Development and Validation of Starkey's Balance Assessment Feature

Justin R. Burwinkel[1], Kristen K. Steenerson[2], Majd Srour[1], Christy Cloninger[1], and Chris Howes[1]

Background

The relationship between hearing difficulties and increased fall risk is well-documented in the literature. One of the studies found that individuals who sought treatment in audiology clinics, specifically, were at a greater risk of falling than their age-matched peers *(Criter & Honaker, 2016). Lin & Ferrucci (2012)* also reported a significant association between the severity of hearing impairment and reports of falls, even when adjusting for demographic, cardiovascular, and vestibular balance function.

Fall risk is understood to be multifactorial, and numerous behavioral, physiological, and pathological mechanisms could underlie the reported associations between hearing impairment and falls, including: comorbid vestibular, neural, or cardiovascular pathology; genetic influences; decreased awareness of the auditory environment; divided attentional resources for locomotion and maintenance of postural balance; and frailty, which could be exacerbated by social isolation, depression, physical inactivity, and cognitive impairment (see Agmon et al. for a review).

Despite the well-established links between both advanced age and hearing impairment with fall risk, routine fall risk assessments are inconsistently performed in hearing clinics and other healthcare settings where older adults are commonly seen *(Howland et al., 2018; Patterson & Honaker, 2014).* Currently, the burden of fall risk screening is largely addressed by primary care providers or gerontologists during Medicare's Annual Wellness Visits (AWV).

However, these visits are underutilized within the healthcare system, with reportedly fewer than 1 in 5 older Americans receiving these no-cost services each year *(Centers for Medicare & Medicaid Services, 2017).* Given that hearing care professionals often see older patients who are at increased risk of falls, they are well-positioned to integrate fall risk screenings and assessments into their routine practices. These additional steps align with a more comprehensive approach to patient care and represent an opportunity to markedly improve patient outcomes, especially since fall risk assessments may not have been conducted with the patient elsewhere.

Moreover, a study surveying primary care providers (PCPs) revealed that while 96% of providers agree that all older adults should be assessed for fall risk, only 14% are familiar with the U.S. Centers for Disease Control and Prevention's (CDC) Stopping Elderly Accidents, Deaths & Injuries (STEADI) Toolkit, which provides a practical framework for effectively implementing the collaborative fall prevention guidelines published by the American Geriatrics Society (AGS) and the British Geriatrics Society (BGS) *(Eckstrom et al., 2017).* In light of this challenge, there is a critical need for accessible tools that can help address gaps in fall risk management, particularly in settings where regular, comprehensive evaluations are not consistently performed.

[1]*Starkey*
[2]*Stanford University School of Medicine, Department of Otolaryngology-Head and Neck Surgery*

The Balance Assessment feature in the My Starkey mobile app addresses this need by offering a user-friendly way for individuals to regularly screen their fall risk and monitor key aspects of their balance, gait, and strength. By integrating advanced sensor technology, artificial intelligence (AI), and the core principles of the CDC's STEADI initiative, it provides clinicians with a reliable tool to support proactive fall risk management between office visits. Regular self-assessment empowers individuals to maintain greater awareness of their balance health and address modifiable risk factors before they escalate, potentially leading to improved patient outcomes and reduced fall-related injuries.

Balance Assessment Feature At a Glance

The Balance Assessment feature in the My Starkey app is a groundbreaking tool designed to help hearing aid users monitor changes in their balance status and understand how modifiable risk factors for falling might affect them personally. Rooted in internationally recognized medical guidelines for fall prevention, it adheres closely to the CDC's recommended screening and functional assessment protocols. The feature begins by administering an electronic version of the STEADI initiative's 12-item *Stay Independent* screening questionnaire. It then utilizes the embedded motion sensors and artificial intelligence of the Starkey Edge AI hearing aids to evaluate the user's performance in functional assessments.

As users progress through the various tasks, the hearing aids' real-time motion sensor data is evaluated by AI algorithms that autonomously score the assessments and meaningfully display the results. These algorithms are designed to evaluate users' movements during the tests similarly to a trained clinician and have been independently validated by several studies conducted at Stanford Medicine's Department of Otolaryngology-Head and Neck Surgery, as detailed in this white paper.

For the first time, Edge AI hearing aids can autonomously score users' performance on commonly recommended assessments for evaluating older adults' functional gait, strength, and balance. This innovative approach helps identify users who may have balance limitations or concerns that warrant further investigation. The user-friendly interface, developed with input and feedback from several groups of intended users and clinical experts, effectively guides users through each phase of the assessment with a combination of written, illustrative, and verbal instructions.

