor in our NAFL example, equating obesity with disease and normal weight with health risks reinforces weight-related stigma.[57] Creators must carefully balance authenticity with sensitivity when designing characters.

Disease risk factors in the dataset are binary-coded, despite existing on a spectrum in reality. For instance, smoking is represented as true or false (see Algorithm 1), whereas the frequency or intensity of smoking could significantly vary. This level of simplification was dictated by the underlying data and was not a deliberate design choice for character generation. Furthermore, certain nuanced details may exceed the capabilities of GenAI. For example, modifying small features such as teeth, which occupy only a few pixels, proved unfeasible. Simplification, therefore, plays a critical role in GenAI-based character design.

Additionally, the dataset is limited to physical and biochemical health factors. Character design could be enhanced by integrating behavioral or socio-economic data, such as stress levels or social support. When such data is not available in the dataset, supplementary research or consultation with treating physicians can provide insights for adding characteristics that align with both the database and the narrative.

9 Conclusion and future research opportunities

Including GenAI in the character design process for data-driven stories may enhance character design and storytelling, uniting creativity and design possibilities. Our research explores semi-automated character design through tailored prompts for data-driven stories to improve health communication competence. Blending human design sensitivities with data analytics techniques, our pipeline enables the iterative production of characters that may contribute to telling more personalized, engaging, and understandable data-driven stories about health and wellness.

9.1 Conclusions

In the following we will provide a concluding overview of our results in the context of our research questions.

Q1 – Authenticity: Is the character authentic to the underlying data? The generated characters were found to authentically reflect the underlying epidemiological data in terms of key attributes like age and body weight. Interpreting emotional states, however, proved to be more challenging. Our evaluation results show, that users searched for additional clues for diseases, such as unhealthy skin color, eye bags, or bad teeth. Including these visible symptoms in the character generation process would create a stronger impression. Employing eye-tracking technology in such assessments can reveal which parts of the image capture users' attention the most.

Q2 – Resemblance: Do the altered images resemble the same person? The resemblance of altered images to the original character was a common critique, though it varied depending on the factor being modified. Changes in emotion did not negatively impact resemblance, whereas modifications to body weight and age often resulted in new character images that lacked sufficient similarity to the original. Notably, with the settings used, image-to-image translation tended to alter not just skin features (e.g., adding wrinkles to signify aging) but also other attributes such as eyes, noses, and occasionally even hair color. While image-to-image translation demonstrated potential for maintaining character consistency, it requires further refinement to ensure coherence during transitions. Specialized tools for ensuring character consistency, e.g., face swap models,[58] could be used in case the character's face is to be preserved. Changing features like the bodyweight may also impact facial features to some degree.

Q3 – Refinement: How to refine the characters to meet design standards? Discussions with design experts highlighted that character design has many variables and preferences, such as artistic styles, facial expressions, and body poses, that differ per person. The evaluation highlighted key strengths and limitations across the styles and suggested additional criteria for designing characters that effectively resonate with target audiences. Torso shots were praised for their ability to visually depict critical medical indicators, such as stomach fat, a significant marker for NAFL (higher scores, e.g., in *C2: Informative and Educational* and *C8: Data-Informed Design*). On the other hand, portraits were preferred for the depiction of showing emotions (higher scores, e.g., in *C4: Empathy and Emotional Connection*). Contextual backgrounds, such as a fast-food restaurant, further enhanced the narrative by illustrating lifestyle factors. However, the background can also be distracting. Comic styles often struggle with maintaining consistency in conveying age and emotional states. This affected audience recognition. Concerns were raised regarding sensitivity and suitability for the target group. On the other hand, comic styles provide a broad range of possibilities to depict different aspects of a disease (e.g., emotions can be emphasized more clearly than with photorealistic representations) and comic characters might increase memorability (higher scores in *C3: Memorable*) but the current images have several design flaws as described before. Additionally, the evaluation highlighted new considerations for character design:

Comprehensibility: To effectively convey the character's health journey (or the portion thereof), it is important to go beyond presenting just two images at the beginning and at the end of the journey. Including transitional images that illustrate gradual changes in between can improve narrative clarity and engagement. Combining photorealistic portraits with torso shots provides a more comprehensive narrative. Portraits foster emotional connection, while torso shots highlight physical health markers, creating a balanced and engaging depiction.

