JMIR Neurotechnology

For research exploring how technologies (e.g. information technology, neural engineering, neural interfacing, clinical data science, robotics, eHealth/mHealth) can be applied in clinical neuroscience (e.g., neurology, neurosurgery, neuroradiology) to prevent, diagnose, and treat neurological disorders.
Volume 4 (2025) ISSN 2817-092X Editors-in-Chief: Pieter Kubben MD, PhD

Contents

Viewpoint

Original Papers

Transforming Perceptions: Exploring the Multifaceted Potential of Generative AI for People With Cognitive Disabilities

Dorit Hadar Souval[1], MA, PhD; Yuval Haber[2], MA, PhD; Amir Tal[3], MA, PhD; Tomer Simon[4], BSc, MPH, PhD; Tal Elyoseph[5], MA; Zohar Elyoseph[6,7], MA, PhD

[1]Max Stern Academic College of Emek Yezreel, Tel Adashim, Israel

[2]Interdisciplinary Studies Unit, PhD Program of Hermeneutics and Cultural Studies, Bar-Ilan University, Ramat Gan, Israel

[3]Tel Aviv University, Tel Aviv, Israel

[4]Microsoft Israel R&D Center, Tel Aviv, Israel

[5]Faculty of Social Welfare and Health Sciences, University of Haifa, Haifa, Israel

[6]Imperial College London, Fulham Palace Rd, London, United Kingdom

[7]Faculty of Education, University of Haifa, Haifa, Israel

Corresponding Author:
Zohar Elyoseph, MA, PhD
Imperial College London, Fulham Palace Rd, London, United Kingdom

Abstract

Background: The emergence of generative artificial intelligence (GenAI) presents unprecedented opportunities to redefine conceptions of personhood and cognitive disability, potentially enhancing the inclusion and participation of individuals with cognitive disabilities in society.

Objective: We aim to explore the transformative potential of GenAI in reshaping perceptions of cognitive disability, dismantling societal barriers, and promoting social participation for individuals with cognitive disabilities.

Methods: This study is a critical review of current literature in disability studies, artificial intelligence (AI) ethics, and computer science, integrating insights from disability theories and the philosophy of technology. The analysis focused on 2 key aspects: GenAI as a social mirror reflecting societal values and biases, and GenAI as a cognitive partner for individuals with cognitive disabilities.

Results: This paper proposes a theoretical framework for understanding the impact of GenAI on perceptions of cognitive disability. It introduces the concepts of GenAI as a "social mirror" that reflects and potentially amplifies societal biases and as a "cognitive copilot" providing personalized assistance in daily tasks, social interactions, and environmental navigation. This paper also presents a novel protocol for developing AI systems tailored to the needs of individuals with cognitive disabilities, emphasizing user involvement, ethical considerations, and the need to address both the opportunities and challenges posed by GenAI.

Conclusions: Although GenAI has great potential for promoting the inclusion and empowerment of individuals with cognitive disabilities, realizing this potential requires a change in societal attitudes and development practices. This paper calls for interdisciplinary collaboration and close partnership with the disability community in the development and implementation of GenAI technologies. Realizing the potential of GenAI for promoting the inclusion and empowerment of individuals with cognitive disabilities requires a multifaceted approach. This involves a shift in societal attitudes, inclusive AI development practices that prioritize the needs and perspectives of the disability community, and ongoing interdisciplinary collaboration. This paper emphasizes the importance of proceeding with caution, recognizing the ethical complexities and potential risks alongside the transformative possibilities of GenAI technology.

(JMIR Neurotech 2025;4:e64182) doi:10.2196/64182

KEYWORDS

generative artificial intelligence; cognitive disability; social participation; AI ethics; assistive technology; cognitive disorder; societal barriers; social inclusion; disability study; social mirror; cognitive partner; empowerment; user involvement; GenAI; artificial intelligence; neurotechnology; neuroinformatics; digital health; health informatics; neuroscience; mental health; computer science; machine learning

Introduction

In the era of generative artificial intelligence (GenAI), traditional notions of personhood and normality are being challenged [1-4]. Technological advances are blurring the boundaries between human and machine capabilities, offering an opportunity to expand the limits of social inclusion and promote change in attitudes toward people with disabilities [1]. As artificial intelligence (AI) systems demonstrate increasingly sophisticated cognitive abilities, they prompt us to reconsider what qualities define personhood and human intelligence. This paper examines the potential of GenAI to disrupt limiting conceptions of morality and humanity, focusing on the implications of GenAI for the social status of people with cognitive disabilities. This paper also proposes a practical toolkit for GenAI development and engineering professionals—product managers, data scientists, and developers—to help incorporate these insights into their work.

Cognitive disability refers to a wide range of impairments affecting cognitive functions such as learning, problem-solving, judgment, communication, and social interaction [5]. Examples of cognitive disabilities include intellectual disability, attention-deficit/hyperactivity disorder, autism spectrum disorders, specific learning disabilities (such as dyslexia), and brain injuries (such as traumatic brain injury or stroke) [5-7]. It is important to emphasize the variety of individuals with cognitive disabilities, each one possessing a unique combination of strengths, impairments, and potential, which means that cognitive disabilities require personalized approaches to intervention. While recognizing the diverse nature of cognitive disabilities and the need for tailored solutions, this paper focuses on the general potential of GenAI to improve the lives of people across the spectrum of cognitive disabilities.

Engaging with the integration of GenAI and individuals with cognitive disabilities is a new direction in the use of technology in the field of disability. The potential for AI to support and empower this population lies in its ability to perform cognitive tasks such as reasoning, planning, decision-making, and communication—areas that are challenging for people with cognitive disabilities [8-10]. The ability of AI to remove barriers and open new paths for inclusive and equitable participation makes it especially relevant for this population [11]. An in-depth analysis of this ability requires examining the philosophical and ethical implications of AI for conceptions of humanity and morality, questions that directly determine how society views and accommodates individuals with cognitive disabilities. These are fundamental inquiries into the nature of intelligence, personhood, consciousness, and human agency, which largely determine the degree of participation and inclusion for this group.

Personhood and AI: An Opportunity for Paradigm Shift

The concept of personhood, which emerged as a central topic in bioethical debates surrounding topics such as abortion, stem cell research, and euthanasia, has evolved into a complex and multifaceted construct that now spans multiple disciplines [12]. Inherently normative in nature, personhood involves value judgments and ethical considerations regarding how we ought to treat and perceive others rather than merely describing observable facts. Personhood is not rooted exclusively in our biology and experiences but in our essence and identity. This identity, however, is not formed in isolation; it is dynamically shaped in an intricate interaction between self-perception and the perception of others and interaction with them. Rosfort [13] argued that this conceptualization of personhood reveals its profoundly relational and social nature, demonstrating how identity and perception of self-worth are inextricably woven into interactions and the broader human context.

The concept of “personhood” has long served as a central criterion in bioethical discussions, determining which entities deserve moral consideration and rights [3]. As a result, this notion has also functioned as a mechanism of exclusion, denying basic rights and opportunities to those deemed cognitively “abnormal” [14].

For example, historically, people with cognitive disabilities were excluded from the public sphere and denied the right to make decisions for themselves [15,16]. Even today, despite significant progress in discourse and work based on the “social model” (an approach that views disability as created by societal barriers rather than by individual impairments alone) [17] and the “minority group model” (which recognizes people with disabilities as a marginalized minority group) [18], exclusion still exists in various aspects of life. People with cognitive disabilities still face barriers to accessing higher education and vocational training because of preconceived notions about their abilities [19]. Despite relevant skills, they have difficulties securing meaningful employment and career advancement opportunities because of social stigma and prejudice [20]. Participation in political or civic decision-making processes, such as voting or community involvement, is limited by discriminatory perceptions of the competence of individuals with cognitive disabilities [21]. They are also excluded from leisure, social, and cultural activities because of a lack of access or restrictive attitudes toward their participation [22].

These exclusion examples illustrate how, as a result of conceptualizing what constitutes a person of merit, individuals with cognitive disabilities are often excluded in the deepest and broadest ways from society. This mechanism is difficult to identify because it operates through our language and the most basic organized mechanisms of any society: law, health care system, education system, and more [23].

Breaking entrenched concepts and perceptions of personhood is challenging because they are deeply embedded in societal structures and norms, but emerging technologies are beginning to challenge these long held beliefs. GenAI offers an opportunity to challenge the definition of personhood perceptions by demonstrating skills previously considered unique to humans [1,4]. Although these capabilities are not yet perfect in AI, their very existence challenges the idea that such traits belong exclusively to the “normal” cognitive function of humans and that social participation is conditional on the presence of these abilities.

The revolutionary potential of GenAI invites us to reexamine the criteria for membership in the moral community and expand them beyond limiting standards. Instead of relying on a narrow model of "correct" cognitive abilities as a prerequisite for rights and participation in society [14], we may adopt, with the assistance of GenAI, a more inclusive view that recognizes human diversity and the inherent value of all individuals, regardless of their abilities [24]. By showcasing the potential of machines to exhibit complex cognitive traits, GenAI challenges the notion that certain abilities are essential for personhood and moral status. It initiates a discourse on the need to redefine our understanding of what it means to be human and to have moral worth, moving away from a focus on cognitive benchmarks and toward a more encompassing vision of human dignity and rights [1,4].

Although AI presents opportunities to challenge our understanding of personhood, there are legitimate concerns about its potential to exacerbate exclusion and narrow definitions of "normal" human cognition. The inherent biases in AI systems, stemming from their training data and algorithmic design [25-28], risk reinforcing and amplifying existing societal prejudices [29]. As AI increasingly influences decision-making processes in areas such as employment, health care, and criminal justice, there is a danger that it could lead to more stringent and narrow criteria for what constitutes "normal" human functioning. This could inadvertently heighten barriers for individuals with cognitive differences, further marginalizing them from full societal participation [30]. Moreover, as AI systems become more sophisticated in mimicking certain human cognitive abilities, there is a risk that societal expectations of human performance might be unrealistically elevated, potentially creating an even more exclusionary standard of "normal" [31]. Thus, while AI challenges our notions of personhood, it simultaneously risks entrenching and exacerbating existing forms of exclusion, highlighting the critical need for ethical AI development and deployment considering diverse human experiences and capabilities. In the following sections, we will explore 2 key areas where GenAI has the potential to drive significant change: GenAI as a social mirror and GenAI as a cognitive partner. These 2 domains highlight the multifaceted impact that GenAI can have on reshaping perceptions, removing barriers, and promoting participation of individuals with cognitive disabilities on the one hand, and exacerbating existing biases and exclusions in society on the other.

Generative AI as a Social Mirror: Opportunity and Challenge

Overview

Vallor's [32] conceptualization of AI as a societal mirror provides a compelling framework for understanding the role of AI in reflecting and potentially amplifying societal biases, particularly concerning cognitive disabilities. This mirror metaphor can be understood as follows: just as a physical mirror reflects the image of what stands before it, AI systems reflect the data, values, and biases present in the society that created them. However, unlike a simple reflection, AI systems can amplify and distort these reflections, much as a funhouse mirror might exaggerate certain features.

This mirror effect illuminates how AI systems, trained on biased data, risk perpetuating existing prejudices against individuals with cognitive differences. AI essentially learns from and then projects back the biases inherent in its training data, potentially reinforcing and spreading these biases further. Paradoxically, this same reflective quality presents a unique opportunity to identify and address longstanding societal biases, rendering implicit prejudices explicit and subject to scrutiny. By closely examining what the AI "reflects back" to us, we can gain insights into biases that might otherwise remain hidden or unacknowledged in society.

Vallor [32] posits that AI systems in general, and GenAI systems in particular, are not merely neutral technological tools but mirrors reflecting the values, norms, and biases prevalent in human society. Given that these systems are constructed upon data and content created by humans, they inherently risk replicating and perpetuating prejudices and discrimination against marginalized groups, including people with cognitive disabilities [27,33].

A study by Gadiraju et al [34] demonstrated this mirroring effect in action. They conducted 19 focus groups with 56 participants with various disabilities who interacted with a dialog model based on a large language model. The researchers found that the model frequently perpetuated harmful stereotypes and narratives about disability. For example, the model often fixated on physical disabilities, particularly wheelchairs, while neglecting other types of disabilities. It also tended to portray people with disabilities as passive, sad, and lonely, reinforcing the misconception that disability is inherently negative. Additionally, the model sometimes produced what participants referred to as "inspiration porn," objectifying people with disabilities as sources of inspiration for nondisabled people.

For example, if the information used to train AI systems contains stereotypical or derogatory expressions toward people with cognitive disabilities, there is a significant risk that these systems might "learn" to adopt discriminatory attitudes. The potential consequences are severe: AI systems could rank individuals with cognitive disabilities as having lower potential in employment or educational contexts, limit their access to certain services, or make biased decisions about them in critical areas such as insurance or credit [35].

When we look into the societal mirror reflected by AI, several possible human responses can be identified. One metaphorical response is "breaking the mirror," representing human resistance to AI use and the insights it presents [36]. While this approach attempts to avoid the uncomfortable truths AI exposes, it risks missing out on the potential benefits and insights AI can offer. Another metaphorical strategy is "cleaning the mirror," where humans attempt to eliminate biases through AI alignment processes [37]. This approach aims to create AI systems aligned with human values and intentions, striving for a bias-free environment. However, it risks producing an artificially sterile system that fails to reflect the complexities of human cognition and interaction, potentially making AI less relevant and less capable of addressing real-world complexities.

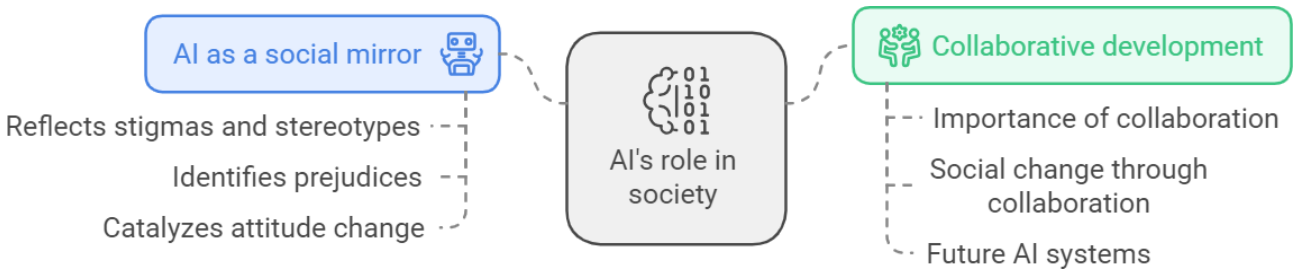
The third and most promising approach involves using reflection as a call to action in the real world. This method requires humans to acknowledge the biases reflected by AI and use this awareness as a catalyst for societal change. It demands active engagement and concrete actions from us as humans to address these issues, both in our AI systems and in society at large [38]. This approach recognizes that if such action is taken, over time, the reflection in the AI mirror itself can change, not as a result of erasing biases in the machine as in the second option, but as a consequence of real societal change that is then differently reflected in the AI mirror.

To implement this approach specifically within the realm of AI development and deployment, we must adopt advanced techniques and ensure inclusive human involvement. As contemporary AI systems increasingly incorporate vast datasets populated from the internet, traditional methods of addressing biases through direct data manipulation, such as the "datasheets" approach proposed by Gebru et al [39], while still valuable in certain contexts, have become more challenging to implement comprehensively. This shift has led to the adoption of complementary techniques that can handle the scale and complexity of modern AI systems such as self-supervised learning [40] and reward modeling [41]. Crucially, these techniques still require human decision-making at key junctures.

To truly address biases and create more equitable AI systems, particularly regarding cognitive disabilities, we must ensure that people with cognitive disabilities are actively involved in these decision-making processes. This collaborative approach aligns with our third strategy, emphasizing real-world action and societal change. By critically examining the biases revealed in AI outputs and involving diverse perspectives in the development process, we can work toward creating more inclusive AI systems. This approach not only helps in developing fairer algorithms and more representative models but also contributes to broader societal change [1,4]. In this way, the AI mirror becomes not just a reflection of our current culture, but a catalyst for the more inclusive society we aspire to create [16,42].

In conclusion, as illustrated in Figure 1, GenAI has the potential to promote social justice and shift perceptions regarding cognitive disabilities. To harness this potential, collaborative work and ongoing effort are required to embed values of accessibility, inclusion, and respect for diversity at the core of technological development. These steps can transform the "reflection in the mirror" into a positive and inclusive image for people with cognitive disabilities, potentially leading to broader societal changes in perception and inclusion.

Figure 1. GenAI as a social mirror: collaborative development for societal change. AI: artificial intelligence; GenAI: generative artificial intelligence.



While this mirror metaphor provides valuable insights, it is important to recognize its limitations. Vallor's conceptualization, though powerful, doesn't fully capture the multifaceted potential of AI, particularly for people with disabilities. It overlooks its capability to actively solve previously intractable problems and enhance accessibility. To provide a more comprehensive understanding, we must expand our view beyond the perception of AI as a mere reflective tool. In the following section, we propose considering AI not only as a mirror but also as a cognitive partner for people with disabilities, emphasizing its potential to actively support and empower individuals with cognitive differences in navigating the world.

Generative AI as a Cognitive Partner for People With Disabilities

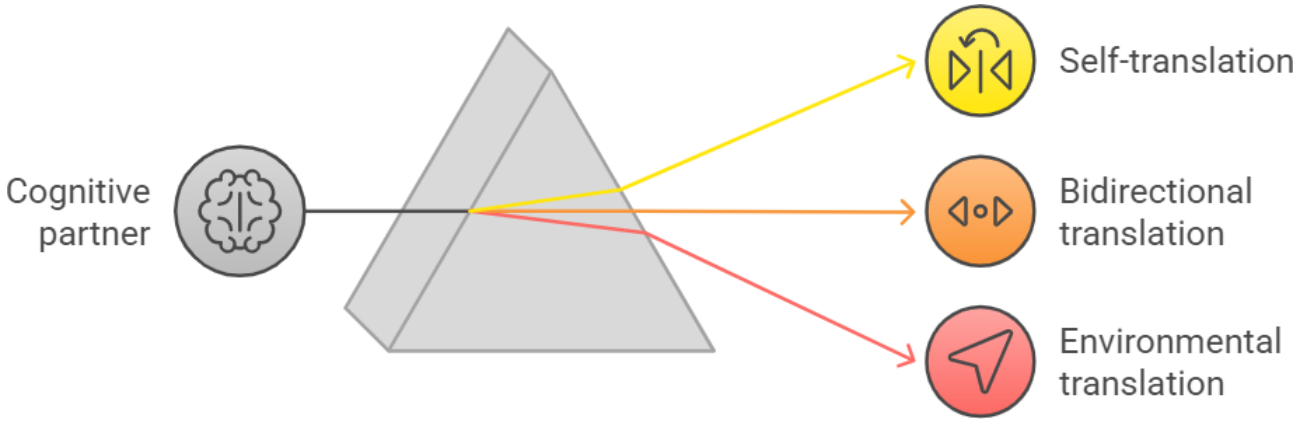
Beyond Vallor's mirror metaphor for AI and its contingent inference on social change for people with cognitive disabilities, a significant potential of GenAI lies in its ability to serve as a "cognitive partner," empowering participation of these people in life domains that were previously blocked or limited for them [43-45]. This partnership can be metaphorically described as a "cognitive copilot" (an AI assistant for complex cognitive tasks), assisting and empowering the individual with tasks requiring complex cognitive functions. For example, GenAI can help a person with cognitive disabilities manage daily tasks such as scheduling, budgeting, or navigating urban spaces by providing personalized reminders, recommendations, and guidance [46,47]. Additionally, it can serve as an advisor in complex social situations, such as interpreting body language [48], suggesting appropriate responses to expressions of anger or mockery from others, or assisting in decision-making [1,49]. In this way, GenAI may act as a kind of "social copilot," providing real-time support and feedback, allowing persons with cognitive disability to expand their circle of social interactions, inclusion, and activities.

One of the outstanding strengths of GenAI is its ability to function as a translator and mediator between languages, concepts, and realities. For people with cognitive disabilities, translation and mediation pose a central challenge in daily life, both in understanding the environment and in expressing themselves in a way others can understand [50]. With its natural learning and processing capabilities, GenAI can bridge these gaps and make information and communication more accessible.

The application of GenAI as a cognitive copilot can focus on 3 main areas (Figure 2):

1. Translating and making the inner world of people with cognitive disability accessible to themselves: GenAI can help people with cognitive disabilities better understand themselves, their thoughts, emotions, and needs. This is achieved by providing explanations and conceptualizations in clear and accessible language, identifying and interpreting emotional states, and suggesting strategies for coping with challenges [50]. GenAI can serve as an "internal translator" that through a process of assistive conceptual scaffolding and cognitive structuring [51] assists individuals in accurate self-understanding and self-expression.
2. Bidirectional translation and mediation in interpersonal communication: By analyzing interpersonal and social information, GenAI can mediate interactions with other people, making it possible to negotiate the complexities inherent in human communication more successfully. The unique contribution of GenAI in this area lies in its ability to bridge the communication gap in both directions, helping the person with cognitive disability understand the social environment, the intentions of others, and the implicit messages in discourse, and making the person's wants, needs, and emotions more accessible to the social environment [1]. For example, on one hand, GenAI can offer interpretations of social cues and recommend appropriate responses, and on the other assist individuals in articulating their thoughts more clearly and presenting their unique perspectives. The technology can serve as a "two-way social translator," enabling people with disability and their environment to better understand each other and promote respectful and equitable communication.
3. Making the physical environment and public spaces accessible: GenAI can act as an "environmental translator," converting complex information about the world into a clear and disability-friendly format. This can include, for example, simplifying official texts, graphically converting numeric data, or creating interactive guides for navigating public spaces [52]. Thus, GenAI models that are open to the public can "see" and "understand" photos and videos and describe their content [1], so that people with cognitive disabilities may gain greater access and independence in managing their lives.

Figure 2. Three main areas of GenAI application as a cognitive partner. GenAI: generative artificial intelligence.



The goal is not to "normalize" individuals with cognitive disabilities or to erase their disability. The cognitive partner metaphor, similar to Vallor's mirror metaphor, can show how the use of AI might exacerbate exclusionary attitudes and further marginalize individuals with disabilities. Therefore, using AI for social change in our attitude toward people with cognitive disabilities means that the aim of this technology should be to enable access to environments and spaces that were previously closed or socially inaccessible to them, while also facilitating the accessibility of these environments to the individuals themselves. The approach should be person-centered, respecting diversity, and tailored to the unique aspirations and needs of everyone, rather than imposing a uniform standard of "proper" functioning.

Serious consideration must be given to the ethical implications of such a close integration between humans and machines, particularly in the areas of autonomy and responsibility. Questions of privacy, data security, and people's ownership of decisions made by AI systems need to be thoroughly examined [52,53]. Robust oversight and regulatory mechanisms must be in place to ensure the responsible and ethical use of AI, safeguarding the rights and well-being of users. This is especially critical when working with vulnerable populations such as people with cognitive disabilities, where protecting individual autonomy is important [27,33].

In conclusion, although AI-based "cognitive copilot" applications for people with cognitive disabilities have the potential to remove barriers, increase participation, and promote equal opportunities across various domains of life, it is essential to proceed with caution. This technology must function as a "translator" to contribute to a more inclusive and equitable society, and we must remain vigilant to its risks. Ensuring that AI development is person-centered, ethically sound, and involves active participation from the disability community is crucial for harnessing its benefits without worsening existing biases and systemic barriers.

Implication for AI Developers and Technologists

GenAI has immense potential to promote inclusion and equality for people with cognitive disabilities but to realize this potential requires a perceptual shift on the part of developers, engineers, researchers, and product managers. Instead of focusing narrowly on "fixing" certain impairments, they must adopt a more holistic approach that views technology as a lever for social integration

and broad improvement in quality of life [54-56]. This involves a transition from regarding GenAI as a mere technical solution to perceiving it as a tool for effecting social change for the population with cognitive disabilities.

In practice, close and ongoing collaboration with people with cognitive disabilities throughout all stages of development is important [57]. Development teams must learn from the unique experiences and needs of individuals with cognitive disability and meaningfully integrate them into the design and construction of GenAI systems and prompts.

Recent research has demonstrated the feasibility and importance of this approach. For example, Newbutt et al [58] conducted a systematic review of studies involving autistic individuals in the design of extended reality technologies. They found that out of 20 studies published between 2002 - 2022, several successfully engaged autistic individuals as active co-designers and cocreators, allowing them to shape the final products according to their needs and preferences. This highlights the growing trend and importance of including the target users in the design process.

This requires a joint definition of goals, adapting user interfaces and user experience to their modes of thinking and communication, and clearly formulating principles of cognitive accessibility from the earliest planning stages [59]. The aspiration is for the empowerment and inclusion of people with cognitive disabilities to be embedded in the core of the technology and in the layer of its use.

Bircanin et al [60] presented a practical approach to including adults with severe intellectual disabilities in co-design through active support. They demonstrated how principles such as “every moment has potential,” “graded assistance,” “little and often,” and “maximizing choice and control” can be applied in design contexts to ensure meaningful participation of individuals with severe cognitive disabilities. This approach provides concrete strategies for AI developers to engage with this population during the development process.

For example, it is important to examine how the prompt-based user interface can be made accessible and adapted to the cognitive and communication characteristics of people with different types of cognitive disabilities. Consideration should be given to whether the development of dedicated products is the right direction or whether personal adaptation at the level of the individual user is preferable [61]. Answering such questions requires ongoing discourse and feedback from the community itself.

Dirks [57] explored the ethical challenges in inclusive software development projects with people with cognitive disabilities. The study emphasized the importance of maximizing choice and control for participants, using a graded assistance approach, and ensuring every moment has potential for meaningful engagement. These principles can guide AI developers in creating more inclusive design processes.

To assist developers and researchers in implementing the principles presented in this paper, we propose a working protocol specifically tailored to the development challenges of GenAI technologies aimed at people with cognitive disabilities. The protocol (Table 1) is based on the model developed by Amershi et al [62], which was formulated following comprehensive research, including a review of academic and industry literature, interviews with experts, and an examination of a wide range of AI-based products. The original model defines 18 general guidelines for designing human-AI interactions across different time frames and stages of interaction. In practice, these guidelines serve as a framework for developing human-centered AI systems, focusing on aspects such as transparency, fairness, reliability, safety, privacy, security, and accountability. Developers and designers use these guidelines to enhance human-AI interaction by implementing practices such as explaining AI decisions to users, designing interfaces that enable user control and feedback, and incorporating mechanisms to identify and mitigate biases [63].

Table . Protocol for designing artifical intelligence (AI) interactions for people with cognitive disabilities.[a]

Stage and dimension		Guidelines for AI interaction with people with cognitive disabilities	Implementation examples
Initial			
	Personal	I1. Identify and adapt to the user's unique cognitive and emotional needs.	I1. Create a personal profile including preferences, abilities, and challenges.
	Interpersonal	I2. Show awareness of the social and cultural context of system use.	I2. Consider the human environment (eg, caregivers or family members) as part of system definition.
During interaction			
	Personal	D1. Provide custom-tailored, gradual, and structured responses to personal needs during use.	D1. Identify difficulties and adapt the level of assistance and feedback in real time.
	Interpersonal	D2. Promote positive and reciprocal communication with the human environment.	D2. Mediate social interactions by simplifying and explaining social cues.
	Environmental	D3. Assist in orientation, navigation, and independent functioning in complex spaces.	D3. Provide detailed instructions and cues on proper conduct in different places.
When the system errs			
	Personal	E1. Handle errors respectfully and in an empowering way, with emphasis on learning and progress.	E1. Provide repeated opportunities to try again, together with verbal encouragement.
	Interpersonal	E2. Involve support persons in the process of learning and correction.	E2. Provide a possibility for a caregiver to assist in problem-solving or making necessary adjustments.
	Environmental	E3. Avoid placing responsibility on the user in complex or unexpected situations.	E3. Make human backup available by default in case of significant problems.
Over time			
	Personal	T1. Continually adapt to the pace of development, learning, and changes in personal needs.	T1. Track progress and adapt tasks and goals accordingly.
	Interpersonal	T2. Show sensitivity to changes in relationships and roles within the support circle.	T2. Update user profiles and access settings based on feedback from the environment.
	Environmental	T3. Show flexibility and adaptability to changing environments and transitions between contexts.	T3. Automatically detect location changes and provide relevant recommendations.
	Collaboration	T4. Actively involve users and stakeholders in the ongoing development of the system.	T4. Provide mechanisms for receiving feedback and involving users in decisions about updates and improvements.

[a]The model for this protocol by Amershi et al [62] is based on extensive research and analysis of a range of artificial intelligence products and defines 18 general guidelines across different stages of interaction. We adapted and extended this model to address specifically the needs and challenges of designing artificial intelligence technologies for people with cognitive disabilities. The protocol incorporates 4 key dimensions: personal, interpersonal, environmental, and collaborative, and provides concrete examples of how these considerations can be integrated throughout the life cycle of the artificial intelligence system. By implementing this protocol, developers can create artificial intelligence tools that empower and enhance the lives of individuals with cognitive disabilities.

Building on the analysis presented in this paper, we expand the model of Amershi et al [62] and adapt it to the 4 central dimensions in which AI systems can assist people with cognitive disabilities: the personal, the interpersonal, the environmental, and the collaborative. For each of these dimensions, we propose guidelines and offer practical examples of how the relevant considerations can be embedded at different stages of the system life cycle, from defining the initial requirements, through ongoing interaction, to continuous adaptation and improvement. The proposed protocol serves as a foundation that requires further development, testing, and investigation, but it can serve as a starting point for discourse and the advancement of best

practices in designing AI systems for individuals with cognitive disabilities.

Conclusion

The emergence of GenAI technologies represents a pivotal moment in reconceptualizing disability and personhood. We suggest that the advent of GenAI challenges assumptions about what qualifies an individual as a "person" and questions the notion that cognitive abilities are the sole determinant of one's rights and societal participation.

In this paper, we explored the transformative potential of GenAI in reshaping perceptions, dismantling barriers, and empowering individuals with cognitive disabilities. By serving as a social mirror [32], AI systems can expose and challenge deeply ingrained biases and prejudices, compelling us to confront the ways we have historically marginalized and excluded the population with cognitive disabilities. Simultaneously, by functioning as a cognitive partner, GenAI may provide unprecedented opportunities for individuals with cognitive disabilities to participate in society.

Realizing this vision requires more than technological innovation, however. It demands a gradual shift in societal attitudes and a sincere effort to involve people with cognitive disabilities in the AI development process, granting them autonomy and recognizing and valuing their abilities. This is where the role of technology professionals and GenAI developers becomes crucial.

The importance of designing AI thoughtfully lies in the understanding that whether we consider AI as a mirror or as a cognitive partner, both metaphors indicate that AI will increasingly mediate how we perceive the world, ourselves, and others, confirming once again McLuhan's [64] statement that "the medium is the message." This means that the significant effect of AI lies not merely in the content we explore through it but in how its very use changes us. Therefore, the design and development of AI tools will profoundly influence the future of human society, how we perceive individuals with disabilities, as well as the rights and social positions they will attain. Therefore, how AI is being shaped now will determine its role in reinforcing existing biases or promoting a more inclusive and equitable society.

The proposed protocol, based on the work by Amershi et al [62], offers a practical framework for implementing these principles as part of GenAI development for people with cognitive disabilities. This paper marks only the beginning of the discussion about GenAI and developmental disabilities, therefore we must remain vigilant regarding the ethical and social implications of GenAI and continue to engage in open, multidisciplinary dialogue about how to harness its potential for the greater good.

The path ahead is complex and challenging, but it is also filled with immense possibilities. As we look toward the future, the evolution of AI from reactive, prompt-based systems to proactive, autopilot models promises to further expand these possibilities, particularly for individuals with cognitive disabilities. These advanced systems, capable of learning user needs and initiating interactions without explicit prompts, could provide more seamless and intuitive support, potentially revolutionizing the way we approach cognitive assistance.

Technological progress also involves an ongoing need for ethical and inclusive development. We must prioritize user autonomy and privacy while maximizing the benefits of technological assistance. This balance is important not only for protecting individual rights but also for ensuring that AI serves the needs of those it aims to support.

By embracing the potential of GenAI while remaining vigilant regarding its ethical implications, researchers, developers, and policy makers can create technologies that not only uplift those who have been historically marginalized but enrich the human experience for us all. In doing so, we may take a step toward a future where technology serves as a platform for inclusivity and empowerment.

Acknowledgments

The authors would like to express their gratitude to the Artificial Third community for promoting multidisciplinary discourse on artificial intelligence in mental health. This community has made possible valuable interactions between researchers in the fields of psychology, disability studies, and artificial intelligence, contributing to the development of this theoretical study.

Conflicts of Interest

The author TS is the chief scientist of R&D at Microsoft Israel. The views and opinions expressed here are those of the authors and do not reflect the official policy or position of Microsoft. TS received no financial compensation for his contribution to this work.

References

1. Elyoseph Z, Shoval DH, Levkovich I. Beyond personhood: ethical paradigms in the generative artificial intelligence era. Am J Bioeth 2024 Jan;24(1):57-59. [doi: 10.1080/15265161.2023.2278546] [Medline: 38236857]
2. Holm S, Lewis J. The ends of personhood. Am J Bioeth 2024 Jan;24(1):30-32. [doi: 10.1080/15265161.2023.2278561] [Medline: 38236874]
3. Blumenthal-Barby J. The end of personhood. Am J Bioeth 2024 Jan;24(1):3-12. [doi: 10.1080/15265161.2022.2160515] [Medline: 36635972]

4. Haber Y, Levkovich I, Hadar-Shoval D, Elyoseph Z. The artificial third: a broad view of the effects of introducing generative artificial intelligence on psychotherapy. JMIR Ment Health 2024 May 23;11:e54781. [doi: 10.2196/54781] [Medline: 38787297]
5. Sutcliffe MS, Radonovich K. Cognitive disabilities in children and adolescents. In: Halpern-Felsher B, editor. Encyclopedia of Child and Adolescent Health: Academic Press; 2023:11-21. [doi: 10.1016/B978-0-12-818872-9.00136-9]
6. Houwen S, Visser L, van der Putten A, Vlaskamp C. The interrelationships between motor, cognitive, and language development in children with and without intellectual and developmental disabilities. Res Dev Disabil 2016 Jun;53-54:19-31. [doi: 10.1016/j.ridd.2016.01.012]
7. Dev P. Introduction to different kinds of cognitive disorders. In: Gupta S, editor. Bio-Inspired Algorithms and Devices for Treatment of Cognitive Diseases Using Future Technologies: IGI Global; 2022:39-55. [doi: 10.4018/978-1-7998-9534-3.ch003]
8. Peltokorpi J, Hoedt S, Colman T, Rutten K, Aghezzaf EH, Cottyn J. Manual assembly learning, disability, and instructions: an industrial experiment. Int J Prod Res 2023 Nov 17;61(22):7903-7921. [doi: 10.1080/00207543.2023.2195957]
9. Heuser P, Letmathe P, Vossen T. Skill development in the field of scheduling: a structured literature review. Eur J Oper Res 2025 Mar;321(3):697-716. [doi: 10.1016/j.ejor.2024.04.005]
10. Rojas M, Balderas DC, Maldonado J, Ponce P, Lopez-Bernal D, Molina A. Lack of verified inclusive technology for workers with disabilities in Industry 4.0: a systematic review. Int J Sustainable Eng 2024 Dec 31;17(1):1-21. [doi: 10.1080/19397038.2024.2328711]
11. Litwin P, Antonelli D, Stadnicka D. Employing disabled workers in production: simulating the impact on performance and service level. Int J Prod Res 2024 Jun 17;62(12):4530-4545. [doi: 10.1080/00207543.2023.2266066]
12. Young G. Personhood across disciplines: applications to ethical theory and mental health ethics. Ethics Med Public Health 2019 Jul;10:93-101. [doi: 10.1016/j.jemep.2019.100407]
13. Rosfort R. Personhood. In: Stanghellini G, Broome M, Fernandez AV, FusarPoli P, Raballo A, Rosfort R, editors. The Oxford Handbook of Phenomenological Psychopathology: Oxford University Press; 2019:335-343. [doi: 10.1093/oxfordhb/9780198803157.013.38]
14. Asch A. Disability, bioethics, and human rights. In: Albrecht GL, Seelman KD, Bury M, editors. Handbook of Disability Studies: Sage Publications; 2001. [doi: 10.4135/9781412976251.n12]
15. Shakespeare T. The social model of disability. In: Davis LJ, editor. The Disability Studies Reader: Routledge; 2006:197-204.
16. Davis LJ, editor. The Disability Studies Reader, 5th edition: Routledge; 2016. [doi: 10.4324/9781315680668]
17. Iacovou M. A contribution towards a possible re-invigoration of our understanding of the social model of disability's potential. Disabil Soc 2021 Aug 9;36(7):1169-1185. [doi: 10.1080/09687599.2020.1787816]
18. Brinkman AH, Rea-Sandin G, Lund EM, et al. Shifting the discourse on disability: moving to an inclusive, intersectional focus. Am J Orthopsychiatry 2023;93(1):50-62. [doi: 10.1037/ort0000653] [Medline: 36265035]
19. Cinquin PA, Guitton P, Sauzéon H. Designing accessible MOOCs to expand educational opportunities for persons with cognitive impairments. Behav Inf Technol 2021 Aug 18;40(11):1101-1119. [doi: 10.1080/0144929X.2020.1742381]
20. Fuentes K, Hsu S, Patel S, Lindsay S. More than just double discrimination: a scoping review of the experiences and impact of ableism and racism in employment. Disabil Rehabil 2024 Feb;46(4):650-671. [doi: 10.1080/09638288.2023.2173315] [Medline: 36724368]
21. Schidel R. Philosophy universal enfranchisement for citizens with cognitive disabilities – a moral-status argument disabilities – a moral-status argument. Crit Rev Int Soc Polit Philos 2023;26(5):658-679. [doi: 10.1080/13698230.2022.2062212]
22. Zhu D, Al Mahmud A, Liu W. Social connections and participation among people with mild cognitive impairment: barriers and recommendations. Front Psychiatry 2023;14:1188887. [doi: 10.3389/fpsyt.2023.1188887] [Medline: 37476544]
23. Foucault M. Discipline and Punish: The Birth of the Prison: Vintage; 1995.
24. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science 2023 Jul 14;381(6654):187-192. [doi: 10.1126/science.adh2586] [Medline: 37440646]
25. Chan CKY, Hu W. Students' voices on generative AI: perceptions, benefits, and challenges in higher education. Int J Educ Technol High Educ 2023;20(1):43. [doi: 10.1186/s41239-023-00411-8]
26. Yusuf A, Pervin N, Román-González M. Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. Int J Educ Technol High Educ 2024;21(1):21. [doi: 10.1186/s41239-024-00453-6]
27. Hadar-Shoval D, Asraf K, Mizrachi 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 2024 Apr 9;11:e55988. [doi: 10.2196/55988] [Medline: 38593424]
28. Hadar-Shoval D, Asraf K, Shinan-Altman S, Elyoseph Z, Levkovich I. Embedded values-like shape ethical reasoning of large language models on primary care ethical dilemmas. Heliyon 2024 Sep 30;10(18):e38056. [doi: 10.1016/j.heliyon.2024.e38056] [Medline: 39381244]
29. Cath C, Wachter S, Mittelstadt B, Taddeo M, Floridi L. Artificial Intelligence and the "Good Society": the US, EU, and UK approach. Sci Eng Ethics 2018 Apr;24(2):505-528. [doi: 10.1007/s11948-017-9901-7] [Medline: 28353045]

30. Whittaker M, Alper M, Bennett CL, et al. Disability, bias, and AI. AI Now Institute. 2019. URL: https://ainowinstitute.org/publication/disabilitybiasai-2019 [accessed 2024-12-24]
31. Cave S, Dihal K. The whiteness of AI. Philos Technol 2020 Dec;33(4):685-703. [doi: 10.1007/s13347-020-00415-6]
32. Vallor S. The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking: Oxford University Press; 2024.
33. Hadar-Shoval D, Elyoseph Z, Lvovsky M. The plasticity of ChatGPT's mentalizing abilities: personalization for personality structures. Front Psychiatry 2023;14:1234397. [doi: 10.3389/fpsyt.2023.1234397] [Medline: 37720897]
34. Gadiraju V, Kane S, Dev S, et al. "I wouldn't say offensive but...": disability-centered perspectives on large language models. Presented at: FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency; Jun 12-15, 2023; Chicago, IL, United States of America. [doi: 10.1145/3593013.3593989]
35. Kidd C, Birhane A. How AI can distort human beliefs. Science 2023 Jun 23;380(6651):1222-1223. [doi: 10.1126/science.adi0248] [Medline: 37347992]
36. Vallor S. The AI mirror: reclaiming our humanity in an age of machine thinking. Presented at: AIES '22: Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society; May 19-21, 2022; Oxford, United Kingdom. [doi: 10.1145/3514094.3539567]
37. Sætra HS, Coeckelbergh M, Danaher J. The AI ethicist's dirty hands problem. Commun ACM 2023 Jan;66(1):39-41. [doi: 10.1145/3529751]
38. Chien J, Danks D. Beyond behaviorist representational harms: a plan for measurement and mitigation. Presented at: FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency; Jun 3-6, 2024; Rio de Janeiro, Brazil. [doi: 10.1145/3630106.3658946]
39. Gebru T, Morgenstern J, Vecchione B, et al. Datasheets for datasets. Commun ACM 2021 Dec;64(12):86-92. [doi: 10.1145/3458723]
40. Kommrusch S, Monperrus M, Pouchet LN. Self-supervised learning to prove equivalence between straight-line programs via rewrite rules. IEEE Trans Softw Eng 2023;49(7):3771-3792. [doi: 10.1109/TSE.2023.3271065]
41. Da Costa L, Sajid N, Parr T, Friston K, Smith R. Reward maximization through discrete active inference. Neural Comput 2023 Apr 18;35(5):807-852. [doi: 10.1162/neco_a_01574] [Medline: 36944240]
42. Canton E, Hedley D, Spoor JR. The stereotype content model and disabilities. J Soc Psychol 2023 Jul 4;163(4):480-500. [doi: 10.1080/00224545.2021.2017253] [Medline: 35037596]
43. Ferland L, Li Z, Sukhani S, Zheng J, Zhao L, Gini M. Assistive AI for coping with memory loss. Presented at: The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18); Feb 2-7, 2018; New Orleans, Louisiana, USA URL: https://cdn.aaai.org/ocs/ws/ws0528/17360-76000-1-PB.pdf [accessed 2024-12-24]
44. Pednekar S, Dhirawani P, Shah R, Shekokar N, Ghag K. Voice-based interaction for an aging population: a systematic review. Presented at: 2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT); Jan 19-20, 2023; Jaipur, India. [doi: 10.1109/ICCT56969.2023.10075801]
45. Robledo-Castro C, Castillo-Ossa LF, Corchado JM. Artificial cognitive systems applied in executive function stimulation and rehabilitation programs: a systematic review. Arab J Sci Eng 2023;48(2):2399-2427. [doi: 10.1007/s13369-022-07292-5] [Medline: 36185593]
46. Huq SM, Maskeliūnas R, Damaševičius R. Dialogue agents for artificial intelligence-based conversational systems for cognitively disabled: a systematic review. Disabil Rehabil Assist Technol 2024 Apr;19(3):1059-1078. [doi: 10.1080/17483107.2022.2146768] [Medline: 36413423]
47. Hueso M, Álvarez R, Marí D, Ribas-Ripoll V, Lekadir K, Vellido A. Is generative artificial intelligence the next step toward a personalized hemodialysis? RIC 2023 Dec 20;75(6):309-317. [doi: 10.24875/RIC.23000162]
48. Elyoseph Z, Refoua E, Asraf K, Lvovsky M, Shimoni Y, Hadar-Shoval D. Capacity of generative AI to interpret human emotions from visual and textual data: pilot evaluation study. JMIR Ment Health 2024 Feb 6;11:e54369. [doi: 10.2196/54369] [Medline: 38319707]
49. Elyoseph Z, Hadar-Shoval D, Asraf K, Lvovsky M. ChatGPT outperforms humans in emotional awareness evaluations. Front Psychol 2023;14:1199058. [doi: 10.3389/fpsyg.2023.1199058] [Medline: 37303897]
50. Kelly C, Cornwell P, Hewetson R, Copley A. The pervasive and unyielding impacts of cognitive-communication changes following traumatic brain injury. Int J Lang Commun Disord 2023;58(6):2131-2143. [doi: 10.1111/1460-6984.12923] [Medline: 37424402]
51. Asman O, Segal-Reich M. Supported decision making "adaptive suit" for non-dominating mental scaffolding. Am J Bioeth - Neurosci 2023;14(3):238-240. [doi: 10.1080/21507740.2023.2243878]
52. Tal A, Elyoseph Z, Haber Y, et al. The artificial third: utilizing ChatGPT in mental health. Am J Bioeth 2023 Oct;23(10):74-77. [doi: 10.1080/15265161.2023.2250297] [Medline: 37812102]
53. Elyoseph Z, Gur T, Haber Y, et al. An ethical perspective on the democratization of mental health with generative AI. JMIR Ment Health. 2024 Oct 17. [doi: 10.2196/58011] [Medline: 39417792]
54. Johansson S, Gulliksen J, Lantz A. User participation when users have mental and cognitive disabilities. Presented at: The 17th International ACM SIGACCESS Conference on computers and accessibility; Oct 26-28, 2015; Lisbon, Portugal. [doi: 10.1145/2700648.2809849]

55. Donati M, Pacini F, Baldanzi L, Turturici M, Fanucci L. Remotely controlled electronic goalkeeper: an example of improving social integration of persons with and without disabilities. Appl Sci (Basel) 2023;13(11):6813. [doi: 10.3390/app13116813]
56. Harb B, Sidani D. Smart technologies challenges and issues in social inclusion – case of disabled youth in a developing country. JABS 2022 Mar 23;16(2):308-323. [doi: 10.1108/JABS-10-2020-0389]
57. Dirks S. Ethical challenges in inclusive software development projects with people with cognitive disabilities: ethische herausforderungen in inklusiven softwareentwicklungsprojekten mit menschen mit kognitiven beeinträchtigungen. Presented at: MuC '22: Proceedings of Mensch und Computer 2022; Nov 4-7, 2022; Darmstadt, Germany. [doi: 10.1145/3543758.3547575]
58. Newbutt N, Glaser N, Francois MS, Schmidt M, Cobb S. How are autistic people involved in the design of extended reality technologies? a systematic literature review. J Autism Dev Disord 2024 Nov;54(11):4232-4258. [doi: 10.1007/s10803-023-06130-3] [Medline: 37715864]
59. Shastri K, Boger J, Marashi S, et al. Working towards inclusion: creating technology for and with people living with mild cognitive impairment or dementia who are employed. Dem (Lon) 2022 Feb;21(2):556-578. [doi: 10.1177/14713012211051885] [Medline: 34749536]
60. Bircanin F, Brereton M, Sitbon L, Ploderer B, Azaabanye Bayor A, Koplick S. Including adults with severe intellectual disabilities in co-design through active support. CHI Conf Hum Factor Comput Syst 2021:1-12. [doi: 10.1145/3411764.3445057]
61. Moreno L, Petrie H, Martínez P, Alarcon R. Designing user interfaces for content simplification aimed at people with cognitive impairments. Univ Access Inf Soc 2023 Mar 24;23(1):1-19. [doi: 10.1007/s10209-023-00986-z] [Medline: 37361672]
62. Amershi S, Weld D, Vorvoreanu M, et al. Guidelines for human-AI interaction. Presented at: CHI '19; May 4-9, 2019; Glasgow, Scotland, United Kingdom. [doi: 10.1145/3290605.3300233]
63. Bingley WJ, Curtis C, Lockey S, et al. Where is the human in human-centered AI? Insights from developer priorities and user experiences. Comput Hum Behav 2023 Apr;141:107617. [doi: 10.1016/j.chb.2022.107617]
64. McLuhan M. Understanding Media: The Extensions of Man: McGraw Hill; 1964.