Research Studies

The Balance Assessment feature was developed through a multi-year collaboration between Starkey and Stanford Medicine, as outlined in Figure 1 *(see top of next page)*. Initially, Starkey collected training data from 71 participants aged 47 to 99 years (mean age = 79.2 years, SD = 9.9) to create algorithms for autonomously scoring gait, strength, and balance tests using motion sensor data from research hearing devices *(Burwinkel et al., 2022)*. A blinded data collection app was then developed, enabling Stanford Medicine's clinical researchers to manually score each test according to the STEADI initiative's assessment protocols, independent of the app's automated scoring. The accuracy of these algorithms was later evaluated by comparing the clinicians' scores with the app's automated results.

Based on these evaluations, Starkey refined the algorithms iteratively and conducted formative usability studies using prototype versions of the feature. Stanford Medicine further explored the feasibility of remote assessments and compared the Balance Assessment feature's performance between in-person, remotely supervised, and independently administered at-home test settings. The final stage of development included a field validation study, comparing participants' performance in supervised lab trials to unsupervised home use, and examining the impact of clinician demonstrations of the feature on user performance.

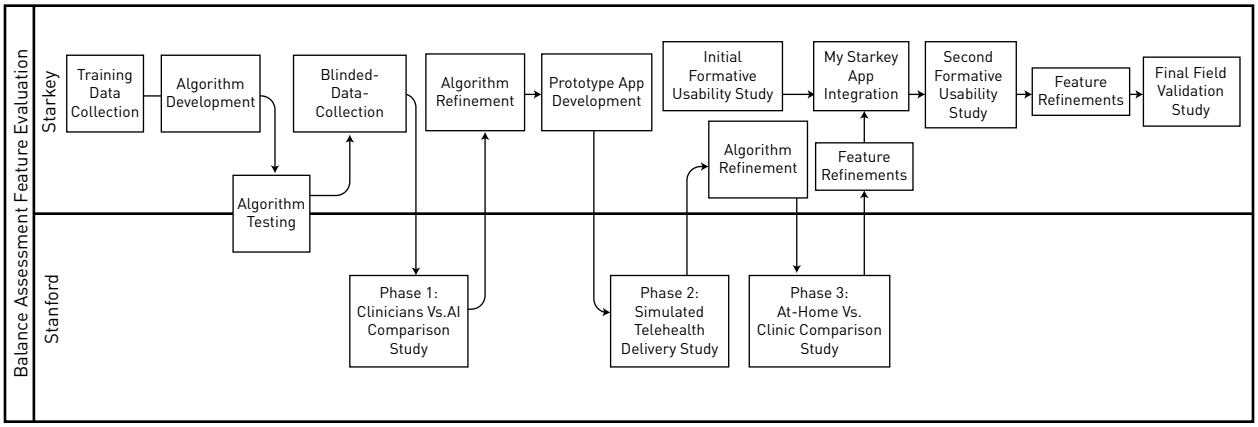

Figure 1. Research and development process of the Balance Assessment feature, a collaboration between Starkey and Stanford Medicine. This included algorithm development, clinician and administration modality validation, usability studies, and a final field validation comparing supervised lab trials with unsupervised home use.

Stanford Medicine Study

Stanford Medicine conducted a three-phase research study evaluating the accuracy and reliability of the Balance Assessment's algorithms (for assessing various fall risk factors) in comparison to evaluations conducted by experienced clinicians using the STEADI initiative's functional gait, strength, and balance test battery. The STEADI initiative's assessment protocols include administering the 4-Stage Balance, 30-Second Chair Stand, and Timed Up and Go (TUG) tests. Each of these tests is recognized for having ecological validity and predictive accuracy for identifying fall risk in older adults.

Phase 1 of the Stanford Medicine study compared the clinicians' observations with the autonomous scoring generated by early versions of the Balance Assessment feature's algorithms. Phase 2 explored the feasibility of using the feature in telehealth. Phase 3 assessed the effectiveness of unsupervised at-home testing and how well the research prototype of the Balance Assessment feature performed outside of a controlled laboratory environment.

Phase 1

The first phase of the study employed a blinded comparative approach to assess fall risk using both clinicians' evaluations and the autonomous assessments derived from hearing aids equipped with motion sensors and AI technologies.

The study was conducted at a tertiary referral center and involved 250 participants aged 55 to 100 years who were identified as having heightened fall risk due to advanced age and/or a history of instability or falls (mean age = 78.4 years, SD = 9.6).

