

Développement et validation de la fonction d'évaluation de l'équilibre de Starkey



Justin R. Burwinkel¹, Kristen K. Steenerson², Majd Srour¹, Christy Cloninger¹ et Chris Howes¹

Contexte

La relation entre troubles de l'audition et risque accru de chute est bien documentée dans la littérature scientifique. Une étude a révélé que les patients traités dans les cliniques d'audiologie, notamment, couraient un plus grand risque de chute que les autres personnes de leur âge (*Criter & Honaker, 2016*). *Lin & Ferrucci (2012)* ont également fait état d'une association importante entre la gravité de la déficience auditive et les déclarations de chutes, même après ajustement des données démographiques, cardiovasculaires et de la fonction d'équilibre vestibulaire.

Le risque de chute est considéré comme multifactoriel, et de nombreux mécanismes comportementaux, physiologiques et pathologiques pourraient expliquer les liens signalés entre déficience auditive et chutes, notamment : pathologie vestibulaire, neuronale ou cardiovasculaire comorbide ; influences génétiques ; diminution de la conscience de l'environnement auditif ; partage des ressources attentionnelles entre la locomotion et le maintien de l'équilibre postural ; fragilité, pouvant être exacerbée par l'isolement social, la dépression, l'inactivité physique et des troubles cognitifs (voir l'analyse d'Agmon et al.).

Malgré les liens bien établis entre âge avancé et déficience auditive d'une part et risque de chute d'autre part, les évaluations régulières du risque de chute ne sont pas effectuées de manière systématique dans les cliniques auditives et autres établissements de soins de santé fréquentés par les personnes âgées (*Howland et al., 2018 ; Patterson & Honaker, 2014*). Aux États-Unis, le dépistage du risque de chute est largement pris en charge par les prestataires de soins primaires ou les gérontologues dans le cadre des visites annuelles de bien-être de Medicare.

Ces visites sont cependant insuffisamment utilisées dans le système de santé, dans la mesure où moins d'un Américain âgé sur cinq bénéficierait chaque année de ces services gratuits (*Centers for Medicare & Medicaid Services, 2017*). Puisque les audioprothésistes voient souvent des patients âgés qui présentent un risque accru de chute, ils sont bien placés pour intégrer les dépistages et les évaluations du risque de chute dans leurs pratiques courantes. Ces tâches supplémentaires s'inscrivent dans une approche plus globale des soins aux patients et constituent une occasion d'améliorer sensiblement les résultats pour les patients, d'autant plus qu'il est possible qu'aucune évaluation du risque de chute n'ait été effectuée ailleurs avec le patient.

En outre, une étude portant sur les prestataires de soins primaires a révélé que si 96 % des prestataires conviennent qu'une évaluation du risque de chute devrait être effectuée pour toutes les personnes âgées, seuls 14 % de ces professionnels connaissent le kit d'outils STEADI des Centers for Disease Control and Prevention (CDC) aux États-Unis. Ce kit fournit un cadre pratique pour appliquer efficacement les directives collaboratives de prévention des chutes publiées par l'American Geriatrics Society (AGS) et la British Geriatrics Society (BGS) (*Eckstrom et al., 2017*). Au vu de ce défi, le besoin d'outils accessibles permettant de combler les lacunes dans la gestion du risque de chute se fait cruellement sentir, surtout dans les contextes où des évaluations régulières et complètes ne sont pas effectuées de manière systématique.

¹Starkey

²École de médecine de l'université Stanford, département d'otorhinolaryngologie et de chirurgie de la tête et du cou

La fonction d'évaluation de l'équilibre de l'application mobile My Starkey répond à ce besoin en offrant un moyen convivial d'évaluer régulièrement le risque de chute et de surveiller des aspects clés de l'équilibre, de la démarche et de la force. Intégrant une technologie de capteurs avancée, l'intelligence artificielle (IA) et les principes fondamentaux de l'initiative STEADI des CDC, elle fournit aux cliniciens un outil fiable pour appuyer la gestion proactive du risque de chute entre les consultations. Une auto-évaluation régulière permet aux gens d'être plus conscients de l'état de leur équilibre et de remédier aux facteurs de risque modifiables avant que la situation ne s'aggrave. Cette démarche peut conduire à une amélioration des résultats pour les patients et à une diminution du nombre de blessures liées aux chutes.

La fonction d'évaluation de l'équilibre en bref

La fonction d'évaluation de l'équilibre de l'application My Starkey est un outil révolutionnaire conçu pour aider les utilisateurs d'aides auditives à surveiller les changements de leur état d'équilibre et à comprendre comment les facteurs de risque de chute modifiables peuvent les affecter personnellement. S'appuyant sur des directives médicales internationalement reconnues en matière de prévention des chutes, elle respecte scrupuleusement les protocoles de dépistage et d'évaluation fonctionnelle recommandés par les CDC. Cette fonction nécessite en premier lieu de remplir une version électronique du questionnaire de dépistage en 12 points de l'initiative STEADI, intitulé *Stay Independent*. La fonction exploite ensuite les capteurs de mouvement intégrés et l'intelligence artificielle des aides auditives Starkey Edge AI pour évaluer les performances de l'utilisateur au cours d'évaluations fonctionnelles.

Au fur et à mesure que les utilisateurs accomplissent les différentes tâches, les données du capteur de mouvement en temps réel des aides auditives sont évaluées par des algorithmes d'intelligence artificielle qui attribuent de manière autonome des notes aux évaluations et affichent les résultats de manière pertinente. Ces algorithmes sont conçus pour évaluer les mouvements des utilisateurs au cours des tests au même titre qu'un clinicien formé. Ils ont été validés de façon indépendante par plusieurs études menées au département d'otorhinolaryngologie et de chirurgie de la tête et du cou de Stanford Medicine, comme le détaille le présent article scientifique.

