Development of a novel clinical outcome assessment: digital instrumental activities of daily living

Abbey Sawyer,[a,d,*] Jamie Brannigan,[b,d] Lisa Spielman,[a] BCI Functional Outcome Measures Group,[e] David Putrino,[a] and Adam Fry[a,c]

[a]Department of Rehabilitation and Human Performance, Icahn School of Medicine at Mount Sinai, New York, NY, United States
[b]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom
[c]Department of Medicine, The University of Melbourne, Melbourne, Australia

eBioMedicine
2025;116: 105732
Published Online xxx
https://doi.org/10.1016/j.ebiom.2025.105732

Summary

Background Digital technology is integral to activities of daily living, particularly instrumental activities of daily living (IADLs). However, tools that accommodate digital performance of IADLs are lacking. The aim of this study was to develop a novel Digital IADL Scale.

Methods The multi-stage methodology included: (i) deductive item generation via a systematic review and assignment to domains using a Delphi process, (ii) inductive item generation via a survey of individuals with lived experience (IWLE) of severe paralysis, (iii) item refinement via item rating surveys of content experts and IWLE, and (iv) focus group discussions with key opinion leaders.

Findings The systematic review identified 1250 IADL items from validated IADL measures, of which 353 met criteria. Deduplication reduced the deductive item set to 77, of which 42 remained following the Delphi process. IWLE generated 152 items, of which 132 met criteria. Deduplication reduced the inductive item set to 41. The combined item pool was reduced to 69 following the item rating surveys. Following focus group feedback, a list of nine domains, containing 37 items, and suggested response scale options are presented.

Interpretation We describe the initial development of a scale to assess functional independence within IADLs that may be completed digitally, which will be submitted to further validation.

Funding Support for this project was provided in kind by the Abilities Research Center. No formal funding was received.

Keywords: Activities of daily living; Digital occupations; Functional independence; Brain computer interface; Paralysis

Introduction

Digital technology is inextricable from modern day-to-day living. With the proliferation of digital technology, many of our everyday occupations have evolved such that previously non-digital activities have come to be performed digitally and new inherently digital activities have become essential within daily life. For example, activities such as shopping and banking have transitioned to digital platforms and new, digital forms of information retrieval and communication have become predominant. Thus, digital technology has become integral to activities of daily living (ADLs)—particularly instrumental ADLs (IADLs).[1] To address this modern reality, IADL assessment methods must accommodate for digital performance of IADLs.

Digital performance of IADLs is now standard and oftentimes no longer optional. In the United States (US), 90% of individuals own a smartphone and 84% report being online at least several times per day.[2] More interactions among young adults now occur online than in person[3] and over half of all adults using the internet report doing so to connect with friends and family digitally.[4] Approximately one fifth of Americans now get their groceries online[5] and the majority perform their banking online.[4] Healthcare and business have also undergone a digital transformation, accelerated by the COVID-19 pandemic. Currently, 39% of US adults engage in telehealth consultations annually,[6] and the use of digital technologies to help manage health is more likely in those with greater healthcare needs.[7] Many

*Corresponding author.
E-mail addresses: abbey.sawyer@mountsinai.org, david.putrino@mountsinai.org (A. Sawyer).
[d]Authors contributed equally.
[e]The complete list of BCI Functional Outcome Measures Group are listed in the Supplementary Material.

Research in context

Evidence before this study

Prior research highlights the growing reliance on digital technology for performing instrumental activities of daily living (IADLs), with digital engagement now being a standard, and often essential, component of daily life. Existing IADL assessment tools have attempted to integrate digital activities, yet they remain inadequate for individuals with severe motor impairments who rely on digital technology for functional independence. The need for a more comprehensive measure is further underscored by emerging assistive technologies, such as brain-computer interfaces, which have the potential to restore digital access for individuals with significant physical disabilities.

Added value of this study

This study employed a multi-stage methodology to develop a scale assessing functional independence in digitally performed IADLs. Items were generated through a systematic review and Delphi process (deductive approach) and a survey of individuals with lived experience of severe paralysis (inductive approach). Item refinement involved expert rating surveys and focus group discussions. The process identified 1250 IADL items, narrowed to 353, and further refined to a final set of 37 items across nine domains.

Implications of all the available evidence

The development of the Digital IADL Scale marks a significant advancement in evaluating functional independence through digital technology, especially for individuals with severe motor impairments. However, further validation—through pilot testing and psychometric analysis—is essential to refine the scale and establish its clinical relevance. Future iterations may benefit from a personalised assessment approach, ensuring adaptability as digital technologies evolve. Additionally, determining an optimal scoring method will be crucial for accurately capturing meaningful changes in digital functional performance. Expanding the scale to address the needs of older adults with cognitive decline could further enhance its applicability across diverse populations.

activities related to employment now happen exclusively online.[8]

Digital exclusion is associated with higher functional dependence.[9] Moreover, reliance on digital devices for functional independence is greater for individuals with physical disability.[10] Individuals with paralysis also report significantly higher satisfaction with their quality of life when they use digital devices to support their daily functioning.[11] Occupational therapy (OT) practice guidelines recognise the necessity of digital technology for improving an individual's functional independence in all domains of occupation, for example; texting a caregiver or participating in a telehealth visit may be included within an individual's rehabilitation plan.[12] However, tools explicitly evaluating independence on IADLs performed digitally are lacking.

Previous attempts have been made to incorporate digital activities into IADL scales. For example, the Advanced IADL Scale added Information and Communication Technology as a new item and modified the traditional item of Money Management to include digital payments using smartphone applications.[13] The Amsterdam IADL Questionnaire also added several technology-related items, including using email, operating a smartphone, and using GPS navigation.[14] However, only some of the items within these scales can be completed digitally with current technology. As such, these instruments remain unsuitable for individuals with severe motor impairments, who may achieve meaningful independence through digital technology. The need for a new instrument is further highlighted by the emergence of novel medical technologies such as brain-computer interfaces (BCIs), which might restore control over digital devices for performance of IADLs in individuals with severe motor impairments.[15,16] A digital IADL scale has been proposed as a potential clinical outcome assessment (COA) for these types of BCI devices[17] and has been considered among potential COAs of implantable BCIs by the US Food and Drug Administration (FDA).[18]

We propose that digital ADLs (dADLs) are not a new concept or distinct category of ADLs. Instead, they may be considered any subset of basic ADLs (BADLs) or IADLs that are performed digitally by an individual. This definition reflects modern performance of ADLs and accommodates future technological developments that will inevitably lead to more ADLs being performed digitally by an increasing number of people.[19]

We describe the development of a Digital IADL Scale comprising items organised by domain and a response scale, achieved through a comprehensive, multi-phase process involving both multidisciplinary experts and individuals with lived experience (IWLE) of severe paralysis. Following the approach of existing ADL measures, we aimed to develop an outcome measure based on the distal concept of daily functioning and participation, rather than on proximal capabilities such as the ability to generate a "click" on a digital device.

