Archives of Sexual Behavior
https://doi.org/10.1007/s10508-026-03518-x

SPECIAL SECTION: ARTIFICIAL INTELLIGENCE AND SEXUALITY

Visual Sexual Health Communication Assisted by AI: The Case of Synthetic Menstruation Images

Nicola Döring[1] · **Anastasiia Shevtsova**[1]

Received: 29 December 2025 / Revised: 11 April 2026 / Accepted: 23 April 2026
© The Author(s) 2026

Abstract

Public communication about sexual and reproductive health and rights (SRHR) relies on both text and images. A recent development is the proliferation of artificial intelligence (AI) image generators, which enable rapid, cost-effective, and highly customizable images with appealing aesthetics based on textual requests (prompts). To examine the opportunities and challenges of AI-generated images in SRHR communication, we used the topic of menstruation and developed six research questions (RQs) anchored in the current state of research on AI-supported visual health communication. These questions addressed key areas of concern, such as unsuccessful image creation (RQ1), objective image flaws (RQ2), lack of diversity of human representation (RQ3), stereotypical portrayals of women (RQ4), use of visual symbols (RQ5), and image valence (RQ6). Our study applied 15 AI image generators, 2 user accounts per generator, 5 menstruation-related prompts, 2 prompt languages, and 2 images per prompt, resulting in a sample of $N = 600$ AI outputs generated in Autumn 2025. All outputs were manually coded using a reliability-tested codebook developed through both deductive and inductive approaches. In 20.0% of cases, menstruation prompts did not produce the requested image but only text output (RQ1). When images were generated, they often (78.1%) exhibited visual and/or textual flaws (RQ2). The images showed a lack of diversity in some human attributes (e.g., body type) but not in others (e.g., skin tone; RQ3). Stereotypical depictions of women were observed but not predominant (RQ4). Common menstruation-related visual symbols (e.g., blood, pain) were identified (RQ5), but most images did not convey a negative portrayal and instead a neutral or positive portrayal of menstruation (RQ6). While results show that initial outputs from AI image generators are often not directly suitable for visual sexual health communication, they might work with targeted refinement. Further research is required to examine the production, characteristics, consumption, and effects of AI-generated images in SRHR communication.

Keywords Artificial Intelligence · Generative artificial intelligence (GenAI) · AI-generated images · Media content analysis · Sexual and reproductive health and rights

Introduction

Sexual and reproductive health and rights (SRHR) are a central component of overall health and human rights (Starrs et al., 2018). To take care of their SRHR (e.g., avoiding sexually transmitted infections and unwanted pregnancies, achieving romantic and sexual well-being), people require accessible health services and appropriate SRHR education and information. In addition to textual information, visual information is important for communicating health topics (Houts et al., 2006): some clinical details are complex and more easily shown in images than described in text (e.g., the insertion and position of an intrauterine device or a tampon). Moreover, in many communication contexts text alone is insufficient; images help attract attention and increase the clarity and persuasiveness of messages. This is why images are used in various SRHR communication contexts such as in press coverage of STI[1] vaccinations, in social media posts for LGBTQ+[2] rights campaigns, in global World AIDS[3] Day

✉ Nicola Döring
nicola.doering@tu-ilmenau.de

[1] Department of Economic Sciences and Media, Technische Universität Ilmenau, Ehrenbergstr. 29, 98693 Ilmenau, Germany

[1] STI = Sexually Transmitted Infection(s).

[2] LGBTQ+ = Lesbian, Gay, Bi, Trans, Queer and further sexual and gender identities.

[3] AIDS = Acquired Immune Deficiency Syndrome.

Published online: 03 November 2026

campaigns, and in sex education brochures on contraception and menstruation.

To encompass all images representing SRHR themes and used for informational and educational purposes, the umbrella term SRHR images or sexual health images is adopted (e.g., Jewitt, 1997; McCarthy et al., 2012). For sexual health actors—such as educators, clinicians, counselors, researchers, and activists—three principal pathways to obtain SRHR images exist:

- *Commissioned, customized images*: Commissioning professional photographers, illustrators, or graphic designers provides the highest degree of control over clinical accuracy, representational diversity, and legal consent, but entails higher costs and longer lead times (e.g., for contracting and briefing). Commissioned, customized images typically have a professional, polished appearance with the highest technical and compositional quality, making them well suited for formal sexual education materials and clinical publications.
- *Licensed stock images*: Selecting images from commercial stock libraries (e.g., Adobe Stock; Getty Images) offers rapid, licensable visuals at moderate expense, yet does not allow customization and risks perpetuating representational gaps and biases. Millions of stock images are available, many of which are categorized and tagged as SRHR-related. Licensed stock images are typically produced by visual communication experts. They have a professional aesthetic, but are sufficiently generic to be reusable across multiple contexts such as press reporting, company websites, and advertisements. Therefore, they are also criticized for their glossy, commercialized, and standardized appearance (Frosh, 2001, 2002).
- *Self-generated images*: Producing images in-house, for example, via smartphone photography, hand drawing, or graphic design software, offers speed, authenticity, and low cost, but requires strict consent procedures when depicting identifiable individuals and expert review when clinical accuracy is necessary. Self-generated images often have an amateur or informal appearance that can enhance perceived realness and credibility in social-media contexts but may lack the refinement needed for printed or clinical materials.

Decisions between these three routes rest on visual communication goals and involve trade-offs among aesthetics, accuracy, timeliness, and budget. Inspecting, for example, the sexual health section of the World Health Organization website,[4] the Instagram account of Planned Parenthood,[5] the online platform of the Asian sexual health provider PULSE,[6] the website of the Global Network of People Living with HIV[7] or the web presence of the African Sex Worker Alliance[8] reveal a variety of public-facing SRHR image types used for educational, informational, commercial, and activist purposes.

A recent development in visual sexual health communication is the broad availability of image generators based on artificial intelligence (AI) technology (Döring et al., 2025). Self-generated images can now also be created with this new class of tools. AI image generators offer the opportunity to generate customized SRHR images for different communication channels (e.g., academic publications, social media, websites, information flyers, brochures). Image generation with these systems is typically swift, low-cost, and highly customizable, and outputs can mimic a professional aesthetic. Text-to-image generators can create cartoons, drawings, paintings, diagrams, photorealistic images,[9] and other types of still images commonly used in SRHR communication based on simple textual requests (i.e., text prompts). Images generated with the help of generative AI tools are typically referred to as AI-generated images (short: AI images) or synthetic images.

The aim of this article is to explore the opportunities and challenges of AI image generators for visual sexual health communication. For this purpose, menstruation was selected as the example SRHR topic to be visualized with AI tools for several reasons: first, menstrual health represents a core domain of sexual and reproductive health and human rights that affects roughly half of the global population but remains surrounded by cultural taboos, stigma, and misinformation (Office of the High Commissioner for Human Rights, 2022; WHO, 2022). Addressing this through accurate, normalizing and inclusive textual and visual health communication supports global health and gender equity goals (Hennegan et al., 2021; Olson et al., 2022). Second, menstruation is a high-impact educational topic where adolescents and adults rely on educational offline and online information. AI image generators can foster menstrual information by producing trustworthy, appealing, and evidence-based educational visuals suitable for digital platforms and beyond (Brown et al., 2024; Evans et al., 2022; Olsson et al., 2024). Third, the topic

[4] https://www.who.int/health-topics/sexual-health.

[5] https://www.instagram.com/plannedparenthood/.

[6] https://www.pulse-health.org/.

[7] https://gnpplus.net/.

[8] https://home.aswaalliance.org/.

[9] AI tools can generate images that mimic the look of actual photography, those images are often called "photorealistic AI images" or shorter "AI photos." However, there is a fundamental difference between a traditional photography that uses a camera to capture a real scene in the physical world, and an AI-generated image mimicking the look of a photo but depicting content without any relation to a real-world scenario.

a) Prompt: "Create an image of an educator teaching reproductive health"



b) Prompt: "Create an image of an abortion rights activist"



c) Prompt: "Create an image of a Black lesbian couple"



d) Prompt: "Create an image of the clitoris"

Is there another image I can try? I'm here to help you create all types of things, but can't make images like that.

Fig. 1 Examples of AI-generated sexual health images and text output (Source: author-created with Gemini 2.5 Flash, October 8, 2025)

offers a valuable avenue to explore the creative and technical potential of AI image generation, as it requires nuanced depictions of bodily processes, diverse identities, and sensitive contexts (Bobel et al., 2020; Döring et al., 2024).

By systematically examining what AI image generators produce when prompted to visualize a stigmatized, diversity-sensitive SRHR topic like menstruation, this study offers a domain-specific content-analytic contribution to explainable AI research (Arrieta et al., 2020): treating the outputs of opaque text-to-image systems as a window into their assumptions and limitations. Such output-centered analysis can inform responsible AI (Arrieta et al., 2020) by providing the kind of empirical, domain-specific knowledge that sexual health communicators need to deploy these tools critically and carefully (WHO, 2024). In doing so, the study may also shed light on AI literacy (Ng et al., 2021) among sexual health professionals, offering insights into how they can use AI image generators to improve visual sexual health communication.

The following sections of this introduction will explain (1) image generation with AI tools, (2) visual communication with AI images, (3) the state of research on AI-generated health images, and (4) the aims of the current study. The introduction is followed by a description of the study method and a presentation of the study results plus their discussion.

Image Generation with AI Tools

Contemporary AI-based text-to-image (short: TTI, T2I) generators translate natural-language prompts (e.g., "create an

image of an educator teaching reproductive health," "create an image of an abortion rights activist," or "create an image of a Black lesbian couple") into synthetic images (see Fig. 1). Technically, these systems combine a text-encoding component with a deep generative image model, typically based on neural-network architectures that learn statistical associations between textual and visual data (Ho et al., 2020; Rombach et al., 2022). Among the most influential architectures are diffusion models and autoregressive models. Diffusion models as implemented, for example, in TTI generators like Midjourney and Stable Diffusion, work like a photograph emerging from fog: the system starts with pure visual noise and gradually clarifies it into a coherent image, guided at each step by the text prompt. Autoregressive models as implemented, for example, in ChatGPT-Image 1 and Gemini 2.5 Flash, work more like a painter filling in a paint-by-numbers canvas: the image is built piece by piece, with each new section informed by what has already been placed.

AI image generators are trained on vast datasets of image-caption pairs, mainly scraped from the Internet, with LAION-5B,[10] a dataset with 5.85 billion image-text pairs, being a primary example (Schuhmann et al., 2022). The outputs of AI image generators depend on the characteristics of the training data. This is why ethical and responsible AI image generators are trained with preprocessed, cleaned, and diverse datasets to ensure, for example, that inaccurate captions and unethical images (e.g., execution images, child sexual abuse material) are eliminated, and that culturally diverse images are included. However, the degree of achieved diversity is an issue of constant debate and auditing. The initiative Queer in AI,[11] for example, collaborates with academic, policy, and industry partners to raise awareness of how AI systems represent sexual and gender diversity and to support community-led, participatory approaches to AI design and governance, including work on diverse and fair AI training data (Organizers of Queer in AI et al., 2023).

To foster ethical and responsible AI image generation, developers often combine training-data curation (e.g., filtering harmful content, increasing diversity) with human-in-the-loop alignment techniques—for example, training a reward model from human annotations and then fine-tuning the generative model so that the AI model better yields desirable outputs and suppresses unwanted ones (e.g., biased, stereotyped or distorted AI images; Liang et al., 2024). These efforts reflect broader principles of responsible AI, particularly fairness, transparency, accountability, inclusivity, and mitigation of representational bias (Arrieta et al., 2020).

Publicly available AI image generators also implement a content moderation layer at runtime: before a generated image is shown, content-filtering components and user-facing guardrails are applied. These systems are designed to block image outputs that could violate ethical norms or platform policies (e.g., prohibition of sexually explicit or violent imagery; Vargas Penagos, 2024). However, in sensitive fields like SRHR communication, algorithmic content moderation frequently results in overblocking. Overblocking means the AI system mistakenly flags and blocks the generation of legitimate sexual health images as supposedly inappropriate content (e.g., "create an image of the clitoris"; see Fig. 1; Döring et al., 2025). AI outputs are also shaped by the user's account location because providers apply region-specific content policies and legal restrictions, and content blocking may therefore differ between world regions.

While AI image generators are technically constrained by their neural network architectures, training data, training processes, and global and regional content-moderation rules, the types of images created are largely determined by the users' inputs. Their text prompts[12] can specify the image form (e.g., an illustration or a photorealistic image), the manifest image content, including people, objects, surroundings (e.g., a White heterosexual couple indoors or a Black lesbian couple outdoors), and the latent image content (e.g., desperate or romantic atmosphere). Users can shape the output through the individual prompts within a particular session. Some AI image generators also allow users to define persistent meta-prompts (e.g., custom instructions such as "always create colorful illustrations with a hand-drawn aesthetic") that are automatically appended to all queries, creating a stable form of personalization. At the same time, image generators can operate without meta-prompts and even with highly minimal prompts (e.g., "create an image of a couple"), where the user provides no guidance regarding form or content and the stochastic model of the AI image generator must determine all details on its own. In these cases, the AI image generator defaults to its most common, average, stereotypical, and safe visual patterns, and avoids risky, rare, or ambiguous content and aesthetic choices.

[10] LAION (= Large-scale Artificial Intelligence Open Network) is a German non-profit that publishes large open image-caption datasets (e.g., LAION-400M with about 400 million images; LIAON-5B with about 5 billion images) which have been used to train popular text-to-image models such as Stable Diffusion (by Stability AI) and DALL-E mini (by OpenAI). Website: https://laion.ai/.

[11] https://www.queerinai.com/.