Methodology

Participants underwent the CDC STEADI initiative's functional gait, strength, and balance test battery while wearing bilateral hearing aids embedded with motion sensors. Each trial was independently interpreted by the hearing aids' algorithms and three clinicians: one present during trials and two who later reviewed video recordings of the same trials. All of the clinicians were blinded to the hearing aids' autonomous scoring. The primary outcomes measured included both the algorithmic scores and the manually determined scores from the three clinicians' observations of the 4-Stage Balance, 30-Second Chair Stand, and TUG test attempts.

Results

On the whole, the researchers observed good agreement between the early Balance Assessment feature's algorithms and clinicians, when evaluating fall risk according to the STEADI initiative's protocols. There were no statistically significant differences between the clinicians' interpretations and the Balance Assessment feature's algorithms for the 4-Stage Balance and TUG tests ($p > 0.05$).

However, a significant difference was observed for the 30-Second Chair Stand test ($t = 10.13$, $p < 0.05$), with a mean difference of -0.8 stand counts. Inter-rater reliability among clinicians was excellent.

Discussion

While the Balance Assessment feature's early algorithms provided accurate scores overall, the consistent underscoring for the 30-Second Chair Stand test highlighted a discrepancy that the researchers attributed to how early the Balance Assessment feature was applying the CDC's scoring guidelines. Specifically, the CDC instructs clinicians to count chair stands initiated before the test time expires as full stands, a consideration that was not initially implemented in the hearing aid scoring algorithm for that test. This finding prompted a refinement of the algorithm for subsequent studies, aiming to better align the Balance Assessment feature's scoring with CDC standards.

Phase 2

Phase 2 of the Stanford Medicine study expanded on the Phase 1 research by assessing fall risk using the STEADI initiative's assessment protocols within a telehealth framework. This phase focused on the feasibility of remote administration of functional gait, strength, and balance tests. This phase was conducted at the same tertiary referral center using the same blinded data collection app as Phase 1. It involved 50 participants aged 57 to 98 years who were at elevated risk of falling due to advanced age or a history of instability or falls (mean age = 78.2 years, SD = 8.1).

Methodology

In this second phase, participants underwent the STEADI initiative's functional assessment test battery while wearing bilateral hearing aids embedded with motion sensors. Unlike in Phase 1, the tests were administered remotely, with a clinician providing direction and supervision via a secure video call.

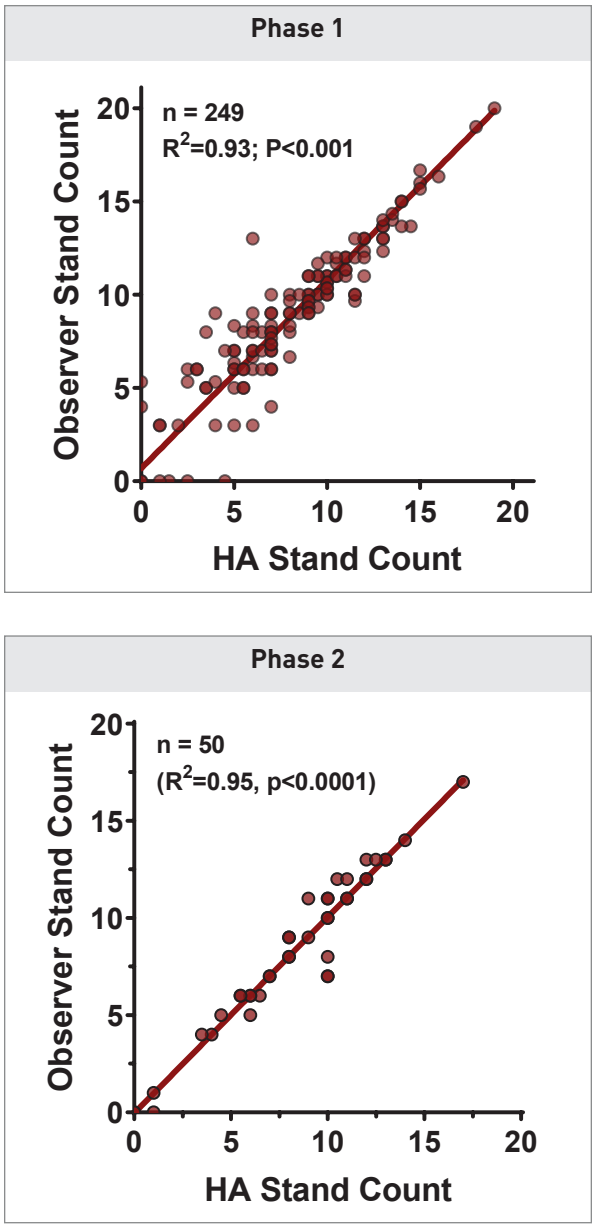

Figure 2. Comparison of the Balance Assessment feature's algorithm performance in the 30-Second Chair Stand test across two phases. The top plot represents Phase 1 data (n=249, r^2=0.93; p<0.001) and the bottom plot represents Phase 2 data (n=50, r^2=0.95; p<0.0001). The y-axes show the total stand counts observed by clinicians, while the x-axes show the total stand counts recorded by the Balance Assessment feature. The closer alignment in Phase 2 indicates improved algorithm accuracy.