Abbreviations

AI: artificial intelligence
GenAI: generative artificial intelligence

Edited by P Kubben; submitted 10.07.24; peer-reviewed by A Yusuf, C Wang; revised version received 22.10.24; accepted 06.11.24; published 15.01.25.

Please cite as:
Hadar Souval D, Haber Y, Tal A, Simon T, Elyoseph T, Elyoseph Z
Transforming Perceptions: Exploring the Multifaceted Potential of Generative AI for People With Cognitive Disabilities
JMIR Neurotech 2025;4:e64182
URL: https://neuro.jmir.org/2025/1/e64182
doi:10.2196/64182

A Game-Based Mechatronic Device for Digital Rehabilitation of Hand Function After a Stroke: Design, Prototyping, and Feasibility Study

Anuprita Kanitkar[1], BPT, MSc, PhD; Nariman Sepehri[2], FCSME, FASME, FCAE, PhD; Ariel Lezen[2]; Sanjay Tejraj Parmar[3], PDF, PhD; Cherry Kit-Fong Hin[1], BPT, MSc; Tony Joseph Szturm[1], BPT, PhD

[1]Department of Physical Therapy, College of Rehabilitation Sciences, University of Manitoba, Winnipeg, MB, Canada

[2]Department of Mechanical Engineering, Price Faculty of Engineering, University of Manitoba, Winnipeg, MB, Canada

[3]SDM College of Physiotherapy, Shri Dharmasthala Manjunatheshwara University, Dharwad, India

Corresponding Author:
Tony Joseph Szturm, BPT, PhD
Department of Physical Therapy, College of Rehabilitation Sciences, University of Manitoba, Winnipeg, MB, Canada

Abstract

Background: This paper presents an easy-to-use, affordable robotic manipulandum device (RMD) equipped with smart monitoring and assistive technologies to engage in game-based exercise and repetitive task practice. The RMD has been designed to enhance a wide range of fine motor manual dexterity skills, including thumb, finger, and wrist movements. By focusing on finger and hand functions, it extends its utility beyond basic reaching or object transfer movements. Various interchangeable 3D-printed therapy handles of different shapes and sizes can be easily attached to the RMD drive shaft. These handle movements can be used to engage with numerous affordable, commercially available computer games, allowing patients to practice tasks that involve varying movement amplitudes, speeds, precision, and cognitive challenges. Additionally, the device is capable of automatically recording and storing the patient's real-time performance data on any given computer, integrating assessment into treatment.

Objective: A pilot study was conducted with 5 patients with stroke to examine the feasibility and benefits of a 6-week game-based exercise program using the proposed device.

Methods: A feasibility study was conducted with 5 participants. Data were collected using the computer game–based upper extremity assessment of manual dexterity and Wolf Motor Function Test (WMFT) before and after the intervention lasting 6 weeks.

Results: The pilot study demonstrated that clients' expectations related to manual dexterity were met. The average improvement in the functional ability score of the WMFT was 14 (SD 3) points, with all participants exceeding the minimal clinically important difference. The average reduction in total time was 30 (SD 14) seconds, with 4 of 5 participants surpassing the minimal clinically important difference. For the computer game–based upper extremity assessment, the average improvement in success rate was 23% (SD 12%), and the average decrease in response time was 105 (SD 44) milliseconds.

Conclusions: Findings revealed acceptable, engaging, game-based, and task-oriented training with a high level of compliance. Substantial improvements from pre- to postintervention were observed using the WMFT and assessments of manual dexterity.

Trial Registration: ClinicalTrials.gov NCT05071885; https://clinicaltrials.gov/study/NCT05071885

(JMIR Neurotech 2025;4:e67779) doi:10.2196/67779

KEYWORDS

stroke; manual dexterity; hand function; poststroke; fine motor; thumb; finger; wrist; movement; motor rehabilitation; assistive technology; smart monitoring; pilot; feasibility; prototyping; prototype; nervous system; nerve; motor neuron

Introduction

Background

Upper extremity (UE) motor impairments and persistent hemiparesis commonly lead to difficulties with manual dexterity after a stroke [1]. Manual dexterity, defined as the ability to manipulate objects, is crucial for many everyday tasks, both for leisure and social interactions. These tasks often require the manipulation of objects that vary widely in physical properties and functional demands, necessitating a high degree of precision [2]. Individuals with chronic sensory-motor deficits in the UE following a stroke can greatly benefit from intensive, well-resourced therapy services [3-6]. A novel approach to enhance patient engagement in therapy is the use of computer

games, which integrate various learning elements and present motor and cognitive challenges. This allows individuals to participate in focused, task-specific activities with a significant number of repetitions [7-10]. Several gaming systems have been used as rehabilitation tools [11]. Various computer input devices have been used to detect arm segments or finger motions. The corresponding motion signals are used to interact with digital avatars or objects [12,13]. However, these game-based exercise programs often fail to adequately address object handling and fine motor function–based object manipulation. Consequently, they do not account for the sensory, tactile, or proprioceptive signals from the hand that are essential for effective goal-directed object manipulation tasks. To enhance the brain's capacity for learning, it is vital to create experiences that improve manual dexterity through guided and repetitive practice of manipulation tasks requiring precision [14-16]. Some game-based rehabilitation systems use handles or joysticks as controllers [17,18], where the handle is manipulated using wrist, elbow, and shoulder motions. However, these systems include only a few custom-made games.

To extend these systems, a cost-effective computer-based gaming platform has already been developed, which integrates various object manipulation tasks with engaging computer game activities. This platform uses a miniature, wireless, inertial-based (IB) computer mouse that directly connects object manipulation with digital gaming [19-22]. The IB mouse can be attached to a wide range of objects with different shapes, sizes, and weights and can be handled using 2-finger, 3-finger, or whole-hand motions as well as wrist, elbow, and shoulder movements. These object manipulation tasks are used to practice diverse, goal-oriented manual dexterity skills while users engage with entertaining computer games. However, this gaming system does not provide movement assistance for patients with limited active range of motion or poor movement control.

Numerous studies have assessed the feasibility and impact of various robotic systems aimed at improving UE functions in patients with stroke [23-30]. Augmented reality game–based devices focus on enhancing the range of motion in the shoulder, elbow, and wrist. However, these devices are not able to detect hand and finger movements with the required amount of precision. The camera-motion and sensor-based devices cannot detect movement with real-life objects. Thus, these devices can only detect active gross movements, neglecting object manipulation. A few robotic devices such as soft or hard gloves and exoskeletons do assist with finger and thumb flexion-extension; however, they primarily feature custom software applications that involve activities performed in digital settings rather than real object manipulations.

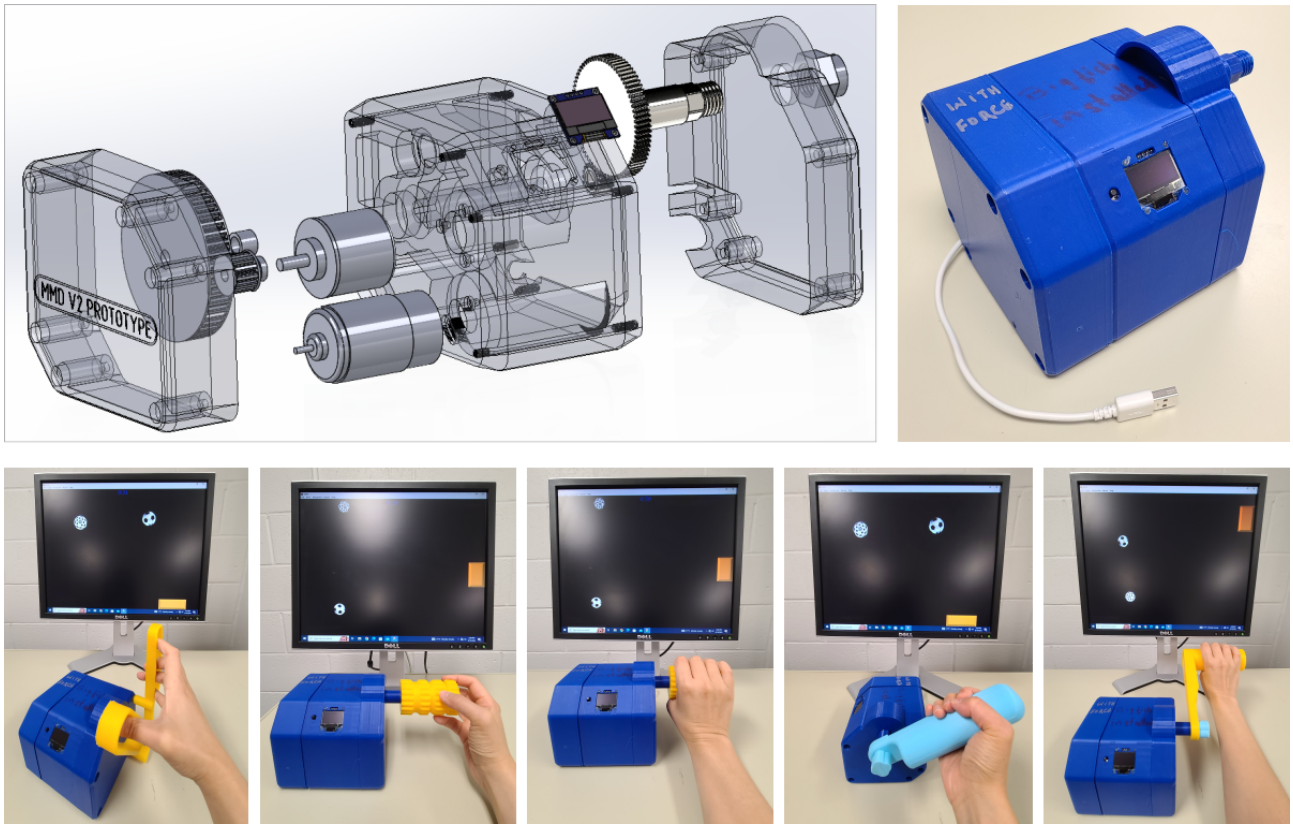
Given the above considerations, a low-cost, portable, multipurpose robotic manipulandum device (RMD) equipped with smart monitoring and assistive technologies for game-based rehabilitation of manual dexterity was developed. The RMD functions as a responsive, high-resolution computer mouse. In this paper, we first describe the RMD hardware and gaming software, its functionality, and related applications to provide both treatment and assessment of recovery programs targeting the manual dexterity of people after a stroke. The objective of this study is to present the results of a proof-of-principle pilot study conducted on 5 patients with stroke to examine the feasibility and benefits of a 6-week game-based exercise program using the RMD. The RMD described in this paper explains an integrated controller to generate forces that can aid voluntary movements necessary during gaming exercises, making it suitable for patients with limited movement control and those with a restricted active range of motion.

Description of the RMD and Software

Referring to Figure 1, the RMD features a compact, integrated 3D-printed chassis that contains the interface board, actuator, sensors, power train, and rotary drive shaft. Various 3D-printed therapy handles of different shapes and sizes can be attached to the shaft. These handles are designed to help users practice a wide range of manual dexterity skills involving thumb and finger movements as well as wrist, elbow, and shoulder functions. The RMD connects to a computer using a standard USB cable. An optical encoder tracks the shaft rotation, which corresponds to the movements of the handle and controls the motion of a computer cursor or game sprite in any single-axis computer game. In this context, the rotation of the shaft is mapped to pixel coordinates on the screen. An Arduino Leonardo microprocessor manages the RMD and its interaction with the games. Additionally, the RMD features a 3-cm LED display that shows several adjustable control parameters, which the user can modify:

- Gameplay orientation: mouse horizontal or vertical motion. Many common and modern video games are played with horizontal game sprite motion, but some require vertical motion.
- Working range: users can select an active range of motion for exercises, for example, from wrist neutral to 10, 20, or 30 degrees of extension or flexion ranges of motion, depending on individual patient needs, and map this to the full-screen mouse position.
- Mouse sensitivity: this setting determines the amount of movement required to navigate the mouse across the entire display range.
- Force: the RMD is designed to facilitate various assistive and resistive movement patterns. One of its applications involves a unidirectional force field mode, where a consistent force is exerted on the output shaft in a specific direction, with both the magnitude and direction adjustable. Many patients exhibit greater impairments in finger and wrist movement in 1 direction (eg, wrist extension), making the assistance of a constant force beneficial. Conversely, the opposite movement (eg, wrist flexion) can be met with a resistive force. This context-sensitive assistive or resistive mode can enhance even minimal voluntary movements in severely affected individuals, creating opportunities for progressive exercise that increases movement demands.

Figure 1. General view of the robotic manipulandum device and examples of various handles used for game-based rehabilitation of manual dexterity.



Since the RMD operates as a USB plug-and-play computer mouse, it is compatible with digitally any commercially available computer game. The inclusion of gaming elements motivates patients, providing an enjoyable way to engage in repetitive movements that are often necessary for rehabilitation. The therapeutic benefits from the types of object manipulation tasks involved vary in physical and anatomical requirements. The selected computer games provide graded responses in movement amplitude, speed, and precision. Table 1 outlines several common computer games that have been extensively tested with the RMD among patients of various ages. Additionally, a specially designed rehabilitation repetitive task practice (RTP) game has been created by the University of Manitoba and validated [31-33]. This simple game records the movements of the computer mouse curser or game paddle to assess the quality of movements. This game automatically tracks patients' goal-directed object manipulation tasks during both local and remote game–based therapy sessions, allowing for performance quantification in each session. This feedback can provide immediate results to the patient and help clinicians monitor progress over time. In practice, RMD-assisted exercises would initially use the RTP game software for therapeutic purposes. The RTP software is customizable, enabling adjustments to all game elements to suit the skill levels of patients with varying degrees of sensory-motor impairments.

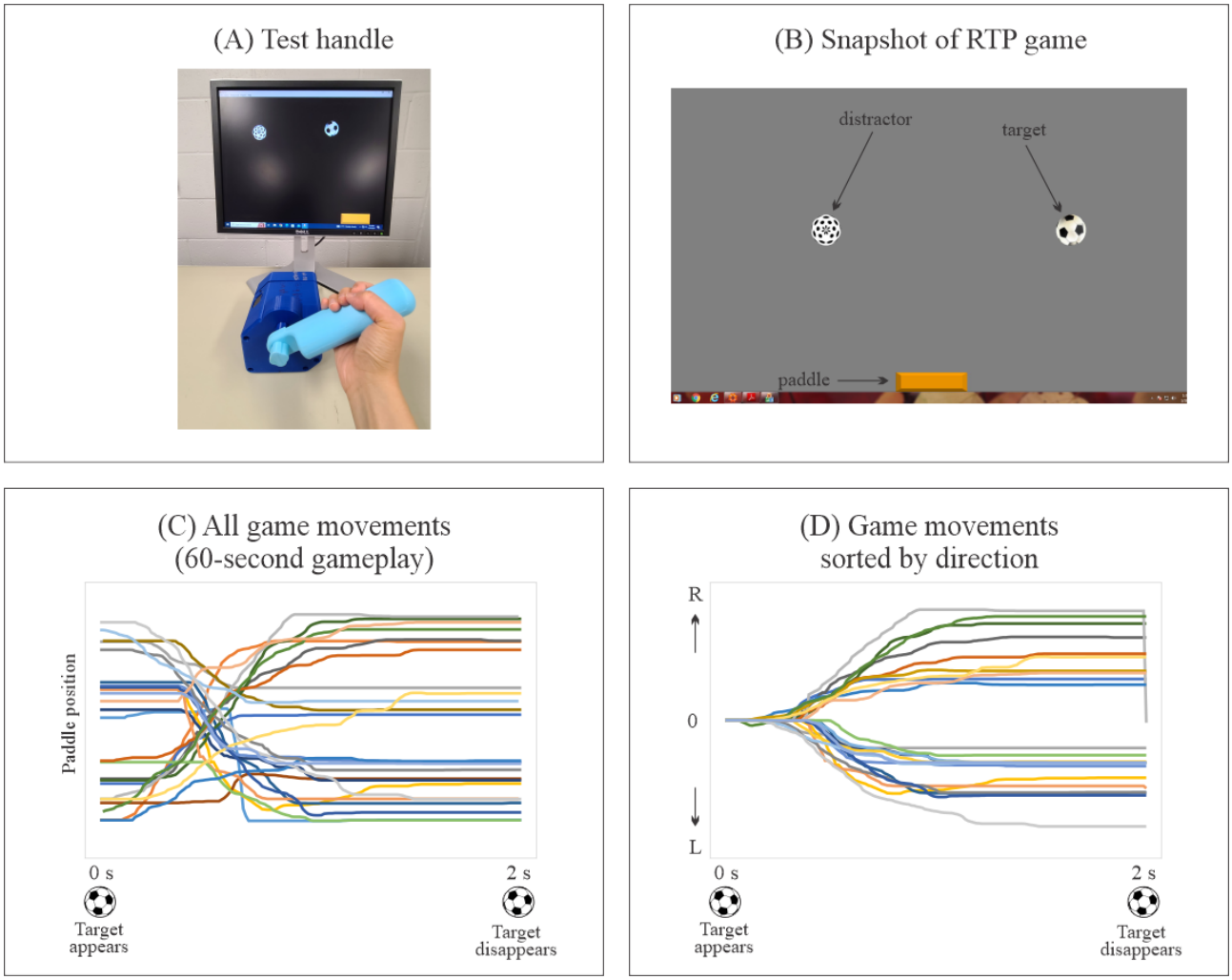
Table . Big Fish Games were used in this study [34].

Game	Axis play	Start difficulty	Response time	Clicker	Precision	Distractor	Type or activity
Abundante	Horizontal	Moderate	Self-paced (time limited)	Yes	Moderate	No	Color matching by directional aiming
Action Ball	Horizontal	Moderate	Fast	Yes	Moderate	Yes	Brick buster
Aqua Ball	Horizontal	Easy	Moderate	Yes	Moderate	Yes	Brick buster
Astro Bugz Revenge	Horizontal	Moderate	Slow	Yes	Moderate	No	Color matching by directional aiming
Birds Town	Horizontal	Moderate	Moderate	Yes	Moderate	No	Color matching by directional aiming
Brave Piglet	Vertical	Easy	Fast	Yes	Low	Yes	Shooting
Bricks of Egypt	Horizontal	Moderate	Fast	Yes	Moderate	Yes	Brick buster
Butterfly Escape	Horizontal	Moderate	Moderate	Yes	Moderate	No	Color matching by directional aiming
Egyptian Ball	Horizontal	Moderate	Fast	Yes	Moderate	Yes	Brick buster
Invadazoid	Horizontal	Moderate	Fast	Yes	Moderate	Yes	Brick buster
Jar of Marbles	Horizontal	Easy	Self-paced (time unlimited)	Yes	Moderate	No	Color matching by directional aiming
Jet Jumper	Horizontal	Difficult	Fast	Yes	High	Yes	Steering and jumping
Luxor HD	Horizontal	Moderate	Moderate	Yes	Moderate	No	Color matching by directional aiming
Ricochet Recharge	Horizontal	Moderate	Fast	Yes	Moderate	Yes	Brick buster

[a]Matching and shooting games require participants to use a small wireless optical computer mouse, pressing the left mouse button when needed. Precision is determined by the size of the paddle and the size of the target objects. Difficulty levels include game speed, the number of distractors, and matching choices.

Figure 2 illustrates a snapshot of the RTP game, highlighting game movement responses when using the RMD. Game objects appear randomly at the top of the display, moving at unpredictable speeds and directions toward the bottom. Players aim to maneuver the game paddle to catch these moving targets, with the RMD handle rotation controlling the paddle's motion. Distractor objects are included to increase challenge and can be toggled on or off. Configurable features include movement speed, precision (eg, sizes of game objects and paddles), movement amplitude, and the incorporation of distractors to assess the interplay between motor and cognitive processing as well as dual-task interference effects. Throughout gameplay, the RTP software logs the timing of each game object's appearance and disappearance, defining game events, along with tracking the position of the paddle and other game objects to establish movement context. Panels C and D in Figure 2 demonstrate typical movement trajectories within the game. Various performance metrics can be captured using the RTP game, offering immediate feedback for both patients and therapists. Additionally, electronic outcome measures are recorded to monitor progress and dose-response relationships in specific exercise programs over time, including success rates (SR), response times, movement durations, accuracy, and movement variability.

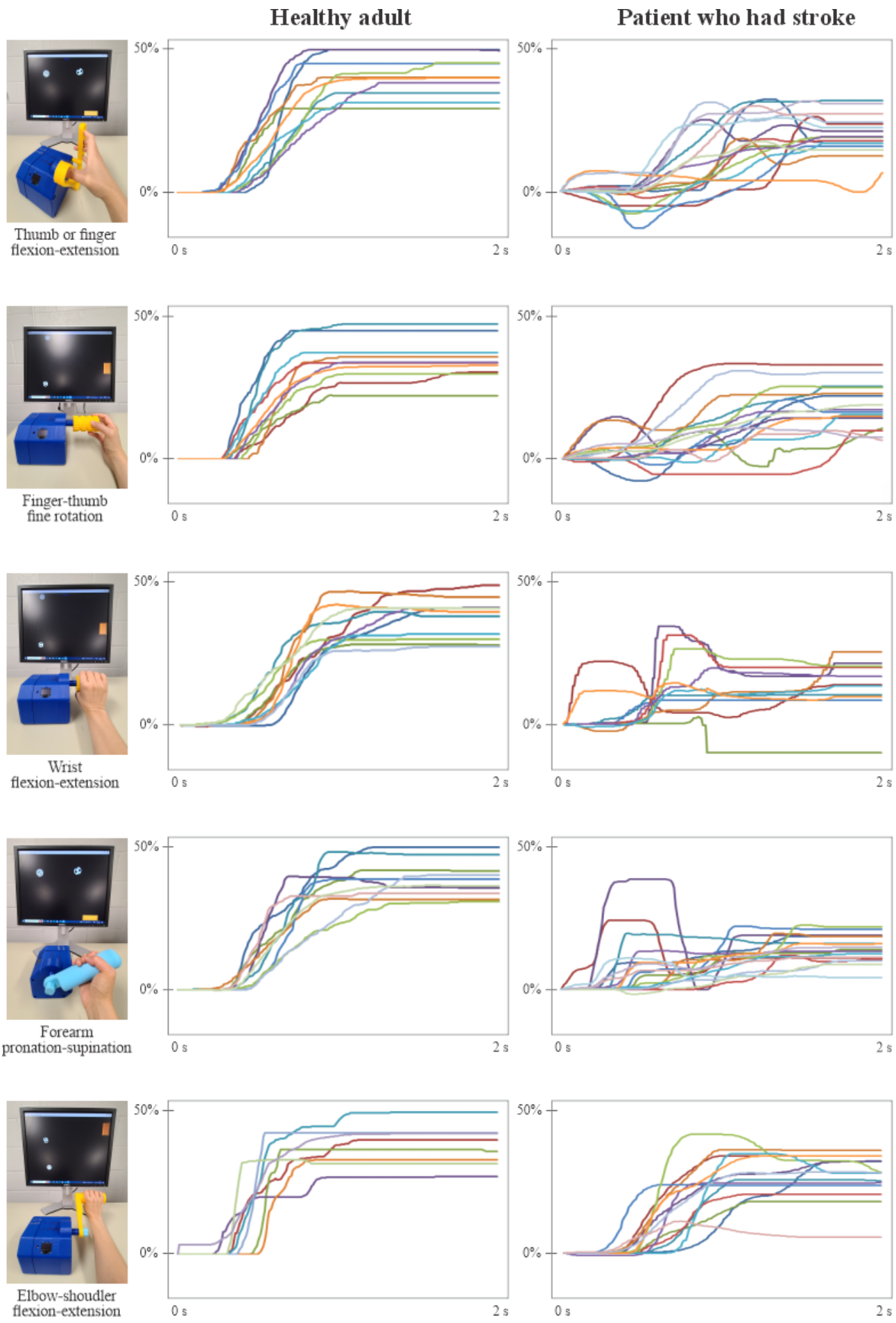
Figure 2. Illustration of RTP game software using the robotic manipulandum device. Panel (A) shows a healthy adult rotating a handle to move the game "paddle" and catch the "target" object while avoiding the "distractor" object. Panel (B) shows a screenshot of the game, where target and distractor objects appear at the top of the display and move to the bottom. Note that the addition of distractors is optional. Panel (C) presents single-game movement trajectories (game paddle coordinates) for all game movement responses in one session. In this example, each game event takes 2 seconds (from target appearance to disappearance), and the game is played for 60 seconds. The location of each successive target appearance is randomized. Approximately half of the 30 game events occur in each direction (leftward or rightward). Panel (D) presents overlay plots of the segmented and sorted game movement trajectories for all 30 game events; upward traces indicate leftward game movements, and downward traces indicate rightward game movements. RTP: repetitive task practice.



It is important to note that with standard commercial games, automatic performance logging is typically unavailable. Therefore, for any training sessions—particularly those conducted at home or remotely—the RTP game developed in-house serves as a valuable resource, providing automated monitoring and quantification of players' motor skills while engaging in a range of game-based exercises for hand and arm coordination (also referred to as telemonitoring).

Figure 3 presents game movement trajectories of a representative able-bodied adult and a patient with stroke playing the RTP game using various handles. As can be seen in the plots, the trajectories of the 5 different manipulation tasks are similar. The SR was 100% for all manipulation tasks. For the patient with stroke, the SR ranged from 50% (thumb-finger flexion-extension) to 80% (elbow flexion-extension). Movement consistency among the 10 - 12 game movement responses of the able-bodied adult was similar, as was movement onset time (MOT). Many of the movement trajectories of the patient with stroke were not smooth, exhibited small amplitudes, and demonstrated several target overshoots. It is also evident that the MOT is delayed in the participant with stroke compared to the able-bodied participant. It is, therefore, seen that the platform presented here is functional and can produce meaningful data for further analysis and treatment decisions.

Figure 3. Repetitive task practice game movement trajectories of an able-bodied adult and a patient with stroke using various robotic manipulandum device handles, as described in Figure 2. The plots show segmented and sorted game movement responses for 1 direction of movement. The y-axis represents movement amplitude as a percentage of screen width (0% to 50%).



Another key feature of the RMD and the associated RTP game software is that it is designed for use at home (telerehabilitation), not just in clinics. At present, the cost of the device is estimated to be less than US $70. The software automatically collects various objective outcome measures to monitor a patient's ongoing progress instantaneously and can be traced over a period of time. These data support the development of sustainable, individualized, long-term rehabilitation protocols. Furthermore, clinical support for home and remote outreach programs can facilitate the creation of more targeted and personalized solutions for patients.

The objective of this pilot study was to evaluate the implementation, usability, acceptability, and benefits of the game-based exercise program using the developed RMD presented in this paper. The experience of participants with stroke who completed a 6-week game-based exercise program was first assessed with semistructured interviews. Interviews were conducted to investigate participants' perspectives and opinions about expectations, acceptability, challenges, and benefits of the game-based exercise program for UE rehabilitation. Quantitative analysis pre- to postintervention was conducted next, which included the Wolf Motor Function Test (WMFT) and a computerized performance-based assessment of manual dexterity.

Methods

Recruitment

Participants were recruited at the clinical rehabilitation research facility of the University of Manitoba. In total, 5 individuals who had a single stroke (onset between 6 months and 5 years) and were aged 40 to 70 years participated in the study. All participants had adequate vision to see images on a standard computer monitor. Exclusion criteria were (1) excessive spasticity of the fingers and wrist (grade 2 and above on the Modified Ashworth Scale [35], (2) significant cognitive impairment (Montreal Cognitive Assessment scores less than 25 [36], and (3) any other neurological disorder except a single stroke before testing.

Ethical Considerations

The University of Manitoba Ethics Board reviewed and approved the study (approval HS25163), and all participants provided informed consent. The consent process ensured participants comprehended the study's objectives, procedures, potential risks, benefits, and their right to discontinue at any time. To maintain participant anonymity, all collected data were anonymized and stored in a secure, locked location. No compensation was provided, and no photographic or video recordings of participants were taken.

Exercise Program

Participants attended 12 treatment sessions twice a week for 6 weeks. Each session lasted 45 minutes. As shown in Figures 1 and 3, a variety of 3D-printed "therapy" handles of different shapes and sizes were used. They were designed to practice a broad range of manual dexterity skills. The exercise programs were established based on the participants' personal goals, the degree of their hemiparesis, and functional status. A typical session involved exercise with 4 to 5 different handles, several computer games, and assistive forces of different magnitudes. Each handle-game-force combination was practiced for 2- to 3-minute intervals and repeated 2 to 3 times. Different handles required different modes of manipulation. Game movement responses were produced by thumb, finger, wrist, or elbow movements. Task demands were adjusted by changing the mouse sensitivity movement range and by adding assistive-resistive forces. Additionally, different games were selected to adjust movement speed and precision. Most participants were competitive and became frustrated if they were not successful in gameplay. Therefore, the difficulty level (movement amplitude, speed, and precision) was adjusted for all combinations of handles, game settings, and game types so that participants were successful in gameplay for at least 60% of the game events or activities. Table 1 presents a list of common computer games used in this study. Big Fish Games are selected based on the level of difficulty participants reported and their personal likes and dislikes. The choices of games presented to them were based on columns 3 to 7 of Table 1. Games with an easy level of difficulty (based on the level of precision required, the presence of distractors, and the type of executive functions required) were introduced before the moderate and difficult games. Task difficulty was also adjusted by increasing the assistive and resistive forces applied to the RMD handles.

The exercises and choice of games were updated on a regular basis, based on the participants' improvements and personal preferences for game selection. Numerous affordable and readily accessible computer video games offer therapeutic benefits. For instance, computer games downloaded from Big Fish Games feature hundreds of arcade-style games across various genres (Table 1). Many of these games align well with the game-based RMD exercise program. In addition to requiring speed and accuracy, these games incorporate several cognitive elements, such as speed versus accuracy dynamics, distractor objects, and object-matching activities. The commercial computer games used in this pilot study are listed in Table 1. The wide variety of games ensures that the individual preferences of participants can be fulfilled. Regularly introducing new games and increasing the difficulty levels can help maintain the challenge, providing the psychological feedback necessary to keep participants engaged and motivated.

Qualitative Analysis

At the end of the 6-week exercise program, all participants were invited to participate in an interview. They were asked a series of open-ended questions, and their responses were documented: (1) when you agreed to participate, how did you hope you would benefit from the therapy program? (2) Were there things about the game or exercise program you liked and things you did not like? (3) What did you think about the computer games that you were asked to play? Did you enjoy the game? Were there games that you did not enjoy? (4) Did you feel that this therapy program helped you? (5) If you were provided with the right settings, would you continue with these exercises?

The duration of the interviews varied among the 5 participants, lasting between 20 and 30 minutes. Participants were invited to share their thoughts, ideas, opinions, and personal experiences

in detail. The analytical framework of interpretive description was used for thematic analysis [37]. All interviews were recorded, and the interviewer's notes and comments were added to the transcriptions separately for triangulation purposes. One researcher (AK) reviewed the translated transcripts and created a coding system, while a second researcher (TJS) oversaw the process and added any additional codes for credibility purposes. A second researcher (TJS) then examined the coded data to identify any unique responses.

The content of each interview was analyzed by paraphrasing, generalizing, and abstracting. A continuous iterative process was maintained until no new themes emerged from the data. The 2 researchers then compared their analyses and resolved any disagreements in a final coding system organized into final themes and subthemes.

Quantitative Analysis

Overview

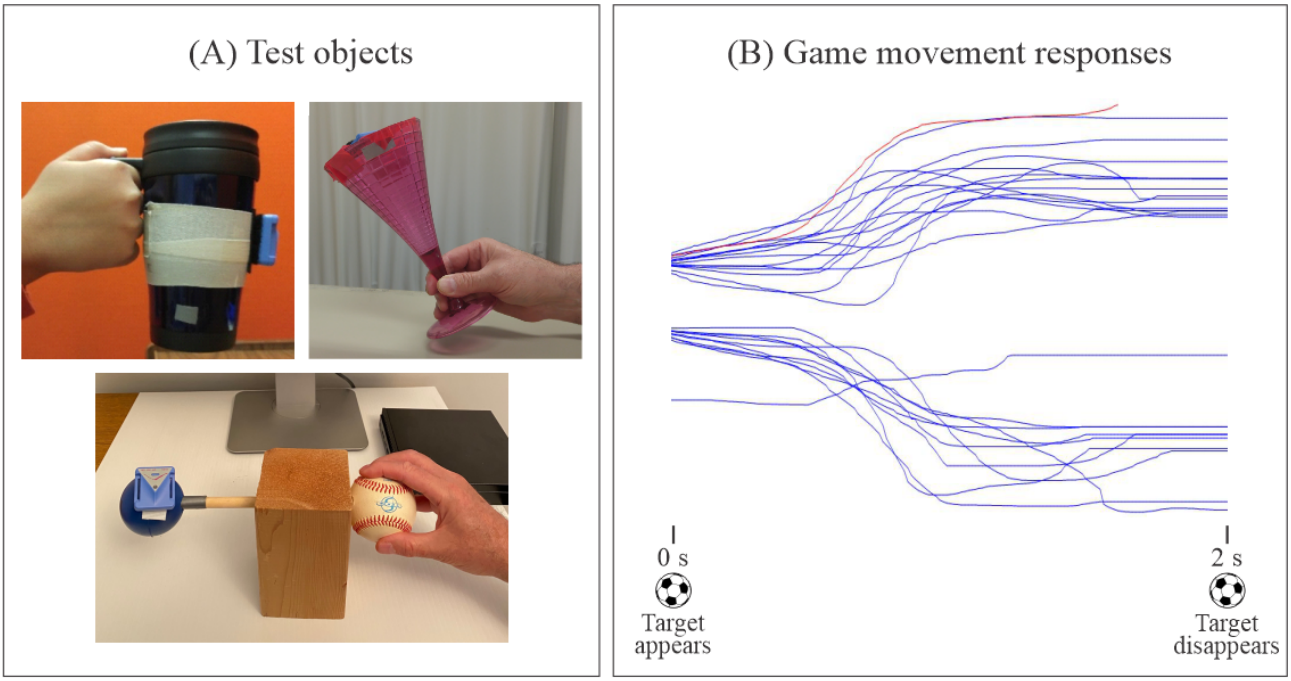
The following outcome measures were obtained before and after the intervention of the 6-week exercise program:

1. Quantitative assessments of UE motor ability were conducted using the WMFT [38,39]. Participants were instructed to complete the 15 tasks of the WMFT within a 120-second time limit, and the time taken to complete each task was recorded. Additionally, the quality of movement for each task was evaluated using an ordinal scale ranging from 0 to 5, where 0 indicates no performance and 5 indicates normal movement. The final WMFT scores were the total time taken for the 15 tasks and the summed movement quality grades of the 15 tasks.
2. The RTP game was used to guide and evaluate different object manipulation tasks. In this application, several test objects with different physical properties and anatomical demands were instrumented with a wireless IB mouse. The rotation of each test object (ie, the instantaneous angular position of the IB mouse) controlled the motion of the game paddle. For a detailed description of the assessment tool, see references [20,33]. All tasks required precision in object manipulation using the finger-thumb or hand palmar surface.

Test Objects

In the context of the RTP game assessment, the following test object manipulation tasks were evaluated, as illustrated in Figure 4: (1) participants grasped a coffee mug to move it with concentric pronation and eccentric supination. (2) Participants held a wine glass between the thumb, index, and middle fingers. It was rotated forward and backward using radial and ulnar deviation. (3) Participants grasped a tennis ball with the thumb and fingertips tethered to a wooden block via a wooden dowel to eliminate the gravity effect. This task required the participant to rotate the tennis ball left and right.

Figure 4. Illustration of the computer game–based upper extremity assessment tool. (A) Three test objects, each equipped with an inertial-based mouse, were used to control the repetitive task practice game paddle movements. (B) Example overlay plots of the segmented and sorted game movement responses for both movement directions: pronation-supination using a coffee mug, ulnar-radial deviation using a wine glass, and leftward-rightward rotation using a tennis ball.



The assessment presented here allows one to determine if there is a transfer of improvement in manual dexterity with objects used in daily life. Moderate to high test-retest reliability of the assessment tool has been reported in a group of 30 patients with stroke [20] and a group of 35 children with cerebral palsy [21].

Test Protocol

Participants were seated with test objects positioned within a comfortable reaching distance on an adjustable-height table. A 50-cm wide computer monitor was placed 1 m in front of them at eye level to perform the assessment game tasks. Participants received a demonstration of the game tasks and were allowed

to practice trials using their unaffected arms. Figure 4 presents typical overlay plots of game movement trajectories for both movement directions for 1 game session.

Outcome Measures

The following outcome measures were derived from the recorded game data of the assessments: SR and average MOT. The percentage of the total number of target objects caught in 1 game trial is the SR. The time from target appearance to the start of the game paddle movement is the average MOT. MOT values are determined for each game movement response. The average is then computed over the group of game movement responses for each direction.

Results

Participants

Table 2 presents the demographic and clinical data of the 5 participants. All participants, who experienced a single stroke, agreed to take part in the study and provided informed consent. They were all right-handed and fully completed the 6-week program, which included 2 exercise sessions per week, each lasting at least 45 minutes.

Table . Demographic and clinical characteristics of participants.

Participant	Age (years)	Sex	Type of stroke	Duration (months)	Affected side	Hand dominance
Participant 1	67	Male	Ischemic	24	Left	Right
Participant 2	68	Male	Ischemic	16	Left	Right
Participant 3	57	Male	Ischemic	12	Left	Right
Participant 4	43	Female	Ischemic	4	Left	Right
Participant 5	51	Male	Hemorrhagic	56	Left	Right

Qualitative Results

The following 4 themes capture the range of participants' experiences and viewpoints regarding the prototyped RMD exercise program: expectations, difficulties with technology, engagement with therapy, and future expectations. Table 3 presents examples of participants' direct quotes for each interview question (theme).

Table . Typical participant responses to interview questions.

Theme	Response
Expectations	• "I get into problems while handling day-to-day things. I often have trouble gauging how much distance and pressure I need. The other day I squeezed the soda cup too hard and spilled everywhere. I am hoping to improve the finer aspects" [Participant 1]. • "My consultant physician told us that I was never going to use my fingers. When we heard about this program, we thought it might help" [Participant 3].
Difficulties with technology	• "Learning how to use the RMD was hard at first. Learning how to move the mouse when my arm is so restricted, you know?" [Participant 3]. • "I am not a tech-savvy person. It took a while to get used to the games and the robot (RMD)" [Participant 4]. • "Coming to therapy twice a week and getting a ride in winter was a lot. But the home-based therapy was not working, so we decided to do it" [Participant 5].
Engagement with therapy	• "I am very competitive. I like that the computer games challenged me. It was fun" [Participant 1]. • "It (conventional therapy) did not show much improvement. It did not seem like it was worth the trouble. I wanted to check out this option (computer games-based protocol) because it sounded new, something fun" [Participant 4]. • "I could comb my hair again. That was something!" [Participant 2]. • "My hand felt completely immobile earlier; now I can use it to support my other hand for different tasks" [Participant 3]. • "I was happy to see that you guys created a steering wheel handle for me to relearn driving" [Participant 5].
Future expectations	• "Honestly, I think we could have done this from home if you had enough of these (RMDs). We can just download these games on my laptop" [Participant 5].

Expectations

All participants indicated that the primary reason for their participation in this exercise program was to improve their hand function, particularly in handling and manipulating objects. One participant agreed to participate because his therapist recommended the program. It is noteworthy that it had been several weeks to over 3 years since the participants last received physiotherapy or occupational therapy.

Difficulties With Technology

In total, 3 of the 5 participants reported that they had not played computer games before. However, they noted that the games were easy to learn. All participants found it intuitive to use the RMD as a game controller. They all considered the exercise program challenging and expressed that it was difficult to play the games by manipulating the RMD handle. Nevertheless, with practice, the exercises became significantly easier. All participants exhibited competitiveness and experienced frustration when they could not successfully play the games. This was taken into account, and the games were carefully selected to match the skill levels of each participant.

Engagement With Therapy

All participants stated that they had previously undergone physiotherapy and occupational therapy for several weeks. They expressed appreciation for the one-to-one therapy sessions, noting that receiving immediate feedback and guidance from the therapist was very helpful. Participants also reported that it was beneficial to know which games to use and why. They indicated a preference for certain games and appreciated the variety available to them during therapy. Furthermore, all participants commented that it was more enjoyable and easier to perform game-based exercises than conventional exercises.

Future Expectations

All participants expressed a desire to continue the program and inquired whether it would be possible to use the device at home.

Quantitative Results

Table 4 presents the pre- and postintervention test scores for the WMFT, highlighting the changes observed from pre- to postintervention. In patients with stroke, the minimal clinically important difference (MCID) for the functional ability score has been reported to range from 3 to 6 points, while the MCID for the total time of the WMFT is 22 seconds [40]. All 5 participants in this study demonstrated postintervention improvements that exceeded the reported MCID for the functional ability measure (with a range of improvement between 9 and 16 points). In total, 4 of the 5 participants exhibited improvements in total time that surpassed the MCID (with a range of improvement between 23 and 28 seconds).

Table . Pre- and postexercise Wolf Motor Function Test scores and magnitude of change.

Participant	Functional ability score (maximum: 75)			Total time (seconds)		
	Pre	Post	Change	Pre	Post	Change
Participant 1	19	28	9	119	71	48
Participant 2	19	34	15	118	84	34
Participant 3	13	27	14	73	62	11
Participant 4	12	28	16	104	71	33
Participant 5	9	23	14	89	65	24
Average (SD)	14 (4)	28 (4)	14 (3)	101 (20)	71 (8)	30 (14)

Figure 5 displays example plots of game movement responses using the 3 test objects, recorded at baseline and after the completion of the 6-week exercise program from different participants. Visual inspection reveals a clear improvement in movement quality, amplitude, and consistency. As indicated in Table 5, substantial improvements were observed in SR and response time for all 5 patients. For SR, the average improvement was 23% (SD 12%), while for response time, there was an average decrease of 105 (SD 44) milliseconds. It is noteworthy that typical response times were approximately 600 milliseconds.

Figure 5. Examples of repetitive task practice game movement responses (1 direction) from different participants using the 3 assessment test objects, taken at pre- and postintervention time periods.

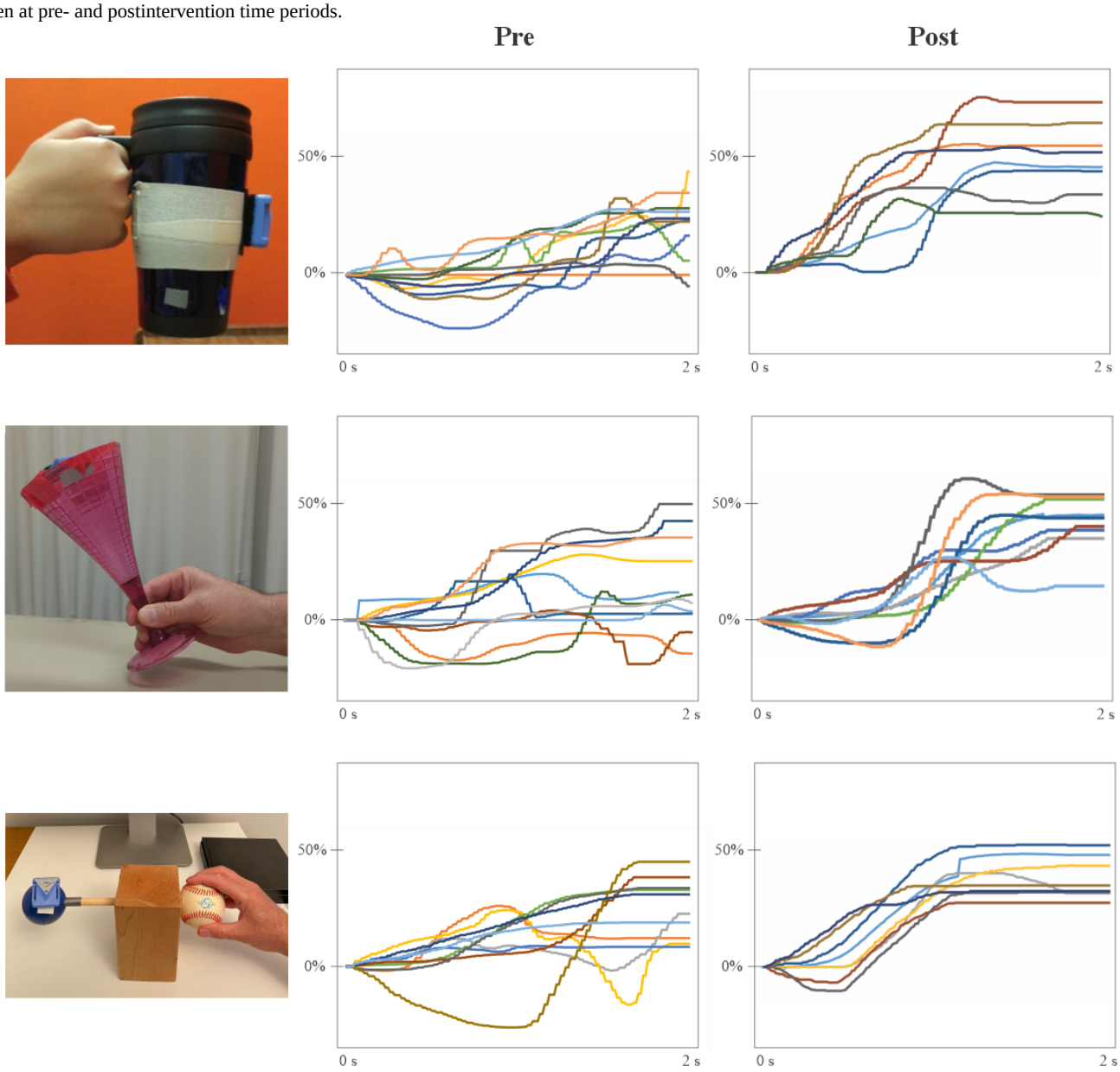


Table . Pre- and postexercise test scores and magnitude change of object manipulation tasks for all participants[a].