C'est une première : les aides auditives Edge AI peuvent évaluer de manière autonome les performances des utilisateurs au cours des tests couramment recommandés pour évaluer la démarche, la force et l'équilibre fonctionnels des personnes âgées. Cette approche innovante permet d'identifier les utilisateurs susceptibles de présenter des troubles de l'équilibre ou qui ont des inquiétudes justifiant un examen plus approfondi. L'interface conviviale, développée en tenant compte des commentaires de plusieurs groupes d'utilisateurs et d'experts cliniques, guide efficacement les utilisateurs durant toutes les phases de l'évaluation en combinant des instructions écrites et orales, ainsi que des illustrations.

Études de recherche

La fonction d'évaluation de l'équilibre a été développée dans le cadre d'une collaboration de plusieurs années entre Starkey et Stanford Medicine, comme l'illustre la figure 1 (*en haut de la page suivante*). Dans un premier temps, Starkey a recueilli des données d'entraînement auprès de 71 participants âgés de 47 à 99 ans (âge moyen = 79,2 ans, écart-type = 9,9) en vue de créer des algorithmes destinés à évaluer de manière autonome les tests de marche, de force et d'équilibre à partir des données des capteurs de mouvement des appareils auditifs utilisés pour ces recherches (*Burwinkel et al., 2022*). Une application de collecte de données en aveugle a ensuite été développée, ce qui a permis aux chercheurs cliniques de Stanford Medicine de noter manuellement chaque test selon les protocoles d'évaluation de l'initiative STEADI, indépendamment de la notation automatisée de l'application. Par la suite, la précision de ces algorithmes a été évaluée en comparant les notes attribuées par les cliniciens avec les résultats automatisés de l'application.

Sur la base de ces évaluations, Starkey a affiné les algorithmes de manière itérative et a réalisé des études formatives d'aptitude à l'utilisation à l'aide de versions prototypes de la fonction. Stanford Medicine a étudié plus en profondeur la faisabilité des évaluations à distance et a comparé les performances de la fonction d'évaluation de l'équilibre pour les tests réalisés en face à face, les tests supervisés à distance et les tests à domicile réalisés de manière autonome. La dernière étape du développement incluait une étude de validation sur le terrain, qui a comparé les performances des participants lors d'essais supervisés en laboratoire à celles d'une utilisation non supervisée à domicile, et qui a examiné l'impact des démonstrations de la fonction par un clinicien sur les performances de l'utilisateur.

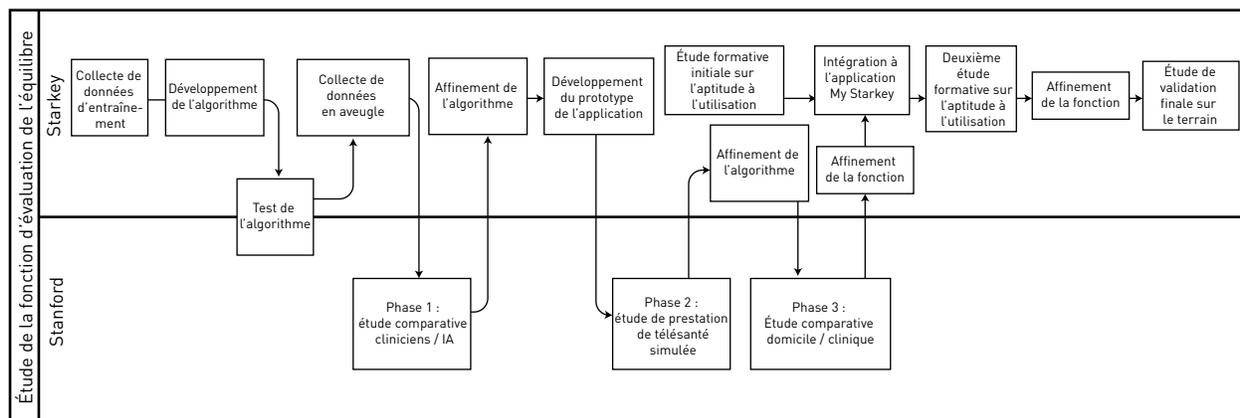


Figure 1. Processus de recherche et développement de la fonction d'évaluation de l'équilibre, une collaboration entre Starkey et Stanford Medicine. Ce processus comprenait le développement d'algorithmes, la validation des modalités cliniques et administratives, des études d'aptitude à l'utilisation et une validation finale sur le terrain comparant les essais supervisés en laboratoire à l'utilisation non supervisée à domicile.

Étude de Stanford Medicine

Stanford Medicine a mené une étude en trois phases pour évaluer la précision et la fiabilité des algorithmes de l'évaluation de l'équilibre (pour divers facteurs de risque de chute) par comparaison avec les évaluations réalisées par des cliniciens expérimentés à l'aide de la batterie de tests fonctionnels de marche, de force et d'équilibre de l'initiative STEADI. Les protocoles d'évaluation de l'initiative STEADI comprennent l'administration d'un test d'équilibre en 4 étapes, d'un test consistant à se lever le plus de fois possible d'une chaise pendant 30 secondes (« 30 secondes assis-debout ») et d'un test chronométré de lever suivi de marche (« TUG »). Tous ces tests sont reconnus pour leur validité écologique et leur précision prédictive en matière d'identification du risque de chute chez les personnes âgées.

La phase 1 de l'étude de Stanford Medicine a comparé les observations des cliniciens avec les notes générées de manière autonome par les premières versions des algorithmes de la fonction d'évaluation de l'équilibre. La phase 2 a examiné la faisabilité de l'utilisation de cette fonction dans le contexte de la télésanté. La phase 3 a évalué l'efficacité des tests à domicile non supervisés et les performances du prototype de recherche de la fonction d'évaluation de l'équilibre en dehors d'un environnement de laboratoire contrôlé.

Phase 1

La première phase de l'étude a adopté une approche comparative en aveugle pour évaluer le risque de chute, en s'appuyant à la fois sur les évaluations des cliniciens et sur les évaluations autonomes provenant d'aides auditives équipées de capteurs de mouvement et de technologies d'intelligence artificielle.

