

## ADVANCED REVIEW



WILEY

# Ethical issues when using digital biomarkers and artificial intelligence for the early detection of dementia

Elizabeth Ford<sup>1</sup> | Richard Milne<sup>2,3</sup> | Keegan Curlewis<sup>4</sup>

<sup>1</sup>Department of Primary Care and Public Health, Brighton and Sussex Medical School, Brighton, UK

<sup>2</sup>Kavli Centre for Ethics, Science and the Public, University of Cambridge, Cambridge, UK

<sup>3</sup>Engagement and Society, Wellcome Connecting Science, Cambridge, UK

<sup>4</sup>Royal Free London NHS Foundation Trust, London, UK

## Correspondence

Elizabeth Ford, Department of Primary Care and Public Health, Brighton and Sussex Medical School, Brighton, UK.  
Email: [e.m.ford@bsms.ac.uk](mailto:e.m.ford@bsms.ac.uk)

## Funding information

National Institute for Health Research Applied Research Collaboration Kent, Surrey and Sussex, Grant/Award Number: 200179; Wellcome Trust, Grant/Award Number: 108413/A/15/D

**Edited by:** Elisa Bertino, Associate Editor and Witold Pedrycz, Editor-in-Chief

## Abstract

Dementia poses a growing challenge for health services but remains stigmatized and under-recognized. Digital technologies to aid the earlier detection of dementia are approaching market. These include traditional cognitive screening tools presented on mobile devices, smartphone native applications, passive data collection from wearable, in-home and in-car sensors, as well as machine learning techniques applied to clinic and imaging data. It has been suggested that earlier detection and diagnosis may help patients plan for their future, achieve a better quality of life, and access clinical trials and possible future disease modifying treatments. In this review, we explore whether digital tools for the early detection of dementia can or should be deployed, by assessing them against the principles of ethical screening programs. We conclude that while the importance of dementia as a health problem is unquestionable, significant challenges remain. There is no available treatment which improves the prognosis of diagnosed disease. Progression from early-stage disease to dementia is neither given nor currently predictable. Available technologies are generally not both minimally invasive and highly accurate. Digital deployment risks exacerbating health inequalities due to biased training data and inequity in digital access. Finally, the acceptability of early dementia detection is not established, and resources would be needed to ensure follow-up and support for those flagged by any new system. We conclude that early dementia detection deployed at scale via digital technologies does not meet standards for a screening program and we offer recommendations for moving toward an ethical mode of implementation.

This article is categorized under:

Application Areas > Health Care

Commercial, Legal, and Ethical Issues > Ethical Considerations

Technologies > Artificial Intelligence

## KEYWORDS

artificial intelligence, dementia, digital biomarkers, early detection, ethics

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *WIREs Data Mining and Knowledge Discovery* published by Wiley Periodicals LLC.

## 1 | INTRODUCTION

Dementia has been identified as a key potential area of application for the development of digital biomarkers and automated tools. However, the development and implementation of these tools at scale also raises a range of ethical questions. In this review, we set out the context of dementia and the scope for early detection, and review some of these ethical questions through the lens of population screening.

### 1.1 | What is dementia?

Dementia is a term referring to a range of disorders characterized by a progressive decline in memory, reasoning, communication, and the ability to carry out everyday activities. These changes are usually associated with progressive structural and chemical changes which damage regions of the brain (Banerjee, 2010). The most common type of dementia is Alzheimer's disease (60%–70% of cases) followed by vascular dementia (20%) and dementia with Lewy Bodies (15%; Alzheimer's Research U. K. Dementia Statistics Hub, 2018). Currently there are no treatments which are curative for dementia, however, a timely diagnosis of dementia can still be valuable particularly for younger patients and their families, as they can be offered a range of supportive or therapeutic interventions, inform themselves about the condition, and plan for their future financial and care needs, which may help them maximize their quality of life (Prince et al., 2011).

The strongest risk factor for dementia is age, with prevalence in all over 65 s of between 6.4% and 7.1%, doubling with every 5-year increase in age to 28.5% at 90 years and older (van der Flier & Scheltens, 2005). Additional, well-evidenced non-modifiable risk factors include sex (dementia is more common in women), and genetics (particularly APOE4; The PHG Foundation, 2019). Potentially modifiable risk factors include early childhood deprivation, lower level of education, midlife hypertension and obesity, late life smoking, depression, sedentary behavior and diabetes, which are thought to all give rise to a higher risk of dementia (The PHG Foundation, 2019). These potentially modifiable risk factors account for about 40% of dementias worldwide (Livingston et al., 2020).

### 1.2 | How dementia is diagnosed at present

Diagnosis and medical care for people with dementia (PwD) typically starts with the recognition of dementia symptoms, often by PwD and their families, followed by a decision to seek help for the symptoms. The next step, in high-income countries with extensive primary healthcare systems, is for the patient to be seen and symptoms recognized by primary care physicians (Wells & Smith, 2016). Cognitive assessment tools such as the Informant Questionnaire on Cognitive Decline in the Elderly are often used to assist healthcare professionals with assessing for potential cognitive decline (National Institute for Health and Care Excellence [NICE], 2018). Patients are then usually referred to specialist services for a full memory assessment and diagnosis. This may take place in neurology, psychiatry, geriatric or multi-disciplinary assessment clinics, and involves batteries of cognitive tests and some form of brain imaging. There is no single operational definition of dementia, so diagnostic thresholds differ in different contexts (Aldus et al., 2020). Post-diagnosis care may revert to the primary care provider or continue in a specialist setting, and may include medication to help symptoms of cognitive dysfunction, social support, or care provision for activities of daily living.

### 1.3 | The dementia diagnosis gap across the world

National dementia strategies such as those implemented in the United Kingdom and elsewhere (Alzheimer Europe, 2021; Alzheimer's Disease International, 2021; Banerjee, 2010) have prioritized increasing rates of dementia diagnosis. This is driven by the assumption that treatment, care and support for dementia-related symptoms can only be offered after a formal diagnosis has been made. UK national dementia strategies, first published in 2009, have aimed to improve diagnostic rates, targeting 67% of PwD having a formal diagnosis recorded in their primary care electronic patient record. This was achieved by 2017, although the estimated rate of diagnosis has subsequently declined (NHS Digital, 2022; NHS England, 2017).

Dementia is thought to be under-diagnosed globally, with one systematic review finding the pooled rate of undetected dementia was 61.7% (Lang et al., 2017), and rates of under-detection higher in China and India compared to North America and Europe. In addition, when dementia is recognized, it is often at a late stage. For example, Aldus et al. (2020) found an average time of 3.5 years from reaching clinical criteria for diagnosis to getting a clinical diagnosis in a sample of PwD in the United Kingdom. They found that PwD living with a partner were twice as likely to be diagnosed than those living in other settings.

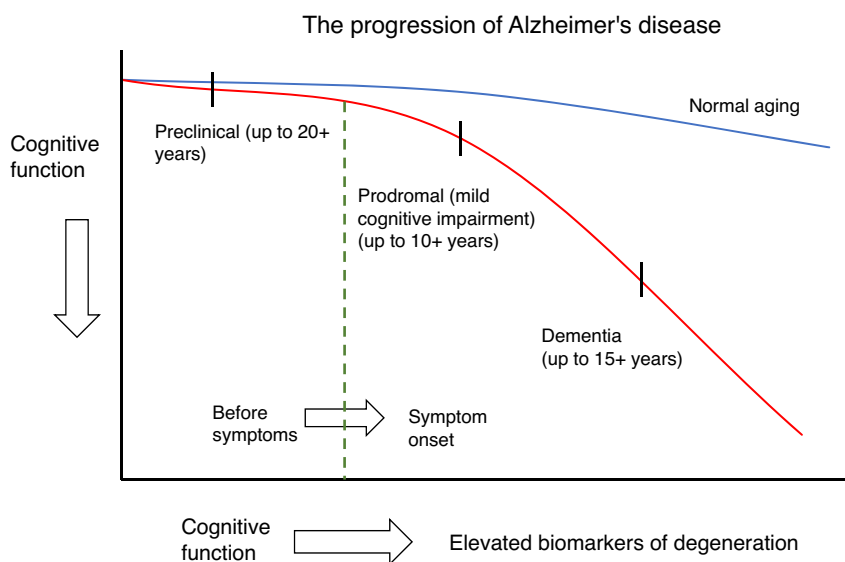
## 1.4 | Why would we want to detect dementia earlier?

Under-detection and late detection have been drivers of efforts to improve the ability to detect dementia more effectively and efficiently. In addition, this is supported by a clinical model of disease that sets out a hypothetical linear and unidirectional model of progression, as shown in the case of Alzheimer's disease in Figure 1.

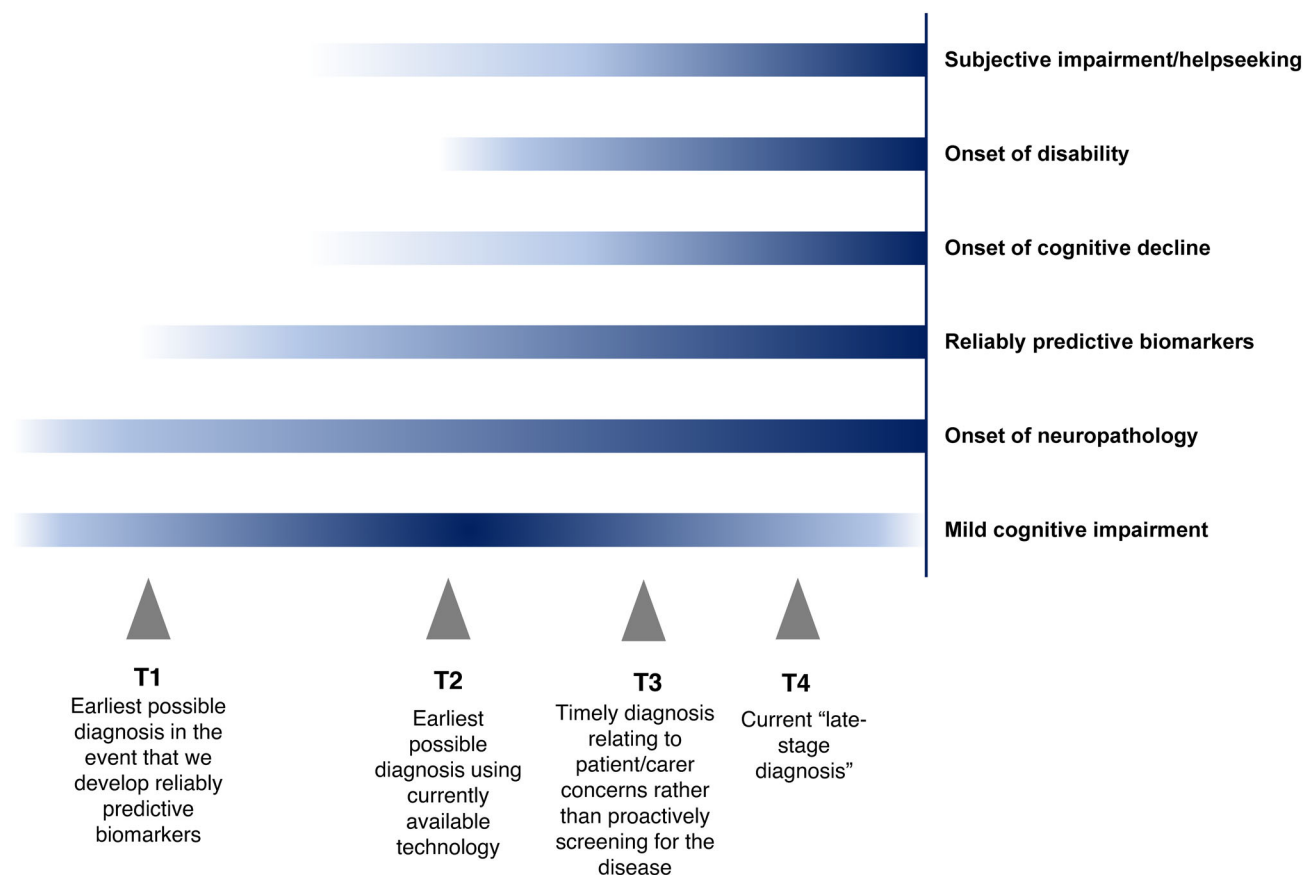
The continuum model suggests that for any intervention to be successful in slowing or preventing dementia, it must be started much earlier than the detection methods currently implemented in clinical practice allow. It posits that the critical period for some kind of effective clinical intervention, and thus the ideal time for diagnosis, may lie between the earliest point at which the diagnosis *could be made*, and the time at which diagnosis *is currently made* (Figure 2). We summarize three reasons for investing in methods to diagnose dementia at earlier time points than is currently possible.