[12] In addition to text prompts, many generative AI models can also work with image prompts—that is, they can generate a new image based on one or more reference images. For the sake of clarity and focus, this article, like many related publications, deliberately concentrates on text prompts. Text prompts without image examples, are also often called "zero-shot prompts," in contrast to "one-shot prompts" or "few-shot prompts," where the text prompt is accompanied by one or more image examples to steer the AI image generator in the intended direction.

Dozens of AI image generators are now publicly available to the broad population. While some offer a free trial plan, many require a paid subscription to generate larger numbers of images. Leading AI chatbots such as ChatGPT[13] by OpenAI and Gemini[14] by Google are integrated with image generators, eliminating the need to visit a dedicated website or use a separate app for AI image generation. Consequently, creating images with AI tools has become an everyday media practice accessible not only to visual communication experts, but also to laypersons and professionals across various fields, including sexual and reproductive health and rights.

Visual Communication with AI Images

Visual communication refers to communication that uses visual aids as opposed to solely verbal and nonverbal communication. It encompasses diverse forms of moving and still images, including videos, animations, paintings, drawings, photographs and charts across different media (Thelander, 2018). Within this landscape, AI-generated or synthetic images represent a new category of visuals that complements existing types of visual communication.

To analyze visual communication processes, four dimensions are relevant (Thelander, 2018), which are applicable to AI-generated imagery: (1) image production, (2) image characteristics, (3) image consumption, and (4) image effects.

Production of AI Images

The image production process includes the communicator's communication goals, choice of generative AI tool(s), prompt engineering and iterative refinement, image selection, and image dissemination. The process of image production to reach certain communication goals is also called visualization (e.g., ideation about how to visualize abstract concepts such as pain, shame, pride, or pleasure) and visual persuasion (e.g., planning how to design images to persuade people to consider HPV[15] vaccination for their children), and can be part of strategic communication (e.g., strategic decisions about condom advertisement design that reflects the company's brand and increases sales). The process of image production with AI tools is always a human-guided process, but AI image generators can help the communicator with visualization ideas, image production, and image refinement (Mahdavi Goloujeh et al., 2024; Oppenlaender et al., 2025). In their text prompt, the communicators can tell the AI tool not only what type of image should be generated (i.e., details about image form and content) but also in which contexts the image will be used (e.g., on social media or in a medical brochure), so that communication goals are considered during image generation and may be reflected in image style.

Characteristics of AI Images

AI-generated images—like any other images—can be characterized along multiple dimensions that correspond to different levels of visual framing (Rodriguez & Dimitrova, 2011; Thelander, 2018). A key distinction is between form (e.g., image type, composition, lighting, perspective, and resolution) and content. Content can be further subdivided into (1) manifest content (explicit, observable elements or denotative themes, e.g., two women sitting on a bench in an urban environment with flowers in the background, holding hands; see Fig. 1), (2) latent content (implicit, interpretive, connotative meanings, e.g., expression of intimacy and joy between the women, the flowers symbolizing romance, beauty, and flourishing), and (3) ideological content (engagement with broader cultural structures, e.g., the acknowledgment, demarginalization, and normalization of same-sex couples; the framing of cities as inclusive and diverse spaces where different identities coexist safely).

AI-generated images are evaluated not only on their form and content but also regarding their degree of artificiality, especially when contrasted with human-made photographs (Aktay, 2022; Velásquez-Salamanca et al., 2025). This includes aspects such as reduced realism, unnaturally glossy or eerie visual qualities, and the presence of AI hallucinations or textual and visual flaws (e.g., generation errors such as distorted human bodies or non-existent objects and misspelled or inaccurate image captions).

Consumption of AI Images

When AI-generated images with their respective characteristics explained above are shared or published in various media and communication settings, image audiences can engage with them in multiple ways spanning cognitive and emotional processes of attention, perception, identification, interpretation and also behavioral processes of active selection or deselection, public evaluation, sharing, and commenting. For example, some social media users actively subscribe to Facebook groups,[16] Instagram accounts, or TikTok channels that publish AI images labeled as such and consciously consume, enjoy, and promote AI imagery. Other users encounter AI images unconsciously because they confuse them with human-made photographs or videos. Still other users identify AI-generated images when they are exposed to them on social

[13] https://chatgpt.com/.

[14] https://gemini.google.com/.

[15] HPV = Human Papillomavirus.

[16] Facebook group "AI Art Universe" with more than 650,000 members https://www.facebook.com/groups/aiartuniverse.

media and actively reject them through ignoring, unsubscribing, disliking, and negative comments. Several studies have investigated whether people can differentiate AI-generated imagery from human-generated photographs and how they evaluate AI images (Califano & Spence, 2024; Döring & Mohseni, 2025; Högemann et al., 2025).

Effects of AI Images

Different effects can result from AI image consumption, which may or may not align with the intentions of the AI image creators and can be described on cognitive, emotional, and behavioral levels (e.g., Chan, 2024, 2025). While effects of AI-generated SRHR images have not yet been systematically investigated, it is plausible to assume that they could elicit cognitive outcomes (e.g., better understanding and recall of different menstrual products), emotional responses (e.g., more positive attitudes toward people living with HIV), and behavioral consequences (e.g., increased intention to take an STI test). Media effects research typically differentiates between short-term effects immediately after the perception of AI images (e.g., temporarily increased interest in the topic of HPV vaccination) and long-term effects resulting from the consumption of large numbers of AI-generated images over time (e.g., distortions of one's body image due to constant exposure to unrealistically clear and smooth skin in photorealistic AI images).

State of Research on AI-Generated Health Images

No study could be identified by the authors that specifically analyzed AI-generated sexual health images. A search in Google Scholar, Web of Science, PubMed, and Scopus did, however, yield 30 studies published in peer-reviewed journals across humanities, social sciences, medicine, and computer science between 2023 and 2025 that examined AI-generated health images more broadly. These studies cluster into three thematic groups: AI images of (1) health professionals, (2) patients, and (3) populations with particular health relevance.

AI images of health professionals: Eight studies investigated how AI image generators visualize health professionals such as "doctors," "surgeons," "anesthesiologists," "otolaryngologists" or "nurses" in photorealistic images (see Table 1). Researchers were mainly interested in representational biases (such as under-representation of women and non-White people as doctors or surgeons; Ali et al., 2023; Campbell et al., 2026) and presentational biases or stereotypes (such as smiling and sexualized presentations of women doctors and women nurses; Reed et al., 2025; Sun et al., 2024). Overall, all eight studies confirmed the expected biases in AI images, albeit with some variation between AI tools, in that DALL-E generated more diverse and less sexualized images than Midjourney (e.g., Cevik et al., 2023; Gisselbaek et al., 2024).

AI images of patients: Twelve studies explored how AI image generators picture patients with different health conditions such as a "cancer patient," a "patient with a disability," a "person with dementia," a "frail older adult" or a "middle-aged Caucasian woman with hypothyroidism" in photorealistic images (see Table 2). Again, researchers identified biases and stereotypes such as the AI portraying cancer patients predominantly as attractive, White, healthy-looking women with a breast cancer awareness ribbon and/or a headband (Senft Everson et al., 2026; Varela-Rodriguez & Plage, 2025; Wang & Wang, 2025). In this way, AI-generated cancer patient imagery makes the diversity of cancers and affected populations, as well as aspects of ill-health and suffering, invisible in favor of promoting an idealized idea of the happy female breast cancer survivor. Three of the twelve studies focused not on photorealistic patient pictures but on medical imagery, such as photorealistic images of surgical operations (Buzzaccarini et al., 2024), AI-generated X-ray images of a healthy lung versus a collapsed lung (Amri & Hisan, 2023), and anatomical illustrations of a healthy heart versus a heart affected by a specific type of congenital heart disease (Temsah et al., 2024). These three analyses showed that the vast majority of AI-generated medical images were anatomically incorrect (see Table 2).

AI images of different populations with particular health relevance: Ten studies examined how AI image generators depict members of different populations with particular health relevance such as a "poor person," a "beautiful woman," a "fat person," "body positivity and teenagers," a "female activist leading a movement" or a "couple with a disability kissing" in photorealistic images (see Table 3). Again, researchers identified different types of biases and stereotypes as well as image flaws. Overall, the studies criticize photorealistic AI images for depicting people of different populations predominantly in a normative way, such as White, young, attractive, slim, and able-bodied (e.g., Bianchi et al., 2023; Sobey, 2025; Westberg & Kvåle, 2024).

Summary: The 30 identified studies on health-related AI images all follow a similar study design: researchers typically choose two or more AI image generators (minimum: 1; maximum: 15) and create a sample of several dozen, hundreds, or thousands of AI images (minimum: 64; maximum: 15,300) using a set of simple and deliberately vague zero-shot text prompts that usually provide no context or visual communication goals (e.g., the AI is prompted to generate an image of "a cancer patient," "a female activist leading a movement," "a couple with a disability kissing") without including the communication purpose or outlet (e.g., for a social media campaign; for a patient-facing brochure; for medical education). As explained above, AI image generators operate with stochastic models, such that the common use of very vague minimal prompts in previous studies has actively steered AI

Table 1 Eight studies on AI-generated images of health professionals

No.	Citation	AI image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Sample of AI images	Main results
1	Alenichev et al. (2023)[a] Alenichev et al. (2024)[a]	Doctors and patients Global health images Africa representation	Midjourney 5.1	"A group of Black African doctors in the style of photojournalism" "African doctors administer vaccines to poor White children in the style of photojournalism"	Not reported in detail (> 300 images)	Generative AI reproduced and amplified colonial and racialized biases Even when prompted with Black African doctors helping White children, AI images often flipped to Black children and White doctors
2	Ali et al. (2023)	Surgeons across eight specialties	DALL-E 2 Midjourney v5.1 Stable Diffusion v2.1	"A photo of the face of a [surgical specialty]" "A photo of the face of a [surgical specialty] in the USA [China / Nigeria]"	3 AI tools × 8 specialties × 100 images per prompt = 2400 images Additional geographic sample: 3 AI tools × 4 prompts (neutral + 3 countries) × 100 images per prompt = 1,200 images	AI images of surgeons overwhelmingly depicted White men AI images consistently underrepresented women across all eight surgical specialties and roles DALL-E 2 was the least biased of the three (showing more racial variety and slightly more women) Country prompts boosted non-White depictions but did not fix gender underrepresentation
3	Campbell et al. (2026)	STEM-related occupations	Adobe Firefly DALL-E 2 Stable Diffusion	"STEM Scientist" "Woman scientist"	2 AI tools × 10 prompts × 2 styles + 1 AI tool × 10 prompts × 1 style Final dataset: 132 + 127 + 32 = 291 images; each double-coded = 582 codes	AI images lacked representativeness in the sense of underrepresentation of women and non-White ethnicities/races AI images presented stereotyped roles (gendered activity cues) and styling (sexualization) AI images had anatomical distortions (e.g., regarding hands/fingers, eyes)
4	Cevik et al. (2023)	Eight types of surgeons	DALL-E 2 Midjourney	"General Surgeon" [surgical specialty]	2 AI tools × 8 prompts × 4 images per prompt = 64 images	AI images reproduced stereotypical representations of surgical professionals (male, light-skinned, slim, 50 + years old) Midjourney produced more biased images than DALL-E

Table 1 (continued)

No.	Citation	AI image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Sample of AI images	Main results
5	Gisselbaek et al. (2024)	Anesthesiologists across six subspecialties	DALL-E Midjourney	"A photo of the face of a general anesthesiologist [anesthesiologist specialty] head of anesthesiology department"	2 AI tools × 6 thematic prompts × 100 images per prompt = 1,200 images	AI images lacked diversity and representativeness (e.g., regarding gender, ethnicity/race, age) AI images reflected discipline stereotypes (e.g., pediatric anesthesiologist more "baby-faced" than cardiac anesthesiologist) DALL-E images were more diverse (race/ethnicity and gender) than Midjourney images
6	Reed et al. (2025)	Nurses	Adobe Firefly ImageFx Midjourney	"Nurse from 1920s" "Nurse from [decade up]" "(Standard) nurse today" "Nurse of the future"	3 AI tools × 12 prompts × 8 images per prompt = 288 images	AI images contained stereotyped/sexualized tropes AI images predominantly depicted nurses as women Adobe Firefly produced greater diversity and less sexualization, Midjourney was the least diverse and most sexualized, with ImageFx in between
7	Shaari et al. (2025)	Different types of otolaryngologists	DALL-E 3 Runway Midjourney	"Portrait of a [demographic] [trait] [rank] [subspecialty] American otolaryngologist"	3 AI tools × 29 prompts × 20 images per prompt = 1740 images	AI images of different types of otolaryngologists predominantly showed White, male persons Midjourney and DALL-E 3 were more diverse than Runway
8	Sun et al. (2024)	People in 153 different occupations including 24 health-related roles	DALL-E 2	Male-dominated occupations: "a mechanical engineer" "a repair worker" Female-dominated occupations: "a mental health counselor" "a childcare worker"	1 AI tool × 153 thematic prompts × 100 images per prompt = 15,300 images	AI images under-represented women in male-dominated occupations and over-represented them in female-dominated occupations, as identified by comparisons with U.S. census data (representational gender bias) AI images showed women more often than men with smiling faces and faces pitched downwards (presentational gender bias) AI images showed more bias than Google Images

Note. In alphabetical order of authors. [a] The same study is reported in two related publications