Cette étude a été menée dans un centre de soins tertiaires de référence et a porté sur 250 participants âgés de 55 à 100 ans identifiés comme présentant un risque de chute accru en raison de leur âge avancé et/ou d'antécédents d'instabilité ou de chutes (âge moyen = 78,4 ans, écart-type = 9,6).

Méthodologie

Les participants ont été soumis à la série de tests fonctionnels de marche, de force et d'équilibre de l'initiative STEADI des CDC, tout en portant des aides auditives bilatérales équipées de capteurs de mouvement. Chaque essai a été interprété de manière indépendante par les algorithmes des aides auditives et par trois cliniciens : un qui était présent pendant les essais et deux qui ont visionné plus tard les enregistrements vidéo de ces mêmes essais. Les cliniciens n'étaient pas été informés de la notation autonome des aides auditives. Les principaux résultats mesurés incluaient les notes calculées par les algorithmes et les notes établies manuellement par les trois cliniciens à partir de leurs observations du test d'équilibre en 4 étapes, du test consistant à se lever le plus de fois possible d'une chaise pendant 30 secondes (« 30 secondes assis-debout ») et du test chronométré de lever suivi de marche (« TUG »).

Résultats

Globalement, les chercheurs ont observé une bonne concordance entre les premiers algorithmes d'évaluation de l'équilibre et les cliniciens, pour l'évaluation du risque de chute selon les protocoles de l'initiative STEADI. Les interprétations des cliniciens et les algorithmes de la fonction d'évaluation de l'équilibre n'ont pas présenté de différences statistiquement significatives pour le test d'équilibre en 4 étapes et le test TUG ($p > 0,05$).

Cependant, une différence significative a été constatée pour le test 30 secondes assis-debout ($t = 10,13$, $p < 0,05$), avec une différence moyenne de -0,8 lever. La corrélation des évaluations entre les différents cliniciens était excellente.

Discussion

Alors que les premiers algorithmes de la fonction d'évaluation de l'équilibre ont globalement produit des notes correctes, la sous-évaluation systématique du test 30 secondes assis-debout a mis en évidence une divergence que les chercheurs ont attribuée à la précocité avec laquelle la fonction d'évaluation de l'équilibre appliquait les directives des CDC en matière de notation. Plus précisément, les CDC recommandent aux cliniciens de compter les levers de chaise effectués avant la fin de la durée du test comme des levers complets, ce qui n'a pas été appliqué dans un premier temps par l'algorithme de notation de l'aide auditive pour ce test. Cette constatation a conduit à affiner l'algorithme pour les études ultérieures, afin de mieux aligner la notation de la fonction d'évaluation de l'équilibre avec les critères des CDC.

Phase 2

La phase 2 de l'étude de Stanford Medicine a étendu les recherches de la phase 1 en évaluant le risque de chute à l'aide des protocoles d'évaluation de l'initiative STEADI dans un cadre de télésanté. Cette phase s'est focalisée sur la faisabilité de l'administration à distance des tests fonctionnels de marche, de force et d'équilibre. Cette phase s'est déroulée dans le même centre de soins tertiaires de référence, avec la même méthode de collecte de données en aveugle que la phase 1. Elle a été menée auprès de 50 participants âgés de 57 à 98 ans qui présentaient un risque élevé de chute en raison de leur âge avancé ou d'antécédents d'instabilité ou de chutes (âge moyen = 78,2 ans, =écart-type = 8,1).

Méthodologie

Au cours de cette deuxième phase, les participants ont été soumis à la série de tests d'évaluation fonctionnelle de l'initiative STEADI tout en portant des aides auditives bilatérales dotées de capteurs de mouvement. À la différence de la phase 1, les tests ont été administrés à distance, un clinicien assurant la direction et la supervision via un appel vidéo sécurisé.