A researcher was present in the lab to monitor each participant for safety concerns but otherwise did not participate in directing or administering the tests. Each participant completed one trial of each task, which was evaluated by the hearing aids' algorithms and scored by a clinician who had access to video recordings of the trials. The clinician was blinded to the hearing aids' scoring, ensuring an unbiased comparison between the algorithm's output and the clinician's assessment. The primary outcomes measured included both the algorithmic scores and the scores from clinician observation for the 4-Stage Balance, 30-Second Chair Stand, and TUG tests.

Results

The results from Phase 2 were fairly consistent with those of Phase 1, showing good overall agreement between the Balance Assessment feature's algorithms and the clinician when evaluating performance of the gait, strength, and balance tests according to the STEADI initiative's assessment protocols.

4-Stage Balance test: Slightly better agreement between the Balance Assessment feature's algorithms and the clinician was observed for the side-by-side, toe-to-instep, and one-foot balance poses, while the tandem stand pose showed poorer agreement. These differences were relatively small and could be attributable to the increased variance stemming from the smaller sample size used in Phase 2 compared to Phase 1.

30-Second Chair Stand test: The Balance Assessment feature's algorithm for scoring the 30-Second Chair Stand test showed significant improvement from Phase 1 ($n=249$, $r^2=0.93$; $p<0.001$) to Phase 2 ($n=50$, $r^2=0.95$; $p<0.0001$), as shown in Figure 2 *(see previous page)*. By Phase 2, the difference between the algorithm's scores and the clinician's observations was no longer statistically significant ($n=50$, $t=0.24$, $p>0.05$).

This outcome highlights the effectiveness of the algorithm refinements made between Phase 1 and Phase 2, resulting in improved alignment between the algorithm and the clinician's count totals.

Timed Up and Go (TUG) test: In Phase 1, the early Balance Assessment feature's algorithms showed no significant difference from clinician assessments of the TUG test ($n=220$, $t=0.94$; $p>0.05$). However, in Phase 2, the algorithms demonstrated a significant difference ($n=48$, $t=2.51$; $p<0.05$), with generally slower TUG times recorded. This difference suggests that the remote administration of the test during Phase 2 may have introduced additional variability or challenges that affected the algorithms' performance compared to direct clinician observation. Nonetheless, this effect appears to have primarily influenced the test's specificity (i.e., its ability to rule out non-fallers) rather than its sensitivity (i.e., its ability to detect potential fallers), indicating that the early Balance Assessment feature remained a viable tool for monitoring a user's TUG test performance and worthy of further exploration, despite the differences observed Figure 3 *(see next page)*.

Discussion

Phase 2 of the Stanford Medicine study demonstrated the feasibility of remote assessment for accurately evaluating gait, strength, and balance in older adults at risk of falls. The remote administration of these tasks also proved safe, with no falls occurring during the participants' performance of the gait, strength, and balance tests.

Similar to Phase 1, the remote scoring accuracy of the algorithms was also generally high. Agreement between the Balance Assessment feature's algorithms and clinician observations was strong for most tasks.

Notably, the accuracy of the strength test improved from Phase 1 to Phase 2 due to the algorithm refinements that were made between Phase 1 and Phase 2 studies. However, some increased variances were observed in the balance tasks and the gait test, likely due to the smaller sample size and potential challenges of remote testing. These findings highlighted the need for clear instructions to be provided visually and verbally.

Overall, Phase 2 demonstrated that remote assessment of gait, strength, and balance could be both accurate and safe, providing a solid foundation for further improvements to the early Balance Assessment feature.

Phase 3

The third phase of the Stanford Medicine study aimed to evaluate fall risk assessments using the STEADI initiative's protocols in participants' home environments utilizing bilateral hearing aids equipped with motion sensors and AI technologies. This study included 50 participants with ages ranging from 56 to 97 (mean age = 76.0, SD=8.3), all of whom were identified for increased fall risk based on self-reported criteria from the STEADI initiative's *Stay Independent* screening criteria.

Methodology

Participants completed three at-home trials: a learning trial followed by two subsequent trials for assessment. Each trial included the 4-Stage Balance, 30-Second Chair Stand, and TUG tests, with assessments conducted using the hearing aids' embedded technologies.