Table 2 Twelve studies on AI-generated images of patients

No.	Citation	AI Image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Number of AI images	Main results
1	Amri and Hisan (2023)	Medical images X-ray images of a normal lung versus a collapsed lung electrocardiogram (ECG) graphs photorealistic images of diseases/wounds	DALL-E	"Pneumothorax X-ray result" "ECG graph of a normal patient" "Ischemic Foot Ulcer of a Diabetic Patient"	Not reported (demonstration study; at least 13 example images/figures)	AI tools can quickly create teaching visuals, providing a cost-effective solution to simulate real-world scenarios AI images were aesthetically convincing but often contained medical inaccuracies AI tools that generate disease/wound imagery should be restricted/controlled
2	Buzzaccarini et al. (2024)	Images of facial plastic surgery Anatomical images of the face	Midjourney	Imagine: "a blepharoplasty surgical operation" "a rhinoplasty surgical operation" "a face anatomy picture for aesthetic surgeon medical training"	Not reported (demonstration study, narrative approach; at least 7 example images/figures)	AI anatomical images were highly fictionalized, far from true anatomical structures or actual surgical operations AI photorealism did not equal correctness and could have deceived non-medical specialists or the public AI images (Midjourney) were unsuitable for medical training or patient consultation and inappropriate for scientific publications
3	Chou et al. (2026)	Patients with four different types of cancers	DALL-E Stable Diffusion	"A photograph of a cancer patient breast cancer patient prostate cancer patient lung cancer patient"	2 AI tools × 4 thematic prompts × 40 images per prompt = 320 images	AI images lacked diversity (e.g., ethnicity/race), with DALL-E producing more racially diverse but less photorealistic outputs AI images partly perpetuated stereotypes (e.g., focus on hair loss/ head covering) Breast-cancer prompts yielded younger, healthier, more positive portrayals
4	Ertman et al. (2026)	People with disabilities in health	DALL-E 3	"Person with a disability" "Patient with a disability" "Doctor with a disability" "Doctor with a disability and a patient without a disability"	4 prompts × 50 images = 200 images	AI image generation often defaulted to a disabled patient and an able-bodied doctor, as observed in a prompt-compliance failure AI images portrayed almost exclusively visible disability, most often a mobility impairment, underrepresenting invisible disabilities

Table 2 (continued)

No.	Citation	AI Image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Number of AI images	Main results
5	Jintaganon et al. (2025)	Older adults Dementia Alzheimer's Disease	DALL-E 3 Midjourney V6	Dementia Prompts: "Photorealistic portrait of a person [man, woman] (living) with dementia [Alzheimer's Disease]" Age Prompts: "Photorealistic portrait of an older person [man, woman]" Control Prompts: "Photorealistic portrait of a person [man, woman]"	2 AI tools × (Age (2 phrasings × 3 groups × 28 each) + Dementia (4 × 3 × 16) + Control (3 × 56) = 1,056 images	AI images portrayed persons with dementia exhibiting more negative affect than older adults or adults with unspecified age 'Living with dementia' generated more positive-affect personas than 'with dementia,' which produced the most negative-affect portrayals DALL-E images were more diverse than Midjourney images
6	Kumar et al. (2024)	Hypothyroidism Horner's syndrome	DALL-E 2 Midjourney	An image depicting a middle-aged Caucasian woman with hypothyroidism An illustrative depiction of a patient displaying Horner's syndrome	2 illustrative images selected from a sample of 240 AI images	AI images depicted characteristic disease symptoms and were useful for medical training in principle, although anatomical and visual inaccuracies were common and required editing, indicating a current limitation
7	Martens et al. (2025)	Older adults Ageism	DALL-E 2	Set of combinations: [man/woman] [with condition] [doing activity] [in setting] "Successful Aging" "Frail Older Adult" "Lively Old Woman" "Old Man Running"	At 2 time points 31 prompts entered 1, 2 or 3 times = 164 images per time point	AI images over-represented White middle-class older adults and under-represented racial/ethnic diversity AI images showed older adults with negative emotions AI images showed no improvement in digital ageism across two years, continuing to depict older adults with stereotyped, negative, and homogeneous portrayals
8	Putland et al. (2023)	Dementia	Stable Diffusion 1.4	"Dementia"	3 visual formats × 19 diffusion samplers × 3 images per prompt = 171 images	AI images showed a distinct lack of visual diversity, portraying predominantly older, light-skinned people, passive and immobile, reinforcing the idea of dementia with old age AI images stereotyped dementia through biomedical reduction (the "diseased brain" over the person)

Table 2 (continued)

No.	Citation	AI Image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Number of AI images	Main results
9	Senft Everson et al. (2026)	Cancer patients Cancer survivors	DALL-E 3 Stable Diffusion	"A photograph of a cancer patient cancer survivor"	2 AI tools × 2 prompts × 40 images per prompt = 160 images	AI images lacked diversity and portrayed cancer survivors predominantly young, White, feminine, happy, and healthy, also often thin and conventionally attractive in neutral contexts AI images of cancer survivors were more positive than of those of cancer patients, and less likely to show illness markers or medical settings
10	Temsah et al. (2024)	Congenital heart diseases (CHD) Normal heart	DALL-E 3	"Draw a 2D accurate illustration of [CHD] to simplify it for medical students, with text in the image to clarify the illustration"	20 types of CHD × 5 images + 10 images of the normal heart = 110 images	AI medical illustrations of congenital heart diseases were mainly rated by experts as anatomically inaccurate, unsuitable for medical education, and visually unattractive
11	Varela-Rodríguez and Plage (2025)	Breast cancer Prostate cancer Pancreatic cancer	DALL-E 2 Stable Diffusion	"[Pancreatic / prostate / breast] cancer" "Pancreatic cancer awareness" "[type] cancer survivor" "[type] cancer patient"	262 images (56 breast cancer, 39 pancreatic cancer, 47 prostate cancer, 120 non-site-specific cancer)	AI images stereotyped cancer by depicting it predominantly as White, female, and heteronormative, narrowing who is seen as a patient or survivor AI images under-represented medical treatments unless explicitly prompted Stable Diffusion showed a wider emotional range, whereas DALL-E 2 tended to generate positive or neutral, polished portrayals
12	Wang and Wang (2025)	Breast cancer advertising images	DALL-E Midjourney 6.1	"Breast cancer" "Breast cancer awareness" "Breast cancer prevention" "Breast cancer screening" "Breast cancer advocacy" "Mammography"	17 prompts × 2 phrasings (with/without "advertisements") × 8 images per prompt = 272 images	AI images presented cancer survivors as young, sexy, slim women, favoring pink and optimistic styling, while ignoring medical information and excluding men

Note. In alphabetical order of authors

Table 3 Ten studies on AI-generated images of different populations with particular health relevance

No.	Citation	AI image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Sample of AI images	Main results
1	Bianchi et al. (2023)	Demographic depictions (traits, occupations, identities)	DALL-E Stable Diffusion	A photo of the face of [DESCRIPTOR] "an attractive person" "a poor person" "a happy family" A photo of the face of [OCCUPATION] "a software developer" "a housekeeper"	Not reported (> 300 images)	AI images amplified racial, gender, and class stereotypes AI images failed to produce counter-stereotypes: even with explicit counter-prompts ("wealthy" or "mansion" aimed at over-association of poverty with Blackness), models often failed to generate the intended outputs and reverted to the stereotypes
2	Elhosary (2026)	Poverty Migration and refugee crisis Social movements Global health Economic instability Terrorism The Gaza war	DALL-E 3	35 neutral, 35 framed, and 230 counter-prompts across seven news topics: "Generate an image of a/an [SUBJECT]" (e.g., "Female activist leading a movement"; "Doctor treating a patient")	300 prompts × 4 images per prompt = 1200 images	AI tool defaulted to male, youth/adults, Western-looking, light skin and slim bodies. Children were completely absent unless explicitly prompted AI images oversimplified non-Western cultures via stereotypical, outdated symbols, victimization and othering AI images exhibited masculine primacy/female subordination (men in authority, women passive/caregiving) AI images sidelined older adults, who appeared infrequently and mostly in negative contexts (e.g., poverty, refugeehood)

Table 3 (continued)

No.	Citation	AI image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Sample of AI images	Main results
3	Freixa et al. (2025)	Demographic depictions Comparison with stock images	Adobe Stock Lexica (Stable Diffusion)	Photography of a smiling person in a landscape in the city on the beach	2 AI tools × 3 prompts × 50 images per prompt = 300 images	AI images illustrated a more balanced gender mix than traditional stock images (Shutterstock, Getty Images), which over-represented women AI images showed only slight gains in African-American presence, while many groups remained largely invisible (e.g., people from Middle East, Africa, South Asia, and Indigenous people) AI images maintained persistent Eurocentric and youth/beauty bias, showing no disability representation, and limited professional diversity
4	Jääskeläinen et al. (2025)	Intersectional representations across social categories	Stable Diffusion	"A color photograph of a [GROUP] (e.g., poverty/wealth; white/blue collar)" "A color photograph of a [IDEOLOGY] (e.g., conservative, liberal)" "A color photograph of people [EVERYDAY SCENARIO] (e.g., doing laundry, relaxing)"	18 prompts × 20 images per prompt (10 randomly selected) = 180 images	AI images reproduced intersecting power systems (racism, sexism, heteronormativity, capitalism, and digital neocolonialism) The AI tool presented a "default person": White, able-bodied, masculine, with masculine subjects framed as authoritative AI images exhibited cultural centrism: Euro-/North America-centric settings and fashion
5	Kenig et al. (2023)	Body aesthetics	Craiyon	[Beautiful/perfect] breasts [Beautiful] woman [Beautiful] man [Beautiful] foot [Beautiful] ear [Beautiful] nose "Eyes" "Dog"	334 images (sampling plan not published)	AI images depicted highly sexualized images of female breasts (e.g., as an "oversized and sexually suggestive organ")

Table 3 (continued)

No.	Citation	AI image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Sample of AI images	Main results
6	Li (2026)	Female cyborg	ERNIE-ViLG 2.0 DALL-E Deep Dream Fotor Jasper Art NightCafe Shuhua Yihui	"Female cyborg" "A female cyborg is doing something at home, reflecting the role she plays at home" "A female cyborg is working, reflecting the role she plays in the workplace" "A female cyborg is fighting, reflecting the role she plays in the fighting"	8 AI tools × 4 thematic prompts × *n* images per prompt (free-quota dependent) = 107 images	AI images overwhelmingly depicted White, female cyborgs with sexual characteristics (overemphasized curves and highlighted breasts) AI images in home/work scenes often reverted to traditional female roles (e.g., cooking/caregiving), while fighting scenes showed empowered, combative agency, revealing context-dependent gendering Chinese AI tools showed muted sexual traits in domestic/work scenes and high-tech empowerment in combat; Western AI tools applied a consistent futuristic aesthetic but more often exaggerated sexual features
7	Martino et al. (2025)	Disabled couples and sexuality	Bing Craiyon Deepai Dezgo Fotor Free Image Generator Freepik Gencraft Hypotenuse AI Imagine.art Picsart Pixir Stable Diffusion VEED.IO Visme	"A couple with a disability kissing"	15 AI tools × up to 9 images per prompt = 69 images	AI images signaled disability mainly via visible aids (wheelchairs, glasses), while invisible disabilities were largely absent AI images of disabled intimacy were mostly heterosexual and White, sidelining queer, interracial, and gender-diverse couples AI images portrayed disabled people as romantic and sexual beings, showing autonomy (kissing, hand-holding, embracing), countering media norms that often infantilize and desexualize them

Table 3 (continued)

No.	Citation	AI image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Sample of AI images	Main results
8	Saumure et al. (2025)	Humor and discrimination	DALL-E 3	Prompts describing a human carrying out an action Make the image "funnier"	150 prompts × 2 research assistants × 2 runs (original vs. "funnier") = 600 images	AI images underrepresented women and people with high body weight even at baseline (pre-humor) When asked to make images "funnier," the AI tool showed more older, heavier, and visually impaired people and fewer racial minorities and women, showing "punch-down" humor, and increasing depictions of stigmatized traits (age, disability, high body weight)
9	Sobey (2025)	Body size	Adobe Firefly Bing Creator Canva Dream Studio Gencraft NightCafe RunwayML Stable Diffusion StarryAI	[Photo / photorealistic image] "a [ADJECTIVE] [SIZE] [PERSON] [doing ACTIVITY] [in SETTING]" fat/overweight/obese/morbidly obese happy/healthy/unhealthy exercising/eating a meal/having fun/dancing in their home/at a party	9 AI tools × 20 prompts × up to 4 images per prompt = 649 images	AI tools defaulted to a thin, able-bodied "normate" without explicit prompts for larger bodies, no fat or disabled bodies appeared AI images with fatness prompts showed more mistakes and anomalies (e.g., stomach "splits," added rolls at joints, malformed crotches) and were whiter, more male, and showed fewer positive facial expressions AI images prompted for fatness contravened content guidelines more often (e.g., nudity/policy breaches)

Table 3 (continued)

No.	Citation	AI image topic(s)	AI image generator(s)	Thematic prompt(s)—selection	Sample of AI images	Main results
10	Westberg and Kvåle (2024)	Teenagers	DALL-E 3	"Generate an image of teenagers doing stuff teenagers normally do" "Body positivity and teenagers" Innovative/surreal prompts	Not reported (at least 44 example images)	AI images depicted teenagers aligned with culturally dominant "good-looking" body norms (slim and fit, none overweight) AI images exhibited conformist diversity, with different skin tones but within a single template (cheerful, tidy, middle-class teens), with roles/power relations rarely changing AI images reflected cultural centrism (Western/European-American settings, clothing) AI images projected promotional positivity by default and pursued innovative surrealization when pushed

Note. In alphabetical order of authors

tools toward generating images with the most common, average, stereotypical, and safe visual patterns.

The most prevalent method of image analysis used in the identified studies is manual quantitative content analysis, while qualitative content analysis (e.g., Wang & Wang, 2025) and automated content analysis (e.g., Sun et al., 2024) were seldom used.

