



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/245208/>

Version: Published Version

Article:

Kaur, L., Griffiths, A.W., Stokes, L. et al. (2026) Dementia carers' perspectives on paper- and artificial intelligence (AI)-based assessment tools: a qualitative think-aloud study using the COM-B framework. *Journal of Geriatric Psychiatry and Neurology*. ISSN: 0891-9887

<https://doi.org/10.1177/08919887261486060>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Dementia Carers' Perspectives on Paper- and Artificial Intelligence (AI)-Based Assessment Tools: A Qualitative Think-Aloud Study Using the COM-B Framework

Journal of Geriatric Psychiatry
and Neurology
2026, Vol. 0(0) 1-16
© The Author(s) 2026



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/08919887261486060
journals.sagepub.com/home/jgp
Mary Ann Liebert
A Part of Sage

Labhpreet Kaur, BSc¹ , Alys Wyn Griffiths, PhD¹ , Louis Stokes, PhD¹,
Judith Harrison, MBChB, PhD², Xingyi Song, PhD³, and
Daniel Blackburn, MBChB, PhD⁴

Abstract

Objectives: Carer observations provide valuable longitudinal and contextual information for dementia assessment, but traditional informant-based questionnaires may limit reporting of nuanced, fluctuating changes. Artificial intelligence (AI)-based tools may offer greater flexibility to report, but their successful uptake depends on carer capability and acceptability. Despite this, little is understood about carers' views on existing informant questionnaires and emerging AI-based tools, therefore, this study explored carers' perspectives on barriers and facilitators influencing engagement with paper- and AI-based assessment tools.

Methods: Ten carers of people living with dementia completed think-aloud interviews, while reviewing the Cambridge Behavioural Inventory-Revised (CBI-R) and two AI-based tools, CognoSpeak™ and LUMEN (Language Usage for Monitoring Early Neurodegeneration). Using a hybrid inductive-deductive approach, themes related to barriers and facilitators to engagement were identified and subsequently mapped onto the Capability, Opportunity and Motivation-Behaviour (COM-B) framework.

Results: Carers valued the CBI-R's simplicity and breadth but found its structure, complex language and limited free-text space restrictive. Verbal, narrative responses felt more natural and less burdensome, therefore, AI-based tools were seen as more engaging and better able to collect richer narratives, with the added advantage of remote access, however, digital literacy, sensory impairments, privacy concerns could limit uptake. Participants stressed that AI should support, not replace, human-led care.

Conclusion: Engagement depended on individual capabilities, available opportunities and motivation; factors fundamental to successful uptake. Future tools should be co-designed to ensure accessibility, cultural sensitivity and clinical relevance.

¹Division of Neuroscience & Neuroscience Institute, Sheffield Institute of Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK

²Translational and Clinical Neurosciences, Biomedical Research Building, Campus for Aging and Vitality, Newcastle University, Newcastle upon Tyne, UK

³School of Computer Science, University of Sheffield, Sheffield, UK

⁴Sheffield Institute of Translational Neuroscience (SITraN), University of Sheffield and Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

Corresponding Author:

Labhpreet Kaur, BSc, Division of Neuroscience & Neuroscience Institute, Sheffield Institute of Translational Neuroscience (SITraN), University of Sheffield, 385a Glossop Road, Sheffield S10 2HQ, UK.

Email: lkaur2@sheffield.ac.uk

Keywords

dementia carers, informant assessment, collateral history, qualitative research, COM-B framework, behaviour change, artificial intelligence, health technology assessment, digital health

Received: April 30, 2026; revised: August 19, 2026; accepted: August 23, 2026

Highlights

What Is Already Known

- Carer collateral histories are essential for contextualising cognitive, behavioural and functional changes in dementia assessment and diagnosis.
- Informant questionnaires can support diagnosis but are often inconsistently integrated into memory services and may be limited by rigid formats, literacy demands and cultural assumptions.
- AI-based dementia screening tools are emerging, but carer acceptability, trust and accessibility remain underexplored.

What This Paper Adds

- Using Think-Aloud interviews and the COM-B framework, this study identified key barriers and facilitators influencing carer engagement with paper- and AI-based assessment tools.
- Carers preferred flexible, verbal and narrative approaches that better captured nuanced everyday changes than fixed-response questionnaires.
- AI-based tools were viewed as promising for remote, richer collateral history collection, but only when privacy, digital literacy and cultural relevance were addressed.

Introduction

Dementia is the fourth leading cause of disability globally among people aged 75 years and above.¹ Over 55 million people live with dementia and prevalence is projected to triple by 2050.² Yet only two-thirds of people with dementia receive a formal diagnosis in the United Kingdom,³ with misdiagnosis rates estimated at 30%.⁴ Around 62% of cases remain undetected globally and one in four patients in Europe and North America receive a late diagnosis.⁵

Early, accurate detection is therefore a public health priority.² Timely diagnosis enables access to interventions, supports advance care planning,^{5,6} and may increase opportunities for research participation, particularly as many disease-modifying trials enrol people with mild cognitive impairment (MCI) or mild dementia.^{7,8} Timely recognition therefore allows sufficient time to discuss

participation. However, current pathways often identify dementia too late to realise these benefits.⁹

Dementia Screening

Dementia assessment incorporates self- and informant reports, cognitive and neuropsychological testing, neuroimaging and functional assessment. Initial screening commonly relies on brief pen-and-paper tools such as the Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA).¹⁰⁻¹⁵ Reported sensitivity for progression from MCI to dementia ranges from 23-89% for the MMSE^{11,12} and 87-93% for the MoCA.^{14,15} However, these tools demonstrate reduced sensitivity for detecting subtle or non-memory-related deficits and are influenced by demographic, educational and cultural biases.¹⁶

Therefore, informant reports provide structured accounts from family members or carers about changes in their relatives' cognition, behaviour and functioning.¹⁷ Carers often identify early changes, with 86% of carers reporting that they were the first to recognise changes compared to 6% of patients and 4% of physicians.¹⁸ Their observations contextualise test performance, helping distinguish neurodegenerative decline from fatigue, sensory impairment or anxiety. Measures such as the Eight-item Informant Interview to Differentiate Aging and Dementia (AD8) and the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) show useful diagnostic sensitivities (AD8 72-91%; IQCODE 80-91%), although specificity varies (AD8 67%; IQCODE 66-85%),¹⁹⁻²¹ while the Cambridge Behavioural Inventory-Revised (CBI-R) assesses a broader range of symptoms and correlates with neuropsychological performance.^{22,23} Although informant reports can be influenced by mood, burden and relationship dynamics,²⁴⁻²⁶ they can improve diagnostic accuracy when used alongside objective measures.²⁷ However, collateral history-taking remains inconsistent and poorly standardised due to brief appointments and variable documentation.²⁸

Emerging Artificial Intelligence

These limitations have driven interest in Artificial Intelligence (AI) to improve dementia screening. Digital adaptations of established tools improve accessibility and

reduce limitations associated with pen-and-paper versions. For example, the electronic MoCA combines voice and drawing inputs with automated scoring,²⁹ and the Amsterdam Instrumental Activities of Daily Living questionnaire automatically scores functional decline.³⁰ Speech is also emerging as a digital biomarker, with machine-learning models identifying acoustic, lexical and linguistic features associated with cognitive decline, reporting accuracies up to 96.8%.³¹ However, high performance often derives from selected samples, small datasets or internally validated models and may not generalise to clinically heterogeneous populations.

Clinical tools are also emerging. CognoSpeak™, an online stratification tool, uses a digital agent to conduct cognitive interviews and has shown sensitivity comparable to traditional assessments.³² Advances in large language models (LLMs) further enhance semantic analysis: NeuroXVocal combines textual and acoustic data to classify speech with a 95.8% F1 score,³³ GPT-based approaches can quantify word-finding difficulties analyse informant reports with substantial agreement with clinicians.³⁴ Such systems may support scalable assessment, reduce travel and administrative burden and collect richer information outside time-limited appointments.

However, the field remains at an early stage of clinical evaluation. Concerns include unrepresentative training data, cross-cultural validity, privacy and the role of automated systems alongside clinical judgement.³⁵ Many models lack independent external validation and may perform differently in routine services, where speech quality, presentation and access to technology are more variable.