### 1.4.1 | Reason 1: Disease modifying treatment—Could we detect changes early enough to prevent, slow down, or reverse the neuropathology?

The long duration and heterogeneity of neurodegenerative process before symptoms show may explain why many drug interventions for dementia have failed to live up to expectations in real-life contexts, as they have been tested on patients with a formal diagnosis of dementia, thus well-established disease (Gauthier et al., 2016; Ritchie et al., 2015). However, should disease-modifying or -reversing drugs become available for some dementias, they may have maximum efficacy when applied before the disease has clinically manifested and extensive and irremediable damage has occurred (Prince et al., 2011). Tools for detection aim to facilitate the identification of individuals who are in this potential zone for intervention and differentiate dementia from normal aging before the symptomatic phase.



**FIGURE 1** The continuum of Alzheimer's disease, where the preclinical phase may last for up to 20 years. Source: Adapted from Laver et al. (2016) and Gale et al. (2018).



**FIGURE 2** Diagram showing how opportunities for intervention intersect with possibilities for detecting dementia. *Source:* Adapted from Dubois et al. (2016).

Tools currently in routine use include structural brain scans (usually MRI, which is preferable to CT scans), which can identify neurodegeneration in the brains of PwD. However, particularly in the case of Alzheimer's disease, there is also increasing use of modalities such as positron emission tomography (PET) imaging, single photon emission computed tomography (SPECT) and cerebrospinal fluid (CSF) analysis. These aim to capture disease progression along the model trajectory of Alzheimer's disease shown in Figure 1, that involves the gradual accumulation of beta-amyloid rich neuritic plaques and neurofibrillary tangles up to 20 years before symptom onset (Gale et al., 2018). PET imaging, CSF markers and future blood-based biomarkers may also allow the identification of neurodegenerative changes in the brains of people who are currently asymptomatic or mildly symptomatic (Hampel et al., 2018). Given that many drugs in development target these biological markers, using them for early detection may be key to getting the best outcomes (Cavedo et al., 2014).

Importantly though, such tests are not definitively diagnostic, as neuritic plaques are also found in the brains of those who die without dementia, such that tests for amyloid or tau levels cannot give a definitive view of whether someone is living with the early stages of dementia or not, rather these tests corroborate a wider clinical picture (Neuropathology Group & Medical Research Council Cognitive Function and Aging Study, 2001). Moreover, despite the approval of Aduhelm by the US FDA, there is evidence that pharmacological effects on amyloid pathology have limited cognitive benefits, and indeed, there are suggestions that such effects are in fact absent (Ackley et al., 2021; Richard et al., 2021). Thus, at the present time it is not clear when, or even if, a successful, disease modifying agent for dementia would become available.

#### 1.4.2 | Reason 2: Identify target populations for trials of dementia treatments

While the first motivation for early detection relates to the identification of those most likely to directly benefit, the second is that without identifying individuals at "early" stages of disease, it is impossible to accurately assess the value of new therapies that may either benefit an individual now or similar patients in the future. As outlined above, it has been

hypothesized that new target compounds for modifying the disease trajectory of dementia will work best at the earlier stages of neurodegeneration (Gauthier et al., 2016; Jack et al., 2018). To evaluate the effectiveness of such compounds, it is necessary to identify patients reliably who have signs of neuropathology, which correspond with early-stage disease, to take part in trials. For example, amyloid biomarkers are increasingly widely used to identify “at-risk” participants for Alzheimer’s clinical trials (Alber et al., 2017), with PET imaging used extensively to assess the eligibility of potential participants in the large Phase III Anti-Amyloid in Asymptomatic Alzheimer’s (A4) trial of Eli Lilly’s solanezumab. “Trial-ready” platforms such as the European Prevention of Alzheimer’s Dementia (EPAD), the Trial Ready Cohort for Prevention of Alzheimer’s Disease (TRC-PAD) and the Japanese Trial-Ready Cohort (J-TRC) aim to facilitate the rapid inclusion of participants with known biomarker status (Aisen et al., 2022). However, these approaches are costly and inefficient when applied at a single trial level—in the case of the A4 study, 71% of people who underwent PET imaging did not meet the inclusion criteria (Meyers & Carrillo, 2020). There is thus a desire for alternative approaches to identify individuals most likely to meet biological inclusion criteria, including the use of blood-based and digital biomarkers for disease detection (Gold et al., 2018).

### 1.4.3 | Reason 3: Help people live well with dementia—Timely access to information, advice, support and appropriate care

Even where no disease-modifying pharmacological treatments exist, there may still be benefits to patients and families in early detection of disease. While the patients’ symptoms are still mild, or even before symptoms emerge, they have the opportunity to clarify the details of their condition with their clinicians and caregivers, understand its prognosis and trajectory, state their wishes for care, and organize their finances for their future needs (Prince et al., 2011). There are also some late-life modifiable risk factors for dementia, such as smoking, physical activity, social isolation and mental health (Prince et al., 2014; The PHG Foundation, 2019). It is possible that if changes associated with future risk of dementia are detected very early, modifying these risk factors could slow disease progression and/or help people with dementia to live longer with a higher quality of life, by, for example, minimizing comorbidities. Several self-management interventions have been trialed with people with early-stage dementia to help improve health behaviors and quality of life (Buettner & Fitzsimmons, 2009; Quinn et al., 2016). To date, however, the evidence for the effects of such programs remains limited. Among those without symptoms, studies of amyloid risk communication have found some effect on intended and actual changes in lifestyle and advance planning (Largent et al., 2020), although the evidence from the communication of genetic risk information suggests that such lifestyle changes may be unlikely without support (Hollands et al., 2016). Importantly, the only substantial clinical trial to date of dementia screening in primary care found no differences in healthcare utilization, advance care planning, or disease recognition by physicians after a year (Fowler et al., 2020).

## 2 | CURRENT TECHNOLOGY FOR EARLY DETECTION OF DEMENTIA

In reviewing technology for early detection, it is important to separate two approaches. The first is risk prediction, that is, the identification of groups of people who are at high risk of developing dementia in the future, but do not currently have detectable or diagnosable disease as defined by current diagnostic criteria. This might include, for example, identifying people who have various lifestyle-related risk factors or who carry the ApoE e4 risk allele that puts them at increased risk of disease. This should be distinguished from early detection, which is the focus of this chapter, in the sense of methods to identify individuals who have markers or signs associated with dementia-causing diseases, earlier than currently happens in the clinic.

### 2.1 | Blood-based versus digital biomarkers

Given the limits of the “fluid” markers described above, there is growing interest in alternative approaches, including blood-based markers, and digital tools, which form our focus.

“Digital” biomarkers are defined as objective, quantifiable physiological and behavioral data that are collected and measured by means of digital devices such as portables (e.g., tablets, smartphones), wearables (smart watches),



implantables, digestibles or passive sensors embedded in homes or cars (Babrak et al., 2019). Digital biomarkers denote the use of a biosensor to collect objective data on a biological (e.g., blood glucose, serum sodium), anatomical (e.g., mole size), or physiological (e.g., heart rate, blood pressure, gait, level of physical activity) parameter followed by the use of algorithms to transform these data into interpretable outcome measures (Dorsey et al., 2017).

## 2.2 | Types of digital biomarkers

There are a range of different out-of-clinic technologies which could generate data to inform algorithms tuned for early detection of cognitive impairment and dementia. These include tools adapted from existing cognitive assessment measures and novel assessments that make use of the capacities of smartphones and similar tools through gamification, virtual and augmented reality and image-based assessments. In addition, a range of more “passive” forms of data collection aim to make use of sensors in wearables, in-home or in-car devices, and data derived from use of non-dedicated technological devices (e.g., capturing patterns of mouse movements during computer use; Piau et al., 2019).

### 2.2.1 | Screening tools on mobile/ICT devices

The first approach to developing digital biomarkers for the early detection of dementia is to put established or validated cognitive screening tests onto apps and websites so that patients can complete these tests on a regular basis (Thabtah, Peebles, et al., 2020), and potentially outside the clinic. Bringing these cognitive tests out of the clinic and into patients' everyday life allows the influence of daily life activities on cognition to be captured, an effect which is not picked up when tests are only administered in the artificial environment of the clinic (Allard et al., 2014). Between 2010 and 2015, 31 computerized instruments were described in the literature, these were presented to patients on PCs, laptops and tablets and using both touchscreen and keyboard inputs; however, validity and reliability of these tests were rather poorly reported (García-Casal et al., 2017).

More recently, tests have been incorporated into mobile phone applications (apps). In their review of dementia screening apps, Thabtah et al. found 20 different apps which presented established cognitive screening tools such as the mini-mental state examination (Thabtah, Peebles, et al., 2020). Almost all had been validated against clinical populations and showed high accuracy (sensitivity range: 38%–100% and specificity range: 40%–100%). Few apps employed machine learning to process the data generated, although some employed artificial intelligence (AI) to analyze drawings (e.g., of a clockface, although no further details were given about AI techniques). However, Thabtah et al. argue that such applications need to focus on the entire cognitive functioning capacity to meet DSM-V diagnosis standards (Thabtah, Mampusti, et al., 2020). They suggest that such comprehensive assessment apps would save time in clinic and speed up formal diagnoses.

