

Document Version

Final published version

Citation (APA)

Figuroa, C. A., Ramos, G., Psihogios, A. M., Ekuban, E., Bansie, P., de Haas, M., Karnik, N. S., Ajilore, O., Anderson, E. E., & Stiles-Shields, C. (2025). Advancing youth co-design of ethical guidelines for AI-powered digital mental health tools. *Nature Mental Health*, 3(8), 870-878. <https://doi.org/10.1038/s44220-025-00467-7>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

**Green Open Access added to [TU Delft Institutional Repository](#)
as part of the Taverne amendment.**

More information about this copyright law amendment
can be found at <https://www.openaccess.nl>.

Otherwise as indicated in the copyright section:
the publisher is the copyright holder of this work and the
author uses the Dutch legislation to make this work public.

Advancing youth co-design of ethical guidelines for AI-powered digital mental health tools

Received: 2 January 2025

Accepted: 24 June 2025

Published online: 31 July 2025

 Check for updates

C. A. Figueroa^{1,2}✉, G. Ramos^{3,4}, A. M. Psihogios^{5,6}, E. Ekuban⁷, P. Bansie⁷, M. de Haas⁷, N. S. Karnik^{8,9}, O. Ajilore¹⁰, E. E. Anderson¹¹ & C. Stiles-Shields^{8,9}

Adolescents and young adults (AYA) often face mental health challenges and are heavily influenced by technology. Digital health interventions (DHIs), leveraging smartphone data and artificial intelligence, offer immense potential for personalized and accessible mental health support. However, ethical guidelines for DHI research fail to address AYA's unique developmental and technological needs and leave crucial ethical questions unanswered. This gap creates risks of either over- or under-protecting AYA in DHI research, slowing progress and causing harm. This Perspective examines ethical gaps in DHI research for AYA, focusing on three critical domains: challenges of passive data collection and artificial intelligence, consent practices, and risks of exacerbating inequities. We propose an agenda for ethical guidance based on bioethical principles autonomy, respect for persons, beneficence and justice, developed through participatory research with AYA, particularly marginalized groups. We discuss methodologies to achieve this agenda, ensuring ethical, youth-focused and equitable DHI research for the mental health of AYA.

The mental health of adolescents and young adults (AYA) has come under great strain due to challenges, such as the COVID-19 pandemic¹, increased loneliness and isolation², social inequity and systemic racism³. AYA are more likely to engage in risky behaviors and experience mental health symptoms, but are less likely to receive mental health care compared with their younger (that is, children <12 years) and older (that is, 26+ years) peers^{4,5}. This combination of heightened risk and limited care access impacts the present and long-term well-being of AYA⁵, underscoring the need for solutions.

Digital health interventions (DHIs), the use of scalable technological tools (for example, apps, websites and wearables) to track or improve mental health outcomes, have the potential to increase mental health-care access for AYA⁶, making services accessible beyond traditional health-care settings⁷. In several age groups, DHIs have shown promising outcomes. For example, small but promising effects of DHIs were found for promoting well-being, relieving anxiety and enhancing protective factors in 11–18-year-olds⁸. Further, a meta-analysis of 6–18-year-olds revealed that DHIs have a moderate treatment effect

¹Technology, Policy and Management, Delft University of Technology, Delft, the Netherlands. ²School of Social Welfare, University of California Berkeley, Berkeley, CA, USA. ³Department of Psychological Science, University of California, Irvine, Irvine, CA, USA. ⁴Departments of Epidemiology and Biostatistics and Medicine, University of California, San Francisco, San Francisco, CA, USA. ⁵Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ⁶Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ⁷Be Ubuntu Foundation, Amsterdam, the Netherlands. ⁸Institute for Juvenile Research, Department of Psychiatry, College of Medicine, University of Illinois Chicago, Chicago, IL, USA. ⁹AI.Health4All Center, College of Medicine, University of Illinois Chicago, Chicago, IL, USA. ¹⁰Department of Psychiatry, University of Illinois Chicago, Chicago, IL, USA. ¹¹Neiswanger Institute for Bioethics and Healthcare Leadership, Stritch School of Medicine, Loyola University Chicago, Maywood, IL, USA. ✉e-mail: c.figueroa@tudelft.nl

on promoting mental health⁹. In 10–24-year-olds, DHIs can be effective in supplementing traditional mental health services or serving as the main treatment^{10,11}. Findings also suggest moderate, although limited, evidence supporting DHIs for improving mental health outcomes among socioeconomically and digitally marginalized youth¹⁰.

However, many unresolved issues remain. First, only a small proportion of existing DHIs are evidence-based, and cost-effectiveness studies are scarce⁶. Further, DHIs suffer from low engagement and high attrition, particularly among AYA^{12,13}, in part because they use a ‘one-size-fits-all’ approach^{12,13} and suffer from unclear ethics and safety practices¹⁴. Further, developing DHIs for AYA requires an understanding of their cognitive development, decision-making processes and unique technological needs, to avoid interventions that may negatively impact AYA’s lives. In this paper, we show how current frameworks on the ethics of digital mental health do not fully address the unique intersection of artificial intelligence (AI), DHIs and the specific needs and vulnerabilities of AYA, and recommend AYA’s involvement in co-creating ethical guidance.

Using AI for persuasive technology for AYA

To overcome issues of engagement and attrition, researchers are increasingly turning to AI and machine learning to design persuasive technology that fits the lives, needs and behaviors of AYA¹⁵. While high-quality evidence is still developing¹⁶, emerging research shows AI-driven DHIs could transform the field of mental health care. For example, AI-based conversational agents (that is, chatbots) have shown a large positive effect on reducing depressive symptoms in adults^{17,18}. Machine-learning algorithms, leveraging real-time data on individuals’ behaviors, such as sleep patterns, physical activity and social interactions, are comparable to or better than human assessors for treatment outcome, adherence and engagement prediction^{7,19}. Machine-learning algorithms also predict suicide risk with small to moderate accuracy^{20,21}.

Importantly, although AI in mental health is at an experimental stage, tools such as AI-based cognitive behavioral therapy chatbots and AI to predict mental health disorders are becoming integrated into mental health care²², although it is unclear how many are specifically for AYA. Further, AYA are increasingly using tools not meant for mental health care, such as publicly available large language models (LLMs), to discuss mental health struggles²³, although the numbers of actual use of LLMs for mental health by AYA are lacking. Emerging research shows that AYA themselves are worried about their ethical use in health care²⁴. With the use of AI in care seeing remarkable growth, alongside increasing calls for using AI in mental health for AYA²⁵ from researchers, health professionals and policymakers, unique ethical challenges arise. As such, new ethical guidelines specific to research with AYA are warranted.

