Internet Interventions 38 (2024) 100775

Contents lists available at ScienceDirect

Internet Interventions

journal homepage: www.elsevier.com/locate/invent

Human centered design of AI-powered Digital Therapeutics for stress prevention: Perspectives from multi-stakeholders' workshops about the SHIVA solution

Marco Bolpagni [a,b,*], Susanna Pardini [b], Silvia Gabrielli [b]

[a] Human Inspired Technology Research Centre, University of Padova, Padova, Italy
[b] Digital Health Research, Center for Digital Health and Wellbeing, Fondazione Bruno Kessler, Trento, Italy

ARTICLE INFO

Keywords:
Digital Therapeutics (DTx)
Human-centered design
Affective computing
Stress prevention
Stakeholders engagement
Virtual reality
Mental well-being

ABSTRACT

Background: AI-powered Digital Therapeutics (DTx) hold potential for enhancing stress prevention by promoting the scalability of P5 Medicine, which may offer users coping skills and improved self-management of mental wellbeing. However, adoption rates remain low, often due to insufficient user and stakeholder involvement during the design phases.
Objective: This study explores the human-centered design potentials of SHIVA, a DTx integrating virtual reality and AI with the SelfHelp+ intervention, aiming to understand stakeholder views and expectations that could influence its adoption.
Methods: Using the SHIVA example, we detail design opportunities involving AI techniques for stress prevention across modeling, personalization, monitoring, and simulation dimensions. Workshops with 12 stakeholders—including target users, digital health designers, and mental health experts—addressed four key adoption aspects through peer interviews: AI data processing, wearable device roles, deployment scenarios, and the model's transparency, explainability, and accuracy.
Results: Stakeholders perceived AI-based data processing as beneficial for personalized treatment in a secure, privacy-preserving environment. While wearables were deemed essential, concerns about compulsory use and VR headset costs were noted. Initial human facilitation was favored to enhance engagement and prevent dropouts. Transparency, explainability, and accuracy were highlighted as crucial for the stress detection model.
Conclusion: Stakeholders recognized AI-driven opportunities as crucial for SHIVA's adoption, facilitating personalized solutions tailored to user needs. Nonetheless, challenges persist in developing a transparent, explainable, and accurate stress detection model to ensure user engagement, adherence, and trust.

1. Introduction

Digital therapeutics (DTx) are defined, in the European Union, as evidence-based therapeutic interventions driven by software to prevent, manage, or treat a medical disorder or disease (Recchia et al., 2020; Refolo et al., 2022). They are patient-facing software applications (e.g., mHealth apps, web apps, chatbots or Virtual Reality environments) that help patients treat, prevent, or manage a condition and have a proven clinical benefit (Refolo et al., 2022). DTx for mental health and well-being have raised particular interest in recent years, with the Food and Drug Administration (FDA) authorizing their prescription for pediatric attention deficit/hyperactivity disorder (Food, U. and (FDA), D. A, 2020), major depressive disorder, and generalized anxiety disorder (MobiHealthNews, 2021). Most DTx for mental health integrates cognitive behavioral therapy (CBT) (Apolinàrio-Hagen et al., 2020) with psychological and psychiatric care, enhancing the scale-up of personalized and comprehensive P5 Medicine (medicine that is personalized, participatory, predictive, preventive and palliative). Recent advances in the development of digital health solutions and artificial intelligence (AI) provide unprecedented opportunities for deploying mental health prevention strategies and programs in different domains, such as stress prevention, with the support of DTx (Ponzo et al., 2020; Roy et al., 2021; McGinnis et al., 2022; Dreyer et al., 2022). AI-powered DTx are commonly reviewed by domain experts before they can be implemented

* Corresponding author at: Human Inspired Technology Research Centre, University of Padova, Padova, Italy.
E-mail addresses: mbolpagni@fbk.eu (M. Bolpagni), spardini@fbk.eu (S. Pardini), sgabrielli@fbk.eu (S. Gabrielli).

https://doi.org/10.1016/j.invent.2024.100775
Received 17 April 2024; Received in revised form 21 August 2024; Accepted 10 September 2024
Available online 14 September 2024
2214-7829/

in a treatment plan. This review process is an example of augmented intelligence or intelligence amplification, wherein AI technology informs and augments rather than replaces healthcare professionals' experience and cognition (Kellogg and Sadeh-Sharvit, 2022; Bazoukis et al., 2022). DTx are considered as rather complex solutions to design (Craig et al., 2008) as they can include multiple components and AI algorithms that adapt provision of support to each person's changing needs. The goal of including these components in DTx solutions is that they can simultaneously provide safe, effective, accessible, sustainable, scalable, and equitable support for individual and population health (Dang et al., 2020). However, accomplishing an effective integration of these components in a DTx presents some challenges. Human-centered design approaches involving all key stakeholders should be adopted in the design and pre-clinical research phases of DTx (Czajkowski et al., 2015) to ensure future adoption and effectiveness of these solutions, overcoming possible conflicts of perspectives among different stakeholders. This paper contributes insights for the human-centered design of a DTx for stress prevention by harnessing the potential of AI techniques in supporting the tailoring of the intervention to the user needs and in providing key support to designers and care providers in enhancing the effectiveness, as well as the ethics soundness, of the DTx treatment. We first present a literature review and design opportunities for AI-powered DTx by using the example of SHIVA (SelfHelp+ Intervention with Virtual reality and Artificial intelligence), a stress prevention DTx based on the SelfHelp+ intervention developed by World Health Organization (WHO) (World Health Organization, 2021), targeting adult users with low or medium stress conditions. We then present results from three multi-stakeholders workshops aimed at eliciting the different views and expectations on stress prevention AI-powered DTx of target users representatives (URs), mental health experts (MHEs) and digital health designers (DHEs), to better understand the possible facilitators and barriers affecting the future adoption of these solutions.