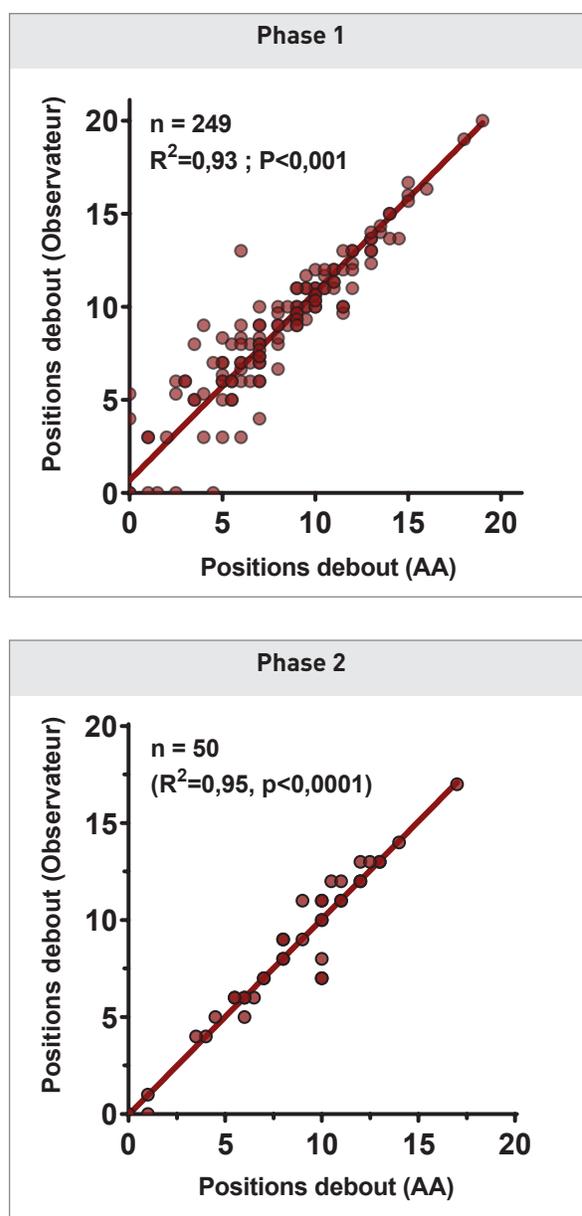


Figure 2. Comparaison des performances de l'algorithme de la fonction d'évaluation de l'équilibre lors de deux phases du test « 30 secondes assis-debout ». Le graphique du haut représente les données de la phase 1 ($n=249$, $r^2=0,93$; $p<0,001$) et le graphique du bas les données de la phase 2 ($n=50$, $r^2=0,95$; $p<0,0001$). L'axe des y indique le nombre total de positions debout observées par les cliniciens, tandis que l'axe des x indique le nombre total de positions debout enregistrées par la fonction d'évaluation de l'équilibre. L'alignement plus marqué dans la phase 2 indique une amélioration de la précision de l'algorithme.

Un chercheur était présent dans le laboratoire pour surveiller chaque participant sur le plan de la sécurité, mais sans participer à la direction ou à l'administration des tests. Chaque participant a effectué un essai pour chaque tâche, qui a été évalué par les algorithmes des aides auditives et noté par un clinicien disposant d'un accès aux enregistrements vidéo des essais. Ce dernier n'a pas été informé de la notation des aides auditives, ce qui a permis une comparaison impartiale entre les résultats de l'algorithme et sa propre évaluation. Les principaux résultats mesurés incluaient les notes calculées par les algorithmes et les notes issues de l'observation par le clinicien du test d'équilibre en 4 étapes, du test 30 secondes assis-debout et du test TUG.

Résultats

Les résultats de la phase 2 étaient assez cohérents avec ceux de la phase 1, indiquant une bonne concordance globale entre les algorithmes de la fonction d'évaluation de l'équilibre et le clinicien pour l'évaluation des performances lors des tests de marche, de force et d'équilibre selon les protocoles d'évaluation de l'initiative STEADI.

Test d'équilibre en 4 étapes : Une concordance légèrement meilleure entre les algorithmes de la fonction d'évaluation de l'équilibre et le clinicien a été constatée pour les postures d'équilibre pieds joints, pieds serrés décalés d'une demi-longueur et sur un seul pied, tandis que la posture d'équilibre pieds sur une ligne a révélé une moins bonne concordance. Ces différences sont relativement faibles et peuvent être attribuées à une augmentation d'écart en raison de la taille plus réduite de l'échantillon utilisé dans la phase 2 par rapport à la phase 1.

Test 30 secondes assis-debout : Pour la notation du test 30 secondes assis-debout, l'algorithme de la fonction d'évaluation de l'équilibre a présenté une amélioration significative entre la phase 1 ($n=249$, $r^2=0,93$; $p<0,001$) et la phase 2 ($n=50$, $r^2=0,95$; $p<0,0001$), comme le montre la figure 2 (voir page précédente). À la phase 2, la différence entre les notes calculées par l'algorithme et les observations du clinicien n'était plus statistiquement significative ($n=50$, $t=0,24$, $p>0,05$).

Ce résultat témoigne de l'efficacité des ajustements apportés à l'algorithme entre la phase 1 et la phase 2, qui ont permis d'améliorer la concordance entre l'algorithme et les comptages du clinicien.

Test TUG : Au cours de la phase 1, les premiers algorithmes de la fonction d'évaluation précoce de l'équilibre n'ont pas présenté de différence significative par rapport aux évaluations du test TUG par les cliniciens ($n=220$, $t=0,94$; $p>0,05$). Cependant, au cours de la phase 2, les algorithmes ont présenté une différence significative ($n=48$, $t=2,51$; $p<0,05$), avec des temps généralement plus lents pour le test TUG. Cette différence suggère que pendant la phase 2, l'administration à distance du test a pu introduire une variabilité ou des difficultés supplémentaires qui ont affecté les performances des algorithmes par rapport à l'observation directe par le clinicien. Cet effet semble toutefois avoir principalement influencé la spécificité du test (sa capacité à exclure les personnes qui ne tombent pas) plutôt que sa sensibilité (sa capacité à détecter les personnes susceptibles de tomber), ce qui montre que la première version de la fonction d'évaluation de l'équilibre constitue un outil fiable pour évaluer les performances d'un utilisateur lors d'un test TUG et qu'elle mérite d'être étudiée davantage, malgré les différences observées à la figure 3 (voir page suivante).

Discussion

La phase 2 de l'étude de Stanford Medicine a prouvé la faisabilité de l'évaluation à distance de la démarche, de la force et de l'équilibre chez les personnes âgées exposées au risque de chute. De plus, l'administration à distance de ces tâches s'est avérée sûre, aucune chute ne s'étant produite pendant que les participants effectuaient les tests de marche, de force et d'équilibre.

Comme pour la phase 1, la précision de la notation à distance par les algorithmes était généralement élevée. Pour la plupart des tâches, la concordance entre les algorithmes de la fonction d'évaluation de l'équilibre et les observations du clinicien était forte.

La précision du test de résistance, notamment, a été améliorée entre la phase 1 et la phase 2 grâce aux affinements de l'algorithme réalisés entre les études de la phase 1 et de la phase 2. Des écarts plus importants ont toutefois été observés pour les exercices d'équilibre et le test de marche, probablement en raison de la taille réduite de l'échantillon et des difficultés potentielles liées à l'utilisation de tests à distance. Ces constatations ont fait ressortir la nécessité de fournir des instructions claires, visuellement et verbalement.

Dans l'ensemble, la phase 2 a montré que l'évaluation à distance de la marche, de la force et de l'équilibre pouvait être précise et sûre, ce qui constitue une base solide pour améliorer la version initiale de la fonction d'évaluation de l'équilibre.