Current DHI ethics frameworks

Several frameworks for DHI ethics have been developed. For example, Laacke et al.²⁶ developed the concept of health-related digital autonomy by analyzing case studies of AI in mental health for social media²⁶. This concept encompasses challenges, such as transparency, user empowerment, and the risks of unintended labeling and stigmatization. In Europe, 16 ethical principles for DHI have been proposed²⁷, consisting of 4 ethical dimensions (including placing DHIs within a framework of humanist values, enabling people to manage their own health data) and inclusive and eco-friendly DHIs. In addition, several checklists for DHI research ethics have been published to guide researchers through various stages of the research process, including informed consent, language regarding data ownership, sharing and monitoring, and accessibility to diverse populations^{28,29}. Psihogios et al.³⁰ published ethical guidance for research in pediatric DHIs (specifically for using remote sensors for pediatric mental health outcomes) related to device selection, data collection and data transparency. In addition, a recent systematic review provides a framework for considerations in AI for

the health of children³¹, but not specific to DHIs and mental health. While these frameworks and tools are an excellent starting point, there is a lack of guidance on DHI research specific to AYA and AI-powered digital mental health.

The current AI regulatory landscape

While new regulations for AI-enabled health tools are under development, the current landscape seems ill-equipped to guide the use of these technologies with AYA. For example, regulations, such as the General Data Protection Regulation and the Health Insurance Portability and Accountability Act, focus on data privacy but fail to address ethical concerns and psychological impacts of AI-driven tools. Medical device regulations primarily cover clinical tools, often excluding wellness apps and non-diagnostic mental health AI interventions³². An exception is the European Union Artificial Intelligence Act of 2024, which categorizes AI systems by risk and considers child-specific protections, providing a significant step toward AI governance³³. However, the current European Union regulatory framework scarcely addresses the unique impact of AI on human interactions and emotions²², let alone those of AYA. The AI Act largely articulates high-level objectives, and associated guidance, standards and policy are needed to ensure its practical implementation in the health-care sector³⁴. Broader frameworks such as those from UNICEF (United Nations Children’s Fund)³⁵, the SRights Children and AI design code³⁶ and the World Economic Forum³⁷ emphasize children’s rights in AI, but they are not specific to DHIs and mental health. In addition, as far as we are aware, none of the professional bodies for child and adolescent mental health (including the American Academy of Child & Adolescent Psychiatry and the European Society for Child and Adolescent Psychiatry) has published ethical guidance for recommendations for AI used in DHI for AYA’s mental health. The American Psychological Association published guidance regarding the use of AI in mental health for clinicians³⁸ and urged the Federal Trade Commission and legislators to put safeguards in place as users increasingly turn to companion apps for mental health support³⁹. In addition, the US National Academy of Medicine published an AI code of conduct⁴⁰ intended to align the field around responsible development and application of AI. This code is also not specific to children and mental health.

Thus, DHI researchers, policymakers, health professionals and bioethicists must advocate stronger ethical standards and policies to protect AYA while empowering them to shape the future of AI-driven mental health interventions.

Aim of this Perspective

This Perspective describes critical issues warranting ethical guidance focusing on mental health and developments in AI research. We base this work on widely accepted ethical principles for biomedical research, including respect for autonomy (ensuring the dignity of AYA as well as their right to make their own decisions^{41,42}), beneficence (minimizing the risk of harm and maximizing the potential for benefits to participants and to science) and justice (fair distribution of risks and benefits of research^{41,42}). We first explain critical ethical issues on the basis of our experiences and the current literature. We then provide preliminary guidance and a research agenda through several discussion sessions between the authors (who have experience in DHI, pediatric psychiatry and psychology, community-based research and ethics) and the AYA community partners who have been involved in projects related to AI and mental health.

AYA are among the most affected by a digital transformation in mental health, yet they have not been involved enough in the decision-making around ethical guidance for DHI research. Recognizing this gap, we issue a call to action for those involved in developing ethical guidance for DHIs. This call to action concerns in particular academic researchers, but it can also include professional regulatory groups and technology developers. We outline targeted methodologies and practical strategies

to ensure AYA's meaningful engagement in these efforts. With this paper, we aim to influence the ethical infrastructure surrounding DHI research and development. By bridging this gap, we hope to ensure that ethical guidelines are based on robust biomedical principles that fit within best practices in research while addressing the needs and wishes of AYA.

Ethical challenges of DHI research with AYA

In the following section, we describe three important ethical issues that require new guidance: (1) complications of passive data collection and AI, (2) consent practices, and (3) risk of exacerbation of injustice. We explain how these issues relate to the bioethical principles of autonomy, beneficence and justice, and provide preliminary recommendations for DHI researchers to address these challenges and better protect AYA in DHI research.

Complications of passive data collection and AI

Research that uses passive data collection and AI poses threats, especially to the respect for AYA's autonomy and beneficence. AYA frequently navigate digital spaces and create vast amounts of data—including technological behavioral interactions (for example, frequency, duration, typing speed and word choices), social communication, physical activity and location. Because these have been linked to mental health outcomes⁴³, collecting these data may help identify AYA at risk and increase their access to mental health care through digital referral. However, researchers lack guidance to account for the complications of such data collection related to the daily activities and experiences of young people⁴⁴.

Continuous tracking of AYA's behavior can lead to researchers receiving information that they may not have expected or may not have a plan for acting on. For example, what if an app indicates an adolescent engages in online cyberbullying, skips school, reports suicidality or suffers severe substance abuse? Are researchers required to report this, and to whom? How should researchers inform AYA and their guardian about their duties to report or not, or when they might break promises of confidentiality? Further, given the rapid evolution of AI methods, it is difficult for anyone to understand the full implications of data-mining personal information. AYA reported willingness to share data if doing so adds value for users and if the data request is not too intimate. However, they also demanded transparency in data usage and control over personalization¹⁵. Researchers lack guidance to determine when data are too intimate to collect and what kind of control AYA should have over personalization of data and could benefit from clarification regarding AYA's expectations for transparency across their development.

AYA may also over-rely on AI, attributing human-like properties, over-exaggerating their expectations of and trust in AI⁴⁵ (for example, believing an AI chatbot assistant is a real friend). This could be especially concerning for vulnerable AYA experiencing loneliness, who may be more likely to use AI chatbots for companionship. For example, a public lawsuit gained notoriety when parents sued an AI company because their teenage son died by suicide after using their chatbot⁴⁶. In addition, recent research shows an empathy gap of LLMs such that current models struggle to respond to unpredictable, abstract and emotional conversations with AYA⁴⁷. Other research has also argued that LLMs do not align perfectly with human values, for example, focusing more on acceptance instead of challenging harmful behaviors, making their use in mental health settings risky⁴⁸. AYA conversing with LLMs may not be able to assess the consequences of interacting with AI and exercise their right to autonomy, leading to AI products not benefiting AYA, and even causing harm^{47,49}.

Emerging evidence suggests that AYA's social interactions with AI can also influence their interpersonal relationships. Conversational agents show potential in teaching AYA politeness and facilitate conflict resolution between parents and children⁵⁰. In addition, AI can enhance peer support by crafting supportive responses and styles²³. Thus, AI tools may enhance peer and familial bonds; however, future research

should clearly identify AYA's diverse exposure to AI and its potential effects on their well-being⁴⁹.