### 2.2.2 | Smartphone-native applications

In addition to porting existing tests into digital formats, a range of digital biomarkers also make use of the novel capabilities afforded by smartphones and tablets. These aim to address concerns that the results of cognitive assessment tests may be confounded by age, educational level, and language ability (Cordell et al., 2013). Digital biomarkers may therefore offer the opportunity to use language-agnostic assessments, as in the Integrated Cognitive Assessment (ICA; Modarres et al., 2021), currently undergoing validation in the UK National Health Service (NHS), or incorporate augmented reality assessments (Meier et al., 2021). Data produced by the ICA was able to detect cognitive impairment with accuracy (AUROC) of 0.81 using a logistic regression model (Modarres et al., 2021). Other researchers have proposed alternative approaches such as gamification (Tong & Chignell, 2014) and virtual reality (Banville et al., 2010; Nolin et al., 2013; Plancher et al., 2012) in which users play games which can assess domains of cognition without relying on language. Games can be used to assess a range of areas of cognition, with high accuracy and acceptability or usability, their discriminative accuracy has been assessed with support vector machine, linear regression and random forest techniques (Valladares-Rodriguez et al., 2017; Valladares-Rodriguez et al., 2019). Accuracy and acceptability factors must both be considered in a test presented as a game. Also of interest are smartphone apps which “passively” measure user activity, such as typing habits, step counts, or location tracking (Sabbagh et al., 2020). While a number of these

apps have moved from research into commercialization and are undergoing validation, none have been widely introduced into health services, and there is limited evidence of their cost-effectiveness (Shore et al., 2022).

## 2.3 | Passively collected data

As well as approaches that rely on the active engagement of the person being assessed, digital biomarkers present the opportunity to draw on the wider streams of data generated by people in the course of their everyday lives and use of technology.

### 2.3.1 | Wearable and mobile sensors

The widespread adoption of smartphones and devices such as the Apple Watch and Fitbit has suggested the possibility of collecting data about an individual's real world cognitive function and providing repeated, real-time measurements taken in real-life settings, as opposed to cognitive assessments and brain scans administered in the clinic only once or at spaced intervals (Dagum, 2018). Major programs, including collaborations between Apple, Evidation and Eli Lilly (Chen et al., 2019), Apple and Biogen (Intuition Study), Alzheimer's Research UK's Early Detection of Neurodegenerative Diseases (EDoN; EDON Initiative, 2021; Frey et al., 2019), RADAR-AD (Muurling et al., 2021) the MRC Deep and Frequent Phenotyping study (Koychev et al., 2019) and the SENDA trial (Müller et al., 2020) are exploring the potential of a range of digital technologies to collect data from community dwelling adults. Measures captured by digital devices may include sleep, neural activity, cognition, speech and language, gait, heart rate, fine motor skills and physical activity, with the aim of collecting a digital picture of individual's daily function, linking fluctuations in memory performance with specific daily life activities or experiences and identifying subtle changes suggestive of cognitive decline (Kourtis et al., 2019). To date, however, algorithms to process sensor or wearables data to make accurate classifications for cognitive impairment or dementia are still in their early stages, but some success has come from using, for example, Extreme Gradient Boosting algorithms, to enable handling of missing data and to produce explainable results, and achieving accuracy of nearly 90% (by AUROC; Chen et al., 2019).

### 2.3.2 | In-home sensors

A further aspiration of digital assessments since the late 1990s has been the assessment of functional performance in real-life environments to help identify and clarify difficulties in situ (Abowd et al., 2002). "Smart" environments have embedded sensors (e.g., infrared, magnetic or electric) or cameras, so that the task of capturing individuals' functional performance is not influenced by the act of observing them, and so that measurements can be taken regularly and "passively," providing longitudinal data to identify change. Developers of such measures also point out that they are unobtrusive and non-intrusive (Piau et al., 2019); once an individual's home is fitted with the sensors, they do not need to remember to put them on or take them out with them, as is the case with wearables or mobile devices.

To date, studies have found evidence of differences between people with cognitive decline and healthy controls on a range of such measures, including walking speed, level of activity (i.e., going out of the house), movement or awakenings during the night, frequency of daily ICT device usage, and regularity of medication intake (Lussier et al., 2019; Rawtaer et al., 2020). Again, the development of algorithms to differentiate people with intact cognition from those with cognitive decline, from in-home sensor-based data, are at early stages, but results are promising. Support vector machines and random forest algorithms have been trialed to discriminate between people with MCI and healthy controls, using at-home sensor data, achieving AUROC of 0.97 (Akl et al., 2015).

### 2.3.3 | In-car sensors

Around one third of drivers with dementia continue to drive in the early stages of the disease, while cognitive deficits are mild (Silverstein, 2008). Making the decision to stop driving can be difficult for individuals, signifying a non-reversible loss of independence and mobility. Advances in sensor, computing and communication technologies mean

that information about a person's driving performance can be collected and analyzed unobtrusively. There may be commonalities in driving performance changes in dementia patients, which could indicate the early onset of dementia. Eby et al. found several differences in driving behaviors in participants with early-stage dementia and an age-matched comparison sample, such as driving more slowly, driving shorter distances, and to fewer destinations (Eby et al., 2012). These differences could potentially be used, against an individual baseline, to detect the early stages of dementia (Bayat et al., 2021).

## 2.4 | The use of AI to accurately interpret data in the clinic

The previous sections have explored out-of-clinic technologies which could provide extra data about patient trajectories of cognitive and physical functioning to aid dementia detection. A further approach is to apply novel analysis methods to data which are already being collected in the clinic (e.g., Electronic Health Records [EHR] data) or in research cohorts such as the Alzheimer's Disease Neuroimaging Initiative (ADNI, 2017), with the aim of finding signals that can differentiate dementia patients from controls, and which may not be obvious to clinicians interpreting the data.

Clinical data can take various forms. The key data sources which have been explored for dementia detection are neuro-imaging data (MRI, PET, SPECT and diffusion-tensor imaging [DTI]; Teipel et al., 2015); EHR data covering medical diagnoses and history, comorbidities, sociodemographic and administrative variables; cognitive screening tests scores and neuropsychological test battery scores; genetic data; and transcribed speech. Many different forms of machine learning have been applied to these different data sources, and techniques can be largely grouped into supervised learning (classification into labeled groups), unsupervised learning (clustering) and deep learning (using multiple layers of nonlinear information processing to recognize feature qualities, generally classifying labeled data, but without revealing the detailed internal processes; Dashwood et al., 2021).

There have been several reviews of machine learning approaches to detect dementia in brain imaging (Ahmed et al., 2019; Pellegrini et al., 2018). Studies using the ADNI dataset, which includes fMRI brain imaging, dominate work developing prediction models for Alzheimer's disease (83% studies in one review on imaging; (Pellegrini et al., 2018) and 43.5% in another review drawing on wider sources of data (Goerdten et al., 2019)), the second most commonly used source is the Open Access Series of Imaging Studies (OASIS; Marcus et al., 2007). Historically, supervised learning using support vector machines (SVM) was the most commonly used approach to classify images into dementia vs controls (59% of studies in Pellegrini et al., 2018). Good performance was achieved with SVM, but predicting brain phenotypes and dementia subclassifications remained challenging (Ahmed et al., 2019). Clustering techniques such as principal component analysis offer some advantages such as low noise sensitivity and increased efficiency, and have shown promise in separating images in terms of brain dysfunction (Ahmed et al., 2019). However, the most accurate classification of brain images has been achieved through the application of "deep learning" to the task of discriminating between dementia patients and controls. Techniques have included feed-forward neural networks (Akhila et al., 2017), including autoencoders such as randomized denoising autoencoder markers (Ithapu et al., 2015), and stacked auto-encoders (Dolph et al., 2017; Ju et al., 2019); deep belief networks (Faturrahman et al., 2017); and convolutional neural networks (Murugan et al., 2021; Sarraf & Tofghi, 2016) including deep ensemble sparse regression networks (Islam & Zhang, 2017), and residual networks (Odusami et al., 2021). In general, deep learning models have achieved highly accurate discrimination between patients with normal cognition, mild cognitive impairment and dementia (ranging 80%–99% accuracy; Pellegrini et al., 2018). Studies have also developed models to predict who is at risk for future accumulation of tau proteins, the accumulation of which is associated with neurodegeneration and cognitive impairment (Giorgio et al., 2022). One of the drawbacks to the use of deep learning is the need for very large sets of labeled data. These models are also limited to detection of dementia in patients who are seen in hospital or specialist settings (and can afford the cost of a brain scan). Furthermore, the reliance on a limited number of datasets (mainly ADNI) presents substantial challenges for the generalizability of current prediction models.

Healthcare systems which generate electronic, longitudinal, routinely collected healthcare service use data on individuals (such as UK primary care) provide the basis for modeling trajectories of cognitive decline over time. Many future risk prediction models have been developed using electronic health records (EHR) data (Barnes et al., 2020; Ben Miled et al., 2020; Fu et al., 2022; McCoy et al., 2019; Nori et al., 2020; Reinke et al., 2022); fewer studies have used EHR data to detect existent dementia cases which were previously undiagnosed. To date, detection models have trialed simple classifiers such as logistic regression, random forest, and SVM (Ford et al., 2019; Jammeh et al., 2018; Shao et al., 2019), with accuracy (measured by AUROC) ranging from 0.74–0.91. Notably the lower accuracy presented by



Ford et al. (AUROC 0.74) was achieved after the team removed features which would indicate that the GP had already recognized a memory problem, and started the underpinning diagnostic process for dementia, for example, recording a memory problem or making a referral to a memory assessment service. When these memory-related features were subsequently included, and notably when free text clinical notes were also included, the model accuracy increased substantially (AUROC 0.84 with memory variables [Ford et al., 2020] and 0.94 with keywords from clinic notes [Ford, Sheppard, et al., 2021]). A review of using deep learning on EHR data (e.g., feed-forward neural networks [FFNN], convolutional neural networks [CNN], and recurrent neural networks [RNN]), found wide paper-to-paper variability, in techniques and model accuracy, but suggested RNN might be the best approach on EHR data to model the sequences of events in longitudinal records (Ayala Solares et al., 2020).

Discrimination between dementia patients and controls has been successfully modeled using scores from cognitive tests and neuropsychiatric tests administered in the clinic. So et al. (2017) reported a two stage screening process using MMSE data to discriminate between normal cognition and cognitive decline, and a second stage using and CERAD-K data (the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease) to classify cognitive decline patients into Alzheimer's or MCI diagnosis. They used a range of methods for example, SVM, Naïve Bayes, Random Forest, Bagging, Bayesian Networks and Multi-layer Perceptron (MLP; a type of feed-forward neural network). In the screening phase using MMSE data, MLP performed highly accurately (F1 measure 0.97), and in the second diagnostic phase, SVM had highest accuracy (F1 measure 0.74). In Chinese data, Zhu et al. (2020) used the Cognitive Abilities Screening Instrument (CASI) and Montreal Cognitive Assessment (MOCA), in random forest, Ada-boost, Logit Boost, neural network, naïve Bayes and SVM models, and found naïve Bayes offered the best discrimination with F1 measure of 0.87.

Some clinical tasks, such as describing pictures, generate patient speech data which can be transcribed. Features can be extracted using natural language processing (NLP) to generate structured data, which can then be fed into ML classifiers. In one study using the Pitt corpus (Becker et al., 1994), using logistic regression models, discrimination between dementia cases and controls was achieved with about 88% accuracy (Roshanzamir et al., 2021). Another study using the ADReSS Challenge dataset (Luz et al., 2021) compared conventional classifiers with a deep learning model incorporating BERT (Bidirectional Encoder Representations from Transformers; a pre-trained, open source deep learning framework for NLP [Devlin et al., 2018]), and found that the BERT deep learning models outperformed the conventional ML models, but by only a small margin (BERT F1 measure 0.81, with SVM performing at 0.79; Balagopalan et al., 2021).