Methods

Overview

This study comprised a multi-stage methodology (Fig. 1), which aligned with FDA draft guidance on the search and development of fit-for-purpose COAs.[20] An overview of the methods is shown in Fig. 1.

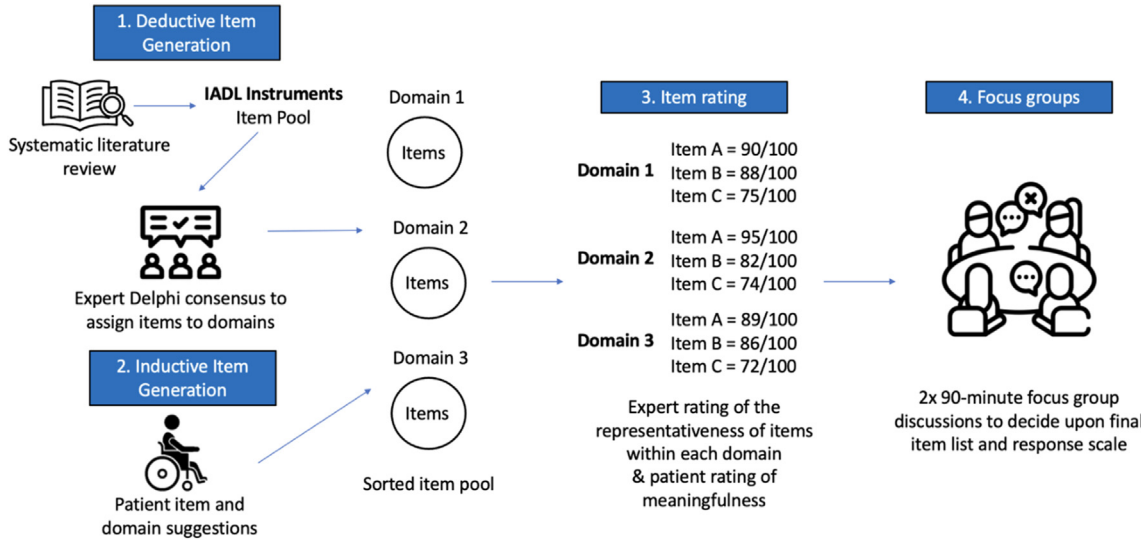

Fig. 1: Overview of multi-stage methodology.

Stage 1. Deductive item generation

Stage 1a. Systematic literature review

We performed a systematic scoping review of the literature, compliant with the Preferred Reporting Items of Systematic Reviews and Meta-Analysis extension for Scoping Reviews (PRISMA-ScR) guidelines.[21] The review was ineligible for registration with PROSPERO due to the absence of a defined clinical outcome. The review objective was to identify all historic validation studies of English-language IADL measures. The review objective was to identify all historic validation studies of English-language IADL measures, where IADL refers to more advanced tasks to support daily living in addition to basic self-care tasks (BADLs). We developed search strategies for three databases (MEDLINE, Embase, CINAHL) (MEDLINE search strategy provided in Fig. 2). Initial searches were performed to assess existing literature and refine the review question. Final search strategies were developed for three databases, MEDLINE, Embase, and CINAHL, using an iterative

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions <1946 to September 23, 2022>

1	Brain-Computer Interfaces/	4230
2	brain computer interface*.mp.	7298
3	brain machine interface*.mp.	1640
4	neuroprosthes*.mp.	1046
5	neural interface*.mp.	1219
6	1 or 2 or 3 or 4 or 5	10033
7	"Activities of Daily Living"/	71216
8	activities of daily living*.mp.	85902
9	ADLs.mp.	2975
10	7 or 8 or 9	86157
11	scale*.mp.	1090896
12	scor*.mp.	1242923
13	instrument*.mp.	986517
14	system*.mp.	5120090
15	questionnaire*.mp.	899606
16	tool*.mp.	910905
17	checklist*.mp.	53026
18	11 or 12 or 13 or 14 or 15 or 16 or 17	8390674
19	6 and 10 and 18	100

Fig. 2: Search strategy for MEDLINE database.

process. To maximise sensitivity, no automated search limits or restrictions were applied. Searches were performed using Ovid (Ovid Technologies, New York, USA) from inception to 1st November 2022. A medical librarian at Mount Sinai Hospital reviewed and provided comments on the searches, which were incorporated into the final strategies. We performed grey literature searching by a combination of citation searching (Fig. 2) and analysis of clinical trial protocols (NIH Clinical Trials Database).

We screened for study eligibility. Studies were included if they were published in English and validated a scale that included Instrumental Activities of Daily Living (IADL) items. Studies were excluded if the scale consisted exclusively of basic Activities of Daily Living (ADL) items or if the publication was a letter, review, editorial, opinion article, or conference abstract. Studies were also excluded if the full text was not available.

Title and abstract screening were completed using Rayyan (Rayyan Systems Inc., Cambridge, USA). Two medically trained reviewers (J.B. and A.S.) performed screening. An initial blinded pilot screen of 50 records was completed to ensure concordance in application of inclusion and exclusion criteria. Decisions were unblinded with discussion between reviewers before proceeding. Both reviewers worked independently and were blinded to each other's decisions until screening was complete. Articles were retrieved for full-text screening and data extraction using a piloted table. This was completed in duplicate by J.B. and A.S. Any differences were reconciled through discussion and consensus. We extracted specific characteristics for each identified IADL scale, including name of IADL scale, validation population(s), IADL items, scale response options, psychometric analysis data, and whether IADL scales were patient- or clinician-reported.