1.1. AI-powered DTx solutions

DTx offer considerable potential across various medical and psychological conditions. This section examines novel AI-powered DTx solutions, emphasizing the role of advanced technologies in enhancing their effectiveness and accessibility. For instance, (Burger et al., 2022) developed an AI-based conversational agent named Luca using the Rasa platform for supporting cognitive therapy. Luca helped participants record their distressing situations, emotions, automatic thoughts, and behaviors and provided supportive feedback with different textual richnesses (low, medium and high). Participants found the task insightful and appreciated the immediate feedback, although some challenges, like difficulties in learning specific techniques suggested by the chatbot, required further human support. Nonetheless, the hypothesized effect of increased feedback on motivation and insight was not supported. Similarly, (Wang et al., 2023) introduced STEF, an AI-driven solution for mental health support relying on empathetic emotional conversations. In order to manage conversations authors, implemented an emotional fusion mechanism and a strategy tendency encoder to make the DTx deliver empathetic and supportive responses. Human evaluations indicated high acceptability, with volunteers rating the STEF agent's responses favorably compared to other models, showcasing its potential in digital mental health interventions. (Egan et al., 2024) instead targets young people with perfectionism by integrating AI with traditional Cognitive Behavioral Therapy for Perfectionism (CBT-P). Utilizing tools like ChatGPT, participants received real-time guidance and support to manage perfectionistic tendencies. The study highlighted the benefits of AI guidance, such as accessibility, non-judgmental support, and reduced stigma, particularly for those with social anxiety. Despite some concerns about AI's limitations, such as the lack of specificity, potential inaccuracies, and reduced empathy compared to humans, the intervention was well-received, indicating its potential to improve mental health support and therapy adherence. (Jeong et al., 2024) developed a gamified deep breathing mobile application for mental health management. The app used affordance-based design and gamification principles to engage users in deep breathing exercises, with machine learning (ML) providing real-time biofeedback. Participants were guided through the breathing process with interactive visual elements, and the system adapted to individual breathing patterns to provide tailored recommendations. Results showed significant improvements in stress indicators and user engagement in the intervention group, highlighting the potential of AI-enhanced DTx to improve mental health outcomes. Lastly, (Jeong et al., 2023) evaluated an AI-powered virtual agent in a DTx mobile application for depression. The agent, inspired by a Welsh Corgi, guided users through CBT and mindfulness meditation using advanced natural language processing techniques for personalized feedback. Conducted over the course of a month with 70 participants, the study found significant reductions in depression levels. Key factors for sustained DTx use included perceived usefulness, user attitudes, and enjoyment. The virtual agent's anthropomorphism and likability were crucial in enhancing user engagement and therapeutic outcomes. As these examples illustrate, AI-powered DTx solutions show great promise across a variety of mental health conditions, offering innovative approaches to therapy and support. However, to further enhance their effectiveness, particularly in stress prevention, it is crucial to address specific design challenges and opportunities.

1.2. Design opportunities in DTx for stress prevention

This section explores key issues and opportunities in developing AI-powered DTx for stress prevention, using emerging literature to highlight design challenges and establish a framework for our workshop's investigation. Due to the absence of a widely recognized framework, we identified recurring themes in the literature and drew insights from several perspective papers discussing AI integration in DTx, including (Palanica et al., 2020; Lee et al., 2023a; Vasdev, 2024; Hu et al., 2024). These papers were selected for their comprehensive analysis of AI applications in DTx, providing a foundation for identifying critical concerns and potential benefits. The key design challenges and opportunities—such as modeling, personalization, monitoring, simulation, and ethical considerations—are summarized in Table 1, which includes a description of each area of interest and examples of potential applications in DTx for stress management.

1.3. The SHIVA intervention

The SHIVA intervention is a DTx for stress prevention currently under development at our research center. SHIVA will serve as a case study to ground our discussion about design opportunities for AI-powered DTx for stress prevention. SHIVA delivers a digital version of the SelfHelp+ intervention (Doing What Matters in Times of Stress) (World Health Organization, 2020; Epping-Jordan et al., 2016; Tol et al., 2020) developed by WHO. The SelfHelp+ intervention is based on the Acceptance and Commitment Therapy (ACT) (Hayes et al., 2011), a form of CBT developed within a theoretical and a philosophical framework that is coherent and evidence-based. It is based on experiential strategies of acceptance and mindfulness, along with a commitment to action and behavior change, to increase psychological flexibility (Hayes and Pierson, 2005). In SHIVA, this intervention is delivered using a serious game environment. A serious game is a game designed for a primary purpose other than pure entertainment; in this case, it aims to provide virtual coaching on stress coping skills to adults with low or moderate levels of stress. When entering the serious game, users can choose a female or male character to receive coaching on stress prevention by means of a series of psycho-educational dialogs, videos and guided exercises (Fig. 1).

The serious game consists of 6 levels or coaching sessions to be

Table 1
Summary of design opportunities in DTx for stress prevention: key areas and example applications.

Area	Description	Example application
Modeling	Digital Twins (DTs) in mental health enables the synthesis of real-time data from wearable sensors and electronic health records, providing a dynamic reflection of psychological and emotional states (Spitzer et al., 2023). This modeling facilitates the identification of behavioral patterns and stress triggers, allowing for predictive analytics and adaptive intervention strategies that enhance therapeutic outcomes.	Developing dynamic models that adapt in real-time to individual stress responses allows for predicting when someone might experience high stress. These models can then inform personalized interventions, like adjusting workloads or suggesting relaxation exercises, based on their predictions.
Monitoring	Continuous monitoring allows for the analysis of both objective physiological data (e.g., heart rate variability, sleep patterns) and subjective feedback (e.g., Ecological Momentary Assessments (Shiffman et al., 2008)). This multiple layer provides critical insights into user engagement, treatment adherence, and the overall effectiveness of interventions, enabling iterative improvements in therapeutic design (Garcia-Ceja et al., 2018).	Evaluating engagement and intervention effectiveness by analyzing usage patterns and physiological data to customize interventions.
Personalization	Advancements of AI-driven personalization leverage predictive behavioral models to dynamically tailor therapeutic interventions to individual user profiles (Shah, 2022). Reinforcement learning and other ML techniques enable real-time optimization of interventions, thus improving engagement and minimizing dropout rates by aligning treatment with user-specific routines and coping mechanisms.	Tailoring interventions based on user behavior, such as delivering mindfulness exercises during commutes or providing coping strategies during high-stress periods.
Simulation	Generative AI facilitates the creation of synthetic datasets that simulate complex life scenarios, overcoming data labeling challenges in AI training (Arora and Arora, 2022). These simulations enable the development of personalized therapeutic models that predict and respond to various stressors, enhancing the adaptability and effectiveness of DTx interventions.	Utilizing synthetic profiles generated through simulation to anticipate varied stress manifestations and behavioral responses, thereby enabling the preemptive adaptation of therapeutic strategies and the projection of well-being trajectories.
Ethics	The ethical deployment of AI-powered DTx necessitates a commitment to transparency, accountability, and inclusivity. This includes addressing algorithmic biases, ensuring equitable access to therapeutic interventions, and maintaining stringent privacy (e.g. GDPR (General Data Protection Regulation (GDPR), 2018)) and data security standards.	Ensuring inclusivity by accommodating users from different socio-economic backgrounds, possibly through device subsidies or compatibility with affordable options, while adhering to strict privacy and security standards.