Participant and test object		Success rate (%)			Response time (milliseconds)		
		Pre	Post	Change	Pre	Post	Change
Participant 1							
	Object 1	42	91	49	723	530	193
	Object 2	65	89	24	478	423	55
	Object 3	51	68	17	792	697	95
	Average (SD)	53 (12)	83 (13)	30 (17)	664 (165)	550 (138)	114 (71)
Participant 2							
	Object 1	50	100	50	528	532	−4
	Object 2	54	62	8	780	702	78
	Object 3	33	100	67	771	421	350
	Average (SD)	46 (11)	87 (22)	42 (30)	693 (143)	552 (142)	141 (185)
Participant 3							
	Object 1	42	64	22	845	811	34
	Object 2	70	75	5	845	771	74
	Object 3	61	73	12	794	805	−11
	Average (SD)	58 (14)	71 (6)	13 (9)	828 (29)	796 (22)	32 (43)
Participant 4							
	Object 1	54	85	31	576	490	86
	Object 2	50	55	5	786.	515	271
	Object 3	66	80	14	784	725	59
	Average (SD)	57 (8)	73 (16)	17 (13)	715 (121)	577 (129)	139 (115)
Participant 5							
	Object 1	66	78	12	823	783	40
	Object 2	72	92	20	806	588	218
	Object 3	56	72	16	760	722	38
	Average (SD)	65 (8)	81 (10)	16 (4)	796 (33)	698 (100)	99 (103)

[a]Values are the average of left and right game movements. Object 1: coffee mug; object 2: wine glass; and object 3: tennis ball (Figure 5).

Discussion

Principal Findings

This paper introduced a rehabilitation device that provides flexible, game-based RTP targeting manual dexterity and includes means to automatically record and assess patients' manual dexterity skills using the RTP software. The 6-week exercise program resulted in clinically significant improvement. In terms of the WMFT, on average, participants showed an improvement of 14 (SD 3) points in functional ability score and a reduction of 30 (SD 14) seconds in total time. Additionally, for the computer game–based UE assessment, the average improvement in success rate was 23% (SD 12%), while the average decrease in response time was 105 (SD 44) milliseconds. The proposed system not only addresses patients' exercise needs but also integrates enjoyment and learning through a gaming platform.

The change in WMFT scores exceeded the MCID for all participants. The WMFT measures daily activities involving fingers, such as picking up small objects and using hand tools. Significant improvements in the WMFT were observed, even though these specific tasks were not practiced during the game-based manipulation program. The WMFT also assesses visual perceptual skills for tasks like stacking blocks and drawing figures. The RMD game tasks, which require precision movements based on visual feedback, showed substantial improvements in both the WMFT tasks and object manipulation tasks in the RTP game. Participants with stroke noted that the game-based exercises were challenging yet engaging and enjoyable.

Handles of different sizes and shapes were used to target precision, goal-directed movements of the thumb, fingers, and wrist as well as combinations of UE movements. In addition to the types of handles used, computer games also possess therapeutic value. Different commercial video games require varying levels of movement speed, accuracy, and amplitude.

For example, participants with severe impairments were able to successfully play computer games when the selected games involved slow movements and low precision (ie, large game paddles and target game objects). Participants with moderate to mild impairments could engage with computer games that required faster speeds and greater precision.

The games also involved various executive cognitive functions, including visual search and spatial processing of moving targets and distractors. The diverse range of games and regular updates to difficulty levels are important for maintaining engagement and challenge.

The RMD is configured to function exactly as a plug-and-play computer mouse and, therefore, can be used to play many commercial computer games. This allows easy access to a large source of commercial games. To meet the needs of each individual, the RMD can be customized to suit specific rehabilitation needs and preferences. This adaptability allows patients to engage with a variety of gaming experiences tailored to their specific motor and cognitive rehabilitation goals. A key feature of the program is to increase the number of repetitions of goal-directed movements at varying speeds and accuracy levels. High intensity and a high number of repetitions are crucial to drive neuroplasticity and functional improvement in patients with stroke [41-45].

Each game-handle combination was played for 3 to 5 minutes, and typically, each game event took approximately 2 seconds. Therefore, participants made 90 to 150 goal-directed game movement responses during this time period. Each session lasted 45 minutes and included 7 to 8 different handles, resulting in several hundred game-handle combinations. The goal-directed game movement responses varied in amplitude, speed, and direction. During gameplay, visual feedback of the game sprite or paddle relative to the game target and distractor objects was used to initiate and guide each contextual game movement response, supporting implicit learning of eye-hand coordination. Additionally, the selected video games featured unpredictable trajectories for game target motion, promoting variable practice.

Interestingly, significant improvements were observed in a participant who was 5 years after a stroke, which was unexpected given that most studies include participants less than 2 years after a stroke. Although some studies have reported significant improvements in UE function 3 to 5 years after a stroke, this finding is based on only 1 participant. Future randomized controlled trials are needed to examine the effectiveness of game-based task-specific exercises for participants 3 to 5 years after a stroke.

Recovery programs can be extensive, involving RTP for many months. A key feature of the RMD is its design for home use (telerehabilitation). In this regard, the cost of the electronic components, motor housing, and handles is less than US $70. Additional costs, several times this amount, will likely be required for the commercialization of the RMD system. The RMD can initially be used in a supervised clinical setting and then transitioned to home use while being monitored by clinicians. The telemonitoring capabilities of the system (ie, RTP game) could allow clinicians to track changes in function and compliance, facilitating the development of sustainable and individualized programs. Prompt clinical assistance for home and remote outreach programs will foster more tailored and effective solutions for patients, facilitating the intended training outcomes. This will require further development to produce a secure content management system for individual electronic game data to be updated and stored for processing as well as to generate queries and reports for registered eHealth stakeholders (eg, therapists, physicians, and third-party insurance providers).

Limitations

The unidirectional force mode, while assisting movement in 1 direction, results in resistance forces in the opposite direction, which may not be desirable. A real-time intelligent control scheme is under development, involving communication between the RMD software and the RTP game. In the upcoming system, the controller will receive coordinates for both the game targets and the paddle, which is controlled by handle rotations. This information about movement directions and amplitudes can then be used. The system will determine the direction and magnitude of the force necessary to rotate the handles effectively to move the game paddle within the RTP game. Notably, this closed-loop assistance can be offered in both movement directions during gameplay. This context-sensitive assistive mode helps facilitate limited voluntary movements in severely affected individuals.

Conclusions

The results of the pilot study indicate the feasibility, acceptability, and positive outcomes of the RMD game–based system for enhancing manual dexterity in people with stroke who have moderate UE motor impairments. The intervention resulted in clinically significant improvements, with all participants showing enhanced performance in the WMFT beyond the MCID. These findings suggest that the system has the potential to advance rehabilitation treatments for finger, thumb, and wrist recovery in people with stroke. The long-term effects of this training on manual dexterity will need to be evaluated in future randomized controlled trials. However, the current findings are encouraging and provide a strong basis for further research and development.

Acknowledgments

This study was supported by grants from the University of Manitoba Collaborative Innovation Research Fund, Research Manitoba Innovation, and the Natural Sciences and Engineering Research Council of Canada.

Conflicts of Interest

None declared.

References

1. Byblow WD, Stinear CM, Barber PA, Petoe MA, Ackerley SJ. Proportional recovery after stroke depends on corticomotor integrity. Ann Neurol 2015 Dec 1;78(6):848-859. [doi: 10.1002/ana.24472] [Medline: 26150318]
2. Sobinov AR, Bensmaia SJ. The neural mechanisms of manual dexterity. Nat Rev Neurosci 2021 Dec 1;22(12):741-757. [doi: 10.1038/s41583-021-00528-7] [Medline: 34711956]
3. Ward NS, Brander F, Kelly K. Intensive upper limb neurorehabilitation in chronic stroke: outcomes from the Queen Square programme. J Neurol Neurosurg Psychiatry 2019 May 1;90(5):498-506. [doi: 10.1136/jnnp-2018-319954] [Medline: 30770457]
4. Cramer SC, Dodakian L, Le V, et al. Efficacy of home-based telerehabilitation vs in-clinic therapy for adults after stroke: a randomized clinical trial. JAMA Neurol 2019 Sep 1;76(9):1079-1087. [doi: 10.1001/jamaneurol.2019.1604] [Medline: 31233135]
5. Mawase F, Cherry-Allen K, Xu J, Anaya M, Uehara S, Celnik P. Pushing the rehabilitation boundaries: hand motor impairment can be reduced in chronic stroke. Neurorehabil Neural Repair 2020 Aug;34(8):733-745. [doi: 10.1177/1545968320939563] [Medline: 32845230]
6. Daly JJ, McCabe JP, Holcomb J, Monkiewicz M, Gansen J, Pundik S. Long-dose intensive therapy is necessary for strong, clinically significant, upper limb functional gains and retained gains in severe/moderate chronic stroke. Neurorehabil Neural Repair 2019 Jul;33(7):523-537. [doi: 10.1177/1545968319846120] [Medline: 31131743]
7. Dale G, Joessel A, Bavelier D, Green CS. A new look at the cognitive neuroscience of video game play. Ann N Y Acad Sci 2020 Mar;1464(1):192-203. [doi: 10.1111/nyas.14295] [Medline: 31943260]
8. Bediou B, Adams DM, Mayer RE, Tipton E, Green CS, Bavelier D. Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. Psychol Bull 2018 Jan;144(1):77-110. [doi: 10.1037/bul0000130] [Medline: 29172564]
9. Choi E, Shin SH, Ryu JK, Jung KI, Kim SY, Park MH. Commercial video games and cognitive functions: video game genres and modulating factors of cognitive enhancement. Behav Brain Funct 2020 Feb 3;16(1):2. [doi: 10.1186/s12993-020-0165-z] [Medline: 32014027]
10. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DML. The effects of gamification on computerized cognitive training: systematic review and meta-analysis. JMIR Serious Games 2020 Aug 10;8(3):e18644. [doi: 10.2196/18644] [Medline: 32773374]
11. Sheng B, Zhao J, Zhang Y, Xie S, Tao J. Commercial device-based hand rehabilitation systems for stroke patients: state of the art and future prospects. Heliyon 2023 Mar;9(3):e13588. [doi: 10.1016/j.heliyon.2023.e13588] [Medline: 36873497]
12. Ma M, Proffitt R, Skubic M. Validation of a Kinect V2 based rehabilitation game. PLoS ONE 2018;13(8):e0202338. [doi: 10.1371/journal.pone.0202338] [Medline: 30142631]
13. Zanatta F, Farhane-Medina NZ, Adorni R, et al. Combining robot-assisted therapy with virtual reality or using it alone? A systematic review on health-related quality of life in neurological patients. Health Qual Life Outcomes 2023 Feb 21;21(1):18. [doi: 10.1186/s12955-023-02097-y] [Medline: 36810124]
14. Lassi M, Dalise S, Bandini A, et al. Neurophysiological underpinnings of an intensive protocol for upper limb motor recovery in subacute and chronic stroke patients. Eur J Phys Rehabil Med 2024 Feb;60(1):13-26. [doi: 10.23736/S1973-9087.23.07922-4] [Medline: 37987741]
15. Salvalaggio S, Gianola S, Andò M, et al. Predictive factors and dose-response effect of rehabilitation for upper limb induced recovery after stroke: systematic review with proportional meta-analyses. Physiotherapy 2024 Dec;125:101417. [doi: 10.1016/j.physio.2024.101417] [Medline: 39395360]
16. Roby-Brami A, Jarrassé N, Parry R. Impairment and compensation in dexterous upper-limb function after stroke. From the direct consequences of pyramidal tract lesions to behavioral involvement of both upper-limbs in daily activities. Front Hum Neurosci 2021;15:662006. [doi: 10.3389/fnhum.2021.662006] [Medline: 34234659]
17. Ghannadi B, Sharif Razavian R, McPhee J. Configuration-dependent optimal impedance control of an upper extremity stroke rehabilitation manipulandum. Front Robot AI 2018;5(NOV):124. [doi: 10.3389/frobt.2018.00124] [Medline: 33501003]
18. Hong R, Li B, Bao Y, Liu L, Jin L. Therapeutic robots for post-stroke rehabilitation. Med Rev (2021) 2024 Feb;4(1):55-67. [doi: 10.1515/mr-2023-0054] [Medline: 38515779]
19. Szturm T, Imran Z, Pooyania S, Kanitkar A, Mahana B. Evaluation of a game based tele rehabilitation platform for in-home therapy of hand-arm function post stroke: feasibility study. PM R 2021 Jan;13(1):45-54. [doi: 10.1002/pmrj.12354] [Medline: 32107868]
20. Szturm T, Imran Z, Pooyania S, et al. A new method to evaluate hand function in clients with stroke: test-retest reliability and discriminative validity. Physiother Res Rep 2024;5:1-7. [doi: 10.15761/PRR.1000136]

21. Kanitkar A, Parmar ST, Szturm TJ, et al. Evaluation of a computer game-assisted rehabilitation program for manual dexterity of children with cerebral palsy: feasibility randomized control trial. PM R 2023 Oct;15(10):1280-1291. [doi: 10.1002/pmrj.12947] [Medline: 36655404]
22. Parmar ST, Kanitkar A, Sepehri N, Bhairannawar S, Szturm T. Computer game-based telerehabilitation platform targeting manual dexterity: exercise is fun. “You are kidding-right?”. Sensors (Basel) 2021 Aug 27;21(17):5766. [doi: 10.3390/s21175766] [Medline: 34502656]
23. Wu J, Cheng H, Zhang J, Yang S, Cai S. Robot-assisted therapy for upper extremity motor impairment after stroke: a systematic review and meta-analysis. Phys Ther 2021 Apr 4;101(4):pzab010. [doi: 10.1093/ptj/pzab010] [Medline: 33454787]
24. Ko MJ, Chuang YC, Ou-Yang LJ, Cheng YY, Tsai YL, Lee YC. The application of soft robotic gloves in stroke patients: a systematic review and meta-analysis of randomized controlled trials. Brain Sci 2023 Jun 2;13(6):900. [doi: 10.3390/brainsci13060900] [Medline: 37371378]
25. Ketkar VD, Wolbrecht ET, Perry JC, Farrens A. Design and development of a spherical 5-bar thumb exoskeleton mechanism for poststroke rehabilitation. J Med Device 2023 Jun 1;17(2):021002. [doi: 10.1115/1.4056864] [Medline: 37152413]
26. Alhamad R, Seth N, Abdullah HA. Initial testing of robotic exoskeleton hand device for stroke rehabilitation. Sensors (Basel) 2023 Jul 12;23(14):6339. [doi: 10.3390/s23146339] [Medline: 37514633]
27. Wang T, Liu Z, Gu J, Tan J, Hu T. Effectiveness of soft robotic glove versus repetitive transcranial magnetic stimulation in post-stroke patients with severe upper limb dysfunction: a randomised controlled trial. Front Neurol 2022;13:887205. [doi: 10.3389/fneur.2022.887205] [Medline: 36712422]
28. Shin S, Lee HJ, Chang WH, Ko SH, Shin YI, Kim YH. A smart glove digital system promotes restoration of upper limb motor function and enhances cortical hemodynamic changes in subacute stroke patients with mild to moderate weakness: a randomized controlled trial. J Clin Med 2022 Dec 10;11(24):7343. [doi: 10.3390/jcm11247343] [Medline: 36555960]
29. Moggio L, de Sire A, Marotta N, Demeco A, Ammendolia A. Exoskeleton versus end-effector robot-assisted therapy for finger-hand motor recovery in stroke survivors: systematic review and meta-analysis. Top Stroke Rehabil 2022 Dec;29(8):539-550. [doi: 10.1080/10749357.2021.1967657] [Medline: 34420498]
30. Ahn SY, Bok SK, Lee JY, et al. Benefits of robot-assisted upper-limb rehabilitation from the subacute stage after a stroke of varying severity: a multicenter randomized controlled trial. J Clin Med 2024 Jan 30;13(3):808. [doi: 10.3390/jcm13030808] [Medline: 38337500]
31. Lockery D, Peters JF, Ramanna S, Shay BL, Szturm T. Store-and-feedforward adaptive gaming system for hand-finger motion tracking in telerehabilitation. IEEE Trans Inf Technol Biomed 2011 May;15(3):467-473. [doi: 10.1109/TITB.2011.2125976] [Medline: 21536526]
32. Kanitkar A, Parmar ST, Szturm TJ, et al. Reliability and validity of a computer game-based tool of upper extremity assessment for object manipulation tasks in children with cerebral palsy. J Rehabil Assist Technol Eng 2021;8:20556683211014023. [doi: 10.1177/20556683211014023] [Medline: 34123406]
33. Imran Z, Szturm T, Pooyania S. Development of a computerized assessment tool for hand-arm function after stroke: test-retest reliability and convergent validity. J Hand Ther 2017 Oct;30(4):e12. [doi: 10.1016/j.jht.2017.06.016]
34. PC games. Big Fish. URL: https://www.bigfishgames.com/us/en/pc-games.html [accessed 2025-02-27]
35. Pandyan AD, Price CIM, Barnes MP, Johnson GR. A biomechanical investigation into the validity of the modified Ashworth Scale as a measure of elbow spasticity. Clin Rehabil 2003 May;17(3):290-293. [doi: 10.1191/0269215503cr610oa] [Medline: 12735536]
36. Dalrymple-Alford JC, MacAskill MR, Nakas CT, et al. The MoCA: well-suited screen for cognitive impairment in Parkinson disease. Neurology (ECronicon) 2010 Nov 9;75(19):1717-1725. [doi: 10.1212/WNL.0b013e3181fc29c9] [Medline: 21060094]
37. Teodoro IPP, Rebouças VDC, Thorne SE, Souza ND, Brito LD, Alencar A. Interpretive description: a viable methodological approach for nursing research. Esc Anna Nery 2018 Mar 19;22(3). [doi: 10.1590/2177-9465-ean-2017-0287]
38. Wolf SL, Catlin PA, Ellis M, Archer AL, Morgan B, Piacentino A. Assessing Wolf Motor Function Test as outcome measure for research in patients after stroke. Stroke 2001 Jul;32(7):1635-1639. [doi: 10.1161/01.str.32.7.1635] [Medline: 11441212]
39. Morris DM, Uswatte G, Crago JE, Cook EW 3rd, Taub E. The reliability of the Wolf Motor Function Test for assessing upper extremity function after stroke. Arch Phys Med Rehabil 2001 Jun;82(6):750-755. [doi: 10.1053/apmr.2001.23183] [Medline: 11387578]
40. Lin KC, Hsieh YW, Wu CY, Chen CL, Jang Y, Liu JS. Minimal detectable change and clinically important difference of the Wolf Motor Function Test in stroke patients. Neurorehabil Neural Repair 2009 Jun;23(5):429-434. [doi: 10.1177/1545968308331144] [Medline: 19289487]
41. Leiker AM, Miller M, Brewer L, Nelson M, Siow M, Lohse K. The relationship between engagement and neurophysiological measures of attention in motion-controlled video games: a randomized controlled trial. JMIR Serious Games 2016 Apr 21;4(1):e4. [doi: 10.2196/games.5460] [Medline: 27103052]
42. Muratori LM, Lamberg EM, Quinn L, Duff SV. Applying principles of motor learning and control to upper extremity rehabilitation. J Hand Ther 2013;26(2):94-102. [doi: 10.1016/j.jht.2012.12.007] [Medline: 23598082]
43. Grefkes C, Fink GR. Recovery from stroke: current concepts and future perspectives. Neurol Res Pract 2020 Jun 16;2:17. [doi: 10.1186/s42466-020-00060-6] [Medline: 33324923]

44. Ciceron C, Sappey-Marinier D, Riffo P, et al. Case report: true motor recovery of upper limb beyond 5 years post-stroke. Front Neurol 2022;13:804528. [doi: 10.3389/fneur.2022.804528] [Medline: 35250813]
45. Fleming MK, Sorinola IO, Roberts-Lewis SF, Wolfe CD, Wellwood I, Newham DJ. The effect of combined somatosensory stimulation and task-specific training on upper limb function in chronic stroke: a double-blind randomized controlled trial. Neurorehabil Neural Repair 2015 Feb;29(2):143-152. [doi: 10.1177/1545968314533613] [Medline: 24803495]

Abbreviations

IB: inertial-based
MCID: minimal clinically important difference
MOT: movement onset time
RMD: robotic manipulandum device
RTP: repetitive task practice
SR: success rate
UE: upper extremity
WMFT: Wolf Motor Function Test

Edited by P Kubben; submitted 21.10.24; peer-reviewed by A Anastasiev, F Noveletto, H Abdullah; revised version received 17.01.25; accepted 05.02.25; published 19.03.25.

Please cite as:
Kanitkar A, Sepehri N, Lezen A, Parmar ST, Hin CKF, Szturm TJ
A Game-Based Mechatronic Device for Digital Rehabilitation of Hand Function After a Stroke: Design, Prototyping, and Feasibility Study
JMIR Neurotech 2025;4:e67779
URL: https://neuro.jmir.org/2025/1/e67779
doi:10.2196/67779

Exploring Speech Biosignatures for Traumatic Brain Injury and Neurodegeneration: Pilot Machine Learning Study

Rahmina Rubaiat[1*], MSc; John Michael Templeton[2], PhD; Sandra L Schneider[3], PhD; Upeka De Silva[4], MEng; Samaneh Madanian[4], PhD; Christian Poellabauer[1*], PhD

[1]Knight Foundation School of Computer and Information Sciences, Florida International University, Miami, FL, United States

[2]Department of Computer Science and Engineering, University of South Florida, Tampa, FL, United States

[3]Department of Communicative Sciences & Disorders, Saint Mary's College, Notre Dame, IN, United States

[4]Department of Computer Science and Software Engineering, Auckland University of Technology, Auckland, New Zealand

[*]these authors contributed equally

Corresponding Author:
Rahmina Rubaiat, MSc
Knight Foundation School of Computer and Information Sciences, Florida International University, Miami, FL, United States

Abstract

Background: Speech features are increasingly linked to neurodegenerative and mental health conditions, offering the potential for early detection and differentiation between disorders. As interest in speech analysis grows, distinguishing between conditions becomes critical for reliable diagnosis and assessment.

Objective: This pilot study explores speech biosignatures in two distinct neurodegenerative conditions: (1) mild traumatic brain injuries (eg, concussions) and (2) Parkinson disease (PD) as the neurodegenerative condition.

Methods: The study included speech samples from 235 participants (97 concussed and 94 age-matched healthy controls, 29 PD and 15 healthy controls) for the PaTaKa test and 239 participants (91 concussed and 104 healthy controls, 29 PD and 15 healthy controls) for the Sustained Vowel (/ah/) test. Age-matched healthy controls were used. Young age-matched controls were used for concussion and respective age-matched controls for neurodegenerative participants (15 healthy samples for both tests). Data augmentation with noise was applied to balance small datasets for neurodegenerative and healthy controls. Machine learning models (support vector machine, decision tree, random forest, and Extreme Gradient Boosting) were employed using 37 temporal and spectral speech features. A 5-fold stratified cross-validation was used to evaluate classification performance.

Results: For the PaTaKa test, classifiers performed well, achieving F_1-scores above 0.9 for concussed versus healthy and concussed versus neurodegenerative classifications across all models. Initial tests using the original dataset for neurodegenerative versus healthy classification yielded very poor results, with F_1-scores below 0.2 and accuracy under 30% (eg, below 12 out of 44 correctly classified samples) across all models. This underscored the need for data augmentation, which significantly improved performance to 60% - 70% (eg, 26 - 31 out of 44 samples) accuracy. In contrast, the Sustained Vowel test showed mixed results; F_1-scores remained high (more than 0.85 across all models) for concussed versus neurodegenerative classifications but were significantly lower for concussed versus healthy (0.59 - 0.62) and neurodegenerative versus healthy (0.33 - 0.77), depending on the model.

Conclusions: This study highlights the potential of speech features as biomarkers for neurodegenerative conditions. The PaTaKa test exhibited strong discriminative ability, especially for concussed versus neurodegenerative and concussed versus healthy tasks, whereas challenges remain for neurodegenerative versus healthy classification. These findings emphasize the need for further exploration of speech-based tools for differential diagnosis and early identification in neurodegenerative health.

(JMIR Neurotech 2025;4:e64624) doi:10.2196/64624

KEYWORDS

speech biosignatures; speech feature analysis; amyotrophic lateral sclerosis; ALS; neurodegenerative disease; Parkinson's disease; detection; speech; neurological; traumatic brain injury; concussion; mobile device; digital health; machine learning; mobile health; diagnosis; mobile phone

Introduction

Overview

The fields of health care and medical diagnostics have witnessed a significant shift toward noninvasive and accessible methods for early detection, assessment, and monitoring of medical conditions. This shift has been driven by technological advancements and growing research interest in digital health solutions [1]. Among these, speech analysis has emerged as a promising avenue, with studies identifying speech as a potential biosignature for a variety of neurodegenerative conditions [2,3]. The ability to reliably distinguish between conditions or detect coexisting disorders is critical for accurate diagnosis, tracking disease progression, and evaluating treatment effectiveness [4].

This pilot study investigates speech-based biosignatures of 2 distinct neurodegenerative conditions, that are, neurodegenerative diseases and mild traumatic brain injuries (mTBIs), specifically concussions. Speech patterns often reflect neurodegenerative health, with specific speech features showing promise for distinguishing between these conditions. The dataset includes individuals with concussions, patients with Parkinson disease (PD), and age-matched healthy controls for both groups (15 samples for each test). These groups were selected to ensure demographic compatibility while addressing the unique speech patterns associated with each condition.

Neurodegenerative diseases, such as PD, are characterized by the progressive loss of neurons in the brain and spinal cord, leading to impairments in motor and cognitive functions [5,6]. PD involves the degeneration of dopaminergic neurons, resulting in clinical symptoms such as tremors, rigidity, bradykinesia, and postural instability [7]. These symptoms worsen over time and lack curative treatments, necessitating reliable diagnostic tools for early intervention [8]. On the other hand, concussions, a form of mTBI, result from sudden trauma to the brain, causing temporary cognitive impairments, disruptions in brain function, and neurochemical changes. Repeated concussions are associated with a heightened risk of neurodegenerative disorders, such as dementia, later in life [9]. Despite their prevalence, approximately 90% of concussions go unreported, leading to inadequate medical attention and potentially catastrophic consequences [10].

Traditional diagnostic methods for neurodegenerative diseases and concussions often rely on observable motor symptoms, such as tremors, gait disturbances, or muscle rigidity, as well as subjective assessments of cognitive impairments [11]. However, emerging research has identified speech as a valuable biomarker for neurodegenerative health. Dysarthria and dysphonia, characterized by changes in articulation and motor speech production, are prevalent in both concussions and neurodegenerative conditions like PD [12-14]. Speech features, such as mel frequency cepstral coefficients (MFCCs), jitter, shimmer, harmonics-to-noise ratio (HNR), and other temporal and spectral attributes, have been shown to correlate with underlying neurodegenerative conditions.

In this study, we analyzed speech data from 2 well-established medical speech tasks, the PaTaKa task and the Sustained Vowel task. These tasks are widely used in clinical settings for assessing speech impairments. The objective of this study is to explore the potential of speech features in differentiating between concussions and neurodegenerative conditions, as well as their respective healthy controls, and to assess the feasibility of using these features as biomarkers for diagnosis. By addressing this objective, we aim to contribute to the development of speech-based diagnostic tools for early and accurate identification of neurodegenerative health conditions.

This study evaluated 37 speech-based features (25 temporal and 12 spectral), applying machine learning models such as support vector machine (SVM), decision tree (DT), random forest (RF), and Extreme Gradient Boosting (XGBoost) to classify between the groups.

The remainder of this paper describes our methodology, feature extraction and analysis, machine learning approaches, and results for the binary classification tasks across the 2 speech tests.

Related Work

Diagnosing brain injuries and neurodegenerative diseases can be challenging; for instance, concussions may present subtle features that are difficult to detect, including using third person witness accounts of the injury, clinical examinations, and laboratory testing, where diagnostic accuracy is not always perfect [15]. Recent work has explored the diagnosis of concussions in athletes using mobile technologies [16] and speech analysis [17,18], while digital assessments, coupled with speech analysis, are also increasingly being used for individuals with neurodegenerative diseases [19]. In a study by Tsanas [19], various speech tasks have been used to distinguish between healthy people and individuals with PD, with relatively high accuracy. Other previous research has investigated the overall symptom severity of individuals with a neurodegenerative condition [11,20], the effectiveness of voice rehabilitation [21], and how to distinguish PD from other conditions such as essential tremor or atypical parkinsonism [22].

The choice of speech task is critical to obtaining speech samples that can be used for subsequent feature extraction and analysis. One commonly used speech task is to ask an individual to produce sustained phonation of vowels. For instance, the study by Mallela et al [23] presents an automatic voice assessment approach for separating healthy individuals from patients with amyotrophic lateral sclerosis (ALS). Although our study focuses exclusively on PD as the representative neurodegenerative condition, references to ALS studies are included to highlight the broader research landscape on neurodegenerative speech biosignatures and their diagnostic significance. Linear discriminant analysis is used to classify phonation, with the most successful model achieving more than 90% accuracy. Similarly, a study by Rueda and Krishnan [24] obtained sustained vowel data from 57 PD patients and 57 healthy individuals, and the study used 5 hierarchical and 1 partition-based clustering techniques to compare and cross-check PD patients at different phases. In some cases, researchers have relied on existing voice recordings, for example, obtained through the Parkinson's Voice Initiative project (the largest speech-PD dataset so far) to analyze voice impairment due to PD [25].

Daudet et al [18] developed a mobile app to diagnose concussions, using data from 47 high-schools and colleges in the Midwest. The study used several speech tasks such as repetition of a sequential motion rate, alternating motion rate, multisyllabic words (words with 4 syllables containing front, middle, and back vowels, and bilabial, alveolar, velar, and glide consonants). The work by Vashkevich et al [26] presented features for detecting pathological changes in acoustic speech signals for ALS diagnosis. It used recordings from 48 people (26 with ALS) and investigated vowel harmony. The features obtained an 88% correct classification performance using linear discriminant analysis. Various speech-based indicators, such as shimmer, jitter, HNR, and other temporal and spectral indicators, have also been explored as dysphonia measures in individuals with neurodegenerative diseases [27]. Finally, in a study by Benba et al [22], the authors investigated the most effective acoustic elements for accurately identifying symptoms of PD, combining shimmer, jitter, pitch, harmonicity, pulses, and voicing by using K-Nearest Neighbor classifiers with different types of kernels (ie, radial basis functions, linear, polynomial, and multilayer perceptron).

Machine learning–based solutions have become the standard for most health care decision-making processes, for example, most previous works focus on differentiating diseased individuals from healthy controls. For example, the work by Tsanas and Arora [28] evaluated 2289 individuals (2023 healthy controls and 246 PD patients) and analyzed 15,227 voice tasks (9994 for healthy controls and 5233 for PD patients). Similarly, the work Bongioanni [29] compared speech-based automatic classification of patients with ALS and healthy people using sustained phoneme generation, diadochokinetic task, and spontaneous speech. They classified voice samples from 25 patients with ALS and 25 healthy participants using SVMs and deep neural networks. More recently, more focus has been given to multiclass scenarios, for example, the study by Benba et al [22] used a Convolutional Neural Network Long Short-term Memory to categorize ALS, PD, and healthy controls. The study analyzed speech data from 60 people, focusing on sentence reading, sound repetition, and sustained vowels.

Though there are studies that had investigate speech features pertaining to neurodegenerative disorders or acquired neurodegenerative disorders like mTBI, there are not many studies exploring speech feature variations between those populations which might co-occur and impact speech production differently.

The aim of this study is to investigate whether distinct speech-based biomarkers, derived from commonly used tasks like the PaTaKa and Sustained Vowel tests, can effectively differentiate between concussed individuals, neurodegenerative conditions (focused on PD), and healthy controls.

Methods

Data Collection

This study focused on 2 widely used speech tasks, the sequential motion rate task (PaTaKa test) and the Sustained Vowel test. The PaTaKa test evaluates speech-motor function by asking participants to take a deep breath and repeatedly articulate “Pa-Ta-Ka” as steadily as possible in 1 breath, providing insights into the rate and precision of sequential articulatory actions.

In the Sustained Vowel test, participants were instructed to sustain the vowel sound “ah” for as long as possible, offering valuable information about voice quality and potential vocal tremor. Both tasks were assigned to four participant groups, that are (1) individuals with concussions, (2) individuals with neurodegenerative conditions (specifically PD), (3) healthy controls age-matched to the concussed group, and (4) healthy controls age-matched to the neurodegenerative group.

Individuals diagnosed with a concussion were evaluated by physicians or athletic trainers using standardized neurocognitive assessment tools, such as ImPACT (Immediate Post-Concussion Assessment and Cognitive Testing) by ImPACT Applications, Inc, SCAT (Sport Concussion Assessment Tool), an open-access tool , and SAC (Standardized Assessment of Concussion) by researchers at the University of North Carolina’s Sports Medicine Research Laboratory, within 48 hours of the suspected injury. Individuals with neurodegenerative conditions (ie, PD) were diagnosed by licensed neurologists or family physicians. All participants with PD were in the early stages of disease progression (Hoehn and Yahr stage 1 - 2) and were assessed using tools such as the MDS-UPDRS (Movement Disorder Society - Unified Parkinson’s Disease Rating Scale) and Hoehn and Yahr Scale.

Healthy controls were divided into two groups: (1) young healthy individuals age-matched to the concussed group and (2) older healthy individuals age-matched to the neurodegenerative group. This separation ensures more accurate comparisons between the groups, minimizing the confounding effects of age-related speech differences.

Participants completed the speech tasks using a mobile app (smartphone or tablet) that provided both visual and auditory instructions. The app also recorded the audio samples digitally for subsequent analysis. Audio data were collected from a total of 235 and 239 participants for the PaTaKa and Sustained Vowel tests, respectively, as shown in Table 1.

Table . Description of collected samples.

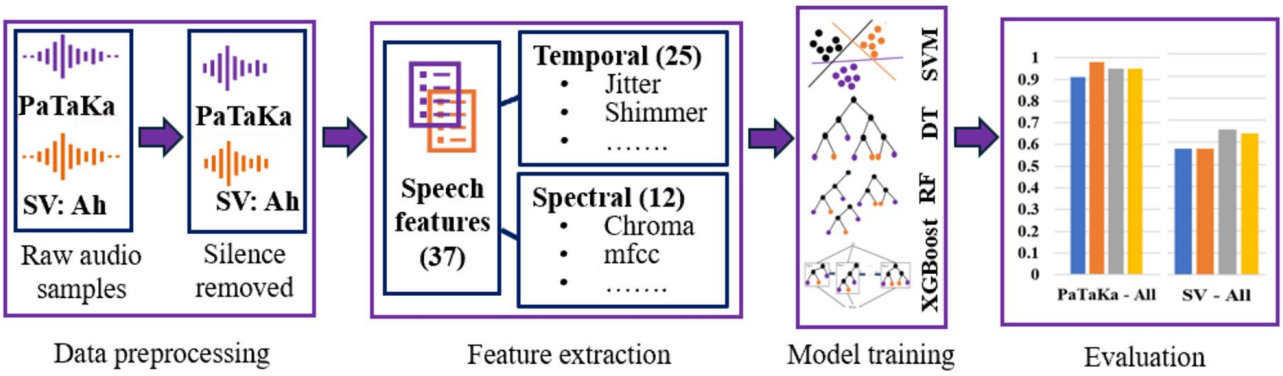
Test name and population		Samples, n	Sex		Age (years), mean (SD)
			Male, n	Female, n	
PaTaKa					
	Concussed	97	86	11	17 (3)
	Healthy control (young)	94	81	13	17 (3)
	Neurodegenerative (PD[a])	29	17	12	63.67 (4.95)
	Healthy control (older)	15	5	10	63.67 (4.95)
Sustained Vowel					
	Concussed	91	82	9	17 (3)
	Healthy control (young)	104	90	14	17 (3)
	Neurodegenerative (PD)	29	17	12	63.67 (4.95)
	Healthy control (older)	15	5	10	63.67 (4.95)

[a]PD: Parkinson disease.

The PaTaKa test dataset includes speech samples from 97 concussed participants, 29 participants with neurodegenerative conditions (ie, PD), 97 age-matched young healthy controls, and 15 age-matched older healthy controls. Similarly, the Sustained Vowel dataset consists of speech samples from 91 concussed participants, 29 participants with neurodegenerative conditions (ie, PD), 91 age-matched young healthy controls, and 15 age-matched older healthy controls.

In the remainder of this section, we describe the 4 key components of the proposed analysis methodology, shown in Figure 1, that are data preprocessing, feature extraction, model training, and evaluation.

Figure 1. Overall visualization of the 4 methodological steps: data preprocessing, feature extraction, model training, evaluation. DT: decision tree; mfcc: mel frequency cepstral coefficient; RF: random forest; SV: Sustained Vowel; SVM: support vector machine; XGBoost: Extreme Gradient Boosting.



Data Preprocessing

The voiced portions of speech signals typically carry the most critical information for analysis. Therefore, to enhance the quality and efficiency of feature extraction, it is essential to eliminate unnecessary components, such as silence intervals and extraneous noise, during the preprocessing phase. In this study, silence intervals were removed at 2 points in each speech recording using the free software developed by Muse group named "Audacity". Specifically, silence was cut from the beginning of the recording to the onset of vocalization and from the offset of vocalization to the end of the recording.

In addition, recordings that did not meet the study's requirements, such as those where participants failed to produce the expected utterances (eg, "PaTaKa" in 1 continuous breath or sustained vowel production without interruptions), were excluded from further analysis. This step ensured a high-quality dataset for feature extraction and classification, thereby improving the reliability of the results.

Data Augmentation

To address the challenges of imbalanced datasets and improve classification performance, data augmentation was applied to specific data subsets, particularly those with limited samples, such as the neurodegenerative (ie, PD) and age-matched healthy datasets. The augmentation process involved adding Gaussian noise to the raw audio signals. The noise factor was set to 0.005 to ensure that the original speech characteristics were preserved

while introducing subtle variations to increase sample diversity. For each audio file, a noise vector was generated using a Gaussian distribution, scaled by the specified noise factor, and added to the original signal. The augmented audio signals were then normalized to ensure they remained within the acceptable amplitude range for further processing.

This step increased the dataset size from 29 PD and 15 healthy samples to 58 PD and 30 healthy samples, resulting in a notable improvement in classification accuracy from under 30% (original data) to 60% - 70% (augmented data).

Feature Extraction

Feature extraction is the process of transforming raw audio data into numerical features while retaining the critical information embedded within the original signal. Among various methods for converting speech into numerical data, temporal and spectral features are widely used in speech-processing research [22,26,27,30]. In these studies, both types of features were extracted using Python's Librosa library [31].

Temporal features describe the changes in an audio signal over time, such as amplitude and pitch variation. This study extracted 25 temporal features, including 4 fundamental frequency measures (eg, mean and SD of F0), 5 jitter measures, 6 shimmer measures, and the HNR. These features provide insights into voice quality and stability, commonly associated with motor speech dysfunctions. The full list and descriptions of these temporal features are provided in Multimedia Appendix 1.

Spectral features analyze the frequency components of the speech signal and are commonly used in applications such as speech recognition and speaker identification. This study extracted 12 spectral features, including MFCC, spectral centroid, chroma features, and spectral flatness. These features capture frequency-domain characteristics that are sensitive to articulation and vocal tract configurations. Detailed descriptions of these spectral features are presented in Multimedia Appendix 1.

All 37 extracted features (25 temporal and 12 spectral) were included in the training and evaluation of machine learning models. By retaining the full feature set, we ensured that potentially valuable information was preserved, particularly given the small sample size. Data augmentation techniques, such as adding noise to the audio samples, were used to improve the robustness of the models and enhance performance, especially for the classification between neurodegenerative and healthy controls, where the original dataset resulted in poor classification performance.

Model Training

In recent years, the trend in digital health care has been to use machine learning models to classify input data (speech samples) into 2 or more classes based on extracted features. In this work, we employed several popular machine learning techniques, such as SVM, DT, RF, and XGBoost [18]. These models were chosen due to their interpretability, robustness, and ability to handle small datasets effectively, which is essential for clinical applications.

SVM, a supervised learning algorithm proposed by Boser et al [32], is grounded in statistical learning theory and is particularly effective for high-dimensional data [33]. It uses hyperplanes and margins to separate data into classes, with its performance being highly dependent on data scaling and the choice of kernel functions. DTs, on the other hand, divide feature space into regions by recursively splitting data and assigning classes to leaf nodes [34]. Despite their simplicity, DTs are prone to overfitting, especially on small datasets.

RFs mitigate this issue by employing an ensemble of DTs trained on bootstrapped datasets, with each tree built using a random subset of features [35]. The final class prediction is based on a majority vote across all trees, which reduces variance and enhances model robustness. Finally, XGBoost, a gradient boosting implementation, constructs DTs sequentially, optimizing performance by correcting errors from previous iterations [36]. It is known for its computational efficiency and scalability, making it a popular choice for structured datasets. For a given sample, the final prediction can be calculated by summing up the scores of overall leaves, which is illustrated in Multimedia Appendix 2.

Given the limited size of our dataset, we prioritized traditional machine-learning models over deep learning methods. While deep learning algorithms have demonstrated exceptional performance on large datasets, their effectiveness diminishes with smaller datasets due to overfitting and computational requirements. Traditional machine learning models, such as SVM and RF, offer superior interpretability, which is critical for clinical decision-making [28]. For instance, the study by Pishgar et al [37] found that on a small voice disorder dataset, SVM outperformed a deep neural network in terms of sensitivity and specificity.

In this study, all 37 extracted features (25 temporal and 12 spectral) were used without any feature selection or filtering. Data augmentation was applied to address the limited sample size, particularly for the neurodegenerative versus healthy dataset, where the augmented dataset improved model performance.

To train and evaluate the machine learning models, we applied a 75 - 25 stratified split of the dataset into training and test sets, ensuring that class distributions were preserved. Stratified 5-fold cross-validation was used to evaluate model performance more reliably, and Grid Search was used to fine-tune hyperparameters for all algorithms.

Evaluation

In this study, we assessed the performance of our classification models using multiple evaluation metrics, with a particular focus on the F_1-score due to its robustness in handling unbalanced datasets. The F_1-score is particularly well-suited for situations where there is an imbalance in the class distribution, as it provides a harmonic mean of precision and recall, balancing the trade-off between these 2 metrics. The F_1-score is defined as follows in Multimedia Appendix 2.

Both precision and recall are crucial in medical applications, where the consequences of false positives or false negatives can

be severe. The F_1-score offers a balanced view of a model's performance when neither precision nor recall can be prioritized over the other. A higher F_1-score (ranging from 0 to 1) indicates a better-performing model.

In addition to the F_1-score, we evaluated our models using precision, recall, and accuracy to provide a comprehensive view of model performance. These metrics helped compare the performance of models across different speech tasks (PaTaKa and Sustained Vowel) and combinations (eg, concussed vs healthy, concussed vs neurodegenerative, neurodegenerative vs healthy). The results section discusses these findings in detail, highlighting the implications of our model's performance for clinical applications.

Ethical Considerations

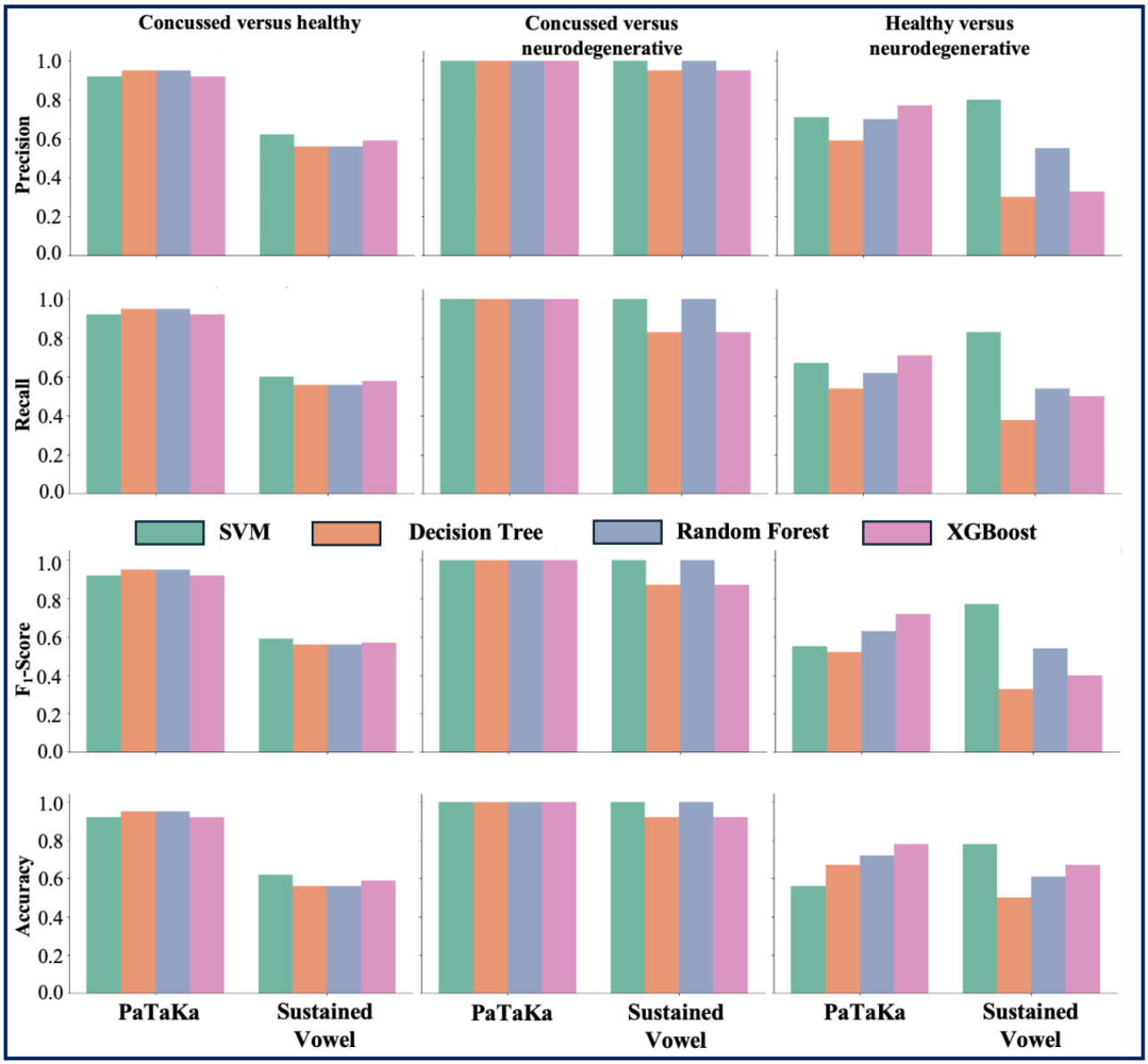
This research was conducted in compliance with ethical standards and approved by the Institutional Review Board at the University of Notre Dame. The approval numbers for this study are 18-01-4338 and 18-01-4340 for PD and concussion, respectively. All participants provided informed consent (Multimedia Appendices 3 and 4), and their confidentiality was ensured throughout the study.