Appeal: The appeal of character design should be considered alongside *C6: Personalization and Audience Relevance*, which focuses on sociodemographic factors. Elements like comic style may vary in appeal depending on the target audience, influenced by factors such as age and culture. Different comic styles might be perceived as more or less appropriate and appealing. Experts recommend exploring archetypes that strike a balance between relatability and generalization while being cautious to avoid stigmatized stereotypes.

Context: Integrating relevant settings, such as a home environment for a female character or a fast-food restaurant for a male character, can enhance narrative immersion and effectively communicate lifestyle factors. However, the chosen background should avoid being overly complex, distracting, or stereotypical to maintain clarity and relevance.

9.2 Future research opportunities

Medical character creation provides individual challenges due to sensitive data in medical contexts. Medical topics evoke emotions, necessitating appropriate, reassuring characters. A character's reflection of the psychological and physical changes caused by disease is vital for aligning with the narrative, as demonstrated in our case. Based on our proof-of-concept, where we utilized GenAI to create data-driven characters for the application case of NAFL, we propose the following future research directions:

1. **Generalizability:** Validation against diverse datasets and character sets can verify the generalizability of our approach. Expand the study to a wider range of disease data, identifying classes of diseases and personas for characters.
2. **Broaden the application area:** Our approach is tailored to epidemiological study data, where the creator can choose from a broad set of patient data entries to find an appropriate base for a character for a data-driven disease story. However, these stories are also relevant beyond epidemiology. Exploring how the pipeline can be applied to other types of data could expand the scope of our approach. In situations where patient data is unavailable, representative personas could be developed in collaboration with physicians.
3. **Explore archetype-based designs:** To support creators with limited experience in character design, providing generalizable archetypes that avoid stigmatized stereotypes while resonating with broader audiences, would be very helpful. These archetypes must be tested across multiple narratives to determine their versatility and impact on audience engagement.

4. **Evaluate style variations:** Identifying appropriate styles for different target groups, such as photorealistic and different artistic styles, can guide the character design process in the future. If a set of popular styles can be identified, model checkpoints can be trained and published that are specialized for these styles.
5. **Providing context:** The influence of context, such as backgrounds in character images, needs to be investigated more thoroughly to identify appropriate settings and designs.
6. **Expanding to video and voice generation:** Lastly, our approach can be expanded to video generation, including the use of an artificial voice.
7. **Tool building and pipeline automation:** In its current form, our proof-of-concept requires user input at several stages of the content generation pipeline. Reducing the manual effort in crafting narrative visualizations will be a key focus of future work. For example, using large language models for generating story arcs and matching characters at several stages of their disease.
8. **Personalized storytelling and digital twins:** Given the current approach one story is crafted for informing a set of patients. With the increasing interest in personalized medicine, crafting stories that match a single patient's data may be of interest. Creating a prototype story that is automatically completed using the patient's data may be a first step in providing a more personalized experience.