unlocked one per week (each session lasts between 20 and 30 min) covering the original SelfHelp+ sections providing coaching on grounding, unhooking, acting on one's values, being kind and making room, with an additional level of coaching on diaphragmatic breathing technique for relaxation. The serious game environment allows patients to learn through educational content, interacting with an avatar (e.g., answering questions), and performing the exercises. During these actions, they can earn points and badges, which serve as incentives and motivation to engage with the material. The next week's session unlocks only if the user has viewed the material from the current week. However, there is no minimum number of interactions required with the exercises to progress to the next week. The guided relaxation session is also designed to be performed by the user with the support of an immersive virtual reality setting (by wearing the Oculus Quest 2 headset) augmented with biofeedback. During the 6 weeks of the intervention the participant is invited to keep exercising for strengthening stress coping and relaxation skills. Additionally, they are provided with a wearable device for stress detection for modeling the user mental condition, with the aim of personalizing the coaching intervention to the user needs, as well as to better assess the effectiveness of the delivered treatment. The stress detection model running on a smartwatch, is designed with the potential future scale-up of the intervention in mind, and exclusively relies on movement and Blood Volume Pulse (BVP) as raw data. These types of data are readily available on most consumer devices in the market and provide enough information to achieve performance comparable to more complex multimodal settings (for instance, those also incorporating electrodermal activity data) (Can et al., 2019). In essence, this model is capable of assessing stress levels by analyzing patterns of cardiac activity, which is reflective of the autonomic nervous system's function. The SHIVA intervention has been designed to allow a self-guided, thus more scalable delivery of evidence based DTx for stress prevention. It can help mental health providers to implement stress prevention strategies in a stepped-care approach, optimizing the available resources for coping with the increasing demand for mental health services. In the context of our ongoing design work, addressing phase 1a (Define) and 1b (Refine) of the ORBIT model (Czajkowski et al., 2015), SHIVA serves also as a research test bed for investigating novel AI-powered functionalities for the delivery of DTx for stress prevention that promise to enhance our understanding of user engagement, adoption and adherence to this type of interventions, for their future human-centered refinement and improvement.

2. Materials and methods

2.1. Workshop objective

Possible reasons for a poor design and subsequent low adoption of DTx for mental well-being is a lack of human-centered design of the solution, as well as lack of support from one or more key categories of stakeholders and future adopters, such as MHEs, DHEs and URs. The expectations and possible concerns of different stakeholders are not always considered or aligned, and in some cases, there are conflicts among their views. The aim of the workshop investigation was to identify stakeholders' perceptions of future AI-powered DTx for stress prevention, using the example of SHIVA prototype to inform the future design, implementation and pre-clinical validation phases of our work. At the same time, the workshop is an opportunity to explore the DTx-related terminology used by various stakeholders and to encourage a more conscious use of it, as suggested by (Smoktunowicz et al., 2020).

2.2. Setting

Three stakeholders workshops were conducted online on the Google Meet platform in March 2024. Each workshop involved 4 participants and at least one representative from each of the three stakeholders categories identified. Prior to the workshop attendance participants

Fig. 1. Example of coaching dialogue delivered in the SHIVA serious game.

received a short document describing the SHIVA prototype, a link to the serious game environment and a video describing the main components of the SHIVA solution.

2.3. Participants

Participants were selected through personal connections and snowballing (e.g. asking participants whether they know another relevant person). A total of 12 participants with different backgrounds attended the workshop, including people representing the views from URs (university students), MHEs with experience in providing psychological support for stress prevention and management, DHEs with expertise in AI techniques (Table 2). All participants provided verbal permission for the information they gave during the workshop to be used for research purposes. They did not receive payments for their participation in the study.

Table 2
Workshop participants' background and position.

ID	Gender	Stakeholder group	Role/position	Organization type
P1	Male	DHE	Lead Computer Scientist	Research Center
P2	Female	DHE	Senior Technologist	Research Center
P3	Female	DHE	Junior Project Manager	University
P4	Female	DHE	Lead Computer Scientist	Company
P5	Male	UR	PhD Student in Mathematics	University
P6	Male	UR	PhD Student in Physics	University
P7	Female	UR	PhD Student in Psychology	Freelance/ University
P8	Male	UR	Master Student in Psychology	University
P9	Female	MHE	Psychotherapist	Freelance
P10	Female	MHE	CBT Psychotherapist	University
P11	Female	MHE	CBT Psychotherapist	Freelance
P12	Female	MHE	CBT Psychotherapist	University

Note. DHE: Digital Health Expert; UR: User Representative; MHE: Mental Health Expert.

2.4. Questions and procedure

On the basis of our analysis of design opportunities for AI-powered DTx for stress prevention, leading to the enhancement of the SHIVA solution through modeling, personalization, monitoring, simulation, and consideration of ethics issues, we focused our workshop on 4 main.

aspects: (1) Benefits and concerns of deploying AI techniques for the processing of data collected during the SHIVA intervention for its future adoption, (2) Perceived role of the wearable devices deployed by SHIVA as optional or essential components of the solution, (3) Expectations regarding the future deployment of SHIVA as a fully unguided digital treatment or as a partially human-guided one, (4) Perceived importance of the transparency, explainability and accuracy of the stress detection model and feedback to the user or mental health provider involved in the adoption of the SHIVA intervention. These aspects were chosen as they provide insights into stakeholders' perspectives on key factors facilitating or preventing the future adoption of DTx, such as SHIVA. We asked participants to interview each other on the 4 topics above (see Appendix A) by using a method inspired by the "Round Robin exercise" (Van Velthoven and Cordon, 2019), facilitating a balanced participation of all stakeholders to the discussion. A total of 4 participants per workshop, carefully balanced by expertise to ensure diversity of background knowledge and perspectives, interviewed each other during the first 40 min of the workshop session, by providing individual responses to the 4 questions (see Appendix A). In the last 20 min of the workshop participants were invited by the facilitators to take part in an open discussion on the topics addressed, by adding any further comment representing their personal perspectives and views on the topics discussed.