Carer Acceptability

Carers are central to dementia recognition, yet obtaining high-quality collateral information can be challenging. Time constraints, questionnaire complexity and uncertainty about reporting observed changes can reduce engagement and data quality.³⁶ AI-based systems may enable remote participation and richer narrative data, but introduce new barriers. Therefore, successful implementation depends not only on technical accuracy, but also carer acceptability, including perceived usefulness, trust, digital literacy, transparency and data agency.³⁷ Cultural relevance and usability are also critical; tools that do not reflect carers' needs and contexts risk poor uptake.³⁸

Carer engagement and acceptability can be conceptualised using the Capability, Opportunity, Motivation-Behaviour (COM-B) model, where behaviour occurs when individuals have sufficient capability to undertake it, opportunity that enables it and motivation.³⁹ Situated within the Behaviour Change Wheel, COM-B

helps to identify barriers and facilitators that can inform intervention development and implementation.³⁹ It has been applied to primary-care dementia screening and management,⁴⁰ dementia care home interventions⁴¹ and understanding health behaviours among informal carers.⁴² It therefore provides a useful framework for comparing engagement with paper- and AI-based assessment tools.

This study used think-aloud interviews and the COM-B framework to explore carers' views of paper- and AI-based assessment tools, identify barriers and facilitators to engagement and develop recommendations for design and implementation.

Materials & Methods

Study Design

This exploratory qualitative study used think-aloud interviews to compare carers' views of the paper-based CBI-R and AI-based CognoSpeak™ and LUMEN.⁴³ Participants interacted with or observed each tool while verbalising real-time thoughts. Tools were presented in a fixed sequence, followed by semi-structured questions on usability, strengths and limitations. Interviews were audio-recorded, transcribed verbatim, anonymised and analysed using a hybrid inductive-deductive thematic approach,⁴⁴ where themes were developed inductively and subsequently mapped onto COM-B.³⁹

Ethical Considerations

Ethical approval was granted by the University of Sheffield (Ref: 058186; May 2024).

Participants and Recruitment

Participants were current or former informal carers of people with dementia, aged 18 or above, English-speaking and able to provide informed consent. Recruitment occurred from July 2024 to January 2025 through charity newsletters, community centres and Join Dementia Research.⁴⁵

Interested carers received a participant information sheet summarising the purpose of the study and their participation, and those who agreed completed a written consent form prior to participation. Seventeen expressed interest, twelve were eligible and consented and ten completed interviews. Non-participation was due to loss of contact, pre-interview withdrawal or ineligibility.

The sample size was guided by the principle of information power,⁴⁶ whereby the required sample size was informed by the specificity of the study aim and the richness and relevance of participants' accounts in addressing the research question. In this study, the focused

aim of understanding informal dementia carers' engagement with three specified assessment tools, alongside recruitment of a specific population with relevant caregiving experience, increased the information power of the sample. The interviews were also in-depth, lasting 85–117 minutes, and generated detailed accounts of usability, acceptability and barriers and facilitators to engagement. As analysis progressed, later interviews reinforced or elaborated on existing codes and themes rather than generating new ones, indicating sufficient thematic depth to address the research question and supporting the adequacy of the final sample.

Materials

The CBI-R is a 45-item paper-based informant questionnaire assessing changes on a 0–4 scale, with higher scores indicating greater impairment.²² It correlates well with clinical severity and differentiates dementia syndromes.⁴⁷ It represented a traditional assessment tool.

CognoSpeak™ is an AI-based stratification tool for early cognitive impairment.^{32,48} A digital agent asks memory-clinic-style questions, and responses are analysed using automatic speech recognition and machine learning models to extract linguistic and acoustic features.³² It represented AI-based speech assessment.

LUMEN (Language Usage for Monitoring Early Neurodegeneration) is a conversational chatbot that gathers collateral history from carers.³⁴ Using the Mistral-7B LLM, it asks open-ended clinical questions and receives typed responses. LUMEN represented an AI-based chatbot tool.

Data Collection

Think-aloud interviews⁴⁹ were used to capture real-time reactions. Participants first completed a brief practice exercise (solving a reasoning task) to familiarise themselves with verbalising thoughts.⁴⁹ They then completed each measure while thinking aloud, followed by semi-structured questions on understanding and usability.

Data Analysis

Codes and themes were developed inductively from participants' accounts,⁵⁰ and then mapped deductively onto COM-B.³⁹ This sequencing allowed carers' accounts to guide the initial inductive analysis before the findings were deductively interpreted through the COM-B framework, reducing the risk of prematurely constraining the analysis.⁵¹

Transcripts were familiarised and reviewed by LK to identify initial codes from participants' accounts.

These resultant codes included perceived barriers, facilitators, emotional responses and clinical value. Codes were compared and grouped into themes. Theme development was reviewed and refined by examining similarities and differences within and across the three tools to identify shared and divergent factors and then mapped onto the COM-B domains and subdomains. These subdomains were therefore used as organising categories rather than as themes generated from the data.

LK led the coding and theme development. Themes, COM-B mapping and interpretations of quotations were further reviewed and refined through ongoing discussion with AWG and LS to consider alternative interpretations, enhance analytic rigour and support a more systematic approach.⁵²

Reflexivity

The team comprised one British-Indian woman, two White British women, two White British men and one Chinese man, all with personal and/or professional experience of dementia. DB and JH were involved in developing CognoSpeak™ and LUMEN, respectively. We recognised that involvement in developing the technologies could shape expectations and interpretations, therefore, data collection, coding and initial theme development were led by LK, AWG and LS, who were independent of tool development and have training and experience in qualitative research. Members involved in tool development contributed complementary perspectives during later theme refinement and COM-B mapping but did not determine the initial codes or themes.

Participants were informed that researchers were members of the wider evaluation team and encouraged to provide critical feedback. During interviews, researchers used minimal, non-directive prompting and avoided defending or explaining features of the tools in response to criticism.⁴³ Positive, negative and ambivalent responses were retained. Reflexive diaries and team discussions were used to acknowledge potential bias. Participants' words are presented verbatim to ensure accurate representation. Nevertheless, we acknowledge that involvement of members in tool development remains a potential source of influence and should be considered when interpreting the findings.

Results

Ten carers participated, with nine interviews conducted online and one in-person, lasting 85–117 minutes (Table 1).

Table 1. Carers' Demographic Information (N = 10)

	Age	Gender	Ethnicity	Years of education	Relationship to care partner	Diagnosis of care partner	Time since diagnosis (years)
Carer 01	72	Female	White British	18	Wife	Alzheimer's disease	4 years
Carer 02	60	Female	White British	16	Daughter	Frontotemporal dementia	5 years
Carer 03	53	Male	White British	-	Ex-Spouse	Dementia	3 years
Carer 04	70	Male	White British	-	Brother	Frontotemporal dementia	4 years
Carer 05	47	Female	White British	-	Daughter	Mother: Alzheimer's disease	12 years
Carer 05	47	Female	White British	-	Daughter	Father: Alzheimer's disease	2 years
Carer 06	61	Female	White British	20	Daughter	Mixed dementia: Alzheimer's in FT region and vascular	3 years
Carer 07	-	Male	African (Somali)	13	Grandson	Dementia	2 years
Carer 08	42	Female	African (Somali)	8	Daughter	Dementia	5 years
Carer 09	69	Female	White British	20	Daughter	Alzheimer's disease	9 years
Carer 10	-	Female	African (Somali)	-	Granddaughter	Dementia	7 years

Where carers supported more than one person with dementia, details are reported separately. "-" indicates missing data. All diagnoses were carer-reported.

Pen-and-Paper Assessment

Carers' engagement with the CBI-R depended on its manageability and ability to represent observed behaviours and changes. Clear presentation and accessible language supported interpretation and confidence, whereas unfamiliar terminology and culturally specific assumptions reduced perceived accuracy and relevance.

Capability

Physical Capability. Carers described the CBI-R as physically "*straightforward*" to complete and requiring no specialist knowledge or skills. Its familiar pen-and-paper format supported independent completion. However, its perceived accessibility was limited by its presentation. Font size, spacing and page layout affected how easily participants could read and engage with the questionnaire, particularly older adults and those with visual impairments. As one carer explained, "*It could be made more user-friendly, maybe bigger writing*" (Carer 07, grandson).

Although the compact two-page format was convenient, several carers felt that the density of information made the questionnaire visually "*busy*," resulting in perceived pressure to complete it quickly rather than pause and reflect on individual items. One carer suggested that the appearance of a lengthy or crowded form might encourage respondents to "*just tick as many things as you want to get out there*" (Carer 07, grandson). Participants therefore suggested larger text, clearer spacing and more room between items to support reflective reporting.

Psychological Capability. Carers valued the simple instructions of the CBI-R, particularly the defined recall

period ("*past month*"), as it provided a clear timeframe against which to judge the frequency of behaviours, supporting accurate completion. They also regarded the breadth of the questionnaire as important because it moved beyond memory difficulties and captured the varied ways in which dementia presents. One carer described it as offering "*more of a 360-degree view*" that reflected her mother's particular dementia rather than "*a stereotypical presentation or the obsession with memory*" (Carer 06, daughter).