Results

Overview

The performance of the models was evaluated using precision, recall, F_1-score, and accuracy across 3 participant combinations (ie, concussed vs healthy, concussed vs neurodegenerative, and neurodegenerative vs healthy) for 2 widely used speech tasks, PaTaKa and Sustained Vowel. The results provide insights into the discriminative ability of each test and highlight the comparative effectiveness of different classifiers in distinguishing between participant groups. While the PaTaKa task generally demonstrated robust performance across all combinations, the Sustained Vowel test showed varying levels of accuracy, particularly for certain groups and classifiers. The performance for each combination and test, along with discussions on their implications are illustrated in Figure 2.

Figure 2. Performance metrics by test type, model, and combination. SVM: support vector machine; XGBoost: Extreme Gradient Boosting.



Concussed Versus Healthy

PaTaKa Test

The models performed exceptionally well, achieving near-perfect precision, recall, F_1-score, and accuracy across all classifiers. DT and RF slightly outperformed SVM and XGBoost, consistently achieving 0.95. There are no sources in the current document across all metrics. These results highlight the PaTaKa test's robustness in distinguishing between concussed and healthy participants.

Sustained Vowel Test

Performance dropped significantly compared with the PaTaKa test. SVM and XGBoost achieved slightly higher metrics, with F_1-scores around 0.59 - 0.62. DT and RF had the lowest performance, with metrics around 0.56. The reduced performance might indicate that sustained vowels are less effective for distinguishing concussed participants from healthy individuals.

Concussed Versus Neurodegenerative

PaTaKa Test

All models performed perfectly, achieving precision, recall, F_1-score, and accuracy of 1.0. This demonstrates the effectiveness of the PaTaKa test for differentiating concussed participants from those with neurodegenerative conditions. Consistency across all classifiers reinforces the reliability of this task for this combination.

Sustained Vowel Test

Similar to the PaTaKa test, most models achieved perfect scores across all metrics. However, DT and XGBoost showed slightly reduced performance, with F_1-scores of 0.87 and accuracy of 0.92. Despite slight variability, the Sustained Vowel test remains a strong indicator for distinguishing these groups.

Neurodegenerative Versus Healthy

PaTaKa Test

Results varied significantly across classifiers. RF and XGBoost outperformed others, achieving F_1-scores of 0.63 and 0.72, respectively. DT and SVM performed poorly, with F_1-scores around 0.52 - 0.55. These results indicate that the PaTaKa test has moderate effectiveness for this group but requires careful classifier selection.

Sustained Vowel Test

Similar trends were observed. XGBoost achieved the highest F_1-score (0.40) and accuracy (0.67), while other models showed significantly lower performance. This underscores the challenge of distinguishing neurodegenerative participants from healthy controls using sustained vowel tasks.

Feature Set Analysis

Understanding the importance of individual features in classification tasks is crucial for interpreting the predictive power of machine learning models. In this study, we examined feature importance across all tests and combinations to identify the most influential speech features contributing to the classification of concussed, neurodegenerative, and healthy individuals. Feature importance was calculated for each model (SVM, DT, RF, and XGBoost) using a combination of metrics, such as Gini importance, SHAP values, or permutation importance, depending on the model.

To identify globally significant features, we analyzed the frequency of features ranked among the top 5 across all 24 tests. A summary of the top 10 most frequent features is presented in Table 2, while Table 3 provides combination-specific feature importance values. The most frequently identified features were temporal and spectral characteristics, which are known to capture both short-term and long-term speech patterns.

Table . Top 10 most frequent features across all tests.

Rank	Feature	Frequency	Mean importance
1	duration	15	0.29
2	zero_crossing_rate	13	0.33
3	spectral_flatness	12	0.30
4	mfcc[a]	11	0.25
5	spectral_bandwidth	7	0.42
6	spectral_centroid	6	0.07
7	spectral_contrast	5	0.07
8	chroma_stft	5	0.32
9	HNR[b]	4	0.06
10	f4_median	4	0.04

[a]mfcc: mel frequency cepstral coefficient.

[b]HNR: harmonics-to-noise ratio.

Table . Combination specific feature importance value.

Combination and test		Feature	Value
Concussed versus healthy			
	PaTaKa	Duration	1.9
	PaTaKa	Zero-crossing rate	0.47
	Sustained Vowel	Spectral flatness	0.12
Concussed versus neurodegenerative			
	PaTaKa	Spectral bandwidth	1.3
	Sustained Vowel	MFCC[a]	2.9
Neurodegenerative versus healthy			
	PaTaKa	HNR[b]	0.43
	Sustained Vowel	Spectral flatness	0.76

[a]MFCC: mel frequency cepstral coefficient.

[b]HNR: harmonics-to-noise ratio.

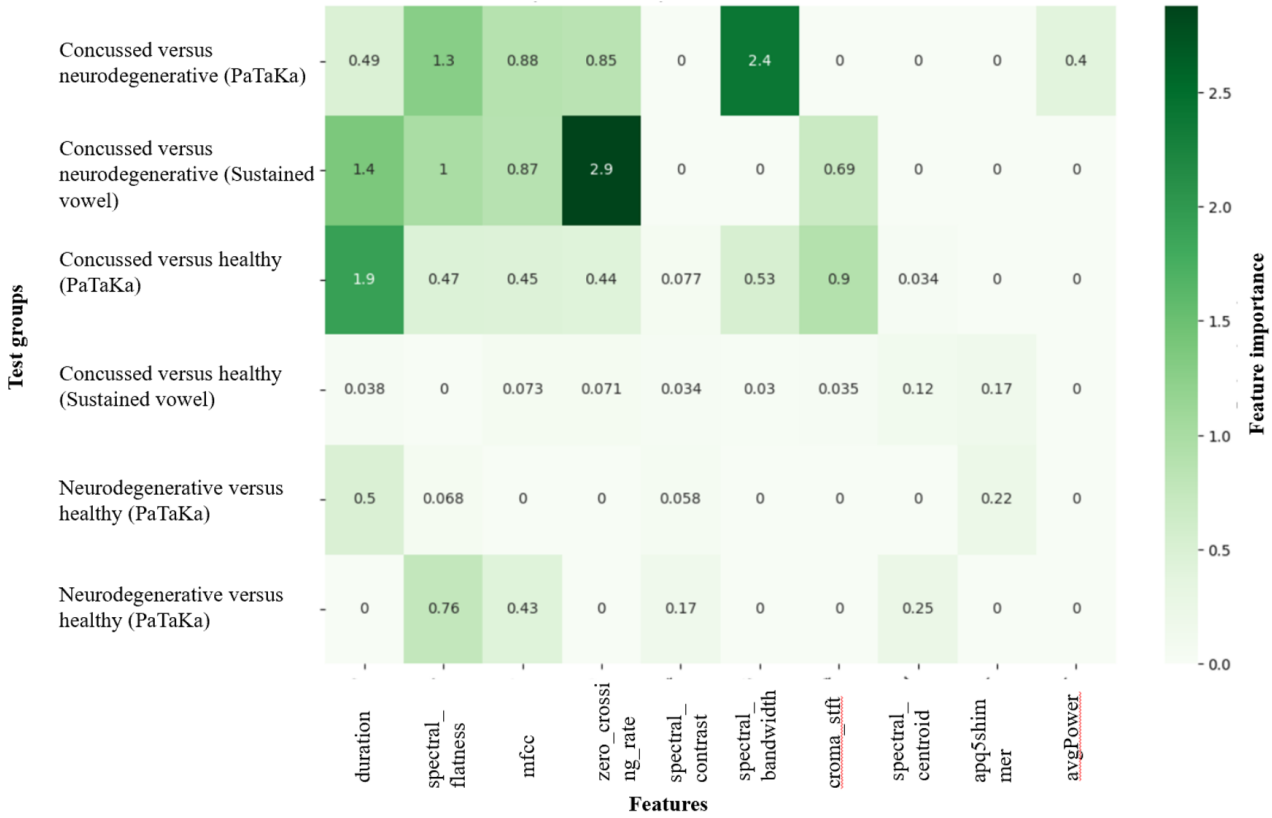
Among the top 10 features, duration, zero-crossing rate, and spectral_flatness were the most influential, appearing consistently across multiple tests and combinations. These features reflect critical aspects of speech production, including articulation rate, periodicity, and frequency smoothness. For instance:

- Duration: This feature provides insights into motor control and speech articulation by measuring the length of utterances.
- Zero-crossing rate: Indicative of voice signal periodicity, this feature is particularly significant in distinguishing voiced and unvoiced speech segments.
- Spectral_flatness: This feature quantifies the uniformity of the speech spectrum, distinguishing between harmonic and noise-like components.

Combination-specific patterns further highlight the variability in feature importance depending on the test (PaTaKa or Sustained Vowel) and the target classification task (concussed vs healthy, concussed vs neurodegenerative, and neurodegenerative vs healthy). For example, (1) in the concussed versus healthy classification, features like mfcc and spectral bandwidth were highly impactful, particularly in the PaTaKa test, (2) in the Concussed concussed versus neurodegenerative classification, spectral_centroid and chroma_stft played a significant role in distinguishing between the 2 groups, and (3) for the neurodegenerative versus healthy classification, features such as f4_median and HNR were key discriminators, particularly in the Sustained Vowel test.

The distribution of feature importance values across combinations and tests is visualized in Figure 3, while the detailed numerical values for each combination and test are available in Table 3. These findings emphasize the variability of feature contributions across different tasks and highlight the importance of task-specific feature analysis for robust classification.

Figure 3. Top 10 most frequent features across all tests. mfcc: mel frequency cepstral coefficients.



Discussion

Principal Findings

The findings of this study provide valuable insights into the use of speech-based features for differentiating between neurodegenerative conditions, particularly mTBI (concussions) and neurodegenerative diseases (eg, PD). By leveraging 2 commonly used speech tasks, the PaTaKa test and the Sustained Vowel test, and a variety of machine learning models, we achieved classification accuracies ranging from 60% to 90%, with RF and XGBoost models consistently outperforming others. In addition, we identified key speech features, such as duration, zero-crossing rate, and spectral flatness, as critical biomarkers for distinguishing between these conditions. These results underscore the potential of speech features as noninvasive biomarkers for neurodegenerative health assessment and highlight the complementary roles of the PaTaKa and Sustained Vowel tests in revealing task-specific and globally significant features.

Key Observations

First, task-specific performance. The PaTaKa test consistently outperformed the Sustained Vowel test across all combinations. This may be attributed to the sequential articulatory movements required in the PaTaKa test, which can better capture subtle motor and speech deficits. For example, in the concussed versus healthy classification, F_1-scores for PaTaKa exceeded 0.9 across all models, whereas the Sustained Vowel test achieved F_1-scores below 0.6 for the same classification. These findings highlight the importance of task selection in speech analysis and suggest that diadochokinetic tasks may provide richer diagnostic information.

Second, model-specific trends. Among the machine learning models, RF and XGBoost consistently performed well, demonstrating their ability to handle complex, nonlinear relationships in speech data. This aligns with previous research highlighting the robustness of ensemble learning methods in biomedical and speech signal processing tasks [38].

Third, the high interpretability of DTs also provides an advantage for clinical applications, particularly in scenarios where transparency is critical for adoption in health care settings.

Fourth, despite its slightly lower performance in some scenarios, DT models remain valuable due to their simplicity and ease of implementation.

Fifth, interestingly, SVMs displayed strong performance in balanced datasets, particularly in the concussed versus neurodegenerative classification, where precision and recall consistently reached 1.0 for the PaTaKa test. This finding is consistent with previous studies showing that SVMs are effective for high-dimensional data, especially when datasets are carefully preprocessed and balanced [39]. The performance of SVM in this classification task further underscores its utility in distinguishing nuanced differences between distinct neurodegenerative conditions using speech features.

Finally, feature importance. The analysis of feature importance revealed that a small subset of features consistently played a dominant role across tests and combinations. Temporal features such as duration and zero-crossing rate were particularly influential, likely reflecting disruptions in motor control and

speech rhythm caused by both concussions and neurodegenerative conditions. Spectral features, including spectral_flatness, mfcc, and spectral_bandwidth, were also critical, highlighting their utility in capturing frequency-domain variations associated with speech pathologies. These results align with previous research, which has emphasized the role of both temporal and spectral features in detecting neurodegenerative impairments.

Comparison With Previous Studies

Our findings corroborate and extend existing literature on speech-based biomarkers for neurodegenerative conditions. Previous research has demonstrated the utility of features such as MFCC and jitter for detecting PD [4], as well as features like zero-crossing rate and duration for identifying concussions [19]. However, this study uniquely emphasizes the differentiation between neurodegenerative diseases like PD and mild traumatic brain injury (eg, concussions), a task that remains relatively underexplored in existing literature.

Furthermore, the inclusion of both PaTaKa and Sustained Vowel tests enables a more comprehensive analysis of task-specific feature relevance. While previous studies have evaluated the diagnostic utility of individual speech tasks (eg, sustained phonation for ALS in studies by Allison et al [13] and Tsanas et al [27]), this work highlights how combining multiple tasks can reveal unique and complementary insights into speech biosignatures associated with diverse neurodegenerative conditions.

In addition to confirming the significance of widely used features such as spectral flatness and zero-crossing rate, our study identifies new combinations of features, including spectral contrast and chroma-based features, as being critical for distinguishing between these groups. These results align with recent advancements in the field, where ensemble learning models, such as RF and XGBoost, are increasingly used to capture the intricate, nonlinear relationships within speech data [23].

By addressing age-related variability and introducing data augmentation to mitigate the challenges of limited datasets, this study not only validates previously established findings but also sets the stage for future research aimed at improving the diagnostic accuracy of speech-based assessments across distinct but potentially overlapping neurodegenerative conditions.

Implications for Clinical Practice

The results of this study highlight several practical implications for clinical applications.

First, noninvasive diagnostics. The reliance on speech features, which can be collected using readily available devices like smartphones, opens up possibilities for remote and noninvasive diagnostics. This is particularly valuable in resource-constrained settings where access to advanced imaging or neurophysiological tests may be limited.

Second, early detection. The ability to detect subtle speech impairments associated with neurodegenerative conditions could enable earlier diagnosis, allowing for timely interventions.

Finally, task selection. The superior performance of the PaTaKa test suggests that it should be prioritized in future speech-based diagnostic protocols, particularly for distinguishing between concussions and neurodegenerative conditions.

Limitations

Despite the promising results, there are several limitations to this study.

First, small dataset—the dataset size, particularly for neurodegenerative diseases, was relatively small. This may limit the generalizability of the findings to larger, more diverse populations.

Second, demographic differences—the age gap between the concussed (younger) and neurodegenerative (older) populations poses a potential confounding factor. While age-matched healthy controls were included, the results could be influenced by inherent age-related differences in speech production.

Third, feature engineering and contextual factors—while the study identified important features, the reliance on manual feature extraction may overlook nuanced patterns. Advanced techniques, such as deep learning–based feature discovery, could reveal hidden characteristics in speech data. Future research should also account for comorbidities and age-related factors, as these can influence speech biosignatures and potentially confound results. Age-normalized datasets and statistical adjustments can further enhance the robustness of classification models.

Future Directions

This study demonstrates the potential of speech-based features to differentiate between concussed, neurodegenerative, and healthy individuals. While promising, the findings also highlight several areas for improvement and expansion, which we aim to address in future work.

First, dataset expansion and diversity. The current dataset includes limited samples from each group, particularly for neurodegenerative diseases. Future studies will expand the dataset to include larger and more diverse populations, ensuring broader generalizability of the results. In addition, we aim to achieve a more balanced age distribution across all participant groups, enabling more robust analyses and minimizing potential biases.

Second, age-related effects. While we mitigated some confounding effects of age by including 2 distinct healthy control groups (age-matched for concussed and neurodegenerative participants), future studies will incorporate more advanced strategies to address age-related variations in speech features. These include (1) explicitly including age as a covariate in statistical models to control its effects and quantify its influence on the results, (2) conducting age-matched subgroup analyses to validate that classification performance is not driven by age-related biases but by the underlying neurodegenerative conditions, and (3) expanding the dataset to improve the representation of younger and older age groups across all conditions.

Third, feature engineering and discovery. While this study focused on predefined temporal and spectral features, advanced deep learning models such as autoencoders or transformer-based models could uncover latent features that may better distinguish between neurodegenerative conditions. In addition, further exploration of task-specific feature relevance could reveal complementary insights into speech patterns for different health conditions.

Fourth, longitudinal data analysis. Future work should explore longitudinal data to track changes in speech biosignatures over time. This would help identify temporal patterns associated with disease progression and recovery, providing valuable insights for monitoring treatment efficacy and early diagnosis.

Fifth, integration with clinical practice. To enhance the clinical utility of this research, future efforts should focus on integrating speech-based diagnostic tools into real-world health care settings. This includes (1) developing user-friendly mobile apps or web applications for noninvasive speech analysis and (2) collaborating with clinicians to validate the models and evaluate their effectiveness in clinical decision making processes.

Finally, evaluation metrics and benchmarking. Expanding the evaluation metrics to include area under the receiver operating characteristic curve and precision-recall curves would provide a more comprehensive understanding of model performance. In addition, benchmarking against existing speech-based models or alternative diagnostic tools could further contextualize the findings and demonstrate the added value of the proposed methods.

By addressing these areas, future research can build upon the findings of this study to further advance the field of speech analysis in neurodegenerative health, improve diagnostic accuracy, and pave the way for noninvasive, scalable diagnostic tools.

Conclusion

This study demonstrates the potential of speech features, particularly those derived from the PaTaKa test, as effective biomarkers for distinguishing between concussed, neurodegenerative, and healthy individuals. By identifying task-specific and globally important features, the findings lay the groundwork for developing noninvasive, speech-based diagnostic tools that can be readily implemented in clinical practice. Further research addressing the study's limitations could pave the way for broader applications of speech analysis in neurodegenerative health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Feature description.
[DOCX File, 17 KB - neuro_v4i1e64624_app1.docx]

Multimedia Appendix 2
Equations.
[DOCX File, 15 KB - neuro_v4i1e64624_app2.docx]

Multimedia Appendix 3
Consent form for participants with neurodegenerative conditions.
[PDF File, 76 KB - neuro_v4i1e64624_app3.pdf]

Multimedia Appendix 4
Consent form for participants with concussions.
[PDF File, 76 KB - neuro_v4i1e64624_app4.pdf]

References

1. Templeton JM, Poellabauer C, Schneider S. Enhancement of neurocognitive assessments using smartphone capabilities: systematic review. JMIR Mhealth Uhealth 2020 Jun 24;8(6):e15517. [doi: 10.2196/15517] [Medline: 32442150]
2. Ramanarayanan V, Lammert AC, Rowe HP, Quatieri TF, Green JR. Speech as a biomarker: opportunities, interpretability, and challenges. Perspect ASHA SIGs 2022 Feb 11;7(1):276-283. [doi: 10.1044/2021_PERSP-21-00174]
3. Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. Evaluation of speech-based digital biomarkers: review and recommendations. Digit Biomark 2020;4(3):99-108. [doi: 10.1159/000510820] [Medline: 33251474]
4. Little MA, McSharry PE, Hunter EJ, Spielman J, Ramig LO. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans Biomed Eng 2009 Apr;56(4):1015-1022. [doi: 10.1109/TBME.2008.2005954] [Medline: 21399744]
5. Gan L, Cookson MR, Petrucelli L, La Spada AR. Converging pathways in neurodegeneration, from genetics to mechanisms. Nat Neurosci 2018 Oct;21(10):1300-1309. [doi: 10.1038/s41593-018-0237-7] [Medline: 30258237]

6. Kiaei M. New hopes and challenges for treatment of neurodegenerative disorders: great opportunities for young neuroscientists. Basic Clin Neurosci 2013;4(1):3-4. [Medline: 25337322]
7. Doherty C. ALS vs Parkinson's: what are the differences? Verywell Health. 2024. URL: https://www.verywellhealth.com/als-vs-parkinsons-6826333 [accessed 2025-02-06]
8. Youdim MB, Riederer P. Understanding Parkinson's disease. Sci Am 1997 Jan;276(1):52-59. [doi: 10.1038/scientificamerican0197-52] [Medline: 8972618]
9. Shively S, Scher AI, Perl DP, Diaz-Arrastia R. Dementia resulting from traumatic brain injury: what is the pathology? Arch Neurol 2012 Oct;69(10):1245-1251. [doi: 10.1001/archneurol.2011.3747] [Medline: 22776913]
10. Delaney JS, Abuzeyad F, Correa JA, Foxford R. Recognition and characteristics of concussions in the emergency department population. J Emerg Med 2005 Aug;29(2):189-197. [doi: 10.1016/j.jemermed.2005.01.020] [Medline: 16029831]
11. O'Sullivan SB, Schmitz TJ, Fulk G. Physical Rehabilitation: FA Davis; 2019.
12. Penttilä N, Korpijaakko-Huuhka AM, Kent RD. Disfluency clusters in speakers with and without neurogenic stuttering following traumatic brain injury. J Fluency Disord 2019 Mar;59:33-51. [doi: 10.1016/j.jfludis.2019.01.001] [Medline: 30641458]
13. Allison KM, Yunusova Y, Campbell TF, Wang J, Berry JD, Green JR. The diagnostic utility of patient-report and speech-language pathologists' ratings for detecting the early onset of bulbar symptoms due to ALS. Amyotroph Lateral Scler Frontotemporal Degener 2017 Aug;18(5-6):358-366. [doi: 10.1080/21678421.2017.1303515] [Medline: 28355886]
14. Aich S, Kim HC, younga K, Hui KL, Al-Absi AA, Sain M. A supervised machine learning approach using different feature selection techniques on voice datasets for prediction of Parkinson's disease. Presented at: 2019 21st International Conference on Advanced Communication Technology (ICACT); Feb 17-20, 2019; PyeongChang, South Korea. [doi: 10.23919/ICACT.2019.8701961]
15. Duffy JR. Motor Speech Disorders: Clues to Neurologic Diagnosis: Humana Press; 2000:35-53. [doi: 10.1007/978-1-59259]
16. Smith R, Chepisheva M, Cronin T, Seemungal BM. Chapter 16 - diagnostic approaches techniques in concussion/mild traumatic brain injury: where are we. In: Hoffer ME, Balaban CD, editors. Neurosensory Disorders in Mild Traumatic Brain Injury: Academic Press; 2019:247-277. [doi: 10.1016/B978-0-12-812344-7.00016-9]
17. Patel S, Grabowski C, Dayalu V, Cunningham M, Testa AJ. Fluency changes due to sports-related concussion. medRxiv. Preprint posted online on Sep 23, 2021. [doi: 10.1101/2021.09.19.21263791]
18. Daudet L, Yadav N, Perez M, Poellabauer C, Schneider S, Huebner A. Portable mTBI assessment using temporal and frequency analysis of speech. IEEE J Biomed Health Inform 2017 Mar;21(2):496-506. [doi: 10.1109/JBHI.2016.2633509] [Medline: 27913365]
19. Tsanas A. New insights into Parkinson's disease through statistical analysis of standard clinical scales quantifying symptom severity. Annu Int Conf IEEE Eng Med Biol Soc 2019 Jul;2019:3412-3415. [doi: 10.1109/EMBC.2019.8856559] [Medline: 31946612]
20. Julien JP. ALS: astrocytes move in as deadly neighbors. Nat Neurosci 2007 May;10(5):535-537. [doi: 10.1038/nn0507-535] [Medline: 17453052]
21. Magrinelli F, Picelli A, Tocco P, et al. Pathophysiology of motor dysfunction in Parkinson's disease as the rationale for drug treatment and rehabilitation. Parkinsons Dis 2016;2016:9832839. [doi: 10.1155/2016/9832839] [Medline: 27366343]
22. Benba A, Jilbab A, Hammouch A. Voice analysis for detecting patients with Parkinson's disease using the hybridization of the best acoustic features. Int J Electr Eng Inform 2016 Mar 30;8(1):108-116. [doi: 10.15676/ijeei.2016.8.1.8]
23. Mallela J, Illa A, Suhas BN, et al. Voice based classification of patients with amyotrophic lateral sclerosis, Parkinson's disease and healthy controls with CNN-LSTM using transfer learning. Presented at: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); May 4-8, 2020; Barcelona, Spain. [doi: 10.1109/ICASSP40776.2020.9053682]
24. Rueda A, Krishnan S. Clustering Parkinson's and age-related voice impairment signal features for unsupervised learning. Adv Data Sci Adapt Data Anal 2018 Apr;10(2):1840007. [doi: 10.1142/S2424922X18400077]
25. Tsanas A, Arora S. Biomedical speech signal insights from a large scale cohort across seven countries: the Parkinson's voice initiative study. 2019 Presented at: 11th International Workshop, Models and Analysis of Vocal Emissions for Biomedical Applications (MAVEBA); Dec 17-19, 2019; Firenze, Italy.
26. Vashkevich M, Azarov E, Petrovsky A, Rushkevich Y. Features extraction for the automatic detection of ALS disease from acoustic speech signals. Presented at: 2018 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA); Sep 19-21, 2018; Poznan, Poland. [doi: 10.23919/SPA.2018.8563414]
27. Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO. Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. IEEE Trans Biomed Eng 2012 May;59(5):1264-1271. [doi: 10.1109/TBME.2012.2183367] [Medline: 22249592]
28. Tsanas A, Arora S. Assessing Parkinson's Disease Speech Signal Generalization of Clustering Results across Three Countries: Findings in the Parkinson's Voice Initiative Study: SCITEPRESS - Science and Technology Publications; 2021:124-131. [doi: 10.5220/0010383001240131]
29. Bongioanni P. Paralysis in ALS patients: an overview on assessment and treatment of motor impairment. Int J Med Biol Front 2012;18(2/3):147.

30. Wall C, Powell D, Young F, et al. A deep learning-based approach to diagnose mild traumatic brain injury using audio classification. PLoS One 2022;17(9):e0274395. [doi: 10.1371/journal.pone.0274395] [Medline: 36170287]
31. McFee B, Raffel C, Liang D, et al. librosa: audio and music signal analysis in Python. Presented at: Python in Science Conference; Jul 6-12, 2015; Austin, Texas. [doi: 10.25080/Majora-7b98e3ed-003]
32. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. Presented at: COLT92; Jul 27-29, 1992; Pittsburgh, PA, United States. [doi: 10.1145/130385.130401]
33. Vapnik V. The Nature of Statistical Learning Theory: Springer Science Business Media; 2013.
34. Rokach L, Maimon O. Data Mining with Decision Trees: Theory and Applications, 2nd edition 2014, Vol. 81.
35. Breiman L. Random forests. Mach Learn 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
36. Abdurrahman G, Sintawati M. Implementation of XGBoost for classification of Parkinson's disease. J Phys: Conf Ser 2020 May 1;1538(1):012024. [doi: 10.1088/1742-6596/1538/1/012024]
37. Pishgar M, Karim F, Majumdar S, Darabi H. Pathological voice classification using mel-cepstrum vectors and support vector machine. arXiv. Preprint posted online on Dec 19, 2018. [doi: 10.48550/arXiv.1812.07729]
38. Difference between Random Forest and XGBoost. Geeks for Geeks. 2024. URL: https://www.geeksforgeeks.org/difference-between-random-forest-vs-xgboost/ [accessed 2025-02-06]
39. Jain M, Narayan S, Balaji P, P BK. Speech emotion recognition using support vector machine. arXiv. Preprint posted online on Feb 3, 2020. [doi: 10.48550/arXiv.2002.07590]

Abbreviations

ALS: amyotrophic lateral sclerosis
DT: decision tree
HNR: harmonics-to-noise ratio
ImPACT: Immediate Post-Concussion Assessment and Cognitive Testing
MDS-UPDRS: Movement Disorder Society - Unified Parkinson's Disease Rating Scale
MFCC: mel frequency cepstral coefficient
mTBI: mild traumatic brain injury
PD: Parkinson disease
RF: random forest
SAC: Standardized Assessment of Concussion
SCAT: Sport Concussion Assessment Tool
SVM: support vector machine
XGBoost: Extreme Gradient Boosting

Edited by P Kubben; submitted 22.07.24; peer-reviewed by H Rajaguru, M Gasmi, R Norel, S Mao; revised version received 26.11.24; accepted 08.01.25; published 12.02.25.

Please cite as:
Rubaiat R, Templeton JM, Schneider SL, De Silva U, Madanian S, Poellabauer C
Exploring Speech Biosignatures for Traumatic Brain Injury and Neurodegeneration: Pilot Machine Learning Study
JMIR Neurotech 2025;4:e64624
URL: https://neuro.jmir.org/2025/1/e64624
doi:10.2196/64624

Effectiveness of Artificial Intelligence–Based Platform in Administering Therapies for Children With Autism Spectrum Disorder: 12-Month Observational Study

Harini Atturu[1*], MBBS, MRCPsych, MSc; Somasekhar Naraganti[2*], MPhil; Bugatha Rajvir Rao[3*], MD

[1]Psychiatry, CARE Hospitals, Hyderabad, India

[2]Department of Psychology, Rainbow Children's Hospital, Hyderabad, India

[3]CognitiveBotics Technologies Private Limited, 6th Floor, TSIIC, MYHOME TWITZA, Raidurg, Telangana, Hyderabad, India

[*]all authors contributed equally

Corresponding Author:
Bugatha Rajvir Rao, MD
CognitiveBotics Technologies Private Limited, 6th Floor, TSIIC, MYHOME TWITZA, Raidurg, Telangana, Hyderabad, India

Abstract

Background: A 12-month longitudinal observational study was conducted on 43 children aged 2 - 18 years to evaluate the effectiveness of the CognitiveBotics artificial intelligence (AI)–based platform in conjunction with continuous therapy in improving therapeutic outcomes for children with autism spectrum disorder (ASD).

Objective: This study evaluates the CognitiveBotics software's effectiveness in supporting children with ASD through structured, technology-assisted learning. The primary objectives include assessing user engagement, tracking progress, and measuring efficacy using standardized clinical assessments.

Methods: A 12-month observational study was conducted on children diagnosed with ASD using the CognitiveBotics AI-based platform. Standardized assessments, include the Childhood Autism Rating Scale (CARS), Vineland Social Maturity Scale, Developmental Screening Test, and Receptive Expressive Emergent Language Test (REEL), were conducted at baseline (T1) and at the endpoint (T2). All participants meeting the inclusion criteria were provided access to the platform and received standard therapy. Participants who consistently adhered to platform use as per the study protocol were classified as the intervention group, while those who did not maintain continuous platform use were designated as the control group. Additionally, caregivers received structured training, including web-based parent teaching sessions, reinforcement strategy training, and home-based activity guidance.

Results: Participants in the intervention group demonstrated statistically significant improvements across multiple scales. CARS scores reduced from 33.41 (SD 1.89) at T1 to 28.34 (SD 3.80) at T2 (P<.001). Social age increased from 22.80 (SD 7.33) to 35.76 (SD 9.09; mean change: 12.96, 56.84% increase; P<.001). Social quotient increased from 53.26 (SD 11.84) to 64.75 (SD 16.12; mean change: 11.49, 21.57% increase; P<.001). Developmental age showed an improvement from 30.93 (SD 9.91) to 45.31 (SD 11.20; mean change: 14.38, 46.49% increase; P<.001), while developmental quotient increased from 70.94 (SD 10.95) to 81.33 (SD 16.85; mean change: 10.39, 14.65% increase; P<.001). REEL scores showed substantial improvements, with receptive language increasing by 56.22% (P<.001) and expressive language by 59.93% (P<.001). In the control group, while most psychometric parameters showed some improvements, they were not statistically significant. CARS scores decreased by 10.62% (P=.06), social age increased by 52.27% (P=.06), social quotient increased by 19.62% (P=.12), developmental age increased by 44.88% (P=.06), and developmental quotient increased by 11.23% (P=.19). REEL receptive and expressive language increased by 34.69% (P=.10) and 40.48% (P=.054), respectively.

Conclusions: Overall, the platform was an effective supplement in enhancing therapeutic outcomes for children with ASD. This platform holds promise as a valuable tool for augmenting ASD therapies across cognitive, social, and developmental domains. Future development should prioritize expanding the product's accessibility across various languages, ensuring cultural sensitivity and enhancing user-friendliness.

(JMIR Neurotech 2025;4:e70589) doi:10.2196/70589

KEYWORDS

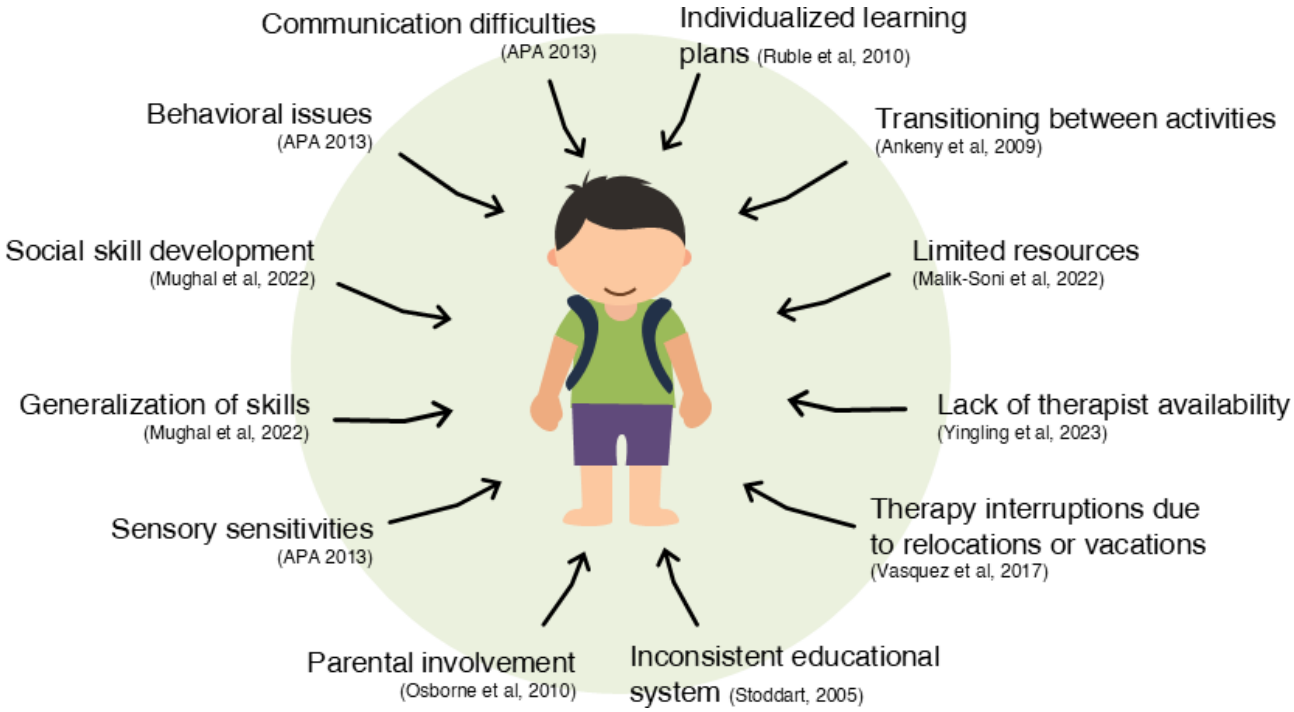
autism spectrum disorder; neurodevelopmental disorders; applied behavior analysis; software; artificial intelligence

Introduction

Autism, otherwise known as autism spectrum disorder (ASD), is a neurodevelopmental disorder with a wide continuum of associated cognitive and neurobehavioral deficits including, but not limited to, 3 core defining features: impairments in social interaction and impairments in verbal and nonverbal communication, combined with restricted and repetitive patterns of behaviors [1]. Such impairments can impede an individual's social level of interaction, learning aptitude, and employability, leading to poor long-term outcomes, difficulties in socializing, poor job performance, and difficulties in activities of daily living [2-5]. The estimated prevalence of ASD has increased from 1 in 10,000 in the 1960s to at least 1 in 36 today [6,7].

The cause for the rise of children diagnosed with ASD is unknown [8]. What is clear is that early and consistent intervention is crucial for positive long-term outcomes [9]. Currently, there are no medical treatments that can effectively cure individuals with ASD, with most interventions involving applied behavioral analysis (ABA), speech and language therapy, and sensory integration to address the core symptoms of ASD [10,11]. To provide adequate and quality therapy to children with autism, a team of trained professionals ranging from pediatricians, child psychiatrists; occupational, behavioral, and speech therapists; psychologists, specialist teachers, and dedicated caregivers are necessary [12]. Providing therapy to children with autism can be rewarding but challenging due to several factors. Figure 1 provides an insight into the challenges faced by the stakeholders in the care and support of children with autism [13-20].

Figure 1. Challenges in providing therapy to individuals with autism spectrum disorder [13-20].



As is, the solution to many of today's challenges may be the leveraging of cutting-edge technologies to enhance autism intervention; these technologies include the use of machine learning, deep learning in artificial intelligence (AI), animated gaming, and data analytics. Computer-assisted interventions (CAIs) are particularly appealing to underresourced schools due to the potential to provide cost-effective individualized instruction and allow teachers to offer concurrent group instruction. Several available CAIs have integrated evidence-based interventions and complement current therapies for individuals with ASD [21].

Research suggests that CAIs, when applied effectively, can enhance learning by fostering four key components of the learning process: (1) active engagement, (2) group participation, (3) regular interaction and feedback, and (4) integration with real-life settings [22]. Furthermore, the convenient access of CAIs among parents and therapists allows ease of access to these technologies right in the palm of their hands [23]. During the recent COVID-19 pandemic, there was significant disruption and reduction in conventional therapies. As a means to continue therapy, many therapists sought to use CAIs, leading to a jump in usage from 15% to 61% [24].

Through the use of intelligent systems–based AI technologies, therapists and parents alike can provide supplementary and consistent therapy to individuals with ASD and enhance outcomes [25-28]. In 2 recent articles, the prospect of integrating AI into standard practices for autism therapy has great potential to improve social and communication outcomes in individuals with autism [29,30].

The integration of video modeling in ABA allows the individual to observe a recorded video of a specific task, gradually enabling independent performance by clearly presenting the instructions and essential stimuli needed to complete the task. Several studies have demonstrated the effectiveness of this strategy across various complex social tasks, such as acquiring conversational skills, commenting, complimenting, and enhancing pragmatic

abilities, as well as initiating and maintaining social relationships [31].

Gaming systems provide a sensory stimulus, where numerous studies have found an attraction factor for participation through a framework or application that provides additional animation and images [32,33]. AI-driven games can improve cognitive skills, social interaction, and emotional regulation. Such games can be modified to the specific needs of individuals with autism, offering personalized learning objectives. Studies have suggested that integrating AI-based interventions into standard therapy can improve the behavioral patterns of children with autism [34,35]. Animation games use engaging animated characters and scenarios to teach essential skills, making learning enjoyable and less stressful for children with autism, thus improving their attention span and resulting in a greater retention of learned skills. Studies using animation-based interventions have observed significant improvements in language acquisition and social skills [36,37]. All these technology-driven solutions have been shown to significantly enhance outcomes and bridge the limitations of therapists and parents in managing challenging behaviors among children with ASD.

As a result, CognitiveBotics, an AI-powered assistive technology, was designed and developed. The platform allows children with autism and their parents and therapists to effortlessly access its program anytime, anywhere, since it only requires a gadget (eg, a laptop or tablet) and access to an internet connection. The development process involved a multidisciplinary approach, combining insights from clinical psychology, child development, and technology experts. The platform provides a "digital" VARK (visual, auditory, read/write, and kinaesthetic) opportunity range to help children acquire social, communication, emotional, and behavioral skills, while automatically recording progress for therapists [38]. For parents, the platform is an easy-to-use digital tool offering training sessions on strategies and techniques, ensuring continuity of therapy at home. For further information on the platform, visit [39].

During the COVID-19 pandemic, a survey was conducted among therapists working with children diagnosed with ASD. Due to the reduction in conventional therapies, the therapists observed a moderate to severe impact on individuals' learning (73%), while parents were impacted emotionally and psychologically (85%). Before the pandemic, only 22% of therapists expressed a willingness to use any digital technology in autism intervention, however, this number tripled to 65% due to the constraints imposed by the lockdown [40]. There was an urgent need for standardizing digital health technologies that can be parent-mediated [41]. An initial pilot study was conducted between November 2020 and April 2021 to assess the software's capabilities using a set of 19 different skills. Throughout the study, the software effectively collected and recorded data during the user interaction, demonstrating its effectiveness in real-time data collecting and analysis [40].

Subsequently, to further evaluate the effectiveness of the CognitiveBotics AI-based platform in augmenting therapies for individuals with ASD, an observational, longitudinal study with an adequate sample size was conducted to assess different domains—the social/emotional, language/communication, and cognitive development of individuals who used the platform for 12 months. The initial study revealed minor glitches, which were promptly addressed, and parents of the individuals expressed a willingness to continue using the app, highlighting its potential impact.

Methods

Overview

The observational, longitudinal study was designed to evaluate the effectiveness of the CognitiveBotics AI-based software over a 12-month period. By understanding the practical challenges and assessing the software's effectiveness, the study provides a foundation for the future development and design of a trial.

The primary objectives of the study are as follows:

1. User engagement: assess the ability of both children and parents to effectively use the software and follow web-based instructions.
2. Progress tracking: evaluate the software's capability to automatically log the child's daily progress and provide visual graphical feedback on the dashboard.
3. Efficacy measurement: using established clinical parameters to evaluate progress at T1 and T2 across multiple measures.

Scoring Systems

Qualified therapists conducted assessments at baseline and at a 1-year follow-up, using the following specific parameters to evaluate progress over time.

The Childhood Autism Rating Scale (CARS) score is a factor analysis–based scale used for assessing the presence and severity of symptoms of autism spectrum disorders [42]. Scores between 30 and 37 are considered as mild to moderate autism and scores between 38 and 60 are considered as a severe level of autism. According to Russell et al [43], CARS has an acceptable level of sensitivity and specificity in Indian populations.

The Vineland Social Maturity Scale (VSMS) scores were compared between groups, assessing changes in social age (SA) and social quotient (SQ). This scale has been used to measure the adaptive behaviors of children with or without ASD by measuring their developmental profile in 8 domains and scoring SA and SQ. Originally developed by Doll in 1935 [44], VSMS was adapted by Malin in 1956 [45] to better suit the Indian population, ensuring its cultural relevance and applicability. This adaptation was further modified by Bharatraj in 1992, incorporating additional changes [46].

The Developmental Screening Test (DST), which measures developmental age (DA) and developmental quotient (DQ), assesses the developmental progress of children across various domains, including motor skills, language, social behavior, and cognitive abilities. It helps in determining the DA and DQ of the participants, which reflects their level of functioning in comparison to typical developmental milestones [47]. Recognizing that many developmental assessments at that time were standardized on Western populations, in 1977, Bharatraj adapted the DST to be more sensitive to the developmental norms of Indian children [48].

The Receptive and Expressive Emergent Language (REEL) test is designed to identify infants and toddlers who have language impairments or who have other disabilities that affect language development. It has 2 core subtests, receptive language age (RLA) and expressive language age (ELA), which are based on caregiver reports and converted into age-equivalent scores. A study conducted with Hindi-speaking children found the REEL assessment to be valid, reliable, and effective in assessing language outcomes [49].

Recruitment

Recruitment for the study took place from January to April 2023 and the completion of the study was 12 months after the last participant was recruited. Parents whose children were diagnosed with ASD and attending Rainbow Hospital in India were identified by the clinical team. Recognizing that individuals with ASD may have a higher chronological age but a lower social or developmental age, participants were accepted if their social or developmental age was between 2 and 18 years. The parent information sheet regarding the study was provided to all identified parents. Parents who expressed interest in their child's participation were contacted by the principal investigator's team. Textbox 1 shows the inclusion, exclusion, and withdrawal criteria of the study.

Textbox 1. Inclusion, exclusion, and withdrawal criteria for participants.

Inclusion criteria

Children who met all the following inclusion criteria were enrolled in the study:

1. Children diagnosed with autism spectrum disorder using assessment scales such as the Childhood Autism Rating Scale.
2. Children aged between 2 and 18 years.
3. Children with associated comorbidities were included on the condition that the child can use the platform.
4. Children with the ability to understand and respond to instructions given in English.
5. Children with access to a device on which the software can be accessed using an internet connection.
6. Children with parents who consented for their child to use the software.

Exclusion criteria

1. Children with parents who were not willing to consent to the study.
2. Children without access to a tablet, computer, or internet connection.
3. Children unable to understand English.

Withdrawal criteria (removal of participants from the therapy or assessment)

Any participant was allowed to voluntarily discontinue participation in the study at any time after giving informed consent and before the completion of the last visit of the study. This would not affect the care provided by their clinical team. The reasons for participant withdrawal were recorded and included but were not limited to the following:

1. Participant was no longer willing to continue in the study.
2. Study termination by sponsor or independent ethics committee.
3. Investigator's discretion (for safety reasons).

When a participant withdrew from the study, the investigator clearly documented the reason in the medical records and completed the appropriate case report form describing the reason for discontinuation. In addition, every effort was made to complete the appropriate assessment.

During this stage, the study objectives and procedures were thoroughly explained, and any questions from the parents were addressed. Informed consent was obtained from those who agreed to participate, and documentation was appropriately maintained. At baseline, clinical assessments including the CARS, DST, VSMS, and REELs were administered. Parent training sessions, conducted either online or offline, were arranged to familiarize parents with the platform and its usage. Parents who had training were granted access to the software and instructed to ensure their children used the software for at least 20 minutes per session, with a minimum of 3 sessions per day over 12 months, followed by home-based activities to reinforce learning. At the beginning of the study, we requested parents to use the software in addition to the standard care they were providing to their children and for ethical reasons did not ask them to stop any other treatments or therapies.

Participants were scheduled for 3 visits during the active study period:

- Visit 1 (day 0, T1): baseline clinical assessments were conducted.
- Visit 2 (6 months): clinical parameters were reassessed.
- Visit 3 (12 months, T2): final clinical assessments were conducted.
- Data from the software tracking the child's progress were collected for statistical analysis at each stage.

Additionally, a follow-up phone call was made every 15 days between the physical visits to verify the child's regular usage of the software and address any concerns. This telephonic follow-up ensured adherence to the study protocol and provided support for parents throughout the trial.