2.5. Data collection and analysis

The authors MB, SG, SP of this study were the facilitators and observers of the workshop. We greeted the participants as they entered the online session and provided an introduction to the workshop objectives and procedure. Before starting the session we asked participants consent to videorecording the workshop discussion. During the workshop, we observed how the participants engaged in the exercise, answered clarification questions, and listened to comments. Post-workshop, we (MB, SG, SP) conducted a thematic analysis based on Braun and Clarke's six-phase framework (Braun and Clarke, 2006). Initially, each author independently reviewed the video transcripts, noting observations and becoming deeply familiar with the content. We employed open coding to systematically identify and label significant information in the data. This process involved generating codes directly from the raw data without any predetermined categories, ensuring that the analysis was grounded in participants' responses. As the analysis progressed, we grouped similar codes into broader themes, capturing the essence of the data and exploring the underlying patterns and relationships. This thematic development was iterative, involving careful review and refinement to ensure that each theme was coherent and distinct. To ensure reliability and validity, discrepancies in coding or theme interpretation were addressed through consensus. We presented our interpretations and evidence, engaging in discussions to reach a shared understanding. This collaborative approach not only resolved differences but also enhanced the depth and richness of the analysis. Finally, we synthesized into a comprehensive narrative that contextualized the findings within the broader research framework.

3. Results

3.1. Overview on key aspects affecting adoption

This section reports the key findings regarding stakeholders' views on the future refinement and adoption of the SHIVA solution (Table 3).

Regarding the benefits and concerns of deploying AI techniques for the processing of data collected during the SHIVA intervention there was an overall optimism about the potential of AI to improve the effectiveness of stress prevention interventions through higher personalization, accurate monitoring and secure data management. However, stakeholders thought there is still the need to address challenges related to transparency and data management to maximize the acceptance and effectiveness of these techniques. For what concerns the perceived role of the wearable devices deployed by SHIVA as optional or essential components of the solution, the discussion revolved around perceived benefits, privacy and comfort concerns. Participants also discussed the need for demonstrated effectiveness of device-supported interventions, accessibility and personalization of treatment, as well as practical and future considerations regarding their use. The overall view reflected a caution in the mandatory adoption of such technologies, with an inclination towards flexibility and adaptability based on individual needs, circumstances and availability. Among wearable devices, the Oculus headset was the one that raised the most doubts among our participants due to its cost and ergonomics. Regarding the expectations on the future deployment of SHIVA as a fully unguided digital treatment or as a partially human-guided one, most participants thought that a human support would be needed to facilitate the initial engagement with the

Table 3
Workshops themes, frequency and examples of supporting quotes. Frequency is the number of quotes supporting the theme. More quotes for each theme are reported in Appendix B.

Theme	Frequency	Examples of supporting quotes
Challenges of Measuring Stress Accurately	13	*"The stress detection should be as accurate as possible, although I know this is a challenge, depending on the quality of the data collected by the wearables" (P10)*
Presence of Human Facilitators	12	*"A facilitator coach is crucial, particularly initially. From my work using biofeedback, I've seen that even with exercises like diaphragmatic breathing, questions arise when the person is left alone. Having support from a professional could facilitate compliance." (P10)*
Key role of transparency and explainability of the stress detection model / feedback	11	*"I think making the feedback transparent and understandable for the average user is crucial because data that is not understood often worries people. An explanation should be neither overly simplistic nor too technical but comprehensive and accessible to everyone, which I believe is essential." (P2)*
Role of wearables in the solution deployment	10	*"Currently, I see it as optional because if the digital intervention enables access to resources that some might not afford for economic reasons, making something like a VR headset mandatory, which is still somewhat a niche device, could limit this potential benefit." (P2)*
AI processing of data: opportunities and concerns	8	*"The idea that AI itself processes data adds a layer of privacy and security, making it safer in my view than having individuals handle such sensitive information. This enhances the benefits of using AI for personal data and assessments, pushing the advantages further." (P9)*
AI and personalization of treatment	8	*"Using AI for providing feedback on stress levels can help to personalize treatment, such as customizing support on virtual reality for breathing exercises based on previous sessions. This may turn exercises perceived as unnecessary by users to be more motivating." (P5)*

intervention, by guaranteeing empathic interaction with the user and possible clarifications about the treatment or the techniques. According to stakeholders, the initial human support in SHIVA would have a double empowering effect, both by gradually fostering independence in using the digital tools and by encouraging a mindset shift from a reliance on a human facilitator to increased psychological autonomy, resulting in users being more in charge of their own stress management. This self-empowerment, combined with technological guidance, would significantly expand the reach of the intervention, as fewer trained therapists would only be needed for short, hands-on coaching sessions. Personalization of the approach according to the user preferences and individual differences, was deemed key to more effectively supporting adoption and adherence to the intervention. For what regards the perceived importance of the transparency, explainability and accuracy of the stress detection model and feedback, from the responses collected, a complex and multifaceted view emerged, stressing the importance of transparency, explainability, accuracy, personalization and technological integration. Ethical considerations and human support in the context of feedback related to stress measurement in digital treatments were also considered fundamental. The participants emphasized the need to make data and processes understandable to users, observing how the clarity and accessibility of the information provided may strongly improve the user experience and increase trust in the innovative solution deployed.