Stage 1b. IADL item data processing

Following data extraction, identified items underwent two sequential phases of independent screening by two medically trained reviewers (A.S. and J.B.) who were blinded to each other's decisions. First, we classified items into BADLs, IADLs, and exclusively cognitive tasks. Given our concept of interest was daily functioning commonly achieved using digital technologies, we focused exclusively on IADLs rather than BADLs. Therefore, items identified as BADLs (e.g., bathing) or cognitive (e.g., comprehending medication instructions) were removed from the item pool. Any discrepancies between reviewers in this phase were reconciled through discussion with an occupational therapist (OT) to reach consensus. We assessed the remaining items for their potential to be executed digitally. If at least one reviewer believed an item could be executed digitally, it remained in the item pool.

BADLs and IADLs have often been considered hierarchically, due in-part to the motor capabilities underpinning BADL assessments being pre-requisite to existing IADL assessments. Independence in IADLs typically implied independence in BADLs, meaning prior IADL assessments have been limited by floor effects where individuals are functionally dependent in BADLs, such as individuals with severe motor impairments. For these individuals, intervention often focused on BADLs. However, access to digital devices via assistive technology, augmentative and alternative communication technology, and BCIs has in some cases reversed this paradigm. These technologies can restore independence in IADLs through digital device control and internet access, without altering capacity for physical movements.[16,17] Therefore, restoration of functional independence may be achieved via some digital technologies and observed within IADLs more easily than BADLs. Dedicated IADL instruments have been utilised since 1969, many of which were derived from the first instrument by Lawton and Brody,[22] which highlighted IADL items as independent construct in functional assessments.[22] More recently, during validation of the CARE Item set, an exploratory factor analysis based upon over 500 assessments identified IADLs as independent factors, distinct from BADLs.[23]

Remaining items were deduplicated by removing both exact duplicates and items describing the same activity with similar wording. Deduplication was conservative at this stage to maximise the number of items and wording choices submitted to subsequent analysis. We subsequently standardised item phrasing by employing the gerund form of verbs to describe each activity, and minimising the number of words per activity, without compromising meaning.[24–26] Standardisation of item phrasing was performed independently by two pairs of blinded reviewers: three clinicians (including A.S. and J.B.) and one biostatistician (L.S.). Initial agreement was reached within each pair of reviewers before a final consensus was reached between pairs.

Stage 1c. Delphi process

To assign each IADL item to an appropriate domain, we conducted a reactive Delphi process using an electronic survey (Supplementary Materials Section 2). Items were presented to a multidisciplinary group of 17 clinicians, engineers, and scientists: four neurologists, two OTs, two physiotherapists, two speech and language pathologists, one physical medicine and rehabilitation physician, one neurosurgeon, one rheumatologist, one psychometrician, and one BCI user interface engineer. We generated a provisional list of nine domains, adapted where possible from the prototypical IADL scale,[22] which was deemed appropriate following piloting by four physicians. Participants could either assign items to a domain in this list or create an additional domain if they felt a different domain more appropriately represented the item. Participants also indicated

"can this item be done digitally". Each item–domain pair with >75% consensus on both item-domain pairing and possibility to perform the item digitally was removed from the Delphi pool and carried through to the next stage. Where >75% respondents believed an item could not be performed digitally, it was removed from the item pool. For the remaining items, we presented each item with the domains suggested by respondents in round one, along with the proportion of participants suggesting each domain. Participants could again select a domain from this list or create a new one if they felt a different domain better represented an item. Items not reaching consensus following the Delphi process were considered individually by three clinicians (including A.S. and J.B.). Two rounds of the Delphi process were performed. Items not reaching consensus were considered individually by three clinicians (including A.S. and J.B.).

Stage 2. Inductive item generation

Stage 2a. Electronic survey and interviews of individuals with lived experience

Twelve IWLEs, comprising patients and caregivers experiencing amyotrophic lateral sclerosis, spinal cord injury, and stroke completed an electronic survey (Supplementary Materials Section 3). Domains identified from the Delphi process (Stage 1C) were presented, and IWLEs were instructed to "list up to give daily activities, or 'items', that would be most important or meaningful to someone with your disability. These daily activities must only relate to activities that can be performed using digital devices". The IWLEs were given the opportunity to propose new domains relevant to categorise daily activities performed using digital devices and accompanying digital activities that were not already included. These steps provided inductive item generation without prompting using existing items. The deductively generated items from the Delphi process were then added to the inductively generated item lists, for each domain, and IWLEs were once again asked to add any further relevant items. If any proposed items were ambiguous or unclear, IWLEs were contacted to determine the intended meaning.

Respondents were also asked a series of questions including self-identifying their gender, age, ethnicity, care arrangements, use of assistive technology and basic physical function (Supplementary Materials Section 4).

Stage 2b. Inductive item processing and combined item grouping

Inductively generated items were deduplicated and reworded according to the same criteria applied during Stage 1b. Following this, deductive and inductive item pools were combined. Items judged to have equivalent meaning were grouped together by two investigators (A.S. and J.B.). Any differences were resolved through discussion. No items were eliminated at this stage.

Stage 3. Item rating and refinement

Stage 3a. Item rating by multidisciplinary experts

Forty multidisciplinary experts completed a survey where they were instructed to rate each item (or item group) to "identify which items are most representative of the ability to function in each domain". The experts comprised 16 physicians, 10 OTs, 10 physiotherapists, two speech and language pathologists, one scientist, and one engineer. Responses were collected using visual analogue scales, where the rightmost end of the line represented items that were most representative of the ability to function in the particular domain, and the leftmost end of the line represented items that were least representative. Each point on the line corresponded to a value from 0 to 100, which was reported to the investigators but hidden from participants. Average values were calculated for each item.

Items were excluded if any two of the following three criteria were met:

- Median item representativeness rating $\leq$50, or
- Upper 95% CI $\leq$50, or
- Alpha improves if item deleted

Stage 3b. Item rating by individuals with lived experience

Twenty IWLEs performed a similar survey where they were instructed to rate each item (or item group) according to which "activities are most meaningful/important in [each domain] to someone with your disability". Respondents were asked to rate each item (or item group) on an 11-point numeric rating scale (0–10), as visual analogue scales can be difficult to use for some individuals with motor impairments. No items were excluded based upon these ratings, however, this did inform later item deduplication and rewording decisions.