Software-Delivered Program

Using tablets or a computer, the platform offers evidence-based therapeutic interventions through a high-quality, patented software program that addresses a broad spectrum of learning difficulties by teaching small, key behaviors incrementally. This aims to improve learning outcomes and developmental progress in individuals with ASD by providing a comprehensive digital platform that supports various learning styles and therapeutic needs. It is designed to personalize learning, adjust difficulty levels, and provide real-time feedback and support to both parents and children.

Upon initially using the platform, parents were registered in the system and requested to complete an auto-generated individualized learning plan (ILP) questionnaire generated by the software. This enabled the software to ascertain the child's current developmental state and learning needs. If there were any difficulties or queries from the parents regarding the questionnaire, a study coordinator was available to assist with the onboarding process. Parents were then requested to attend a webinar session, where an interactive orientation on the software and its features was given, and any queries were addressed. Additionally, parents received a user manual and a navigation video for reference. Participation in this webinar session was mandatory before an ILP was assigned to the child.

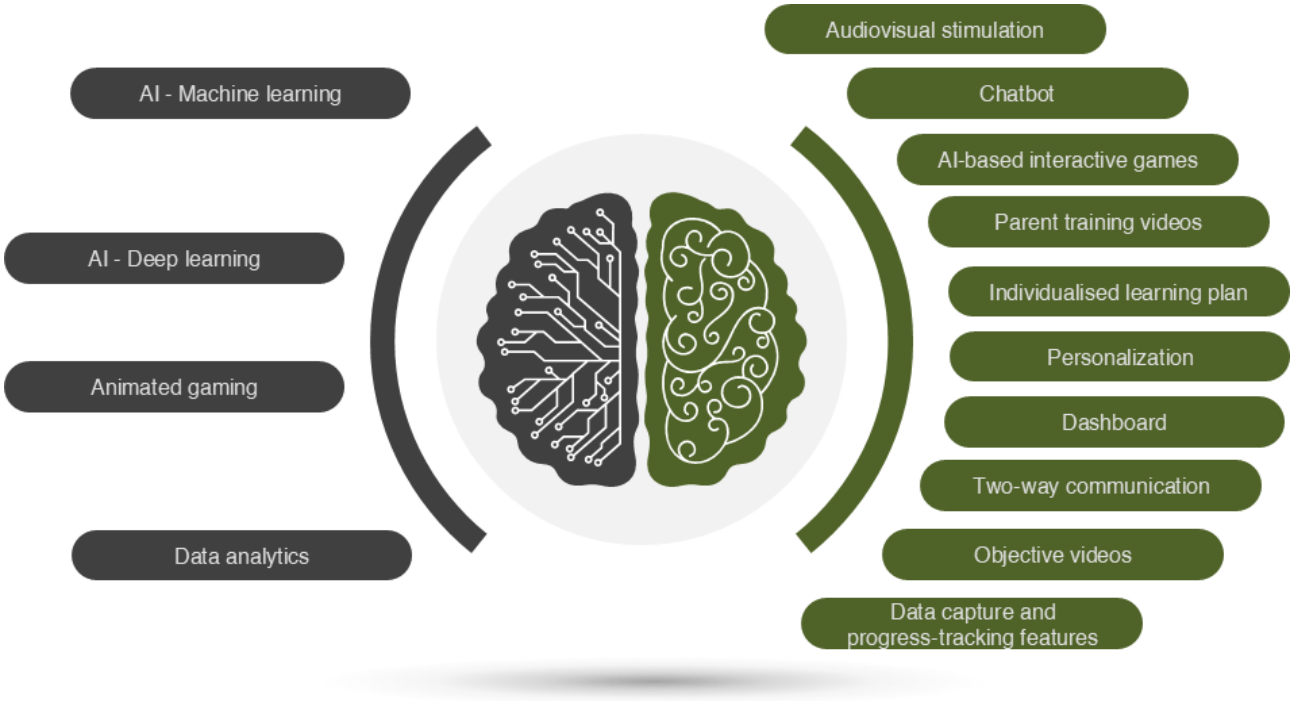
Based on the parental responses and child assessments, an ILP consisting of 3 target goals was generated by AI models focusing on 4 domains (social/emotional, language/communication, cognitive, and movement/physical development). Table 1 contains the lesson plan within the software and its advantages in providing adjunct therapy to children with ASD. The content is personalized and mapped to individual learning objectives, guided by therapist-defined developmental goals.

Table . Lesson plan structure and associated advantages of the platform.

Goal/skill domain	Task/learning objective	Methodology and advantages
Eye contact/attention	Looking at the object	Gamified, visually engaging content designed for children with neurodiverse profiles. Encourages sustained visual attention through interactive elements.
Eye contact/attention	Responding to name	Multimodal cues and visual prompts enhance auditory responsiveness and social awareness.
Imitation skills	Imitating arm, leg, or facial movements	Structured video models guide imitation in a low-anxiety, judgment-free digital space.
Cognitive skills	Number identification, shape recognition	Tasks scaffold foundational academic concepts in a playful, exploratory manner.
Communication/language	Labeling objects, requesting help	Activities promote expressive and receptive communication. Coviewing with caregivers enhances language modeling.

Before engaging in any lessons, parents were encouraged to watch the objective videos to improve the reasoning of mastering each goal. A practice session was available for skill reinforcement; however, the scores in these practice sessions were not recorded for progression to the next stage. Each daily practice session lasted 20 minutes, after which the software automatically concluded the learning session and redirected the child to the dashboard. If the caregiver determined that the child was prepared for an additional session, they had the option to initiate a new session., Overall, there are 227 activities or tasks organized under goals. Figure 2 presents the technologies and features of the CognitiveBotics platform.

Figure 2. The technologies and features in the software. AI: artificial intelligence.



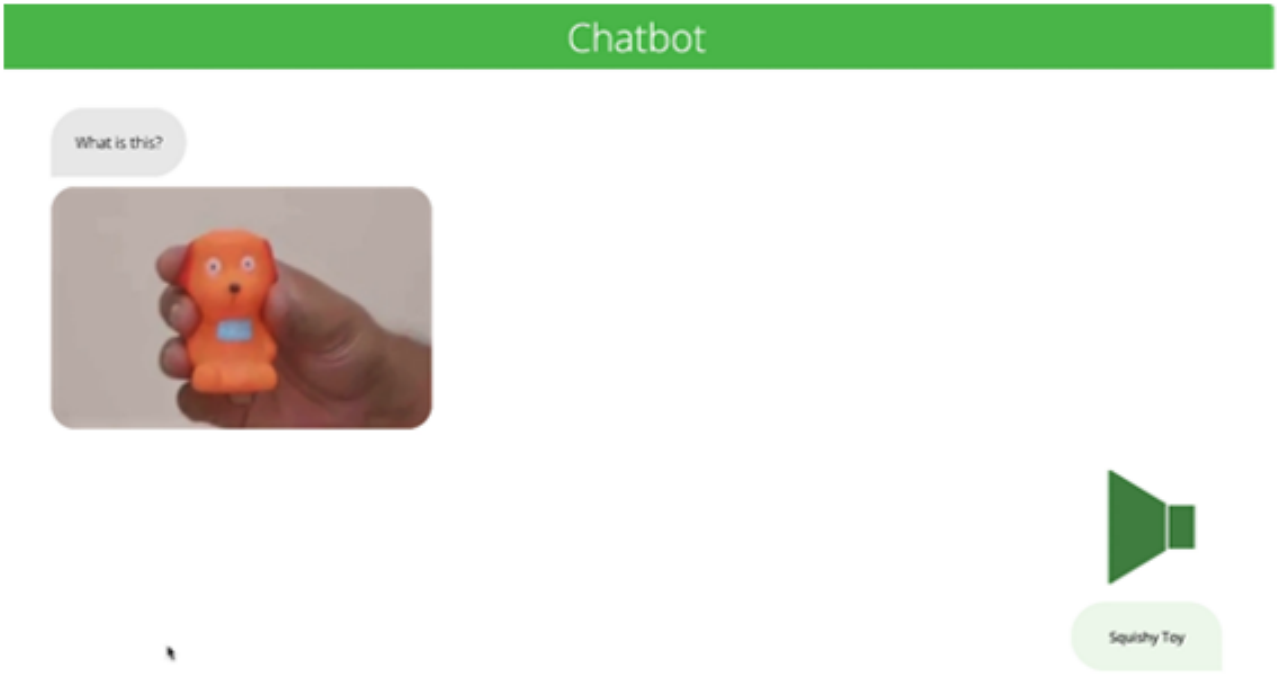
The session begins with the caregiver launching the daily schedule on the CognitiveBotics app. This schedule presents a sequence of personalized tasks aligned with the child's developmental goals. Each task is supported by engaging, gamified digital content designed specifically for children with ASD. Caregivers are encouraged to coview and participate in the learning process, fostering emotional bonding and reinforcing engagement through shared experience. Alternatively, under parent supervision, the child may explore the content independently, depending on their comfort and developmental level.

Once the child achieved 3 goals, a new ILP with a new set of 3 goals was created. To achieve each goal, the child is taught through 4 modalities:

- Audiovisual stimulation: Concepts are introduced through video modeling with interactive questions embedded within the content, increasing with complexity across four levels (level 0, 1, 2, and 3). Prompts are provided to guide the child's learning and are gradually reduced as the child becomes more proficient.
- Chatbot: This feature uses interactive questions to reinforce learning and promote generalization. The feature is particularly effective in fostering verbal engagement and enhancing the child's communication skills. An example of a chatbot goal is given in Figure 3.
- AI-based interactive games: Learning is facilitated through AI-driven interactive games that are tailored to each child's learning style, making the learning engaging and adaptive to individual needs.
- Home-based parent training videos: To support home-based activities, parents are provided with instructional videos that demonstrate how to apply the skills learned by their child in various settings, thus reinforcing learning outside the therapy center. The child's performance is assessed using 3 metrics captured by the software: first-time rights (accuracy of initial responses), correct questions (total number of correctly answered questions), and number of questions attempted (total engagement with the learning material). Once the lesson is mastered, the software automatically assigns the next set of goals.

If a child is not progressing toward their goals, the system proactively alerts the parents and therapists. Separately, parents are instructed to record a video of the lesson and submit it to the study coordinator or therapist team for review. In response, therapists will simplify the web-based goals to better suit the child's needs. Should the child continue to struggle, parents will receive a notification prompting them to resubmit the ILP checklist. Following this, the system will reassign 3 new goals, which will be carefully verified by therapists to ensure they align with the child's learning trajectory.

Figure 3. A screenshot of a lesson and an example transcript of a child-software interaction.



Software: *"Hi [Child's Name]! Let's name the object"*
(*[Object] appears on the screen...*)
Software: *[Child's name]!* What is this?
(Child should respond "Squishy toy"
(*If the child does not respond, the caregiver points at [object] on the screen*)
Software: *What is this?*
(*If the child still does not respond...*)
Software: *"Squishy Toy". [Child's Name] say "Squishy toy."*
(*If the child responds...*)
Software: *"Great job, [Child's Name]! Let's continue!"*
(*If the child did not respond...*)
Software: *"Great job, [Child's Name]!*
(Software moves on to the next goal and repeats this lesson the next day until the child masters the goal.

Other Core Features of the Platform

Other core features of the CognitiveBotics platform include the following:

- ILP progression: The software adjusts the level of difficulty of the ILP based on the child's progress, providing necessary assistance and notifications to parents and therapists.
- Personalization: Personalization is a unique feature, where all learning goals are delivered in a personalized and customized manner, tailored to the specific needs of each child. During interactive sessions, the system personalizes by using the child's name while asking the interactive questions, drawing the child's attention.
- Dashboard: A daily progress graph is displayed on the child's dashboard, which is accessible to both parents and therapists, offering real-time insights into the child's development.
- Two-way communication: The software includes a fun activity that detects and encourages body part interactions, in addition to occupational therapy tasks, promoting overall development from a young age.
- Objective videos: Parents are empowered through videos that outline the objectives of each task, enabling them to actively participate in and support their child's learning.
- Data capture and progress tracking features: Aim to automate monitoring and capture the child's progress based on key learning principles—attention, retention, and

generalization, such as “eye gaze detection.” These data are presented in a user-friendly format on a dashboard, facilitating easy comprehension for both parents and therapists.

Fidelity of Implementation Data

The fidelity of implementation was assessed via a multitiered approach to ensure attendance to the session lessons. The software has an automated session notification and progress tracker to prompt parents to complete assigned goals within the learning plan. To progress to the next learning level, mandatory successive mastering of goals is required. This ensures that all lesson components were completed as intended. Additionally, therapist-led monitoring and follow-up calls were conducted to monitor progress, reinforce engagement with the intervention, and address any caregiver-reported concerns to ensure fidelity.

Caregivers underwent a structured training program on reinforcement strategies aimed at ensuring consistency in their interactions with the child beyond software-guided sessions. This training equipped caregivers with evidence-based behavioral techniques that align with the principles of ABA and developmental learning models, such as immediate reinforcement or reward systems. Furthermore, to encourage parental involvement, caregivers were provided zero-fee in-person therapy sessions at the center, on the condition their child is actively engaged with the platform.

Lastly, software usage was collected at the back end, tracking metrics such as log-in frequency, time spent on lessons, and completion rates. This allowed the software programmer to evaluate the platform utilization and adherence. Any deviations from the lesson plans were brought to the attention of the therapist. Together, these mechanisms ensured consistent implementation and provided opportunities for timely intervention when necessary.

Statistical Analysis

After completion of the study, the data were analyzed to compare the effectiveness of the CognitiveBotics platform between the intervention and control groups. For each group and clinical assessment parameter, the mean scores and standard deviations were calculated at 2 stages: the start of the study (T1) and the end of the study (T2). The mean change and percentage mean change from T1 to T2 were also computed. To determine the statistical differences, the *P* values were calculated using the Mann-Whitney *U* test, with a *P* value of <.05 being considered as statistically significant.

Ethical Considerations

This study was conducted in accordance with the study protocol, the New Drugs and Clinical Trials Rules 2019 issued by the Government of India, the ethical principles that have their origin in the Declaration of Helsinki (64th World Medical Association General Assembly, Fortaleza, Brazil, October 2013), the International Council for Harmonisation Good Clinical Practice, and all applicable local regulatory requirements. The investigators agreed to conduct the study according to the principles of the International Council for Harmonisation Good Clinical Practice, as well as in accordance with the ethical principles that have their origin in the Declaration of Helsinki, the protocol, and all national, state, and local laws or regulations. The medical care given to and medical decisions made on behalf of study participants were always the responsibility of a principal (site) investigator. Each individual involved in conducting the study was qualified by education, training, and experience to perform his or her respective task(s).

Informed consent was obtained from the parents or legal guardians of all participants. The study details were thoroughly explained, including the study’s purpose and procedures and the voluntary nature of participation. Parents were informed that they and their children were free to withdraw from the study at any time, with no impact on their routine activities or any other services received. As this study included human participants, the collection of data from medical records, as well as software usage, it adheres to all institutional ethical guidelines. Ethical approval for this observational study was obtained from the Institutional Ethics Committee of the Rainbow Children’s Medicare (registration number EC/RENEW/INST/2021/10510).

Before any collection of data, the study protocol, participant information sheets, and informed consent forms were reviewed and approved. The data were maintained throughout the study, with all reports and communications relating to participants being kept confidential. Names and other identifiable details were removed, and all records were coded using unique identification acronyms. No images or video recordings of participants are included in the manuscript. No monetary compensation was provided to the participants or their families. However, participants in both the intervention and control groups received free access to the software platform, as well compensation for travel expenses when coming to the center for assessments.

Results

Participant Selection and Characteristics

The results of this study examine the impact and utility of the CognitiveBotics platform for children with ASD over a 12-month observational period. Key outcomes focus on quantitative measures of behavioral, developmental, and language-based parameters. An intervention versus control analysis was performed, organized by baseline (T1) and end-of-study (T2), to ascertain the software’s impact across multiple functional and developmental domains, namely CARS, VSMS, DST, and REEL scores. This approach provided structured insights into the software’s influence on each parameter and allowed for comparative analysis of outcomes over time.

Figure 4 illustrates the study’s recruitment and retention flow. Of an initial total of 88 enrolled participants, 43 completed the study, while 35 continued to use the software for the entire 1-year duration, and 5 did not use the software but participated in the 1-year follow-up assessments, and were categorized as the control group. A further 3 participants were labeled as outliers and were excluded from further analysis. Table 2 shows

the key baseline demographic characteristics of the 40 participants who completed the study.

Figure 4. Flowchart of participants in the study.

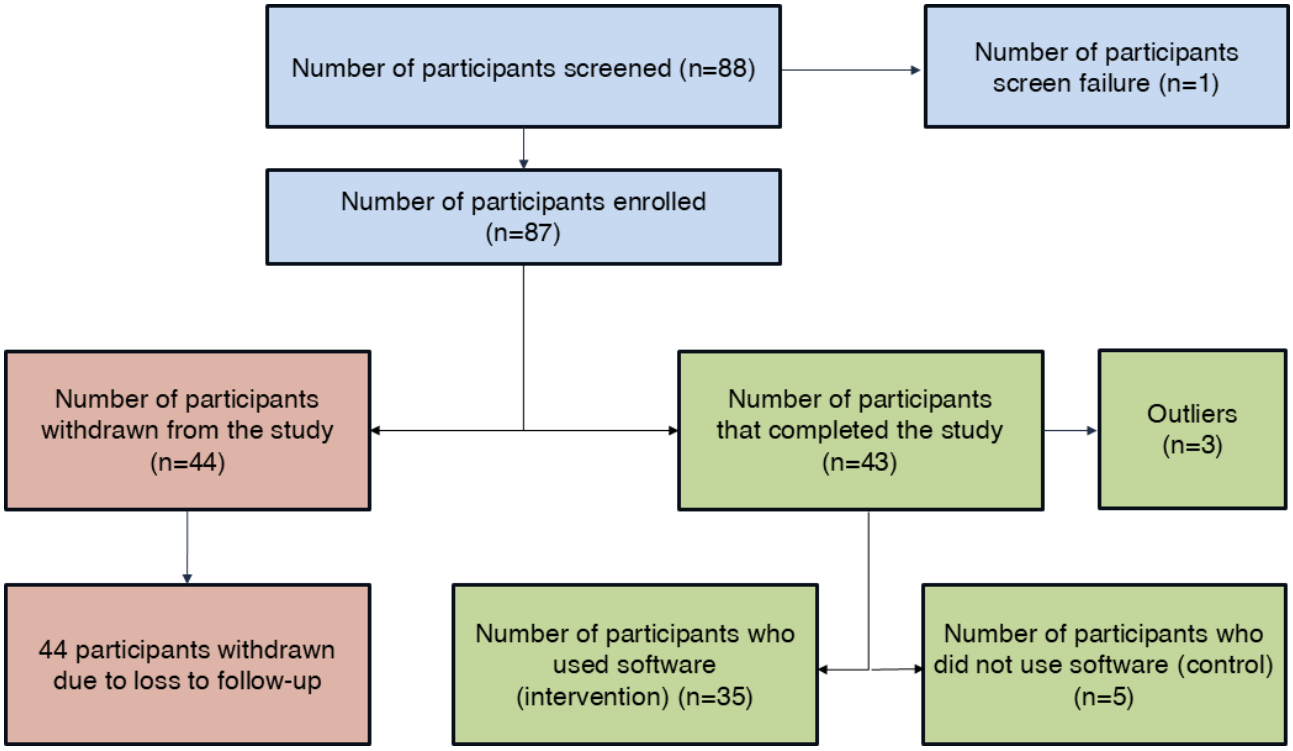


Table . Comparison of baseline demographics of participants in the intervention and control groups.

Parameter and statistics		Intervention (n=35)	Control (n=5)	Overall (n=40)
Age (years)				
	Mean (SD)	43.71 (SD 15.48)	44.60 (SD 14.98)	43.83 (SD 15.23)
	Median	39.00	39.00	39.00
	Quantile	31.50; 52.00	33.00; 54.00	31.75; 54.50
	Range	25.00 - 87.00	31.00 - 66.00	25.00 - 87.00
Gender, n (%)				
	Male	33 (94)	3 (60)	36 (90)
	Female	2 (6)	2 (40)	4 (10)

The participants in the intervention group were stratified into 3 developmental groups based on chronological age:

- Toddler group (n=12): children aged 2 - 3 years
- Preschool group (n=15): children aged 4 - 6 years
- School-aged group (n=8): children aged 7 - 8 years

The purpose was to assess the impact of the intervention across different developmental ages, considering variations in cognitive, language, and social skills.

Based on the study location, the majority of participants were of South Indian descent and from families with a higher educational background. All participants showed delays across multiple developmental domains, necessitating structured therapeutic intervention. Their academic skill levels in reading, writing, and mathematics were rudimentary, with significant challenges observed in social/emotional, language/communication, cognitive, and movement/physical development.

Intervention and Control Group–Based Analysis Using Different Parameters

The study evaluated outcome measures in the intervention and control groups across T1 (baseline) and T2 (12 months), assessing CARS, SA, SQ, DA, DQ, and REEL scores.

Table 3 shows the outcome measures of 35 participants in the intervention group, which were compared across T1 and T2. For the CARS score, there was a significant decrease from 33.41 (SD 1.89) at T1 to 28.34 (SD 3.80) at T2, showing a mean change of 5.07 and a percentage change of 15.18% (P<.001).

Table . Comparison of outcome measures in the intervention group only at baseline (T1) and end of study (T2).

Parameters	Intervention group (n=35)				
	T1[a], mean (SD)	T2[b], mean (SD)	Mean change	Mean change, %	*P* value[c]
CARS[d]	33.41 (1.89)	28.34 (3.80)	5.07	15.18	<.001
SA[e]	22.80 (7.33)	35.76 (9.09)	12.96	56.84	<.001
SQ[f]	53.26 (11.84)	64.75 (16.12)	11.49	21.57	<.001
DA[g]	30.93 (9.91)	45.31 (11.20)	14.38	46.49	<.001
DQ[h]	70.94 (10.95)	81.33 (16.85)	10.39	14.65	<.001
RLA[i]	22.09 (8.94)	34.51 (14.93)	12.42	56.22	<.001
ELA[j]	18.69 (8.52)	29.89 (15.60)	11.20	59.93	<.001

[a]T1: start of the study.

[b]T2: end of the study.

[c]*P* value was calculated using the Mann-Whitney *U* test.

[d]CARS: Childhood Autism Rating Scale.

[e]SA: social age.

[f]SQ: social quotient.

[g]DA: developmental age.

[h]DQ: developmental quotient.

[i]RLA: receptive language age.

[j]ELA: expressive language age.

In the SA score, there was a significant improvement from 22.80 (SD 7.33) at T1 to 35.76 (SD 9.09) at T2, with a mean change of 12.96 and a percentage change of 56.84% (*P*<.001).

In the SQ score, there was an improvement from 53.26 (SD 11.84) at T1 to 64.75 (SD 16.12) at T2, with a mean change of 11.49 and a percentage change of 21.57% (*P*<.001).

In the DA score, there was an improvement from 30.93 (SD 9.91) at T1 to 45.31 (SD 11.20) at T2, showing a mean change of 14.38 and a percentage change of 46.49% (*P*<.001).

In the DQ score, there was an improvement from 70.94 (SD 10.95) at T1 to 81.33 (SD 16.85) at T2, showing a mean change of 10.39 and a percentage change of 14.65% (*P*<.001).

In the REEL score, the RLA showed a substantial increase from 22.09 (SD 8.94) at T1 to 34.51 (SD 14.93) at T2, with a mean change of 12.42 and a percentage change of 56.22% (*P*<.001). Similarly, the ELA exhibited a significant increase from 18.69 (SD 8.52) to 29.89 (SD 15.60), showing a mean change of 11.20 and a percentage change of 59.93% (*P*<.001).

Table 4 shows the outcome measures of 5 participants in the control group, which were compared across T1 and T2. For the CARS score, there was a significant decrease from 33.90 (SD 1.24) at T1 to 30.30 (SD 3.68) at T2, showing a mean change of 3.6 and a percentage change of 10.62% (*P*=.06).

Table . Comparison of outcome measures in the control group only at baseline (T1) and end of study (T2).

Parameters	Control group (n=5)				
	T1[a], mean (SD)	T2[b], mean (SD)	Mean change	Mean change, %	*P* value[c]
CARS[d]	33.90 (1.24)	30.30 (3.68)	3.6	10.62	.06
SA[e]	21.41 (5.44)	32.60 (8.24)	11.19	52.27	.06
SQ[f]	49.13 (5.45)	58.77 (14.73)	9.64	19.62	.12
DA[g]	28.30 (6.69)	41.00 (7.04)	12.7	44.88	.06
DQ[h]	65.60 (11.68)	72.97 (7.22)	7.37	11.23	.19
RLA[i]	19.60 (7.13)	26.40 (9.53)	6.80	34.69	.10
ELA[j]	16.80 (4.60)	23.60 (6.23)	6.80	40.48	.054

[a]T1: start of the study.

[b]T2: end of the study.

[c]*P* value is calculated using Mann-Whitney *U* test.

[d]CARS: Childhood Autism Rating Scale.

[e]SA: social age.

[f]SQ: social quotient.

[g]DA: developmental age.

[h]DQ: developmental quotient.

[i]RLA: receptive language age.

[j]ELA: expressive language age.

In the SA score, there was a significant improvement from 21.41 (SD 5.44) at T1 to 32.60 (SD 8.24) at T2, with a mean change of 11.19 and a percentage change of 52.27% (*P*=.06).

In the SQ score, there was an improvement from 49.13 (SD 5.45) at T1 to 58.77 (SD 14.73) at T2, with a mean change of 9.64 and a percentage change of 19.62% (*P*=.12).

Similarly, in the DA score, there was an improvement from 28.30 (SD 6.69) at T1 to 41.00 (SD 7.04) at T2, showing a mean change of 12.7 and a percentage change of 44.88% (*P*=.06).

In the DQ score, there was an improvement from 65.60 (SD 11.68) at T1 to 72.97 (SD 7.22) at T2, showing a mean change of 7.37 and a percentage change of 11.23% (*P*=.19).

In the REEL score, the RLA showed a substantial increase from 19.60 (SD 7.13) at T1 to 26.40 (SD 9.53) at T2, with a mean change of 6.80 and a percentage change of 34.69% (*P*=.10). The ELA exhibited an increase from 16.80 (SD 4.60) to 23.60 (SD 6.23), showing a mean change of 6.80 and a percentage change of 40.48% (*P*=.054).

Overall, the intervention group presented substantial improvements across all outcome measures, particularly in CARS, SA, and language scores (RLA and ELA), with the majority of these changes reaching statistical significance. This indicates that the platform may enhance social, cognitive, and language outcomes in the intervention group. In contrast, the control group of 5 participants showed positive changes but with less significance and the changes were statistically weaker across measures.

Discussion

Principal Findings

This study demonstrated that CognitiveBotics, an AI-powered assistive technology, has made significant gains in developmental and social parameters over the course of 12 months in children diagnosed with autism. Both parents and therapists have reported minimal negative behavioral changes while using the platform, including screen addiction and sleep disturbances. In intervention versus control analysis, there were significant improvements in the intervention group, particularly in those with higher baseline levels of functioning, underlining the efficacy of the software in reducing autism severity and enhancing developmental skills in children with ASD. Accompanied by highly significant *P* values, the intervention group showed an improvement in symptoms, as well as marked enhancements in social skills, developmental age, and language abilities.

The CognitiveBotics software, like many other available ABA-assistive technologies, was observed to have various benefits and advantages specifically for individuals with ASD [50]. Supported in laptops and tablets, the platform is commonly available, affordable, and socially acceptable, making it an ideal tool for parent-mediated interventions [51,52]. Using the platform, parents played a crucial role in supporting their children's learning, observing better improvements compared to the control group using only traditional therapy. The software helps enhance attention span and motivation during learning activities, offering engaging, interactive experiences that increase children's participation [53,54].

Within a learning environment, the software increases interaction and participation and improves the learning process [55]. Additionally, the software provides real-time feedback on key skills and is customizable to focus on individual needs, similar to the benefits seen in the Picture Exchange Communication System and other visual aids, texts, and sounds [56,57]. The portability of the devices can allow parents to provide learning at times when the child is most receptive, despite the unavailability of therapists. Furthermore, parent-implemented technologies can be the most readily and affordably deployed, and such assistive technology enables parents to offer the most opportunities for social contact [58]. The software incorporates interactive games that improve social-emotional functioning and behavior. The interactive feature allowed the participants to recognize emotions, use deconfliction strategies, collaborate with others, and address issues like greeting known people like teachers or neighbors. In a recent study, parents who used social skills programs incorporating features similar to those in the CognitiveBotics platform found significant improvements in social skills and reductions in problematic behaviors, in contrast to those in the control group [59].

There may be certain shortfalls with the use of ABA assistive technologies, but as with any problem, there are solutions that can overcome such shortfalls. The first area of concern is increased screen time, possibly leading to restricted or repetitive behaviors, lack of socializing, and concerns over metabolic and sleep disturbances [60,61]. In such circumstances, CognitiveBotics has incorporated a preset screen time feature of 20 minutes, after which the session concludes and takes the user to the dashboard. It is also advisable to provide minimal access in a group setting to reduce potential isolation [62]. Devices may also be misused to view passive content, in which case supervised coviewing with parents is advised [63]. Furthermore, the choice of content has to be predetermined, whereby highly interactive and engaging media is most beneficial to the child as it promotes engagement, motivation, and learning outcomes [64]. Another issue is the potential for tantrums if the device is removed. As is the case in other situations, when access to preferred items is interrupted, parents and therapists should be trained to control such behaviors.

In recent years, there have been numerous studies on the proposed use of tablets or computers in autism interventions. A meta-analysis conducted by Sandbank et al [65], reviewed 252 separate trials examining the efficacy of technology in autism interventions. The findings suggest an overall improvement in social communication skills and reductions in difficult behaviors, particularly when used by parents. This aligns with the intentions behind the CognitiveBotics platform, which aims to support individuals with autism and their families. Furthermore, a low incidence of adverse events reported when using such interventions supports adoption of the software in both home and clinical settings.

Novack et al [66] conducted a study to assess the effectiveness of mobile apps on the principles of ABA, particularly in assessing the impact on the receptive language skills of individuals. Randomized into an immediate-treatment or a delayed-treatment control group, the results indicated significant improvements in receptive language skills in the former group. However, the study had limitations, particularly with the absence of psychometric parameters to assess outcomes. Although improvements in receptive language skills were observed, the study is incomplete. Our 12-month study demonstrated how CognitiveBotics leverages AI to improve receptive language skills, offering prolonged benefits using personalized ABA-based interventions and addressing limitations in traditional psychometric assessments. Another study aimed at addressing social engagement by using a proposed 3D complex facial expression recognition system to recognize facial emotions; it found that, in 3 weeks, users had a marked improvement in identifying facial cues compared with the control group, with surprise and shy expressions being the easiest to identify [67]. Similarly, CognitiveBotics contains activities that enable children to better recognize and respond to social and emotional cues, significantly boosting their social communication skills within a short intervention period.

A study conducted in Saudi Arabia assessed the effectiveness of AI-driven apps in a traditional education setup. Apps such as “My School” and “Alfaz” were chosen for their adaptive and interactive content that aligned with the academic curriculum. Participants who received 60-minute sessions twice weekly for 5 weeks showed significant improvements in reading and math skills compared to those in the control group [68]. Similarly, our software incorporates real-time feedback, task adaptation, and data-driven insights to ensure that children receive targeted, engaging, and effective support, ultimately enhancing their cognitive and functional independence.

Lastly, a meta-analysis conducted by Moon et al [23] aimed to review the effectiveness of mobile apps in the treatment of individuals with ASD. After a review of 1100 randomized controlled trials, only 7 studies were deemed suitable for further analysis, suggesting a very methodological approach. Using the Mullen Scales of Early Learning, the results favored the intervention group, indicating a significant improvement in the participants’ early learning and developmental outcomes compared to control groups. Moreover, the analysis found minimal heterogeneity (P>.10) across different studies or no significant evidence of publication bias. Correspondingly, our platform aligns with these findings by offering a technology-based, interactive tool specifically designed to enhance learning and developmental progress in individuals with autism. With an emphasis on providing individualized interventions that target key skills, CognitiveBotics uses validated clinical parameters to monitor improvements, reducing inaccuracies, similar to the studies highlighted in Moon’s analysis [23].

Limitations of the Study

Although evidence from our longitudinal study shows significant improvement in outcome measures for individuals with ASD using the software, a few limitations have to be discussed. First, the small sample size of 40 participants is a critical limitation, suggesting inadequate generalization of the findings. However, most studies regarding children with autism often face challenges in recruiting adequate numbers of participants. Limited research has explored effective strategies for efficiently

recruiting participants with ASD, a challenge that poses a barrier to larger and more comprehensive studies in this field [69].

Second, the participants were recruited from a single center and predominantly came from literate and urban families. Such a demographic is not representative of the entire population of individuals with ASD, particularly in India. The benefits observed in using the software may not translate to individuals with a lower socioeconomic status or those located in rural areas, who may face different challenges and have different needs. Further studies should be conducted to include participants from rural areas and various socioeconomic backgrounds. This includes incorporating features that reflect local languages and cultural sensitivities to ensure the software is relevant and effective for a wider range of users.

Third, the study experienced a 59% attrition rate, which could be attributed to several factors, including language barriers or the demanding schedules of caregivers, which may have limited their ability to fully engage with the platform. Such high levels of attrition are commonly observed in digital therapeutics for mental health. Similarly, a recent meta-analysis found more than half of the users discontinued using smartphone apps aimed at treating depressive symptoms [70].

Finally, while randomized controlled trials are considered the gold standard for assessing the effectiveness of interventions, their feasibility in such a population remains challenging. To address this, future research should explore methodologies that balance scientific rigor with practical implementation to further validate the software's effectiveness among different subgroups.

Conclusions

This 12-month study demonstrated that the CognitiveBotics platform delivering parent-mediated interventions significantly improved multiple developmental and social parameters in participants. Furthermore, it highlights that these digital technologies using audiovisuals, AI-based interactive games, animation games, and chatbots have an attraction factor that keeps the interest of children with ASD. Particularly, the incorporation of AI into digital technology has been shown to enhance social communication skills, especially in younger participants with learning difficulties, helping them reach their specific learning objectives.

Most assistive technologies are not intended to satisfy the needs of individuals with ASD as a whole, as they have variable needs. Despite being in its infancy, such digital technologies have been proposed to address the wide array of learning needs and work on the core symptoms of ASD. Further research must be conducted to include a larger number of children with different levels of social and developmental delays and ASD severity along with regional, linguistic, and sociocultural variations.

In conclusion, the promising results of this study underscore the potential of AI software interventions in revolutionizing holistic support for children with ASD. As these technologies continue to evolve, aligning the software not just to the needs of the child but also to those of parents and therapists offers hope for more personalized and effective strategies for not just children on the autism spectrum but also all neurodiverse children.

Acknowledgments

The authors are grateful to the participants and their families for their cooperation, support, and commitment. Furthermore, the authors would like to acknowledge Dr Pravalika Deti, Mr Varada Chaitanya, Ms Srilekha Gayatri, and Dr Keerthana Tummuri for their contribution to the research. This study was funded and sponsored by CognitiveBotics Technologies Private Limited, which provided the financial support necessary for study implementation, data collection, and analysis.

Conflicts of Interest

The primary author (HA) is on the Advisory Board of CognitiveBotics. HA was also actively involved in designing the study methodology and contributed to drafting and revising the manuscript. The principal investigator (SN) conducted research at the study site and received an honorarium for overseeing the study's execution. The corresponding author (BRR) is currently employed at CognitiveBotics Technologies Private Limited.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2013. URL: https://doi.org/10.1176/appi.books.9780890425596 [accessed 2025-04-21]
2. Legg H, Tickle A. UK parents' experiences of their child receiving a diagnosis of autism spectrum disorder: a systematic review of the qualitative evidence. Autism 2019 Nov;23(8):1897-1910. [doi: 10.1177/1362361319841488] [Medline: 30995082]
3. Özerk G, Özerk K, Silveira-Zaldivara T. Developing social skills and social competence in children with autism. iejee 2021 Mar 1;13(3):341-363. [doi: 10.26822/iejee.2021.195]
4. Hendricks DR, Wehman P. Transition from school to adulthood for youth with autism spectrum disorders. Focus Autism Other Dev Disabl 2009 Jun;24(2):77-88. [doi: 10.1177/1088357608329827]
5. Howlin P, Moss P, Savage S, Rutter M. Social outcomes in mid- to later adulthood among individuals diagnosed with autism and average nonverbal IQ as children. J Am Acad Child Adolesc Psychiatry 2013 Jun;52(6):572-581. [doi: 10.1016/j.jaac.2013.02.017] [Medline: 23702446]

6. Lotter V. Epidemiology of autistic conditions in young children. Soc Psychiatry 1966 Dec;1(3):124-135. [doi: 10.1007/BF00584048]
7. Autism Spectrum Disorder (ASD). Data and statistics on autism spectrum disorder. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/autism/data-research/index.html [accessed 2025-04-13]
8. Emberti Gialloreti L, Curatolo P. Autism spectrum disorder: why do we know so little? Front Neurol 2018;9:670. [doi: 10.3389/fneur.2018.00670] [Medline: 30174643]
9. Franz L, Goodwin CD, Rieder A, Matheis M, Damiano DL. Early intervention for very young children with or at high likelihood for autism spectrum disorder: an overview of reviews. Dev Med Child Neurol 2022 Sep;64(9):1063-1076. [doi: 10.1111/dmcn.15258] [Medline: 35582893]
10. Shenoy MD, Indla V, Reddy H. Comprehensive management of autism: current evidence. Indian J Psychol Med 2017;39(6):727-731. [doi: 10.4103/IJPSYM.IJPSYM_272_17] [Medline: 29284801]
11. Case-Smith J, Weaver LL, Fristad MA. A systematic review of sensory processing interventions for children with autism spectrum disorders. Autism 2015 Feb;19(2):133-148. [doi: 10.1177/1362361313517762] [Medline: 24477447]
12. Bowman KS, Suarez VD, Weiss MJ. Standards for interprofessional collaboration in the treatment of individuals with autism. Behav Anal Pract 2021 Dec;14(4):1191-1208. [doi: 10.1007/s40617-021-00560-0] [Medline: 34868822]
13. Ruble LA, McGrew J, Dalrymple N, Jung LA. Examining the quality of IEPs for young children with autism. J Autism Dev Disord 2010 Dec;40(12):1459-1470. [doi: 10.1007/s10803-010-1003-1] [Medline: 20373007]
14. Malik-Soni N, Shaker A, Luck H, et al. Tackling healthcare access barriers for individuals with autism from diagnosis to adulthood. Pediatr Res 2022 Apr;91(5):1028-1035. [doi: 10.1038/s41390-021-01465-y] [Medline: 33767375]
15. Yingling ME, Ruther MH, Dubuque EM, Mandell DS. County-level variation in geographic access to Board Certified Behavior Analysts among children with autism spectrum disorder in the United States. Autism 2021 Aug;25(6):1734-1745. [doi: 10.1177/13623613211002051] [Medline: 33740869]
16. Stoddart KP. Introduction to Asperger syndrome: a developmental lifespan perspective. In: Children, Youth, and Adults with Asperger Syndrome: Integrating Multiple Perspectives 2005:9-32 URL: https://www.researchgate.net/profile/Kevin-Stoddart-2/publication/273260232_Introduction_to_Asperger_Syndrome_A_Developmental_Lifespan_Perspective/links/55a544ea08ae00cf99c95791/Introduction-to-Asperger-Syndrome-A-Developmental-Lifespan-Perspective.pdf [accessed 2025-04-21]
17. Vasquez S, Brewer A, Leon Y, Vasquez J. The effects of advance notice on problem behavior occasioned by interruptions of an ongoing activity in a young girl with autism. Behav Anal Pract 2017 Dec;10(4):417-421. [doi: 10.1007/s40617-017-0187-7] [Medline: 29214139]
18. Osborne LA, Reed P. Stress and self-perceived parenting behaviors of parents of children with autistic spectrum conditions. Res Autism Spectr Disord 2010 Jul;4(3):405-414. [doi: 10.1016/j.rasd.2009.10.011]
19. Mughal S, Faizy RM, Saadabadi A. Autism spectrum disorder. In: StatPearls: Treasure Island (FL): StatPearls Publishing; 2024.
20. Ankeny EM, Wilkins J, Spain J. Mothers' experiences of transition planning for their children with disabilities. Teaching Exceptional Children 2009 Jul;41(6):28-36. [doi: 10.1177/004005990904100604]
21. Fletcher-Watson S. A targeted review of computer-assisted learning for people with autism spectrum disorder: towards a consistent methodology. Rev J Autism Dev Disord 2014 Jun;1(2):87-100. [doi: 10.1007/s40489-013-0003-4]
22. Roschelle JM, Pea RD, Hoadley CM, Gordin DN, Means BM. Changing how and what children learn in school with computer-based technologies. Future Child 2000;10(2):76-101. [doi: 10.2307/1602690] [Medline: 11255710]
23. Moon SJ, Hwang J, Hill HS, et al. Mobile device applications and treatment of autism spectrum disorder: a systematic review and meta-analysis of effectiveness. Arch Dis Child 2020 May;105(5):458-462. [doi: 10.1136/archdischild-2019-318258] [Medline: 31848147]
24. Atturu H, Lakhani S. Impact of COVID-19 on autism spectrum disorder management - a therapist perspective. Eur Psychiatr 2021 Apr;64(S1):S104-S104. [doi: 10.1192/j.eurpsy.2021.301]
25. Pruneti C, Coscioni G, Guidotti S. Evaluation of the effectiveness of behavioral interventions for autism spectrum disorders: a systematic review of randomized controlled trials and quasi-experimental studies. Clin Child Psychol Psychiatry 2024 Jan;29(1):213-231. [doi: 10.1177/13591045231205614] [Medline: 37803891]
26. Scarcella I, Marino F, Failla C, et al. Information and communication technologies-based interventions for children with autism spectrum conditions: a systematic review of randomized control trials from a positive technology perspective. Front Psychiatry 2023;14:1212522. [doi: 10.3389/fpsyt.2023.1212522] [Medline: 37547218]
27. Kohli M, Kar AK, Bangalore A, Ap P. Machine learning-based ABA treatment recommendation and personalization for autism spectrum disorder: an exploratory study. Brain Inform 2022 Jul 25;9(1):16. [doi: 10.1186/s40708-022-00164-6] [Medline: 35879626]
28. Li G, Zarei MA, Alibakhshi G, Labbafi A. Teachers and educators' experiences and perceptions of artificial-powered interventions for autism groups. BMC Psychol 2024 Apr 11;12(1):199. [doi: 10.1186/s40359-024-01664-2] [Medline: 38605422]
29. Iannone A, Giansanti D. Breaking barriers-the intersection of AI and assistive technology in autism care: a narrative review. J Pers Med 2023 Dec 28;14(1):41. [doi: 10.3390/jpm14010041] [Medline: 38248742]

30. Lampos V, Mintz J, Qu X. An artificial intelligence approach for selecting effective teacher communication strategies in autism education. NPJ Sci Learn 2021 Sep 1;6(1):25. [doi: 10.1038/s41539-021-00102-x] [Medline: 34471124]
31. Frolli A, Ricci MC, Bosco A, et al. Video modeling and social skills learning in ASD-HF. Children (Basel) 2020 Dec 8;7(12):279. [doi: 10.3390/children7120279] [Medline: 33302423]
32. Penev Y, Dunlap K, Husic A, et al. A mobile game platform for improving social communication in children with autism: a feasibility study. Appl Clin Inform 2021 Oct;12(5):1030-1040. [doi: 10.1055/s-0041-1736626] [Medline: 34788890]
33. Almurashi H, Bouaziz R, Alharthi W, Al-Sarem M, Hadwan M, Kammoun S. Augmented reality, serious games and picture exchange communication system for people with ASD: systematic literature review and future directions. Sensors (Basel) 2022 Feb 7;22(3):1250. [doi: 10.3390/s22031250] [Medline: 35161995]
34. Behnamnia N, Kamsin A, Ismail MAB, Hayati SA. A review of using digital game-based learning for preschoolers. J Comput Educ 2023 Dec;10(4):603-636. [doi: 10.1007/s40692-022-00240-0]
35. Meneses do Rêgo AC, Araújo-Filho I. Leveraging artificial intelligence to enhance the quality of life for patients with autism spectrum disorder: a comprehensive review. Clinicmed 2024;5(5):28-38. [doi: 10.24018/clinicmed.2024.5.5.350]
36. Atherton G, Cross L. The use of analog and digital games for autism interventions. Front Psychol 2021;12:669734. [doi: 10.3389/fpsyg.2021.669734] [Medline: 34447328]
37. Moore M, Calvert S. Brief report: vocabulary acquisition for children with autism: teacher or computer instruction. J Autism Dev Disord 2000 Aug;30(4):359-362. [doi: 10.1023/a:1005535602064] [Medline: 11039862]
38. Prithishkumar IJ, Michael SA. Understanding your student: using the VARK model. J Postgrad Med 2014;60(2):183-186. [doi: 10.4103/0022-3859.132337] [Medline: 24823519]
39. CognitiveBotics. URL: https://cognitivebotics.com/ [accessed 2025-04-21]
40. Atturu H, Thumkunta BR, Lakhani S, Naraganti SP. P.0046 Learning from e-learning – option for personalized intervention in autism children. Eur Neuropsychopharmacol 2021 Dec;53:S34. [doi: 10.1016/j.euroneuro.2021.10.051]
41. Nuske HJ, Mandell DS. Digital health should augment (not replace) autism treatment providers. Autism 2021 Oct;25(7):1825-1827. [doi: 10.1177/13623613211043368] [Medline: 34448659]
42. Schopler E, Reichler RJ, DeVellis RF, Daly K. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). J Autism Dev Disord 1980 Mar;10(1):91-103. [doi: 10.1007/BF02408436] [Medline: 6927682]
43. Russell PSS, Daniel A, Russell S, et al. Diagnostic accuracy, reliability and validity of Childhood Autism Rating Scale in India. World J Pediatr 2010 May;6(2):141-147. [doi: 10.1007/s12519-010-0029-y] [Medline: 20490769]
44. Doll EA. A genetic scale of social maturity. American Journal of Orthopsychiatry 1935;5(2):180-190. [doi: 10.1111/j.1939-0025.1935.tb06339.x]
45. Malin AJ. Vineland social maturity scale: Nagpur adaptation. : Indian Psychological Corporation; 1971.
46. Bharatraj J. Vineland social maturity scale-Indian adaptation: enlarged version. : Mysore: Swayamsiddha Prakashanam; 1992.
47. Frankenburg WK, Dodds JB. The Denver developmental screening test. J Pediatr 1967 Aug;71(2):181-191. [doi: 10.1016/s0022-3476(67)80070-2] [Medline: 6029467]
48. Bharatraj J. Developmental screening test. : Mysore: Swayam Sidha Prakashanam; 1977.
49. Kumar V, Mehta R. Adaptation and validation of receptive expressive emergent Language Test-3: evidence from Hindi speaking children with cochlear implant. Int J Pediatr Otorhinolaryngol 2020 May;132:109891. [doi: 10.1016/j.ijporl.2020.109891] [Medline: 32006860]
50. Grynszpan O, Weiss PLT, Perez-Diaz F, Gal E. Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. Autism 2014 May;18(4):346-361. [doi: 10.1177/1362361313476767] [Medline: 24092843]
51. Light J, McNaughton D. Supporting the communication, language, and literacy development of children with complex communication needs: state of the science and future research priorities. Assist Technol 2012 Mar;24(1):34-44. [doi: 10.1080/10400435.2011.648717]
52. Kagohara DM, van der Meer L, Ramdoss S, et al. Using iPods and iPads in teaching programs for individuals with developmental disabilities: a systematic review. Res Dev Disabil 2013 Jan;34(1):147-156. [doi: 10.1016/j.ridd.2012.07.027] [Medline: 22940168]
53. Knight V, McKissick BR, Saunders A. A review of technology-based interventions to teach academic skills to students with autism spectrum disorder. J Autism Dev Disord 2013 Nov;43(11):2628-2648. [doi: 10.1007/s10803-013-1814-y] [Medline: 23543292]
54. Stockall N, Dennis LR. Using pivotal response training and technology to engage preschoolers with autism in conversations. Interv Sch Clin 2014 Mar;49(4):195-202. [doi: 10.1177/1053451213509486]
55. Qahmash AIM. The potentials of using mobile technology in teaching individuals with learning disabilities: a review of special education technology literature. TechTrends 2018 Nov;62(6):647-653. [doi: 10.1007/s11528-018-0298-1]
56. Odunukwe CN. Using modern technology to enhance learning of students with autism spectrum disorders. Presented at: ADECT 2019 Proceedings; May 22-24, 2019; Abuja FCT, Nigeria URL: https://open.library.okstate.edu/adect/chapter/using-modern-technology-to-enhance-learning-of-students-with-autism-spectrum-disorders/ [accessed 2025-04-13]
57. Massaro DW, Bosseler A. Perceiving speech by ear and eye: multimodal integration by children with autism. J Dev Learn Disord 2003 Oct:111-144.