These issues raise concerns that researchers may violate the bioethical principles of autonomy and beneficence. At the same time, as stated in UNICEF's foundations for child-centered AI, "the opportunities that AI systems bring to mental health care, such as increased personalization and access, need to be fully leveraged when it is appropriate to use AI systems"³⁵. As such, not using novel technologies could also undermine these same principles.

Research involving AI must adopt safeguards to ensure that potential harms are minimized and benefits maximized, while also protecting AYA's autonomy. This requires providing concise and clear education prior to and across decision points, such as sharing data, giving AYA options to opt out of data collection or to choose between personalized and non-personalized app versions. This must be communicated in a youth-friendly way (Table 1). Researchers must continue to ask key questions, such as how can we make AYA aware of AI's capabilities and risks and involve AYA in answering these questions? In parallel, to understand if, when and how to intervene when learning of certain risky behaviors, design teams should, besides AYA, also consult stakeholders, such as schools, caregivers and health professionals. Some steps can also be taken via regional laws and regulations. For example, California now has a state law that AI chatbots must state that they are not human every 3 hours⁵¹. These types of safeguards can help minimize the risks of AI to the autonomy of AYA.

Researchers must also consider the trade-offs between AI ethical decision points. Giving AYA complete control over their interactions with AI would allow full autonomy, yet it could compromise the principle of beneficence by introducing risks, for example, allowing AYA to bypass important digital mental health intervention functions (for example, ignoring alerts) or even crucial intervention components (for example, not designing a safety plan for suicidal ideation). In addition, full transparency may facilitate the potential re-identification (that is, de-anonymization) of AYA⁵². Future research should focus on understanding how to best balance the bioethical principles of autonomy and beneficence when using passive data collection and AI in DHIs for mental health.

Consent practices are too complex

Primary parental consent, which is the current legal standard to protect AYA under the age of 18, can threaten AYA's autonomy and impede justice. Human participant review panels have long recognized that adolescents have some capability to make independent decisions, but the default for research has been parental or guardian consent alongside adolescent assent. In some situations, requiring parental consent may increase the risk of harm. Some AYA use digital tools to learn about identity (for example, LGBTQ+), gain access to peers and find support⁵³. Requiring caregivers' consent may alert them to behavior they believe to be punishable, leading to risks for AYA, from discomfort to personal safety or housing instability⁵⁴.

Consent practices can also limit the potential for a diverse and inclusive sample. For example, during the COVID-19 pandemic, a DHI study for depression waived parental consent to be more accessible, and showed a higher representation of AYA with sexual minority identities and racial and ethnic minority backgrounds³⁵. Thus, waiving parental consent may lead to more inclusive research samples, promoting the bioethical principle of justice, and can benefit AYA's mental health while preventing the risk of harm (beneficence) for some AYA.

We also acknowledge the known beneficial role of parent participation in AYA mental health treatment. Especially for younger AYA, a mental health app might provide features that encourage them to share concerns with their parents or suggest joint activities without compromising personal privacy, which is a facilitator for AYA's engagement in DHIs⁵⁶. However, such a tool could also let the young user

Table 1 | Preliminary guidance for researchers for AI-based DHI research with AYA

Ethical issue	Questions	Ethical recommendations
Passive data collection and AI complicates autonomy and beneficence	<p>What kind of control should AYA have over personalization (for example, basing intervention components on their data and (sensed) behavior)?</p> <p>How do researchers explain AI benefits and limitations to AYA?</p> <p>Do researchers have any responsibility to intervene when collecting data about risky behaviors (for example, skipping school, cyberbullying, suicidality or building ‘too intimate’ relationships with AI)?</p> <p>What level of transparency do adolescents versus young adults expect regarding usage of their data?</p> <p>How can researchers consider the changing needs of AYA, considering brain development, impulse control and cognition?</p>	<p>Give options to opt out of data collection and personalization (for example, having a personalized or non-personalized version)</p> <p>For each study, involve AYA advisory boards to guide ethical responses to detected behaviors. For example, if ‘risky behavior’ is detected, action mechanisms, such as warnings, links to resources, a reporting system to flag harmful content, contacting parents in the case of young AYA, or messaging to underscore that AI is not a human</p> <p>Researchers should partner with AYA, and other stakeholders, such as caregivers, schools and mental health professionals, in determining these red flags and developing appropriate response pathways</p> <p>Develop data ownership rules based on AYA’s developmental phase and their interactions in the online world</p> <p>Invest in youth-friendly materials for AI literacy and trust, such as videos or reels, vlogs and quizzes, to explain AI being used, the data being collected, and what research participation means</p> <p>To explain AI, use creative methods of connecting with AYA, such as metaphors, storytelling and peer explanations (that is, describing AI to each other)</p>
Consent practices impede autonomy and justice	<p>When does a requirement to obtain parental consent in research put adolescents at increased risk?</p> <p>When is it feasible and safe to waive parental/guardian consent in DHI research?</p> <p>How can AYA be ensured they have the right to contribute to research about their lives?</p>	<p>Work with institutional review boards to determine appropriate circumstances for waiving the requirement for parental or guardian consent, using alternatives to proxy consent (involve other trusted professionals, such as teachers or youth workers for consent), models of graduated consent or minor self-consent if parental consent is contra-indicated</p> <p>Involve AYA and caregivers in the development of additional safeguards for AYA research participants when parental consent can be appropriately waived (for example, participant advocate)</p>
Risk of exacerbation of inequities impedes justice	<p>How do we value the knowledge and lived experiences of AYA equally with clinical and professional expertise, considering the power imbalances between researchers and AYA?</p> <p>How can inclusive design and recruitment strategies, specific to AYA, be ensured across all research phases?</p> <p>How do we address the impact of social inequalities and structural issues for AYA on mental health?</p>	<p>Determine where in the DHI design cycle (for example, planning, development, testing or implementation) AYA can contribute meaningfully</p> <p>Over-recruit AYA from minoritized backgrounds, for example, through community-based partnerships and key trusted AYA in the community</p> <p>Use frameworks of diverse, equitable and inclusive design, considering cultural adaptation, diverse representation, internet access and digital literacy⁶⁴. Provide AYA with socially complex needs free access to evidence-based DHIs</p> <p>Provide DHI support that also assists with tackling social needs, for example, social support, financial stability and education</p>

make the decision as to whether they want to involve their guardian. To determine when this is warranted, both parents and AYA should be involved in these decisions, for example, through advisory boards and co-creation sessions.