Phase 3

La troisième phase de l'étude de Stanford Medicine avait pour but d'évaluer le risque de chute en appliquant les protocoles de l'initiative STEADI dans l'environnement domestique des participants, en utilisant des aides auditives bilatérales équipées de capteurs de mouvement et de technologies d'intelligence artificielle. Cette étude a porté sur 50 participants âgés de 56 à 97 ans (âge moyen = 76,0 ans, écart-type = 8,3 ans), tous identifiés comme exposés à un risque accru de chute sur la base de critères autodéclarés issus des critères de dépistage *Stay Independent* de l'initiative STEADI.

Méthodologie

Les participants ont effectué trois essais à leur domicile : un essai d'apprentissage suivi de deux essais d'évaluation. Chaque essai incluait le test d'équilibre en 4 étapes, le test 30 secondes assis-debout et le test TUG, les évaluations étant effectuées au moyen des technologies intégrées dans les aides auditives.

Résultats

Les tests à domicile ont révélé une concordance variable d'un essai à l'autre pour le test d'équilibre en 4 étapes, allant de 76 % à 87 % pour les trois positions debout mentionnées dans les critères de réussite/d'échec de l'initiative STEADI, après suppression des essais pour lesquels aucune note n'avait été attribuée en raison d'une erreur de l'utilisateur avec la version de recherche de l'application.

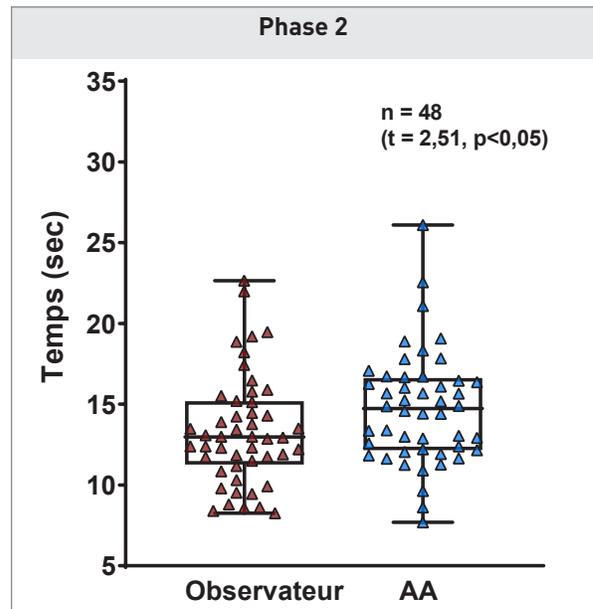
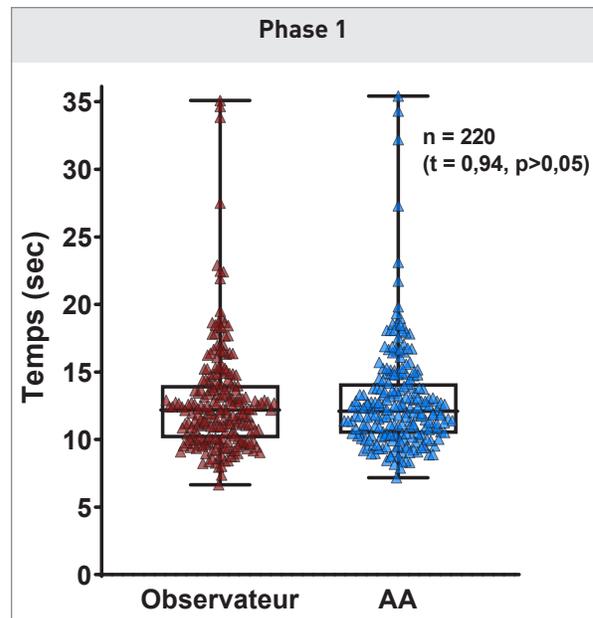


Figure 3. Comparaison des temps du test TUG enregistrés par les cliniciens [Observateur] et les algorithmes de la fonction d'évaluation de l'équilibre [AA] pendant la phase 1 (à gauche) et la phase 2 (à droite). Les diagrammes en boîte représentent la distribution des temps pour le test TUG, mesurés en secondes, avec des points de données dispersés. Les diagrammes en boîte mettent en évidence la tendance centrale et la variabilité des temps pour les deux phases du test TUG. Une différence significative est observée dans les performances de l'algorithme au cours de la phase 2 par rapport à l'alignement plus étroit constaté au cours de la phase 1.

Le test 30 secondes assis-debout a présenté initialement une fiabilité modérée entre test et répétition du test. Des erreurs d'interprétation similaires se sont toutefois produites lorsque l'une des tentatives du participant aboutissait à zéro lever alors qu'une autre tentative aboutissait à une note. L'ajustement de ces erreurs d'interprétation a révélé une grande fiabilité ($n=41$, $R2=0,83$, $p < 0,001$) pour le test 30 secondes assis-debout. Le test TUG a lui aussi montré une fiabilité modérée, avec des erreurs d'interprétation pour environ 20 % des essais. Après correction des erreurs d'interprétation, la fiabilité s'est améliorée de manière significative ($n=40$, $R2=0,84$, $p < 0,001$) pour le test TUG.

Discussion

L'étude de la phase 3 a mis en évidence des difficultés notables dans la conduite des évaluations à domicile du risque de chute à l'aide d'une application mobile. Les chercheurs ont qualifié d'« erreurs d'interprétation » un pourcentage important d'essais, ce qui indique que les utilisateurs ont eu des difficultés avec l'application et que certains essais n'ont pas été pris en compte dans les résultats. Des rapports anecdotiques des participants ont en outre mis en évidence des problèmes de facilité d'utilisation de la version de recherche de l'application mobile.

Malgré ces difficultés, l'étude a montré que les protocoles d'évaluation de l'initiative STEADI pouvaient être appliqués à domicile en toute sécurité avec des aides auditives. Point important, aucun cas de chute ou de trouble de l'équilibre n'a été signalé lors des évaluations auto-administrées, ce qui suggère que le protocole est adapté à une utilisation à domicile par cette population.