Completing the questionnaire also prompted carers to reflect systematically on changes they might otherwise have accepted or overlooked in everyday life. In particular, it encouraged them to consider whether a behaviour represented a dementia-related change or ordinary variation: "*It makes you question behaviour, is it dementia or is it normal?*" (Carer 08, daughter). This reflective process helped to strengthen their ability to identify and communicate changes over time. However, confidence was reduced when important aspects of the person's presentation were absent, including relational changes such as "*how affectionate they were.*"

The Likert scale had mixed effects on psychological capability. Some carers found it intuitive and helpful for quantifying how often behaviours occurred, however, others believed the fixed categories were too rigid, especially when behaviours fluctuated over time. Distinctions between response options such as "*daily*" and "*constantly*" were not always clear, causing frustration and concern about accurate reporting.

Language and terminology presented a further constraint. Although the questionnaire was manageable for most participants, carers considered some instructions dense and identified terms such as "*orientation*" and

“catnaps” as difficult for people with lower literacy or those who did not speak English as a first language: *“Some people might be disadvantaged by the level of language and structure. You’ve got to have a high level of literacy”* (Carer 09, daughter). Carers therefore recommended simpler wording, explanations of unfamiliar terms and translated versions to allow carers from different linguistic backgrounds to participate.

Participants also questioned whether some items reflected culturally specific assumptions about appropriate behaviours. One participant commented, *“We make tactless remarks, but that’s from a Western standpoint; for us it’s the norm”* (Carer 08, daughter). This created uncertainty about whether carers were reporting and judging behaviours against culturally defined expectations, highlighting the role of cultural framing in interpreting items.

Opportunity

Physical Opportunity. Specific design elements influenced carers’ ability to provide information that accurately reflected their relative’s state. Carers valued the inclusion of a *“not applicable”* response as it reduced pressure to answer irrelevant items. However, the limited space provided for additional comments constrained their ability to explain complex, fluctuating experiences. As one participant stated, a small text box would have made her *“less expressive,”* leading her to reduce her account to *“a few sentences, a sentence at best”* (Carer 10, granddaughter). Participants therefore suggested larger comment boxes to support fuller explanations, although some cautioned that excessive blank space could appear intimidating and discourage completion.

Carers also emphasised that written responses needed to be discussed with clinicians. Standardised scores could not always convey the severity, context or variability of behaviours and, when interpreted in isolation, risked producing an inaccurate impression. One participant explained that, based on the questionnaire alone, *“you might think my dad’s better than reality,”* because important details *“may not be obvious from the answers given”* (Carer 05, daughter). Follow-up conversations therefore provided an important opportunity to clarify responses and enabled carers to translate their observations into clinically meaningful information.

Overall, carers were knowledgeable and highly motivated to contribute, yet were unable to communicate these observations adequately when response options were restrictive or opportunities for discussion were absent. Engagement was further undermined when questionnaires were not collected at all or visibly integrated into diagnostic pathways.

Social Opportunity. The CBI-R provided a socially acceptable and validating space to disclose sensitive or

stigmatised behaviours. By naming these behaviours explicitly, the questionnaire signalled that they were clinically relevant and could be discussed openly. One participant recalled feeling *“relieved”* that she could acknowledge, *“yes, she does do that, this is real, it’s embarrassing”* (Carer 02, daughter).

However, carers with limited English literacy or formal education could struggle to complete the questionnaire, even when they had the greatest day-to-day knowledge of the person being assessed. In some families, responsibility therefore shifted to younger relatives who were more confident in English but less directly involved in care. As one grandson explained, older relatives *“have more interaction with our grandma and would have better answers,”* but *“because of their lack of language, they pass it over to us”* (Carer 07, grandson). This created a tension between the capability to complete the questionnaire and the experiential knowledge required to answer it accurately, highlighting the need for clinicians to speak to carers in preferred languages and provide translated versions.

Carers also felt that standardised formats failed to account for individual personality and premorbid behaviours, instead *“packing everyone in one systematic space.”* Generic questions and rigid scoring made it difficult to distinguish symptoms from baseline. One daughter criticised the application of *“generic testing”* that gave *“no attention to my mum’s character as an individual and what was normal for her”* (Carer 09, daughter).

Motivation

Automatic Motivation. Emotional responses strongly shaped motivation and subsequently, engagement. Frustration with long diagnostic delays increased motivation to participate in any assessment that might accelerate the process. One carer, reflecting on a 9-month wait for diagnosis, felt that *“it should have been quicker”* (Carer 10, granddaughter). In this context, the questionnaire could offer a sense of progress and an opportunity to contribute actively to the assessment.

However, the perceived consequences of carers’ answers also generated anxiety. Some worried that rating behaviours too mildly might prevent their relative from receiving a diagnosis, whereas rating them too severely might misrepresent or stigmatise the person. One daughter described questioning whether an answer would mean that her mother *“won’t get a diagnosis”* or would *“look worse than she is,”* making completion *“really anxiety-provoking”* (Carer 06, daughter). These concerns could lead carers to minimise or exaggerate symptoms rather than report them as accurately as possible. Minimisation was sometimes driven by affection, loyalty or a desire to protect the person being assessed. As one participant explained, a carer may try to protect their partner because

“it might feel horrible to say that” about them (Carer 02, daughter). Emotional capability was therefore important: carers needed to feel able to acknowledge distressing or stigmatised behaviours, while also trusting that their responses would be interpreted sensitively and in context.

The CBI-R’s visual design also triggered immediate negative responses. The dense, black-and-white design was described as “boring” or overwhelming and could encourage rushed completion or avoidance.

Reflective Motivation. Alongside emotional reactions, carers made judgements about the value, relevance and purpose of the CBI-R. Many appreciated its comprehensive approach, recognising it reflected the complex emotional, behavioural and routine-based changes associated with dementia. One carer described the questionnaire as attempting to obtain “a holistic view of the person emotionally, routines and things” (Carer 07, grandson). This breadth enhanced its perceived relevance and encouraged carers to consider aspects of their relative’s behaviour that might otherwise have been overlooked.

Being invited to contribute also increased motivation by signalling that their experiences mattered and contributed to their relatives’ care: “It’s good to gather feedback from people that are involved. You can improve things” (Carer 07, grandson). Recognition of carer expertise strengthened reflective motivation while also increasing social opportunity by legitimising their role within the assessment process.

However, motivation declined when carers doubted whether questionnaires truly informed clinical decisions. Questionnaires introduced late in the diagnostic pathway, collected without discussion or followed by no visible action were seen as procedural rather than meaningful. One carer described the process as “pointless” when she believed that her concerns were “not going anywhere,” leading her to report that “everything’s good” instead (Carer 10, granddaughter). Overall, when the CBI-R felt consequential, carers engaged thoughtfully but when perceived as tokenistic, motivation and accuracy of responses decreased.

Artificial Intelligence-Based Tools

Across the accounts, engagement with AI-based tools was shaped by interface design, digital literacy and attitudes towards AI.

Capability

Physical Capability. CognoSpeak’s speech-based format was widely appreciated as it reduced the physical and cognitive effort associated with written responses. One carer described the interaction as “a flow of conversation rather than actual questions,” while another emphasised

that “all you’re doing is just speaking into it, you don’t have to write” (Carer 10, granddaughter). The conversational format therefore appeared to make the assessment feel more natural and accessible, particularly for users who may find typing or writing difficult. However, speech was not universally accessible. Visual, hearing or speech impairments could limit access: “people would find that difficult, not because they’ve got dementia; people lose their capacity to speak and read aloud” (Carer 09, daughter).

Psychological Capability. Psychological capability was shaped by digital literacy, interface clarity and linguistic and cultural fit. Younger carers generally expressed confidence with technology: “I grew up in the era of technology, so it’s easy” (Carer 07, grandson). At the same time, they recognised that older carers may require additional support: “tend to happen to older people, the technology aspect would be challenging for them” (Carer 07, grandson). Digital literacy was therefore viewed as a potential source of inequality, with users who had less experience of technology likely to require additional support. Nevertheless, both tools were generally described as simple and visually uncluttered. CognoSpeak’s conversational style evoked familiar face-to-face interactions with one carer comparing it to approaching a receptionist and asking, “How do I go about doing that?” (Carer 10, granddaughter). This familiarity appeared to reduce uncertainty about how to interact with the tool and strengthened users’ confidence in completing the assessment.

CognoSpeak’s verbal interaction was preferred for older adults and expected to produce “better answers” and “capture insights better” (Carer 05, daughter), whereas LUMEN’s typed input was considered restrictive for those with limited typing skills, less confident in grammar and written expression or difficulty capturing fleeting, detailed or emotional thoughts. Despite these differences, carers valued the ability to provide open-ended responses rather than being limited to fixed-choice questionnaires. This flexibility allowed them to describe the person they cared for in their own words and reduced the impression that there was a single correct answer: “I would have the opportunity to say something about her rather than being made to fit into a box of yes or no; it’s not a test” (Carer 06, daughter).