Results

At-home testing revealed varying trial-to-trial agreement for the 4-Stage Balance test, ranging from 76% to 87% agreement for the three standing positions used in the STEADI initiative's pass/fail criteria after removing trials where no score was assigned due to user error with the research version of the application.

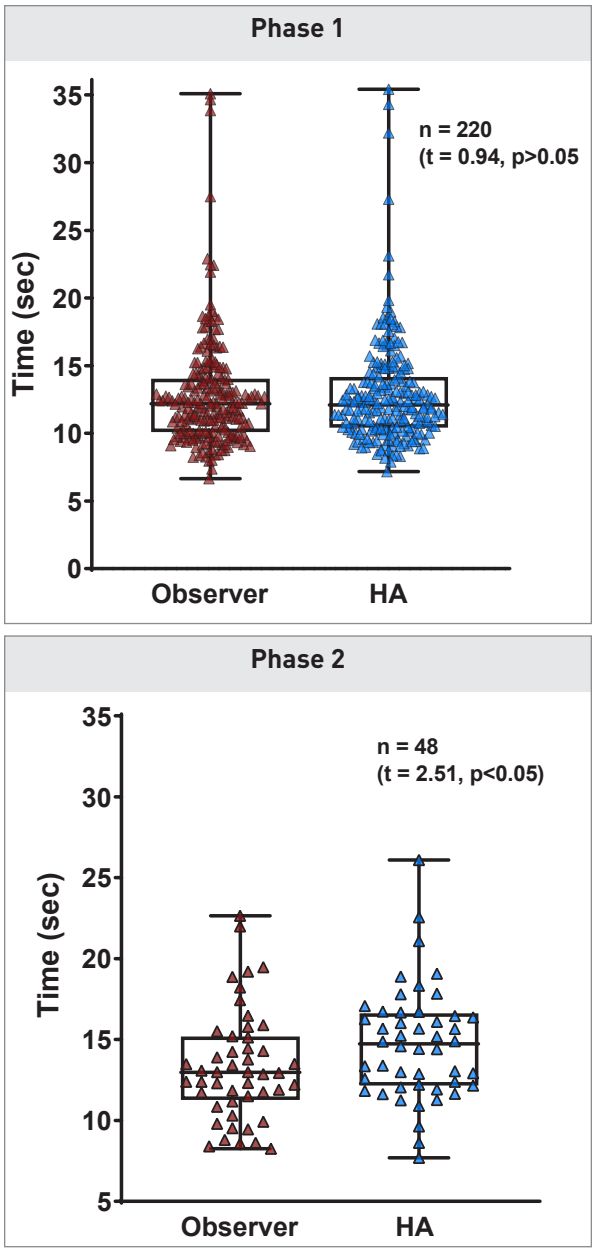

Figure 3. Comparison of TUG test times recorded by clinicians (Observer) and the Balance Assessment feature's algorithms (HA) during Phase 1 (left) and Phase 2 (right). Each box plot represents the distribution of TUG times measured in seconds, with individual data points scattered over the plots. The box plots highlight the central tendency and variability in TUG times for both phases, illustrating the significant difference in algorithm performance observed in Phase 2 compared to the closer alignment observed in Phase 1.

The 30-Second Chair Stand test initially showed moderate test-retest reliability. However, similar "misreads" occurred when one of the participants' attempts recorded zero stands while another attempt recorded a measured score. Adjusting for these misreads demonstrated strong reliability ($n=41$, $R2=0.83$, $p < 0.001$) for the 30-Second Chair Stand test. Similarly, the TUG test exhibited moderate reliability, with misreads affecting about 20% of trials. After correcting for misreads, reliability improved significantly ($n=40$, $R2=0.84$, $p < 0.001$) for the TUG test.

Discussion

The Phase 3 study highlighted notable challenges in conducting at-home fall risk assessments using a mobile application. Researchers characterized a notable percentage of trials as "misreads", indicating users had difficulties with the application that led to some trials missing scores. Anecdotal reports from participants further underscored the usability concerns with the research version of the mobile application.

Despite these challenges, the study demonstrated that the STEADI initiative's assessment protocols could be safely implemented using hearing aids in home settings. Importantly, no reports of falls or balance issues occurred during self-administration of the assessments, suggesting that the protocol is suitable for at-home use by this population.

Formative Usability Testing

The Balance Assessment feature was developed through an extensive co-development process involving iterative feedback from users and clinicians. Starkey's objective was to create a tool that is user-friendly, effective, and reliable for hearing aid users to assess their balance and maintain independence. This process included multiple stages of formative usability testing, in addition to algorithm performance testing, to help refine the feature's design and functionality based on users' insights and experiences.