58. Pickard K, Rowless S, Ingersoll B. Understanding the impact of adaptations to a parent-mediated intervention on parents' ratings of perceived barriers, program attributes, and intent to use. Autism 2019 Feb;23(2):338-349. [doi: 10.1177/1362361317744078] [Medline: 29172633]
59. Beaumont R, Walker H, Weiss J, Sofronoff K. Randomized controlled trial of a video gaming-based social skills program for children on the autism spectrum. J Autism Dev Disord 2021 Oct;51(10):3637-3650. [doi: 10.1007/s10803-020-04801-z] [Medline: 33389304]
60. Nakshine VS, Thute P, Khatib MN, Sarkar B. Increased screen time as a cause of declining physical, psychological health, and sleep patterns: a literary review. Cureus 2022 Oct;14(10):e30051. [doi: 10.7759/cureus.30051] [Medline: 36381869]
61. Parent J, Sanders W, Forehand R. Youth screen time and behavioral health problems: the role of sleep duration and disturbances. J Dev Behav Pediatr 2016 May;37(4):277-284. [doi: 10.1097/DBP.0000000000000272] [Medline: 26890562]
62. Pennington RC. Computer-assisted instruction for teaching academic skills to students with autism spectrum disorders: a review of literature. Focus Autism Other Dev Disabl 2010 Dec;25(4):239-248. [doi: 10.1177/1088357610378291]
63. Domoff SE, Radesky JS, Harrison K, Riley H, Lumeng JC, Miller AL. A naturalistic study of child and family screen media and mobile device use. J Child Fam Stud 2019 Feb;28(2):401-410. [doi: 10.1007/s10826-018-1275-1] [Medline: 31105418]
64. Alotaibi MS. Game-based learning in early childhood education: a systematic review and meta-analysis. Front Psychol 2024;15:1307881. [doi: 10.3389/fpsyg.2024.1307881] [Medline: 38629045]
65. Sandbank M, Bottema-Beutel K, Crowley LaPoint S, et al. Autism intervention meta-analysis of early childhood studies (Project AIM): updated systematic review and secondary analysis. BMJ 2023 Nov 14;383:e076733. [doi: 10.1136/bmj-2023-076733] [Medline: 37963634]
66. Novack MN, Hong E, Dixon DR, Granpeesheh D. An evaluation of a mobile application designed to teach receptive language skills to children with autism spectrum disorder. Behav Anal Pract 2019 Mar;12(1):66-77. [doi: 10.1007/s40617-018-00312-7] [Medline: 30918771]
67. Cheng Y, Luo SY, Lin HC, Yang CS. Investigating mobile emotional learning for children with autistic spectrum disorders. Int J Dev Disabil 2016 Aug 8;64(1):25-34. [doi: 10.1080/20473869.2016.1206729] [Medline: 34141288]
68. Alsolami AS. The effectiveness of using artificial intelligence in improving academic skills of school-aged students with mild intellectual disabilities in Saudi Arabia. Res Dev Disabil 2025 Jan;156:104884. [doi: 10.1016/j.ridd.2024.104884] [Medline: 39631141]
69. Ahmed KL, Simon AR, Dempsey JR, Samaco RC, Goin-Kochel RP. Evaluating two common strategies for research participant recruitment into autism studies: observational study. J Med Internet Res 2020 Sep 24;22(9):e16752. [doi: 10.2196/16752] [Medline: 32969826]
70. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419. [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]

Abbreviations

ABA: applied behavioral analysis
AI: artificial intelligence
ASD: autism spectrum disorder
CAI: computer-assisted interventions
CARS: Childhood Autism Rating Scale
DA: developmental age
DQ: developmental quotient
DST: Developmental Screening Test
ELA: expressive language age
ILP: individualized learning plan
REEL: Receptive and Expressive Emergent Language Test
RLA: receptive language age
SA: social age
SQ: social quotient
VSMS: Vineland Social Maturity Scale

Edited by P Kubben; submitted 26.12.24; peer-reviewed by EA Ashaat, M Dunn; revised version received 19.03.25; accepted 21.03.25; published 28.04.25.

Please cite as:
Atturu H, Naraganti S, Rao BR
Effectiveness of Artificial Intelligence–Based Platform in Administering Therapies for Children With Autism Spectrum Disorder: 12-Month Observational Study
JMIR Neurotech 2025;4:e70589
URL: https://neuro.jmir.org/2025/1/e70589
doi:10.2196/70589

Reduction of Anxiety-Related Symptoms Using Low-Intensity Ultrasound Neuromodulation on the Auricular Branch of the Vagus Nerve: Preliminary Study

Izzy Kohler, DPT; Jon Hacker, BS; Ethan Martin, BS

NeurGear, 295 Warrington Dr, Rochester, NY, United States

Corresponding Author:
Izzy Kohler, DPT
NeurGear, 295 Warrington Dr, Rochester, NY, United States

Abstract

Background: Neuromodulation of the auricular branch of the vagus nerve using low-intensity focused ultrasound (LIFU) is an emerging mode of treatment for anxiety that could provide a complementary or alternative treatment modality for individuals who are refractory to conventional interventions. The proposed benefits of this technology have been largely unexamined with clinical populations. Further research is required to understand its clinical potential and use in improving and managing moderate to severe symptoms.

Objectives: The aim of this study was to do a preliminary investigation into the efficacy, safety, and usability of the wearable headset that delivers LIFU to the auricular branch of the vagus nerve for the purpose of alleviating anxiety disorder symptoms.

Methods: This study was a pre-post intervention study design for which we recruited 28 participants with a Beck Anxiety Inventory score of 16 points or greater. Participants completed 5 minutes of treatment daily consisting of LIFU neuromodulation delivered to the auricular branch of the vagus nerve. Participants did this for a period of 4 weeks. Assessments of anxiety symptom severity (Beck Anxiety Inventory), depression symptom severity (Beck Depression Inventory), posttraumatic stress disorder symptom severity (Post Traumatic Stress Disorder Checklist for the *Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]), and sleep quality (Pittsburgh Sleep Quality Index) were taken prior to starting treatment and weekly for 4 weeks of treatment. Usability and safety were also assessed using an exit questionnaire and adverse event logging.

Results: After completing 4 weeks of LIFU neuromodulation to the auricular branch of the vagus nerve, the average Beck Anxiety Inventory score decreased by 14.9 (SD 10.6) points (Cohen d=1.06; P<.001), the average Beck Depression Inventory score decreased by 10.3 (SD 7.8) points (Cohen d=0.81; P<.001), the average Post Traumatic Stress Disorder Checklist for the *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition) score decreased by 20.0 (SD 20.5) points (Cohen d=0.94; P<.001), and the average Pittsburgh Sleep Quality Index score decreased by 2.2 (SD 3.1) points (Cohen d=0.65; P=.001). On the exit questionnaire, participants rated the treatment highly for ease of use, effectiveness, and worthiness of the time invested. Only 1 adverse event was reported throughout the entire trial, which was mild and temporary.

Conclusions: This preliminary study provided justification for further research into the efficacy, safety, and feasibility of using LIFU to modulate the auricular branch of the vagus nerve and reduce the symptoms of anxiety, depression, and posttraumatic stress disorder.

Trial Registration: ClinicalTrials.gov NCT06574971; https://clinicaltrials.gov/study/NCT06574971

(JMIR Neurotech 2025;4:e69770) doi:10.2196/69770

KEYWORDS

low-intensity focused ultrasound; auricular branch of the vagus nerve; anxiety; depression; posttraumatic stress disorder

Introduction

Anxiety is the “anticipation of real or imagined future threat or danger” [1], which manifests itself with a mix of emotional signals, such as hyperarousal and panic, and physiological ones, including increased heart rate, shortness of breath, sweating, and chest pain [2]. The emotional and physiological responses experienced with anxiety result from the activation of the hypothalamus, which engages the sympathetic nervous system (SNS) [3,4]. This sympathetic activation is adaptive in short bursts and enables us to handle threats and stressors, but in anxiety disorders, the SNS may be overly sensitive or chronically activated, leading to distress and health challenges over time [5,6]. Clinically significant anxiety symptoms are disproportionate to the future threat, endure after it has passed, and cause substantial distress or incapacitation [1,7]. The etiology of anxiety disorders is complex, with heritability ranging from 30% to 67% depending on the research study and

anxiety disorder type [1]. However, trauma, chronic stress, and other environmental factors play an important role in the development of maladaptive anxiety [7].

The complex etiology of anxiety opens opportunities for intervention at multiple points in the course of the illness from a variety of disciplines. There are also several multidisciplinary approaches that offer a more holistic care plan. The primary goal of preventative strategies is to lower the risk of developing disordered anxiety responses prior to onset. Preventative psychoeducational interventions for adolescents and adults have been shown to reduce the risk of anxiety onset [8] with small to moderate effect sizes [1,8]; however, studies of these interventions tend to end their follow-ups after only 9 months, so the long-term stability of their benefits after intervention completion is still in question [1]. Once an active anxiety disorder has developed, psychotherapeutic treatments for it range in intensity from self-guided programs to highly intense weekly sessions with a licensed therapist. Self-guided treatments derived from evidence-based psychotherapies are more effective than active controls but show smaller effect sizes than therapist-guided programs [9]. Cognitive behavioral therapy is widely considered to be the gold standard for anxiety disorder treatment, particularly in adults, although Haller et al [10] found mindfulness-based cognitive therapy and acceptance and commitment therapy to be similar in efficacy. In recent years, virtual psychotherapy modalities have emerged as a compromise that balances the convenience of self-help approaches and the rigor and guidance of a traditional in-person therapy session. Thus, recent advances in telehealth have paved the way for approaches that afford convenience and accessibility without a loss of efficacy [11,12].

Pharmacotherapy is similar in efficacy to psychotherapy, and both pharmacotherapy and psychotherapy are considered first-line treatments for anxiety disorders in most standard care plans [1]. Selective serotonin reuptake inhibitors, serotonin and norepinephrine reuptake inhibitors, benzodiazepines, antipsychotics, and β-blockers are all used to treat anxiety. Despite this wealth of options, anxiety disorders remain chronic and refractory to treatment in many individuals, with 15% - 40% achieving less than 50% remission in symptoms [13]. Studies of combinations of psychotherapeutic and pharmacological approaches to anxiety treatment are sparse, leaving confusion surrounding which combinations are most efficacious [1]. Taken as a whole, while current neurobiological and psychosocial treatment approaches to anxiety disorders are sufficient for a large portion of affected individuals, there is still a substantial proportion of patients who would benefit from additional treatment options.

Low-intensity focused ultrasound (LIFU) is an emerging mode of treatment for anxiety that could provide an alternative treatment modality. LIFU can stimulate or inhibit neural activity, depending on the parameters of the energy applied to neural tissue. Also referred to as acoustic neuromodulation, the use of LIFU to modulate the activity of neural structures is a promising method for noninvasive treatment of neurological disorders [14]. While the majority of investigations featuring LIFU neuromodulation have primarily focused on modulation of neural structures within the central nervous system, disorders affecting the peripheral nervous system stand to benefit from this powerful tool as well [15]. LIFU neuromodulation of the peripheral nervous system is accomplished through a nonthermal, noncavitation bioeffect produced by setting the parameters to the intermediate intensity range. At intensities between 1 and 200 W/cm^2, ultrasound is able to noninvasively and reversibly enhance peripheral neural activities by activating low-threshold mechanosensitive nerve endings, opening mechanosensitive ion channels to evoke action potentials [15]. Ultrasound of intermediate intensity also enhances the neural activity of peripheral nerve axons, leading to increased nerve conduction velocities in both A- and C-type fibers, which is likely caused by mechanical gating of other ion channels [16]. In addition, enhanced neural activity could be attributed to a direct effect of acoustic radiation forces on the lipid-bilayer neural membrane. Plausible mechanisms for this include a transient capacitive current from rapid changes of local membrane capacitance and transmembrane pore formation to allow sodium and potassium ions to pass through [15,16].

The vagus nerve, also known as cranial nerve X, is the longest cranial nerve and its branches enable the organs to adjust to the demands of a person's internal state and external environment. The vagus nerve is a primary component of the parasympathetic nervous system, which, paired with the SNS, constitutes the autonomic nervous system [4,17]. Normally, sympathetic and parasympathetic nerve pathways act synergistically to create a state of equilibrium appropriate to meet the demands of the current internal state and external challenges. Disruption of the balance of sympathetic and parasympathetic activity in favor of sympathetic activity is one indicator of anxiety disorders [4,18].

The many branches of the vagus nerve are increasingly seen as pathways for promoting or restoring health and ameliorating the physiologic unease that gives rise to anxiety and other negative mental states [19]. The vagus nerve operates bidirectionally, meaning states of homeostasis and calm can be induced from the bottom up or the top down. The brain can use cognitive strategies to dissipate states of bodily unease (top down) or activate vagal nerve pathways to create psychological comfort and a sense of safety (bottom up) [20]. In addition to its role in regulating the parasympathetic nervous system, the vagus nerve also projects to the amygdala and hippocampus, both of which are important to extinction learning techniques commonly used in the treatment of anxiety and posttraumatic stress disorder (PTSD) [21,22]. Stimulation of the vagus nerve can downregulate sympathetic activity, restoring visceral order and psychological calm [23,24].

Early research into the clinical applications of vagus nerve stimulation (VNS) primarily centered on epilepsy and depression [17], but the vagus nerve is an attractive target for antianxiety therapies as well. In addition to its role in regulating the parasympathetic nervous system, the vagus nerve also projects to the amygdala and hippocampus, both of which are important to extinction learning techniques commonly used in the treatment of anxiety and PTSD [21,22]. Preliminary clinical studies have demonstrated VNS's therapeutic applications to treatment-resistant anxiety disorders [23] and long COVID-19

symptoms [25]. Physiological changes as an effect of VNS are also well known in the literature. Wittbrodt et al [26] discovered that transcutaneous cervical VNS increased activation of the anterior cingulate and hippocampus during exposure to traumatic scripts. Lamb et al [27] found that transcutaneous auricular vagal nerve stimulation (taVNS) improved respiratory sinus arrhythmia and skin conductance during exposure to physical and emotional stress. Bremner et al [28] found that transcutaneous cervical VNS decreased inflammatory markers and sympathetic tone while increasing medial prefrontal function during exposure to trauma-specific and neutral stressors.

While VNS is traditionally done electrically, ultrasound's noninvasiveness and specificity make it ideal for VNS [29]. Ultrasound has been successfully used for vagus nerve neuromodulation in rats [30] and for peripheral nerve [29] and suborgan [31] stimulation in humans. With a recent study showing the feasibility of transauricular VNS as an at-home intervention [20,22], transauricular ultrasound VNS has emerged as a noninvasive, yet potentially effective, at-home treatment for the management of anxiety symptoms. In response to this, we have developed a wearable headset with an ultrasound transducer that delivers LIFU to the auricular branch of the vagus nerve that can be used at home for treatment of anxiety symptoms. The purpose of this study was to do a preliminary investigation into the efficacy, safety, and usability of the wearable headset that delivers LIFU to the auricular branch of the vagus nerve for the purpose of alleviating the symptoms of anxiety. Because depression [32] and PTSD [33] frequently co-occur with anxiety, we also investigated the efficacy of transauricular ultrasound VNS for alleviating the symptoms of depression and PTSD in individuals with anxiety.

Methods

Study Design

This was a pre-post-intervention study in which all participants received the intervention daily, at home, for a period of 4 weeks. The clinical trial is registered at ClinicalTrials.gov [NCT06574971]. Informed consent was obtained from each of the 28 participants prior to screening. All activities were completed remotely and a ZenBud device with a user manual and participant instructions was shipped to each participant's home. Participants completed 5 minutes of LIFU to the auricular branch of the vagus nerve each day using the ZenBud device. Treatment could be completed at any convenient time of day and did not have to be completed at the same time every day, as long as the treatment was completed within every 24-hour period. Assessments were completed on the web on the day before the first treatment session and then weekly. The final assessment was completed on the day of the final treatment after the final treatment session. The battery of assessments included 4 validated clinical outcome measures: Beck Anxiety Inventory (BAI), Beck Depression Inventory (BDI), PTSD Checklist for *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition) (*DSM-V*) (PC5), and Pittsburgh Sleep Quality Index (PSQI). The details of these assessments are further described in the data collection section.

Participant Recruitment

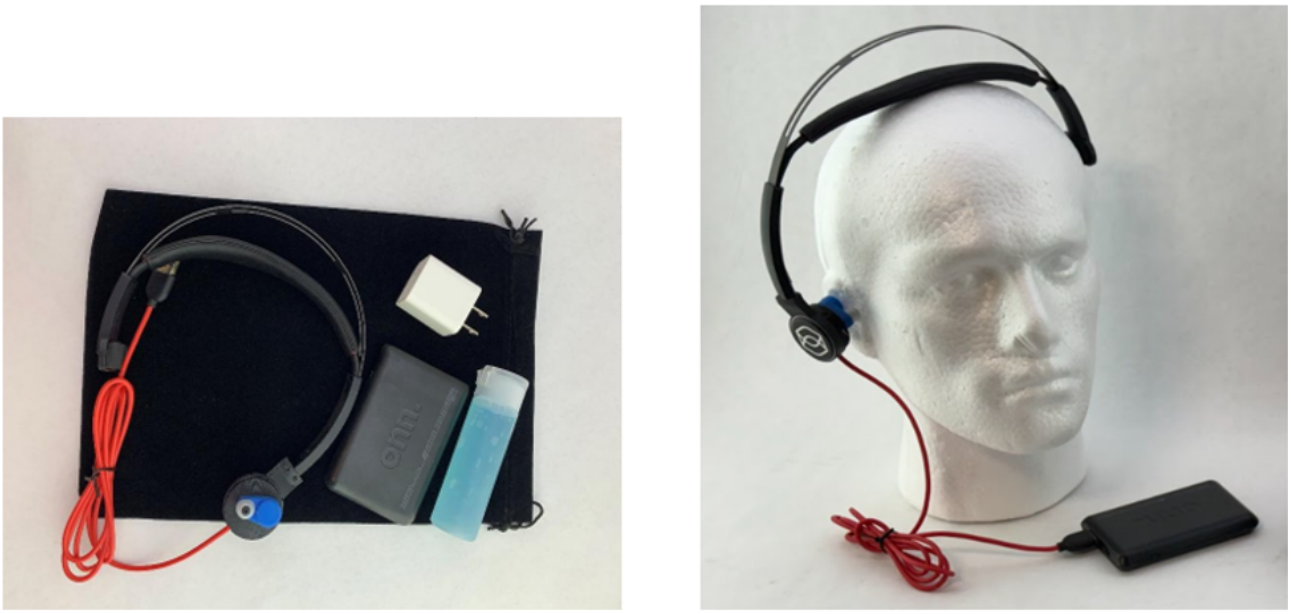
Adults in the United States were recruited through web-based social media advertising mentioning a study investigating a new treatment for anxiety disorders. Interested individuals filled out a study registration form containing only contact information and were then contacted by a member of the research team via email with further details of the study and a link to sign the informed consent. Upon completion of the informed consent, candidates were then screened for inclusion and exclusion criteria using web-based questionnaires. Interested individuals were included if they scored 16 or higher on the BAI, were older than 18 years, and did not have any additional conditions that were contraindications for VNS or ultrasound. Conditions that were contraindications for VNS included a history of vagotomy, heart arrhythmias, schizophrenia, or rapid cycling bipolar disorder. Conditions that were contraindications for ultrasound included presence of a pacemaker, pregnancy, active cancer, decreased sensation or open wounds in the ear, ear infection, or metal implants in or around the ear. A BAI score of 16 was chosen as the cutoff threshold because a score of 16 or higher in the BAI classifies an individual as having moderate to severe anxiety symptoms [34]. We did not exclude individuals who were receiving other treatments for their anxiety as long as the treatment was not initiated or ceased within the past month.

A total of 100 individuals completed the interest form, 63 signed the informed consent and were screened, and 28 were enrolled in the study. Each participant was assigned a unique identifier code so that participant information could be managed in a confidential manner throughout the study and the data could be deidentified upon completion of the study. Only the principal investigator and the study coordinator had access to the unique identifier code assignments.

Ultrasound Device

ZenBud, the device used for this trial, is a proprietary Conformité Européenne–compliant over-the-ear wearable headset that was developed by NeurGear (Figure 1A and B). The ZenBud delivers LIFU to the auricular branch of the vagus nerve through several layers of skin. The ZenBud is designed to mimic a standard headset so that users can integrate the use of the device into their routine with minimal effort and discomfort. When the user plugs the ZenBud device into the battery pack it immediately turns on. There is a hardware limit in the circuitry so that the device shuts down after running for 29 minutes, limiting the duration of use. The ZenBud device specifications include a center frequency of 5.3 MHz, a pulse repetition frequency of 41 Hz, a duty cycle of 50%, and an average intensity of 1.03 MPa.

Figure 1. (A) The ZenBud headset, powerpack, power brick, and bottle of gel. The ultrasound transducer is located in the round earpiece on the right side of the headset. (B) The ZenBud device as depicted properly placed on a model human head.

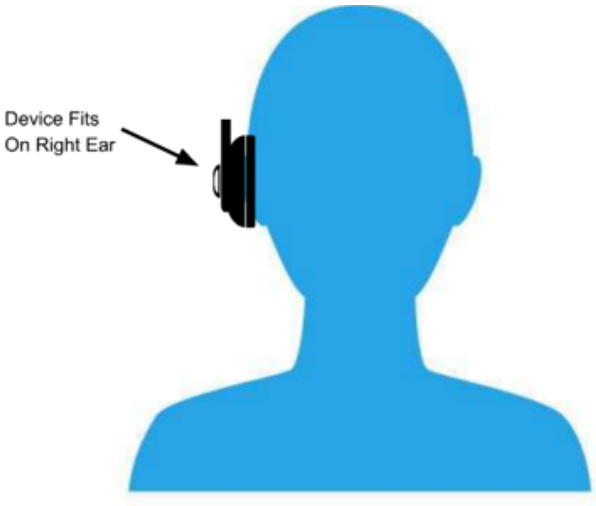


(A) (B)

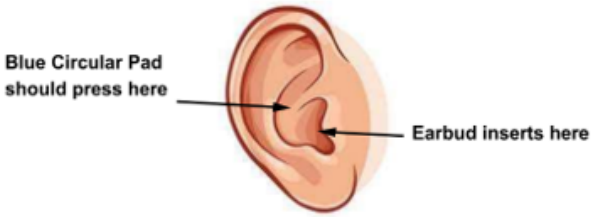
A detailed instruction manual was provided in the package with every device. A copy of the manual is provided as Multimedia Appendix 1. The participants were instructed to use the device once a day for 5 minutes unless instructed otherwise by a health care professional. There were no stipulations set for the time of day that treatment could be completed and participants were free to choose a time that was convenient for them. For step-by-step set up and use, participants were instructed to apply a pea-sized amount of the aqua sonic gel to the blue part of the device located directly above the headset (Figure 2A), position the blue circular pad against the skin just above the ear canal (Figure 2C), adjust the headset until they felt a moderate pressure (without pain) just above the ear canal where the blue circular pad was positioned (Figure 2B), and begin stimulation by plugging the USB cable into the battery pack (Figure 2D). Once the headset is plugged into the battery pack the device starts working and a low humming noise can be heard. The manual instructs users to listen for the humming sound to indicate that the device is working properly.

Figure 2. Images extracted from the ZenBud user manual depicting step-by-step setup and operation of the device. (A) Application of the ultrasonic gel. (B) Placement of the headset with the headset located over the right ear. (C) Correct placement of the headset on the ear. (D) Treatment is started upon inserting the USB cable into the battery pack.

(A)



(B)



(C)

(D)

Data Collection

Assessments were done using a battery of 4 validated clinical outcome measures. These were taken on the day before the first treatment session, weekly, and on the day of the final treatment session after the final treatment session was completed. The following 4 clinical outcome measures were used.

Beck Anxiety Inventory

The BAI is a rating scale used to evaluate the severity of anxiety symptoms in individuals aged 17 years and older. It contains 21 self-report items that reflect common physiological symptoms of anxiety such as numbness or tingling, feeling hot, and trembling. Participants indicate how much they have been bothered by each symptom, from “not at all” to “severely,” using a 4-point Likert scale. The item scores are then summed, with possible scores ranging from 0 to 63. A total score of 0 - 7 is classified as minimal anxiety, 8 - 15 as mild, 16 - 25 as moderate, and 26 - 63 as severe [35,36]. The BAI has a Cronbach α value of 0.91, a good test-retest reliability (κ=0.65, 95% CI 0.61-0.69), and correlates moderately (Pearson *r*=0.51) with the revised Hamilton Anxiety Rating Scale (HAM-A) [34,35,37,38].

Beck Depression Inventory Version II

Depression and anxiety are highly comorbid, with 60% of patients with anxiety disorders also having depression [32]. Long-term activation of the stress response may explain this overlap [39], implying that inhibiting overactivation of the stress response may alleviate depressive symptoms in addition to anxiety and stress. The BDI-II is a valid and reliable self-report measure for depression that quantifies depressive symptoms over the last week [40]. For each of the 21 items, respondents are asked to choose the statement they most agree with out of a group of 4 choices. Each statement corresponds to a score ranging from 0 to 3 and total scores range from 0 to 63 [35,41,42]. The scores are classified as minimal depression (0 - 13), mild depression (14-19), moderate depression (20-28), and severe depression (29-63) [38,41]. The BDI is positively correlated with the Hamilton Depression Rating Scale with a Pearson *r* of 0.71, showing good agreement. The test was also shown to have a high 1-week test-retest reliability (Pearson *r*=0.93), suggesting that it was not overly sensitive to daily variations in mood and high internal consistency (α=.91) [38,41].

PTSD Checklist for DSM-V (PCL-5)

While the *DSM-V* does not classify PTSD as an anxiety-related disorder, both PTSD and anxiety disorders involve dysregulation in neural structures dealing with fear, arousal, and anticipation of future threats [33]. Thus, there is reason to believe that VNS simulation could be beneficial for PTSD-related symptoms. The PCL-5 is a self-report questionnaire that helps assess the presence and severity of PTSD symptoms. The PCL-5 can be used to screen for PTSD, assist in making a provisional diagnosis, and monitor symptoms over time [43] . The measure asks participants to rate how much they were bothered by certain PTSD symptoms over the past month on a 5-point Likert scale ranging from "not at all" to "extremely" [44]. Total scores range from 0 to 60 and scores ranging from 31 to 33 are widely accepted as the cutoff for diagnosing PTSD [45]. In a systematic review of PCL-5 validation studies, Forkus et al [45] concluded that the full 20-item version showed good to excellent internal consistency across studies (Cronbach α values ranging from 0.83 to 0.97) and acceptable temporal stability (correlations ≥0.60) across time points within 1 - 5 weeks of one another. Scores were also moderately to highly correlated with other measures of PTSD as well as measures of anxiety, depression, suicidal ideation, and sleep.

Pittsburgh Sleep Quality Index

Anxiety and sleep disturbance are frequently co-occurring [46] such that sleep disturbance is a *DSM-V* criterion for generalized anxiety disorder. Studies have found correlations between BAI scores and subjective sleep quality among college students [47], indicating that measuring sleep quality could provide insight into the burden of anxiety on well-being. The PSQI is a validated and widely used global measure of sleep quality [48,49]. It comprises 19 self-report items and 5 items to be reported by a sleeping partner, but the 19 self-report items are commonly used on their own in research contexts [50]. The different items call for responses in different formats (bedtimes, number of hours, Likert scales, etc), thus the instrument is scored with the use of 7 component scores that are summed for 1 total score ranging from 0 to 21 [48]. The original creators of the PSQI found that a score of 5 or greater differentiated between "good" and "poor" sleepers, with a sensitivity of 89.6% and a specificity of 86.5% [48]. Research since has generally supported the validity of this cutoff. Mollayeva et al [49] did a meta-analysis of the psychometric properties of the PSQI and found that it showed acceptable internal reliability for within-group comparisons across studies (Cronbach α values ranging from 0.70 to 0.83). They also found that intraclass correlations for PSQI scores across timepoints met the cutoffs for use in within-group comparisons (0.70 or greater) [49].

Exit Survey

In addition to the clinical outcome measures, participants also completed an exit survey on the final day of the trial. This survey asked questions regarding overall satisfaction with the treatment, impact on daily functioning and quality of life, ease of use, symptom improvement, side effects, and how quickly effects from the treatment were perceived to be felt. The purpose of this questionnaire was to provide further insight into the perceived experiences of the participants during the treatment period, which is important information for full and complete understanding of the treatment's impact.

Adverse Event Tracking

Adverse events (AEs) and device deficiencies were documented and categorized in accordance with ISO14155:2020. These AEs were documented based on reports provided by the participants through email or on the exit survey. The investigators closely tracked the AEs and their resolution throughout the study. Each AE was categorized by type and seriousness according to the definitions provided in ISO14155. Whether an AE was related to the device or procedures was also distinguished. All available details for each AE were recorded in the participant case report forms, including relationship to the investigational device, severity (mild, moderate, or severe), onset date, resolution status, any action taken, and if there were any sequelae. For the causality assessment of all AEs, the MDCG 2020-10/1 guideline was followed. This guidance is specifically aimed at severe adverse events; however, it was extrapolated to all AEs for this study.

According to MDCG 2020-10/1, causal relatedness was defined as an AE associated with the investigational device beyond reasonable doubt. Probably device-related was defined as having a relationship with the use of the investigational device that seems relevant or the event cannot be reasonably explained by another cause. "Possibly device related" was defined as having a relationship with the use of the investigational device that was weak but cannot be ruled out completely. "Not device related" was defined as an event not having a temporal relationship with the device or not following a known response pattern to the device. The AEs were then further classified into mild, moderate, or severe categories. Mild severity AEs correspond to awareness of easily tolerated and mildly irritating signs or symptoms, with no or minimal loss of time from normal activities; these symptoms are transient and do not require therapy or a medical evaluation. Moderate cases are events that introduce a low level of inconvenience or concern to the participant and may interfere with daily activities; moderate experiences may cause some interference with functioning. Severe cases are events that substantially interrupt the participant's normal daily activities and generally require systemic drug therapy or other treatment; these events are usually incapacitating.

Statistical Analysis

The primary and secondary end points of the study are thoroughly described in the "Data Collection" section. These end points included pre- to posttreatment changes from baseline to the end of treatment at 4 weeks for the BAI as the primary end point and the BDI, PCL-5, and PSQI as secondary end points. Baseline scores were defined as the BAI, BDI, PCL-5, and PSQI scores on the first day of treatment, prior to the first treatment session. The within-group analyses were based on a per-protocol estimand and tested with paired 2-tailed *t* tests, where the normality assumption was confirmed with the Shapiro-Wilk test and α value was set to .05. The effect sizes reported in this paper are based on Cohen *d* and calculated as the mean score at the end of treatment minus the mean score at baseline, divided by the pooled SD for the 2 scores. The use of

per-protocol estimand ensured that the changes in outcome measures within each treatment arm were reflective of scenarios where the participants used the treatment as directed and thus included only the participants who were compliant to treatment. The usage criteria for inclusion in the per-protocol analysis was set at 5-29 minutes of treatment per day 6-7 days per week across the intended 4-week treatment period. There were only 2 missing scores, 1 in week 2 and 1 in week 3. Because these data are a time series that exhibits a trend line and the number of missing values was very small, these were filled using a linear interpolation between the score from the previous week and the score from the following week. There was no missing baseline or final scores.

To determine the appropriate sample size a power analysis was performed assuming a dependent *t* test with a significance level of 5%, power of 80%, and moderate effect size of 0.6 between pairs. This gave us a necessary sample size of 25 participants. Accounting for a potential dropout rate of 20% gave us a target sample size of 30 participants. All analyses were performed using GraphPad Prism 10.3.0 (507; Dotmatics).

Ethical Considerations

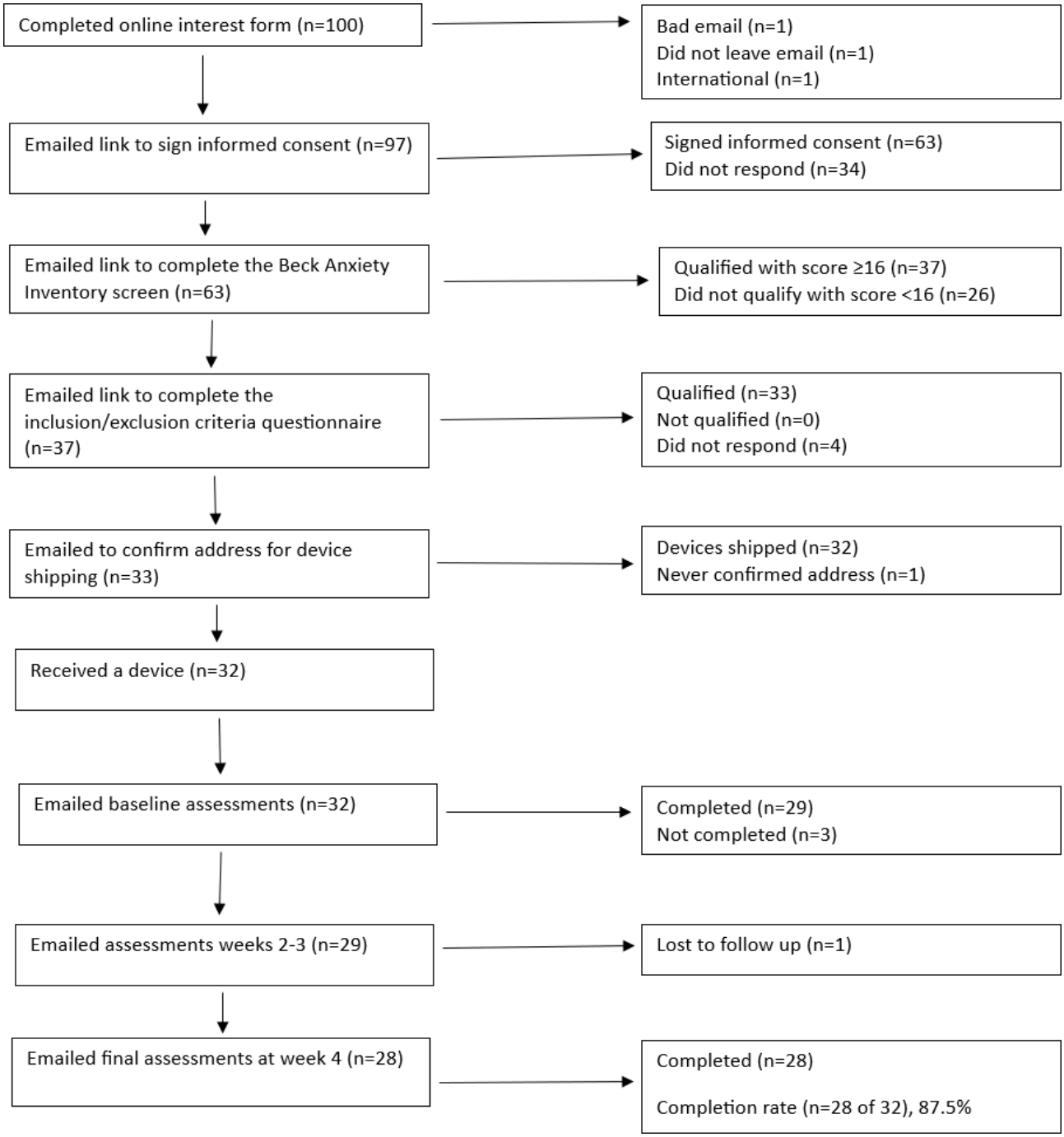
Ethical approval for this trial was obtained from the WIRB-Copernicus Group (WCG) institutional review board (reference no. 20233919), and the study was conducted in compliance with the principles outlined in the Declaration of Helsinki. Signed and documented informed consent was obtained from all participants prior to starting the study. For their time, participants were given a US $25.00 Amazon gift card.

Results

Study Participants

Between October 22, 2023, and October 2, 2024, 100 individuals completed the web-based interest form (Figure 3). A total of 63 participants consented to the trial, with 26 of these not satisfying the criteria of having a BAI score of 16 or greater, 4 not responding to requests to complete the screening questionnaire, and 1 not responding to requests for confirmation of their shipping address. A total of 32 participants were shipped a device, with 3 of these not responding to requests to complete the baseline assessments and 1 participant failing to respond to requests to take the reassessments after week 2. In total, 28 participants completed all LIFU sessions and weekly assessments (28/32, 87.5%). Data for all 28 participants who completed the trial are included in the analysis.

Figure 3. Flowchart of study participants through the trial.



The average age of the participants was 48.1 (SD 15.6) years. The group was heavily weighted toward women, with 22 women and 6 men. The National Institutes of Health reports that generalized anxiety affects approximately 2.7% of American adults, with women experiencing the disorder at a higher rate (3.4%) versus men (1.9%), making the fact that the sample contained a higher percentage of women a reflection of actual population distributions. The self-reported average duration of time suffering with anxiety was 16.5 (SD 11.8) years. There were also 8 participants currently receiving treatment for their anxiety and 20 who were not receiving any treatment.

Beck Anxiety Inventory

After 4 weeks of treatment with the ZenBud, the average BAI score decreased by 14.9 (SD 10.6) points from 26.5 (SD 12.5) to 11.5 (SD 11.1) (Figure 4). This change in score was both statistically significant (P<.001, 2-tailed dependent t test) and clinically meaningful. While there is no consistently defined definition of clinical improvement for the BAI, based on the categorical definitions of severity for the scores, there was a great deal of progression into decreased severity levels of anxiety throughout the treatment period. As seen in Figure 5, at the start of the study, 22 participants had BAI scores in the moderate or severe anxiety ranges and only 6 participants had BAI scores in the mild or minimal severity ranges. After 4 weeks of using the ZenBud, 22 participants had BAI scores into the mild or minimal severity rages, and only 6 participants had scores in the moderate or severe ranges. In terms of Cohen d, the effect size was large at 1.06.

Figure 4. The progression of Beck Anxiety Inventory scores through 4 weeks of treatment with ZenBud. The thin lines represent each individual participant. The thick line represents the group mean.

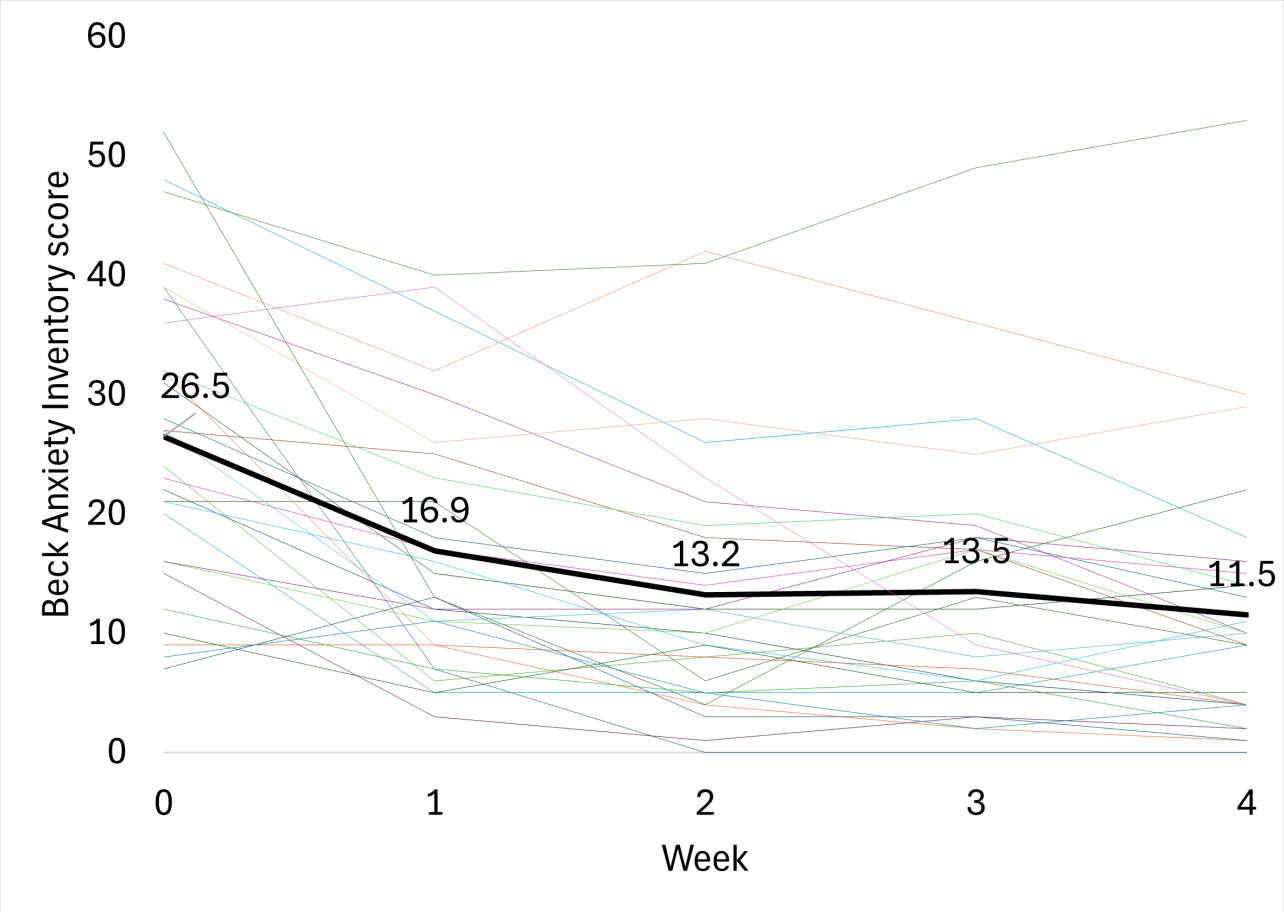
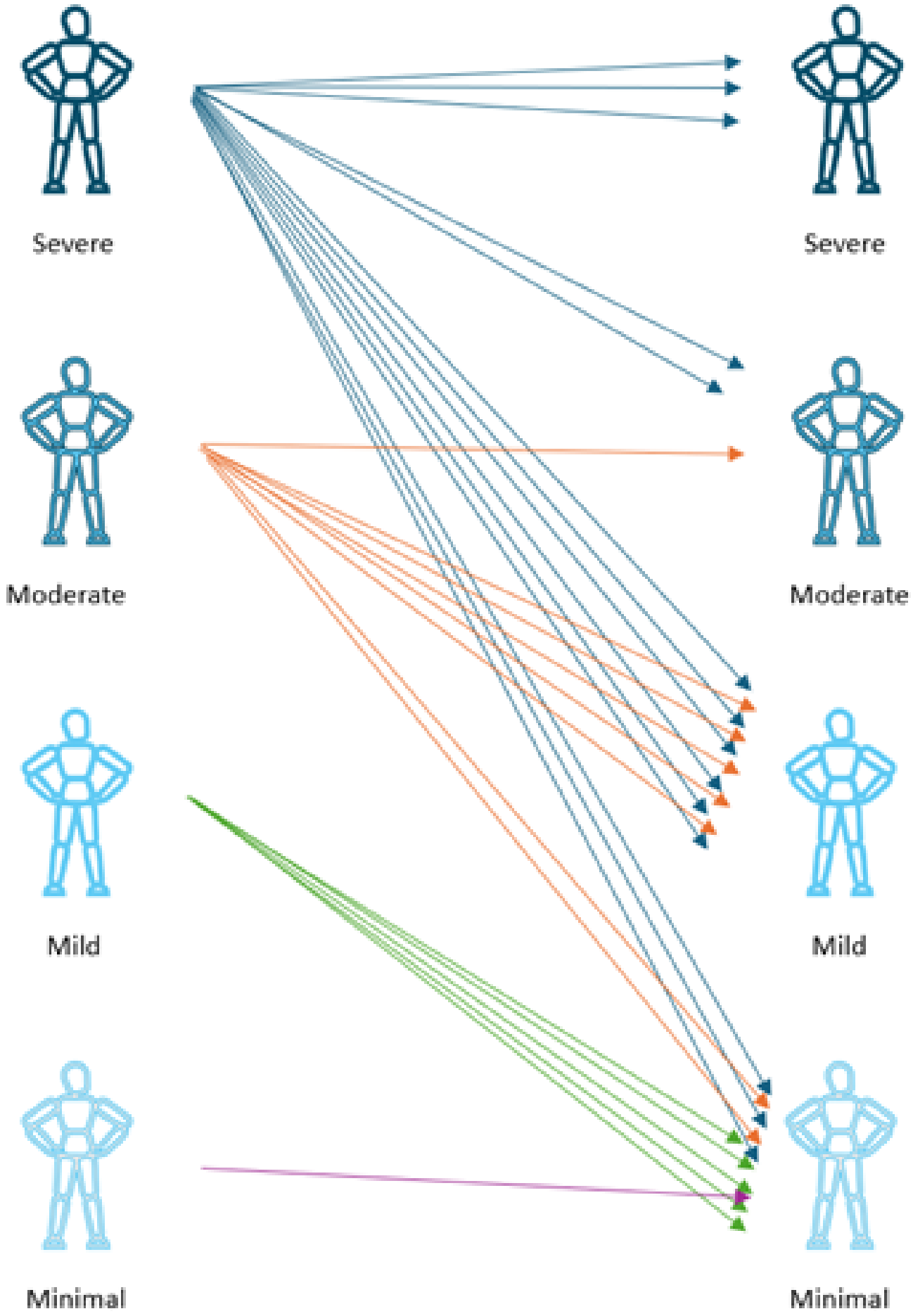


Figure 5. Categorical movement across degrees of severity based on the Beck Anxiety Inventory (BAI) definitions. At the start of the study, 20 participants had BAI scores in the moderate or severe anxiety ranges and only 6 participants had BAI scores in the mild or minimal severity ranges. After 4 weeks of using the ZenBud, 20 participants had BAI scores into the mild or minimal severity ranges, and only 6 participants had scores in the moderate or severe ranges.