However, linguistic and cultural accessibility remained concerns: “You need a certain level of English to be able to deal with this” (Carer 07, grandson). Some assessment tasks were also viewed as lacking cross-linguistic validity. For example, a letter-fluency task was questioned because “P doesn’t exist in my language, Somali” (Carer 10, granddaughter). Other questions, including those concerning employment or engagement with Western news,

were considered poorly aligned with the experiences of some communities. Reflecting on an employment-related question, one carer observed that *“a lot of women in the Somali community don’t usually work,”* and questioned how informative the item would be and how appropriate it was to ask (Carer 10, granddaughter).

Opportunity

Physical Opportunity. Physical opportunity was shaped by features of the user interface, practicality and technical quality of the tools. For example, CognoSpeak’s conversational, *“game-like”* tasks were described as engaging, whereas LUMEN’s interface was criticised as *“dull,” “wordy,”* and reminiscent of *“hospital forms,”* reducing engagement. These design features affected more than usability: interfaces that appeared overly clinical or text-heavy were perceived as less inviting and could reduce users’ motivation to continue. Participants also had mixed reactions to CognoSpeak’s avatar; some valued the presence of a visible *“person,”* which made the interaction more familiar and conversational, while others were distracted by the ‘robotic’ voice, repetitive phrasing and imperfect synchronisation between the avatar’s speech and facial movements. One carer remarked that *“the mouth speaking with the questions was slightly out”* and found this *“off-putting”* (Carer 05, daughter). Technical imperfections could therefore undermine the sense of social presence that the avatar was intended to create, reducing trust and engagement even when the tool was otherwise straightforward to use.

Nevertheless, AI-based tools were valued for the option to complete at home, reducing the time and effort required to prepare for and travel to appointments: *“there’s a lot of pre-prep to leave for appointments; imagine how easy it would be to set up a laptop at home, where you’re more relaxed”* (Carer 10, granddaughter). Carers also suggested that home-based assessments would allow them to provide a more representative account of their relative’s difficulties, as they might feel less anxious and more able to express their insights and observations without the person being assessment present. However, the benefits of remote access depended on the availability of suitable devices, reliable internet connections and practical assistance: *“some of these people don’t have internet at home; they’re your target group”* (Carer 08, daughter). Remote delivery could therefore reduce or reinforce inequalities.

Despite these barriers, carers believed AI-based tools could help address delays within overstretched diagnostic pathways by enabling earlier assessment and faster access to support. Reflecting on their own experience, one carer suggested that such a tool *“might have given us a diagnosis quicker,”* enabling *“quick access”* to beneficial services (Carer 07, grandson). This perceived potential increased motivation to engage with the tools, particularly

when carers saw them as a route to timely help rather than as an additional administrative task. However, participants stressed that these benefits would depend on effective publicity and integration into clinical pathways.

Social Opportunity. AI-based tools enabled more active carer participation. Several preferred them to traditional paper-based approaches: *“I say, always do CognoSpeak with me rather than paper tests”* (Carer 02, daughter). This preference appeared to reflect both the accessibility of the conversational format and the opportunity it provided for carers to contribute observations in a less restrictive way.

Carers consistently emphasised the importance of having their knowledge recognised. One participant highlighted that *“carers need to be given the opportunity to talk,”* adding that *“even a less than perfect solution is better than nothing”* (Carer 09, daughter). Their involvement was seen as diagnostically valuable as they could contextualise responses, identify discrepancies and provide information that the person being assessed might omit or be unable to recall. Another participant explained, carers could *“validate the information and fill in gaps,”* thereby strengthening *“the method, diagnosis and treatment”* (Carer 07, grandson). Some also suggested that AI could collect accounts from multiple relatives or supporters and provide *“a wealth of observation”* over time (Carer 02, daughter).

Generational and cultural attitudes towards conversational AI also mattered. Some regarded it as a *“young person’s thing,”* potentially alienating older adults. Others expressed unease with AI substituting human interaction. CognoSpeak’s avatar partly mitigated these concerns by creating a more human-like interaction. Some participants described it as *“calming”* and less intimidating than traditional assessment because it removed subtle forms of perceived judgement that occur during face-to-face consultations. One carer recalled that clinicians’ facial expressions during clinical appointments had left them wondering *“Did I say something wrong?”* (Carer 07, grandson). The absence of such reactions was seen as potentially encouraging greater openness, although this benefit depended on the avatar appearing responsive and credible. Small gestures and acknowledgements such as *“thank you”* created an impression that the avatar was listening, whereas praise such as *“you’re doing very well”* was considered *“condescending”* and inauthentic.

The possibility of avatar customisation was viewed positively, allowing alignment with personal or cultural preferences. One carer commented that *“it’s nice to be able to see somebody who looks like you”* (Carer 10, granddaughter), and participants suggested that some users might feel more comfortable speaking to an avatar of the same gender or cultural background. Representation therefore increased social opportunity by promoting

familiarity and a sense of inclusion among communities that may feel marginalised within healthcare settings. However, avatar selection exposed assumptions about race, identity and trust. Some participants questioned whether the available choices encouraged simplistic cultural categorisation, with one carer observing that there were “*an awful lot of presumptions*” involved in expecting her to choose an avatar because it was “*white and a woman*” (Carer 09, daughter). Other responses reflected racialised and discriminatory beliefs, including an assumption that selecting a white avatar would ensure “*a normal face that I’m likely to get normal language and relate to*” (Carer 04, brother). Such accounts highlight that avatar customisation does not operate in a socially neutral context and that although it may enhance comfort, it can also activate prejudice.

By contrast, carers’ willingness to engage with LUMEN was often shaped by previous experiences with commercial chatbots, often associated with delays, repetitive and ineffective exchanges. One carer compared LUMEN to “*agonising, lengthy discussions with the bank chatbots*,” where a person would resolve the problem faster (Carer 01, wife).

Motivation

Automatic Motivation. Emotional responses ranged from enjoyment and reassurance to irritation and fear. Person-centred prompts about interests focused on identity and strengths rather than solely on cognitive deficits: “*I started smiling when he said, ‘What are your hobbies?’*” (Carer 08, daughter). However, “*tell me about yourself*,” was too broad: “*I would just not say anything. So perhaps break it down more rather than ‘tell me about yourself’*” (Carer 04, brother). The emotional effect of a prompt was therefore closely linked to psychological capability: clear structure supported responses, whereas ambiguity discouraged participation.

As previously mentioned, acknowledgements such as “*thank you*” enhanced reassurance that responses were received, however repeated praise became less convincing over time: “*he’s beginning to irritate me more; he’s got no normal human mannerisms, just patronising*” (Carer 01, wife). Suggestions that the avatar should “*smile more*” or “*move his arms otherwise he just looks like a robot*,” reflected a desire for more natural social cues.

Interaction that more closely resembled real-life consultation made participants feel calmer and more willing to engage, in particular, seeing CognoSpeak’s avatar created a sense of reciprocity and reduced the impression that users were simply supplying information to a machine: “*You can see someone, feel like you are talking to someone, like someone’s listening*” (Carer 07, grandson).

Emotional responses were also shaped by broader fears about AI, privacy, hacking and the security of sensitive health information reduced motivation: “*I’m scared. It only needs a few bad actors and the whole thing goes awry*,” adding that if hackers could access highly secure institutions, they could “*beat the security on my phone*” (Carer 01, wife). These anxieties reduced automatic motivation even when potential clinical benefits were acknowledged.

Reflective Motivation. Reflective motivation centred on perceived usefulness, reliability and appropriate role of AI in healthcare. Carers’ sense of duty to act in their relative’s best interests encouraged use of tools that could improve diagnostic accuracy, speed up diagnosis or support future planning. Carers particularly valued the potential for AI to collect richer narratives that would otherwise be difficult to collect during time-limited clinical appointments. LUMEN was therefore seen as potentially enabling clinicians to base decisions on “*rich narrative history rather than fragmented reports*” (Carer 02, daughter). This capacity to gather a “*360-degree view*” strengthened the perceived usefulness of the tools and reinforced carers’ motivation to contribute.

However, confidence was conditional. Although many participants were optimistic about earlier detection and more efficient diagnostic pathways, they also feared misinformation, inaccurate conclusions and excessive clinical reliance on automated outputs: “*I remain suspicious in terms of misinformation as long as its peer proven*” (Carer 09, daughter).