Stage 4. Focus group discussion

Stage 4a. Focus group 1

Eight experts, including three IWLE, attended a 90-min virtual focus group, moderated by two investigators (A.S. and D.P.). The group consisted of three patient advocacy organisation leaders (one OT who lives with a spinal cord injury, one physician who was a caregiver to an individual with ALS, and one individual who lives with spinal muscular atrophy), two further OTs, three BCI experts: one neurologist, one cognitive neuroscientist, and one speech and language pathologist. All items were presented to participants, categorised by domain. Item reduction rules and decisions were presented. Scores were also presented from both the expert (relevance) and IWLE (importance) rating tasks.

Focus group discussion was transcribed (L.S.), thematically analysed (J.B.) and cross-checked (A.S.). Based upon the feedback from the focus group

participants, IADL items were further refined to address suggestions and concerns. These suggestions related to item redundancy or phrasing to ensure relevance to individuals with severe motor impairment. We also revised language throughout the scale to be patient centric, shifting the focus from care providers to the individuals themselves, at the request of the focus groups. In cases where a majority of investigators agreed that an item was inappropriate or needed modifying, the item was modified or omitted accordingly. Conflicts with no majority agreement were resolved by discussion. The refined list was then assessed by four blinded investigators and an OT (including A.S., J.B., L.S., and A.F.) to ensure appropriate specificity of each item. A priori criteria stipulated that items must be related to a clearly defined task, so that items were not excessively broad, and not so specific that they represent a niche case of an already appropriate item. Conflicts were resolved by discussion.

Stage 4b. Focus group 2

The second 90-min virtual focus group involved eight experts, six of whom participated in the first focus group. The group consisted of two patient advocacy organisation leaders (one OT who lives with a spinal cord injury, one physician who was previously a caregiver to an individual with ALS), two further OTs, one speech and language pathologist, one health policy expert, and one disability rights advocate (for accessible technology). Discussion was moderated by A.S. The item decisions following the first focus group were presented, providing an opportunity for participants to confirm or contest decisions. A proposed response scale, based upon the response scale from the Continuity Assessment Record and Evaluation (CARE) Item Set,[23] was presented, prompting discussion about the merits of this scale and alternatives.

Statistics

The analysis is as described in the Methods section. No other statistical analyses were performed.

Ethics

Ethical approval was granted by the Institutional Review Board at Mount Sinai Hospital, New York, USA (23-00847). Each participant in the study was asked for written informed consent, and their data was anonymised for reporting. Any potentially identifiable information was removed prior to reporting and publishing the findings.

Role of funders

Support for this project was provided in kind by the Abilities Research Center, which is a sub group of the Department of Rehabilitation and Human Performance, Icahn School of Medicine at Mount Sinai relevant to author A.S., D.P. and A.F. As such, the Abilities Research Center had direct involvement in the study design, data collection, data analyses, interpretation, or writing of the report.

Results

The results align with the FDA draft guidance on the search and development of fit-for-purpose COAs (Supplementary Materials Section 5 Table S1; FDA Draft Guidance to Summarise Rationale and Support for the Proposed Instrument).[20]

Stage 1. Deductive item generation

Following screening, the systematic review yielded 71 validated IADL scales. A total of 1250 items were extracted from these scales. Of these items, 224 were excluded on the basis of being purely cognitive and 260 were BADLs and removed. A further 413 items were removed as they were deemed unable to be performed digitally in their entirety. Remaining items were deduplicated, reducing the number of items to 77. Further detail on the screening and data extraction process is presented in Fig. 3.

Following round 1 of the Delphi process, 24 items were retained as there was consensus on both the domain they were allocated to and their ability to be performed digitally, and eight items were eliminated as there was consensus that they could not be performed digitally. Following round 2, consensus on allocation to domain and ability to be performed digitally was reached on a further 18 items. Following two rounds of electronic Delphi surveys, it was felt that a further round would not yield consensus for the remaining items. Of the 27 items not reaching consensus, 13 did not reach consensus on allocation to domain, six did not reach consensus on ability to be performed digitally, and eight met consensus on neither criteria. Items not reaching consensus were considered individually (A.S. and J.B.), and in each case the item was excluded because it lacked specificity for consistent interpretation and domain assignment (e.g., *Operating devices*) or involved compound activities (e.g., *Accessing entertainment and information*) resulting in multiple potential domains for assignment, with roughly equal allocations within the Delphi responses. Compound activities were subsequently split into their individual activities for consideration as distinct items, however, in each case the individual activities were considered duplicates of existing items already assigned to domains by consensus and were therefore removed. Ultimately, 42 IADL items were generated deductively. The full dataset for Stage 1 can be found in the Supplementary Tables S2–S8 (named Literature Review—Supplementary Table S2, Item Exclusion—Supplementary Table S3, Item Domains—Supplementary Table S4, Deduplication—Supplementary Table S5, Deduplicated Items—Supplementary Table S6,