Institutional research boards need clearer guidance on when it is appropriate to waive parental consent in DHI research, when alternative models of consent may be sought, and what protections can ensure adolescents’ rights and safety. The majority of AYA seem to be able to comprehend the purpose and nature of research, research risks and benefits, and understanding increases with age⁵⁷. Parental consent practices vary internationally. In the USA, Poland and Portugal, for example, consent is required for adolescents up to the age of 18 years. By contrast, in Hungary and the Netherlands, AYA can independently consent from the age of 16, and in Croatia and the Czech Republic, they can consent from the age of 12 (ref. 57). These differences highlight the arbitrary nature of legal age thresholds and show that more flexibility in consent practices is needed.

Adolescents could be encouraged to ‘opt in’ for guardian consent should they choose to involve their caregiver, as successfully done in our own work⁵⁸. In addition, proxy alternatives, such as assent combined with oversight by a trusted adult or youth worker, or minor self-consent may be allowed in lower-risk contexts (for example, when accessing a DHI with previously demonstrated effectiveness and a low risk profile)⁵⁹. Research consent rules should probably parallel local and regional laws or policies for medical and mental health consent. For example, in the state of Illinois, youth aged 12 are legally permitted to consent for mental health treatment except medications. Institutional research boards in Illinois could approve digital mental health research in parallel with state law for treatment. However, objections to youth-led consent include concerns about minors’ cognitive ability, risks of exposure to harmful interventions and parental rights⁶⁰. Importantly, although we provide examples of methods successfully used in research contexts, the field still

lacks data to determine which strategies are better at protecting AYA’s safety while also promoting their autonomy, which should be examined in future research.

The risk of exacerbation of injustice

In DHI research, there are additional risks that compromise the bioethical principle of justice compared with non-digital research. First, DHI research has a history of exclusive design and recruitment strategies across various research phases^{61–63}. For example, 70% of youth DHI studies are conducted with university students as the primary sample. Further, most studies do not consider inclusive design practices related to access to the technology necessary to benefit from DHIs, culturally robust content or representation of diverse identities⁶⁴. Inclusive design guides have been developed for DHI in general, such as the US Agency of Healthcare Research and Quality⁶⁵. However, they are not specific to AYA or the use of AI, and it is unclear whether they are commonly used. This lack of guidance warrants the development of new inclusivity standards.

A crucial component of these standards is youth engagement in the design cycle of DHIs (that is, planning, development and testing implementation). In non-DHI research, youth engagement positively impacts research design, recruitment, data collection, analysis and dissemination⁶⁶. It also enhances the youth-friendliness and validity of research, improves the usability of tools, and increases the accessibility of consent forms and questionnaires⁶⁶. Although guidelines exist for involving individuals at all stages of DHI research—ranging from design studies and prototyping to controlled trials and implementation⁶⁵—there is variation in how ‘meaningful participation’ is interpreted and applied across these stages. Moreover, achieving meaningful engagement is not without challenges. For example, inexperienced youth facilitators may compromise data quality, and AYA might interpret findings through the lens of their personal experiences, potentially limiting generalizability⁶⁶.

The ambiguity surrounding youth participation complicates efforts to promote justice throughout the DHI research process and exacerbates the problem of ‘epistemic injustice’: when the lived experiences of AYA with health challenges and technology use are undervalued in favor of clinical, academic or technical expertise. This issue becomes even more pronounced when AI is involved because AI is often dealt with at the expert level, and the role of young people in AI governance is still unclear⁶⁷.

Further, there are concerns that using AI in DHIs could exacerbate existing health and social inequities. On the one hand, AI could enhance health equity by its potential to help overcome human decision-making, which is often clouded by biases⁶⁸. On the other hand, AI often shows gender biases⁶⁹ and is not able to uphold different cultural values⁷⁰. AI and machine learning have exhibited ineffective clinical recommendation and perpetuations of racism⁷¹. For example, the National Eating Disorders Association chatbot, developed to provide support to AYA with restrictive eating disorders, dispensed dieting and weight-loss advice instead of evidence-based information⁷². In addition, LLMs’ responses to Black individuals consistently showed lower empathy than for any other demographic group when asked to act as a peer support therapist⁷³. These findings highlight a critical issue: AI tools risk perpetuating mental health inequities if they are trained on historical data embedded with social biases.

To promote the bioethical principle of justice, researchers should strive to over-recruit AYA from minoritized backgrounds to ensure that DHIs are tailored to their needs. To further ensure equal access, providing AYA from minoritized groups with free access to DHIs—regardless of their eligibility for research participation—could be considered as a strategy to mitigate disparities in DHI outcomes, as done in some of our previous work⁷⁴. In addition, DHIs should focus not only on clinical symptoms but also on the factors related to the social determinants of health, such as social support, financial stability for AYA and education⁷⁵. Youth engagement, inclusive design standards and equitable access are essential to promoting justice.

Why we should view AYA as experts on an unknowable topic

On the basis of the ethical challenges identified above, we provide preliminary guidance for researchers in Table 1. Although we recognize this space requires a high degree of technology and data literacy and expertise in ethics, we argue that the lived experiences of AYA are needed for a transformation in the field. Crucially, through participatory methods, AYA can help us anticipate potential ethical dilemmas that could arise, and youth-specific values, in areas of research that might not be immediately obvious to researchers. Youth engagement is key to reforming youth mental health services and health policy⁷⁶ and has a strong link with outcomes, especially for marginalized youth⁷⁷. Emerging evidence shows that AYA themselves want to be involved in DHI research (including AI)⁷⁸. Yet it is unclear what, if any, voice AYA have had in the development of ethical guidelines for research in general⁶⁰, and particularly for DHI research. Therefore, our first recommendation for developing sound ethical guidance is that AYA should participate in co-designing ethical guidance.

The need for participation is particularly crucial for AYA with minoritized identities and experiences who have been repeatedly and systemically excluded from designs of technologies⁷⁹, DHI development⁶² and ethical guidance design⁸⁰. AYA who identify with one or more minoritized identities or marginalized experiences (that is, socially complex needs) should be over-represented as AYA experts invited to the planning and evaluation table to ensure that ethical guidance is better designed for generalizability⁸¹.

To over-recruit young people from marginalized backgrounds, several strategies could be considered. First, building partnerships with community centers where young people from marginalized backgrounds engage in activities. For example, we have conducted

co-design sessions with AYA embedded into already planned activities in these spaces (for example, personal development workshops)⁸². Similarly, the AYA authors of this paper became involved in AI and mental health research through a partnership that the lead author formed with a community organization (Be Ubuntu) for young people in socially challenging circumstances. Recruitment can also be conducted by AYA themselves. For example, the AYA authors of this paper conduct mental health research in their neighborhoods through word-of-mouth recruitment, which leads to reaching youth who may not be as active online.

Second, culturally matched research personnel, linguistically appropriate materials—including whether parents or guardians speak a different language than their AYA—and a personalized recruitment process with endorsement from trusted community figures can improve engagement. For example, in a DHI trial for racially and ethnically minoritized groups exposed to race-based stressors, we designed recruitment materials that depicted members of the communities we intended to reach (for example, African American AYA) and used language directly addressing issues relevant to these groups (for example, have you ever experienced discrimination or stress related to being a person of color?), and key community members (for example, the director of the community wellness center) promoted the intervention. This led to meeting our initial recruitment goal in just 1 month and doubling our sample size to meet the demand for this program⁷⁴.