Tests formatifs d'aptitude à l'utilisation

La fonction d'évaluation de l'équilibre a été mise au point dans le cadre d'un processus de codéveloppement étendu impliquant un retour d'information itératif de la part des utilisateurs et des cliniciens. Starkey avait pour objectif de créer un outil convivial, efficace et fiable, qui permette aux utilisateurs d'aides auditives d'évaluer leur équilibre et de conserver leur indépendance. Ce processus a comporté plusieurs étapes de tests formatifs d'aptitude à l'utilisation, en plus des tests de performance des algorithmes, afin d'affiner la conception et le fonctionnement de la fonction d'après les avis et les expériences des utilisateurs.

Méthodologie

Avant son intégration à l'application My Starkey, la conception de la fonction d'évaluation de l'équilibre – y compris l'interface utilisateur, le flux de navigation et le contenu pédagogique – a été inspirée par des observations directes et des entretiens structurés avec des utilisateurs potentiels. Les premiers tests d'aptitude à l'utilisation visaient à comprendre les préférences des utilisateurs, à identifier les obstacles à l'accomplissement des tâches et à vérifier que des instructions claires étaient fournies aux utilisateurs.

Suite au développement d'un prototype de l'application, le département d'otorhinolaryngologie et de chirurgie de la tête et du cou de Stanford Medicine a mené des études d'évaluation de l'efficacité de la version initiale de la fonction dans des environnements de télésanté et à domicile. Ces études, menées avec des aides auditives et le prototype de la fonction d'évaluation de l'équilibre en complément des protocoles de l'initiative STEADI, ont identifié plusieurs aspects susceptibles d'être améliorés. Ces améliorations possibles comprenaient la réduction des erreurs d'utilisation qui avaient abouti à des essais non notés et l'amélioration de la gestion de la notation des essais partiels. Des problèmes tels que les erreurs d'interprétation dues à certains comportements de l'utilisateur, l'application sur smartphone au début des tests d'équilibre, ont également été identifiés, donnant lieu à des ajustements de l'interface utilisateur et à des affinements de l'algorithme dans le but d'améliorer la robustesse et la gestion des erreurs.

Phase 4

Une autre série de tests formatifs d'aptitude à l'utilisation a été menée avec une version candidate de la fonction d'évaluation de l'équilibre intégrée à l'application My Starkey. Cette phase a porté sur un groupe diversifié de 15 utilisateurs d'aides auditives présentant différents niveaux d'expérience avec les applications pour smartphones. Les participants se sont servis de la fonction d'évaluation de l'équilibre dans un environnement simulant un salon pour remplir une version électronique du questionnaire *Stay Independent* et effectuer des tests fonctionnels de marche, de force et d'équilibre, tout en portant des aides auditives Edge AI.

Ils ont également reçu un kit d'outils utilisables pour mesurer le chemin à parcourir pour le test TUG. Tout au long de l'étude, ils ont été surveillés pour voir s'ils effectuaient les exercices correctement.

Résultats

Les retours d'information faisant suite à la deuxième série de tests formatifs d'aptitude à l'utilisation ont révélé une nette amélioration de l'expérience utilisateur. La plupart des participants ont trouvé les instructions faciles à suivre et ont compris les résultats, et beaucoup ont fait part de leur souhait d'utiliser cette fonction chez eux. Cependant, trois aspects clés ont été identifiés comme devant être améliorés :

1. Les participants auraient souhaité davantage de contexte concernant l'objectif de chaque exercice et les mesures effectuées.
2. Les instructions illustrées fournies pour chaque exercice n'ont pas toujours été remarquées par les utilisateurs.
3. Les instructions relatives au test 30 secondes assis-debout et au test TUG restaient floues pour certains, ce qui a donné lieu à des malentendus.

Discussion

Les tests formatifs d'aptitude à l'utilisation ont joué un rôle essentiel pour l'affinement de la fonction d'évaluation de l'équilibre. Les problèmes initiaux, par exemple les essais non notés en raison d'erreurs de l'utilisateur, ont été résolus grâce à des ajustements algorithmiques et à des instructions plus claires. Bien que des progrès significatifs aient été accomplis, certaines difficultés d'utilisation ont persisté, notamment la manière dont les instructions ont été communiquées et le contexte fourni pour chaque test. Ces difficultés persistantes ont mis en évidence l'importance de recueillir continuellement les commentaires des utilisateurs afin d'optimiser l'efficacité de la fonction. Ce processus itératif d'amélioration a été essentiel pour le déploiement en toute fiabilité de la fonction d'évaluation de l'équilibre lors de l'étude finale de validation sur le terrain, qui visait à tester la robustesse de la fonction dans des environnements réels non supervisés.

Validation finale sur le terrain

L'étude finale de validation sur le terrain a été conçue pour évaluer l'aptitude à l'utilisation et l'efficacité de la fonction d'évaluation de l'équilibre dans des environnements supervisés et non supervisés.

L'objectif principal de l'étude consistait à déterminer si la fonction produisait des résultats fiables et cohérents dans différents contextes et si les démonstrations effectuées par les cliniciens avaient une incidence sur la facilité d'utilisation ou les performances des participants.

Méthodologie

Quatorze participants ont été recrutés parmi la base de données de participants aux recherches Starkey et ont été divisés en deux groupes : un groupe (n=7) a bénéficié d'une démonstration en personne par un audioprothésiste sur la façon d'utiliser la fonction d'évaluation de l'équilibre, tandis que l'autre groupe (n=7) s'est seulement vu montrer comment accéder à la fonction dans l'application et a été invité à suivre les instructions de manière autonome.