Beck Depression Inventory

After 4 weeks of treatment with the ZenBud, the average BDI score decreased by 10.3 (SD 7.8) points from 24.2 (SD 10.5) to 13.9 (SD 12.6) (Figure 6). Similar to results seen for the BAI, this change in score was both statistically significant (P<.001; 2-tailed dependent t test) and clinically meaningful. A 17% reduction in score on the BDI is considered clinically meaningful [3]. Based on this definition, as seen in Table 1, 71.4% (20/28) of participants demonstrated a clinically meaningful reduction in score by the end of the trial. In terms of Cohen d, the effect size was large at 0.81.

Figure 6. The progression of Beck Depression Inventory scores through 4 weeks of treatment with ZenBud. The thin lines represent each individual participant. The thick line represents the group mean.

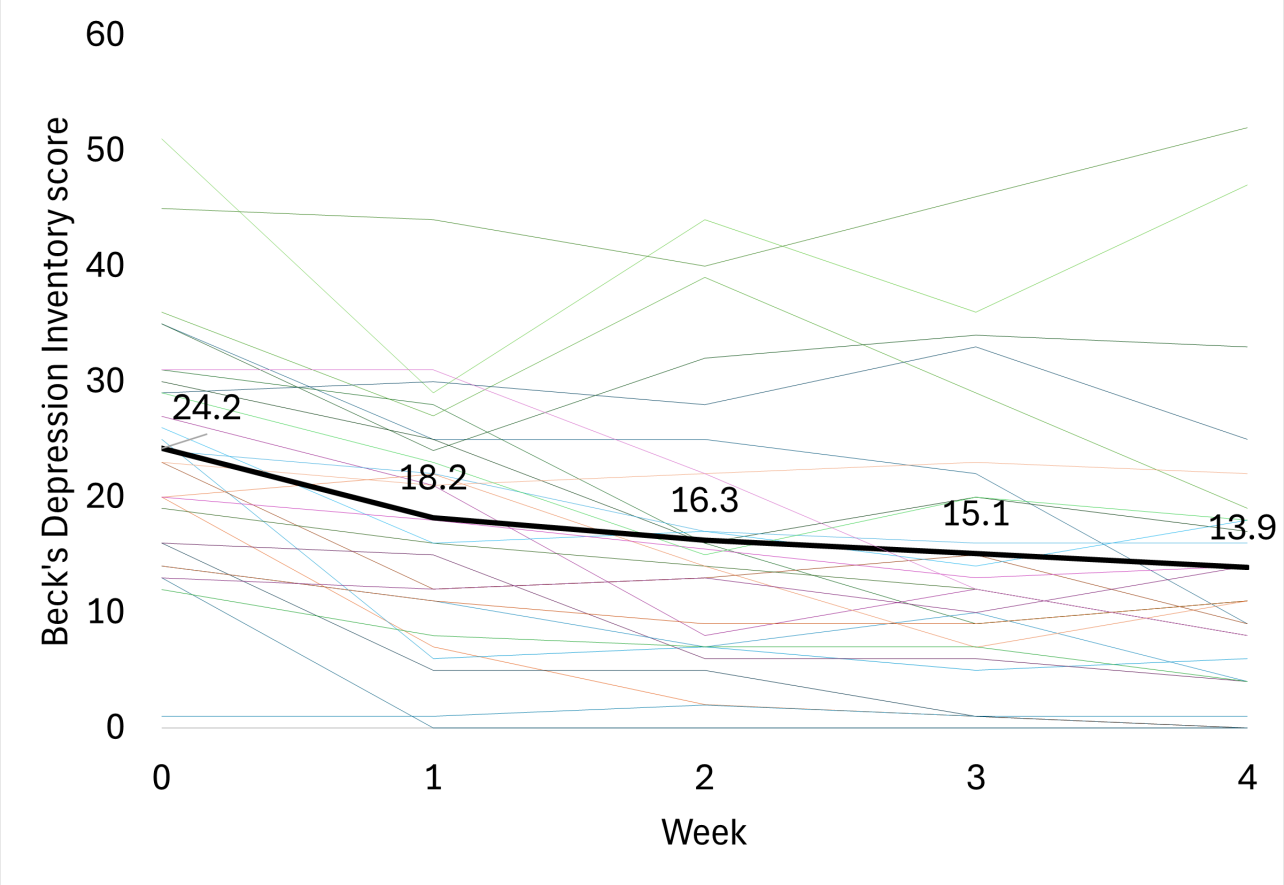


Table . The number of participants who experienced clinically significant reductions in Beck Depression Inventory score following 4 weeks of treatment with the ZenBud.

Degree of score change	Participants, n (%)
Clinical decrease	20 (71)
Nonclinical decrease	5 (18)
Nonclinical increase	3 (11)

Post Traumatic Stress Disorder Checklist for the *DSM-V*

After 4 weeks of treatment with the ZenBud, the average PCL-5 score decreased by 20.0 (SD 20.5) points from 38.8.8 (SD 18.0) to 18.8 (SD 18.9) (Figure 7). Similar to results seen for the BAI and BDI, this change in score was both statistically significant (P<.001; 2-tailed dependent t test) and clinically meaningful. A 10-point reduction in score on the PCL-5 is considered clinically meaningful [43,51]. Based on this definition, as seen in Table 2 , 71.4% (20/28) of participants demonstrated a clinically meaningful reduction in score by the end of the trial. In terms of Cohen d, the effect size was large at 0.94.

Figure 7. The progression of PCL-5 scores through 4 weeks of treatment with ZenBud. The thin lines represent each individual participant. The thick line represents the group mean.

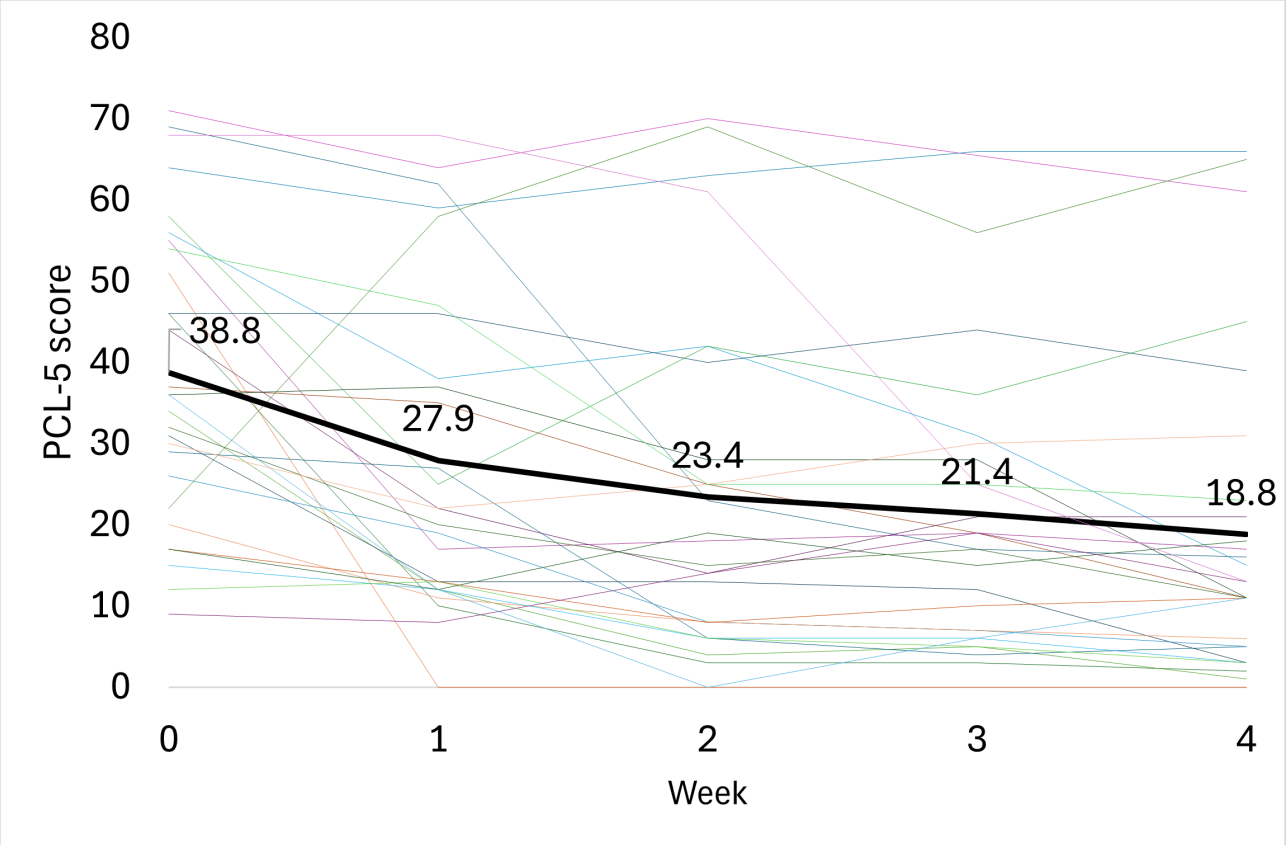


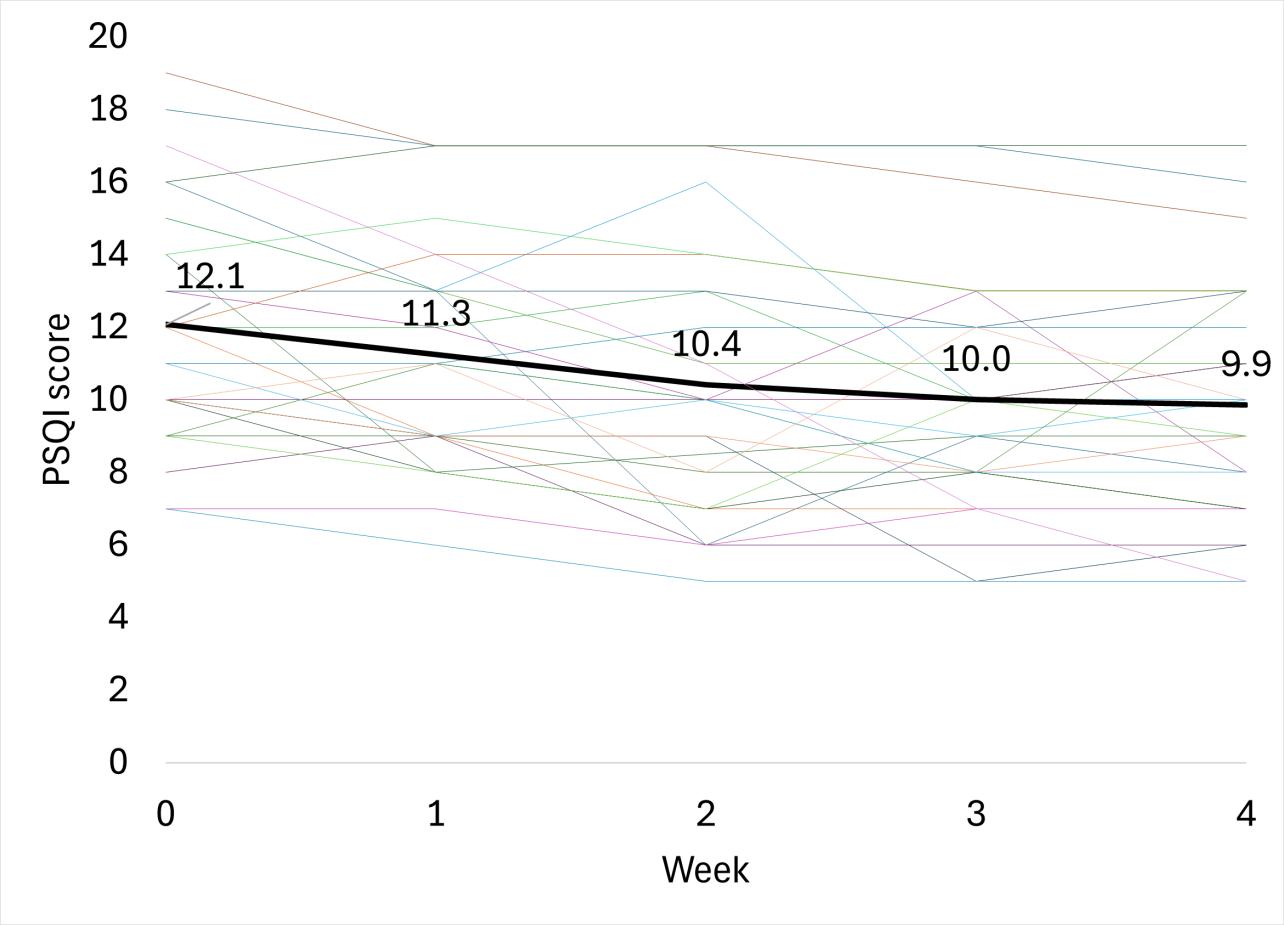
Table . The number of participants who experienced clinically significant reductions in PCL-5 score following 4 weeks of treatment with the ZenBud.

Degree of score change	Participants, n (%)
Clinical decrease	20 (71)
Nonclinical decrease	3 (11)
Nonclinical increase	4 (14)
Clinical increase	1 (4)

Pittsburgh Sleep Quality Index

After 4 weeks of treatment with the ZenBud, the average PSQI score decreased by 2.2 (SD 3.1) points from 12.1 (SD 3.2) to 9.9 (SD 3.2) (Figure 8). While this change in score was statistically significant (P=.001; 2-tailed dependent t test), it was not clinically meaningful. In terms of Cohen d, the effect size was medium at 0.65.

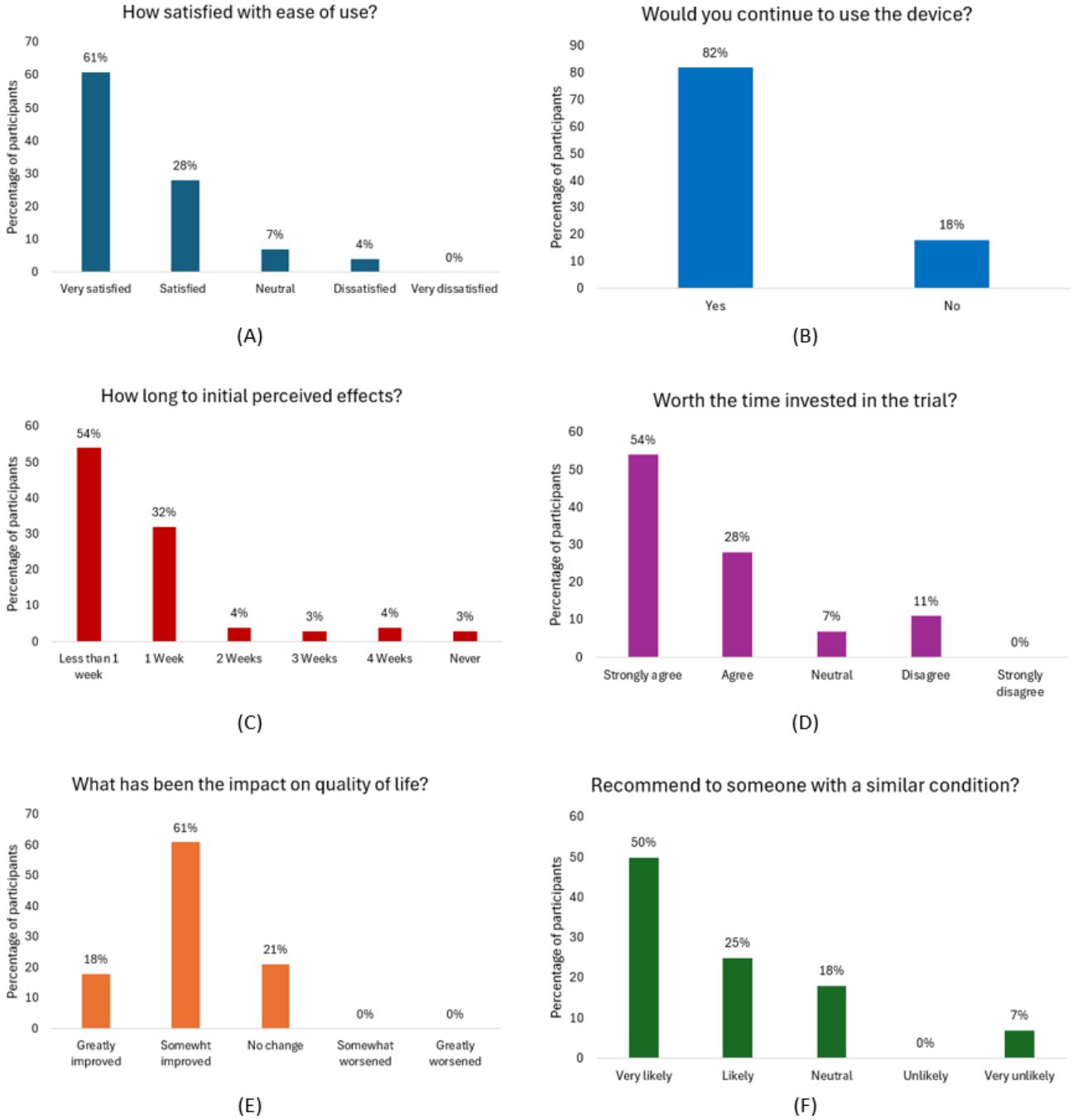
Figure 8. The progression of PSQI scores through 4 weeks of treatment with ZenBud. The thin lines represent each individual participant. The thick line represents the group mean. PSQI: Pittsburgh Sleep Quality Index.



Satisfaction and Acceptability

After the final treatment and assessment, battery participants completed an exit survey asking questions regarding satisfaction with the treatment, acceptability, and quality-of-life impact. When asked about satisfaction with ease of use, 89.3% (25/28) of participants responded with very satisfied or satisfied (Figure 9A). In addition, 82.1% (23/28) of participants reported that they would continue using the device if offered the opportunity (Figure 9B). When asked whether the treatment was worth the time invested in the trial, 82.1% (23/28) of participants strongly agreed or agreed that the time invested was worth it (Figure 9D). When asked about the impact on quality of life, 78.6% (22/28) of participants reported that the treatment somewhat or greatly impacted their quality of life (Figure 9E). When asked how long it took to feel initial effects, 53.6% 15/28) of participants noticed effects in less than 1 week and 32.1% (9/28) felt initial effects by 1 week (Figure 9C). When asked whether they would recommend the treatment to someone with a similar condition, 75.0% (21/28) of participants responded with very likely or likely (Figure 9F).

Figure 9. Results of the exit survey. (A) Responses of the participants when asked "How satisfied were you with the ease of using the device?" (B) Responses of the participants when asked "Would you continue using this device for treatment?" (C) Responses of the participants when asked "How quickly did you feel the effects of the ZenBud device during your trial?" (D) Responses of the participants when asked "Do you feel the device was worth the time invested in the trial?" (E) Responses of the participants when asked "How did the device impact your overall quality of life?" (F) Responses of the participants when asked "How likely are you to recommend this device to others with similar conditions?."



Adverse Events

Only 1 AE was reported throughout the duration of the trial. On the exit survey following completion of the 4 weeks of treatment, 1 participant reported that the treatment would make them feel jittery for a short period of time afterward. This effect was short-lived and classified as a mild AE that was probably device related. The participant reported that this side effect was not enough of an effect to make them stop treatment or drop out of the study. Overall, the high satisfaction rates as described in the "Satisfaction and Acceptability" section combined with the low rate of AE support a strong benefit-to-risk profile for the ZenBud. However, this study was done with a small sample size and these results need to be further validated with a larger sample size.

Discussion

Principal Findings

The main objective of this study was to provide preliminary evidence of the efficacy, safety, and usability of the ZenBud for treating symptoms of anxiety in humans. Overall, the study represents one of the first clinical trials supporting the safety,

patient tolerability, and efficacy of using LIFU to the auricular branch of the vagus nerve for the treatment of anxiety symptoms.

Among the 28 participants, 92.9% (26/28) demonstrated improvements in anxiety symptoms, 89.3% (25/28) demonstrated improvements in depression symptoms, 82.1% (23/28) demonstrated a reduction in symptoms of PTSD, and 65.5% (18/28) demonstrated improvements in sleep quality after 4 weeks of treatment. The average score reduction on the BAI was clinically meaningful at 14.9 points (SD 10.6, $P<.001$; 2-tailed dependent t test), reflecting a general movement from severe anxiety symptoms to mild [35,36]. The average score reduction on the BDI was clinically meaningful at 10.3 points (SD 7.8, $P<.001$; 2-tailed dependent t test), which was a 42.6% decrease in score, far greater than the 17% clinically meaningful threshold [3]. The average score reduction on the PCL-5 was clinically meaningful at 20.0 points (SD 20.5, $P<.001$; 2-tailed dependent t test) [43]. It is also noteworthy to mention that the PCL-5 is commonly used to determine whether an individual meets a provisional diagnosis of PTSD and requires further assessment to confirm the diagnosis. The cutoff score for meeting the criteria for a provisional PTSD diagnosis is 31 - 33. Based on using a cutoff score of 32, at the start of the study 18 participants exceeded the threshold score for a provisional PTSD diagnosis. Upon completion of the study, 14 of these participants (77.8%) had dropped their score below the threshold score of 32 and no longer met the requirements for a provisional PTSD diagnosis. The average score reduction on the PSQI was 2.2 (SD 3.1, $P=.001$; 2-tailed dependent t test) which, while statistically significant, was not clinically meaningful, indicating that the improvements in anxiety, depression, and PTSD symptoms did not carry over into improved sleep quality. The effect sizes were also large for the BAI (Cohen $d=1.06$), BDI (Cohen $d=0.81$), and PCL-5 (Cohen $d=0.94$) indicating that the observed score improvements were substantial enough to have a meaningful impact beyond just statistical significance.

The extent of improvement in anxiety, depression, and PTSD observed in this study is comparable with the clinically meaningful results reported in other clinical trials featuring noninvasive VNS as a treatment intervention. Srinivasan et al [52] conducted a randomized controlled trial of taVNS with 60 retired schoolteachers who had been diagnosed with anxiety during the COVID-19 pandemic. The participants did 30-minute sessions 4 times per week (16 total sessions) and demonstrated significantly greater reductions in Generalized Anxiety Disorder-7 (GAD-7) scores and salivary cortisol levels compared with control group participants. Zhang et al [53] investigated the effect of taVNS on anxiety symptoms and neural functioning in 30 individuals with Parkinson disease and anxiety compared with 30 controls with no anxiety. They treated patients with Parkinson disease with taVNS for 2 weeks and measured progress using the HAM-A and nerve activation in the bilateral prefrontal cortex during a verbal fluency task. After 2 weeks of taVNS treatment, the group demonstrated a significant decrease in HAM-A scores ($P<.001$) and increased activation of the left triangle portion of the inferior frontal gyrus. Ferreira et al [54] treated college students with chronic anxiety with a week of taVNS. Immediately postintervention and 2 weeks postintervention the students demonstrated substantial reductions in pain perception, Beck Anxiety Inventory scores, and masseter activation. Rong et al [55] treated 91 patients with mild to moderate depression with taVNS for 30 minutes twice a day for 12 weeks. Upon completion of treatment the average reduction in score in the 24-item Hamilton Depression Rating Scale (HAM-D-24) was both statistically significant and clinically meaningful, the responder rate was 80%, and the remission rate was 39%. In our study, we saw similar results in only 4 weeks, making an investigation into longer treatment periods with LIFU an important area of future research.

The results of this study are also consistent with the results of studies investigating the use of transcranial focused ultrasound (tfUS) targeting the amygdala for the treatment of generalized anxiety disorders. Mahdavi et al [56] recruited 25 participants with treatment-refractory generalized anxiety disorder and treated them with tfUS targeting the right amygdala for 8 weekly 10-minute sessions. The results showed an average reduction in BAI score of 12.88 (SD 10.42) points and an average reduction in HAM-A scores of −12.64 (SD 12.51). Chou et al [57] recruited 30 healthy individuals and compared activation of the amygdala, hippocampus, and dorsal anterior cingulate cortex during a fear task after treating them with active or sham tfUS targeting the left amygdala. They found decreased activation of the amygdala ($P=.04$), hippocampus ($P=.05$), and dorsal anterior cingulate ($P=.02$) in the active tfUS group when compared with the sham. They also found decreased amygdala-insula ($P=.03$) and amygdala-hippocampal ($P=.01$) resting state functional connectivity and increased amygdala-ventromedial prefrontal cortex ($P=.05$) resting state functional connectivity.

Limitations

While the results of this study are optimistic, this study was preliminary and suffers from several limitations. This study did not feature a control group, making it impossible to quantify the possible impact of a placebo effect or distinguish the specific effects of the ZenBud device from other factors that may have influenced the results. The lack of a control group also limits the ability to directly compare the efficacy of the ZenBud with other interventions. Other than participant reports, there was also no objective way of determining the exact amount of time the device was used by each participant. While the majority of participants were not receiving treatment during the study, there was no control over concurrent therapeutic modalities participants were receiving. The lack of control for these additional therapies may have influenced the results, making it difficult to attribute the observed effects exclusively to the ZenBud device. Further research with larger sample sizes, control groups, control over concurrent treatment modalities, and physiological measurements needs to be done to validate these findings and further negate the possibility of placebo effects.

Conclusions

This preliminary study provided justification for further research into the efficacy, safety, and feasibility of using LIFU to modulate the auricular branch of the vagus nerve and reduce the symptoms of anxiety, depression, and PTSD. Given the wide prevalence of anxiety disorders, depression, and PTSD,

and the shortfalls of current treatment options, this novel treatment approach has potential to meaningfully improve patient outcomes and continued research is warranted.

Conflicts of Interest

IK is the chief science officer for NeurGear. JH is the chief executive officer for NeurGear. EM declares no conflicts of interest.

Multimedia Appendix 1
ZenBud user manual.
[PDF File, 1446 KB - neuro_v4i1e69770_app1.pdf]

References

1. Penninx BW, Pine DS, Holmes EA, Reif A. Anxiety disorders. Lancet 2021 Mar 6;397(10277):914-927. [doi: 10.1016/S0140-6736(21)00359-7] [Medline: 33581801]
2. Chand SP, Marwaha R. StatPearls. Anxiety.: Treasure Island, FL: StatPearls; 2023 Apr 24. URL: https://www.ncbi.nlm.nih.gov/books/NBK470361 [accessed 2025-04-04]
3. Button KS, Kounali D, Thomas L, et al. Minimal clinically important difference on the Beck Depression Inventory—II according to the patient's perspective. Psychol Med 2015 Nov;45(15):3269-3279. [doi: 10.1017/S0033291715001270] [Medline: 26165748]
4. Tasker JG, Joëls M. The synaptic physiology of the central nervous system response to stress. In: Russell J, Shipston M, editors. Neuroendocrinology of Stress: John Wiley & Sons, Ltd; 2015:43-70. [doi: 10.1002/9781118921692]
5. Holwerda SW, Luehrs RE, Gremaud AL, et al. Relative burst amplitude of muscle sympathetic nerve activity is an indicator of altered sympathetic outflow in chronic anxiety. J Neurophysiol 2018 Jul 1;120(1):11-22. [doi: 10.1152/jn.00064.2018] [Medline: 29537916]
6. Giulio P. Maladaptive stress: theoretical, neurobiological and clinical profiles. Arch Depress Anxiety 2021;7(1):001-007. [doi: 10.17352/2455-5460.000057]
7. Ströhle A, Gensichen J, Domschke K. The diagnosis and treatment of anxiety disorders. Dtsch Arztebl Int 2018 Sep 14;155(37):611-620. [doi: 10.3238/arztebl.2018.0611] [Medline: 30282583]
8. Stockings EA, Degenhardt L, Dobbins T, et al. Preventing depression and anxiety in young people: a review of the joint efficacy of universal, selective and indicated prevention. Psychol Med 2016 Jan;46(1):11-26. [doi: 10.1017/S0033291715001725] [Medline: 26315536]
9. Fischer R, Bortolini T, Karl JA, et al. Rapid review and meta-meta-analysis of self-guided interventions to address anxiety, depression, and stress during COVID-19 social distancing. Front Psychol 2020;11:563876. [doi: 10.3389/fpsyg.2020.563876] [Medline: 33192837]
10. Haller H, Breilmann P, Schröter M, Dobos G, Cramer H. A systematic review and meta-analysis of acceptance- and mindfulness-based interventions for DSM-5 anxiety disorders. Sci Rep 2021 Oct 14;11(1):20385. [doi: 10.1038/s41598-021-99882-w] [Medline: 34650179]
11. Chi D, Zhang Y, Zhou D, Xu G, Bian G. The effectiveness and associated factors of online psychotherapy on COVID-19 related distress: a systematic review and meta-analysis. Front Psychol 2022;13:1045400. [doi: 10.3389/fpsyg.2022.1045400] [Medline: 36438323]
12. Moghimi E, Stephenson C, Agarwal A, et al. Efficacy of an electronic cognitive behavioral therapy program delivered via the online psychotherapy tool for depression and anxiety related to the COVID-19 pandemic: pre-post pilot study. JMIR Ment Health 2023 Dec 25;10:e51102. [doi: 10.2196/51102] [Medline: 37993984]
13. Garakani A, Murrough JW, Freire RC, et al. Pharmacotherapy of anxiety disorders: current and emerging treatment options. Front Psychiatry 2020;11:595584. [doi: 10.3389/fpsyt.2020.595584] [Medline: 33424664]
14. Baek H, Pahk KJ, Kim H. A review of low-intensity focused ultrasound for neuromodulation. Biomed Eng Lett 2017 May;7(2):135-142. [doi: 10.1007/s13534-016-0007-y] [Medline: 30603160]
15. Collins MN, Mesce KA. A review of the bioeffects of low-intensity focused ultrasound and the benefits of A cellular approach. Front Physiol 2022;13:1047324. [doi: 10.3389/fphys.2022.1047324] [Medline: 36439246]
16. Feng B, Chen L, Ilham SJ. A review on ultrasonic neuromodulation of the peripheral nervous system: enhanced or suppressed activities? Appl Sci (Basel) 2019 Apr 2;9(8):1637. [doi: 10.3390/app9081637] [Medline: 34113463]
17. Johnson RL, Wilson CG. A review of vagus nerve stimulation as A therapeutic intervention. J Inflamm Res 2018;11:203-213. [doi: 10.2147/JIR.S163248] [Medline: 29844694]
18. Martin EI, Ressler KJ, Binder E, Nemeroff CB. The neurobiology of anxiety disorders: brain imaging, genetics, and psychoneuroendocrinology. Psychiatr Clin North Am 2009 Sep;32(3):549-575. [doi: 10.1016/j.psc.2009.05.004] [Medline: 19716990]
19. Kaniusas E, Kampusch S, Tittgemeyer M, et al. Current directions in the auricular vagus nerve stimulation I—a physiological perspective. Front Neurosci 2019;13:854. [doi: 10.3389/fnins.2019.00854] [Medline: 31447643]

20. Butt MF, Albusoda A, Farmer AD, Aziz Q. The anatomical basis for transcutaneous auricular vagus nerve stimulation. J Anat 2020 Apr;236(4):588-611. [doi: 10.1111/joa.13122] [Medline: 31742681]
21. Breit S, Kupferberg A, Rogler G, Hasler G. Vagus nerve as modulator of the brain-gut axis in psychiatric and inflammatory disorders. Front Psychiatry 2018;9:44. [doi: 10.3389/fpsyt.2018.00044] [Medline: 29593576]
22. Noble LJ, Gonzalez IJ, Meruva VB, et al. Effects of vagus nerve stimulation on extinction of conditioned fear and post-traumatic stress disorder symptoms in rats. Transl Psychiatry 2017 Aug 22;7(8):e1217. [doi: 10.1038/tp.2017.191] [Medline: 28892066]
23. George MS, Ward HE Jr, Ninan PT, et al. A pilot study of vagus nerve stimulation (VNS) for treatment-resistant anxiety disorders. Brain Stimul 2008 Apr;1(2):112-121. [doi: 10.1016/j.brs.2008.02.001] [Medline: 20633378]
24. Gurel NZ, Wittbrodt MT, Jung H, et al. Transcutaneous cervical vagal nerve stimulation reduces sympathetic responses to stress in posttraumatic stress disorder: a double-blind, randomized, sham controlled trial. Neurobiol Stress 2020 Nov;13:100264. [doi: 10.1016/j.ynstr.2020.100264] [Medline: 33344717]
25. Badran BW, Huffman SM, Dancy M, et al. A pilot randomized controlled trial of supervised, at-home, self-administered transcutaneous auricular vagus nerve stimulation (taVNS) to manage long COVID symptoms. Res Sq 2022 Jun 21:rs.3.rs-1716096. [doi: 10.21203/rs.3.rs-1716096/v1] [Medline: 35765566]
26. Wittbrodt MT, Gurel NZ, Nye JA, et al. Noninvasive cervical vagal nerve stimulation alters brain activity during traumatic stress in individuals with posttraumatic stress disorder. Psychosom Med 2021;83(9):969-977. [doi: 10.1097/PSY.0000000000000987] [Medline: 34292205]
27. Lamb DG, Porges EC, Lewis GF, Williamson JB. Non-invasive vagal nerve stimulation effects on hyperarousal and autonomic state in patients with posttraumatic stress disorder and history of mild traumatic brain injury: preliminary evidence. Front Med (Lausanne) 2017;4:124. [doi: 10.3389/fmed.2017.00124] [Medline: 28824913]
28. Bremner J, Gurel N, Wittbrodt M, et al. Non-invasive vagal nerve stimulation paired with stress exposure in posttraumatic stress disorder (PTSD). Brain Stimul 2019 Mar;12(2):438. [doi: 10.1016/j.brs.2018.12.417]
29. Riis T, Kubanek J. Effective ultrasonic stimulation in human peripheral nervous system. IEEE Trans Biomed Eng 2022 Jan;69(1):15-22. [doi: 10.1109/TBME.2021.3085170] [Medline: 34057888]
30. Juan EJ, González R, Albors G, Ward MP, Irazoqui P. Vagus nerve modulation using focused pulsed ultrasound: potential applications and preliminary observations in a rat. Int J Imaging Syst Technol 2014 Mar 1;24(1):67-71. [doi: 10.1002/ima.22080] [Medline: 25165410]
31. Cotero V, Fan Y, Tsaava T, et al. Noninvasive sub-organ ultrasound stimulation for targeted neuromodulation. Nat Commun 2019 Mar 12;10(1):952. [doi: 10.1038/s41467-019-08750-9] [Medline: 30862827]
32. Bentley KH, Bernstein EE, Wallace B, Mischoulon D. Treatment for anxiety and comorbid depressive disorders: transdiagnostic cognitive-behavioral strategies. Psychiatr Ann 2021 Aug;51(5):226-230. [doi: 10.3928/00485713-20210414-01] [Medline: 34433988]
33. Williamson JB, Jaffee MS, Jorge RE. Posttraumatic stress disorder and anxiety-related conditions. Continuum (Minneap Minn) 2021 Dec 1;27(6):1738-1763. [doi: 10.1212/CON.0000000000001054] [Medline: 34881734]
34. Fydrich T, Dowdall D, Chambless DL. Reliability and validity of the Beck Anxiety Inventory. J Anxiety Disord 1992 Jan;6(1):55-61. [doi: 10.1016/0887-6185(92)90026-4]
35. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]
36. Beck AT, Steer RA. Manual for the Beck Anxiety Inventory: The Psychological Corporation; 1990.
37. Bardhoshi G, Duncan K, Erford BT. Psychometric meta‐analysis of the English version of the Beck Anxiety Inventory. Jour of Counseling & Develop 2016 Jul;94(3):356-373 [FREE Full text] [doi: 10.1002/jcad.12090]
38. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev 1988 Jan;8(1):77-100. [doi: 10.1016/0272-7358(88)90050-5]
39. Maletic V, Robinson M, Oakes T, Iyengar S, Ball SG, Russell J. Neurobiology of depression: an integrated view of key findings. Int J Clin Pract 2007 Dec;61(12):2030-2040. [doi: 10.1111/j.1742-1241.2007.01602.x] [Medline: 17944926]
40. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry 1961 Jun;4:561-571. [doi: 10.1001/archpsyc.1961.01710120031004] [Medline: 13688369]
41. Beck AT, Steer RA, Ball R, Ranieri WF. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. J Pers Assess 1996 Dec;67(3):588-597. [doi: 10.1207/s15327752jpa6703_13] [Medline: 8991972]
42. Wang YP, Gorenstein C. Psychometric properties of the Beck Depression Inventory—II: a comprehensive review. Braz J Psychiatry 2013;35(4):416-431. [doi: 10.1590/1516-4446-2012-1048] [Medline: 24402217]
43. Blanchard BE, Johnson M, Campbell SB, et al. Minimal important difference metrics and test-retest reliability of the PTSD Checklist for DSM-5 with a primary care sample. J Trauma Stress 2023 Dec;36(6):1102-1114. [doi: 10.1002/jts.22975] [Medline: 37845820]
44. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]

45. Forkus SR, Raudales AM, Rafiuddin HS, Weiss NH, Messman BA, Contractor AA. The Posttraumatic Stress Disorder (PTSD) Checklist for DSM-5: a systematic review of existing psychometric evidence. Clin Psychol (New York) 2023 Mar;30(1):110-121. [doi: 10.1037/cps0000111] [Medline: 37378352]
46. Cox RC, Olatunji BO. A systematic review of sleep disturbance in anxiety and related disorders. J Anxiety Disord 2016 Jan;37:104-129. [doi: 10.1016/j.janxdis.2015.12.001] [Medline: 26745517]
47. K.p S, Joseph N, Sreenivasan GK. Sleep quality and anxiety level among college students. Int J Res Rev 2021;8(10):271-276. [doi: 10.52403/ijrr.20211036]
48. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
49. Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. The Pittsburgh Sleep Quality Index as a screening tool for sleep dysfunction in clinical and non-clinical samples: a systematic review and meta-analysis. Sleep Med Rev 2016 Feb;25:52-73. [doi: 10.1016/j.smrv.2015.01.009] [Medline: 26163057]
50. Fabbri M, Beracci A, Martoni M, Meneo D, Tonetti L, Natale V. Measuring subjective sleep quality: a review. Int J Environ Res Public Health 2021 Jan 26;18(3):1082. [doi: 10.3390/ijerph18031082] [Medline: 33530453]
51. Held P, Smith DL, Bagley JM, et al. Treatment response trajectories in a three-week CPT-based intensive treatment for veterans with PTSD. J Psychiatr Res 2021 Sep;141:226-232. [doi: 10.1016/j.jpsychires.2021.07.004] [Medline: 34246978]
52. Srinivasan V, Ruthuvalan V, Raja S, et al. Efficacy of vagal nerve stimulation on anxiety among elderly retired teachers during COVID-19 pandemic. Work 2024;79(2):645-652. [doi: 10.3233/WOR-230356] [Medline: 38489208]
53. Zhang H, Shan AD, Wan CH, et al. Transcutaneous auricular vagus nerve stimulation improves anxiety symptoms and cortical activity during verbal fluency task in Parkinson's disease with anxiety. J Affect Disord 2024 Sep 15;361:556-563. [doi: 10.1016/j.jad.2024.06.083] [Medline: 38925314]
54. Ferreira LMA, Brites R, Fraião G, et al. Transcutaneous auricular vagus nerve stimulation modulates masseter muscle activity, pain perception, and anxiety levels in university students: a double-blind, randomized, controlled clinical trial. Front Integr Neurosci 2024;18:1422312. [doi: 10.3389/fnint.2024.1422312] [Medline: 39051059]
55. Rong P, Liu J, Wang L, et al. Effect of transcutaneous auricular vagus nerve stimulation on major depressive disorder: a nonrandomized controlled pilot study. J Affect Disord 2016 May;195:172-179. [doi: 10.1016/j.jad.2016.02.031] [Medline: 26896810]
56. Mahdavi KD, Jordan SE, Jordan KG, et al. A pilot study of low-intensity focused ultrasound for treatment-resistant generalized anxiety disorder. J Psychiatr Res 2023 Oct 19;168:125-132. [doi: 10.1016/j.jpsychires.2023.10.039] [Medline: 39491902]
57. Chou T, Deckersbach T, Guerin B, et al. Transcranial focused ultrasound of the amygdala modulates fear network activation and connectivity. Brain Stimul 2024;17(2):312-320. [doi: 10.1016/j.brs.2024.03.004] [Medline: 38447773]

Abbreviations

AE: adverse event
BAI: Beck Anxiety Inventory
BDI: Beck Depression Inventory
***DSM-V*:** *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition)
HAM-A: Hamilton Anxiety Rating Scale
LIFU: low-intensity focused ultrasound
PSQI: Pittsburgh Sleep Quality Index
PTSD: posttraumatic stress disorder
SNS: sympathetic nervous system
taVNS: transcutaneous auricular vagal nerve stimulation
tfUS: transcranial focused ultrasound
VNS: vagus nerve stimulation

Edited by P Kubben; submitted 07.12.24; peer-reviewed by MA Hefny, R Galo; revised version received 07.03.25; accepted 12.03.25; published 01.05.25.

Please cite as:
Kohler I, Hacker J, Martin E
Reduction of Anxiety-Related Symptoms Using Low-Intensity Ultrasound Neuromodulation on the Auricular Branch of the Vagus Nerve: Preliminary Study
JMIR Neurotech 2025;4:e69770
URL: https://neuro.jmir.org/2025/1/e69770
doi:10.2196/69770

Exploring Remote Monitoring of Poststroke Mood With Digital Sensors by Assessment of Depression Phenotypes and Accelerometer Data in UK Biobank: Cross-Sectional Analysis

Stephanie J Zawada[1], PhD, MS; Ali Ganjizadeh[2], MD; Gian Marco Conte[2], MD, PhD; Bart M Demaerschalk[1,3], MD, MSc; Bradley J Erickson[2,4], MD, PhD

[1]Mayo Clinic College of Medicine and Science, 5777 E. Mayo Boulevard, Phoenix, AZ, United States

[2]AI Laboratory, Department of Radiology, Mayo Clinic, Rochester, MN, United States

[3]Department of Neurology, Division of Cerebrovascular Diseases, Mayo Clinic, Phoenix, AZ, United States

[4]Mayo Clinic College of Medicine and Science, Rochester, MN, United States

Corresponding Author:
Stephanie J Zawada, PhD, MS
Mayo Clinic College of Medicine and Science, 5777 E. Mayo Boulevard, Phoenix, AZ, United States

Abstract

Background: Interest in using digital sensors to monitor patients with prior stroke for depression, a risk factor for poor outcomes, has grown rapidly; however, little is known about behavioral phenotypes related to future mood symptoms and if patients with and without previously diagnosed depression experience similar phenotypes.

Objective: This study aimed to assess the feasibility of using digital sensors to monitor mood in patients with prior stroke with a prestroke depression diagnosis (DD) and controls. We examined relationships between physical activity behaviors and self-reported depression frequency.

Methods: In the UK Biobank wearable accelerometer cohort, we retrospectively identified patients who had previously suffered a stroke (N=1603) and conducted cross-sectional analyses with those who completed a subsequent depression survey follow-up. Sensitivity analyses assessed a general population cohort excluding previous stroke participants and 2 incident cohorts: incident stroke (IS) and incident cerebrovascular disease (IC).

Results: In controls, the odds of being in a higher depressed mood frequency category decreased by 23% for each minute spent in moderate - to - vigorous physical activity (odds ratio 0.77, 95% CI 0.69 - 0.87; *P*<.001). This association persisted in both general cohorts and in the IC control cohort.

Conclusions: Although moderate - to - vigorous physical activity was linked with less frequent depressed mood in patients with prior stroke without DD, this finding did not persist in DDs. Thus, accelerometer-mood monitoring may provide clinically useful insights about future mood in patients with prior stroke without DDs. Considering the finding in the IC cohort and the lack of findings in the IS cohorts, accelerometer-mood monitoring may also be appropriately applied to observing broader cerebrovascular disease pathogenesis.

(JMIR Neurotech 2025;4:e56679) doi:10.2196/56679

KEYWORDS

depression; cerebrovascular disease; remote monitoring; stroke; accelerometers; mobile phone

Introduction

Overview

Depression is an established risk factor for poor outcomes after a stroke and transient ischemic attack (TIA), including subsequent stroke and other cerebrovascular diseases (CeVDs) [1,2]. Although poststroke depression (PSD) affects roughly one-third of patients with stroke, screening for depression in patients after a stroke is not routine, with less than 10% of patients with stroke screened [3]. Furthermore, it remains unclear when follow-up PSD screening should occur, as current research suggests that not all patients will experience PSD symptoms immediately after a stroke and, for those who do, the majority will experience recurrent depression episodes in the years after a stroke [4]. A reason for this gap in screening is the shortage of neurologists, particularly those with diagnostic training in identifying PSD [5]. Accelerated by the widespread adoption of personal mobile devices, from computers to smartwatches, it is critical to investigate the potential of such devices to collect meaningful data outside of clinical settings, aiding clinicians in identifying depressed mood in patients with

stroke—and, potentially, those most at risk for subsequent stroke and CeVDs [6].

Background

The prevalence of PSD remains unknown, partly due to its heterogeneous nature, spanning unique somatic, behavioral, cognitive, motivational, and emotional components [7]. The severity of its manifestation ranges from mild symptoms to clinical-grade depression, the former of which relies on self-reported scoring methods inherently subject to bias, especially in patients with cognitive impairment for whom self-reported surveys may not be reliable [8]. Although clinician-administered assessments, like the Montgomery–Åsberg Depression Rating Survey (MADRS), offer gold-standard assessments of symptoms, nurse and physician shortages complicate the routine administration of such instruments [9].

In some survivors, depression may emerge alongside the incipient pathogenesis of cerebrovascular dysfunction, while for others, depression may be a reaction to being conscious of cognitive impairment or the putative manifestation of silent cerebral infarcts [10,11]. As such, individual depression phenotypes may vary greatly across survivors with identical survey summary scores. Although investigations into the associations between stroke location within the brain and self-reported depression survey scores have yielded inconclusive results, a recent cross-sectional study of patients with prior stroke (n=200) found that symptoms assessed by MADRS correlate with specific macrostructural characteristics [12]. Considering that clinician-administered assessments, like MADRS, are more accurate than self-reported survey scores in patients with prior stroke, the need for a modified approach to monitoring patients with stroke for depression emerges.