Finally, ensuring adequate remuneration and offering DHIs for free are also essential for accessibility and fairness. In the DHI trial for racially and ethnically minoritized groups exposed to race-based stressors, in addition to being financially compensated, all participants eventually received free subscriptions to the DHI regardless of whether they were randomized to the intervention or control group⁷⁴. Previously developed remuneration guidelines can serve as a reference for researchers interested in engaging AYA in these programs⁸³.

Methods that could be used to develop new ethical guiding principles together with AYA

In Table 2 we provide an overview of participatory methods for research with AYA. Human-centered design (HCD) is a group of mixed-methods approaches for developing a product (for example, DHI) to be used by individuals in their daily lived experience^{84,85}. However, HCD methods have limitations, including limited guidance for involving minoritized AYA and averting biases, and a limited focus on ethical and social issues (such as the risk of AI amplifying inequalities)⁸⁶. In line with calls for HCD methodologies to be rethought and more purposefully grounded in anti-racism, justice, equity, diversity and inclusion^{61,86}, using inclusive methodologies in the development, evaluation and application of ethical guiding principles must be done purposefully. For example, we may need to combine HCD with youth participatory action research approaches⁸⁷, which involves enquiry based on youths’ lived experiences, participatory collaboration in research and active intervention to improve the lives of youth and their communities⁸⁸. In addition, ethical values could be explicitly integrated using methods such as value-sensitive design, a framework for considering the values of different stakeholders when designing systems, to uncover value tensions (such as between AYA, developers and parents)⁸⁹. These approaches have been scarcely used in DHIs⁸⁶.

Examples of inclusive methodologies

With regard to AI, examples of inclusive methodologies are emerging. One study used a community-based natural language processing method to inform machine-learning-based tactics to increase the recruitment of trial participants who identify as racially and ethnically minoritized⁹⁰. A group model for engaging underrepresented groups in research was developed by both researchers and participants, which involved establishing a stakeholder board and various co-creation sessions.

Table 2 | Examples of methods to develop new ethical guiding principles together with AYA

Example	Description	Benefits	Challenges
Focus groups	A small group of people discussing a topic or issues defined by a researcher.	Gathering new ideas, gain an in-depth understanding of social issues	Power imbalances ⁹¹ between youth and researchers, difficult to voice individual opinions, social biases (for example, conformity) ⁹⁴
Co-design (also called HCD or user-centered design)	Approach that grounds product development in information collected about end users of products ⁸⁴ . Uses a range of activities, such as role playing, design sketching, mind mapping storytelling, personas, user journeys, prototyping, thinking aloud protocol and cognitive walkthroughs ⁸⁵	Co-design can improve DHIs through integrating the user's real-life experiences and needs. By giving users a say in the design process, it is regarded an ethical and pragmatic approach ⁸⁵	Choose between in-person and virtual Power imbalances and Zoom fatigue, if the latter May have a limited reach (for example, no guidance for involving minoritized AYA) Limited guidance on averting biases unknown to users and researchers, and a narrow contextual focus
Youth participatory action research (AYA as co-researchers)	AYA are involved in the research process as collaborators and co-researchers. Youth participate in the process and receive training in subjects such as methods, ethics, data collection, analysis and dissemination. They are also included in evaluation and reflection cycles throughout the research	Empowering for youth fosters agency and self-efficacy Avoid 'listening' to AYA in a tokenistic manner and interpreting their views in their absence ⁸⁷ Designed to enable justice and social change ⁸⁷	Requires extensive training and support; maintaining youth engagement can be challenging Youth may not really be given any new tools to improve their lives
Youth community advisory board	Asking AYA what they need, having them drive research questions rather than vice versa	AYA ownership of the process ensures relevance to youth needs, builds trust and engagement	Time consuming, requires sustainable funding sources and long-term commitment
Ethnographic approaches ⁹⁴	This includes methods such as photovoice and videovoice	Captures lived experiences in context, good for capturing marginalized voices	Time consuming AYA may not act naturally
Digital participatory action research ⁹⁵	Using technology as a core tool within youth participatory action research. For example, AYA create podcasts, design videos or map the neighborhood ⁹⁵	AYA learn and apply new digital tools or techniques that increase their confidence and sense of ownership	Time consuming May be impacted by a digital divide (not all children have equal digital skills)
Value-sensitive design	A framework for translating abstract values, such as well-being, into design features, so that AYA experience those values in practical and personally relevant ways. Can include co-design, participatory design and stakeholder analysis	Method to explicitly integrate values of ethical importance (for example, privacy, autonomy and fairness) into the design process	Resource and time intensive Can be difficult to integrate in practice because it can involve abstract and sometimes conflicting values ⁹⁶

Other research used a participatory design approach to co-create an ethics framework for AI in education based on moving past deficit-oriented views of youth toward understanding and re-imagining how AI can be a useful collaborator ('how can I as a researcher support ideal relationships in classrooms, and what possibilities for AI does that open up?')⁹¹. In our own work, we used a combination of HCD and participatory research to understand the wishes of AYA for mental health apps, providing them with training to co-analyze their own data⁸². In another study, AYA participated in designing a suicide prevention social media campaign, using co-design methods the team adapted for young people, such as providing guidance, examples, simple explanations and fewer time restrictions⁹². These approaches could serve as examples for the development of ethical principles in this field.

Building on current participatory methods, we can ensure a paradigm shift toward a youth-focused DHI ethical guidance co-created with diverse AYA. However, achieving this shift will require sustained time, funding and commitment to participatory research with young people⁹³ and a shared understanding of the critical need for ethical guidance in AI-based DHI.

Conclusion

Current ethical guidance for DHI for AYA has not been developed from a youth-centered perspective and risks both over- and under-protection in DHI research. In the face of a behavioral health crisis for AYA, time is of the essence to define and iterate new ethical guiding principles for DHI research with these dynamic age groups. Lessons from youth participatory action research⁸⁸ and HCD methodologies⁶¹ should be used in complement to the grounding application

of foundational research principles of respect for persons, beneficence and justice^{41,42}. Interdisciplinary collaboration (for example, bioethicists, members of institutional review boards, behavioral health specialists and caregivers) is necessary to co-design new guidance for DHI research. Most importantly, AYA themselves, particularly those from minoritized backgrounds, must guide the formation and subsequent evaluations and iterations of ethical guiding principles for AYA. Their insights into an unknowable new frontier are the most important to both protect and empower AYA in a growing body of DHI research.