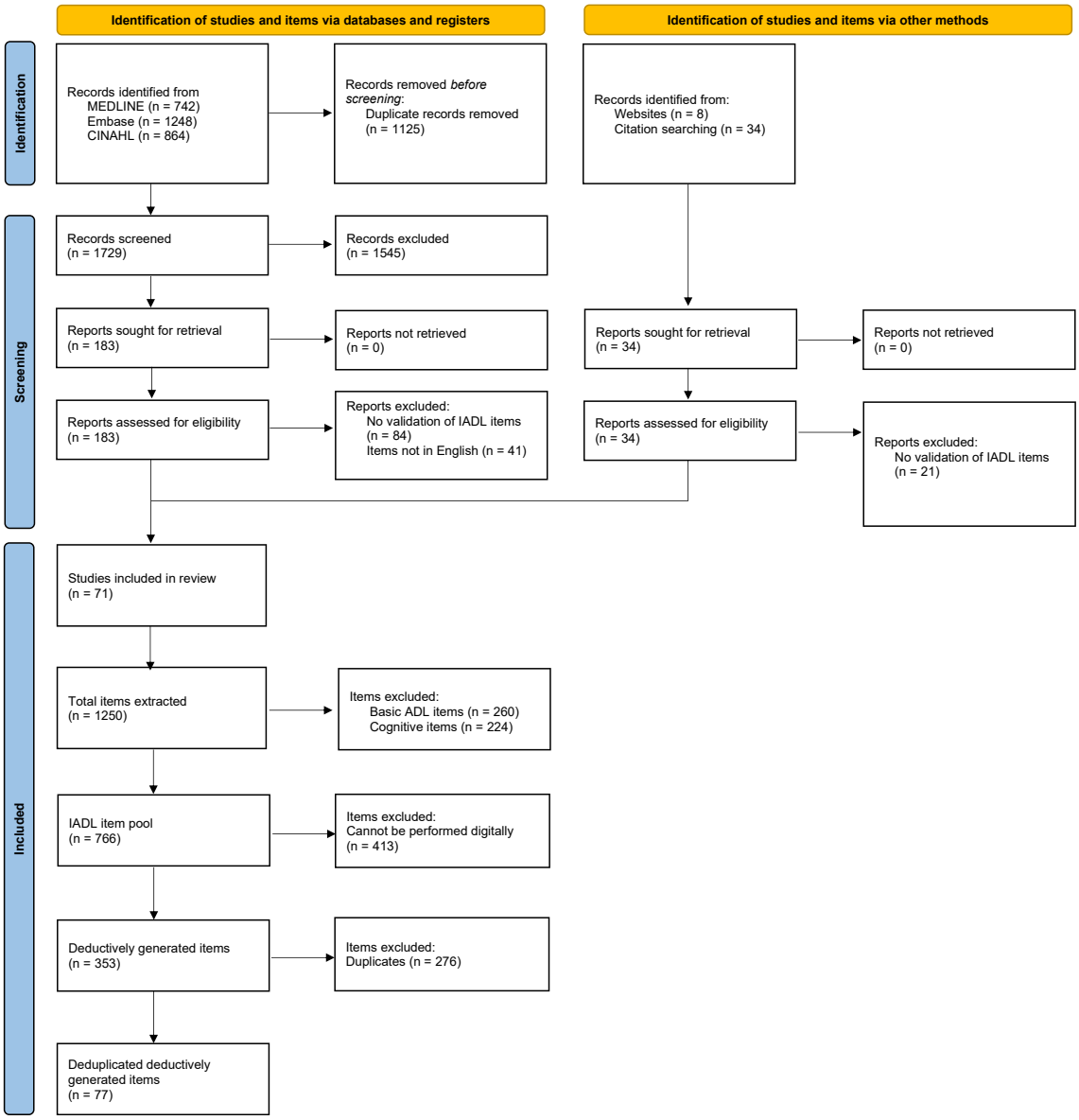

Fig. 3: Adapted PRISMA flow diagram to show the systematic review process for deductive item generation.

Rephrasing—Supplementary Table S7, Delphi—Supplementary Table S8).

Stage 2. Inductive item generation

Twelve IWLEs responded to our inductive item generation survey. Respondents self-identified their gender as male in 91.7% (n = 11), and were aged 48 ± 13 years (mean ± standard deviation). Aetiology of paralysis was spinal cord injury (n = 5, 41.7%), amyotrophic lateral sclerosis (n = 4, 33.3%), and stroke (n = 3, 25.0%). Seven (58.3%) reported severe distal upper limb motor impairments, and nine (66.7%) were unable to walk independently. Two (16.7%) IWLEs reported relying primarily on non-verbal communication, two (16.7%) reported relying on assistive communication devices, two (16.7%) reported that they communicated verbally, but their speech was not always clear, and the remaining six (50.0%) reported no impairment in verbal communication. All participants reported living in a private residence, with care primarily provided by family full time (n = 2, 16.7%), by family part time (n = 2, 16.7%), by professionals full time (n = 4, 33.3%), by professionals part time (n = 3, 25.0%), or no care was required (n = 1, 8.3%). Most IWLEs (n = 7, 58.3%)

reported using digital devices at an hourly frequency. The remaining respondents reported digital device use of more than once per day (n = 3, 25.0%), about once per day (n = 1, 8.3%), or several times per week (n = 1, 8.3%). Assistive technologies currently used by respondents included voice operated devices (n = 3, 25%), eye-tracking devices (n = 2, 16.7%), joysticks (n = 2, 16.7%), and head pointers/trackers (n = 2, 16.7%). One IWLE (8.3%) reported using 'Other' devices and indicated the device was an iPhone and two (16.7%) reported not using any assistive technologies.

Respondents inductively generated an additional 152 items to consider for inclusion in the item pool and one new domain—*Work*. Twenty of the items were excluded because they were explicitly physical tasks, with no current potential to be completed in their entirety using digital devices. The remaining 132 items were combined with the pool of deductively generated items.

Forty-six items were combined to reduce repetition if it was determined that there was equivalent meaning by the investigators. In the eight instances an inductively generated item was assigned to a different domain by IWLEs than by participants of the Delphi process, the items were retained in each domain for exclusion based on their rating of domain representativeness (Stage 3). This resulted in a combined pool of 83 items: 42 deductively generated and 41 inductively generated; an equal balance between the two item generation processes. After Stages 1 and 2, no items remained within the domain of *Transport* and this domain was removed. This resulted in a total of nine domains to which all items were each assigned (Table 1). The full dataset from Stage 2 can be found in the Supplementary Tables S9 and S10 (named IWLE Survey and Processing).

Digital IADL domains
Communication
Information retrieval
Healthcare
Smart environment control
Organisation
Managing finances
Shopping
Leisure
~~Transport~~[b]
Work[a]

The final nine domains differed to the original nine domains presented to participants in the Delphi process. [a]Work was added during the inductive item generation in stage 2. [b]Transport was removed as no items remained after item processing in stages 1 and 2. The rationale for this exclusion is further elaborated in the discussion.