In recent years, objective data from portable and wearable sensors have demonstrated the feasibility of augmenting self-reported mood surveys outside of clinics, a promising approach for monitoring patients with symptomatic and asymptomatic deteriorating brain health outside of standardized, clinical environments [13-20]. In addition, accelerometer measures of behavior have established a difference in PA engagement stratified by depression severity, highlighting the need for a thoughtful approach to PSD screening and monitoring that ensures patients with emerging or mild depression symptoms, unlike those with previous documented depressive episodes, are not neglected [21].

While triaging patients with PSD for preventative intervention could yield clinically meaningful functional recovery outcomes, the potential of such an approach for preventing future CeVD diagnoses remains to be seen. Numerous studies have found that depressive symptoms are associated with an increased risk of subsequent CeVD, from acute CeVDs, like stroke and TIA, to more chronic conditions, like cerebral arterial stenosis and vascular dementia [11,22-24]. Furthermore, recent research suggests daily functioning and cognitive changes may be observable up to 10 years before some types of CeVD [25]. Thus, particular attention should be paid to behavioral patterns in patients with PSD to elucidate phenotypes with predictive potential for functional outcomes and neurologic disorders.

Previous Work

Blending self-reported assessments of phenomena, like mood, recorded through web browsers and smartphone apps, with passive sensor data, like that from wearable accelerometers, is gaining popularity in real-world settings [26,27]. Numerous pilot studies have demonstrated the potential for wearable and minimally invasive sensors to detect neurologic conditions; however, these tools have neither been validated in population cohorts nor combined with survey sampling of mood [28].

Early-stage evidence suggests that monitoring lifestyle behavior and mood in PSD is feasible [29-31]. The results of a small longitudinal study (n=40) suggest that self-reported moderate-to-vigorous physical activity (MVPA) before stroke is associated with improved mobility and self-care as well as decreased discomfort after stroke [32]. While the study did not sample mood outside of clinical environments, Reinholdsson et al [33] used self-report surveys to expand on the above findings, demonstrating that patients who engage in higher levels of prestroke physical activity (PA) experienced less severe PSD compared with patients who were physically inactive .

In addition, current literature on accelerometers in PSD suggests that distinct behavioral patterns may identify patients with depression within the first year after a stroke. In a 2022 prospective observational study of recently discharged patients with minor ischemic stroke (n=76), participants wore accelerometers in-hospital for 1 week. Analyses revealed that only increased sedentary behavior (SB) and reduced light physical activity (LPA) were linked with more intense depression, assessed through a written Geriatric Depression Scale survey, 3 months after hospitalization in this older adult cohort [34]. In a small pilot study (n=40) of stroke survivors, MVPA was linked with positive mood [35]. Although extensive research has confirmed links between sleep disorders and both depression and incident CeVD (IC), no research has observed both depressive symptoms and objectively measured sleep after stroke [36,37]. Furthermore, no previous accelerometer research into PSD beyond the first year of stroke recovery has been published.

Goal of This Study

The goal of this study is twofold: first, to investigate potential associations between objectively measured behavior and future depression frequency in patients with prior stroke assessed by a remote approach and second, to explore whether that association varies between patients with prior stroke with a prestroke depression diagnosis (DDs) and those without (controls).

We conducted a cross-sectional analysis with the UK Biobank (UKBB), the most extensive lifestyle and mood cohort to date, assessing the relationships between accelerometer-measured sleep, SB, LPA, and MVPA and a subsequent depression descriptor (depressed mood frequency). Given that depression before stroke may yield behavioral phenotypes distinct from those emergent in participants without a prestroke depression diagnosis, we created 2 cohorts of patients with prior stroke: those with a clinical depression diagnosis before stroke and those without. As this analysis focuses on participants who may

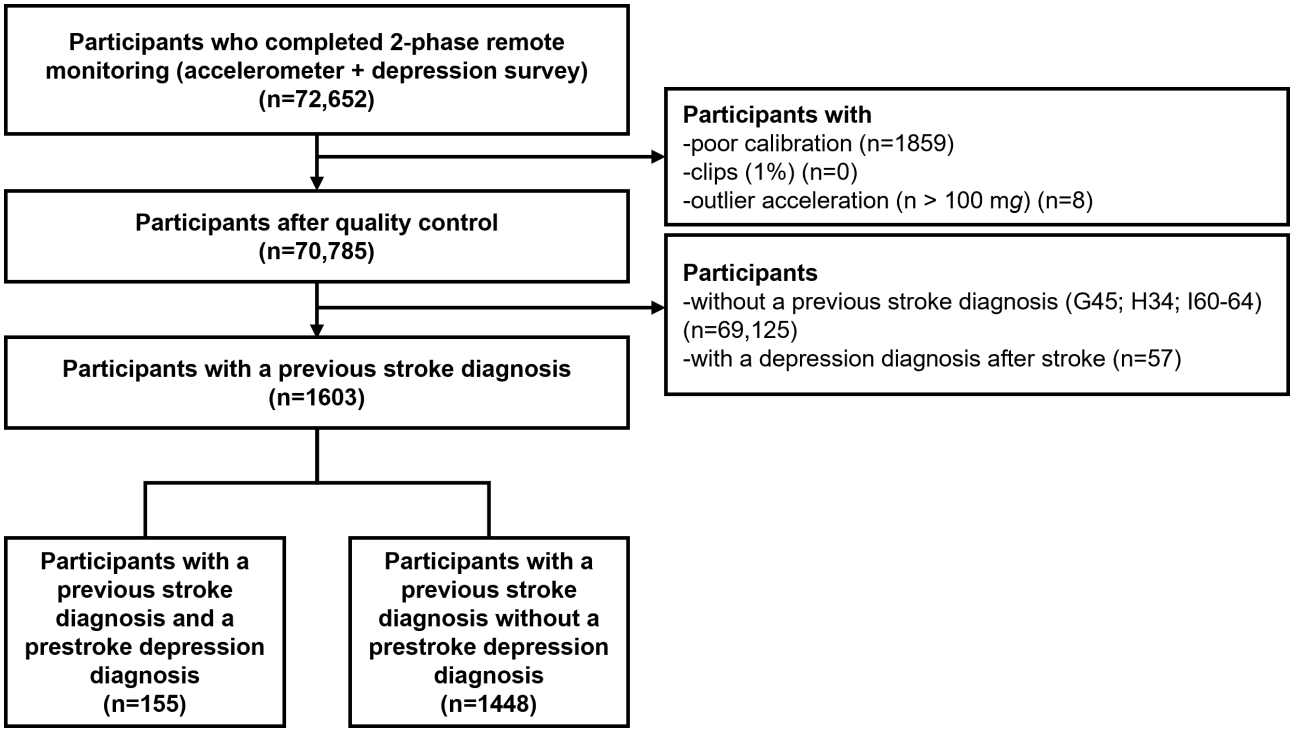
develop or have undiagnosed PSD, participants whose PSD diagnosis was recorded were excluded. Adjusting for age, sex, ethnicity, multiple relevant comorbidities, and time elapsed between accelerometer monitoring and depression survey submission, we hypothesized that increased LPA and MVPA time would be associated with a reduction in the odds of being in a more frequent depressed mood category while increased SB time would be associated with a rise in the odds of being in a more frequent depressed mood category. Considering the established relationship between sleep and depressed mood, we created a binary variable (yes or no) for guideline-recommended sleep (7 - 9 h/d). We hypothesized that guideline-recommended sleep would be associated with a reduction in the odds of being in a more frequent depressed mood category.

Methods

Recruitment

The UKBB enrolled middle-aged (40 - 69 y) participants (N=502,364) at 22 assessment centers across the United Kingdom at a baseline assessment (2006 - 2010), which included in-person interviews, touchscreen surveys, and physical examinations to extract lifestyle and environmental data used in this study. Although all baseline participants (n=502,151) were invited, only 72,652 enrolled in the 1-week accelerometer study (2013 - 2015) and completed the depression frequency survey (2016 - 2017). Hospital and other diagnostic registries were linked to enrolled participants.

Figure 1. Classification algorithm for participant cohorts.



Data Collection

Accelerometer study participants were instructed to wear the Axivity AX3 commercial accelerometer wristwatch continuously on their dominant arm for 1 week. The depressed mood frequency question was administered through a link accessible on smartphone, tablet or PC browsers as part of the standardized Patient Health Questionnaire-2 (PHQ-2) survey: “Over the past two weeks, how often have you felt down, depressed, or hopeless?” Responses were ordinal scores indicating the

Participant Cohorts

Among participants who completed both remote monitoring components, those with dementia (n=23) were excluded. Quality control filtering demonstrated by Madjedi et al [38] was applied (n=70,785), which excluded those with outlier acceleration (>100 mg), more than 1% of readings exceeding ±8 *g* (clips), accelerometer wear time less than 3 days, and missing data for at least one 60-minute interval throughout 24-hour periods. Only participants with a previous stroke, including ischemic stroke, hemorrhagic stroke, and TIA (G45), were included (n=1660). Retinal artery occlusion (H34) was included as a stroke, as it is now considered a type of acute ischemic stroke [39]. Participants who were diagnosed with depression after stroke but before the accelerometer study (n=57) were excluded.

Among those meeting the inclusion criteria (n=1603), participants were divided into two cohorts: (1) those with a prestroke depression diagnosis at accelerometer study commencement (n=155) and (2) controls, that is, those without a prestroke depression diagnosis (n=1448) (Figure 1). No participants were diagnosed with depression between the accelerometer study and the follow-up depression survey.

Participants with a history of depression (*International Classification of Diseases, Tenth Revision* [*ICD-10*] codes F32-39) comprised the depression diagnosis (DDs) cohort. Definitions (*ICD-10* codes) used for inclusion and exclusion criteria as well as diagnostic classification are available in Multimedia Appendix 1.

frequency of depressed mood, with 1="Not at all"; 2="Several days"; 3="More than half of days"; and 4="Nearly every day."

Permanent covariables were obtained at baseline visit, including sex and ethnicity. For each participant, age at the time of accelerometer study was calculated. Time-to-assessment was individually calculated by subtracting the accelerometer start date from the date of submitting the depressed mood survey. Comorbidity diagnoses before the accelerometer study were obtained from linked patient and hospital databases.

Statistical Analysis

To compare continuous and categorical covariables, the Mann-Whitney U test and χ^2 test, respectively, were used. A cross-sectional analysis using ordinal logistic regression to investigate the association between objective behavior predictors and the ordinal outcome variable, depressed mood frequency over the past 2 weeks, was conducted on data obtained at the accelerometer study and remote follow-up survey.

For both DD and control cohorts, separate models were fitted to evaluate whether the role of objective behavior predictors in depressed mood frequency differed between cohorts.

Analyses were performed in R (R Foundation for Statistical Computing), using *polr* from the library MASS. The effect sizes of objective behavior predictors, adjusted for confounders, on depressed mood frequency were plotted as odds ratios with 95% CIs. The Likelihood Ratio Test was used to obtain all *P* values and associated CIs. $P<.05$ was statistically significant.

Sensitivity Analysis

Three sensitivity analyses (also using ordinal logistic regression models), each considering DDs and controls, were performed using UKBB data. First, a general population dataset wasf generated. This included all participants eligible for inclusion in the accelerometer study and follow-up depression frequency survey who did not have a previous stroke diagnosis.

Next, participants with an initial IC diagnosis (after the depression frequency survey) were filtered into a separate dataset. Ordinal logistic regression models were fitted to assess the relationships between objective behavior predictors and depressed mood frequency. Finally, participants in the IC cohort who had an IS diagnosis were filtered into a separate dataset, and ordinal logistic regression models were fitted to assess the target relationship. The investigation of IC as a composite end point reflects updated understanding of stroke as sharing etiology with other neurologic rather than circulatory system disorders, as defined in the most recent *International Classification of Diseases, Eleventh Revision* (*ICD-11*) [37].

For each filtered cohort, sample characteristics were obtained for review.

Ethical Considerations

National Health Service Research Ethics Committee (11/NW/0382) granted ethical approval for the UKBB population cohort study. Informed consent was obtained from all UK Biobank participants under National Health Service National Research Ethics Service (Ref 11/NW/0382). All UKBB data are deidentified.

Results

Study Characteristics

For participants in the 2-stage remote monitoring study (Table 1), the DDs had a higher proportion of women compared with controls (58.7% vs 40.7%). On average, DDs were younger (64 vs 66 y), slept slightly longer (9.2 vs 9.0 h/d), spent slightly less time in MVPA (29.3 vs 37.3 min/d) and SB (580.1 vs 583.6 min/d), and spent slightly more time in LPA (281.4 vs 278.2 min/d).

Table . Baseline characteristics of patients with previous stroke.

		Prestroke depression	Controls	*P* value
Number of participants, n		155	1448	
Age, mean (SD)		64 (7)	66 (6.5)	<.001
Gender, n (%)				
	Men	64 (41.3)	859 (59.3)	<.001
Race, n (%)				
	White	153 (98.7)	1418 (97.9)	.68
Sleep, mean (SD)		9.2 (1.8)	9.0 (1.8)	<.001
Sleep (7 - 9 h/d), n (%)		71 (45.8)	736 (50.8)	.27
SB[a], mean (SD)		580.1 (114.4)	583.6 (112.8)	<.001
LPA[b], mean (SD)		281.4 (106.9)	278.2 (102.4)	<.001
MVPA[c], mean (SD)		29.3 (31.2)	37.3 (33.0)	<.001
Time-to-assessment, mean (SD)		1.8 (0.7)	1.8 (0.6)	<.001
Diabetes, n (%)		21 (13.5)	130 (9.0)	.09
Hyperlipidemia, n (%)		76 (49)	615 (42.5)	.14
Hypertension, n (%)		155 (100)	1448 (100)	1
Multiple strokes, n (%)		40 (25.8)	348 (24.0)	.70
Time since most recent stroke, mean (SD)		7.8 (6.4)	9.8 (8.8)	<.001

[a]SB: sedentary behavior.

[b]LPA: light physical activity.

[c]MVPA: moderate-to-vigorous physical activity.

All participants had a hypertension diagnosis. The average time between accelerometer study start and depressed mood survey submission (time-to-assessment) was 1.8 years for both cohorts.

The average time from the initial stroke to the accelerometer study commencement was less for DDs than controls (7.8 vs 9.8 y).

Among DDs, 9 participants slept less than 7 hours while 75 slept more than 9 hours. In the control group, 79 participants slept less than 7 hours while 633 slept more than 9 hours.

Cross-Sectional Analysis

No significant association persisted in both the DD and control cohorts (Table 2). In controls, for each minute spent in MVPA per day, the odds of being in a higher depressed mood frequency category decreased by 23% (*P*<.001).

Table . Ordinal logistic regression assessing objective behavior predictors and depressed mood frequency.

Previous stroke participants	Prestroke depression		Controls	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value
Sleep (7 - 9 hr/d)	0.49 (0.23 - 1.03)	.06	0.88 (0.66 - 1.19)	.41
SB[b] (min/d)	1.00 (1.00 - 1.01)	.10	1.00 (1.00 - 1.00)	.63
LPA[c] (min/d)	1.00 (0.99 - 1.00)	.20	1.00 (1.00 - 1.00)	.35
MVPA[d] (min/d)	0.86 (0.64 - 1.17)	.33	0.77 (0.69 - 0.87)	<.001

[a]OR: odds ratio.

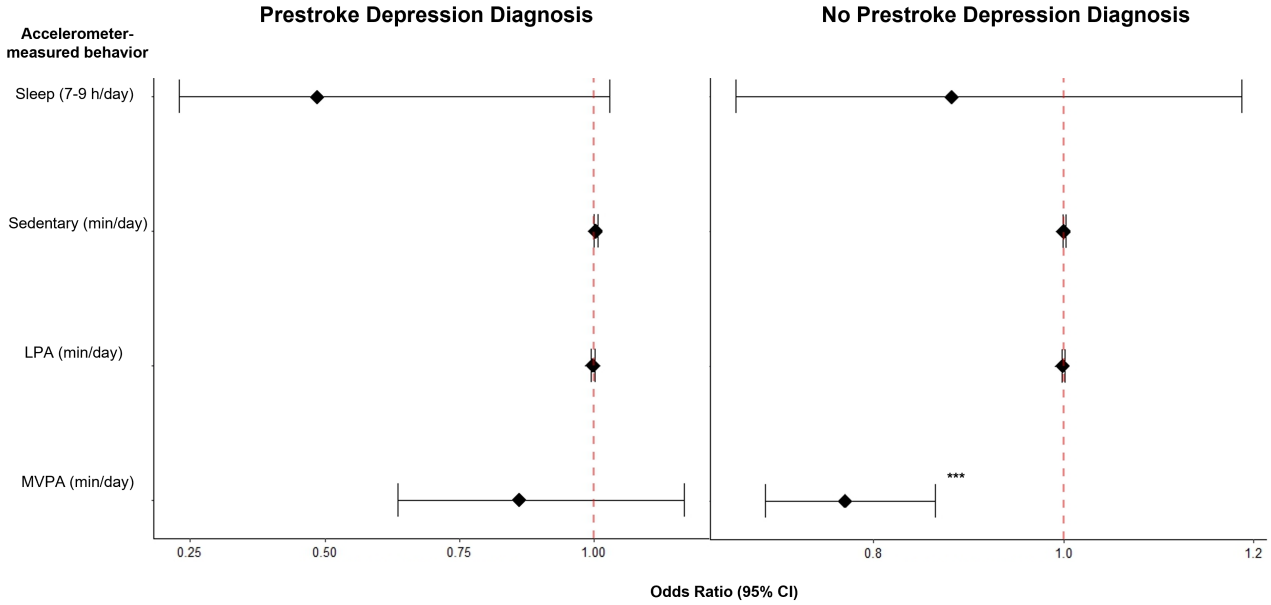
[b]SB: sedentary behavior.

[c]LPA: light physical activity.

[d]MVPA: moderate-to-vigorous physical activity.

Models were adjusted for age, sex, ethnicity, time-to-assessment, hyperlipidemia diagnosis, and diabetes diagnosis. Odds ratios (ORs) with 95% CIs for frequency of depressed mood are reported (Figure 2). ORs above 1 correspond to an increase in the accelerometer-measured behavior associated with increased depressed mood frequency.

Figure 2. Forest plot of odds ratios for depressed mood frequency by accelerometer-measured behavior comparing participants with prestroke depression diagnosis (DDs) and control cohorts. LPA: light physical activity; MVPA: moderate-to-vigorous physical activity. *** denotes statistical significance.



Sensitivity Analysis

Study Characteristics

In each filtered cohort (Multimedia Appendix 2), DDs were younger than controls (general cohort: 60 vs 62 y; IS: 64 vs 66 y; IC: 65 vs 67 y) and had a greater proportion of females (69.4% vs 56.8%; 60.0% vs 45.6%; 61.8% vs 45.1%). In the general population cohort, DDs had a greater proportion of White participants (97.7% vs 97%). On average, DDs also spent less time across cohorts in MVPA (35.0 vs 42.9 min/d; 31.0 vs 39.5 min/d; 32.0 vs 38.4 min/d), less time in LPA (295.0 vs 300.3 min/d; 286.7 vs 291.0 min/d; 287.5 vs 287.7 min/d), and more time asleep (9.1 vs 8.9 h/d; 9.1 vs 9.0 h/d; 9.04 vs 8.98 h/d).

While DDs in the general cohort spent slightly less time, on average, in SB than controls (564.3 vs 564.4 min/d), DDs in the IS and IC cohorts spent more time sedentary on average (577.1 vs 569.9 min/d; 578.2 vs 575.3 min/d).

In the general cohort, DDs had a higher proportion of diabetes (4.4% vs 2.9%) and hyperlipidemia (17.9% vs 14.9%) diagnoses and a lower proportion of participants with optimal sleep duration per day (49.5% vs 55.4%).

For the IS cohort, the average time from the completion of the depression survey to first stroke diagnosis was slightly more for DDs (1.9, SD 0.7 y) than controls (1.8, SD 0.6 y). In the IC cohort, the average time from the completion of the depression survey to first CeVD diagnosis was similarly more for DDs (1.9, SD 0.7 y) than controls (1.8, SD 0.7 y).

Cross-Sectional Analysis

In the general model (Multimedia Appendix 3), for each minute spent in MVPA, the odds of being in a higher depressed mood frequency category decreased by 18.4% (P<.001) and 13.5% (P<.001) for DDs (n=6096) and controls (n=62,589), respectively.

Also in the general model, specific only to controls, getting guideline-recommended sleep hours (7 - 9 h) each day was associated with a decreased odds of being in a higher depressed mood frequency category (5.3%; P=.02).

No significant associations were identified for those in the IS-only cohort (Multimedia Appendix 4).

For the final sensitivity analysis (Multimedia Appendix 5), assessing only those participants with an IC diagnosis, including stroke, the odds of being in a higher depressed mood frequency category decreased by 12.2% for each minute increase in MVPA (P=.03), only in controls (n=1526).

Discussion

Principal Findings

This investigation partially supports the hypothesis that objective behavior predictors would be associated with future depressed mood frequency. Although we found no significant associations between depressed mood frequency and SB, LPA or sleep for patients with prior stroke, regardless of prestroke depression diagnosis, we did find that the odds of being in a higher depressed mood frequency category decreased for each minute spent in MVPA; however, this association was only observed in participants without a prestroke depression diagnosis. This finding supports the exploratory aim of this manuscript, suggesting that participants with prestroke depression may experience different behavioral patterns compared to those without a prestroke depression diagnosis. Such a finding can potentially help clinicians tailor programs monitoring patients at risk of PSD.

The sensitivity analysis in the general cohort corroborates established findings that MVPA confers a protective effect on

mood, regardless of previous depression diagnosis. The lack of findings for the sensitivity analysis including only IS cases may be driven by the small sample sizes; however, the lack of findings also brings into question the potential for accelerometers to capture clinically actionable aberrations in patients before a stroke. Given that the protective effect of MVPA on depressed mood frequency was observed in the control cohort of patients with IC, accelerometer monitoring may be more appropriately directed to assess a broader range of neurologic changes, not just those linked with strokes.

Overall, the results suggest that accelerometer-based monitoring of behavior linked to depressed mood frequency may help clinicians identify patients who would benefit from resource-intensive screening, like the MADRS assessment. The sensitivity analyses support a separate approach for monitoring patients with a previous depression diagnosis, or more severe depression, compared to those with no documented depression or mild undiagnosed depression. When applied to predictive monitoring, a remote accelerometer-mood survey approach may be useful in cohorts of patients without a previous depression diagnosis, considering that patients with IC without clinical depression may experience observable behavior and mood changes before a CeVD diagnosis while their clinically depressed counterparts may not.

Limitations

A chief limitation of this study is that self-report data, like the depressed mood frequency survey, are subject to inaccuracies. Self-reported bias in survey responses may lead to misclassification of depressive symptom frequency and could influence different time-dependent results in our cohorts. Furthermore, the frequency of depression measures was not obtained by a clinician-graded protocol but, rather, by a survey questionnaire. Also, as the accelerometer study was only administered for one week and, on average, over a year before the follow-up mood survey, the impact of time between the objective measures and follow-up could have introduced substantial changes. The lack of associations observed for DDs may be due to the small sample size of participants with a previous depression diagnosis across cohorts. Moreover, the accelerometer study was only 1-week long and, therefore, may not generalize well to accurately represent busier or less busy weeks for patients. Accelerometer data collected on weekends versus weekdays may be distinct; however, this was not considered in this study.

The dichotomous investigation of clinically depressed and control patients are study strengths. In addition, UKBB participants were primarily White, limiting the generalizability of our findings outside of European populations. This UKBB study also primarily included participants aged 60 years and older and, as such, may not generalize well to young or middle-age adult populations. The majority of DDs were female across cohorts, a frequent finding in studies; however, male patients are less likely to seek out mental health resources, and the cohort stratification may be impacted by this.

Also, in the main analysis of previous stroke patients, participants diagnosed with clinical-grade depression after first stroke were excluded from this analysis. Considering the long gap in time from initial stroke to accelerometer study commencement, participants with a more immediate PSD diagnosis may either exhibit more intense symptoms or experience an underlying pathogenesis distinct from participants whose PSD symptoms are mild or emerge in the years after stroke.

Combining stroke types together as a single end point, as was done in the main analysis as well as the IS sensitivity analysis, may not consider unique characteristics of each stroke type and, as such, generated no significant results. Sleep was also assessed as a daily composite value, without consideration for time spent in a nap or broken sleep throughout the day. Together, these 2 limitations may have introduced confounding effects when considering sleep and depressed mood frequency, as previous research has shown short and long sleep to be associated with increased risk of intracerebral hemorrhage and ischemic stroke, respectively [40]. Furthermore, considering that all participants in our cohorts were hypertensive, MVPA's protective effect on depressive mood frequency may occur through improved cardiovascular health, rather than by conferring direct cerebral effects.

Comparison With Previous Work

No previous study assessed objective behavior measures and self-reported depressed mood frequency in patients with prior stroke years after their initial diagnosis. A key problem inherent in accelerometer research is that adherence to study designs is less-than-satisfactory for most studies [41]. This study also excluded participants with a more immediate PSD diagnosis, considering only those with prestroke depression diagnoses or those with no or mild depression after stroke. A self-report survey study of recent patients with prior stroke found that patients with high levels of PA before a stroke experienced less severe PSD [33]. Although our study could not confirm this analysis due to the design of the UKBB study, we extended those results by confirming that MVPA confers a protective effect on mood before a CeVD diagnosis in patients without a previous depression diagnosis, but not before a stroke-only diagnosis.

One plausible explanation for the lack of association between MVPA and depressed mood frequency in DDs may be that stroke survivors with a previous depression diagnosis have persistently deficient levels of brain-derived neurotropic factor (BDNF), a trophic factor released after exercise that is linked with improved mood benefits. It is well established that stroke patients in general have lower levels of BDNF, a marker of poor functional recovery [42]. The lack of a link between improved mood and MVPA in DDs may be driven by a less intense "exercise high" due to reduced or impaired BDNF function. In addition, other contributing factors, such as time spent in MVPA or neuroinflammation, may play a role in modulating BDNF expression in DDs. Of note, the lack of a significant association between MVPA and depressed mood frequency in participants with a previous depression diagnosis may be attributed to less time spent in MVPA compared with controls across all cohorts (patients with prior stroke, general population, IS, and IC). Time spent in MVPA may need to exceed a time threshold in

participants with previous depression diagnoses to improve mood.

The significant findings for IC cases, compared with the lack of findings for IS-only cases, are consistent with the updated *ICD-11* classification of CeVDs as a type of brain disease with shared etiology, rather than circulatory system disorders [37]. The protective effect of guideline-recommendation sleep (7 - 9 h/d) only observable in controls in the general cohort corroborates established work; however, the lack of associations across other cohorts may be explained by high levels of individual variability in sleep patterns, that is, nighttime disturbances, insomnia, and so on, previously identified in patients with depression as well as those at high risk of stroke [43,44].

A small pilot study of patients with minor ischemic stroke that found SB was positively associated with depression intensity and LPA was inversely associated with depression intensity [34]. Considering that this accelerometer study was conducted within the first 3 months after hospital discharge, our results extend these findings to look at mood in the years after a stroke. For instance, SB and LPA may be significant to monitor in the months after a stroke, while MVPA may be appropriate to monitor in the years after a stroke. Alternatively, MVPA may be less useful to monitor in minor ischemic stroke cases.

Using a larger dataset, our study builds on the feasibility demonstration of a small real-world study with patients with prior stroke, years after diagnosis, collecting one week of accelerometer data and ecological momentary assessments [45]. The results of our general cohort analysis considering participants without a previous depression diagnosis align with those from Sarris et al [46], who found that self-reported optimal sleep and PA were linked with decreased frequency of depressed mood in UKBB participants.

Conclusions

Our results highlight the importance of encouraging MVPA in patients with prior stroke without a depression diagnosis. Patients with prior strokes may be able to minimize short- and long-term disability and improve outcomes by proactively managing depressive symptoms. Applying MVPA to improve mood provides the added benefits of exercise-induced inflammation reduction and enhanced vascular elasticity while simultaneously reducing the risk of developing comorbidities and arterial stenosis or occlusion [47].

Considering that the only significant associations in the main analysis and incident sensitivity analyses were those that involved MVPA, it calls into question whether using accelerometer and depressed mood frequency survey data together can help clinicians identify patients who would benefit from remote monitoring, that is, this approach may generate more noise than signal over time. This study only considered a brief (1-week) accelerometer study, and over a year, on average, eclipsed between the *in situ* accelerometer study and the remote mood follow-up survey. Since neither the main analysis (previous stroke cohort) nor the incident sensitivity analyses resulted in significant associations for participants with a previous depression diagnosis, this underscores the need for additional research to determine whether this type of monitoring strategy can generate clinically actionable insights in participants with a previous depression diagnosis. Behavioral monitoring with accelerometer data and self-report surveys may not be helpful in patients with severe, or clinical-grade, depression. Future research should consider large sample sizes, longitudinal study designs, and analyze results stratified by time-to-diagnosis. Relevant to remote monitoring researchers, our findings highlight behavioral differences for those developing exploratory programs and clinically meaningful digital endpoints.

Overall, the cross-sectional analyses offer a robust perspective into the appropriateness of depression monitoring by digital sensors, using accelerometer wristwatches and smartphone, tablet, or PC-linked sensors. These insights offer clinical teams a strategy for translating digital health data, in this case, objective and subjective behavior measures, into scientifically valid frameworks for investigation. Future monitoring of patients at risk of different CeVD types, including those with a previous stroke diagnosis, should expand on our strategy and use both active and passive data to investigate relationships between objective digital sensor data and subsequent mood reports in patients diagnosed with and screened for depression. Based on our exploratory analysis, the potential for longitudinal data from objective sensors to predict mood appears feasible. In addition, PSD researchers should aim to characterize behavior measures linked with depressed mood across defined and clinically meaningful time periods, such as in the 3-month routine monitoring period after a stroke or TIA, considering that observable behaviors may evolve as CeVD or other neurologic disorder pathogenesis progresses.

Acknowledgments

This research has been conducted using the UK Biobank Resource under Application Number 91159. This research was funded by Mayo Clinic. SJZ was supported by a predoctoral fellowship in value assessment from the PhRMA Foundation and a National Center for Advancing Translational Sciences TL1 training grant (5TL1TR002380-05).

Disclaimer

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Data Availability

All data are publicly available, upon research approval access, from UK Biobank [48]. The datasets generated during and analyzed during this study are available from the corresponding author on reasonable request. Analysis code is available [49].

Authors' Contributions

SJZ completed the study design and manuscript drafting. BJE, BMD, and AG provided clinical expertise and contributed to manuscript editing. GMC provided expertise for obtaining data access and designing the study. AG conducted statistical review.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definitions for classifying patients.
[DOCX File, 17 KB - neuro_v4i1e56679_app1.docx]

Multimedia Appendix 2
Sample characteristics across sensitivity cohorts.
[DOCX File, 18 KB - neuro_v4i1e56679_app2.docx]

Multimedia Appendix 3
Ordinal logistic regression assessing objective behavior predictors and depressed mood frequency in the general cohort.
[DOCX File, 15 KB - neuro_v4i1e56679_app3.docx]

Multimedia Appendix 4
Ordinal logistic regression assessing objective behavior predictors and depressed mood frequency in incident stroke cohorts.
[DOCX File, 15 KB - neuro_v4i1e56679_app4.docx]

Multimedia Appendix 5
Ordinal logistic regression assessing objective behavior predictors and depressed mood frequency in incident cerebrovascular disease cohorts.
[DOCX File, 15 KB - neuro_v4i1e56679_app5.docx]

References

1. Broomfield NM, Quinn TJ, Abdul-Rahim AH, Walters MR, Evans JJ. Depression and anxiety symptoms post-stroke/TIA: prevalence and associations in cross-sectional data from a regional stroke registry. BMC Neurol 2014 Oct 1;14:198. [doi: 10.1186/s12883-014-0198-8] [Medline: 25269762]
2. Medeiros GC, Roy D, Kontos N, Beach SR. Post-stroke depression: a 2020 updated review. Gen Hosp Psychiatry 2020;66:70-80. [doi: 10.1016/j.genhosppsych.2020.06.011] [Medline: 32717644]
3. Selvaraj S, Arora T, Casameni Montiel T, et al. Early screening for post-stroke depression, and the effect on functional outcomes, quality of life and mortality: a protocol for a systematic review and meta-analysis. BMJ Open 2021 Aug 17;11(8):e050451. [doi: 10.1136/bmjopen-2021-050451] [Medline: 34404715]
4. Liu L, Marshall IJ, Pei R, et al. Natural history of depression up to 18 years after stroke: a population-based South London Stroke Register study. Lancet Reg Health Eur 2024 May;40:100882. [doi: 10.1016/j.lanepe.2024.100882] [Medline: 38745986]
5. Dall TM, Storm MV, Chakrabarti R, et al. Supply and demand analysis of the current and future US neurology workforce. Neurology (ECronicon) 2013 Jul 30;81(5):470-478. [doi: 10.1212/WNL.0b013e318294b1cf] [Medline: 23596071]
6. Brambilla M, Cerasetti M, Pepe F, et al. Comparison of Oxford Cognitive Screen and Montreal Cognitive Assessment feasibility in the stroke unit setting. a pilot study. Cereb Circ Cogn Behav 2021;2:100021. [doi: 10.1016/j.cccb.2021.100021] [Medline: 36324706]
7. Capaldi II VF. Emerging strategies in the treatment of poststroke depression and psychiatric distress in patients. PRBM 2010;3:109. [doi: 10.2147/PRBM.S10035]
8. Nakling AE, Aarsland D, Næss H, et al. Cognitive deficits in chronic stroke patients: neuropsychological assessment, depression, and self-reports. Dement Geriatr Cogn Disord Extra 2017 Sep 14;7(2):283-296. [doi: 10.1159/000478851]
9. Ahmed NN, Reagu S, Alkhoori S, Cherchali A, Purushottamahanti P, Siddiqui U. Improving mental health outcomes in patients with major depressive disorder in the Gulf States: a review of the role of electronic enablers in monitoring residual symptoms. J Multidiscip Healthc 2024;17:3341-3354. [doi: 10.2147/JMDH.S475078] [Medline: 39010931]

10. Lyketsos CG, Kozauer N, Rabins PV. Psychiatric manifestations of neurologic disease: where are we headed? Dialogues Clin Neurosci 2007;9(2):111-124. [doi: 10.31887/DCNS.2007.9.2/clyketsos] [Medline: 17726911]
11. van Sloten TT, Sigurdsson S, van Buchem MA, et al. Cerebral small vessel disease and association with higher incidence of depressive symptoms in a general elderly population: the AGES-Reykjavik study. Am J Psychiatry 2015 Jun;172(6):570-578. [doi: 10.1176/appi.ajp.2014.14050578] [Medline: 25734354]
12. Krick S, Koob JL, Latarnik S, et al. Neuroanatomy of post-stroke depression: the association between symptom clusters and lesion location. Brain Commun 2023;5(5):fcad275. [doi: 10.1093/braincomms/fcad275] [Medline: 37908237]
13. Moses JC, Adibi S, Shariful Islam SM, Wickramasinghe N, Nguyen L. Application of smartphone technologies in disease monitoring: a systematic review. Healthcare (Basel) 2021 Jul 14;9(7):889. [doi: 10.3390/healthcare9070889] [Medline: 34356267]
14. Wouters F, Gruwez H, Vranken J, et al. The potential and limitations of mobile health and insertable cardiac monitors in the detection of atrial fibrillation in cryptogenic stroke patients: preliminary results from the REMOTE trial. Front Cardiovasc Med 2022;9:848914. [doi: 10.3389/fcvm.2022.848914] [Medline: 35498000]
15. Phienphanich P, Tankongchamruskul N, Akarathanawat W, et al. Automatic stroke screening on mobile application: features of gyroscope and accelerometer for arm factor in FAST. Presented at: 2019 41st Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); Jul 23-27, 2019; Berlin, Germany p. 4225-4228. [doi: 10.1109/EMBC.2019.8857550]
16. Capela NA, Lemaire ED, Baddour N, Rudolf M, Goljar N, Burger H. Evaluation of a smartphone human activity recognition application with able-bodied and stroke participants. J Neuroeng Rehabil 2016 Jan 20;13:5. [doi: 10.1186/s12984-016-0114-0] [Medline: 26792670]
17. Beerten SG, Proesmans T, Vaes B. A heart rate monitoring app (FibriCheck) for atrial fibrillation in general practice: pilot usability study. JMIR Form Res 2021 Apr 7;5(4):e24461. [doi: 10.2196/24461] [Medline: 33825692]
18. Forster SD, Gauggel S, Loevenich R, et al. A microanalysis of mood and self-reported functionality in stroke patients using ecological momentary assessment. Front Neurol 2022;13:854777. [doi: 10.3389/fneur.2022.854777] [Medline: 35665036]
19. Forster SD, Gauggel S, Petershofer A, Völzke V, Mainz V. Ecological momentary assessment in patients with an acquired brain injury: a pilot study on compliance and fluctuations. Front Neurol 2020;11:115. [doi: 10.3389/fneur.2020.00115] [Medline: 32194494]
20. Lenaert B, van Kampen N, van Heugten C, Ponds R. Real-time measurement of post-stroke fatigue in daily life and its relationship with the retrospective Fatigue Severity Scale. Neuropsychol Rehabil 2022 Jul;32(6):992-1006. [doi: 10.1080/09602011.2020.1854791] [Medline: 33297839]
21. Helgadóttir B, Forsell Y, Ekblom Ö. Physical activity patterns of people affected by depressive and anxiety disorders as measured by accelerometers: a cross-sectional study. PLoS One 2015;10(1):e0115894. [doi: 10.1371/journal.pone.0115894] [Medline: 25585123]
22. Murphy RP, Reddin C, Rosengren A, et al. Depressive symptoms and risk of acute stroke: INTERSTROKE case-control study. Neurol (ECronicon) 2023 Apr 25;100(17):e1787-e1798. [doi: 10.1212/WNL.0000000000207093] [Medline: 36889922]
23. Everson-Rose SA, Roetker NS, Lutsey PL, et al. Chronic stress, depressive symptoms, anger, hostility, and risk of stroke and transient ischemic attack in the multi-ethnic study of atherosclerosis. Stroke 2014 Aug;45(8):2318-2323. [doi: 10.1161/STROKEAHA.114.004815] [Medline: 25013018]
24. Avan A, Hachinski V. Stroke and dementia, leading causes of neurological disability and death, potential for prevention. Alzheimers Dement 2021 Jun;17(6):1072-1076. [doi: 10.1002/alz.12340] [Medline: 34057294]
25. Heshmatollah A, Dommershuijsen LJ, Fani L, Koudstaal PJ, Ikram MA, Ikram MK. Long-term trajectories of decline in cognition and daily functioning before and after stroke. J Neurol Neurosurg Psychiatry 2021 Nov;92(11):1158-1163. [doi: 10.1136/jnnp-2021-326043] [Medline: 34230107]
26. Johnson SA, Karas M, Burke KM, et al. Wearable device and smartphone data quantify ALS progression and may provide novel outcome measures. NPJ Digit Med 2023 Mar 6;6(1):34. [doi: 10.1038/s41746-023-00778-y] [Medline: 36879025]
27. Winer JR, Lok R, Weed L, et al. Impaired 24-h activity patterns are associated with an increased risk of Alzheimer's disease, Parkinson's disease, and cognitive decline. Alz Res Therapy 2024;16(1):35. [doi: 10.1186/s13195-024-01411-0]
28. Zawada SJ, Aissa NH, Conte GM, et al. In situ physiologic and behavioral monitoring with digital sensors for cerebrovascular disease: a scoping review. Mayo Clin Proc Dig Health 2023 Jun;1(2):139-160. [doi: 10.1016/j.mcpdig.2023.03.007]
29. Jean FAM, Swendsen JD, Sibon I, Fehér K, Husky M. Daily life behaviors and depression risk following stroke: a preliminary study using ecological momentary assessment. J Geriatr Psychiatry Neurol 2013 Sep;26(3):138-143. [doi: 10.1177/0891988713484193] [Medline: 23584854]
30. Johnson EI, Sibon I, Renou P, Rouanet F, Allard M, Swendsen J. Feasibility and validity of computerized ambulatory monitoring in stroke patients. Neurology (ECronicon) 2009 Nov 10;73(19):1579-1583. [doi: 10.1212/WNL.0b013e3181c0d466] [Medline: 19901250]
31. Sibon I, Lassalle-Lagadec S, Renou P, Swendsen J. Evolution of depression symptoms following stroke: a prospective study using computerized ambulatory monitoring. Cerebrovasc Dis 2012;33(3):280-285. [doi: 10.1159/000334663] [Medline: 22285959]

32. Tanaka H, Kitamura G, Tamura M, et al. Pre-stroke physical activity is associated with post-stroke physical activity and sedentary behavior in the acute phase. Sci Rep 2023;13(1):21298. [doi: 10.1038/s41598-023-48232-z]
33. Reinholdsson M, Palstam A, Jood K, S Sunnerhagen K. Associations between pre-stroke physical activity levels and health-related quality of life 3 months after stroke: a registry-based study (part of PAPSIGOT). Int J Stroke 2023 Dec;18(10):1178-1185. [doi: 10.1177/17474930231184367]
34. Ashizawa R, Honda H, Yoshizawa K, Kameyama Y, Yoshimoto Y. Association between physical activity levels and depressive symptoms in patients with minor ischemic stroke. J Stroke Cerebrovasc Dis 2022 Sep;31(9):106641. [doi: 10.1016/j.jstrokecerebrovasdis.2022.106641] [Medline: 35834937]
35. Lau SCL, Tabor Connor L, Baum CM. Motivation, physical activity, and affect in community-dwelling stroke survivors: an ambulatory assessment approach. Ann Behav Med 2023 Apr 12;57(4):334-343. [doi: 10.1093/abm/kaac065] [Medline: 36732938]
36. Alvaro PK, Roberts RM, Harris JK. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. Sleep 2013 Jul 1;36(7):1059-1068. [doi: 10.5665/sleep.2810] [Medline: 23814343]
37. Ramasubbu R. Relationship between depression and cerebrovascular disease: conceptual issues. J Affect Disord 2000;57(1-3):1-11. [doi: 10.1016/s0165-0327(99)00101-9] [Medline: 10708811]
38. Madjedi KM, Stuart KV, Chua SYL, et al. The association of physical activity with glaucoma and related traits in the UK Biobank. Ophthalmology 2023 Oct;130(10):1024-1036. [doi: 10.1016/j.ophtha.2023.06.009] [Medline: 37331483]
39. Mac Grory B, Schrag M, Biousse V, et al. Management of central retinal artery occlusion: a scientific statement from the American Heart Association. Stroke 2021 Jun;52(6):e282-e294. [doi: 10.1161/STR.0000000000000366] [Medline: 33677974]
40. Lu H, Wu PF, Li RZ, Zhang W, Huang GX. Sleep duration and stroke: a Mendelian randomization study. Front Neurol 2020;11:976. [doi: 10.3389/fneur.2020.00976]
41. Barak S, Wu SS, Dai Y, Duncan PW, Behrman AL. Locomotor Experience Applied Post-Stroke (LEAPS) Investigative Team. Adherence to accelerometry measurement of community ambulation poststroke. Phys Ther 2014;94(1):101-110. [doi: 10.2522/ptj.20120473]
42. Chaturvedi P, Singh AK, Tiwari V, Thacker AK. Brain-derived neurotrophic factor levels in acute stroke and its clinical implications. Brain Circ 2020;6(3):185-190. [doi: 10.4103/bc.bc_23_20] [Medline: 33210043]
43. Lim JA, Yun JY, Choi SH, Park S, Suk HW, Jang JH. Greater variability in daily sleep efficiency predicts depression and anxiety in young adults: estimation of depression severity using the two-week sleep quality records of wearable devices. Front Psychiatry 2022;13:1041747. [doi: 10.3389/fpsyt.2022.1041747] [Medline: 36419969]
44. Ma J, Ma N, Zhang L, Xu L, Liu X, Meng G. Association of total sleep duration variability with risk of new stroke in the middle-aged and elderly Chinese population. BMC Neurol 2024 Jun 25;24(1):217. [doi: 10.1186/s12883-024-03727-8] [Medline: 38918750]
45. Lau SCL, Connor LT, King AA, Baum CM. Multimodal ambulatory monitoring of daily activity and health-related symptoms in community-dwelling survivors of stroke: feasibility, acceptability, and validity. Arch Phys Med Rehabil 2022 Oct;103(10):1992-2000. [doi: 10.1016/j.apmr.2022.06.002] [Medline: 35780826]
46. Sarris J, Thomson R, Hargraves F, et al. Multiple lifestyle factors and depressed mood: a cross-sectional and longitudinal analysis of the UK Biobank (N = 84,860). BMC Med 2020 Nov 12;18(1):354. [doi: 10.1186/s12916-020-01813-5] [Medline: 33176802]
47. Prior PL, Suskin N. Exercise for stroke prevention. Stroke Vasc Neurol 2018 Jun;3(2):59-68. [doi: 10.1136/svn-2018-000155] [Medline: 30191075]
48. UK Biobank. URL: https://www.ukbiobank.ac.uk/ [accessed 2024-12-18]
49. Szawada-mayo/jmir_56679. GitHub. URL: https://github.com/szawada-mayo/jmir_56679 [accessed 2024-12-13]

Abbreviations

BDNF: brain-derived neurotropic factor
CeVD: cerebrovascular disease
DD: depression diagnosis
IC: incident cerebrovascular disease
***ICD-10*:** *International Classification of Diseases, Tenth Revision*
***ICD-11*:** *International Classification of Diseases, Eleventh Revision*
IS: incident stroke
LPA: light physical activity
MADRS: Montgomery–Åsberg Depression Rating Survey
MVPA : moderate-to-vigorous physical activity
OR: odds ratio
PA: physical activity
PHQ-2: Patient Health Questionnaire-2
PSD: poststroke depression

SB: sedentary behavior
TIA: transient ischemic attack
UKBB: UK Biobank

Edited by P Kubben; submitted 23.01.24; peer-reviewed by E Harahsheh, N Allen; revised version received 23.10.24; accepted 14.11.24; published 10.01.25.

Please cite as:
Zawada SJ, Ganjizadeh A, Conte GM, Demaerschalk BM, Erickson BJ
Exploring Remote Monitoring of Poststroke Mood With Digital Sensors by Assessment of Depression Phenotypes and Accelerometer Data in UK Biobank: Cross-Sectional Analysis
JMIR Neurotech 2025;4:e56679
URL: https://neuro.jmir.org/2025/1/e56679
doi:10.2196/56679

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org