The main goal of most studies was the demonstration of representational and presentational biases or mistakes in AI images. Hence, the background of the studies is typically aligned with theories of AI bias and representational and presentational harm. Image analysis is focused on aspects such as diversity of depicted people in terms of gender, ethnicity/race/skin tone, age or (dis)ability to evaluate over- or under-representation in the sense of representational bias. Furthermore, to measure presentational biases, image analysis focuses on visual stereotypes, for example gender stereotypes in the sense that women are shown more often smiling, with a tilted head or in a sexualized way than men. The analysis of AI-generated image mistakes includes distortions of the human form (e.g., extra fingers) and factual inaccuracies in anatomical images (e.g., non-existing or distorted anatomical structures or incorrect in-image text labels).

As their main result, most studies point to a relatively high prevalence of mistakes and representational and presentational biases in the analyzed AI-generated images. As a countermeasure, researchers mostly demand that AI image generators be trained with more diverse data sets to produce more diverse and less stereotyped visual outputs. Only four of the 30 studies concluded that communicators need to (a) strengthen their prompting skills to provide the AI with goal-aligned, specific prompts and (b) refine AI images several times so as to produce images that align with their communication goals, such as producing helpful educational material (Campbell et al., 2026; Elhosary, 2026; Kumar et al., 2024; Temsah et al., 2024).

The Current Study

The current exploratory study investigates, for the first time, the characteristics of AI-generated sexual health images using the example of menstruation-themed images. It follows prior quantitative research designs that produce and analyze researcher-generated AI images, but advances this work by adopting a visual communication perspective and grounding image generation in a specific use case embedded in the prompts, namely AI image use in an educational brochure. The study is anchored in the state of research on AI image characteristics (see Tables 1, 2, 3) and related theories of AI-bias and representational and presentational harms. It aims to answer six preregistered research questions presented below (preregistration is available at https://osf.io/h5pn3/).

Regarding the issue of generative AI tools' tendency to overblock SRHR-related prompts (Döring et al., 2025; Vargas Penagos, 2024), we asked:

RQ 1: How often do AI image generators succeed or fail to generate menstruation-themed images?

Concerning the issue of prevalent AI hallucinations, including textual and visual inaccuracies in AI images (Amri & Hisan, 2023; Buzzaccarini et al., 2024; Campbell et al., 2026; Döring et al., 2025; Temsah et al., 2024; Westberg & Kvåle, 2024), we asked:

RQ 2: How prevalent are visual and textual flaws in AI-generated menstruation-themed images?

Relating to the issue of prevalent representational biases in terms of under-representation of certain genders, age groups, ethnicities/races/skin tones, or body types (e.g., Ali et al., 2023; Campbell et al., 2026; Westberg & Kvåle, 2024), we asked:

RQ 3: How diverse are representations of people in AI-generated menstruation-themed images?

Pertaining to the issue of prevalent presentational biases and in particular gender stereotypes (Bianchi et al., 2023; Elhosary, 2026; Wang & Wang, 2025), we asked:

RQ 4: How gender stereotypical are depictions of women in AI-generated menstruation-themed images?

In connection with the demands of visualization and the use of recognizable menstruation symbols, such as blood, pain and menstrual products (Bobel et al., 2020; Döring et al., 2024), we asked:

RQ 5: Which visual symbols are used in AI-generated menstruation-themed images?

With respect to the problem of stigmatization of menstruation in many social and cultural contexts and the call for normalization (Bobel et al., 2020; Döring et al., 2024), we asked:

RQ 6: How prevalent are negative versus positive portrayals of menstruation in AI-generated menstruation-themed images?

Together, these six research questions delineate a structured framework for examining AI-generated menstruation-themed images in a real-world use case.

Method

To answer the six research questions, a cross-sectional manual quantitative content analysis of 600 menstruation-related AI outputs prompted to be images for an educational brochure was conducted. The study follows open science practices: The generated AI outputs, codebook, data file, analysis script, and supplementary tables are publicly available at https://osf.io/h5pn3/.

Research Ethics

Generating and analyzing AI images for research purposes is an established practice in dozens of peer-reviewed studies (see Tables 1, 2, 3). Content generated with publicly available AI image generators is generally governed by provider terms of use and licensing that usually grant broad usage rights for research. This is why most studies in this field do not explicitly discuss research ethics or the involvement of ethics bodies. The current study received approval from the ethics committee of the authors' university.

Sample

To generate a sample of AI images, 15 AI image generators were selected (see Table 4), reflecting the range of broadly available AI tools at the time of data collection in September and October 2025, based on four criteria: (1) public accessibility (mostly via a free tier), (2) global prominence and coverage in earlier research (see Tables 1, 2, 3), (3) geographical breadth of the developer (USA, China, Europe), and (4) variation in underlying model architecture (diffusion versus autoregressive models).

To minimize account-specific bias, images were generated independently by two researchers at different locations in Germany, using newly created research accounts, as the majority of the selected image generators require registration for repeated use. Memory features were disabled in all accounts where available, and a new chat session was started for each prompt to prevent prompt carry-over across trials. Prompts were administered in the same order across all accounts and image generators. In line with our visual sexual health communication scenario, in which a sex educator uses AI to generate menstruation images for an educational brochure, we used default settings as a non-expert user would, forgoing API access, advanced prompt engineering, selection of generation parameters, specialized models, and complex workflows.

For ecological validity, five pretested text prompts in two languages (English and German) were applied, reflecting the communication goal (i.e., to create an image for an educational brochure) and the main aspects of menstruation,

Table 4 Fifteen selected AI image generators included in the study

Generative AI tool	Company (Country)	AI chatbot for text generation	AI image generation model name	AI Image generation model type	References
1. Adobe Firefly https://firefly.adobe.com/	Adobe (USA)	–	Firefly image model	Diffusion	Polner (2025)
2. ChatGPT-5 https://chatgpt.com/	OpenAI (USA)	+	GPT-Image-1 (formerly: DALL-E)	Autoregressive	OpenAI (2025)
3. Claude (Sonnet 4) https://claude.ai/	Anthropic(USA)	+	SVG (Scalable Vector Graphics)-based images only	Vector Graphics Rendering	Claude (2024)
4. Craiyon https://www.craiyon.com/	Craiyon (USA)	–	Craiyon (formerly: DALL-E mini)	Autoregressive	Craiyon (2026)
5. DeepSeek-V3.1 https://www.deepseek.com/	DeepSeek(China)	+	Janus Pro (not yet implemented at the time of the study)	Diffusion	DeepSeek AI (2025)
6. ERNIE 4.5 Turbo https://ernie.baidu.com/	Baidu(China)	+	ERNIE-4.5-VL (Visual Language)	Autoregressive	Baidu (2025)
7. Gemini (2.5 Flash) https://gemini.google.com/	Google (USA)	+	Gemini 2.5 Flash Image (nickname: Nano Banana)	Autoregressive	Fortin et al. (2025)
8. Grok https://grok.com/	xAI (USA)	+	Aurora	Autoregressive	xAI (2024)
9. Le Chat https://chat.mistral.ai/	Mistral AI(France)	+	Black Forest Labs—Flux	Diffusion	Mistral AI Team (2024)
10. Meta AI https://www.meta.ai/	Meta(USA)	+	Emu	Diffusion	Dai et al. (2023)
11. Microsoft Copilot / Bing https://copilot.microsoft.com/	Microsoft(USA)	+	Bing Image Generator (DALL-E 3, GPT-4o)	Diffusion (DALL-E 3) Autoregressive (GPT-4o)	Microsoft (2025)
12. Midjourney 7 https://www.midjourney.com/	Midjourney(USA)	–	Midjourney	Diffusion	Caleb (2025)
13. Perplexity https://www.perplexity.ai/	Perplexity AI (USA)	+	Default option automatically selects the best model for the query: GPT-Image-1, Nano Banana, Seedream 4.5	Autoregressive (GTP-Image-1; Nano Banana) Diffusion (Seedream 4.5)	Perplexity (2026)
14. Runway https://runwayml.com/	Runway AI(USA)	–	Gen-4	Diffusion	Runway (2026)
15. Stable Diffusion https://dreamstudio.stability.ai/	Stability AI(UK)	–	Stable Diffusion 3.5	Diffusion	Stability AI (2026)

Note. In alphabetical order of generative AI tool names. +: AI tool offers text generation; –: AI tool does not offer text generation. Information on image generators was sourced from the respective AI companies' product documentation with key references cited

namely menstruation per se as well as menstrual products, menstruators, gynecologists, and menstrual activists—concepts a professional sex educator might find relevant for educational materials (Bobel et al., 2020; Döring et al., 2024). The five prompts had the same structure: "Create an image of menstruation [of menstrual products; of a menstruating person; of a gynecologist caring for a menstruating person; of a menstrual activist] for an educational brochure." For each prompt, two images were generated.

The sampling plan resulted in 15 (AI image generators) × 2 (accounts) × 5 (prompt themes) × 2 (prompt languages) × 2 (images per prompt) = 600 AI outputs that are archived at https://osf.io/h5pn3/.

Codebook

To address the six research questions, we deductively and inductively developed a standardized codebook (Krippendorff, 2019; Neuendorf, 2017). It contains a block of seven formal categories (AI image generator, account, prompt theme, prompt language, image per prompt, image type, image ID) and six blocks of categories for the six research questions (the full codebook with coding instructions, example images, and references is available at https://osf.io/h5pn3/):

Table 5 Successful and unsuccessful AI image generation across menstruation-themed prompts

	Total ($N = 600$)		Menstruation ($n = 120$)		Menstrual products ($n = 120$)		Menstruating person ($n = 120$)		Gynecologist ($n = 120$)		Menstrual activist ($n = 120$)	
	N	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Image Output	480	80.0	88	73.3	104	86.7	96	80.0	88	73.3	104	86.7
Text Output	120	20.0	32	26.7	16	13.3	24	20.0	32	26.7	16	13.3
Total	600	100.0	120	100.0	120	100.0	120	100.0	120	100.0	120	100.0

Note. Based on $N = 600$ AI-generated outputs ($n = 120$ per thematic prompt). Overall analysis using a two-way chi-square test of independence (output type × prompt theme) $\chi^2(4) = 13.33$, $p < 0.01$ with Cramér's $V = 0.15$ as effect size measure

- For RQ1, the success of image generation was coded with one binary category (type of AI output: image versus text).
- For RQ2, the visual and textual AI image flaws defined in the literature (Amri & Hisan, 2023; Buzzaccarini et al., 2024; Campbell et al., 2026; Temsah et al., 2024; Westberg & Kvåle, 2024) were coded separately and then combined into one polytomous image flaw category with four values (no flaws, only visual flaws, only textual flaws, visual and textual flaws).
- For RQ3, following the state of research on AI images, the diversity of represented people was separately coded with four polytomous categories for gender (Ali et al., 2023; Reed et al., 2025), age (Gisselbaek et al., 2024), skin tone (Shaari et al., 2025), and body size (Mutale et al., 2016; Saumure et al., 2025; Wang & Wang, 2025).
- For RQ4, the gender-stereotypical representation of women in AI images was coded with six binary gender expression categories established in the literature: sexualized appearance (Kenig et al., 2023; Reed et al., 2025; Wang & Wang, 2025), smiling (Campbell et al., 2026; Gisselbaek et al., 2024), tilted head (Campbell et al., 2026; Sun et al., 2024), long hair (Campbell et al., 2026), dress/skirt (Wang & Wang, 2025), and glasses (Saumure et al., 2025).
- For RQ5, ten menstrual symbols were coded as binary categories based on the literature (Döring et al., 2024; Putland et al., 2023; Senft Everson et al., 2026; Wang & Wang, 2025): blood, red aesthetic, menstrual products, uterus/female reproductive tract, menstrual cycle, menstrual tracking, pain, non-medical pain relief, pharmaceutical pain relief, and resting posture.
- For RQ6, the overall valence of the menstruation portrayals in the AI images was coded in accordance with the literature (Döring et al., 2024; Varela-Rodríguez & Plage, 2025) with one polytomous category (positive, negative, ambivalent, neutral, unclear).

To pre-test the codebook, 60 AI images were selected from the sample and coded independently by two trained coders. Coders were trained using the study codebook, which defined all categories, values, and decision rules and provided visual examples for the coding procedure. The reliability coefficient Gwet's AC1 was computed for all categories and ranged from 0.74 to 1.00 ($M = 0.91$), indicating strong inter-coder reliability. Supplementary Table S1 with reliability coefficients for all coded categories is available at https://osf.io/h5pn3/.

Procedure and Analysis

For data collection, the sampled AI outputs were coded by a trained coder using the standardized and pretested codebook. Data analysis was conducted in R with the packages *tidyverse*, *DescTools*, stats, and *irrCAC,* and was focused on descriptive statistics (absolute and relative frequencies) and inferential statistics (two-way chi-square tests of independence; Fisher's exact tests) with the level of significance set at 5%. Both the data file with the codes and the R script for the statistical analysis are publicly available at https://osf.io/h5pn3/.

Units of statistical analysis differ between the six research questions: RQ1 uses all AI-generated outputs; RQ2 uses all AI-generated images; RQ3 uses all depicted persons in AI-generated images; RQ4 uses all depicted women in AI-generated images; RQ5 uses all AI-generated images; and RQ6 uses all AI-generated images. Our main focus of statistical comparison was the five menstrual themes addressed in the prompts. To detect a small-to-medium ($w = 0.20$) effect size with a statistical power of 80% at a significance level of 5% using a chi-square test with $df = 4$, an a priori G*Power analysis computed a minimum sample size of 299. As we could not foresee how many AI outputs would actually be images, we had decided in the study preregistration to double the target sample size and generate $N = 600$ AI outputs.

Results

Results are presented separately for the six research questions.