Les deux groupes ont participé à des essais supervisés d'évaluation de l'équilibre dans un environnement de laboratoire. Cela a été le cas pour le groupe « avec démonstration » avant de procéder aux évaluations à domicile, tandis que le groupe « sans démonstration » a effectué les essais supervisés à son retour de l'essai sur le terrain. Il a ainsi été possible d'évaluer si la démonstration préalable influençait les performances des participants et leur compréhension de la fonction lors d'une utilisation à domicile non supervisée.

Fait notable, un participant de chaque groupe (n=2) n'a pas terminé les essais à domicile : l'un d'eux a par erreur visionné des vidéos d'exercices d'équilibre dans une autre section de l'application My Starkey, et l'autre avait oublié comment accéder à la fonction d'évaluation de l'équilibre. Un troisième participant (n=1) a déclaré avoir effectué les essais à domicile, mais les données d'utilisation de l'application étaient manifestement absentes du cloud pour ces essais, ce qui a empêché l'inclusion de ce participant dans l'analyse.

Des données quantitatives ont été recueillies à partir des résultats du test d'équilibre, du test 30 secondes assis-debout et du test TUG, et à partir des réponses des participants au questionnaire *Stay Independent*. Les données des essais supervisés et non supervisés ont été analysées à l'aide du test t de Welch, et les résultats globaux ont été comparés au moyen du test du Khi-deux d'indépendance afin d'évaluer la cohérence et la fiabilité des tentatives d'évaluation de l'équilibre. Les différences de performances entre le groupe avec démonstration (n=5) et le groupe sans démonstration (n=6) ont été analysées de la même manière.

Résultats

Les résultats de l'étude de validation sur le terrain n'ont pas révélé de différences significatives dans les notes obtenues au test d'équilibre, au test 30 secondes assis-debout, ni dans les réponses au questionnaire *Stay Independent* entre les tentatives supervisées et non supervisées, ni entre les groupes avec démonstration et sans démonstration pour la plupart des mesures. Comme le résume le tableau 1 (*voir page suivante*), les analyses statistiques réalisées à l'aide du test t de Welch et des tests du Khi-deux ont révélé des notes et des résultats cohérents pour la plupart des mesures, ce qui indique qu'il n'y a pas de différences notables dans les performances imputables à l'environnement de test ou à la méthode d'instruction. Une différence significative a toutefois été observée entre les groupes avec et sans démonstration dans les notes obtenues au questionnaire *Stay Independent* (test t de Welch : $t=-2,93$, $p=0,01$; Khi-deux : $\chi^2=7,37$, $p=0,01$). Cette constatation découle probablement de caractéristiques intrinsèques imprévues chez les participants, le groupe avec démonstration ayant signalé un plus grand nombre de facteurs de risque de chute auto-identifiés (chutes antérieures, problèmes d'équilibre ou préoccupations relatives à la stabilité de la marche), par rapport au groupe sans démonstration. Cela semble davantage lié aux vulnérabilités perçues par les participants eux-mêmes plutôt qu'à la seule méthode d'instruction.

Discussion

L'étude finale de validation sur le terrain a mis en évidence la robustesse de la version finale de la fonction d'évaluation de l'équilibre dans des environnements supervisés et non supervisés, confirmant sa fiabilité pour une utilisation à distance dans chacune des quatre sous-tâches. Même si des différences d'écart entre les sujets ont été observées, en particulier dans les tentatives non supervisées et parmi les participants du groupe sans démonstration par rapport au groupe avec démonstration, ces écarts n'ont pas atteint le seuil de signification statistique. La fonction a toujours fourni des résultats précis, sans qu'aucune chute ni aucun incident de sécurité n'aient été signalés, ce qui confirme la sécurité de son utilisation dans des environnements non supervisés.

Si les performances ont été globalement homogènes, quelques participants ont éprouvé des difficultés à accéder à la fonction et à l'utiliser de manière autonome chez eux, ce qui suggère qu'un accompagnement pédagogique supplémentaire pourrait être bénéfique pour les nouveaux utilisateurs. Ces conclusions mettent en évidence l'avantage potentiel d'un accompagnement par l'audioprothésiste lors de la première utilisation, en particulier pour les utilisateurs peu familiarisés avec les outils numériques.

Tableau 1. Résumé des tests statistiques (test t de Welch et Khi-deux) comparant les tentatives supervisées et non supervisées, ainsi que les groupes avec démonstration et sans démonstration pour les notes des tests d'équilibre, assis-debout, TUG et les notes du questionnaire Stay Independent. Des différences significatives n'ont été constatées que pour la comparaison entre les groupes avec démonstration et sans démonstration en ce qui concerne le questionnaire Stay Independent.

Mesure	Comparaison	Test	Statistiques du test	Valeur-p	Différence significative ? (p<0,05)
Notes d'équilibre	Note supervisé / non supervisé	Test t de Welch	t = -1,05	0,3	Non
	Résultat supervisé / non supervisé	Khi-deux	$\chi^2 = 0,04$	0,85	Non
	Note avec démo / sans démo	Test t de Welch	t = 1,35	0,19	Non
	Résultat avec démo / sans démo	Khi-deux	$\chi^2 = 1,99$	0,16	Non
Notes assis-debout	Note supervisé / non supervisé	Test t de Welch	t = 1,06	0,95	Non
	Résultat supervisé / non supervisé	Khi-deux	$\chi^2 = 0,05$	0,83	Non
	Note avec démo / sans démo	Test t de Welch	t = 1,71	0,9	Non
	Résultat avec démo / sans démo	Khi-deux	$\chi^2 = 0,96$	0,33	Non
Notes TUG	Note supervisé / non supervisé	Test t de Welch	t = 1,70	0,1	Non
	Résultat supervisé / non supervisé	Khi-deux	$\chi^2 = 0,56$	0,45	Non
	Note avec démo / sans démo	Test t de Welch	t = 1,47	0,15	Non
	Résultat avec démo / sans démo	Khi-deux	$\chi^2 = 1,39$	0,24	Non
Questionnaire Stay Independent	Note supervisé / non supervisé	Test t de Welch	t = -0,24	0,81	Non
	Résultat supervisé / non supervisé	Khi-deux	$\chi^2 = 0,02$	0,9	Non
	Note avec démo / sans démo	Test t de Welch	t = -2,93	0,01	Oui
	Résultat avec démo / sans démo	Khi-deux	$\chi^2 = 7,37$	0,01	Oui