Finally, some teams have trialed modeling multi-modal clinic data from different sources, such as MRI, genetic and neurological examination data using deep learning (e.g., James et al., 2021; Qiu et al., 2022; Venugopalan et al., 2021). Some approaches have used staged models, to separate out cognitively impaired individuals first and then derive a refined diagnosis of AD versus MCI on the selected patients in a second stage (Qiu et al., 2022; Zhou et al., 2019). Some studies have shown that integrating multi-modal data, for example genetic data, clinical test data, plus imaging, as well as using deep learning rather than “shallow” models, produces better performance than single modality data (Venugopalan et al., 2021).

### 3 | WHAT ARE THE POTENTIAL ETHICAL ISSUES?

Many of the early detection methods using digital biomarkers that are described above incorporate machine learning or AI methods. Their implementation in an ethical and responsible manner thus presents challenges related to the source of data and method of analysis, and the field of application.

Guidance for reflection on “AI ethics” in the clinical environment has been set out by a number of bodies, most notably the World Health Organisation and the European Commission (European Commission High-Level Expert Group on AI, 2019; World Health Organisation, 2021). The approaches commonly rest on broadly agreed principles that include core principles of medical ethics—benevolence, nonmaleficence, justice, respect for autonomy—and a specific emphasis on transparency, explainability, interpretability, intelligibility, responsibility and accountability (Solomonides et al., 2022). As these technologies are integrated in health care, models of AI governance thus emphasize the need for fairness, trustworthiness, transparency and accountability (Reddy et al., 2020). Such considerations are particularly important in light of the unprecedented amount of information about users that is captured by digital technologies, the potential invasion of privacy that results, and potential harms from data breaches (Perakslis & Ginsburg, 2021).

Guidance on the ethics and governance of AI tools has practical relevance for their successful implementation. For example, there is some evidence that clinicians are more likely to trust AI which is explainable, that is, it is clear to the clinician which input features were weighted in the AI model to justify the output (Reddy et al., 2020). However, new deep learning models (usually neural networks) are currently much less explainable than other forms of machine learning, yet to date, give the highest accuracy of dementia early detection models, creating a tension for clinical implementation. The integration of AI tools into the clinical workflow also raises questions about accountability and fairness (Reddy et al., 2020), such as in attributing liability for errors resulting from the use of AI devices in the clinical context or conversely, from their non-use. Robust guidance is needed over the circumstances in which a clinician would or should feel confident to dismiss the AI result, or how they can best explain a “blackbox AI” output to a patient.

However, in addition to wider questions around AI in healthcare, the implementation of automated methods for the early detection of dementia requires engagement with domain-specific concerns. These include both the *disease context* and the nature of the *clinical practice or service* into which tools will be introduced. In this case, these are the prediction of dementia and the implementation of screening programs.

Debates about the ethics of dementia risk prediction have been current since at least the identification of the ApoE e4 allele associated with increased susceptibility to Alzheimer's disease in the early 1990s (Corder et al., 1993; Post, 1996). These debates have primarily focused on the value of knowing or not knowing test results and the impact of this knowledge on an individual's view of their future. These discussions have evolved through deliberations around the use of tools such as PET amyloid imaging or cerebrospinal fluid biomarkers and remain vital as digital biomarkers emerge (Harms et al., 2022; Karlawish, 2011; Molinuevo et al., 2016; Piau et al., 2019).

Digital biomarkers, however, present slightly different challenges, not least because of the possibility they offer of detecting signs of neurodegenerative disease before the onset of the symptoms at scale, using widely available technology and in a way that may not be directly controlled or initiated by a clinician. Instead, either a patient-user or the developer of the biomarker technology may have control over its use. These factors point to the need for a robust framework for contemplating both the practical and ethical questions associated with digital tools for early detection.

One robust approach to considering these interlinked questions of ethics and practice is provided by the lens of population screening. In their seminal 1968 report for the World Health Organisation (Wilson & Jungner, 1968), Wilson and Jungner described the object of screening for disease as *to discover those among the apparently well who are in fact suffering from disease*, a conceptualization which fits well the paradigm of early detection. They set out 10 principles in the form of normative statements, which have become paradigmatic in evaluating whether a screening program is desirable, and whether it is likely to achieve the ethical goal of producing more benefit than harm:

1. The condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a “once and for all” project.

While these principles often form the basis for governmental or quasi-governmental programs for the assessment and governance of screening programs (Sturdy et al., 2020), they have wider relevance to any systematic attempt to identify those with hitherto unidentified disease. As such, they are highly relevant to early detection programs—whether established by health systems, health insurers or in other settings, for example by technology developers. To date, this wider relevance has rarely been recognized in discussions of digital health, including in the case of dementia (Gunnarson et al., 2021). Below, we consider how these principles apply for the paradigm of dementia early detection using digital biomarkers and automated detection methods, and emphasize the ethical issues brought to light by each principle.

### 3.1 | The condition sought should be an important health problem

It is clear from the increasing numbers of people with dementia globally, the aging population across the world, and the cost or burden of high-quality care for people living with dementia, that dementia constitutes an important health problem. Consequently, in the case of symptomatic disease, this condition is largely met. However, as discussed below, this is more complex in the case of detection prior to the emergence of symptoms or in mildly symptomatic disease, which, if widely accepted as health problems, may cover a substantial population worldwide (Brookmeyer et al., 2018; Gustavsson et al., 2017). In order to establish this, and as elaborated below, it is critical that criteria (4) and (7) related to the understanding and identification of the disease trajectory, from early or latent stage through to full dementia, are met.

### 3.2 | There should be an accepted treatment for patients with recognized disease

Screening programs can only be justified if there is an accepted treatment for patients with recognized disease. Unless an acceptable treatment can *currently* be provided which improves the prognosis of the disease, and which gives better outcomes due to the condition being found at an earlier stage, there can be no advantage to the patient of intensive case-detection in an otherwise well population (Wilson & Jungner, 1968).

As discussed above, no treatments that effectively modify disease trajectory for dementia are available (Ackley et al., 2021; Richard et al., 2021). Consequently, large scale detection programs do not yet meet this criterion. As a result, while there may be cases in which individuals opt in to something approximating early detection, notably when choosing to take part in a clinical trial (Kim et al., 2015), it would be unethical to proceed with systematic or population-based early detection of dementia on the basis only that a disease-modifying treatment may become available in the future.

This is also an important consideration for risk prediction: before widespread programs involving individual risk prediction models are rolled out beyond research settings it is essential that evidenced risk reduction strategies at an appropriate life stage are available (Weatherby & Agius, 2018). As the World Health Organisation guidelines on risk reduction recognize, evidence in most areas of risk reduction for dementia is currently moderate or low, particularly for older and mildly symptomatic populations (World Health Organisation, 2019).

### 3.3 | Facilities for diagnosis and treatment should be available to all

Screening programs identify a range of people who are likely to have the disease, for whom additional follow up tests are needed to establish a diagnosis. Resources must be allocated in the health service at the time of the implementation of an early detection program to manage the flow of patients referred for diagnostic assessment. However, mental health services in many health systems are drastically underfunded (Bannister, 2021) and cannot cope with current demand (World Health Organisation, 2017). While the hope is that early detection will reduce healthcare costs in the long term (see 9 below), current systems are unable to absorb a wave of patients channeled through to memory assessment services following screening. Consequently, for early detection programs to be effective and ethical there should first be a sufficient commitment of resources to dementia diagnostic and support services (Brayne & Kelly, 2019). This is an important factor in the development of digital biomarkers which may be inexpensive to administer but impose costs throughout the health system.

Additional consideration should go toward the psychological support needed when dealing with information about personal risk of dementia. At the current time, dementia is a progressive disease with no disease-modifying treatment options and with implications for future health and mental capacity. To date, evidence related to the impact of receiving predictive information related to dementia while asymptomatic (e.g., based on genetic or biomarker tests) has been accompanied by counseling and professional support before and after testing (Goldman et al., 2011; Harkins et al., 2015). While such support may not necessarily be required long-term, there is limited evidence of the benefit versus harm consequences of unsupported return of results (Frederiksen et al., 2021). Further, upon receiving results, people may also require and (reasonably) expect some level of support, monitoring and follow-up (Milne et al., 2018); a service that again should be factored into cost predictions for the systemic impact of digital biomarkers (Schmutte et al., 2022).

### 3.4 | (4) There should be a recognizable latent or early symptomatic stage and (7) the natural history of the condition, including development from latent to declared disease, should be adequately understood

Contemporary models of Alzheimer's disease posit a linear model of disease characterized by early and prodromal phases of disease that can be identified by changes in biomarkers or subtle alterations in cognition, mild cognitive impairment, and/or psychological and behavioral symptoms. While this linear model holds at a population level, the future trajectory of any individual patient's cognition from mild impairment to dementia cannot yet be accurately predicted (Hu et al., 2017). Patients in the early stages of disease may be labeled as having mild cognitive impairment (MCI), subjective memory complaints (SMC) or subjective cognitive decline (SCD). Each of these confers an elevated risk of developing dementia in the future (Slot et al., 2019). However, many patients in these categories will stay stable or show improvements, with only around 10% of those labeled as having MCI being diagnosed with dementia in each following year (Bruscoli & Lovestone, 2004), and up to 18% reverting to normal cognition (Canevelli et al., 2016). Machine learning and AI analysis techniques could likely contribute to understanding which patients with MCI are at highest risk of conversion to dementia; however, tests to gather the data needed for current predictive methods remain invasive and expensive (Eckerström et al., 2013; Gelosa & Brooks, 2012; Gomar, 2011; Zhang & Shen, 2012).

In sum, while there is a prodromal or latent phase in dementia, this needs to be better understood before the introduction of screening tests. It is not easy to determine the future course of disease for an individual with mild cognitive impairment. What makes one person move from MCI to dementia and another revert to normal cognition is not well understood.

### 3.5 | (5) There should be a suitable test or examination

The principle of there being a “suitable” test is arguably the hardest to meet in the paradigm of early detection of dementia. A “suitable test” should be (a) highly accurate, (b) minimally invasive and (c) low cost.

Digital devices are clearly “suitable tests” in that they are minimally invasive and potentially low cost. However, although Wilson and Jungner suggest that “simplicity and ready application” (Wilson & Jungner, 1968, p. 33) may be enough to make a test suitable, in the case of dementia the potential consequences of screening for individuals and populations emphasize persistent and significant challenges related to accuracy.

As outlined above, where prevalence is low, the accuracy of a screening test, usually given as sensitivity and specificity, must be extremely high not to produce more false positives than true positives. For example, a dementia detection method rolled out on all over 65-year olds in the United Kingdom, where dementia prevalence is 7.1% (NHS England, 2017), with 95% sensitivity and 90% specificity (representative of some of the best current digital biomarker methods), would have a positive predictive value (i.e., the proportion of positive results which are true positives) of 42%. Thus 58% of individuals given a positive result would subsequently be found not to have dementia.