Acknowledgments: The authors thank Till Ittermann of the University of Greifswald for fruitful discussions of data and design decisions.
Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: None declared.
Conflict of interest: The authors state no conflict of interest.
Research funding: This work is supported in part by the University of Bergen and the Trond Mohn Foundation in Bergen \#813558, Visualizing Data Science for Large Scale Hypothesis Management in Imaging Biomarker Discovery (VIDI) Project.
Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Mittenentzwei, S.; Weiß, V.; Schreiber, S.; Garrison, L.; 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), 123–135.
2. Neeley, L.; Barker, E.; Bayer, S. R.; Maktoufi, R.; Wu, K. J.; Zaringhalam, M. Linking Scholarship and Practice: Narrative and Identity in Science. *Front. Commun.* **2020**, *5*, 35.
3. Midjourney. Midjourney, 2024. https://www.midjourney.com (accessed 2024-05-3).
4. Stability, A. I. Stable Diffusion, 2024. https://stability.ai/stable-image (accessed 2024-05-3).
5. Liu, P.; Yuan, W.; Fu, J.; Jiang, Z.; Hayashi, H.; Neubig, G. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Comput. Surv.* **2023**, *55* (9), 1–35.
6. Budich, B.; Garrison, L. A.; Preim, B.; Meuschke, M. Reflections on AI-Assisted Character Design for Data-Driven Medical Stories. In *Eurographics Workshop on Visual Computing for Biology and Medicine*, 2023; pp. 1–5.
7. Ittermann, T.; Haring, R.; Wallaschofski, H.; Baumeister, S. E.; Nauck, M.; Dörr, M.; Lerch, M. M.; Meyer zu Schwabedissen, H. E.; Rosskopf, D.; Völzke, H. Inverse Association Between Serum Free Thyroxine Levels and Hepatic Steatosis: Results from the Study of Health in Pomerania. *Thyroid* **2012**, *22* (6), 568–574.
8. Lee, B.; Riche, N. H.; Isenberg, P.; Carpendale, S. More Than Telling a Story: Transforming Data into Visually Shared Stories. *IEEE Comput. Graph. Appl.* **2015**, *35* (5), 84–90.
9. Meuschke, M.; Garrison, L. A.; Smit, N. N.; Bach, B.; Mittenentzwei, S.; Weiß, V.; Bruckner, S.; Lawonn, K.; Preim, B. Narrative Medical Visualization to Communicate Disease Data. *Comput. Graph.* **2022**, *107*, 144–157.
10. Garrison, L.; Meuschke, M.; Fairman, J.; Smit, N. N.; Preim, B.; Bruckner, S. An Exploration of Practice and Preferences for the Visual Communication of Biomedical Processes. In *Eurographics Workshop on Visual Computing for Biology and Medicine*, 2021; pp. 1–12.
11. Gillmann, C.; Smit, N. N.; Gröller, E.; Preim, B.; Vilanova, A.; Wischgoll, T. Ten Open Challenges in Medical Visualization. *IEEE Comput. Graph. Appl.* **2021**, *41* (5), 7–15.
12. Ghosh, A. K.; Joshi, S. Enhancing Physician's Toolkit: Integrating Storytelling in Medical Practice. *J. Assoc. Physicians India* **2021**, *69* (7), 11–12.
13. McCurdy, K. Visual Storytelling in Healthcare: Why We Should Help Patients Visualize Their Health. *Inf. Visual.* **2016**, *15* (2), 173–178.
14. Mittenentzwei, S.; Preim, B.; Meuschke, M. Why, What, and How to Communicate Health Information Visually: Reflections on the Design Process of Narrative Medical Visualization. In *Eurographics Workshop on Visual Computing for Biology and Medicine*; Garrison, L., Jönsson, D., Eds., 2024.
15. Garrison, L. A.; Meuschke, M.; Preim, B.; Bruckner, S. Current Approaches in Narrative Medical Visualization. In *Approaches for Science Illustration and Communication*; Springer International Publishing: Cham, 2023; pp. 95–116.