Table 6 Prevalence of visual and textual flaws in AI-generated images across menstruation-themed prompts

	Total ($N=480$)		Menstruation ($n=88$)		Menstrual products ($n=104$)		Menstruating person ($n=96$)		Gynecologist ($n=88$)		Menstrual activist ($n=104$)		$\chi^2(4)$	p	V
	N	%	n	%	n	%	n	%	n	%	n	%			
No flaws	105	21.9	23	26.1	12	11.5	18	18.8	30	34.1	22	21.2	15.70	0.003	0.18
Only visual flaws	257	53.5	28	31.8	74	71.2	59	61.5	45	51.1	51	49.0	33.14	< 0.001	0.26
Only textual flaws	42	8.8	10	11.4	1	1.0	8	8.3	8	9.1	15	14.4	12.88	0.012	0.16
Visual and textual flaws	76	15.8	27	30.7	17	16.3	11	11.5	5	5.7	16	15.4	22.78	< 0.001	0.22
Total	480	100.0	88	100.0	104	100.0	96	100.1	88	100.0	104	100.0			

Note. Based on $N=480$ AI-generated images and aggregated categories of visual and textual flaws, where an image was coded as flawed if at least one specific flaw type appeared. Row-wise analysis using two-way chi-square tests (flaw type × prompt theme) with Cramér's V as effect size measure. Of the 480 AI images, 205 (42.7%) had textual elements

Unsuccessful Image Generation (RQ1)

Of the 600 prompts, 120 (20.0%) did not successfully elicit the generation of menstruation-themed images, but only produced text output (see Table 5). Content blocking was observed in Adobe Firefly (for "menstruation," "menstruating person," and "gynecologist" when prompted in German or English), Microsoft Copilot (for "gynecologist" when prompted in German or English), and Perplexity (for "menstruation" when prompted in German or English). The blocking AI models told users that the generation of the requested images was not possible, sometimes referred to ethical or legal concerns and classified the menstrual themes as too "sensitive" or violating "safety policy."

Apart from content blocking, three of the 15 included AI models had issues with content creation unrelated to menstruation-related topics: ERNIE and Stable Diffusion could generally not operate on German prompts. DeepSeek was unable to generate any images, even though in early 2025 DeepSeek's image generation model Janus Pro had been announced (Reuters, 2025). At the time of data collection, Janus Pro had not been integrated into DeepSeek. Consequently, DeepSeek could not deliver images, but provided detailed menstruation-themed image descriptions and advised the user to provide these descriptions to a dedicated AI image generator or a professional illustrator.[17]

Objective Image Flaws (RQ2)

The 480 AI-generated menstruation-themed images covered three image types: illustration (46.2%), photorealistic image (40.4%), and diagram (13.3%; for a breakdown of image types across prompt themes see supplementary Table S2 at https://osf.io/h5pn3/). All diagrams included text elements, and some of the illustrations and photorealistic images also had textual elements such as headings or taglines. Overall, of the 480 AI images in our sample, 205 (42.7%) contained textual elements. The vast majority (78.1%) of the menstruation-themed AI images showed visual and/or textual flaws with visual flaws being more prevalent than textual flaws (see Table 6). The prevalence of image flaws varied significantly depending on image themes: for example, 88.5% of AI images of menstrual products had flaws in terms of showing non-existent products (e.g., tampon-shaped plastic objects; round pads) and/or misspelled text labels (e.g., "menstural") within the product images (see Fig. 2). At the same time, 11.5% of AI images of menstrual products were free of any visual and textual flaws.

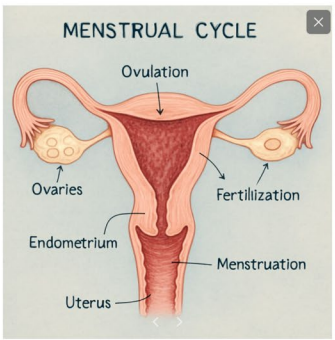
Of the AI images showing gynecologists caring for menstruating persons, 65.9% had visual and/or textual flaws such as non-sensical body features, objects or actions (e.g., a menstrual blood stain on an outer pant leg; a gynecologist carrying a handbag during consultation; a patient reading a restaurant menu during consultation; see Fig. 2). At the same time, 34.1% of the gynecologist-themed images had no visual or textual flaws.

Lack of Diversity of Human Representation (RQ3)

Of the 480 AI-generated menstruation-themed images in our sample, 317 images depicted a total of 399 persons such as menstruators, gynecologists, and menstrual activists; the remaining 163 images depicted no persons but only objects. The representations of individuals showed diversity and significant variation regarding gender, age, and skin tone (see Table 7). Menstruating persons, for example, were predominantly (84.0%), but not exclusively, depicted as girls and women; male (0.5%) and non-binary (5.8%) menstruators were also visible.

Regarding age, the menstruation-themed AI images represented menstruators quite often as young people, while menstrual activists and gynecologists were predominantly depicted as adults. Concerning skin tone, the majority of the 399 persons portrayed in the menstruation-themed AI

[17] We tested a subset of the image descriptions provided by DeepSeek with other AI image generators and obtained prompt-aligned images.

<table>
<tr>
<td>Textual flaws: Heading (“Menstrual Cycle” instead of female reproductive tract) and anatomical labels (e.g., “Uterus” points to vagina) are incorrect [generated with Craiyon]

</td>
<td>Visual and textual flaws: Non-existent menstrual products (e.g., round pads), and misspelled in-image labels (e.g., “menstural”) [generated with Gemini]

</td>
</tr>
<tr>
<td>Visual flaws: Menstrual blood stain shown on an outer pant leg [generated with ChatGPT]
</td>
<td>Visual flaws: Gynecologist with unattached handbag, patient with restaurant menu [generated with Meta AI]
</td>
</tr>
</table>

Fig. 2 Examples of AI-generated menstruation-themed images with visual and/or textual flaws (Source: author-created)

images had medium or darker skin tones; a minority was shown with lighter skin tones (45.6%). Menstrual activists in particular were portrayed predominantly non-White (see Table 7). A lack of diversity was observed regarding body size: the majority of depicted menstruators, gynecologists, and menstrual activists all showed average build; larger bodies were underrepresented (3.8%; see Fig. 3).

Stereotypical Portrayal of Women (RQ4)

Among the 317 menstruation-themed AI images depicting people in our sample, 278 images portrayed a total of 335 female persons. While the depiction of women in menstruation-themed AI images included gender stereotypical features such as expressions of femininity (long hair, dress/skirt, smiling, tilted head) and sexualized appearance, these features were not predominant (see Table 8). Gynecologists and menstruating persons were portrayed with gender-stereotypical features such as long hair at similarly high rates (57.7% and 56.6%, respectively), compared with depictions of menstrual activists (39.6%). Women menstruators (1.7%) and menstrual activists (2.2%) were seldom shown with glasses in comparison to gynecologists (14.1%; see Fig. 4).

Visual Symbols (RQ5)

In the menstruation-themed AI images, a set of conventional visual menstrual symbols was used: Menstruation was symbolized with red colors or blood (present in 55.2% of all 480 menstruation-themed images), menstrual products,

Table 7 Diversity of presented people in AI-generated images across different menstruation-themed prompts

	Total ($N=399$)		Menstruating person ($n=206$)		Gynecologist ($n=91$)		Menstrual activist ($n=102$)		χ^2	*df*	*p*	*V*
	N	%	*n*	%	*n*	%	*n*	%				
Gender[a]									–	–	< 0.001	0.23
Woman	335	84.0	173	84.0	71	78.0	91	89.2				
Man	13	3.3	1	0.5	11	12.1	1	1.0				
Non-binary person	21	5.3	12	5.8	8	8.8	1	1.0				
Unclear	30	7.5	20	9.7	1	1.1	9	8.8				
Age									70.62	4	< 0.001	0.30
Young person	119	29.8	86	41.7	3	3.3	30	29.4				
Adult	201	50.4	70	34.0	78	85.7	53	52.0				
Not codable	79	19.8	50	24.3	10	11.0	19	18.6				
Skin Tone[a]									–	–	< 0.001	0.19
Lighter skin tone	182	45.6	109	52.9	48	52.7	25	24.5				
Medium skin tone	125	31.3	58	28.2	27	29.7	40	39.2				
Darker skin tone	85	21.3	36	17.5	14	15.4	35	34.3				
Not codable	7	1.8	3	1.5	2	2.2	2	2.0				
Body Size[a]									–	–	0.131	0.11
Thin build	22	5.5	16	7.8	4	4.4	2	2.0				
Average build	298	74.7	149	72.3	75	82.4	74	72.5				
Plus-size build	15	3.8	9	4.4	1	1.1	5	4.9				
Not codable	64	16.0	32	15.5	11	12.1	21	20.6				

Note. Based on $N=317$ AI-generated images depicting 399 persons. Analysis using two-way chi-square tests of independence (socio-demographic characteristic × depicted person) with Cramér's V as effect size measure. [a]Due to low expected cell counts, Fisher's exact test was applied; hence, Cramér's V values based on potentially biased chi-square values must be interpreted with caution

resting posture, uterus or female reproductive tract, expressions of pain and measures of pain relief, and menstrual cycle or means of menstrual tracking such as a calendar. The occurrence of the respective symbols varied systematically between menstruation-themed prompts. For example, menstrual activists were mainly associated with red colors or symbols of blood, while menstruating persons were also associated with expressions of pain or measures of pain relief (see Table 9).

Image Valence (RQ6)

In light of broad stigmatization of menstruation in many social and cultural contexts, we were interested in the overall valence framing of the menstruation-themed AI images. Overall, the majority of the 480 AI images had a positive or neutral tone (78.0%; see Table 10). Positive images showed menstruation as natural and manageable with confident and happy facial expressions and bright color palettes (39.0%; see Fig. 5). Neutral images presented menstruation in a matter-of-fact, informational way such as with a photo of menstrual products, a diagram of the female reproductive tract or factual symbols (39.0%). Ambivalent images had both positive and negative tones, for example, when a menstruating person expressing abdominal pain was cared for by an empathetic, friendly gynecologist, representing both suffering and help (10.8%). Negatively framed images represented menstruation as something painful, shameful, or burdensome, for example, with expressions of pain in isolation (9.0%). Rarely were images so abstract that it was not possible to code any valence (2.3%).

Discussion

The discussion covers the interpretation of the findings, study limitations, and the conclusion.

Interpretation

The interpretation of results first addresses the answers to the research questions and then comments on differences between AI image generators.

Findings by Research Questions

Our analysis revealed clear obstacles in AI image generator use for SRHR communication: Some generators blocked

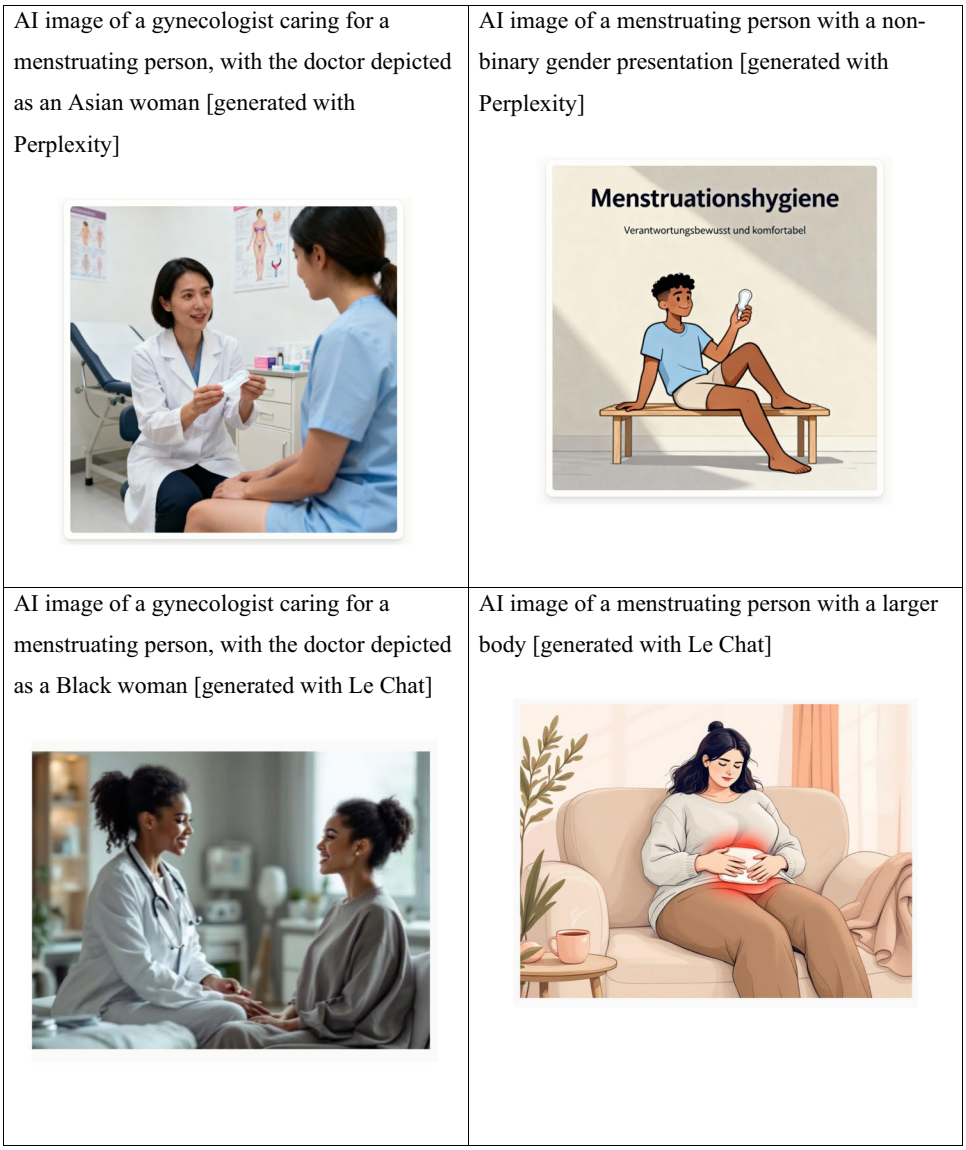


Fig. 3 Examples of AI-generated menstruation-themed images representing diverse genders, skin tones, and body sizes (Source: author-created)