References

1. Chavira, D. A., Ponting, C. & Ramos, G. The impact of COVID-19 on child and adolescent mental health and treatment considerations. *Behav. Res. Ther.* **157**, 104169 (2022).
2. Lee, C. M., Cadigan, J. M. & Rhew, I. C. Increases in loneliness among young adults during the COVID-19 pandemic and association with increases in mental health problems. *J. Adolesc. Health* **67**, 714–717 (2020).
3. Cotton, N. K. & Shim, R. S. Social determinants of health, structural racism, and the impact on child and adolescent mental health. *J. Am. Acad. Child Adolesc. Psychiatry* **61**, 1385–1389 (2022).
4. Copeland, W. E. et al. Increase in untreated cases of psychiatric disorders during the transition to adulthood. *Psychiatr. Serv.* **66**, 397–403 (2015).
5. Stroud, C., Walker, L. R., Davis, M. & Irwin, C. E. Jr. Investing in the health and well-being of young adults. *J. Adolesc. Health* **56**, 127–129 (2015).

6. Lehtimäki, S., Martić, J., Wahl, B., Foster, K. T. & Schwalbe, N. Evidence on digital mental health interventions for adolescents and young people: systematic overview. *JMIR Ment. Health* **8**, e25847 (2021).
7. Ferrante, M., Esposito, L. E. & Stoeckel, L. E. From palm to practice: prescription digital therapeutics for mental and brain health at the National Institutes of Health. *Front. Psychiatry* **15**, 1433438 (2024).
8. Wright, M., Reitegger, F., Cela, H., Papst, A. & Gasteiger-Klicpera, B. Interventions with digital tools for mental health promotion among 11–18 year olds: a systematic review and meta-analysis. *J. Youth Adolesc.* **52**, 754–779 (2023).
9. Chen, T., Ou, J., Li, G. & Luo, H. Promoting mental health in children and adolescents through digital technology: a systematic review and meta-analysis. *Front. Psychol.* **15**, 1356554 (2024).
10. Piers, R., Williams, J. M. & Sharpe, H. Review: can digital mental health interventions bridge the ‘digital divide’ for socioeconomically and digitally marginalised youth? A systematic review. *Child Adolesc. Ment. Health* **28**, 90–104 (2023).
11. Chen, B. et al. Comparative effectiveness and acceptability of internet-based psychological interventions on depression in young people: a systematic review and network meta-analysis. *BMC Psychiatry* **25**, 321 (2025).
12. Jahedi, F., Fay Henman, P. W. & Ryan, J. C. Personalization in digital psychological interventions for young adults. *Int. J. Hum. Comput. Interact.* **40**, 2254–2264 (2024).
13. Hightow-Weidman, L. B., Horvath, K. J., Scott, H., Hill-Rorie, J. & Bauermeister, J. A. Engaging youth in mHealth: what works and how can we be sure? *MHealth* **7**, 23 (2021).
14. Zhu, S., Wang, Y. & Hu, Y. Facilitators and barriers to digital mental health interventions for depression, anxiety, and stress in adolescents and young adults: scoping review. *J. Med. Internet Res.* **27**, e62870 (2025).
15. Götzl, C. et al. Artificial intelligence-informed mobile mental health apps for young people: a mixed-methods approach on users’ and stakeholders’ perspectives. *Child Adolesc. Psychiatry Ment. Health* **16**, 86 (2022).
16. Hornstein, S., Zantvoort, K., Lueken, U., Funk, B. & Hilbert, K. Personalization strategies in digital mental health interventions: a systematic review and conceptual framework for depressive symptoms. *Front. Digit. Health* **5**, 1170002 (2023).
17. Li, H., Zhang, R., Lee, Y.-C., Kraut, R. E. & Mohr, D. C. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *npj Digit. Med.* **6**, 236 (2023).
18. Heinz, M. V. et al. Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI* <https://doi.org/10.1056/Aloa2400802> (2025).
19. Gutierrez, G., Stephenson, C., Eadie, J., Asadpour, K. & Alavi, N. Examining the role of AI technology in online mental healthcare: opportunities, challenges, and implications, a mixed-methods review. *Front. Psychiatry* **15**, 1356773 (2024).
20. Ehtemam, H. et al. Role of machine learning algorithms in suicide risk prediction: a systematic review-meta analysis of clinical studies. *BMC Med. Inform. Decis. Mak.* **24**, 138 (2024).
21. Atmakuru, A. et al. Artificial intelligence-based suicide prevention and prediction: a systematic review (2019–2023). *Inf. Fusion* **114**, 102673 (2025).
22. Goodin, P., Van Kolfschooten, H. B. & Centola, F. *Artificial Intelligence in Mental Healthcare* (Mental Health Europe, 2025).
23. Young, J. et al. The role of AI in peer support for young people: a study of preferences for human- and AI-generated responses. In *Proc. CHI Conference on Human Factors in Computing Systems (CHI ’24)* 1006 (Association for Computing Machinery, 2024).
24. Schaaff, C. et al. Youth perspectives on generative AI and its use in health care. *J. Med. Internet Res.* **27**, e72197 (2025).
25. Smith, K. A. et al. Digital mental health: challenges and next steps. *BMJ Ment. Health* **26**, e300670 (2023).
26. Laacke, S., Mueller, R., Schomerus, G. & Salloch, S. Artificial intelligence, social media and depression. A new concept of health-related digital autonomy. *Am. J. Bioeth.* **21**, 4–20 (2021).
27. Seroussi, B. & Zablit, I. Implementation of digital health ethics: a first step with the adoption of 16 European ethical principles for digital health. *Stud. Health Technol. Inform.* **310**, 1588–1592 (2024).
28. Nebeker, C., Gholami, M., Kareem, D. & Kim, E. Applying a digital health checklist and readability tools to improve informed consent for digital health research. *Front. Digit. Health* **3**, 690901 (2021).
29. Nebeker, C., Torous, J. & Bartlett Ellis, R. J. Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Med.* **17**, 137 (2019).
30. Psihogios, A. M., King-Dowling, S., Mitchell, J. A., McGrady, M. E. & Williamson, A. A. Ethical considerations in using sensors to remotely assess pediatric health behaviors. *Am. Psychol.* **79**, 39–51 (2024).
31. Chng, S. Y. et al. Ethical considerations in AI for child health and recommendations for child-centered medical AI. *npj Digit. Med.* **8**, 152 (2025).
32. Mello, M. M. & Cohen, I. G. Regulation of health and health care artificial intelligence. *JAMA* **333**, 1769–1770 (2025).
33. Tavery, T. Regulating AI in mental health: ethics of care perspective. *JMIR Ment. Health* **11**, e58493 (2024).
34. van Kolfschooten, H. & van Oirschot, J. The EU Artificial Intelligence Act (2024): implications for healthcare. *Health Policy* **149**, 105152 (2024).
35. Dignum, V., Penagos, M., Pigmans, K. & Vosloo, S. *Policy Guidance on AI for Children* (UNICEF, 2021); <https://www.unicef.org/innocenti/reports/policy-guidance-ai-children>
36. Atabey, A. et al. *Children & AI Design Code* (5Rights Foundation); <https://5rightsfoundation.com/children-and-ai-code-of-conduct/>
37. *Artificial Intelligence for Children* (World Economic Forum, 2022); https://www3.weforum.org/docs/WEF_Artificial_Intelligence_for_Children_2022.pdf
38. *APA’s AI Tool Guide for Practitioners* (APA, 2024); <https://www.apaservices.org/practice/business/technology/tech-101/evaluating-artificial-intelligence-tool>
39. *Urging the Federal Trade Commission to Take Action on Unregulated AI* (APA, 2025); <https://www.apaservices.org/advocacy/news/federal-trade-commission-unregulated-ai>
40. National Academy of Medicine. *An Artificial Intelligence Code of Conduct for Health and Medicine: Essential Guide for Aligned Action* (National Academies Press, 2025); <https://nap.nationalacademies.org/catalog/29087/an-artificial-intelligence-code-of-conduct-for-health-and-medicine>
41. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research: Appendix* (Department of Health, Education, and Welfare, National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1978).
42. Cassell, E. J. The principles of the Belmont report revisited. How have respect for persons, beneficence, and justice been applied to clinical medicine? *Hastings Cent. Rep.* **30**, 12–21 (2000).
43. McNeilly, E. A. et al. Adolescent social communication through smartphones: linguistic features of internalizing symptoms and daily mood. *Clin. Psychol. Sci.* **11**, 1090–1107 (2023).
44. Stiles-Shields, C. et al. mHealth uses and opportunities for teens from communities with high health disparities: a mixed-methods study. *J. Technol. Behav. Sci.* **8**, 282–294 (2022).