16. Kleinau, A.; Stupak, E.; Mörth, E.; Garrison, L.; Mittenentzwei, S.; Smit, N. N.; Lawonn, K.; Bruckner, S.; Gutberlet, M.; Preim, B.; Meuschke, M. Is There a Tornado in Alex's Blood Flow? A Case Study for Narrative Medical Visualization. In *Proc. of EG Workshop on VCBM*, 2022; pp. 11–22.
17. Mittenentzwei, S.; Garrison, L. A.; 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*, 229–238.
18. Cavazza, M.; Charles, F.; Mead, S. J. Interacting with Virtual Characters in Interactive Storytelling. In *Proc. of Int. Conf. on Autonomous Agents and Multiagent Systems*, 2002; pp. 318–325.
19. Mariani, I.; Ciancia, M. Character-Driven Narrative Engine. Storytelling System for Building Interactive Narrative Experiences. In *Proc. of the Digra Int. Conf.: Game, Play, and the Emerging Ludo-Mix*; Digital Games Research Association, 2019; pp. 1–19.
20. Sheldon, L. *Character Development and Storytelling for Games*; CRC Press: Boca Raton, Florida, 2022.
21. Isbister, K. *Better Game Characters by Design*; CRC Press: Boca Raton, Florida, 2006.
22. Chen, M.; Bell, R. A.; Taylor, L. D. Persuasive Effects of Point of View, Protagonist Competence, and Similarity in a Health Narrative about Type 2 Diabetes. *J. Health Commun.* **2017**, *22* (8), 702–712.
23. Lee, N. *Encyclopedia of Computer Graphics and Games*; Springer International Publishing: Cham, 2024.
24. Dasu, K.; Kuo, Y.-H.; Ma, K.-L. Character-Oriented Design for Visual Data Storytelling. *IEEE Trans. Visual. Comput. Graph.* **2024**, *30* (1), 98–108.
25. Smith, E.; Fiasson, J.; Bhasme, J. Leonardo.ai, 2024. https://leonardo.ai (accessed 2024-05-3).
26. Hu, E. J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; LoRA, W. C. Low-Rank Adaptation of Large Language Models. In *International Conference on Learning Representations*, 2022. https://openreview.net/forum?id=nZeVKeeFYf9.
27. Civit AI, Inc. Civitai, 2024. https://civitai.com (accessed 2024-05-3).
28. AUTOMATIC1111. Stable Diffusion Web UI, 2024. https://github.com/AUTOMATIC1111/stable-diffusion-webui (accessed 2024-05-3).
29. So, W.; Bogucka, E. P.; Šćepanović, S.; Joglekar, S.; Zhou, K.; Quercia, D. Humane Visual AI: Telling the Stories Behind a Medical Condition. *IEEE Trans. Visualization Comput. Graphics* **2020**, *27* (2), 678–688.
30. Gritti, P. The Bio-Psycho-Social Model Forty Years Later: A Critical Review. *J. Psychosoc. Syst.* **2017**, *1* (1), 36–41.
31. Kluwer, W. UpToDate: Industry-Leading Clinical Decision Support — wolterskluwer.com, 2023. https://www.wolterskluwer.com/en/solutions/uptodate (accessed 2024-05-3).
32. Healthline. Healthline: Medical Information and Health Advice You Can Trust — healthline.com, 2023. https://www.healthline.com/ (accessed 2024-05-3).
33. Cleveland Clinic. Medical Questions & Answers | Cleveland Clinic — my.clevelandclinic.org, 2023. https://my.clevelandclinic.org/health (accessed 2024-05-3).
34. Journal of the American Medical Association. Patient Information — jamanetwork.com, 2023. https://jamanetwork.com/collections/6258/patient-information (accessed 2024-05-3).
35. Adobe Inc. Adobe Creative Cloud, 2019. https://www.adobe.com/creativecloud.html (accessed 2024-05-3).
36. Savage Interactive Pty Ltd. Procreate, 2023. https://procreate.com/ipad (accessed 2024-05-3).
37. Pettersson, J. Basics of Character Design. In *Larp Design: Creating Role-Play Experiences*; Landsforeningen Bifrost: Copenhagen, 2019; pp. 193–204.