Table 1: **The IADL domains following the multi-stage methodology.**

Stage 3. Item refinement and rating

Forty multidisciplinary experts completed an item rating survey to assess the representativeness of the 83 deductively and inductively generated items to their assigned domains. The mean professional experience of respondents was 10.9 years. Eight of the 83 items (9.6%) were eliminated as the same item was rated as more representative in another domain. Six of the remaining 75 items (8.0%) were found to have low relevance within their domain due to meeting at least two of the three a priori statistical criteria and were excluded. This left a total of 69 items with a "representativeness" rating of 71.6 ± 13.9 out of 100.

Twenty IWLEs rated the "importance or meaningfulness" of all items. The highest-rated items were 'Answering email', 'Talking with family and friends', and 'Obtaining information', achieving scores of 8.85 ± 2.46, 8.75 ± 2.34, and 8.65 ± 2.43 out of 10, respectively. The lowest-rated items were 'Making trades and transfers' (2.75 ± 2.86), 'Using reminders to take medication' (3.00 ± 3.49), and 'Scheduling a ride-hailing service' (3.05 ± 3.79). However, these low rated items also exhibited greater variability in meaningfulness, with each item receiving a high rating from individual IWLEs. The average rating for all items was 6.60 ± 3.47. The full dataset for Stage 3 can be found in Supplementary Tables S11–S13 (named Expert Rating, Processed Items and IWLE Rating, respectively).

Stage 4. Focus group discussion

Inductive thematic analysis of the first focus group resulted in 38 codes assigned to sections of the transcribed discussion, which were grouped into five themes (see Supplementary Materials Section 6 Table S14; Focus Group Inductive Thematic Analysis Results). After addressing this feedback, a second focus group was held from which a further 28 codes were generated from the transcribed discussion. Two additional themes were observed. The full dataset can be found in the Supplementary Table S15 (Focus Group Edits). In response to focus group feedback, the item list was further reduced from 69 to 37 items.

Discussion

Consensus-Based adjustments to IADL domain and item selection

Throughout each stage of instrument development, we required that each item represents an IADL activity that could be completed in its entirety using digital methods and that independently completing each activity in its entirety would represent independence within the domain of interest. Following these criteria, two conventional IADL domains; *Food preparation* and *Transport*, were excluded as all activities initially identified within these domains were deemed inappropriate or considered representative of a different domain in their digital form. For example, whilst a person with severe

motor impairments may be able to order food online, they would remain unable to prepare the food for eating, such as transferring a delivered meal from the front door to a plate, through digital means. Items relating to ordering food online were therefore considered important but more appropriately categorised within the *shopping* domain than *food preparation*. Similarly, scheduling a taxi online would not represent independence in *Transport* if the individual cannot transfer into and out of the vehicle digitally. Instead, the activity of scheduling transport was recognised as important and retained within the domain of *Organisation*. Power wheelchair control was considered but felt to represent *Functional Mobility*, which is a BADL and beyond the scope of this initial scale. It was also recognised that full digital control of other vehicles may be on the horizon, however, it was agreed that the domain of *Community Mobility* was not yet applicable for most individuals who may use this scale in the near-term. Future updates and revisions may expand beyond the currently included domains as technology advances.

Conversely, less well established IADL domains were included by consensus between content experts and IWLE. This includes organisation (e.g. 'setting an alarm or reminder'), information retrieval (e.g. 'using a search engine'), and smart environment control. *Gaming* was also advocated for inclusion, as an item, by focus group participants. This represents one of the most widespread forms of leisure activity in the 21st century, and the restored ability to play video games has been cited as highly meaningful by participants in early feasibility studies of BCIs (Rogers, 2016).[27] Furthermore, gaming can improve mental wellbeing and provide opportunities to maintain social relationships online (Jones et al., 2014).[28] Items related to paid employment, such as teleconferencing were also included. This reflects the increasing dependence upon digital technologies in the workplace, and the expansion in remote working since the COVID-19 pandemic.

Development of a digital IADL response scale

We evaluated existing approaches to scoring ADLs as a basis for a potential dIADL response scale. A binary scale similar to that used by Lawton and Brody (Lawton and Brody, 1969)[22] was considered to provide insufficient resolution for measuring meaningful change in independence on dIADLs, which was supported by the focus group.

The CARE Item Set, previously validated by CMS,[23] was identified as the predominant method of evaluating independence in ADLs in clinical practice in the US. The corresponding 6-point response scale was appraised, and some updates were considered for the assessment of dIADLs. Specifically, measurement of performance achieved during acute observations following careful setup (current Level 5) might overestimate typical performance or lead to ceiling effects. For instance, some technologies that enable digital performance of IADLs are not reliable throughout the day, leading to increased reliance on caregivers, meaning the assessment of performance immediately following setup may not best reflect an individual's functional dependence.[29,30] Independence across a typical day was therefore considered of greater clinical importance than observed performance in test situations. This view was supported by the focus groups and is consistent with prior literature.[30,31] Assessing independence across a typical day also follows the precedence of other common assessments of functional independence, such as the Barthel Index[32] and UK Functional Independence Measure and Functional Assessment Measure (FIM + FAM),[33–35] as demonstrated by the following instructions from the FIM + FAM.

> *The person is scored on what they actually do, on a day-to-day basis, not what they could or might be able to do, in different circumstances*

A change to patient first language was also advocated by the focus group. A provisional Digital IADL response scale incorporating these updates to the CARE Item Set response scale is presented in the Supplementary Materials Section 7 Table S16 (Proposed Digital IADL Response Scale and Comparison to the CMS CARE Item Set). In each case, updates were suggested and/or supported by the focus group participants.

Alternatively, the rigorous process of item generation described herein may enable the current set of domains and items to be used in combination with existing, ideographic instruments. This would not require the development of a new response scale. For example, the current pool of items, organised into their respective domains, might be used as a framework providing some guidance for appropriate selection of occupational problems specific to digital IADLs using a tool such as the Canadian Occupational Performance Measure.[36]

Ultimately, the response scale chosen for this scale will be required to directly capture the construct we sought to measure — changes in functional independence, across a typical day within an individual's typical environment, achieved through the use of digital technology.

Future use and validation of the instrument

The Digital IADL Scale under development requires further validation prior to use in clinical practice or clinical trials. This should include a pilot study, followed by a psychometric analysis to evaluate the underlying structure of the scale and ensure that it accurately captures the construct of functional independence achieved via digital performance of IADLs. Such analyses may also explore whether the dIADL assessment is sufficient when conducted at the domain-rather than item-level.