45. Movahed, S. V. & Martin, F. Ask me anything: exploring children's attitudes toward an age-tailored AI-powered chatbot. Preprint at <https://arxiv.org/abs/2502.14217> (2025).
46. Duffy, C. 'There are no guardrails.' This mom believes an AI chatbot is responsible for her son's suicide. *CNN Business* <https://edition.cnn.com/2024/10/30/tech/teen-suicide-character-ai-lawsuit/> (2024).
47. Kurian, N. 'No, Alexa, no!': designing child-safe AI and protecting children from the risks of the 'empathy gap' in large language models. *Learn. Media Technol.* <https://doi.org/10.1080/17439884.2024.2367052> (2024).
48. Hadar-Shoval, D., Asraf, K., Mizrahi, Y., Haber, Y. & Elyoseph, Z. Assessing the alignment of large language models with human values for mental health integration: cross-sectional study using Schwartz's theory of basic values. *JMIR Ment. Health* **11**, e55988 (2024).
49. Mansfield, K. L. et al. From social media to artificial intelligence: improving research on digital harms in youth. *Lancet Child Adolesc. Health* **9**, 194–204 (2025).
50. Xu, Y., Prado, Y., Severson, R. L., Lovato, S. & Cassell, J. in *Handbook of Children and Screens* 611–617 (Springer Nature, 2025).
51. Wong, Q. California lawmakers consider possible dangers of AI chatbots. *Los Angeles Times* (9 April 2025); <https://www.govtech.com/artificial-intelligence/california-lawmakers-consider-possible-dangers-of-ai-chatbots>
52. Sanderson, C., Douglas, D. & Lu, Q. Implementing responsible AI: tensions and trade-offs between ethics aspects. In *Proc. 2023 International Joint Conference on Neural Networks (IJCNN)* 1–7 (IEEE, 2023).
53. "Without It, I Wouldn't Be Here Today" (Born This Way Foundation, 2025); <https://bornthisway.foundation/research/lgbtq-young-people-online-communities-2025/>
54. Mustanski, B. Ethical and regulatory issues with conducting sexuality research with LGBT adolescents: a call to action for a scientifically informed approach. *Arch. Sex. Behav.* **40**, 673–686 (2011).
55. Chleider, J. L., Dobias, M., Sung, J., Mumper, E. & Mullarkey, M. C. Acceptability and utility of an open-access, online single-session intervention platform for adolescent mental health. *JMIR Ment. Health* **7**, e20513 (2020).
56. Liverpool, S. et al. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. *J. Med. Internet Res.* **22**, e16317 (2020).
57. Samdal, O. et al. Encouraging greater empowerment for adolescents in consent procedures in social science research and policy projects. *Obes. Rev.* **24**, e13636 (2023).
58. Stiles-Shields, C. et al. Digital mental health screening, feedback, and referral system for teens with socially complex needs: protocol for a randomized controlled trial integrating the Teen Assess, Check, and Heal System into pediatric primary care. *JMIR Res. Protoc.* **14**, e65245 (2025).
59. Hein, I. M. et al. Informed consent instead of assent is appropriate in children from the age of twelve: policy implications of new findings on children's competence to consent to clinical research. *BMC Med. Ethics* **16**, 76 (2015).
60. Crane, S. & Broome, M. E. Understanding ethical issues of research participation from the perspective of participating children and adolescents: a systematic review. *Worldviews Evid. Based. Nurs.* **14**, 200–209 (2017).
61. Stiles-Shields, C. et al. A call to action: using and extending human-centered design methodologies to improve mental and behavioral health equity. *Front. Digit. Health* **4**, 848052 (2022).
62. Ramos, G. & Chavira, D. A. Use of technology to provide mental health care for racial and ethnic minorities: evidence, promise, and challenges. *Cogn. Behav. Pract.* **29**, 15–40 (2022).
63. Whitehead, L., Talevski, J., Fatehi, F. & Beauchamp, A. Barriers to and facilitators of digital health among culturally and linguistically diverse populations: qualitative systematic review. *J. Med. Internet Res.* **25**, e42719 (2023).
64. Ramos, G., Ponting, C., Labao, J. P. & Sobowale, K. Considerations of diversity, equity, and inclusion in mental health apps: a scoping review of evaluation frameworks. *Behav. Res. Ther.* **147**, 103990 (2021).
65. Hatef, E., Hudson Scholle, S., Buckley, B., Weiner, J. P. & Austin, J. M. Development of an evidence- and consensus-based Digital Healthcare Equity Framework. *JAMIA Open* **7**, ooa136 (2024).
66. McCabe, E. et al. Youth engagement in mental health research: a systematic review. *Health Expect.* **26**, 30–50 (2023).
67. Zhou, B. et al. *Youth Perspectives and Recommendations for the United Nations's High-Level Advisory Board on Artificial Intelligence*. Foundational Papers of the United Nations High-Level Advisory Body on AI (United Nations, 2023).
68. Gurevich, E., El Hassan, B. & El Morr, C. Equity within AI systems: what can health leaders expect? *Healthc. Manage. Forum* **36**, 119–124 (2023).
69. Kotek, H., Dockum, R. & Sun, D. Gender bias and stereotypes in large language models. In *Proc. ACM Collective Intelligence Conference 12–24* (Association for Computing Machinery, 2023).
70. Kharchenko, J., Roosta, T., Chadha, A. & Shah, C. How well do LLMs represent values across cultures? Empirical analysis of LLM responses based on Hofstede cultural dimensions. Preprint at <https://arXiv:2406.14805> (2024).
71. Karnik, N. S., Afshar, M., Churpek, M. M. & Nunez-Smith, M. Structural disparities in data science: a prolegomenon for the future of machine learning. *Am. J. Bioeth.* **20**, 35–37 (2020).
72. Wells, K. An eating disorders chatbot offered dieting advice, raising fears about AI in health. *Shots* <https://www.npr.org/sections/health-shots/2023/06/08/1180838096/an-eating-disorders-chatbot-offered-dieting-advice-raising-fears-about-ai-in-health> (2023).
73. Gabriel, S., Puri, I., Xu, X., Malgaroli, M. & Ghassemi, M. in *Findings of the Association for Computational Linguistics: EMNLP 2024* (eds Al-Onaizan, Y. et al.) 2206–2221 (Association for Computational Linguistics, 2024).
74. Ramos, G. et al. App-based mindfulness meditation for people of color who experience race-related stress: protocol for a randomized controlled trial. *JMIR Res. Protoc.* **11**, e35196 (2022).
75. Thomas et al. Leveraging data and digital health technologies to assess and impact social determinants of health (SDoH): a state-of-the-art literature review. *Online J. Public Health Inform.* **13**, E14 (2021).
76. McGorry, P. D. et al. The Lancet Psychiatry Commission on youth mental health. *Lancet Psychiatry* **11**, 731–774 (2024).
77. Bailey, K. et al. Benefits, barriers and recommendations for youth engagement in health research: combining evidence-based and youth perspectives. *Res. Involv. Engagem.* **10**, 92 (2024).
78. Zidar, T., Morrow, E. M. & Stockley, R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: a systematic scoping review and agenda for design justice. *Health Expect.* **24**, 1072–1124 (2021).
79. Schueller, S. M., Hunter, J. F., Figueroa, C. & Aguilera, A. Use of digital mental health for marginalized and underserved populations. *Curr. Treat. Options Psychiatry* **6**, 243–255 (2019).
80. Juntunen, C. L. et al. Centering equity, diversity, and inclusion in ethical decision-making. *Prof. Psychol. Res. Pr.* **54**, 17–27 (2023).