As noted, while digital biomarkers offer some promise, their accuracy is not yet well established. Few digital health technologies have been rigorously shown to be equivalent to standard, accepted measurement devices (Perakslis & Ginsburg, 2021). The more accurate AI approaches rely on invasive brain imaging data and/or blood/CS fluid biomarkers and very few of the passive digital biomarker methods can yet offer sensitivities and specificities over 90%. As such, these tests do not *currently* meet the requirements of being highly accurate. In fact, the widespread and indiscriminate implementation of these methods to automatically detect dementia from multiple passively generated sources of data without knowledge of prior health status may lead to an inflation of the false positive risk (Ienca et al., 2018). In implementing digital screening tools therefore, those responsible should consider the potential rate of false positive results and the consequent risk of harm to these individuals, and plan how to mitigate it.

The possibility that implementation of AI will widen health inequalities must also be addressed and monitored (Walsh et al., 2020). There is a risk that algorithms for early detection will work less accurately when making predictions on some groups than others, and that this may not be evident to those using them. Such algorithmic bias is inherited from the social patterns reflected in training data, including how data is created or collected, and is increasingly well evidenced in medical applications beyond dementia (Obermeyer et al., 2019). Such bias refers to any application of algorithms that compounds existing inequities in socioeconomic status, race, ethnic background, religion, gender, disability, or sexual orientation.



Many of these existing inequalities are evident in the case of dementia and dementia research. Hence, for example, in the United States, African-Americans are at greater risk of dementia than whites and Asian-Americans (Mayeda et al., 2016), as are those living in areas of the United Kingdom with higher levels of deprivation (Matthews et al., 2016), and immigrants to the Netherlands with non-Western ethnic backgrounds (Parlevliet et al., 2016). All have substantially higher risks of developing dementia, but may be under-captured in both clinical and research data, due to systemic inequities in healthcare. People from minority ethnic groups in western countries are also under-represented in widely used datasets such as ADNI, in part because of differences in how people engage with clinical services and research. Research datasets, sampled from different populations, are also heterogeneous and present challenges for validation and replication (Birkenbihl, Salimi, Fröhlich, Japanese Alzheimer's Disease Neuroimaging Initiative, & Alzheimer's Disease Neuroimaging Initiative, 2022). Unsurprisingly, many dementia risk models—even developed without automated approaches—already do not translate across cultural settings (Stephan et al., 2020). These limitations and inequalities may be compounded in the case of automated systems.

### 3.6 | (6) The test should be acceptable to the population

As suggested above, the acceptability of a dementia screening test to the target population is by no means a given. People have mixed views on whether they would want to know if they currently have dementia (Corner & Bond, 2004; Marzanski, 2000) or are at increased risk of developing it in the future, with some research suggesting that only patients with severe dementia uniformly want to know their diagnosis (Jha et al., 2001).

Overall, it has been estimated that between half and 90% of people may be interested in knowing the results of a “predictive” or “reliable” test of whether they are likely to develop dementia, with numbers varying across countries (Caselli et al., 2014; Gooblar et al., 2015; Ott et al., 2016; Wikler et al., 2013). In the United Kingdom, two waves of Alzheimer Research UK's Dementia Attitudes Monitor have suggested that about three-quarters of people would like to be told their personal risk of getting dementia in the future—and a similar proportion would be willing to use smartphone apps or wearable devices that might help them understand their dementia risk (Alzheimer Research UK Dementia Statistics Hub, 2021). However, given the current limitations of knowledge described above, information provided at this time would be highly uncertain—a factor that decreases interest in learning personal risk (Gooblar et al., 2015; Milne et al., 2018).

Dementia remains highly stigmatized—half of UK adults say that it is the condition they fear getting most (Alzheimer Research UK Dementia Statistics Hub, 2021). In the case of older adults without overt symptoms, some express fear and worry about dementia, feeling that “there was little that could be done” (Corner & Bond, 2004), and as a result were unlikely to seek health care advice or a diagnosis for memory problems. For some, this may result in fatalism—the view that dementia is an inevitable part of aging, or that nothing can be done to reduce risk, views which evidence suggests may be particularly common among lower socioeconomic groups and racial and ethnic minority groups (Alzheimer Research UK Dementia Statistics Hub, 2021). In this context, a screening program would pick up more people who would not otherwise seek a diagnosis—and for whom the perceived value of a diagnosis may be less (Brayne & Kelly, 2019).

There are few concrete examples of screening programs in practice. In a large screening trial conducted in Indiana, 38% of those eligible to participate refused to do so, while two-thirds of people who did participate and had positive results on screening for cognitive impairment refused subsequent diagnostic evaluation (Fowler et al., 2020). Patients refusing further testing were more likely to live alone and have a perceived stigma of dementia, while those who refused to participate raised concerns about potential emotional distress and fear of losing independence. This echoes concerns about losing health insurance cover, driving privileges, or employment and worry about the effect on family finances and emotions described in earlier work (Couteur et al., 2013). Others have raised concern about privacy and disclosure issues of personal data about dementia (Ienca et al., 2018). In practice, therefore, population screening for dementia may not be acceptable to the general public (Martin et al., 2015).

Moreover, the test or program of early detection should also be acceptable to clinicians who might have responsibility for administering or interpreting the test and discussing the result (see also below (8)). For clinicians, early detection tools that allow them to effectively identify those with memory complaints or mild cognitive impairment who are most likely to develop dementia may have value (Saunders et al., 2022). Clinicians may also perceive benefits associated with possibilities for life planning of healthcare, finances and care preferences (Schweda et al., 2018). However, they are also aware of risks associated with stigmatization, and discrimination in insurance, employment or by family members, as



well as concerns about psychological distress in the absence of therapy (Schweda et al., 2018). In the United Kingdom, General Practitioners (GPs, also known as Primary Care or Family Physicians) are responsible for recognizing memory problems, referring the patient for further memory assessment and supporting the patient's holistic health care needs following diagnosis. It is most likely that any early detection method using digital biomarkers would be rolled out in primary care, in the community, or by private companies. However, GPs report resistance to making a dementia diagnosis early in the disease course (Iliffe et al., 2003) and adopt new predictive technologies in the clinic only when they meet certain criteria (Ford, Edelman, et al., 2021). GPs are more likely to use artificial intelligence as clinical decision support (CDS) when they are offered training in how to use it, the CDS is developed from a transparent evidence base, GPs have control over when to deploy it, and there is a clear communication guide and clinical pathway with which to follow up the result (Ford, Edelman, et al., 2021). GPs are also concerned about how communicating a bleak diagnosis may affect their relationship with their patient (Dhedhi et al., 2014). However, where GPs could have a well-resourced diagnostic service and a dementia support package to offer screen-positive patients, they would be, in the majority, favorable about systematic early dementia detection (Thyrian et al., 2016; Thyrian et al., 2018). A large majority of clinicians (86%) would also want clinical guidelines to accompany the testing program, with explicit information about risks and benefits of testing (Schweda et al., 2018). As a result, and as similarly concluded by Martin et al. (2015), clinician attitudes are likely to match those of the general public. Indeed, 69% of clinician-respondents felt a screening program would only be useful or acceptable if there were effective treatment options (Schweda et al., 2018).

### 3.7 | (8) There should be an agreed policy on whom to treat as patients

Who is considered as a patient in the context of dementia care has undergone a shift over several decades, reflected in evolving diagnostic criteria. Thus, while those with symptomatic dementia continue to be seen as the patients of memory services, since the early 2000s they have increasingly been joined by those understood to have mild cognitive impairment. More recently, the adoption of criteria delineating “preclinical” disease stages has led to the expanded global estimates discussed under (1) above.

However, both MCI and the most recent diagnostic classifications have proved to be contentious in both scientific and clinical practice, as well as being unclear for patients themselves (Gomersall et al., 2015; Saunders et al., 2022; Swallow, 2020). There is consequently uncertainty about whether and how the results of early detection should be communicated to patients, who should do this, and how patients should be supported to make decisions about what to do as a result. Thus, the clinical pathway associated with screening is not at all clear at the current time.

### 3.8 | (9) The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole

The balance of costs to the whole health and care system is critical question for the early detection of diseases leading to dementia, and for the development of technologies that aim toward this goal. As Wilson and Jungner pointed out in their elaboration of screening criteria, “*It is often considered that the detection of disease by screening will be economical of a country's resources*” (p. 35). Arguments in favor of early detection thus often pivot around the substantial economic cost of dementia care, particularly in later stages of disease. In the United Kingdom, the Alzheimer's Society, using figures from the MODEM study, estimates the cost of dementia care at £34.7 billion and projects an increase to £94.1 billion in 2040 (Wittenberg et al., 2019). In 2015, the global cost of dementia was estimated to be US\$818 billion (Wimo et al., 2017).

It is important to note that the bulk of these costs are associated with direct or informal care. While some of these may be amenable to technological intervention (Knapp et al., 2016), they may not be *directly* affected by early detection tools. However, the ability to delay and slow the progression of dementia and control comorbid conditions may be facilitated by early detection, and this may have potentially significant economic impacts (Mattap et al., 2022). The US Alzheimer's Association, for example, has estimated that an intervention that delays the onset of dementia by 5 years may save \$220 billion over that time (Alzheimer's Association, 2015). A more recent estimate, based on a projection of a 30% relative reduction of baseline progression rates from MCI to mild dementia and from mild to moderate dementia,

suggests a \$2.62 trillion benefit over 20 years in the United States, much made up of the gain in quality-adjusted life-years (Prados et al., 2022).

However, in the case of specific tools for screening and case-finding, the health economic arguments are currently limited by other factors discussed above, crucially the availability (and timing) of treatments, evidence about longitudinal disease trajectories and patient acceptability (Barnett et al., 2014; Gustavsson et al., 2017). As such, it is not clear how strong the economic arguments are in favor of early detection.

In addition to the wider economic impact of early detection, digital tools for early detection present additional questions related to consumer expenditure on early detection and its financial consequences for both those individuals and health systems. Tools for risk assessment and early detection are already being incorporated into health insurance programs and offered as apps to older adults—for example, in the EASIIT collaboration between pharmaceutical company Eisai and digital testing firm CogState (Cogstate, 2020).

The adoption of apps by consumers shows an emerging direct-to-consumer screening market. It highlights both the potential for lifestyle prevention but also the exploitation of older adults' concerns about dementia—in the case of dementia, concerns that were evidenced in the substantial penalties handed out by the US Federal Trade Commission to the company Lumosity for preying “on consumers' fears about age-related cognitive decline, suggesting their games could stave off memory loss, dementia, and even Alzheimer's disease” (US Federal Trade Commission, 2016).

The proliferation of digital screening tests also has potential implications for healthcare systems. In the United Kingdom, these concerns have resulted in a joint position statement from the UK Royal College of General Practitioners and the British Medical Association on screening by organizations which have not been approved by the UK National Screening Committee (Royal College of General Practitioners, 2019). This statement highlights the growing private sector offer of screening tests unsupported by high quality trials, with 91% of doctors reporting having seen patients in NHS time to discuss private health screening results; with only 13% feeling that this was an appropriate use of NHS resources. It concluded that such discussions “can be an inappropriate use of NHS resources and can have a potentially significant negative impact on primary care.” As such, it recommended that GPs consider informing patients why such tests are not offered through the NHS, that the government and NHS ensure that health care providers are not put under obligation to follow up the results of these tests, and that regulators remind private providers of their responsibility not to put the NHS at risk of subsidizing private care and ensure private clinics arrange follow-up at their own cost.