3.2. Aligned and differing views of stakeholders

In exploring the diverse perspectives on the SHIVA intervention, it became evident that while stakeholders shared some key viewpoints, significant differences also emerged, reflecting the varied priorities and concerns of each group. The perspectives of different stakeholder groups revealed both common ground and varying opinions on critical aspects of stress measurement and intervention delivery. There was a strong consensus on the importance of providing users with accurate, transparent, and explainable stress measurement models. It was highlighted that the level of detail in the feedback should be tailored to individual user preferences. AI-based data analysis during the intervention was seen as key to supporting the personalization of a DTx, potentially increasing user adoption. However, opinions diverged on the use of wearables and VR devices. While some stakeholders, including URs and DHEs, saw these tools as beneficial for enhancing engagement and personalization, concerns were raised about privacy issues and the potential cost barriers. URs and DHEs noted that wearables could facilitate a more customized experience, while MHEs emphasized that the mandatory use of such devices could contribute to increased stress levels in some users who are particularly prone to over-focusing on numbers and constant monitoring. The role of a human coach in the intervention also elicited mixed views. Many stakeholders agreed that a human coach could support initial engagement and ensure the correct performance of relaxation exercises. However, a MHE pointed out that a human coach might inadvertently increase stress in some users, such as those with social phobias. URs suggested that a self-guided approach could generally improve the accessibility and scalability of the DTx. Nevertheless, they strongly preferred a human-guided approach when forced to choose, citing concerns about the quality of emotional support from existing technology and the effectiveness of a fully unguided system. Differences were also observed in preferences for treatment delivery modalities. DHEs generally favored a partially human-guided approach, valuing the role of human interaction in boosting engagement and adherence. MHEs expressed openness to fully unguided treatments, provided that the AI systems are robust and reliable enough to guide users independently, recognizing that some users may prefer less human interaction.

4. Discussion

4.1. Principal findings

This paper reports on findings from a workshop investigation with the aim to identify stakeholders' views on key factors facilitating or preventing the future adoption of AI-powered DTx for stress prevention, such as the SHIVA solution. The workshops led to 4 main insights. First, the deployment of AI-based data processing techniques during the treatment was perceived by stakeholders to bring more benefits than concerns in terms of support to the personalization of the intervention to the user needs, within a trustful and privacy preserving environment ensured by the automatic nature of the processing. However, the existing literature presents a mixed picture regarding this aspect. Some studies support the stakeholders' claims, while others present conflicting views. For instance, (Shan et al., 2022) analyzed reviews of publicly available AI-based mental health chatbots and found that users generally trust these apps, report positive experiences, and express satisfaction. In contrast, (Jain et al., 2024) found that awareness of AI involvement can alter user perceptions of interactions, with humans often perceived as more authentic than AIs. Although they go in different directions, both studies (Shan et al., 2022; Jain et al., 2024) discussed various factors that might influence these perceptions, such as users' education, experiences, biases, and attitudes towards automation, suggesting a complex mechanism in building trust in AI powered DTx solutions. These varying perspectives indicate that trust in AI is not solely dependent on individual interactions but is influenced by broader societal and psychological factors. Building on this foundation, the concept of institutional trust emerges as a central element. Stakeholders pointed out that providing information about the institution promoting the DTx might significantly reduce apprehensions related to surveillance, such as the fear of being monitored, and could increase trust in AI. This insight, supported by findings from (Lee et al., 2022), highlights the importance of transparency in fostering trust during the deployment of interventions like SHIVA. Moreover, the implementation of AI-driven data analysis of the user interaction with SHIVA would be valuable for designers to inform the refinement and adaptation of the solution to the particular needs and preferences of the different target populations, as well as for mental health clinicians to inform the care they provide and better monitor how users engage with the intervention. This methodology has been previously applied to the development of virtual coaching solutions by (Akker et al., 2021) and is now emerging as a powerful framework to guide the development of modern DTx interventions personalized on real-world data gathered directly from the patients. Second, the wearable devices used in SHIVA were considered key to enable a more precise and objective deployment of the intervention, supporting a better assessment of the evolution of health trajectories of users during the weeks of treatment. However, several stakeholders expressed caution towards the mandatory use of these devices, especially the Oculus VR headset, citing potential barriers to adoption due to cost when the SHIVA intervention is rolled out on a large scale. However, one possible solution to overcome this barrier could be the adoption of low cost VR equipment like Google Cardboard, which has been previously evaluated in pain management research (Tong et al., 2016; Patel et al., 2020) and in developing skills for individuals with autism (Miller et al., 2020; Schmidt et al., 2021). Third, the possibility of deploying SHIVA as a blended intervention, supported also by a human facilitator in the initial phases of the treatment was considered by our stakeholders functional to prevent dropouts and support engagement, especially for those users who are less familiar with digital health solutions and may be less motivated to comply with fully self-guided forms of intervention. This result aligns with the growing body of literature suggesting that DTx, particularly those enhanced by artificial intelligence, can be viewed as complementary to more traditional interventions and require, at least for now, the support of a human facilitator (Burger et al., 2022; Egan et al., 2024). Additionally, as suggested by (Lee et al., 2023b), the role of the human coach may be interestingly augmented and supported by the analytics provided by SHIVA, leading to a more fine-tuned and effective deployment of the intervention. Fourth, the key requirements of the stress detection modeling and feedback, in terms of transparency, explainability and accuracy were discussed at length by our stakeholders during the workshops. Ensuring transparency of the AI-driven stress detection component of SHIVA is critical at all stages of the AI model design, training, refinement, testing, as well as in the final deployment of SHIVA in real-world settings. Devising ethical principles and standards for accounting the role of distinct algorithmic processes for providing a more in-depth understanding of how AI-driven stress detection operates will be critical (Ananny and Crawford, 2018; Giovanelli et al., 2023). This is a main challenge that future research should address to raise the quality of the stress detection solutions provided and their adoption by stakeholders.

4.2. Implications for AI-based DTx development

The insights gained from the SHIVA project provide important considerations for the development of AI-based DTx solutions. Based on the findings presented above, key recommendations for researchers and developers are outlined in Table 4.

Table 4
Suggestions for the development of AI-based DTx solutions.