Methodology

Prior to being integrated into the My Starkey app, the design of the Balance Assessment feature—including the user interface, navigational flow, and instructional content—was guided by direct observations and structured interviews with potential users. Early usability testing aimed to understand user preferences, identify barriers to task completion, and ensure clear instructions were provided to users.

Following the development of a prototype app, Stanford Medicine's Department of Otolaryngology-Head and Neck Surgery was able to conduct studies that evaluated the early feature's effectiveness in telehealth and at-home environments. These studies, using hearing aids and the prototype Balance Assessment feature to complete the STEADI initiative protocols, identified several areas for improvement. These included reducing the user errors that had led to unscored trials and improving the handling of partial-trial scoring. Issues like misreads due to certain user behaviors, smartphone app as the balance tests began, were also identified, prompting adjustments to the user interface and algorithm refinements for enhanced robustness and error-handling.

Phase 4

A subsequent round of formative usability testing was performed with a candidate version of the Balance Assessment feature integrated into the My Starkey app. This phase involved a diverse group of 15 hearing aid users with varying levels of experience with smartphone applications. Participants used the Balance Assessment feature in a simulated living room environment to complete an electronic version of the *Stay Independent* questionnaire and functional gait, strength, and balance tests, while wearing Edge AI hearing aids.

They were also provided with a "junk drawer" containing tools they could select to measure the walking path for the tug test. Their adherence to tasks was monitored throughout the study.

Results

Feedback from the second round of formative usability testing indicated a marked improvement in user experience. Most participants found the instructions easy to follow and understood the results, with many expressing willingness to use the feature at home. However, three key areas for further improvement were identified:

1. Participants wanted more context regarding the purpose of each exercise and the measurements being taken.
2. The illustrative instructions provided for each exercise were not always noticed by users.
3. Instructions for the 30-Second Chair Stand and TUG tests were still unclear to some, leading to misunderstandings.

Discussion

Formative usability testing was pivotal in refining the Balance Assessment feature. Initial issues, such as unscored trials resulting from user errors, were resolved through algorithmic adjustments and clearer instructions. While significant progress was made, some usability challenges persisted—particularly related to how instructions were conveyed, and the context provided for each test. These persistent challenges emphasized the critical role of gaining continuous user feedback to optimize the feature's effectiveness. This iterative process of refinement was essential before the Balance Assessment feature could be confidently deployed in the final field validation study, which aimed to test the feature's robustness in unsupervised, real-world environments.

Final Field Validation

The final field validation study was designed to evaluate the Balance Assessment feature's usability and effectiveness in both supervised and unsupervised environments. The primary objective of the study was to determine whether the feature produced reliable and consistent results across different settings and whether clinician-led demonstrations affected usability or participant performance.

Methodology

Fourteen participants were recruited from Starkey's research participant database and were divided into two groups: one group (n=7) received an in-person demonstration from a hearing care professional on how to use the Balance Assessment feature, while the other group (n=7) was shown only how to access the feature within the app and asked to follow its instructions independently.

Both groups completed supervised Balance Assessment trials in a lab environment. The Demonstration Group did so before attempting the assessments at home, while the No Demonstration Group completed the supervised trials after returning from their field trial. This allowed the evaluation of whether prior demonstration influenced participant performance and understanding of the feature during unsupervised at-home use.

Notably, one participant from each group (n=2) did not complete the at-home trials: one participant mistakenly engaged with balance exercise videos in another section of the My Starkey app, while the other forgot how to access the Balance Assessment feature. A third participant (n=1) claimed to have completed the at-home trials, but their app usage data was notably absent from the cloud for those trials, precluding that participant's inclusion in the analysis.

Quantitative data was gathered from the Balance, 30-Second Chair Stand, and TUG test scores, as well as each participant's *Stay Independent* questionnaire responses. Both supervised and unsupervised trial score data were analyzed using Welch's t-Test, and the overall outcomes were compared using the Chi-Square test of independence to assess the consistency and reliability of the balance assessment attempts. Performance differences between the Demonstration Group (n=5) and No Demonstration Group (n=6) were analyzed similarly.

Results

The results of the field validation study showed no significant differences in Balance, Chair Stand, and TUG scores, or *Stay Independent* responses between unsupervised and supervised attempts, nor between the Demonstration and No Demonstration groups across most measures. As summarized in Table 1 *(see next page)*, statistical analyses using Welch's t-Test and Chi-Square tests revealed consistent test scores and outcomes across most measures, indicating no notable differences in performance due to test environment or method of instruction. However, a significant difference was observed between the Demonstration and No Demonstration groups in the *Stay Independent* questionnaire scores (Welch's t-Test: $t=-2.93$, $p=0.01$; Chi-Square: $x^2=7.37$, $p=0.01$). This observation likely stemmed from uncontrolled intrinsic participant characteristics, with the Demonstration Group reporting a higher number of self-identified fall risk factors—such as previous falls, balance issues, or concerns about walking stability—compared to the No Demonstration Group. This appears to be related more to the participants' self-perceived vulnerabilities rather than the instructional method alone.