### 3.9 | (10) Case-finding should be a continuing process and not a “once and for all” project

The final consideration for those developing digital and automated tools for early detection relates to the longevity of the projects they facilitate. For screening to be of maximum value, it must continue to pick up the people who will benefit most from it through regular offers of examination. This is one argument in favor of the development of ongoing surveillance and health checks based in primary care or community settings.

However, for digital devices this criterion also raises questions associated with access, sustainability and obsolescence. For a digital tool to be able to deliver large-scale, repeated assessments, it must be accessible to the widest number of people using available technology, supported by future technology and/or compatible with future tools. This is often little acknowledged, but was evident in the roll-out of digital contact tracing devices for Covid-19; the largest implementation of digital health solutions to date, but one which highlighted the existence of a digital divide both in terms of device ownership and internet access and in terms of the currency of these devices (Ada Lovelace Institute & The Health Foundation, 2021). This divide risks limiting the accessibility of technology for those who are most in need. There is commonly a tension for developers of digital health technology, who in some cases must choose between taking time to run trials to gain evidence for effectiveness or non-inferiority but risking obsolescence, or getting the product to market while it is still current and compatible with the devices currently in use but without much of an evidence-base.

## 4 | CONCLUSION

In this paper we have discussed the potential reasons for aiming to detect dementia earlier than is currently achieved, the emerging digital technologies which could achieve this, and the issues which arise when considering applying the

early detection paradigm to a wide population. We have applied the lens of ethical screening programs to dementia early detection using digital biomarkers, and have found that efforts toward early detection of dementia fall short on a number of points.

The first issue is whether there is or can be any tangible benefit to patients or the population in identifying dementia in a pre- or early-symptomatic population. This is currently doubtful, given the lack of treatments which change the course or outcome of the disease. No potential reasons for earlier detection are currently supported by a strong evidence base. Screening principles require there to be an available treatment which would change the course of the illness if the illness is detected at an early stage compared to a late stage. There is currently no good evidence that any dementia intervention changes the disease trajectory or prognosis. Many other screening programs show little to no benefit over a population, due to the harms associated with false positive results, and sometimes little benefit to outcomes of detecting asymptomatic disease (UK Government, 2021; World Health Organisation, 2020). Given the low evidence base for benefits for individuals of earlier detection of dementia, there is unlikely to be benefit across a wider population.

There is a prodromal phase for dementia during which patients experience some cognitive deficits but do not meet criteria for dementia diagnosis. However, it is neither clear nor certain that patients with mild cognitive impairment will progress to dementia, and some patients will regain normal cognitive function. While invasive tests of fluid biomarkers and brain scans can identify neuropathology prior to onset of symptoms there is no direct correspondence between detected neuropathology and level of cognitive impairment. Until we can accurately predict which patients will go on to develop symptomatic dementia from mild cognitive impairment or from detected neuropathology, it is difficult to say that there is a clear latent phase of dementia in which to focus disease-modifying treatment efforts.

A potential reason to detect disease earlier is so that people can be aware of their condition and receive a prognosis at a time when they are able to plan and express their wishes for the future. This may be beneficial for family caregivers and patients, but this benefit would likely only accrue for people who were keen to know their dementia status. By extending early detection to a wider group of people, there is the risk of including people who do not wish to know their disease status or prognosis (Brayne & Kelly, 2019). If these technologies are state or health insurance funded and rolled out to a population, many individuals may be unwilling to engage or may not want to know their dementia status. If technologies are direct-to-consumer products offered by private companies, there is a potential risk of entrenching inequalities by means of cementing the digital divide between people who embrace both new technologies and health knowledge and awareness, and those who do not.

We reviewed the current and emerging technologies utilizing digital biomarkers or applying machine learning techniques to clinical data to detect patients with dementia at earlier time points. Screening criteria require that there should be a suitable and acceptable test for the condition being screened.

We characterized “suitable” as being low cost, non-invasive and highly accurate; three conditions which are currently hard to meet in the current digital biomarker domain. While applying deep learning to extensive batteries of clinic data has started to produce highly accurate discrimination between dementia and cognitively intact patients, the kind of noisy data generated from less-invasive and less costly passive sensors or wearables does not allow such accurate classification. The risk of generating many false positives with associated real risk of harm, must be seriously considered in this paradigm, given the current low accuracy of non-invasive means of data collection. A key challenge relates to the potential for algorithmic bias. Reducing this will require datasets that are more representative of the target clinical population, and attentive to the social, cultural and economic factors that shape clinical care and thereby generate the data on which AI is trained or validated. This requires commitment of suitable resources, and a recognition that socially and ethnically diverse research teams may be necessary to effectively identify flaws and potential bias in the design and functionality of automated systems.

It is by no means a given that digitally enabled screening of the population for dementia would be an acceptable activity; not to patients, and neither to clinicians. One population who may benefit from improved characterization of prognosis are those patients with an MCI diagnosis. Knowing their future dementia status might be of real benefit to this cohort, and the chance of accurate prediction, while minimizing false positive risk, is more likely in a focused patient group at high risk.

Even without an approved disease-modifying treatment, PwD have a number of care needs greater than equivalent age patients without dementia. Psychological screening and support would need to be available to help patients navigate what is currently a bleak diagnosis and prognosis, and responsive social care services would be needed to meet social, domestic and occupational care needs. However, there is a workforce crisis in social care in countries like the United Kingdom (Dromey & Hochlaf, 2018), which the Covid-19 pandemic has severely exacerbated (Pearson

et al., 2022). There is a real risk of sleepwalking into the roll out of “exciting” new technological screening or detection opportunities without proper planning and resourcing of mental health and social care services to receive the increase in workload that these technological innovations generate.

## 4.1 | Strengths and limitations

Our approach to this review was strengthened by use of screening framework, this helped us to identify important ethical domains to consider when a new screening or early detection paradigm is emerging. However, a limitation is that our review has not followed systematic reviewing practice, we have covered too many domains for a single search string to pick up the range of literature which was needed to cover the issues in this field.

## 4.2 | Next steps and final conclusions

We note a number of future activities in the field of dementia early detection which would be valuable in progressing this field in an ethical way. First, involving diverse patients, carers and clinicians in the development of these technologies would ensure that progress in this domain is acceptable and meets a patient or clinician need. It would also make it more likely that technologies will fit with clinician workflow and/or produce data which is valuable for, and interpretable by, clinicians (Ford, Edelman, et al., 2021). Second, it would benefit implementation of technologies if it is clear from the outset who these technologies are targeting. When considering the potential audience, issues of bias and inequality can then be evaluated. Developers should have clear strategies to prevent widening health inequalities via digital inequity.

### BOX 1 Recommendations for ethical practice in early detection of dementia

*The early detection of dementia using widely available digital devices and automated systems should be considered as a potential form of population screening and those developing such tools should consider the relevance of ethical screening criteria.*

Early detection of dementia may identify a large number of people with mild cognitive impairment for whom it is not currently possible to determine if they will progress to dementia. Early detection models should ideally focus on accurate prediction of who will progress to dementia from MCI.

It is not ethical to implement population level early detection of dementia until:

1. There is a treatment available which changes the course of the disease and which is more effective when started earlier in the disease course;
2. We can accurately determine how likely someone is to progress to dementia in the prodromal or MCI phase;
3. There is a highly accurate detection method which is minimally invasive and cost-effective;
4. An appropriately resourced health service pathway supports early detection, including (i) clinician training in using the detection method and communicating results, and (ii) diagnostic, psychological support and treatment services resourced to meet the additional need generated by the early detection program.

In addition, the following principles should be followed (cf. Arribas-Ayllon, 2011):

5. The decision to undertake early detection of dementia should be voluntary on the part of the patient, free of coercion and based on informed consent; and the result should be kept confidential.
6. The potential for psycho-social harm to the patient which could arise as a result of early detection should be assessed in each case.
7. Any deployment of early detection methods should be monitored for fairness, so that the method does not contribute to health inequalities.

Third, it must be clear to users of new technologies what benefits, if any, they can expect from using the technology. At a time where there is no disease-modifying treatment for dementia, the potential for changing the course of their condition is small, and supportive care for dementia is often poorly resourced. For users to make an informed decision to use the technology, they should be clear on what the outcome may (or may not) lead to. It may also be preferable to regulate companies providing these digital tools, ensuring that they provide appropriate follow up for their clients who are flagged as having possible dementia.

Finally, we hope to have shown that frameworks for population screening are both relevant and useful for considering the ethical and practical challenges associated with digital technologies for early disease detection. While discussions of these challenges often focus on the process of diagnosis within a clinical setting, this wider population health approach can offer a valuable, tried and tested framework for those involved in developing and implementing the tools discussed here.

We summarize our recommendations for ethical practice in this field in Box 1.

### 4.3 | Final conclusions

The evidence-base for benefits from early detection of dementia using digital biomarkers and automated methods is weak, and the possibility of harm is tangible. Before these technologies become common place, it should be clear how newly detected patients with dementia can and should be supported and how technologies can be rolled out without exacerbating digital health inequalities.

#### AUTHOR CONTRIBUTIONS

**Elizabeth Ford:** Conceptualization (equal); formal analysis (equal); project administration (equal); writing – original draft (equal). **Richard Milne:** Conceptualization (supporting); writing – original draft (supporting); writing – review and editing (equal). **Keegan Curlewis:** Data curation (supporting); investigation (supporting); writing – review and editing (supporting).

#### ACKNOWLEDGMENTS

We thank Ms. Katie Goddard who kindly proofread the manuscript.

#### FUNDING INFORMATION

Elizabeth Ford acknowledges funding from the NIHR Applied Research Collaboration Kent Surrey and Sussex (grant number NIHR200179). Views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. Richard Milne acknowledges funding from the Wellcome Trust (grant number 108413/A/15/D).

#### CONFLICT OF INTEREST STATEMENT

The authors have declared no conflicts of interest for this article.

#### DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

#### ORCID

Elizabeth Ford  <https://orcid.org/0000-0001-5613-8509>

Richard Milne  <https://orcid.org/0000-0002-8770-2384>

Keegan Curlewis  <https://orcid.org/0000-0002-2285-3798>

#### RELATED WIREs ARTICLES

[Computational resources in healthcare](#)

[Integrative biomarker discovery in neurodegenerative diseases](#)

#### FURTHER READING

Dashwood, M., Churchhouse, G., Young, M., & Kuruvilla, T. (2021). Artificial intelligence as an aid to diagnosing dementia: An overview. *Progress in Neurology and Psychiatry*, 25(3), 42–47.