Area	Suggestion
Personalization	Utilize AI systems to deliver highly personalized therapeutic interventions, tailoring treatments to individual user needs for increased effectiveness.
Privacy	Develop privacy-preserving data processing frameworks where sensitive user data is processed locally, reducing privacy concerns and building user trust.
Transparency and Trust	Ensure AI models are reliable and transparent. Provide clear and comprehensive information on AI operations and data management to enhance trust.
Flexibility	Adopt a modular design approach for DTx interventions to allow adaptable and flexible treatment protocols, meeting individual user needs.
Inclusivity and Accessibility	Prioritize the use of widely available, cost-effective wearable devices to make the technology accessible and practical for long-term use across user groups.
Support Systems	Integrate support facilitators like human coaches, FAQ systems, or AI-driven coaching bots from early stages to enhance user experience, engagement and effectiveness.
Stakeholder Engagement	Engage stakeholders, including users, providers, and regulatory bodies, early and consistently to ensure alignment with end-user needs and expectations.

These recommendations highlight the critical areas that must be addressed to ensure the successful integration of AI in DTx solutions, ultimately leading to more effective, trustworthy, and accessible therapeutic interventions.

4.3. Limitations and future research

There is growing evidence suggesting that AI techniques can empower DTx for mental health and well-being, supporting their wider adoption by users and health providers. However, due to the complexity of behavior change interventions for mental health, the design of dedicated AI-powered DTx raises open questions and challenges deserving a deeper investigation by the broader research community. The contribution of this work is to have shed some light on the most promising steps to take in the development of DTx for stress prevention by starting in-depth discussion with key stakeholders in the field. Main limitations include that our workshops involved a limited number of participants and addressed challenges related to the design and implementation of SHIVA, a stress prevention solution currently in the design and refinement stage. The selection process for participants relied on personal connections and snowball sampling, which may have influenced the diversity of perspectives. This method could potentially lead to the inclusion of participants with similar viewpoints and experiences, which might not fully capture the range of insights needed. Additionally, there was a gender imbalance among the different stakeholder groups, which might have subtly influenced the perspectives shared during the workshop. For example, the underrepresentation of males in some groups (e.g., MHEs) and females in others (e.g., URs) could have affected the discussions, reflecting more the predominant experiences of the over-represented groups. Despite these limitations, it is our intention to further carry on our discussion with stakeholders in the next phases of SHIVA development and validation, to ensure a responsible and effective use of AI-powered DTx for their larger adoption in mental healthcare settings. Future research should aim to use more systematic and randomized selection methods to ensure a more balanced and representative sample of stakeholders. This will help to mitigate potential biases and provide a more comprehensive understanding of the challenges and opportunities in the development and implementation of AI-powered DTx.

5. Conclusion

This paper contributes insights from multi-stakeholders workshops aimed at uncovering the most promising steps to take in the human-centered design of AI-powered DTx for stress prevention, by starting from the current design and refinement of the SHIVA solution. Overall, stakeholders' views were aligned in finding AI-driven opportunities key to facilitate the future adoption of SHIVA by supporting the achievement of a better personalization of the solution to the user needs and preferences. However, important challenges were identified in the need to realize a transparent, explainable and accurate stress detection model and feedback for the SHIVA users, fostering their full engagement, adherence and trust in the solution over the intervention deployment.

Based on these insights, the SHIVA project has outlined the following roadmap:

1. Feedback-driven refinements (e.g., making wearables optional, adopting a modular approach, enhancing AI models for better personalization, and introducing an LLM-based, human-like chatbot for user coaching to improve adherence and motivation).
2. Development of a new stress detection model that is explainable and addresses the challenge of ensuring user trust.
3. Piloting with stakeholders, to assess the appropriateness of the refined SHIVA solution.
4. Conducting a feasibility study to test the refined SHIVA solution in real-world settings.

AI statement

During the preparation of this work the authors used ChatGPT (GPT-4o) to improve readability and language. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This work was funded by Hub Life Science – Digital Health (LSH-DH) PNC-E3-2022-23683267 - Project DHEAL-COM, Ministry of Health (Italy) under the Piano Nazionale Complementare al PNRR Ecosistema Innovativo della Salute, Code PNC-E.3.

CRediT authorship contribution statement

MB, SP and SG contributed substantially to the conception and design of the study, to the acquisition of data and to the editing of the manuscript. All authors contributed to the article and approved the submitted version.

Declaration of competing interest

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

Acknowledgements

None.

Appendix A. Workshops details

A.1. Procedure details

1. *Preparation and Distribution of Materials:* Prior to the workshops, participants received a short document describing the DTx prototype, a link to the serious game environment, and a video detailing the main components of the DTx solution.
2. *Conducting the Workshops on Meet:* Three stakeholders workshops were held online via Google Meet in March 2024. Each workshop included four participants and at least one representative from each of the three identified stakeholder categories (details are provided in Appendix A.4).
   a. *Individual Interviews:* The first 40 min of each workshop were dedicated to individual interviews where each of the four participants interviewed another participant. Individual responses to the four predefined questions (reported in Appendix A.3) were collected.
   b. *Open Discussion:* In the last 20 min of each workshop, participants were encouraged by the facilitators to engage in an open discussion about the topics covered. They were invited to add further comments, representing their personal perspectives and views on the discussed topics.

Role of facilitators: The authors of the study served as facilitators and observers of the workshop. They welcomed the participants into the online session and provided an introduction (example in page 2) to the workshop's objectives and procedures. Throughout the workshop, they observed how participants engaged with the exercise, responded to questions, and listened to comments.

A.2. Introduction to the workshop

The facilitators welcomed all participants, emphasizing the value of their contribution to guide the future evolution of the SHIVA solution. Participants were introduced to the workshop phases: (a) individual interviews and responses to the workshop questions (10 min per answer), (b) collective discussion on the topics addressed. It was clarified that during phase (a) each participant would pose one of the questions to another participant and then answer to the question posed by their fellow participant in return. This procedure was meant to ensure a balanced contribution of each participant to the workshop discussion. Before starting with phase (a) the facilitators answered all clarifying questions about the description of the SHIVA solution shared with participants prior to the workshop.

A.3. Questions

Q1 Benefits and concerns of deploying AI techniques for the processing of data
"In your view, the deployment of AI techniques for the processing of data collected during the DTx (for monitoring, user profiling, etc…) brings mainly benefits or concerns for the future adoption of this DTx?"