Judgements about usefulness and credibility were also shaped by cultural relevance. Formal prompts such as “*tell me about yourself*” were criticised for sounding like a job interview rather than a supportive health assessment, prompting “*Is that appropriate?*” (Carer 07, grandson). Questions misaligned with linguistic, cultural or social realities undermined both answerability and confidence in conclusions.

Overall, carers were cautiously optimistic, provided AI supported rather than replaced clinicians: “*People need people*” (Carer 09, daughter). Engagement was therefore motivated by the prospect of combining the efficiency and information-gathering capacity of AI with the empathy and judgement of human clinicians (Table 2).

Discussion

This study explored dementia carers’ views on a traditional pen-and-paper questionnaire (CBI-R) and two AI-based tools (CognoSpeak™ and LUMEN), using the COM-B framework³⁹ to understand barriers and facilitators to engagement. Paper questionnaires were perceived as familiar and straightforward but also

Table 2. Facilitators and Barriers for the CBI-R, CognitoSpeak and LUMEN Mapped Across Capability, Opportunity and Motivation (Carer Perspectives)

	Pen-and-paper assessments	Artificial Intelligence (AI)-based tools	
COM-B framework	Cambridge Behavioural Inventory-Revised	CognitoSpeak™	LUMEN
Capability	+ Familiar, easy to complete; broad domains supports reflection - Dense layout; jargon and ambiguous response options; limited narrative space; language/cultural mismatch.	+ Speech reduced writing demands; Conversational and open-ended - Sensory impairments; digital literacy; English-only	+ Typing allows considered responses; Open-text responses allow detail - Typing, literacy and digital literacy demands; text-heavy interface
Opportunity	+ Familiar format; Validates carer input; “Not applicable” option - Limited comment boxes; rigid categories; language barriers may shift completion to less-involved relatives	+ Remote avoids reduced travel; Engaging tasks and avatar may support comfort - Requires device/internet; technical imperfections; low awareness and avatar-related concerns may reduce trust	+ Remote screening and scalable; Written responses can be reflective; Supports multiple perspectives. - Requires device/internet; low-engagement interface; negative chatbot associations
Motivation	+ Perceived as holistic; Carers value being asked to contribute to assessment. - Anxiety, guilt and protectiveness may bias scoring; “boring” format reduces engagement; Unclear clinical impact	+ Conversational interaction and perceived potential for earlier assessment encourage engagement - Robotic or patronising feedback; trust conditional on clinical evidence	+ Perceived potential to capture narratives missed in clinics - Chatbot fatigue; limited warmth; concerns about accuracy and validation

rigid, linguistically demanding and culturally narrow. AI-based tools enabled more flexible, narrative and remote reporting, but concerns about digital literacy, privacy and loss of human contact limited engagement. Across both formats, participants emphasised accessibility, cultural sensitivity and human-led care.

Pen-and-Paper Assessments

Carers valued the CBI-R's clear recall period, multi-domain coverage and recognised the diagnostic utility of informant reports, consistent with evidence that they provide ecologically valid insights into everyday functioning and complement cognitive tests.⁵³ Tools such as the AD8 and IQCODE demonstrate good sensitivity (88-97% and 80-91% respectively) and moderate specificity (64-81% and 66-85% respectively),²¹ and can predict conversion from MCI to dementia, while correlating with neuropsychological decline and biomarkers.⁵⁴ The CBI-R similarly demonstrates strong concurrent validity.⁵⁵

However, carers struggled to compress nuanced, context-dependent behaviours into a single 0-4 rating, increasing cognitive effort and concerns about inaccurate reporting. Given existing carer burden, such complexity risks disengagement and may contribute to lower specificity reported in informant measures, (IQCODE-16 as low as 38%),²² resulting in false positives and unnecessary referrals.⁵⁶

Structural and linguistic barriers were prominent. Dense instructions, small font and technical terminology (eg, “orientation”, “catnaps”) could exclude carers with limited literacy or English proficiency. This reflects evidence that lengthy questionnaires have low completion rates.⁴⁷ Many tools also assume literate, English-speaking respondents, so lower education and non-native language use can therefore affect performance and classification.⁵⁷ Cultural mismatch may further undermine validity, where items reflecting Western lifestyle norms do not match an individual's usual activities or background, particularly in rural or low-income contexts, potentially inflating apparent decline.⁵⁸

Carers valued the CBI-R for legitimising their observations, but questioned whether responses meaningfully informed diagnostic decisions. This perceived disconnect between effort and impact represents an opportunity-level barrier and may reduce motivation to engage. Conversely, frustration with diagnostic delays and a desire to contribute to better care encouraged participation. However, anxiety, guilt and protective instincts sometimes led to over- or under-reporting, consistent with literature that informant ratings can be influenced by carer mood, burden and relationship dynamics.^{24,25} Together, these influences highlight the importance of parallel patient- and informant-reporting.

Artificial Intelligence-Based Tools

Participants appreciated CognoSpeak and LUMEN for reducing physical writing demands and allowing more natural, narrative-driven responses. Remote participation could also reduce travel burden and provide opportunities to contribute information outside time-limited appointments. Conversational AI may therefore offer a more flexible way of collecting collateral histories.^{59,60} Related technologies are already used in smart assistants (Amazon Alexa and Google Home) to support daily routines and are being explored for safety monitoring and carer alerts, although their relevance to dementia assessment requires separate evaluation.⁶⁰

However, capability barriers were significant. Sensory impairments, low digital literacy, slow typing skills and culturally inappropriate tasks were viewed as exclusionary. These concerns reflect wider debates on bias and equity in AI, as many dementia speech models are trained on small, unrepresentative datasets,³⁵ or collected in controlled environments that do not reflect real-world settings. Models trained in English often perform poorly in other languages and dialects, and the “digital divide” presents a major equity concern. Older adults and socioeconomically disadvantaged carers face disproportionate barriers to access and use.^{61,62} Without careful design and support, AI risks excluding those it aims to serve.

Nevertheless, AI enhances opportunity by enabling remote participation and allowing carers to share insights through open-ended prompts. However, limited access to devices, low awareness and concerns that it lacked sufficient human-like qualities to foster trust were barriers. Overall, trust emerged as central, with carers raising concerns about surveillance, data misuse and transparency in how outputs were generated.⁶⁰ Many felt “*not ready*”, suggesting acceptance may increase as people become more digitally familiar.⁶¹ Emotional authenticity also influenced trust and engagement: avatars were described as both “*calming*” and “*condescending*”. Therefore, warmth and anthropomorphism may support interaction, but poorly executed social cues undermined credibility.⁵⁹

Motivation was strengthened by optimism that AI might support earlier identification of cognitive changes and timely diagnosis. Previous speech-based models have reported 78–89% accuracy in distinguishing Alzheimer’s disease from healthy controls using spontaneous speech tasks,⁶³ while baseline speech has shown prognostic potential, predicting cognitive decline over two years (F1 = 66.7%) and future MMSE scores.⁶⁰ Multimodal approaches combining acoustic and linguistic features have also shown promise in predicting progression from MCI to dementia.⁶⁴ However, these findings do not establish the clinical effectiveness of CognoSpeak or LUMEN, nor

demonstrate that routine use would improve diagnostic accuracy, time-to-diagnosis or patient and carer outcomes.

Scepticism and fear of AI replacing human-led care reduced motivation, echoing wider ethical debates regarding bias, consent, transparency and data repurposing.^{65,66} Speech and free-text data may reveal identity or other sensitive characteristics beyond their intended clinical purpose, requiring clear governance of data storage, access, secondary use and model training.⁶⁷ Algorithmic bias may also reinforce inequalities where training data do not adequately represent different languages, cultural groups, educational backgrounds or rarer dementia subtypes.⁶⁸ False-positive results could cause anxiety and unnecessary investigation, whereas false-negative results could delay diagnosis and access to support. AI outputs should therefore be treated as probabilistic indicators and interpreted with meaningful human oversight.⁶⁹

Overall, carers valued traditional tools for their clarity but criticised their rigidity, while AI enabled richer narratives and remote access, but introduced new concerns around equity, trust and privacy.

Implications for Design

The usability of pen-and-paper assessments depends on clinical content and accessibility. Plain language, accessible formatting and items anchored to an individual’s previous abilities and routines may reduce difficulties with comprehension and rating. Explaining how informant responses contribute to diagnosis may also strengthen perceived usefulness and encourage reflective completion. These priorities are summarised in [Table 3](#).

AI-based tools require additional attention to digital access, transparency and governance. Multimodal participation and supported onboarding may reduce barriers associated with sensory impairment or limited digital literacy. Multilingual and culturally co-designed tasks are also needed to avoid reproducing inequalities present in paper-based assessments. Trust may be strengthened by transparent explanations of data use and explicit positioning of AI as supporting clinical decision-making (see [Table 4](#)).