81. Stiles-Shields, C., Ramos, G., Ortega, A. & Psihogios, A. M. Increasing digital mental health reach and uptake via youth partnerships. *npj Ment. Health Res.* **2**, 9 (2023).
82. Figueroa, C. et al. Skepticism and excitement when co-designing just-in-time mental health apps with minoritized youth. Preprint at SSRN <https://doi.org/10.2139/ssrn.5039034> (2024).
83. Afkinich, J. L. & Blachman-Demner, D. R. Providing incentives to youth participants in research: a literature review. *J. Empir. Res. Hum. Res. Ethics* **15**, 202–215 (2020).
84. Lyon, A. R., Brewer, S. K. & Areán, P. A. Leveraging human-centered design to implement modern psychological science: return on an early investment. *Am. Psychol.* **75**, 1067–1079 (2020).
85. Malloy, J., Partridge, S. R., Kemper, J. A., Braakhuis, A. & Roy, R. Co-design of digital health interventions with young people: a scoping review. *Digit. Health* **9**, 20552076231219117 (2023).
86. van Velsen, L., Ludden, G. & Grünloh, C. The limitations of user- and human-centered design in an eHealth context and how to move beyond them. *J. Med. Internet Res.* **24**, e37341 (2022).
87. Khawaja, J., Bagley, C. & Taylor, B. Breaking the silence: critical discussion of a youth participatory action research project. *JCPP Adv.* **4**, e12283 (2024).
88. Leman, A. M. et al. An interdisciplinary framework of youth participatory action research informed by curricula, youth, adults, and researchers. *J. Res. Adolesc.* **35**, e13007 (2024).
89. Figueroa, C., Marin, L., de Reuver, M. and Jaff, M. Translating well-being values into design features of social media platforms: a value sensitive design approach. Preprint at OSF https://osf.io/preprints/osf/wxtj7_v1 (2024).
90. Farmer, N. et al. Use of a community advisory board to build equitable algorithms for participation in clinical trials: a protocol paper for HoPeNET. *BMJ Health Care Inform.* **29**, e100453 (2022).
91. Chang, M. A. et al. Co-design partners as transformative learners: imagining ideal technology for schools by centering speculative relationships. In *Proc. CHI Conference on Human Factors in Computing Systems* 1–15 (ACM, 2024).
92. Thorn, P. et al. Developing a suicide prevention social media campaign with young people (the #chatsafe project): co-design approach. *JMIR Ment. Health* **7**, e17520 (2020).
93. Teixeira, S., Augsberger, A., Richards-Schuster, K. & Sprague Martinez, L. Participatory research approaches with youth: ethics, engagement, and meaningful action. *Am. J. Community Psychol.* **68**, 142–153 (2021).
94. Bevan et al. Practitioner review: co-design of digital mental health technologies with children and young people. *J. Child Psychol. Psychiatry* **61**, 928–940 (2020).
95. Cockerham, D. Participatory action research: building understanding, dialogue, and positive actions in a changing digital environment. *Educ. Technol. Res. Dev.* **72**, 2763–2791 (2024).
96. Gerdes, A. & Frandsen, T. F. A systematic review of almost three decades of value sensitive design (VSD): what happened to the technical investigations? *Ethics Inf. Technol.* **25**, (2023).

Acknowledgements

C.A.F. is supported through the ‘High Tech for a Sustainable Future’ capacity-building program of the 4TU Federation in the Netherlands. The funders had no role in the content of the manuscript. G.R. is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under award number T32HL166114. N.S.K.’s work on this project is supported, in part, by the J. Usha Raj Endowed Professorship at the University of Illinois Chicago. C.S.-S.’s work is supported in part by a grant from the National Institute of Mental Health (K08 MH125069). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or any of the institutions named here.

Author contributions

C.S.-S. drafted the initial version of the paper. C.A.F. led the development of subsequent drafts and facilitated discussion sessions among the authors. C.A.F., C.S.-S., E.E.A., G.R., N.S.K., A.M.P. and O.A. participated in at least one group discussion. P.B. and E.E. engaged in a dedicated session with C.A.F. and met with her individually to discuss their contributions, which they additionally provided in writing. C.A.F. also held a separate meeting with M.d.H. to incorporate their input. C.A.F., C.S.-S., E.E.A., G.R., N.S.K., A.M.P., O.A., P.B., E.E. and M.d.H. contributed to editing and approved the final version of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to C. A. Figueroa.

Peer review information *Nature Mental Health* thanks Marcello Ienca, Tamar Tavory and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature America, Inc. 2025