Discussion

The final field validation study underscored the final Balance Assessment feature's robustness across both supervised and unsupervised environments, confirming its reliability for remote use in each of the four sub-tasks. Although differences in inter-subject variance were observed—particularly in unsupervised attempts and among participants in the No Demonstration Group compared to those who received provider-led demonstrations—these variances did not reach statistical significance. The feature consistently delivered accurate results, with no falls or safety incidents reported, reinforcing the safety of its use in unsupervised settings.

While task performance was largely consistent, a few participants encountered difficulties accessing and using the feature independently at home, suggesting that additional instructional support may benefit first-time users. These findings highlight the potential advantage of the hearing care professional's guidance guidance during initial use, especially for users less familiar with digital tools.

Table 1. Summary of statistical tests (Welch's t-Test and Chi-Square) comparing supervised and unsupervised attempts, as well as Demonstration vs No Demonstration groups across Balance Scores, Chair Stand Scores, TUG Scores, and Stay Independent Questionnaire Scores. Significant differences were only found in the Stay Independent questionnaire for the Demonstration vs No Demonstration group comparison.

Measure	Comparison	Test	Test Statistic	p-value	Significant Difference? (p≤0.05)
Balance Scores	Unsupervised vs Supervised Score	Welch's t-Test	t = -1.05	0.3	No
	Unsupervised vs Supervised Outcome	Chi-Square	$x^2 = 0.04$	0.85	No
	Demonstration vs No Demonstration Score	Welch's t-Test	t = 1.35	0.19	No
	Demonstration vs No Demonstration Outcome	Chi-Square	$x^2 = 1.99$	0.16	No
Chair Stand Scores	Unsupervised vs Supervised Score	Welch's t-Test	t = 1.06	0.95	No
	Unsupervised vs Supervised Outcome	Chi-Square	$x^2 = 0.05$	0.83	No
	Demonstration vs No Demonstration Score	Welch's t-Test	t = 1.71	0.9	No
	Demonstration vs No Demonstration Outcome	Chi-Square	$x^2 = 0.96$	0.33	No
TUG Scores	Unsupervised vs Supervised Score	Welch's t-Test	t = 1.70	0.1	No
	Unsupervised vs Supervised Outcome	Chi-Square	$x^2 = 0.56$	0.45	No
	Demonstration vs No Demonstration Score	Welch's t-Test	t = 1.47	0.15	No
	Demonstration vs No Demonstration Outcome	Chi-Square	$x^2 = 1.39$	0.24	No
Stay Independent Questionnaire	Unsupervised vs Supervised Score	Welch's t-Test	t = -0.24	0.81	No
	Unsupervised vs Supervised Outcome	Chi-Square	$x^2 = 0.02$	0.9	No
	Demonstration vs No Demonstration Score	Welch's t-Test	t = -2.93	0.01	Yes
	Demonstration vs No Demonstration Outcome	Chi-Square	$x^2 = 7.37$	0.01	Yes

Summary

The Balance Assessment feature within the My Starkey app represents a significant advancement towards improving and managing balance in older adults with hearing loss. By integrating the CDC's STEADI initiative with advanced sensor technology and AI, this feature helps hearing aid users easily track and maintain good balance through personalized, user-friendly assessments that measure their gait, strength, and balance abilities. These assessments can promote healthier lifestyle choices by raising awareness of modifiable risk factors for falls, supporting greater independence and ultimately contributing to a reduction in fall risk.

Frequent assessments may empower users to detect changes in their balance earlier, enabling them to take timely action and address potential issues before they worsen. For those who may not receive regular assessments from healthcare providers, this feature offers a reliable and accessible tool to monitor balance from the comfort of their home, keeping them actively engaged in managing their health. Collectively, these benefits help users stay aware of their balance health and make informed decisions that may reduce their risk of falling.

The studies presented in this white paper confirm that the Balance Assessment feature is highly effective for remote use, demonstrating both safety and reliability even in unsupervised settings. However, hearing care professional involvement remains essential in helping users maximize the feature's full potential, providing guidance to ensure accurate assessments and facilitating timely conversations about interventions when balance concerns arise. This partnership between technology and hearing care professionals ensures that users not only benefit from the My Starkey app's capabilities but also receive the necessary support to manage their balance proactively and effectively.