Q2 DTx and role of wearables
"Do you think that the wearable devices involved in this DTx (stress monitoring device, VR headset) should be required or optional components of the future DTx deployment?"

Q3 Expectations about treatment delivery modality
"Do you think that this DTx would better fit a future deployment as a fully unguided treatment or as a (partially) human guided one (coach/facilitator)?"

Q4 Perceived importance of the transparency, explainability and accuracy of the stress detection model
"How far do you think it is important to make transparent and explainable the stress detection feedback related to the DTx treatment to the user and/or to the human coach/facilitator? Additionally, how accurate should this feedback/model be?"

A.4. Groups composition

Individuals were divided into the following groups according to their characteristics in order to obtain heterogeneous groups of stakeholders:
Group #1: P2 (DHE), P5 (UR), P7 (UR), P10 (MHE
Group #2: P1 (DHE), P6 (UR), P11 (MHE), P12 (MHE)
Group #3: P3 (DHE), P4 (DHE), P8 (UR), P9 (MHE)

A.5. Rotation sequence

Table A.5
Rotation sequence.

Group	Question	Ask	Answer
G1	Q1	P7	P5
	Q2	P5	P7
	Q3	P2	P10
	Q4	P10	P2
G2	Q1	P6	P1
	Q2	P11	P12
	Q3	P12	P11
	Q4	P1	P6
G3	Q1	P4	P3
	Q2	P9	P8
	Q3	P8	P9
	Q4	P3	P4

Appendix B. Extended version of Table 3

Table B.6
Extended version of Table 3 [more quotes].

Theme	Examples of supporting quotes
Challenges of measuring stress accurately	*"The stress detection should be as accurate as possible, although I know this is a challenge, depending on the quality of the data collected by the wearables" (P10)* *"Understanding stress is complex because it involves multiple factors. Physical signs like heart rate, sweating, or muscle tension can indicate stress, but they aren't definitive on their own, as these indicators may have other causes." (P1)*
Presence of human facilitators	*"A facilitator coach is crucial, particularly initially. From my work using biofeedback, I've seen that even with exercises like diaphragmatic breathing, questions arise when the person is left alone. Having support from a professional could facilitate compliance." (P10)* *"Offering a self-guided option could make it more accessible, but having a coach might make the intervention more effective and reduce dropout risk. A self-guided approach could lower barriers for those hesitant to see a psychologist." (P5)*
Key role of transparency and explainability of the stress detection model / feedback	*"I think making the feedback transparent and understandable for the average user is crucial because data that is not understood often worries people. An explanation should be neither overly simplistic nor too technical but comprehensive and accessible to everyone, which I believe is essential." (P2)* *"Transparency is crucial, especially when it comes to stress. It's important to provide patients with a clear explanation of the measurements taken. By helping them understand what they are experiencing during stressful situations, such as elevated heart rates and other physiological changes, we can foster greater self-awareness." (P9)*
Role of wearables in the solution deployment	*"Currently, I see it as optional because if the digital intervention enables access to resources that some might not afford for economic reasons, making something like a VR headset mandatory, which is still somewhat a niche device, could limit this potential benefit." (P2* *"Wearables play a key role in collecting data for stress detection, but they should be optional. Not everyone can afford devices like VR headsets, and making them mandatory could exclude some users. It's better to offer them as an optional enhancement to the experience." (P3)*
AI processing of data: opportunities and concerns	*"The idea that AI itself processes data adds a layer of privacy and security, making it safer in my view than having individuals handle such sensitive information. This enhances the benefits of using AI for personal data and assessments, pushing the advantages further." (P9* *"AI can analyze complex data sets more efficiently than humans. However, the lack of transparency in how AI makes its decisions can be a barrier to user acceptance. Clear guidelines and explanations can help mitigate these concerns." (P10)*
AI and personalization of treatment	*"Using AI for providing feedback on stress levels can help to personalize treatment, such as customizing support on virtual reality for breathing exercises based on previous sessions. This may turn exercises perceived as unnecessary by users to be more motivating." (P5)* *"AI-driven personalization can make stress management interventions more effective by tailoring activities to individual needs. For instance, AI can adjust the intensity of exercises based on a user's stress levels, providing a more customized approach." (P11)*

References

Akker, H. op den, et al., 2021. Digital therapeutics: virtual coaching powered by artificial intelligence on real-world data. Frontiers in Computer Science 3, 750428.

Ananny, M., Crawford, K., 2018. Seeing without knowing: limitations of the transparency ideal and its application to algorithmic accountability. New Media Soc. 20 (3), 973–989.

Apolinàrio-Hagen, J., et al., 2020. Cognitive behavioral therapy, mindfulness-based cognitive therapy and acceptance commitment therapy for anxiety disorders: integrating traditional with digital treatment approaches. Anxiety disorders: Rethinking and understanding recent discoveries 291–329.

Arora, A., Arora, A., 2022. Generative adversarial networks and synthetic patient data: current challenges and future perspectives. Future Healthcare Journal 9 (2), 190–193.

Bazoukis, G., et al., 2022. The inclusion of augmented intelligence in medicine: a framework for successful implementation. Cell Reports Medicine 3, 1.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qual. Res. Psychol. 3 (2), 77–101.

Burger, F., et al., 2022. Using a conversational agent for thought recording as a cognitive therapy task: feasibility, content, and feedback. Frontiers in Digital Health 4, 930874.

Can, Y.S., et al., 2019. Continuous stress detection using wearable sensors in real life: algorithmic programming contest case study. Sensors 19 (8), 1849.

Craig, P., et al., 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. Bmj 337.

Czajkowski, S.M., et al., 2015. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol. 34 (10), 971.

Dang, A., et al., 2020. Role of digital therapeutics and the changing future of healthcare. J. Family Med. Prim. Care 9 (5), 2207–2213.

Dreyer, R.P., et al., 2022. PreScription DigitaL ThErapEutic for patients with insomnia (SLEEP-I): a protocol for a pragmatic randomised controlled trial. BMJ Open 12 (8), e062041.

Egan, S.J., et al., 2024. A pilot study of the perceptions and acceptability of guidance using artificial intelligence in internet cognitive behaviour therapy for perfectionism in young people. Internet Interv. 35, 100711.