Considerations for Future Clinical Pathways

Future studies could evaluate hybrid pathways in which carers provide collateral histories remotely before appointments, potentially from more than one informant. Participants anticipated that this could allow clinicians to review concerns in advance and use face-to-face time for more targeted discussion. Multi-informant workflows may also reduce bias and burden by collating multiple carers’ perspectives, aligning with calls for digital triage to

Table 3. Implications for the Design of Pen-and-Paper Assessments

Identified Limitations	Implications for Tool Design
Complex and technical language	• Use brief, plain language prompts and provide accessible definitions or examples where needed • Pilot with diverse groups (eg, language, cultural backgrounds, age, accessibility, socioeconomic background) • Use rigorous translation procedures
Culturally narrow assessment items	• Anchor questions to the individual's previous functioning and usual routines, rather than culturally specific standards • Adapt using cultural co-design, forward/back translation and cognitive debriefing
Rigid or Ambiguous responses options	• Use clearly differentiated response anchors and brief worked examples • include options such as “not applicable” or “never performed this activity” • Provide free-text space to add detailed additional comments
Unclear Purpose and clinical Use	• Explain why the information is being collected, how it will be reviewed and how it contributes to clinical decision-making • Provide appropriate reassurance about value of responses (eg, progress trackers, reassurance statements) to enhance motivation and support completion
Unengaging visual design	• Accessible formatting, including large fonts, adequate spacing and clear section headings • Use of colour to enhance engagement, with different options to align with preferences

Table 4. Implications for the Design of AI-Based Assessments Tools

Identified Limitations	Implications for Tool Design
Complex or lengthy Prompts	• Use brief, plain language prompts • Provide definitions, examples or optional help functions
Culturally or linguistically inappropriate task	• Provide multilingual interfaces and culturally appropriate task alternatives • Use cultural co-design, translation-back translation and cognitive debriefing procedures
Digital literacy, sensory and accessibility barriers	• Offer multimodal input options (eg, speech, typing and file upload). • Implement accessible design features such as larger fonts, visual cues and adjustable audio • Provide supported in-clinic routes, onboarding, training and access to appropriate equipment where required
Emotional design and interaction style	• Enable appropriate personalisation (eg, avatar) • Use natural pacing and context-sensitive feedback • Clearly disclose that the system is non-human and avoid design features that imply human equivalence
Uncertainty about accuracy, data Use and clinical influence	• Explain what is measured, how data are stored and used and how outputs contribute to clinical decisions • Present relevant validation evidence • Ensure that results are reviewed by appropriately trained clinicians, helping to position AI as supporting, rather than replacing, clinical judgement
Limited awareness and Uptake	• Integrate tools into established clinical pathways and provide information and onboarding through health and community services • Maintain non-digital participation routes to support equitable access

improve diagnostic timeliness and efficiency.⁷⁰ However, prospective comparative studies would be required to determine whether digitally collected collateral histories improve information quality, diagnostic decision-making, time-to-diagnosis, workflow burden or patient and carer experiences compared with existing practice.

Equity should be central to pathway evaluation. Offering at-home or supported in-clinic routes, digital literacy support, accessible onboarding, interpreters and

access to appropriate devices may reduce exclusion, particularly among older carers, those with sensory impairments, non-English speakers and people with limited access to technology.

Future studies should also assess how clinicians interpret and use AI-generated information, including whether training is needed for consistent and appropriate interpretations. Pathways should prioritise transparency and user control by explaining what data is collected, how

outputs are generated and how information contributes to clinical decisions. Providing carers with opportunities to review or amend their responses before submission may support trust, although effects on data quality and engagement require evaluation.

AI should therefore be embedded within care rather than positioned as a standalone solution. The COM-B model, central to the Behaviour Change Wheel (BCW),³⁹ identifies what people need to engage with a behaviour and guides appropriate intervention functions.⁴² Using the BCW framework, potential implementation could combine education (carer and clinician onboarding), enablement (multi-informant access, safeguarding flags) and environmental restructuring (on-site kiosks, interpreters), to support equitable participation. However, whether AI-based tools actually improve the quality of collateral histories, streamline clinical workflows or reduce diagnostic delays remains to be established through prospective implementation and effectiveness studies.

Limitations

This study involved a small, UK-based sample that was predominantly White British (70%) and English-speaking, limiting transferability. Minority ethnic perspectives, non-English speakers, carers from low-income or rural settings and those supporting rarer dementia subtypes were underrepresented. Although years of education were recorded, detailed socioeconomic circumstances and prior familiarity with technology was not collected, limiting our ability to contextualise findings relating to digital literacy. Caregiving intensity, co-residence and a standardised measure of carer burden were also not assessed. These factors may influence available time, emotional responses, perceptions of assessment and willingness or ability to engage with digital technologies, limiting interpretation of the findings. Future research should address some of these limitations to ensure that the full caregiving context is presented.

Recruitment through Join Dementia Research may have favoured participants more comfortable with research and technology, potentially overestimating AI acceptability. During the interview, all participants reviewed the tools in the same fixed sequence, which may have introduced order, contrast or carryover effects. Interviews were also relatively long, lasting 85–117 minutes, so fatigue may have affected responses to tools presented later.

Future Directions

Future work should prioritise co-design and usability testing with carers from diverse linguistic, cultural and socioeconomic backgrounds, followed by prospective

validation and implementation studies examining diagnostic accuracy, time to diagnosis, clinician workload, carer burden, acceptability, cost and equity. Psychometric testing across languages and cultural groups should precede widespread implementation, alongside exploration of low-resource delivery models (eg, offline capture with later upload). Research should also examine how best to communicate validation, limitations and data use, and how different explainability strategies influence trust and engagement.

Conclusion

This study identified carer-perceived barriers and facilitators to engagement with paper- and AI-based dementia assessment tools. Carers emphasised their role as active partners in assessment; collateral histories should therefore be collected systematically, considered alongside other clinical information and interpreted within human-led care. Their perspectives should also inform the design, implementation and evaluation of future tools to reduce exclusion and ensure alignment with caregiving realities.

Acknowledgements

This work was funded by the NIHR BRC, part of LK's PhD studentship. We thank the carers who generously shared their time and experiences. We also thank Research Champions, Muse Jama and Sahra Abdi and Join Dementia Research for their support with recruitment.

ORCID iDs

Labhpreet Kaur  <https://orcid.org/0009-0002-5509-4270>

Alys Wyn Griffiths  <https://orcid.org/0000-0001-9388-9168>

Ethical Considerations

Ethical approval was granted by the University of Sheffield (Ref: 058186; May 2024).

Consent to Participate

Informed consent was obtained from all participants prior to their involvement in the study.

Consent for Publication

All participants provided consent for publication of the findings.

Author Contributions

LK: Conceptualisation; Methodology; Formal analysis; Investigation; Writing. AWG: Conceptualisation; Methodology; Formal analysis; Review & Editing. LS: Formal analysis; Review & Editing. JH: Supervision; Review & Editing. XS: Supervision; Review & Editing. DB: Conceptualisation; Supervision; Review & Editing.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Institute for Health and Care Research (NIHR) Sheffield Biomedical Research Centre (BRC) as part of LK's PhD studentship. The views expressed are those of the authors and not necessarily those of the NIHR.

Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: JH serves as an Expert Reviewer for the Medicines and Healthcare Products Regulatory Agency (MHRA) and as Clinical Advisor for Akkrivia Health. JH is the chief investigator for LUMEN, and DB is clinical co-lead of CognoSpeak, therefore both involved in the development of these tools. The other authors declare no conflicts of interest.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author (LK) upon reasonable request.