The reliability and clinical utility of the final scale should also be established, prior to using this scale in clinical settings.

In addition to the items identified thus far, future use of the Digital IADL Scale may incorporate an ideographic component, as suggested by the focus group. Recognising that not all patients may engage with every item in the scale as part of their daily routines, this approach allows for a personalised assessment that better reflects each individual's specific digital activities. Moreover, as digital technologies evolve, the relevance of certain items may change, necessitating the addition of new items or the revision of existing ones. The flexibility to adapt the scale over time is important to maintaining its relevance and utility in a fast-evolving digital landscape. If an ideographic component is piloted and validated, this might also contribute to the continued refinement of the item set through further inductive item generation. While a similar number of items were generated through the inductive and deductive processes, the inductive item generation involved input from only twelve IWLEs. Some potentially relevant items (e.g., two-factor authentication) were also identified after item generation, rating, and processing were already complete, and have therefore not been included in this initial item set.

The optimal method for calculating an individual's score has not yet been established. Regardless of the response scale used, a simple summative score across all items may not be appropriate, as some items may not be appropriate or applicable to every patient. Instead, calculating the mean score of applicable items within each domain may provide a more valid assessment of an individual's functional independence in digital IADLs. Other scoring strategies such as directly scoring at the domain level or considering a selection of items might also be considered, and should be a subject for further validation. For the CARE Item Set[23] or modified CARE Item Set response scale (see Supplementary Materials Section 7 Table S16), which is a clinician-reported outcome measure, a change of one point is expected to represent a meaningful change on that item.[23] A minimally important change on aggregate scoring would require further analysis.

This preliminary scale (see Instructions for such in the Supplementary Materials Section 8) was developed with a primary focus on individuals with severe motor impairments. As a result, it may not yet address the needs specific to older adults with cognitive decline, who also face significant barriers to digital participation. These barriers may expand outside of the scope of this study and include issues of digital literacy as the basis to impaired digital functional performance. Expanding the scale to include this population might be considered in future validation.

Our study was limited by a small sample of IWLEs participating in the inductive item generation. Scale piloting is yet to be performed, and further validation is required. These limitations represent the subject of future study. Additionally, an individual's baseline digital literacy was identified as a potential confound when using this tool to assess the effects of interventions addressing independence in dIADLs, which should be accounted for by appropriate study design.

Appropriate next steps include pilot and validation process. Although preliminary feedback has been gathered, further validation—encompassing pilot testing including user experience, reliability testing (internal consistency, test-retest reliability) and validity (construct, criterion and discriminant) testing is required.

This study provides the initial development of a Digital IADL Scale, which represents the first instrument to specifically assess independence on IADLs performed digitally. A multi-stage methodology involving clinicians, researchers, and IWLE, produced a list of nine IADL domains and 37 items. A potential response scale was adapted from the CARE Item Set to emphasise functional independence across a typical day, however, the optimal method for scoring remains undetermined. This instrument may address an important gap in the assessment of functional independence achieved digitally by serving as a novel COA for clinical trials of assistive technologies or BCI devices for individuals with severe motor impairment. The scale's ultimate utility will be elucidated following further validation.

Contributors

AS and JB contributed to the literature review, study design, data collection and interpretation, access to and verification of data, and the write-up of the manuscript. LS led the data analysis and provided statistical oversight, as well as access to and verification of data, and contributed to manuscript editing. DP assisted with the conceptualisation, review of data, and editing of the manuscript. AF led the conceptualisation and study design, overseeing access to and verification of data, data interpretation, and the write-up. All authors read and approved the final version of the manuscript. Additionally, the BCI Functional Outcome Measures Collaborative Group participated in completing surveys (e.g., Delphi or Item Rating), interviews, and/or contributed to the focus groups.

Data sharing statement

The full Scale can be provided upon request to the authors. The selected domains and items are presented in the Supplementary Table S17 (Domains and Items).

Declaration of interests

AF disclosed stock options for Synchron Inc. JB received consulting fees for Synchron Inc. and UK Advanced Research and Invention Agency. LS received consulting fees for Synchron Inc.

Acknowledgements

We acknowledge support for this project provided in kind by the Abilities Research Center. We acknowledge the contributions from Maria Nardozzi, OTD, OTR/L, who provided clinical insights throughout this study. We acknowledge contributions from Professor Todd Little, who reviewed and provided input on our methodology. We acknowledge the assistance of medical librarian Gloria Willson from Icahn School of Medicine at Mount Sinai during the systematic literature review. A complete list of collaborators in the BCI Functional Outcome Measures Group is provided in the Supplementary Materials

Section 1. These collaborators participated in data collection for surveys or focus groups.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2025.105732.