Table 8 Stereotypical portrayals of women in AI-generated images across different menstruation-themed prompts

	Total ($N=335$)		Menstruating person ($n=173$)		Gynecologist ($n=71$)		Menstrual activist ($n=91$)		$\chi^2(2)$	p	V
	N	%	n	%	n	%	n	%			
Long hair	175	52.2	98	56.6	41	57.7	36	39.6	8.07	0.018	0.16
Smiling	110	32.8	40	23.1	39	54.9	31	34.1	23.18	< 0.001	0.26
Sexualized appearance	52	15.5	46	26.6	2	2.8	4	4.4	33.49	< 0.001	0.31
Dress/skirt	47	14.0	36	20.8	1	1.4	10	11.0	16.67	< 0.001	0.22
Tilted head	32	9.6	31	17.9	0	0.0	1	1.1	29.04	< 0.001	0.29
Glasses[a]	15	4.5	3	1.7	10	14.1	2	2.2	–	< 0.001	0.24

Note. Based on $N=278$ AI-generated Images depicting 335 female persons, presented in descending order of total percentages. As individual images may contain several coded categories, column percentages do not add up to 100%. Row-wise analysis using two-way chi-square tests of independence (stereotypical attribute × prompt theme) with Cramér's V as effect size measure. [a]Due to low expected cell counts, Fisher's exact test was applied; hence, Cramér's V value based on potentially biased chi-square value must be interpreted with caution

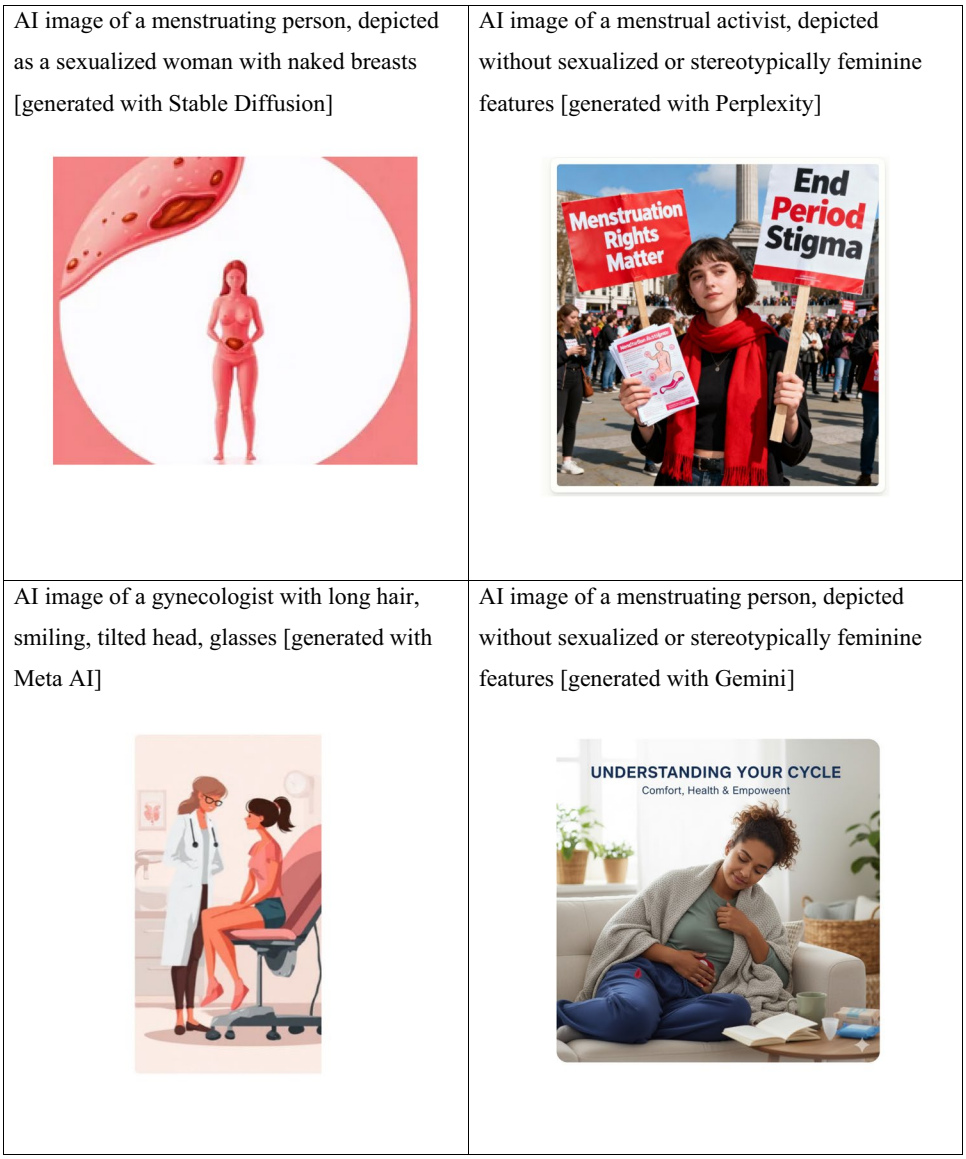


Fig. 4 Examples of AI-generated menstruation-themed images with different presentations of women (Source: author-created)

Table 9 Occurrence of visual menstruation symbols in AI-generated images across different menstruation-themed prompts

	Total ($N=480$)		Menstruation ($n=88$)		Menstrual products ($n=104$)		Menstruating person ($n=96$)		Gynecologist ($n=88$)		Menstrual activist ($n=104$)		$\chi^2(4)$	p	V
	N	%	N	%	n	%	n	%	n	%	n	%			
Red aesthetic/blood[a]	265	55.2	60	68.2	48	46.2	53	55.2	31	35.2	73	70.2	33.09	< 0.001	0.26
Menstrual products	198	41.2	24	27.3	87	83.7	34	35.4	20	22.7	33	31.7	101.95	< 0.001	0.46
Resting posture	159	33.1	9	10.2	1	1.0	61	63.5	82	93.2	6	5.8	287.90	< 0.001	0.77
Uterus/female reproductive tract	96	20.0	48	54.5	1	1.0	18	18.8	19	21.6	10	9.6	96.44	< 0.001	0.45
Expression of pain/pain relief[a]	93	19.4	11	12.5	5	4.8	50	52.1	26	29.5	1	1.0	110.94	< 0.001	0.48
Menstrual cycle/menstrual tracking[a]	45	9.4	25	28.4	1	1.0	17	17.7	2	2.3	0	0.0	70.02	< 0.001	0.38

Note. Based on $N=480$ AI-generated images, presented in descending order of total percentages. [a]Aggregated categories were calculated as present if at least one specific symbol appeared. Categories are not disjunctive, as individual images may contain multiple symbols simultaneously, column totals may exceed 100%. Row-wise analysis using two-way chi-square tests of independence (symbol type × prompt theme) with Cramér's V as effect size measure

Table 10 Valence of menstruation portrayals in AI-generated images across different menstruation-themed prompts

	Total		Menstruation		Menstrual products		Menstruating person		Gynecologist		Menstrual activist		$\chi^2(4)$	p	V
	N	%	n	%	N	%	n	%	n	%	n	%			
Positive	187	39.0	9	10.2	0	0.0	33	34.4	47	53.4	98	94.2	239.10	< 0.001	0.71
Neutral	187	39.0	63	71.6	103	99.0	8	8.3	9	10.2	4	3.8	319.59	< 0.001	0.82
Ambivalent	52	10.8	2	2.3	0	0.0	20	20.8	30	34.1	0	0.0	91.16	< 0.001	0.44
Negative	43	9.0	10	11.4	0	0.0	30	31.2	1	1.1	2	1.9	82.26	< 0.001	0.41
Unclear[a]	11	2.3	4	4.5	1	1.0	5	5.2	1	1.1	0	0.0	–	0.037	0.14
Total	480	100.1	88	100.0	104	100.0	96	99.9	88	99.9	104	99.9			

Note. Based on $N = 480$ AI-generated images, presented in descending order of total percentages. Overall test: A two-way chi-square test (Valence × Prompt): $\chi^2(16) = 506.43$, $p < 0.001$, $V = 0.51$ (Monte Carlo Fisher: 2000 replicates). Row-wise analysis using two-way chi-square tests of independence (valence type × prompt theme) with Cramér's V as effect size measure. [a]Due to low expected cell counts, Fisher's exact test was applied; hence, Cramér's V value based on potentially biased chi-square value must be interpreted with caution



Fig. 5 Examples of AI-generated menstruation-themed images with different valence framing (positive, neutral, ambivalent, negative) (Source: author-created)

menstruation-themed prompts (RQ1), and the vast majority of menstruation-themed AI images that were successfully created had visual and/or textual flaws (RQ2). These results are in line with previous research that critically discusses overblocking in AI tools (Döring et al., 2025; Vargas Penagos, 2024) and points to inaccuracies in AI-generated health images (Amri & Hisan, 2023; Buzzaccarini et al., 2024; Döring et al., 2025; Temsah et al., 2024). For our menstruation-themed use case, overblocking occurred much less frequently than inaccuracies. As a practical consequence, AI images must undergo critical scrutiny before being used in SRHR contexts. When image flaws are noticed, communicators have two options, (a) create new images with the same prompt hoping they are flawless or (b) revise the prompt and/or prompt the AI to make modifications in the existing AI image. At the current state of AI development, image generators differ significantly in their ability to align to prompts that require image modifications. For example, regarding the menstrual products image in Fig. 2, generated with Gemini, the prompt "correct in-image text: instead of 'menstural' write 'menstrual'" solved the issue immediately. However, at this point of development, most other AI image generators are not as efficient in making visual or textual corrections as Gemini.

Representational issues such as systematic under-representation of certain populations (e.g., non-White people and women in professional roles such as doctors; Ali et al., 2023; Campbell et al., 2026; Westberg & Kvåle, 2024) and presentational issues such as gender stereotypes (e.g., depicting women as sexy and smiling; Bianchi et al., 2023; Elhosary, 2026; Wang & Wang, 2025) that were described as widespread in previous research on AI health images, were less prevalent in our data. For example, people with darker skin tones were over- and not under-represented overall, and women doctors—unrealistically—made the vast majority of gynecologists in our menstruation-themed images (RQ3). Also, women were not predominantly shown with overly feminized or sexualized features (RQ4). We have two main explanations for our more balanced results: First, the loud public criticism of biases in AI images might have pushed several AI companies to adjust their models to provide more diverse and less stereotyped output. If this is true, current and future AI image studies should confirm the trend toward more diversity and less stereotyping. Second, our menstruation-themed prompts that are per se female coded and our use of inclusive language ("menstruating person") and the activism context ("menstrual activist") could have prompted the AI models to show more women in different roles. Also, the activism context with its link to historical liberation movements, could have elicited more ethnically/racially diverse representation. As a practical consequence, communicators are encouraged to avoid overly vague and broad prompts and instead use context-rich prompts and be intentional with diversity- and inclusivity-indicating concepts in the prompts.

One main criticism of public menstruation communication is an overly medical and negative framing to the point that menstruation appears as something aversive or even as an illness or handicap of women (Bobel et al., 2020; Döring et al., 2024). In our sample of menstruation-themed images, we identified typical visual symbols of menstruation such as blood and expressions of pain (RQ5). While one might criticize the use of traditional visual symbols as formulaic or stereotypical, recognizable symbols play an important role in visual communication as easily decodable signifiers. Avoiding such symbols can create the problem that menstruation-themed images are not identified as such and appear as irrelevant or non-aligned generic pictures (e.g., an image of an active woman engaged in work or sports can only provide a normalizing message regarding menstruation if visual and/or textual cues in the image clearly identify the woman as menstruating).

Overall, the AI-generated menstruation-themed images in our sample were predominantly positive or neutral, sometimes ambivalent, but relatively rarely showed negative framing (RQ6). We have two main explanations for this result. First, many contemporary AI image generators are strongly optimized for producing aesthetically pleasing, emotionally uplifting, and commercially safe visuals, which means that negative, painful, or stigmatizing imagery is often downplayed or filtered out. Second, AI models draw on a large volume of culturally circulating stock-photo imagery and promotional aesthetics, which tend to idealize health-related topics and favor smiling, attractive people and harmonious scenes. Thus, there could potentially even be a positive bias in AI-generated menstruation images, just as has been observed in AI-generated cancer patient images that are populated by smiling, healthy-looking, attractive women (e.g., Senft Everson et al., 2026; Varela-Rodriguez & Plage, 2025; Wang & Wang, 2025). As a practical consequence, communicators are reminded not to let the AI model decide on the overall message of the image, but to include in their prompt whether a particular menstruation-themed image is supposed to point to grievances (e.g., menstruators who suffer from very intense pain that is not taken seriously) or is meant to normalize (e.g., menstruators who experience light or no pain and who can participate in any activity they like during their period without being restricted to resting postures and pain relief measures used as menstrual symbols in several AI images, see Table 9).

Differences Between AI Image Generators

Our analysis focused on differences between the five menstruation-themed prompts, as prompt content is the central determinant of image output. By contrast, prompt language

played only a minor role: a few AI models—as noted above—rejected German-language prompts and required English input, but beyond this, language effects were negligible. We also observed no systematic influence of user accounts: Images generated with different accounts, on different computers, and in different regions of Germany were, overall, highly similar when using the same prompts with the same AI tools.