- Dorsey, E. R., Papapetropoulos, S., Xiong, M., & Kiebertz, K. (2017). The first frontier: Digital biomarkers for neurodegenerative disorders. *Digital Biomarkers*, 1(1), 6–13.
- Ienca, M., Vayena, E., & Blasimme, A. (2018). Big data and dementia: Charting the route ahead for research, ethics, and policy. *Frontiers in Medicine*, 5, 13.
- Weatherby, T. J. M., & Agius, M. (2018). Ethical and organisational considerations in screening for dementia. *Psychiatry Danubina*, 30(Suppl 7), 463–468.
- Wilson, J. M. G., & Jungner, G. (1968). *Principles and practice of screening for disease*. [http://apps.who.int/iris/bitstream/handle/10665/37650/WHO\\_PHP\\_34.pdf?sequence=17](http://apps.who.int/iris/bitstream/handle/10665/37650/WHO_PHP_34.pdf?sequence=17)

## REFERENCES

- Abowd, G. D., Bobick, A. F., Essa, I. A., Mynatt, E. D., & Rogers, W. A. (2002). *The aware home: A living laboratory for technologies for successful aging*. Paper presented at the Proceedings of the AAAI-02 Workshop "Automation as Caregiver".
- Ackley, S. F., Zimmerman, S. C., Brenowitz, W. D., Tchetchen, E. J. T., Gold, A. L., Manly, J. J., Mayeda, E. R., Filshten, T. J., Power, M. C., Elahi, F. M., & Brickman, A. M. (2021). Effect of reductions in amyloid levels on cognitive change in randomized trials: Instrumental variable meta-analysis. *BMJ*, 372, n159.
- Ada Lovelace Institute, & The Health Foundation. (2021). *The data divide: Public attitudes to tackling social and health inequalities in the COVID-19 pandemic and beyond*. [https://www.adalovelaceinstitute.org/wp-content/uploads/2021/03/The-data-divide\\_25March\\_final-1.pdf](https://www.adalovelaceinstitute.org/wp-content/uploads/2021/03/The-data-divide_25March_final-1.pdf)
- Ahmed, M. R., Zhang, Y., Feng, Z., Lo, B., Inan, O. T., & Liao, H. (2019). Neuroimaging and machine learning for dementia diagnosis: Recent advancements and future prospects. *IEEE Reviews in Biomedical Engineering*, 12, 19–33. <https://doi.org/10.1109/RBME.2018.2886237>
- Aisen, P. S., Jimenez-Maggiore, G. A., Rafii, M. S., Walter, S., & Raman, R. (2022). Early-stage Alzheimer disease: Getting trial-ready. *Nature Reviews Neurology*, 18, 389–399.
- Akhila, J., Markose, C., & Aneesh, R. (2017). *Feature extraction and classification of dementia with neural network*. Paper presented at the 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT).
- Akl, A., Taati, B., & Mihailidis, A. (2015). Autonomous unobtrusive detection of mild cognitive impairment in older adults. *IEEE Transactions on Biomedical Engineering*, 62(5), 1383–1394.
- Alber, J., Lee, A., Menard, W., Monast, D., & Salloway, S. (2017). Recruitment of at-risk participants for clinical trials: A major paradigm shift for Alzheimer's disease prevention. *The Journal of Prevention of Alzheimer's Disease*, 4(4), 213–214.
- Aldus, C., Arthur, A., Dennington-Price, A., Millac, P., Richmond, P., Fox, C., Robinson, L., Stephan, B., Brayne, C., & Savva, G. (2020). Undiagnosed dementia in primary care: A record linkage study. *Health Services and Delivery Research*, 8(20), 1–108.
- Allard, M., Husky, M., Catheline, G., Pelletier, A., Dilharreguy, B., Amieva, H., Pérès, K., Foubert-Samier, A., Dartigues, J. F., & Swendsen, J. (2014). Mobile technologies in the early detection of cognitive decline. *PLoS One*, 9(12), e112197. <https://doi.org/10.1371/journal.pone.0112197>
- Alzheimer Europe. (2021). *National Dementia Strategies*. <https://www.alzheimer-europe.org/policy/national-dementia-strategies>
- Alzheimer Research UK Dementia Statistics Hub. (2021). *Public attitudes towards dementia*. <https://www.dementiastatistics.org/attitudes/>
- Alzheimer's Association. (2015). *Changing the trajectory of Alzheimer's disease: How a treatment by 2025 saves lives and dollars*. <https://www.alz.org/media/documents/changing-the-trajectory-r.pdf>
- Alzheimer's Disease International. (2021). *Dementia plans*. <https://www.alzint.org/what-we-do/policy/dementia-plans/>
- Alzheimer's Disease Neuroimaging Initiative. (2017). *Alzheimer's Disease Neuroimaging Initiative*. <https://adni.loni.usc.edu/>
- Alzheimer's Research U. K. Dementia Statistics Hub. (2018). *Different types of dementia*. <https://www.dementiastatistics.org/statistics/different-types-of-dementia/>
- Arribas-Ayllon, M. (2011). The ethics of disclosing genetic diagnosis for Alzheimer's disease: Do we need a new paradigm? *British Medical Bulletin*, 100, 7–21. <https://doi.org/10.1093/bmb/ldr023>
- Ayala Solares, J. R., Dilettta Raimondi, F. E., Zhu, Y., Rahimi, F., Canoy, D., Tran, J., Pinho Gomes, A. C., Payberah, A. H., Zottoli, M., Nazarzadeh, M., Conrad, N., Rahimi, K., & Salimi-Khorshidi, G. (2020). Deep learning for electronic health records: A comparative review of multiple deep neural architectures. *Journal of Biomedical Informatics*, 101, 103337. <https://doi.org/10.1016/j.jbi.2019.103337>
- Babarak, L. M., Menetski, J., Rebhan, M., Nisato, G., Zinggeler, M., Brasier, N., Baerenfaller, K., Brenzikofer, T., Baltzer, L., Vogler, C., Gschwind, L., Schneider, C., Streiff, F., Groenen, P. M. A., & Miho, E. (2019). Traditional and digital biomarkers: Two worlds apart? *Digital Biomarkers*, 3(2), 92–102. <https://doi.org/10.1159/000502000>
- Balogopalan, A., Eyre, B., Robin, J., Rudzicz, F., & Novikova, J. (2021). Comparing pre-trained and feature-based models for prediction of Alzheimer's disease based on speech. *Frontiers in Aging Neuroscience*, 13, 635945.
- Banerjee, S. (2010). Living well with dementia—Development of the national dementia strategy for England. *International Journal of Geriatric Psychiatry*, 25(9), 917–922.
- Bannister, R. (2021). Underfunded mental healthcare in the NHS: The cycle of preventable distress continues. *BMJ*, 375, n2706. <https://doi.org/10.1136/bmj.n2706>
- Banville, F., Nolin, P., Lalonde, S., Henry, M., Dery, M.-P., & Villemure, R. (2010). Multitasking and prospective memory: Can virtual reality be useful for diagnosis? *Behavioural Neurology*, 23, 209–211. <https://doi.org/10.3233/BEN-2010-0297>

- Barnes, D. E., Zhou, J., Walker, R. L., Larson, E. B., Lee, S. J., Boscardin, W. J., Marcum, Z. A., & Dublin, S. (2020). Development and validation of eRADAR: A tool using EHR data to detect unrecognized dementia. *Journal of the American Geriatrics Society*, 68(1), 103–111. <https://doi.org/10.1111/jgs.16182>
- Barnett, J. H., Lewis, L., Blackwell, A. D., & Taylor, M. (2014). Early intervention in Alzheimer's disease: A health economic study of the effects of diagnostic timing. *BMC Neurology*, 14(1), 1–9.
- Bayat, S., Babulal, G. M., Schindler, S. E., Fagan, A. M., Morris, J. C., Mihailidis, A., & Roe, C. M. (2021). GPS driving: A digital biomarker for preclinical Alzheimer disease. *Alzheimer's Research & Therapy*, 13(1), 1–9.
- Becker, J. T., Boiler, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The natural history of Alzheimer's disease: Description of study cohort and accuracy of diagnosis. *Archives of Neurology*, 51(6), 585–594.
- Ben Miled, Z., Haas, K., Black, C. M., Khandker, R. K., Chandrasekaran, V., Lipton, R., & Boustani, M. A. (2020). Predicting dementia with routine care EMR data. *Artificial Intelligence in Medicine*, 102, 101771. <https://doi.org/10.1016/j.artmed.2019.101771>
- Birkenbihl, C., Salimi, Y., Fröhlich, H., Japanese Alzheimer's Disease Neuroimaging Initiative, & Alzheimer's Disease Neuroimaging Initiative. (2022). Unraveling the heterogeneity in Alzheimer's disease progression across multiple cohorts and the implications for data-driven disease modeling. *Alzheimer's & Dementia*, 18(2), 251–261.
- Brayne, C., & Kelly, S. (2019). Against the stream: Early diagnosis of dementia, is it so desirable? *BJPsych Bulletin*, 43(3), 123–125.
- Brookmeyer, R., Abdalla, N., Kawas, C. H., & Corrada, M. M. (2018). Forecasting the prevalence of preclinical and clinical Alzheimer's disease in the United States. *Alzheimer's & Dementia*, 14(2), 121–129.
- Bruscoli, M., & Lovestone, S. (2004). Is MCI really just early dementia? A systematic review of conversion studies. *International Psychogeriatrics*, 16(2), 129–140. <https://doi.org/10.1017/S1041610204000092>
- Buettner, L. L., & Fitzsimmons, S. (2009). Promoting health in early-stage dementia: Evaluation of a 12-week course. *Journal of Gerontological Nursing*, 35(3), 39–49. <https://doi.org/10.3928/00989134-20090301-02>
- Canevelli, M., Grande, G., Lacorte, E., Quarchioni, E., Cesari, M., Mariani, C., Bruno, G., & Vanacore, N. (2016). Spontaneous reversion of mild cognitive impairment to normal cognition: A systematic review of literature and meta-analysis. *Journal of the American Medical Directors Association*, 17(10), 943–948. <https://doi.org/10.1016/j.jamda.2016.06.020>
- Caselli, R. J., Langbaum, J., Marchant, G. E., Lindor, R. A., Hunt, K. S., Henslin, B. R., Dueck, A. C., & Robert, J. S. (2014). Public perceptions of presymptomatic testing for Alzheimer disease. *Mayo Clinic Proceedings*, 89(10), 1389–1396.
- Cavedo, E., Lista, S., Khachaturian, Z., Aisen, P., Amouyel, P., Herholz, K., Jack, C. R., Jr., Sperling, R., Cummings, J., Blennow, K., O'Bryant, S., Frisoni, G. B., Khachaturian, A., Kivipelto, M., Klunk, W., Broich, K., Andrieu, S., de Schotten, M. T., Mangin, J. F., ... Hampel, H. (2014). The road ahead to cure Alzheimer's disease: Development of biological markers and neuroimaging methods for prevention trials across all stages and target populations. *The Journal of Prevention of Alzheimer's Disease*, 1(3), 181–202. <https://doi.org/10.14283/jpad.2014.32>
- Chen, R., Jankovic, F., Marinsek, N., Foschini, L., Kourtis, L., Signorini, A., Pugh, M., Shen, J., Yaari, R., Maljkovic, V., & Sunga, M. (2019). *Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams*. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery and data mining. <https://doi.org/10.1145/3292500.3330690>
- Cogstate. (2020). *Eisai and Cogstate expand agreement for global development and commercialization of digital cognitive assessment technologies*. <https://www.cogstate.com/eisai-cogstate-expand-agreement-global-commercialization-digital-cognitive-assessments/>
- Cordell, C. B., Borson, S., Boustani, M., Chodosh, J., Reuben, D., Verghese, J., Thies, W., Fried, L. B., & Medicare Detection of Cognitive Impairment Workgroup. (2013). Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare annual wellness visit in a primary care setting. *Alzheimer's & Dementia*, 9(2), 141–150. <https://doi.org/10.1016/J.JALZ.2012.09.011>
- Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G., Roses, A. D., Haines, J. L., & Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, 261(5123), 921–923.
- Corner, L., & Bond, J. (2004). Being at risk of dementia: Fears and anxieties of older adults. *Journal of Aging Studies*, 18(2), 143–155. <https://doi.org/10.1016/j.jaging.2004.01.007>
- Couteur, D. G. L., Doust, J., Creasey, H., & Brayne, C. (2013). Political drive to screen for pre-dementia: Not evidence based and ignores the harms of diagnosis. *BMJ*, 347, f5125. <https://doi.org/10.1136/bmj.f5125>
- Dagum, P. (2018). Digital biomarkers of cognitive function. *Npj Digital Medicine*, 1(1), 10. <https://doi.org/10.1038/s41746-018-0018-4>
- Dashwood, M., Churchhouse, G., Young, M., & Kuruvilla, T. (2021). Artificial intelligence as an aid to diagnosing dementia: An overview. *Progress in Neurology and Psychiatry*, 25(3), 42–47. <https://doi.org/10.1002/pnp.721>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.
- Dhedhi, S. A., Swinglehurst, D., & Russell, J. (2014). 'Timely' diagnosis of dementia: What does it mean? A narrative analysis of GPs' accounts. *BMJ Open*, 4(3), e004439.
- Dolph, C. V., Alam, M., Shboul, Z., Samad, M. D., & Iftekharuddin, K. M. (2017). *Deep learning of texture and structural features for multiclass Alzheimer's disease classification*. Paper presented at the 2017 International Joint Conference on Neural Networks (IJCNN).
- Dorsey, E. R., Papapetropoulos, S., Xiong, M., & Kiebert, K. (2017). The first frontier: Digital biomarkers for neurodegenerative disorders. *Digital Biomarkers*, 1(1), 6–13. <https://doi.org/10.1159/000477383>