References

1 Fessler EB, Brown RT, Miller RK. Rebooting instrumental activities of daily living for the 21st century. *Ann Intern Med*. 2022;175(2):278–279. https://doi.org/10.7326/M21-3065.
2 Pew Research Center. *Mobile device ownership and internet usage statistics*. Washington, DC: Pew Research Center; 2023.
3 Gómez-Baya D, Grasmeijer AJ, López-Bermúdez E, Gaspar de Matos M, Mendoza R. Positive youth development and internet use in a sample of Spanish adolescents. *Front Pediatr*. 2022;10:842928. https://doi.org/10.3389/fped.2022.842928.
4 Statista. *Global smartphone ownership statistics*. Hamburg: Statista; 2023.
5 U.S. Department of Agriculture. *Annual report on nutritional guidance and accessibility in the U.S*. Washington, DC: USDA; 2023.
6 Chang E, Penfold RB, Berkman ND. Patient characteristics and telemedicine use in the US. *JAMA Netw Open*. 2024;7(3):e243354. https://doi.org/10.1001/jamanetworkopen.2024.3354.
7 Melrose RJ, Brommelhoff JA, Narvaez T, et al. The use of Information and Communication technology when completing instrumental activities of daily living. *Comput Hum Behav*. 2016;63:471–479. https://doi.org/10.1016/j.chb.2016.05.045.
8 Amankwah-Amoah J, Khan Z, Wood G, Knight G. COVID-19 and digitalization: the great acceleration. *J Bus Res*. 2021;136:602–611. https://doi.org/10.1016/j.jbusres.2021.08.011.
9 Lu X, Yao Y, Jin Y. Digital exclusion and functional dependence in older people: findings from five longitudinal cohort studies. *eClinicalMedicine*. 2022;54:101708. https://doi.org/10.1016/j.eclinm.2022.101708.
10 Baumgartner A, Rohrbach T, Schönhagen P. "If the phone were broken, I'd be screwed": media use of people with disabilities in the digital era. *Disabil Soc*. 2019;38(1):73–97. https://doi.org/10.1080/09687599.2021.1916884.
11 Rigby P, Ryan SE, Campbell KA. Electronic aids to daily living and quality of life for persons with tetraplegia. *Disabil Rehabil Assist Technol*. 2011;6(3):260–267. https://doi.org/10.3109/17483107.2010.522678.
12 American Occupational Therapy Association. *Occupational therapy practice framework: domain and process*. 4th ed. Bethesda, MD: AOTA; 2022.
13 Lai FH, Tong AY, Fung AW, et al. Information communication technology as instrumental activities of daily living for aging-in-place in Chinese older adults with and without cognitive impairment: the validation study of advanced instrumental activities of daily living scale. *Front Neurol*. 2022;13:746640. https://doi.org/10.3389/fneur.2022.746640.
14 Stringer G, Leroi I, Sikkes SAM, Montaldi D, Brown LJE. Enhancing "meaningfulness" of functional assessments: UK adaptation of the Amsterdam IADL questionnaire. *Int Psychogeriatr*. 2021;33(1):39–50. https://doi.org/10.1017/S1041610219001881.
15 Patrick-Krueger KM, Burkhart I, Contreras-Vidal JL. The state of clinical trials of implantable brain–computer interfaces. *Nat Rev Bioeng*. 2024;3:1–18.
16 Mitchell P, Lee SC, Yoo PE, et al. Assessment of safety of a fully implanted endovascular brain-computer interface for severe paralysis in 4 patients: the Stentrode with thought-controlled digital switch (SWITCH) study. *JAMA Neurol*. 2023;80(3):270–278.
17 Fry A, Chan HW, Harel NY, Spielman LA, Escalon MX, Putrino DF. Evaluating the clinical benefit of brain-computer interfaces for control of a personal computer. *J Neural Eng*. 2022;19(2):021001.
18 Sawyer A, Chetty N, McMullen DP, et al. Building consensus on clinical outcome assessments for BCI devices. A summary of the 10th BCI society meeting 2023 workshop. *J Neural Eng*. 2024;22. https://doi.org/10.1088/1741-2552/ad7bec.
19 Mois M, Rogers T. Trends in wearable technology for chronic disease management. *J Chronic Illn Manag*. 2024;14(3):95–106.
20 U.S. Food and Drug Administration. *Patient-focused drug development: selecting, developing, or modifying fit-for-purpose clinical outcome assessments*. Silver Spring, MD: FDA; 2022 [cited 2025 Apr 3]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-selecting-developing-or-modifying-fit-purpose-clinical-outcome.
21 Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467–473. https://doi.org/10.7326/M18-0850.
22 Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9(3):179–186.
23 Centers for Medicare & Medicaid Services. *CARE item set validation. Care management systems: a comprehensive overview*. Baltimore, MD: CMS; 2023.
24 Fanego T. The gerund in early modern english: evidence from the Helsinki corpus. *Folia Linguist Hist*. 1996;30(17 1-2):97–152. https://doi.org/10.1515/flih.1996.17.1-2.97.
25 World Health Organization. *International classification of functioning, disability and health (ICF)*. Geneva: WHO; 2001.
26 Ustün TB, Chatterji S, Kostanjsek N, et al, WHO/NIH Joint Project. Developing the World Health Organization disability assessment schedule 2.0. *Bull World Health Organ*. 2010;88(11):815–823. https://doi.org/10.2471/BLT.09.067231.
27 Tabacof L, Dewil S, Herrera JE, Cortes M, Putrino D. Adaptive esports for people with spinal cord injury: new frontiers for inclusion in mainstream sports performance. *Front Psychol*. 2021;12:612350.
28 Jones CM, Scholes L, Johnson D, Katsikitis M, Carras MC. Gaming well: links between videogames and flourishing mental health. *Front Psychol*. 2014;5:260. https://doi.org/10.3389/fpsyg.2014.00260.
29 Benge JF, Aguirre A, Scullin MK, et al. Digital methods for performing daily tasks among older adults: an initial report of frequency of use and perceived utility. *Exp Aging Res*. 2024;50(2):133–154. https://doi.org/10.1080/0361073X.2023.2172950.
30 Sawyer A, Cooke L, Breyman E, et al. Meeting the needs of people with severe quadriplegia in the 21st century: the case for implanted brain–computer interfaces. *Neurorehabil Neural Repair*. 2024;38(11-12):877–886. https://doi.org/10.1177/15459683241282783.
31 Zhang Q, Karunanithi M, Bradford D, van Kasteren Y. Activity of Daily Living assessment through wireless sensor data. *Annu Int Conf IEEE Eng Med Biol Soc*. 2014;2014:1752–1755.
32 Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *J Chronic Dis*. 1965;18(1):47–55.
33 Hall KM, Hamilton BB, Gordon WA, Zasler ND. Characteristics and comparisons of functional assessment indices: disability rating scale, functional independence measure, and functional assessment measure. *J Head Trauma Rehabil*. 1993;8(2):60–74.
34 Turner-Stokes L, Siegert RJ. A comprehensive psychometric evaluation of the UK FIM+FAM. *Disabil Rehabil*. 2013;35(22):1885–1895. https://doi.org/10.3109/09638288.2013.766271.
35 UK Rehabilitation Institute. *Functional independence measure (FIM) and functional assessment measure (FAM) guidelines*. London: UK Rehabilitation Institute; 2020.
36 Law M, Baptiste S, McColl M, Opzoomer A, Polatajko H, Pollock N. The Canadian occupational performance measure: an outcome measure for occupational therapy. *Can J Occup Ther*. 1990;57(2):82–87. https://doi.org/10.1177/000841749005700207.