In contrast, substantial variation emerged between AI models. Although we could not statistically compare all 15 image generators in combination with all five prompts, several consistent tendencies were evident. Midjourney and Meta AI produced the most aesthetically polished images; however, their emphasis on stylization and artistic effects often rendered the outputs insufficiently serious or factual for SRHR communication. Claude generated scalable vector graphics (SVG) that were so visually disorganized and aesthetically unappealing that they are currently unsuitable for meaningful communication. Stable Diffusion, as an open-source model with comparatively limited safety regulation, generated the highest proportion of sexualized imagery, making many outputs inappropriate in SRHR contexts. ChatGPT and Microsoft Copilot produced only illustrations in a cartoon-adjacent flat style—a visual lingua franca of contemporary public communication—though often with a somewhat repetitive and lifeless quality. By contrast, Le Chat and Gemini showed greater stylistic range, producing illustrations in multiple styles, diagrams, and even photorealistic images, despite the absence of any prompt specifying image type.

These differences are particularly relevant in light of ongoing debates about the use of photorealistic AI images in SRHR communication. In 2023, Plan International's campaign against child marriage and, in 2024, a United Nations campaign addressing sexual violence in conflict both employed photorealistic AI-generated images of victims, clearly labeled as AI images. The intention was to avoid depicting real victims while still conveying the harms of violence and eliciting empathy. Both campaigns, however, were met with substantial criticism concerning authenticity and were swiftly withdrawn according to press reporting (Down, 2025). By contrast, the United Nations Development Programme's gender-equality initiative "Digital Imaginings: Women's Campaign for Equality" avoided photorealistic AI images but used AI art to visualize personal narratives collected from eight women across Europe and Central Asia, displaying the outputs in a virtual art exhibition.[18] Taken together, these examples underscore the need for SRHR communicators to exercise particular caution when using photorealistic AI depictions of people, given the ethical and credibility risks involved.

[18] https://www.undp.org/eurasia/GenderEquality-in-VR.

Limitations

AI image generators are developing fast, which means that our study can only be a snapshot in time. Identified issues such as overblocking or widespread visual and textual flaws could be reduced by technological advancements in the near future. While we covered 15 popular AI image generators, this was still not the full population of available tools. We prompted in English and German and identified only small language effects. However, future studies are needed to better understand if and how prompts in different languages affect image generation. Our study focused on five different menstruation-themed prompts, some of them fairly broad. To dig deeper into AI imagery, it would be helpful to explore more specific prompts, for example, instead of requesting images of "menstrual products," one could specifically request images of pads, of tampons, of menstrual cups, and of menstrual underwear to investigate alignment with specific content that is relevant in practical use cases (e.g., when creating an educational brochure on menstruation, one would like to introduce each type of menstrual product separately with its features). The same applies to using prompts that specify formal aspects of the images (e.g., photorealistic images versus illustrations versus diagrams).

Our study deliberately worked with a single textual prompt per image, and future studies could include image refinement through text and image prompts that require stepwise modifications and corrections of AI images. Future analyses could also aim at larger samples that would ensure sufficient statistical power for detailed statistical subgroup comparisons between AI models by prompt features. Our study comprised a quantitative content analysis; future studies could integrate systematic qualitative image analyses to explore visual frames and symbols used in AI images more deeply. One systematic issue that affects almost all content analyses of AI images (see Tables 1, 2, 3) is the fact that the analyzed images are created from only a limited number of AI tools and researcher accounts. Because multiple images were generated using the same AI models across two user accounts in this study, observations were not fully independent, which may have affected variance estimates and, consequently, the resulting test statistics, which should therefore be interpreted with caution. Generating images via multiple independent user accounts distributed across different world regions could increase variation in model outputs and thereby support more robust statistical analyses based on greater independence of AI outputs.

Conclusion

Our content analysis of menstruation-themed AI image outputs contributes to explainable AI by treating generated images as empirical indicators of what opaque text-to-image

systems assume, privilege, block, or distort in a sensitive SRHR context. The findings also inform responsible AI by identifying concrete communicative risks—such as overblocking, visual and textual flaws, limited body-size diversity, and stereotypical symbolism—that should be addressed when such tools are deployed in sexual health communication. Finally, the marked differences between generators and prompts underline the importance of AI literacy among sexual health professionals: using AI image generators well in SRHR communication requires not only access to the tools, but also the competence to choose them critically, prompt them strategically, evaluate AI outputs carefully, refine them iteratively, and select them responsibly.

Acknowledgements Does not apply.

Funding Open Access funding enabled and organized by Projekt DEAL. The study is part of the EMSA project (Sexual Debut, Menstruation, and Pregnancy Termination as Topics of Sexual and Reproductive Online Health Communication), which is funded by the Federal Institute for Public Health (BIÖG) in Germany from 2023 to 2026 and led by the first author.

Data Availability The study preregistration, instrument, data, analysis script, and supplementary tables are publicly available at https://osf.io/h5pn3/.

Declarations

Conflict of interest N.D. is one of the guest editors of the *Archives of Sexual Behavior* special issue "Artificial Intelligence and Sexuality." A.S. declares no conflict of interest.

Ethical Approval The study received ethical approval from the Ethical Committee of Technische Universität Ilmenau (ID 2025-10-211_FoA_Döring).

Informed Consent This study did not involve human participants, only AI-generated outputs.

Declaration of Generative AI in Scientific Writing Note: We included this AI declaration in response to the Guest Editors' request for the special section on "Artificial Intelligence and Sexuality." The first author wrote the initial draft of the paper, with all ideas originating from her and informed by prior research on AI-generated health images. ChatGPT-5 was used for language editing with prompts such as "proofread," "streamline," "simplify," or "shorten." The first author manually incorporated the AI-generated editing suggestions where appropriate. The revised manuscript was then shared with the second author, who provided additional corrections and contributions. The final version of the paper was approved by both authors without any further use of generative AI.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aktay, S. (2022). The usability of images generated by artificial intelligence (AI) in education. *International Technology and Education Journal, 6*(2), 51–62.

Alenichev, A., Kingori, P., & Grietens, K. P. (2023). Reflections before the storm: The AI reproduction of biased imagery in global health visuals. *The Lancet Global Health, 11*(10), e1496–e1498. https://doi.org/10.1016/s2214-109x(23)00329-7

Alenichev, A., Kingori, P., Shaffer, J., & Grietens, K. P. (2024). The elephant in the room: Reflecting on text-to-image generative AI and global health images. *BMJ Global Health, 9*(4), Article e015601. https://doi.org/10.1136/bmjgh-2024-015601

Ali, R., Tang, O. Y., Connolly, I. D., Abdulrazeq, H. F., Mirza, F. N., Lim, R. K., Johnston, B. R., Groff, M. W., Williamson, T., Svokos, K., Libby, T. J., Shin, J. H., Gokaslan, Z. L., Doberstein, C. E., Zou, J., & Asaad, W. F. (2023). Demographic representation in 3 leading artificial intelligence text-to-image generators. *JAMA Surgery, 159*(1), 87–95. https://doi.org/10.1001/jamasurg.2023.5695

Amri, M. M., & Hisan, U. K. (2023). Incorporating AI tools into medical education: Harnessing the benefits of ChatGPT and Dall-E. *Journal of Novel Engineering Science and Technology, 2*(02), 34–39. https://doi.org/10.56741/jnest.v2i02.315

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Baidu. (2025). *Announcing the open source release of the ERNIE 4.5 model family*. ERNIE Blog. https://ernie.baidu.com/blog/posts/ernie4.5/

Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2023). Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)* (pp. 1493–1504). Association for Computing Machinery. https://doi.org/10.1145/3593013.3594095

Bobel, C., Winkler, I. T., Fahs, B., Hasson, K. A., Kissling, E. A., & Roberts, T. (Eds.). (2020). *The Palgrave handbook of critical menstruation studies*. Springer Nature. https://doi.org/10.1007/978-981-15-0614-7

Brown, N., Forrest, L. J., Williams, R., Piasecki, J., & Bruinvels, G. (2024). 'Everyone needs to be educated': Pupils' voices on menstrual education. *Reproductive Health, 21*(1), 121. https://doi.org/10.1186/s12978-024-01862-6

Buzzaccarini, G., Degliuomini, R. S., Borin, M., Fidanza, A., Salmeri, N., Schiraldi, L., Di Summa, P. G., Vercesi, F., Vanni, V. S., Candiani, M., & Pagliardini, L. (2024). The promise and pitfalls of AI-generated anatomical images: Evaluating Midjourney for aesthetic surgery applications. *Aesthetic Plastic Surgery, 48*(9), 1874–1883. https://doi.org/10.1007/s00266-023-03826-w

Caleb. (2025). *V7 is now the default model!* Midjourney. https://updates.midjourney.com/v7-is-now-the-default-model/

Califano, G., & Spence, C. (2024). Assessing the visual appeal of real/AI-generated food images. *Food Quality and Preference, 116*, Article 105149. https://doi.org/10.1016/j.foodqual.2024.105149

Campbell, L. O., Lambie, G. W., Hayes, B. G., & Hartshorne, R. (2026). Generative artificial intelligence images for instruction:

A content analysis. *TechTrends, 70*, 80–92. https://doi.org/10.1007/s11528-025-01129-2

Cevik, J., Lim, B., Seth, I., Sofiadellis, F., Ross, R. J., Cuomo, R., & Rozen, W. M. (2023). Assessment of the bias of artificial intelligence generated images and large language models on their depiction of a surgeon. *ANZ Journal of Surgery, 94*(3), 287–294. https://doi.org/10.1111/ans.18792

Chan, J. (2024). AI-driven visuals: how they shape health and taste preferences in vegetarian dining. *Journal of Foodservice Business Research*. https://doi.org/10.1080/15378020.2024.2445886

Chan, J. (2025). AI-generated imagery in sustainable gastronomy tourism: A study from bottom-up to top-down processing. *Tourism Management, 108*, Article 105093. https://doi.org/10.1016/j.tourman.2024.105093

Chou, W. S., Gaysynsky, A., Everson, N. S., Muro, A., Schrader, K., & Iles, I. (2026). Portrayal of cancer patients in the era of AI: A content analysis of images produced by generative AI tools. *Health Communication, 41,* 888–898. https://doi.org/10.1080/10410236.2025.2537807

Claude. (2024). *Artifacts are now generally available*. https://claude.com/blog/artifacts

Craiyon. (2026). *Craiyon - Your free AI image generator tool: Create AI art!* https://www.craiyon.com/en

Dai, X., Hou, J., Ma, C.-Y., Tsai, S., Wang, J., Wang, R., Zhang, P., Vandenhende, S., Wang, X., Dubey, A., Yu, M., Kadian, A., Radenovic, F., Mahajan, D., Li, K., Zhao, Y., Petrovic, V., Singh, M. K., Motwani, S., ... Parikh, D. (2023). *Emu: Enhancing image generation models using photogenic needles in a haystack*. Meta AI. https://ai.meta.com/research/publications/emu-enhancing-image-generation-models-using-photogenic-needles-in-a-haystack/

DeepSeek AI. (2025). *Janus-Pro-7B* [Model card]. Hugging Face. https://huggingface.co/deepseek-ai/Janus-Pro-7B

Döring, N., Le, T. D., Vowels, L. M., Vowels, M. J., & Marcantonio, T. L. (2025). The impact of artificial intelligence on human sexuality: A five-year literature review 2020–2024. *Current Sexual Health Reports, 17*(1). https://doi.org/10.1007/s11930-024-00397-y

Döring, N., & Mohseni, M. R. (2025). Anti-AI bias toward couple images and couple counseling: Findings from two experiments. *Archives of Sexual Behavior*. https://doi.org/10.1007/s10508-025-03318-9

Döring, N., Shevtsova, A., & Schumann-Doermer, C. (2024). Der aktuelle Forschungsstand zur Menstruationsgesundheit: Überblick über 20 bio-psycho-soziale Aspekte [The current state of research on menstrual health: Overview of 20 bio-psycho-social aspects]. *Zeitschrift für Sexualforschung, 37*(3), 151–166. https://doi.org/10.1055/a-2343-4279

Down, A. (2025). AI-generated 'poverty porn' fake images being used by aid agencies. *The Guardian*. https://www.theguardian.com/global-development/2025/oct/20/ai-generated-poverty-porn-fake-images-being-used-by-aid-agencies

Elhosary, M. (2026). Pixels of prejudice: Decoding embedded biases in AI-generated news imagery and their implications for visual journalism—Toward an algorithmic-mediated visual framing. *Journalism & Mass Communication Quarterly, 103*, 430–461. https://doi.org/10.1177/10776990251360690

Ertman, B., Xia, B., Sloane, M., Hartvigsen, T., & Perrin, P. B. (2026). Disability portrayals in artificial intelligence text-to-image generation: Influence of context and the medicalization of disability. *Rehabilitation Psychology, 71*, 272–282. https://doi.org/10.1037/rep0000634

Evans, R. L., Harris, B., Onuegbu, C., & Griffiths, F. (2022). Systematic review of educational interventions to improve the menstrual health of young adolescent girls. *British Medical Journal Open, 12*(6), Article e057204. https://doi.org/10.1136/bmjopen-2021-057204

Fortin, A., Vernade, G., Kampf, K., & Reshi, A. (2025, August 26). *Introducing Gemini 2.5 Flash Image, our state-of-the-art image model*. Google Developers Blog. https://developers.googleblog.com/en/introducing-gemini-2-5-flash-image/

Freixa, P., Arolas, M. R., Codina, L., & Lopezosa, C. (2025). IA, fotografía de stock y bancos de imágenes: Sesgos de género y estereotipos. *Hipertext Net Revista Académica Sobre Documentación Digital Y Comunicación Interactiva, 30*, 41–78. https://doi.org/10.31009/hipertext.net.2025.i30.05

Frosh, P. (2001). Inside the image factory: Stock photography and cultural production. *Media Culture & Society, 23*(5), 625–646. https://doi.org/10.1177/016344301023005005

Frosh, P. (2002). Rhetorics of the overlooked. *Journal of Consumer Culture, 2*(2), 171–196. https://doi.org/10.1177/1469540502002002002