38. Wang, Y.; Wang, F.; Cheng, Y.; Zhao, C.; Zheng, Z. The Study of Characters Design in prg Educational Games. In *IEEE International Conference on Computer Science and Information Technology*, 2009; pp. 44–47.
39. Ballon, R. *Breathing Life into Your Characters*; Penguin: City of Westminster, London, 2009.
40. Kuntjara, H.; Almanfaluthi, B. Character Design in Games Analysis of Character Design Theory. *J. Games Game Art Gamification* **2017**, *2* (2), https://doi.org/10.21512/jggag.v2i2.7197.
41. Klemm, P.; Glaßer, S.; Lawonn, K.; Rak, M.; Völzke, H.; Hegenscheid, K.; Preim, B. Interactive Visual Analysis of Lumbar Back Pain-What the Lumbar Spine Tells About Your Life. In *Proc. of IVAPP*, Vol. *2*, 2015; pp. 85–92.
42. Palmer, L. J. Uk Biobank: Bank On It. *Lancet* **2007**, *369* (9578), 1980–1982.
43. Hofman, A.; Brusselle, G. G. O.; Murad, S. D.; van Duijn, C. M.; Franco, O. H.; Goedegebure, A.; Ikram, M. A.; Klaver, C. C. W.; Nijsten, T. E. C.; Peeters, R. P.; Stricker, B. H. C.; Tiemeier, H. W.; Uitterlinden, A. G.; Vernooij, M. W. The Rotterdam Study: 2016 Objectives and Design Update. *Eur. J. Epidemiol.* **2015**, *30* (8), 661–708.
44. Jansen, B. J.; Jung, S.-G.; Nielsen, L.; Guan, K. W.; Salminen, J. How to Create Personas: Three Persona Creation Methodologies with Implications for Practical Employment. *Pac. Asia J. Assoc. Inf. Syst.* **2022**, *14*, 1–28.
45. Die Ernährungs-Docs | ARD. Fettleber heilen: Mit Ernährung natürlich gegensteuern und abnehmen | Die Ernährungs-Docs | NDR, 2022. https://www.youtube.com/watch?v=eTz-W73HNZU&t=54s (accessed 2024-05-3).
46. Die Ernährungs-Docs | ARD. Fettleber heilen mit Gemüse, Eiweiß und Vollkornprodukten | Die Ernährungs-Docs | ARD Gesund, 2023. https://www.youtube.com/watch?v=Q-mQdIC1lkY&t=355s (accessed 2024-05-3).
47. HarkerTechnologies. Ht Photorealism v4.1.7, 2023. https://civitai.com/models/114956/ht-photorealism-v417.
48. Hao, Y. Towards Data Science, 2024. https://towardsdatascience.com/image-to-image-translation-69c10c18f6ff (accessed 2024-05-3).
49. Wichmann, H.-E.; Kaaks, R.; Hoffmann, W.; Jöckel, K.-H.; Greiser, K.; Linseisen, J. Die Nationale Kohorte. *Bundesgesundheitsblatt — Gesundheitsforsch. — Gesundheitsschutz* **2012**, *55* (6–7), 781–789.
50. Federal Employment Agency. German Classification of Occupations 2010 — Revised Version 2020, 2024. https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010-Fassung2020/Arbeitsmittel/Arbeitsmittel-Nav.html (accessed 2024-01-29).
51. Willcox, G. The Feeling Wheel: A Tool for Expanding Awareness of Emotions and Increasing Spontaneity and Intimacy. *Trans. Anal. J.* **1982**, *12* (4), 274–276.
52. Zovya. Rev Animated v1.2.2, 2024. https://civitai.com/models/7371/rev-animated (accessed 2024-05-3).

53. Wall, S. *Medical Graphics and Graphic Medicine*; Springer: Nature Switzerland, 2023; pp. 23−40.
54. LONGD. Inksketch v1.5, 2024. https://civitai.com/models/84494/a-comic-sketches-style.
55. RealtimeBoard, Inc. Miro. https://miro.com/.
56. Snap Inc. Snapchat. https://www.snapchat.com.
57. Puhl, R. M.; Heuer, C. A. Obesity Stigma: Important Considerations for Public Health. *Am. J. Publ. Health* **2010**, *100* (6), 1019−1028.
58. Peng, B.; Fan, H.; Wang, W.; Dong, J.; Lyu, S. A Unified Framework for High Fidelity Face Swap and Expression Reenactment. *IEEE Trans. Circ. Syst. Video Technol.* **2022**, *32* (6), 3673−3684.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/icom-2024-0041).