Epping-Jordan, J.E., et al., 2016. Self-Help Plus (SH+): a new WHO stress management package. World Psychiatry 15 (3), 295.

Food, U. and (FDA), D. A, 2020. FDA permits marketing of first game-based digital therapeutic to improve attention function in children with ADHD. https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-first-game-based-digital-therapeutic-improve-attention-function-children-adhd.

Garcia-Ceja, E., et al., 2018. Mental health monitoring with multimodal sensing and machine learning: a survey. Pervasive and Mobile Computing 51, 1–26.

General Data Protection Regulation (GDPR), 2018. https://gdpr-info.eu/.

Giovanelli, A., et al., 2023. Supporting adolescent engagement with artificial intelligence–driven digital health behavior change interventions. J. Med. Internet Res. 25, e40306.

Hayes, S.C., Pierson, H., 2005. Acceptance and Commitment Therapy. Springer.

Hayes, S.C., et al., 2011. Acceptance and Commitment Therapy: The Process and Practice of Mindful Change. Guilford press.

Hu, P., et al., 2024. Computing and Artificial Intelligence in Digital Therapeutics.

Jain, G., et al., 2024. Revealing the source: how awareness alters perceptions of AI and human-generated mental health responses. Internet Interv. 36, 100745.

Jeong, H., et al., 2023. Virtual agents in DTx: focusing on usability and therapeutic effectiveness. Electronics 13 (1), 14.

Jeong, H., et al., 2024. Deep breathing in your hands: designing and assessing a DTx mobile app. Frontiers in Digital Health 6, 1287340.

Kellogg, K.C., Sadeh-Sharvit, S., 2022. Pragmatic AI-augmentation in mental healthcare: key technologies, potential benefits, and real-world challenges and solutions for frontline clinicians. Frontiers in Psychiatry 13, 990370.

Lee, H., et al., 2022. Understanding privacy risks and perceived benefits in open dataset collection for mobile affective computing. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 6 (2), 1–26.

Lee, U., et al., 2023a. Toward data-driven digital therapeutics analytics: literature review and research directions. IEEE/CAA Journal of Automatica Sinica 10 (1), 42–66.

Lee, U., et al., 2023b. Data-driven digital therapeutics analytics. In: 2023 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, pp. 386–388.

McGinnis, E., et al., 2022. A digital therapeutic intervention delivering biofeedback for panic attacks (PanicMechanic): feasibility and usability study. JMIR Formative Research 6 (2), e32982.

Miller, I.T., et al., 2020. Virtual reality air travel training using apple iPhone X and google cardboard: a feasibility report with autistic adolescents and adults. Autism in adulthood 2 (4), 325–333.

MobiHealthNews, 2021. Happify Health rolls out first PDTx for depression and anxiety. https://www.mobihealthnews.com/news/happify-health-rolls-out-first-pdtx-depression-and-anxiety.

Palanica, A., et al., 2020. The need for artificial intelligence in digital therapeutics. Digital biomarkers 4 (1), 21–25.

Patel, P., et al., 2020. Low-cost virtual reality headsets reduce perceived pain in healthy adults: a multicenter randomized crossover trial. Games for health journal 9 (2), 129–136.

Ponzo, S., et al., 2020. Efficacy of the digital therapeutic mobile app BioBase to reduce stress and improve mental well-being among university students: randomized controlled trial. JMIR mHealth and uHealth 8 (4), e17767.

Recchia, G., et al., 2020. Digital therapeutics-what they are, what they will be. Acta Sci Med Sci 4 (3), 1–9.

Refolo, P., et al., 2022. Ethics of digital therapeutics (DTx). Eur. Rev. Med. Pharmacol. Sci. 26, 18.

Roy, A., et al., 2021. Clinical efficacy and psychological mechanisms of an app-based digital therapeutic for generalized anxiety disorder: randomized controlled trial. J. Med. Internet Res. 23 (12), e26987.

Schmidt, M., et al., 2021. Evaluation of a spherical video-based virtual reality intervention designed to teach adaptive skills for adults with autism: a preliminary report. Interact. Learn. Environ. 29 (3), 345–364.

Shah, V., 2022. AI in mental health: predictive analytics and intervention strategies. Journal Environmental Sciences And Technology 1 (2), 55–74.

Shan, Y., et al., 2022. Public trust in artificial intelligence applications in mental health care: topic modeling analysis. JMIR Human Factors 9 (4), e38799.

Shiffman, S., et al., 2008. Ecological momentary assessment. Annu. Rev. Clin. Psychol. 4 (1), 1–32.

Smoktunowicz, E., et al., 2020. Consensus statement on the problem of terminology in psychological interventions using the internet or digital components. Internet Interv. 21, 100331.

Spitzer, M., et al., 2023. Digital twins and the future of precision mental health. Frontiers in Psychiatry 14, 1082598.

Tol, W.A., et al., 2020. Guided self-help to reduce psychological distress in South Sudanese female refugees in Uganda: a cluster randomised trial. Lancet Glob. Health 8 (2), e254–e263.

Tong, X., et al., 2016. The design of an immersive mobile virtual reality serious game in cardboard head-mounted display for pain management. In: Pervasive Computing Paradigms for Mental Health: 5th International Conference, MindCare 2015, Milan, Italy, September 24–25, 2015, Revised Selected Papers 5. Springer, pp. 284–293.

Van Velthoven, M.H., Cordon, C., 2019. Sustainable adoption of digital health innovations: perspectives from a stakeholder workshop. J. Med. Internet Res. 21 (3), e11922 (issn: 1438-8871).

Vasdev, N., et al., 2024. Navigating the future of health care with AI-driven digital therapeutics. Drug Discov. Today 29 (9), 104110.

Wang, Q., et al., 2023. Enhancing the conversational agent with an emotional support system for mental health digital therapeutics. Frontiers in Psychiatry 14, 1148534.

World Health Organization, 2020. Doing What Matters in Times of Stress: An Illustrated Guide. World Health Organization, Geneva, Switzerland isbn: 978-92-4-000392-7. url: https://www.who.int/publications/i/item/9789240003927.

World Health Organization, 2021. SelfHelp+: a group-based stress management course for adults. https://www.who.int/publications/i/item/9789240035119. Geneva, Switzerland.