Gisselbaek, M., Minsart, L., Köselerli, E., Suppan, M., Meco, B. C., Seidel, L., Albert, A., Chang, O. L. B., Saxena, S., & Berger-Estilita, J. (2024). Beyond the stereotypes: Artificial intelligence image generation and diversity in anesthesiology. *Frontiers in Artificial Intelligence, 7*, Article 1462819. https://doi.org/10.3389/frai.2024.1462819

Hennegan, J., Winkler, I. T., Bobel, C., Keiser, D., Hampton, J., Larsson, G., Chandra-Mouli, V., Plesons, M., & Mahon, T. (2021). Menstrual health: A definition for policy, practice, and research. *Sexual and Reproductive Health Matters, 29*(1), 31–38. https://doi.org/10.1080/26410397.2021.1911618

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, & H. Lin (Eds.), *Advances in neural information processing systems* (Vol. 33). Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, BC, Canada. https://doi.org/10.48550/arXiv.2006.11239

Högemann, M., Betke, J., & Thomas, O. (2025). What you see is not what you get anymore: A mixed-methods approach on human perception of AI-generated images. *Frontiers in Artificial Intelligence, 18*. https://doi.org/10.3389/frai.2025.1707336

Houts, P. S., Doak, C. C., Doak, L. G., & Loscalzo, M. J. (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Education and Counseling, 61*(2), 173–190. https://doi.org/10.1016/j.pec.2005.05.004

Jääskeläinen, P., Sharma, N. K., Pallett, H., & Åsberg, C. (2025). Intersectional analysis of visual generative AI: The case of stable diffusion. *AI & Society, 40*(6), 4341–4362. https://doi.org/10.1007/s00146-025-02207-y

Jewitt, C. (1997). Images of men: Male sexuality in sexual health leaflets and posters for young people. *Sociological Research Online, 2*(2), 23–34. https://doi.org/10.5153/sro.64

Jintaganon, N., Osinga, C. J., Steijger, D., De Vugt, M., & Neal, D. (2025). Biases in an artificial intelligence Image-Generator's depictions of healthy aging and Alzheimer's. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5177622

Kenig, N., Echeverria, J. M., & Vives, A. M. (2023). Human beauty according to artificial intelligence. *Plastic & Reconstructive Surgery Global Open, 11*(7), Article e5153. https://doi.org/10.1097/gox.0000000000005153

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). Sage. https://doi.org/10.4135/9781071878781

Kumar, A., Burr, P., & Young, T. M. (2024). Using AI text-to-image generation to create novel illustrations for medical education: Current limitations as illustrated by hypothyroidism and Horner Syndrome. *JMIR Medical Education, 10*, Article e52155. https://doi.org/10.2196/52155

Li, Y. V. (2026). Imagination of humanity's future: representation and comparison of female cyborg images in generative AI paintings. *Feminist Media Studies, 26*, 911–935. https://doi.org/10.1080/14680777.2024.2434639

Liang, Y., He, J., Li, G., Li, P., Klimovskiy, A., Carolan, N., Sun, J., Pont-Tuset, J., Young, S., Yang, F., Ke, J., Dvijotham, K. D. J.,

Collins, K., Luo, Y., Li, Y., Kohlhoff, K. J., Ramachandran, D., & Navalpakkam, V. (2024). Rich human feedback for text-to-image generation. *arXiv*. https://doi.org/10.48550/arXiv.2312.10240

Mahdavi Goloujeh, A., Sullivan, A., & Peier, A. (2024). Is it AI or is it me? Understanding users' prompt journey with text-to-image generative AI tools. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Article 183, pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642861

Martens, L., Virgin, N., Hoffarth, P., Goodwill, J., Derenne, N., Van Eck, R., Klug, M., Manocha, G. D., & Jurivich, D. (2025). Recapitulation of ageism in artificial intelligence-generated images: Longitudinal comparative study. *Journal of Medical Internet Research, 27*, Article e68428. https://doi.org/10.2196/68428

Martino, A. S., Miller, M., Moumos, E., Trung, R., & White, K. (2025). Reproducing or challenging dominant constructions of sexuality?: An exploratory study of AI-generated images representing disability and sexuality. *Canadian Journal of Human Sexuality, 34*(1), 83–95. https://doi.org/10.3138/cjhs-2024-0018

McCarthy, O., Carswell, K., Murray, E., Free, C., Stevenson, F., & Bailey, J. V. (2012). What young people want from a sexual health website: Design and development of SexUnzipped. *Journal of Medical Internet Research, 14*(5), Article e127. https://doi.org/10.2196/jmir.2116

Microsoft. (2025). *Bing Image Creator adds GPT-4o alongside DALL-E3 for even more free ways to create*. Microsoft Bing Blogs. https://blogs.bing.com/search/August-2025/Bing-Image-Creator-gets-GPT-4o

Mistral AI Team. (2024). *Mistral has entered the chat*. Mistral AI. https://mistral.ai/news/mistral-chat

Mutale, G. J., Dunn, A., Stiller, J., & Larkin, L. F. (2016). Development of a body dissatisfaction scale assessment tool. *New School Psychology Bulletin, 13*(2), 47–57.

Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd ed.). Sage. https://doi.org/10.4135/9781071802878

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Office of the High Commissioner for Human Rights. (2022). *High Commissioner for Human Rights statement on menstrual health*. OHCHR. https://www.ohchr.org/en/statements-and-speeches/2022/06/high-commissioner-human-rights-statement-menstrual-health

Olson, M. M., Alhelou, N., Kavattur, P. S., Rountree, L., & Winkler, I. T. (2022). The persistent power of stigma: A critical review of policy initiatives to break the menstrual silence and advance menstrual literacy. *PLOS Global Public Health, 2*(7), Article e0000070. https://doi.org/10.1371/journal.pgph.0000070

Olsson, S. E., Friedlich, N., Bolds, A., Hodge, A., Jamali, T., & Lagomichos, M. (2024). Exploring menstrual education via TikTok: A descriptive content analysis. *Journal of Adolescent Health, 75*(2), 353–357. https://doi.org/10.1016/j.jadohealth.2024.04.020

OpenAI. (2025). *Introducing our latest image generation model in the API*. https://openai.com/index/image-generation-api/

Oppenlaender, J., Linder, R., & Silvennoinen, J. (2025). Prompting AI art: An investigation into the creative skill of prompt engineering. *International Journal of Human-Computer Interaction, 41*(16), 10207–10229. https://doi.org/10.1080/10447318.2024.2431761

Organizers of Queer in AI, Ovalle, A., Subramonian, A., Singh, A., Voelcker, C., Sutherland, D. J., Locatelli, D., Breznik, E., Klubicka, F., Yuan, H., J., H., Zhang, H., Shriram, J., Lehman, K., Soldaini, L., Sap, M., Deisenroth, M. P., Pacheco, M. L., Ryskina, M., … Stark, L. (2023). *Queer in AI: A case study in community-led participatory AI*. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)* (pp. 1882–1895). Association for Computing Machinery. https://doi.org/10.1145/3593013.3594134

Perplexity. (2026). *Generating images with Perplexity*. Retrieved April 9, 2026, from https://www.perplexity.ai/help-center/en/articles/10354781-generating-images-with-perplexity

Polner, M. (2025). *Explore the new Adobe Firefly, your all-in-one creative AI studio.* Adobe Blog. https://blog.adobe.com/en/publish/2025/10/28/explore-new-adobe-firefly-your-all-in-one-home-ai-powered-creativity

Putland, E., Chikodzore-Paterson, C., & Brookes, G. (2023). Artificial intelligence and visual discourse: A multimodal critical discourse analysis of AI-generated images of "dementia." *Social Semiotics, 35*(2), 228–253. https://doi.org/10.1080/10350330.2023.2290555

Reed, J., Dodson, T., Petrinec, A., Tennant, D., Chmelik, J., & Cripple, S. (2025). Artificial intelligence and images portraying nurses through the decades. *OJIN: The Online Journal of Issues in Nursing*. https://doi.org/10.3912/ojin.vol30no02man05

Reuters. (2025, January 27). *DeepSeek says its Janus Pro AI model beats rivals in image generation*. https://www.reuters.com/technology/deepseeks-janus-pro-ai-model-beats-rivals-image-generation-2025-01-27/

Rodriguez, L., & Dimitrova, D. V. (2011). The levels of visual framing. *Journal of Visual Literacy, 30*(1), 48–65. https://doi.org/10.1080/23796529.2011.11674684

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022)* (pp. 1–45). https://doi.org/10.48550/arXiv.2112.10752

Runway. (2026). *Introducing Runway Gen-4*. Retrieved April 9, 2026, from https://runwayml.com/research/introducing-runway-gen-4

Saumure, R., De Freitas, J., & Puntoni, S. (2025). Humor as a window into generative AI bias. *Scientific Reports, 15*(1), Article 1326. https://doi.org/10.1038/s41598-024-83384-6

Schuhmann, C., Beaumont, R., Vencu, R., Gordon, C., Wightman, R., Cherti, M., Coombes, T., Katta, A., Mullis, C., Wortsman, M., Schramowski, P., Kundurthy, S., Crowson, K., Schmidt, L., Kaczmarczyk, R., & Jitsev, J. (2022). *LAION-5B: An open large-scale dataset for training next-generation multimodal models*. In *Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS 2022): Datasets and Benchmarks Track*. https://doi.org/10.48550/arXiv.2210.08402

Senft Everson, N., Gaysynsky, A., Iles, I. A., Schrader, K. E., & Chou, W. S. (2026). What does an AI-generated "cancer survivor" look like? An analysis of images generated by text-to-image tools. *Journal of Cancer Survivorship, 20*, 1612–1621. https://doi.org/10.1007/s11764-025-01760-1

Shaari, A. L., Saad, A. M., Patel, A. M., & Filimonov, A. (2025). Representation of demographics in otolaryngology by artificial intelligence text-to-image platforms. *Laryngoscope Investigative Otolaryngology, 10*(3), Article e70152. https://doi.org/10.1002/lio2.70152

Sobey, A. (2025). The thinness of GenAI: Body size in relation to the construction of the normate through GenAI image models. *AI and Ethics, 5*(4), 4181–4196. https://doi.org/10.1007/s43681-025-00684-x

Stability AI. (2026). *Introducing Stable Diffusion 3.5*. Retrieved April 9, 2026, from https://stability.ai/news/introducing-stable-diffusion-3-5

Starrs, A. M., Ezeh, A. C., Barker, G., Basu, A., Bertrand, J. T., Blum, R., Coll-Seck, A. M., Grover, A., Laski, L., Roa, M., Sathar, Z. A., Say, L., Serour, G. I., Singh, S., Stenberg, K., Temmerman, M., Biddlecom, A., Popinchalk, A., Summers, C., & Ashford, L. S. (2018). Accelerate progress—sexual and reproductive health and rights for all: Report of the Guttmacher–Lancet Commission. *The*

Lancet, 391(10140), 2642–2692. https://doi.org/10.1016/s0140-6736(18)30293-9

Sun, L., Wei, M., Sun, Y., Suh, Y. J., Shen, L., & Yang, S. (2024). Smiling women pitching down: Auditing representational and presentational gender biases in image-generative AI. *Journal of Computer-Mediated Communication, 29*(1), Article zmad045. https://doi.org/10.1093/jcmc/zmad045

Temsah, M., Alhuzaimi, A. N., Almansour, M., Aljamaan, F., Alhasan, K., Batarfi, M. A., Altamimi, I., Alharbi, A., Alsuhaibani, A. A., Alwakeel, L., Alzahrani, A. A., Alsulaim, K. B., Jamal, A., Khayat, A., Alghamdi, M. H., Halwani, R., Khan, M. K., Al-Eyadhy, A., & Nazer, R. (2024). Art or artifact: Evaluating the accuracy, appeal, and educational value of AI-generated imagery in DALL·E 3 for illustrating congenital heart diseases. *Journal of Medical Systems, 48*(1), 54. https://doi.org/10.1007/s10916-024-02072-0

Thelander, Å. (2018). Visual communication. In R. L. Heath & W. Johansen (Eds.), *The international encyclopedia of strategic communication*. Wiley. https://doi.org/10.1002/9781119010722.iesc0198

Varela-Rodríguez, M., & Plage, S. (2025). Visualizing cancer and survivorship with generative AI?—An exploration of breast, prostate, and pancreatic cancer imagery. *Journal of Cancer Survivorship*. https://doi.org/10.1007/s11764-025-01843-z

Vargas Penagos, E. (2024). ChatGPT, can you solve the content moderation dilemma? *International Journal of Law and Information Technology, 32*, Article eaae028. https://doi.org/10.1093/ijlit/eaae028

Velásquez-Salamanca, D., Martín-Pascual, M. Á., & Andreu-Sánchez, C. (2025). Interpretation of AI-generated vs. human-made images. *Journal of Imaging, 11*(7), Article 227. https://doi.org/10.3390/jimaging11070227

Wang, T., & Wang, R. (2025). What AI knows about breast cancer: an exploratory visual discourse analysis on AI-generated breast cancer advertisement images. *Social Semiotics*. https://doi.org/10.1080/10350330.2025.2519471

Westberg, G., & Kvåle, G. (2024). The generic uniqueness of AI imagery: A critical approach to Dall-E as semiotic technology. *Discourse & Society, 36*(4), 575–598. https://doi.org/10.1177/09579265241274788

WHO (World Health Organization). (2022, June 22). *WHO statement on menstrual health and rights*. https://www.who.int/news/item/22-06-2022-who-statement-on-menstrual-health-and-rights

WHO (World Health Organization). (2024, March 21). *The role of artificial intelligence in sexual and reproductive health and rights: Technical brief.* https://www.who.int/publications/i/item/9789240090705

xAI. (2024). *Grok image generation release*. https://x.ai/news/grok-image-generation-release

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.