References

- Browne B, Nuriye K, Moore KJ, Sampson EL, Davies N. Defining end of life in dementia: a systematic review. *Palliat Med.* 2021;35(10):1733-1746. doi:10.1177/02692163211025457.
- Xue C, Kowshik SS, Lteif D, et al. AI-based differential diagnosis of dementia etiologies on multimodal data. *Nat Med.* 2024;30:1-13. doi:10.1038/s41591-024-03118-z.
- Couch E, Co M, Albertyn CP, Prina M, Lawrence V. A qualitative study of informal carer perceptions of the benefits of an early dementia diagnosis. *BMC Health Serv Res.* 2024;24(1):508. doi:10.1186/s12913-024-10957-6.
- Davies-Abbott I, Anthony BF, Jackson K, Windle G, Edwards RT, Rare Dementia Support RDS Research Team. The diagnostic pathway experiences of people living with rare dementia and their family carers: a cross-sectional mixed methods study using qualitative and economic analyses. *Int J Environ Res Publ Health.* 2024;21(2):231. doi:10.3390/ijerph21020231.
- Visser W, Eersel V, Hester LH, Perry MB, van Munster BC. Doing the right thing? General practitioners' considerations in achieving a timely dementia diagnosis. *Int J Geriatr Psychiatry.* 2024;39(9):e6142. doi:10.1002/gps.6142.
- Liss JL, Seleri Assunção S, Cummings J, et al. Practical recommendations for timely, accurate diagnosis of symptomatic Alzheimer's disease (MCI and dementia) in primary care: a review and synthesis. *J Intern Med.* 2021; 290(2):310-334. doi:10.1111/joim.13244.
- van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in early Alzheimer's disease. *N Engl J Med.* 2023;388(1):9-21. doi:10.1056/NEJMoa2212948.
- Sims JR, Zimmer JA, Evans CD, et al. Donanemab in early symptomatic alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. *JAMA.* 2023;330(6):512-527. doi:10.1001/jama.2023.13239.
- Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention and care: 2024 report of the Lancet standing Commission. *Lancet.* 2024;404(10452):572-628. doi:10.1016/s0140-6736(24)01296-0.
- Truong QC, Cervin M, Choo CC, et al. Examining the validity of the mini-mental State Examination (MMSE) and its domains using network analysis. *Psychogeriatrics.* 2023; 24(2):259-271. doi:10.1111/psyg.13069.
- Arevalo-Rodriguez I, Nadja S, Roqué M, et al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Libr.* 2015;2015(3): CD010783. doi:10.1002/14651858.cd010783.pub2.
- Gallegos M, Morgan ML, Cervigni M, et al. 45 years of the mini-mental state examination (MMSE): a perspective from ibero-america. *Dementia & Neuropsychologia.* 2022;16(4): 384-387. doi:10.1590/1980-5764-dn-2021-0097.
- Arvanitakis Z, Shah RC, Bennett DA. Diagnosis and management of dementia: review. *JAMA.* 2019;322(16): 1589. doi:10.1001/jama.2019.4782.
- Islam N, Hashem R, Gad M, et al. Accuracy of the Montreal Cognitive Assessment tool for detecting mild cognitive impairment: a systematic review and meta-analysis. *Alzheimer's & Dementia.* 2023;19(7):3235-3243. doi:10.1002/alz.13040.
- Danquah MO, Yan E, Lee JW, et al. The utility of the Montreal cognitive assessment (MoCA) in detecting cognitive impairment in surgical populations – a systematic review and meta-analysis. *J Clin Anesth.* 2024;97:111551. doi:10.1016/j.jclinane.2024.111551.
- Ratcliffe LN, McDonald T, Robinson B, Sass JR, Loring DW, Hewitt KC. Classification statistics of the Montreal Cognitive Assessment (MoCA): are we interpreting the MoCA correctly? *Clin Neuropsychol.* 2022;37(3):562-576. doi:10.1080/13854046.2022.2086487.
- Galvin JE. Using informant and performance screening methods to detect mild cognitive impairment and dementia. *Current Geriatrics Reports.* 2018;7(1):19-25. doi:10.1007/s13670-018-0236-2.
- Speechly CM, Bridges-Webb C, Passmore E. The pathway to dementia diagnosis. *Med J Aust.* 2008;189(9):487-489. doi:10.5694/j.1326-5377.2008.tb02140.x.
- Chen H-H, Sun F-J, Yeh T-L, et al. The diagnostic accuracy of the Ascertain Dementia 8 questionnaire for detecting cognitive impairment in primary care in the community, clinics and hospitals: a systematic review and meta-analysis. *Fam Pract.* 2017;35(3):239-246. doi:10.1093/fampra/cmz098.
- Taylor-Rowan M, Nafisi S, Owen R, et al. Informant-based screening tools for dementia: an overview of systematic

- reviews. *Psychol Med.* 2021;53:1-10. doi:[10.1017/s0033291721002002](https://doi.org/10.1017/s0033291721002002).
21. Chingono J, Creavin ST, Fish M, et al. Informant accuracy of IQCODE, AD8 and GPCOGi for diagnosis of dementia: does your friend know best? *BMC Prim Care.* 2025;26(1):112. doi:[10.1186/s12875-025-02745-w](https://doi.org/10.1186/s12875-025-02745-w).
 22. Wedderburn C, Wear H, Brown J, et al. The utility of the Cambridge Behavioural Inventory in neurodegenerative disease. *J Neurol Neurosurg Psychiatr.* 2008;79(5):500-503. doi:[10.1136/jnnp.2007.122028](https://doi.org/10.1136/jnnp.2007.122028).
 23. Tsang RSM, Diamond K, Mowszowski L, Lewis SJG, Naismith SL. Using informant reports to detect cognitive decline in mild cognitive impairment. *Int Psychogeriatr.* 2012;24(6):967-973. doi:[10.1017/s1041610211002900](https://doi.org/10.1017/s1041610211002900).
 24. Jorm AF. The informant questionnaire on cognitive decline in the elderly (IQCODE): a review. *Int Psychogeriatr.* 2004;16(3):275-293. doi:[10.1017/s1041610204000390](https://doi.org/10.1017/s1041610204000390).
 25. Quinn TJ, Fearon P, Noel-Storr AH, Young C, McShane R, Stott DJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the detection of dementia within community dwelling populations. *Cochrane Database Syst Rev.* 2021;2021(7). doi:[10.1002/14651858.cd010079.pub3](https://doi.org/10.1002/14651858.cd010079.pub3)
 26. Brodaty H, Donkin M. Family carers of people with dementia. *Dialogues Clin Neurosci.* 2009;11(2):217-228. doi:[10.31887/dcms.2009.11.2/hbrodaty](https://doi.org/10.31887/dcms.2009.11.2/hbrodaty).
 27. Yim SJ, Yi D, Byun MS, et al. Screening ability of subjective memory complaints, informant-reports for cognitive decline and their combination in memory clinic setting. *Psychiatry Investigation.* 2017;14(5):640-646. doi:[10.4306/pi.2017.14.5.640](https://doi.org/10.4306/pi.2017.14.5.640).
 28. Neubauer K, Brant H, Thomas C, Biggs S, Hill E, Redwood S. Taking a collateral history: the missing piece of the puzzle. *BMJ.* 2023;382(382):e076462. doi:[10.1136/bmj-2023-076462](https://doi.org/10.1136/bmj-2023-076462).
 29. Kantithammakorn P, Punyabukkana P, Pratanwanich PN, Hemrungronj S, Chunharas C, Wanvarie D. Using automatic speech recognition to assess Thai speech language fluency in the Montreal cognitive assessment (MoCA). *Sensors.* 2022;22(4):1583. doi:[10.3390/s22041583](https://doi.org/10.3390/s22041583).
 30. Sikkes SAM, de Lange-de Klerk ESM, Pijnenburg YAL, et al. A new informant-based questionnaire for instrumental activities of daily living in dementia. *Alzheimer's Dementia.* 2012;8(6):536-543. doi:[10.1016/j.jalz.2011.08.006](https://doi.org/10.1016/j.jalz.2011.08.006).
 31. Sharafeldeen A, Keowen J, Shaffie A. Machine learning approaches for speech-based Alzheimer's detection: a comprehensive survey. *Computers.* 2025;14(2):36. doi:[10.3390/computers14020036](https://doi.org/10.3390/computers14020036).
 32. Pahar M, Tao F, Mirheidari B, et al. CognoSpeak: an automatic, remote assessment of early cognitive decline in real-world conversational speech. In: *2025 IEEE Symposium on Computational Intelligence in Health and Medicine (CIHM)*; 2025:1-7. doi:[10.1109/cihm64979.2025.10969487](https://doi.org/10.1109/cihm64979.2025.10969487).
 33. Ntampakis N, Diamantaras K, Chouvarda I, Tsolaki M, Argyriou V, Sarigiannidis P. NeuroXVocal: detection and explanation of Alzheimer's disease through non-invasive analysis of picture-prompted speech. *Lect Notes Comput Sci.* 2025;410-419. doi:[10.1007/978-3-032-05185-1_40](https://doi.org/10.1007/978-3-032-05185-1_40).
 34. Tang SL, Robertson A, Liang H, Taylor J-P, Harrison J. LUMEN: prototype conversational AI to streamline dementia assessments. *BJPsych Open.* 2025;11(S1):S16. doi:[10.1192/bjo.2025.10185](https://doi.org/10.1192/bjo.2025.10185).
 35. Treder MS, Lee S, Tsvetanov KA. Introduction to large Language Models (LLMs) for dementia care and research. *Frontiers in Dementia.* 2024;3:1385303. doi:[10.3389/frdem.2024.1385303](https://doi.org/10.3389/frdem.2024.1385303).
 36. Gaugler JE, Millenbah AN, Jensen CJ, Bailey D, Yam H, Baker ZG. Connecting Dementia carers to recommended services: care to plan. *J Gerontol Nurs.* 2023;49(10):5-11.
 37. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. *Gerontol.* 2018;58(3):467-477. doi:[10.1093/geront/gnw249](https://doi.org/10.1093/geront/gnw249).
 38. Gitlin LN, Hodgson NA. Who should assess the needs of and care for a dementia patient's carer? *The AMA Journal of Ethic.* 2016;18(12):1171-1181. doi:[10.1001/journalofethics.2016.18.12.ecas1-1612](https://doi.org/10.1001/journalofethics.2016.18.12.ecas1-1612).
 39. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci.* 2011;6(1):42. doi:[10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42).
 40. Gong N, Yang D, Zou J, et al. Exploring barriers to dementia screening and management services by general practitioners in China: a qualitative study using the COM-B model. *BMC Geriatr.* 2023;23(1):55. doi:[10.1186/s12877-023-03756-x](https://doi.org/10.1186/s12877-023-03756-x).
 41. Maidment DW, Heyes R, Gomez R, Coulson NS, Wharrad H, Ferguson MA. Evaluating a theoretically informed and cocreated Mobile health educational intervention for first-time hearing aid users: qualitative interview study. *JMIR mHealth uHealth.* 2020;8(8):e17193. doi:[10.2196/17193](https://doi.org/10.2196/17193).
 42. Guzman J, Aguiñaga S. The COM-B model: a cross-sectional survey assessing capability, opportunities, and motivation to follow the MIND diet among informal female carers of people with Alzheimer's disease and related dementias. *Frontiers in Dementia.* 2024;3:1451310. doi:[10.3389/frdem.2024.1451310](https://doi.org/10.3389/frdem.2024.1451310).
 43. Charters E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal.* 2003;12(2):68-82. doi:[10.26522/brocked.v12i2.38](https://doi.org/10.26522/brocked.v12i2.38).
 44. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods.* 2006;5(1):80-92. doi:[10.1177/160940690600500107](https://doi.org/10.1177/160940690600500107).