Acknowledgements

We would like to extend our gratitude to Matthew Fitzgerald, Robert K. Jackler, Bryn Griswold, Soumya Venkitakrishnan, Deborah M. Kado, David A. Fabry, Achintya K. Bhowmik, Amit Shahar, Matan Sivan, Roy Rozenman, Andy Lin and Archelle Georgiou for their invaluable contributions to the development of the Balance Assessment feature, and to the thoughtful design and execution of the research studies presented in this white paper.

References

1. Agmon, M., Lavie, L., & Doumas, M. (2017). The Association between Hearing Loss, Postural Control, and Mobility in Older Adults: A Systematic Review. *Journal of the American Academy of Audiology*, 28(6), 575–588. https://doi.org/10.3766/jaaa.16044
2. Burwinkel, J. R., Sivan, M., Rozenman, R., & Georgiou, A. (2022, April). Preliminary evaluation of ear-wearable postural stability monitoring technology. EU Falls Festival, Leuven, Belgium.
3. Centers for Medicare & Medicaid Services. (2017). Beneficiaries Utilizing Free Preventive Services by State, 2016. Centers for Medicare & Medicaid Services. https://downloads.cms.gov/files/Beneficiaries%20Utilizing%20Free%20Preventive%20Services%20by%20State%20YTD%202016.pdf
4. Criter, R. E., & Honaker, J. A. (2016). Audiology patient fall statistics and risk factors compared to non-audiology patients. *International Journal of Audiology*, 55(10), 564–570. https://doi.org/10.1080/14992027.2016.1193235
5. Eckstrom, E., Parker, E. M., Lambert, G. H., Winkler, G., Dowler, D., & Casey, C. M. (2017). Implementing STEADI in Academic Primary Care to Address Older Adult Fall Risk. *Innovation in Aging*, 1(2), igx028.
6. Howland, J., Hackman, H., Taylor, A., O'Hara, K., Liu, J., & Brusch, J. (2018). Older adult fall prevention practices among primary care providers at accountable care organizations: A pilot study. PLoS One, 13(10), e0205279. https://doi.org/10.1371/journal.pone.0205279
7. Lin, F. R., & Ferrucci, L. (2012). Hearing Loss and Falls Among Older Adults in the United States. *Archives of Internal Medicine*, 172(4), 369. https://doi.org/10.1001/archinternmed.2011.728
8. Patterson, J. N., & Honaker, J. A. (2014). Survey of Audiologists' Views on Risk of Falling Assessment in the Clinic. *Journal of the American Academy of Audiology*, 25(4), 388–404. https://doi.org/10.3766/jaaa.25.4.10

Author Biographies

Justin R. Burwinkel, Au.D., is a Senior Research Audiologist at Starkey. Dr. Burwinkel's research has contributed to the development of various non-audio sensor applications for use with hearing instruments. He has been granted numerous patents for inventions related to fall risk management, artificial intelligence applications in audiology, and advancements in hearing aid connectivity. Dr. Burwinkel earned both his B.S. and Au.D. at the University of Cincinnati. He currently teaches as an Adjunct Professor in the Salus at Drexel University Osborne College of Audiology and is a collaborator with the University of Cincinnati's FETCHLAB.

Kristen K. Steenerson, MD is a board-certified neurologist with fellowship training in vestibular neurology. She graduated cum laude from Claremont McKenna College, received her MD at the University of Utah and completed residency at Mayo Clinic in Arizona in neurology. There, she discovered the unmet need in balance disorders and vertigo, motivating her to pursue a fellowship in otoneurology at Barrow Neurological Institute. She joined Stanford with positions in both Otolaryngology-Head and Neck Surgery and Neurology with the goal of jointly addressing the junction of inner ear and brain disorders.

Majd Srour is a Senior Software Engineer II at Starkey, where he has spent over six years in the Advanced Development department, focusing on developing proof-of-concepts, prototypes, and new technologies that are now integrated into the company's hearing aids. Prior to joining Starkey, Majd gained valuable experience working for startups and corporations, including Intel Corporation.

Christy Cloninger, Ph.D., is a Senior Human Factors Engineer at Starkey. Dr. Cloninger earned her B.A. in Psychology from Grinnell College and her M.S. and Ph.D. in Neurobiology and Anatomy from the University of Rochester Medical Center. She additionally has 8 years of experience in the field of human factors, conducting formative and summative studies to ensure safe and effective use of both medical devices and consumer products.

Chris Howes is a Principal Software Product Manager at Starkey. Chris has worked at Starkey for over 25 years across all areas of product research and development. His primary focus is the design and development of mobile and cloud software for all aspects of hearing aid interaction. Chris holds multiple granted patents across a wide range of concepts and was awarded Starkey Inventor of the year in 2016.