Synthèse

La fonction d'évaluation de l'équilibre incluse dans l'application My Starkey représente une avancée significative dans l'amélioration et la gestion de l'équilibre chez les personnes âgées souffrant de perte auditive. S'appuyant sur l'initiative STEADI des CDC, sur des capteurs sophistiqués et sur l'intelligence artificielle, cette fonction permet aux utilisateurs d'aides auditives de surveiller et de conserver facilement un bon équilibre grâce à des évaluations personnalisées et conviviales qui mesurent leur démarche, leur force et leurs capacités d'équilibre. Ces évaluations contribuent à des choix de vie plus sains en sensibilisant aux facteurs de risque de chute modifiables, en favorisant une plus grande indépendance et en contribuant finalement à une réduction du risque de chute.

Des évaluations fréquentes permettent aux utilisateurs de détecter plus tôt les changements dans leur équilibre, de sorte à pouvoir prendre des mesures opportunes et à répondre aux problèmes potentiels avant qu'ils ne s'aggravent. Les personnes qui ne bénéficient pas d'évaluations régulières de la part de prestataires de soins de santé profitent ainsi d'un outil fiable et accessible leur permettant de contrôler leur équilibre depuis le confort de leur domicile, et de rester activement impliquées dans la gestion de leur santé. Tous ces avantages aident les utilisateurs à connaître l'état de leur équilibre et à prendre des décisions éclairées susceptibles de réduire le risque de chute.

Les études présentées dans cet article scientifique confirment que la fonction d'évaluation de l'équilibre est très efficace dans le cadre d'une utilisation à distance, et qu'elle est à la fois sûre et fiable, même dans des environnements non supervisés. Cependant, l'implication de l'audioprothésiste reste essentielle pour permettre aux utilisateurs d'exploiter au maximum le potentiel de cette fonction, en prodiguant des conseils pour réaliser des évaluations précises et en favorisant des conversations opportunes sur d'éventuelles interventions lorsque des problèmes d'équilibre apparaissent. Ce partenariat entre la technologie et les professionnels de l'audition assure aux utilisateurs qu'ils bénéficieront non seulement des fonctionnalités de l'application My Starkey, mais aussi d'un soutien précieux pour gérer leur équilibre de manière proactive et efficace.

Remerciements

Nous souhaitons exprimer notre gratitude à 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 et Archelle Georgiou pour leur précieuse contribution au développement de la fonction d'évaluation de l'équilibre, ainsi qu'à la conception et à la conduite judicieuses des études de recherche présentées dans cet article scientifique.

Sources

1. Agmon, M., Lavie, L., & Dumas, 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., & Brusck, 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>

Biographies des auteurs



Justin R. Burwinkel, Au.D., est audiologiste chercheur principal chez Starkey. Les recherches du Dr Burwinkel ont contribué au développement de diverses applications de capteurs non acoustiques destinés à une utilisation avec des appareils auditifs. Il est titulaire de nombreux brevets pour des inventions liées à la gestion du risque de chute, aux applications de l'intelligence artificielle en audiologie et aux progrès dans le domaine de la connectivité des aides auditives. Le Dr Burwinkel a obtenu une licence et un doctorat à l'Université de Cincinnati. Il enseigne actuellement en tant que professeur adjoint au Salus at Drexel University Osborne College of Audiology et collabore avec le FETCHLAB de l'Université de Cincinnati.



Kristen K. Steenerson, MD est neurologue certifiée et spécialisée en neurologie vestibulaire. Elle a obtenu son diplôme avec mention au Claremont McKenna College, a validé son doctorat en médecine à l'université de l'Utah et a effectué son internat en neurologie à la Mayo Clinic, en Arizona. C'est là qu'elle a pris conscience des besoins non satisfaits en matière de troubles de l'équilibre et de vertiges, ce qui l'a incitée à suivre une formation en otoneurologie au Barrow Neurological Institute. Elle a rejoint l'université Stanford pour occuper des postes en otorhinolaryngologie et chirurgie de la tête et du cou, ainsi qu'en neurologie, dans le but d'aborder simultanément le lien entre troubles de l'oreille interne et troubles neurologiques.



Majd Srour est ingénieur logiciel senior II chez Starkey, où il a travaillé plus de six ans au sein du département de développement avancé, en se focalisant sur le développement de preuves de concept, de prototypes et de nouvelles technologies qui sont aujourd'hui intégrées dans les aides auditives de l'entreprise. Avant de rejoindre Starkey, il a acquis une expérience précieuse en travaillant pour des start-ups et des grandes entreprises, notamment Intel Corporation.



Christy Cloninger, Ph.D., est ingénieure principale en facteurs humains chez Starkey. Le Dr Cloninger est titulaire d'une licence en psychologie du Grinnell College ainsi que d'une maîtrise et d'un doctorat en neurobiologie et anatomie du centre médical de l'université de Rochester. Elle compte également 8 ans d'expérience dans le domaine des facteurs humains, ayant dirigé des études formatives et sommatives visant à garantir l'utilisation sûre et efficace de dispositifs médicaux et de produits de consommation.



Chris Howes est chef de produit logiciel principal chez Starkey. Il travaille chez Starkey depuis plus de 25 ans et a exploré tous les domaines de la recherche et du développement de produits. Il se consacre principalement à la conception et au développement de logiciels mobiles et cloud pour tous les aspects de l'interaction avec les aides auditives. Il détient plusieurs brevets couvrant un large éventail de concepts et a reçu le prix Starkey de l'inventeur de l'année en 2016.