45. Kotting P, Smith A, O'Hare MB, et al. A national open-access research registry to improve recruitment to clinical studies. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*. 2021;7(1):e12221. doi:[10.1002/trc2.12221](https://doi.org/10.1002/trc2.12221).
46. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies. *Qual Health Res*. 2016; 26(13):1753-1760. doi:[10.1177/1049732315617444](https://doi.org/10.1177/1049732315617444).
47. Wear HJ, Wedderburn CJ, Mioshi E, et al. The Cambridge behavioural inventory revised. *Dementia & Neuropsychologia*. 2008;2(2):102-107. doi:[10.1590/s1980-57642009dn20200005](https://doi.org/10.1590/s1980-57642009dn20200005).
48. Mirheidari B, Blackburn D, Harkness K, et al. An avatar-based system for identifying individuals likely to develop dementia. In: *Proceedings of Interspeech 2017, Stockholm, Sweden, 20-24 August 2017*. ISCA; 2017:3147-3151. doi:[10.21437/Interspeech.2017-690](https://doi.org/10.21437/Interspeech.2017-690).
49. Birch PDJ, Whitehead AE. Investigating the comparative suitability of traditional and task-specific think aloud training. *Percept Mot Skills*. 2020;127(1):202-224. doi:[10.1177/0031512519882274](https://doi.org/10.1177/0031512519882274).
50. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101. doi:[10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa).
51. Moffat A, Cook EJ, Chater AM. Examining the influences on the use of behavioural science within UK local authority public health: qualitative thematic analysis and deductive mapping to the COM-B model and Theoretical Domains Framework. *Front Public Health*. 2022;10:1016076. doi:[10.3389/fpubh.2022.1016076](https://doi.org/10.3389/fpubh.2022.1016076).
52. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. *Int J Qual Methods*. 2017;16(1):1-13. doi:[10.1177/1609406917733847](https://doi.org/10.1177/1609406917733847).
53. Rabin LA, Guayara-Quinn CG, Nester CO, Ellis L, Paré N. Informant report of practical judgment ability in a clinical sample of older adults with subjective cognitive decline, mild cognitive impairment and dementia. *Aging Neuropsychol Cognit*. 2022;29: 139-157. doi:[10.1080/13825585.2020.1859081](https://doi.org/10.1080/13825585.2020.1859081).
54. Rueda AD, Lau KM, Saito N, et al. Self-rated and informant-rated everyday function in comparison to objective markers of Alzheimer's disease. *Alzheimer's Dementia*. 2015;11(9):1080-1089. doi:[10.1016/j.jalz.2014.09.002](https://doi.org/10.1016/j.jalz.2014.09.002).
55. Nelson A, Russell LL, Peakman G, et al. The CBI-R detects early behavioural impairment in genetic frontotemporal dementia. *Ann Clin Transl Neurol*. 2022;9(5):644-658. doi:[10.1002/acn3.51544](https://doi.org/10.1002/acn3.51544).
56. Yang C-H, Lin Y-T, Hsieh MH, Hwang T-J. A re-evaluation study and literature review on AD8 as a screening tool for dementia. *PLoS One*. 2025;20(5): e0321570. doi:[10.1371/journal.pone.0321570](https://doi.org/10.1371/journal.pone.0321570).
57. Sayegh P, Knight BG. Cross-cultural differences in dementia: the Sociocultural Health Belief Model. *Int Psychogeriatr*. 2013;25(4):517-530. doi:[10.1017/s104161021200213x](https://doi.org/10.1017/s104161021200213x).
58. Khobragade P, Nichols E, Meijer E, et al. Performance of the Informant Questionnaire on Cognitive Decline for the Elderly (IQCODE) in a nationally representative study in India: the LASI-DAD study. *Int Psychogeriatr*. 2022;36: 1-11. doi:[10.1017/S1041610222000606](https://doi.org/10.1017/S1041610222000606).
59. Xygykou A, Ang CS, Siriaraya P, et al. MindTalker: navigating the complexities of AI-Enhanced social engagement for people with early-stage dementia. 2024;96:1-15. doi:[10.1145/3613904.3642538](https://doi.org/10.1145/3613904.3642538).
60. Lima MR, Horrocks S, Daniels S, Lampthey M, Harrison M, Vaidyanathan R. The role of conversational AI in ageing and dementia care at home: a participatory Study. 2023; 571-578. doi:[10.1109/ro-man57019.2023.10309459](https://doi.org/10.1109/ro-man57019.2023.10309459).
61. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. *J Gerontol: Series B*. 2020;76(3):e99-e103. doi:[10.1093/geronb/gbaa098](https://doi.org/10.1093/geronb/gbaa098).
62. Spanakis P, Heron P, Walker L, et al. Use of the internet and digital devices among people with severe mental ill health during the COVID-19 pandemic restrictions. *Front Psychiatr*. 2021;12:732735. doi:[10.3389/fpsy.2021.732735](https://doi.org/10.3389/fpsy.2021.732735).
63. Luz S, Haider F, Fromm D, MacWhinney B. Detecting cognitive decline using speech only: the ADReSSo Challenge. *arXiv.org*. 2021. <https://arxiv.org/abs/2104.09356>. Accessed 2 Feb. 2026.
64. Tao F, Mirheidari B, Pahar M, et al. Early dementia detection using multiple spontaneous speech prompts: the PROCESS Challenge. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2025)*. IEEE; 2025:1-2. doi:[10.1109/ICASSP49660.2025.10889017](https://doi.org/10.1109/ICASSP49660.2025.10889017).
65. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *Npj Digital Medicine*. 2023;6(1):120. doi:[10.1038/s41746-023-00873-0](https://doi.org/10.1038/s41746-023-00873-0).
66. Saxena RR, Khan A. Assessing the practicality of designing a comprehensive intelligent conversation agent to assist in dementia care. In: *Paper Presented at the 18th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2025), HEALTHINF, SCITEPRESS*; 2025:655-663. doi:[10.5220/0013191600003911](https://doi.org/10.5220/0013191600003911).
67. Bélisle-Pipon J-C. Why we need to be careful with LLMs in medicine. *Front Med*. 2024;11:1495582. doi:[10.3389/fmed.2024.1495582](https://doi.org/10.3389/fmed.2024.1495582).
68. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. doi:[10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342).
69. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med*. 2022;28(5):924-933. doi:[10.1038/s41591-022-01772-9](https://doi.org/10.1038/s41591-022-01772-9).
70. NHS England. NHS Topol Review: progress on the recommendations 2019. 2019. <https://www.england.nhs.uk/topol-review-progress-on-the-recommendations-2019/>. Accessed: 2 February 2026.