- Dromey, J., & Hochlaf, D. (2018). *Fair care: A workforce strategy for social care*. Institute for Public Policy Research <https://www.ippr.org/files/2018-11/fair-care-workforce-strategy-nov18-summary.pdf>
- Dubois, B., Padovani, A., Scheltens, P., Rossi, A., & Dell'Agnello, G. (2016). Timely diagnosis for Alzheimer's disease: A literature review on benefits and challenges. *Journal of Alzheimer's Disease*, 49(3), 617–631.
- Eby, D. W., Silverstein, N. M., Molnar, L. J., LeBlanc, D., & Adler, G. (2012). Driving behaviors in early stage dementia: A study using in-vehicle technology. *Accident Analysis & Prevention*, 49, 330–337. <https://doi.org/10.1016/j.aap.2011.11.021>
- Eckerström, C., Olsson, E., Bjerke, M., Malmgren, H., Edman, Å., Wallin, A., & Nordlund, A. (2013). A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia. *Journal of Alzheimer's Disease*, 36(3), 421–431. <https://doi.org/10.3233/JAD-122440>
- EDON Initiative. (2021). *First digital data collected in global Alzheimer's wearable project*. <https://edon-initiative.org/first-digital-data-collected-in-global-alzheimers-wearable-project/>
- European Commission High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Faturrahman, M., Wasito, I., Hanifah, N., & Mufidah, R. (2017). *Structural MRI classification for Alzheimer's disease detection using deep belief network*. Paper presented at the 2017 11th International Conference on Information & Communication Technology and System (ICTS).
- Ford, E., Edelman, N., Somers, L., Shrewsbury, D., Lopez Levy, M., van Marwijk, H., Curcin, V., & Porat, T. (2021). Barriers and facilitators to the adoption of electronic clinical decision support systems: A qualitative interview study with UK general practitioners. *BMC Medical Informatics and Decision Making*, 21(1), 193. <https://doi.org/10.1186/s12911-021-01557-z>
- Ford, E., Rooney, P., Oliver, S., Hoile, R., Hurley, P., Banerjee, S., van Marwijk, H., & Cassell, J. (2019). Identifying undetected dementia in UK primary care patients: A retrospective case-control study comparing machine-learning and standard epidemiological approaches. *BMC Medical Informatics and Decision Making*, 19(1), 248. <https://doi.org/10.1186/s12911-019-0991-9>
- Ford, E., Sheppard, J., Oliver, S., Rooney, P., Banerjee, S., & Cassell, J. A. (2021). Automated detection of patients with dementia whose symptoms have been identified in primary care but have no formal diagnosis: A retrospective case-control study using electronic primary care records. *BMJ Open*, 11(1), e039248.
- Ford, E., Starlinger, J., Rooney, P., Oliver, S., Banerjee, S., van Marwijk, H., & Cassell, J. (2020). Could dementia be detected from UK primary care patients' records by simple automated methods earlier than by the treating physician? A retrospective case-control study. *Wellcome Open Research*, 5(120), 120.
- Fowler, N. R., Perkins, A. J., Gao, S., Sachs, G. A., & Boustani, M. A. (2020). Risks and benefits of screening for dementia in primary care: The Indiana University cognitive health outcomes investigation of the comparative effectiveness of dementia screening (IU CHOICE) trial. *Journal of the American Geriatrics Society*, 68(3), 535–543.
- Frederiksen, K. S., Nielsen, T. R., Appollonio, I., Andersen, B. B., Riverol, M., Boada, M., Ceccaldi, M., Dubois, B., Engelborghs, S., Frölich, L., & Hausner, L. (2021). Biomarker counseling, disclosure of diagnosis and follow-up in patients with mild cognitive impairment: A European Alzheimer's disease consortium survey. *International Journal of Geriatric Psychiatry*, 36(2), 324–333.
- Frey, A.-L., Karran, M., Jimenez, R., Baxter, J., Adeogun, M., Chan, D., & Hinds, C. (2019). *Harnessing the potential of digital technologies for the early detection of neurodegenerative diseases*. OSF Pre-Prints. <https://doi.org/10.31219/osf.io/u31249z31215>
- Fu, S., Ibrahim, O. A., Wang, Y., Vassilaki, M., Petersen, R. C., Mielke, M. M., St Sauver, J., & Sohn, S. (2022). Prediction of incident dementia using patient temporal health status. In *MEDINFO 2021: One world, one health—Global partnership for digital innovation* (pp. 757–761). IOS Press.
- Gale, S. A., Acar, D., & Daffner, K. R. (2018). Dementia. *The American Journal of Medicine*, 131(10), 1161–1169. <https://doi.org/10.1016/j.amjmed.2018.01.022>
- García-Casal, J. A., Franco-Martín, M., Perea-Bartolomé, M. V., Toribio-Guzmán, J. M., García-Moja, C., Goñi-Imizcoz, M., & Csipke, E. (2017). Electronic devices for cognitive impairment screening: A systematic literature review. *International Journal of Technology Assessment in Health Care*, 33(6), 654–673.
- Gauthier, S., Albert, M., Fox, N., Goedert, M., Kivipelto, M., Mestre-Ferrandiz, J., & Middleton, L. T. (2016). Why has therapy development for dementia failed in the last two decades? *Alzheimer's & Dementia*, 12(1), 60–64. <https://doi.org/10.1016/j.jalz.2015.12.003>
- Gelosa, G., & Brooks, D. J. (2012). The prognostic value of amyloid imaging. *European Journal of Nuclear Medicine and Molecular Imaging*, 39(7), 1207–1219. <https://doi.org/10.1007/s00259-012-2108-x>
- Giorgio, J., Jagust, W. J., Baker, S., Landau, S. M., Tino, P., & Kourtzi, Z. (2022). A robust and interpretable machine learning approach using multimodal biological data to predict future pathological tau accumulation. *Nature Communications*, 13(1), 1887. <https://doi.org/10.1038/s41467-022-28795-7>
- Goerden, J., Čukić, I., Danso, S. O., Carrière, I., & Muniz-Terrera, G. (2019). Statistical methods for dementia risk prediction and recommendations for future work: A systematic review. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 5, 563–569.
- Gold, M., Amatniek, J., Carrillo, M. C., Cedarbaum, J. M., Hendrix, J. A., Miller, B. B., Robillard, J. M., Rice, J. J., Soares, H., Tome, M. B., & Tarnanas, I. (2018). Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 4, 234–242.
- Goldman, J. S., Hahn, S. E., Catania, J. W., Larusse-Eckert, S., Butson, M. B., Rumbaugh, M., Strecker, M. N., Roberts, J. S., Burke, W., Mayeux, R., & Bird, T. (2011). Genetic counseling and testing for Alzheimer disease: Joint practice guidelines of the American College of Medical Genetics and the National Society of Genetic Counselors. *Genetics in Medicine*, 13(6), 597–605.

- Gomar, J. J. (2011). Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. *Archives of General Psychiatry*, 68(9), 961–969. <https://doi.org/10.1001/archgenpsychiatry.2011.96>
- Gomersall, T., Astell, A., Nygård, L., Sixsmith, A., Mihailidis, A., & Hwang, A. (2015). Living with ambiguity: A metasynthesis of qualitative research on mild cognitive impairment. *The Gerontologist*, 55(5), 892–912.
- Gooblar, J., Roe, C. M., Selsor, N. J., Gabel, M. J., & Morris, J. C. (2015). Attitudes of research participants and the general public regarding disclosure of Alzheimer disease research results. *JAMA Neurology*, 72(12), 1484–1490.
- Gunnarson, M., Kapeller, A., & Zeiler, K. (2021). Ethico-political aspects of conceptualizing screening: The case of dementia. *Health Care Analysis*, 29(4), 343–359.
- Gustavsson, A., Green, C., Jones, R. W., Förstl, H., Simsek, D., de Vulpillieres, F. D. R., Luthman, S., Adlard, N., Bhattacharyya, S., & Wimo, A. (2017). Current issues and future research priorities for health economic modelling across the full continuum of Alzheimer's disease. *Alzheimer's & Dementia*, 13(3), 312–321.
- Hampel, H., O'Bryant, S. E., Molinuevo, J. L., Zetterberg, H., Masters, C. L., Lista, S., Kiddle, S. J., Batrla, R., & Blennow, K. (2018). Blood-based biomarkers for Alzheimer disease: Mapping the road to the clinic. *Nature Reviews Neurology*, 14(11), 639–652. <https://doi.org/10.1038/s41582-018-0079-7>
- Harkins, K., Sankar, P., Sperling, R., Grill, J. D., Green, R. C., Johnson, K. A., Healy, M., & Karlawish, J. (2015). Development of a process to disclose amyloid imaging results to cognitively normal older adult research participants. *Alzheimer's Research & Therapy*, 7(1), 1–9.
- Harms, R. L., Ferrari, A., Meier, I. B., Martinkova, J., Santus, E., Marino, N., Cirillo, D., Mellino, S., Catuara Solarz, S., Tarnanas, I., & Szoek, C. (2022). Digital biomarkers and sex impacts in Alzheimer's disease management—Potential utility for innovative 3P medicine approach. *EPMA Journal*, 13, 1–15.
- Hollands, G. J., French, D. P., Griffin, S. J., Prevost, A. T., Sutton, S., King, S., & Marteau, T. M. (2016). The impact of communicating genetic risks of disease on risk-reducing health behaviour: Systematic review with meta-analysis. *BMJ*, 352, i1102. <https://doi.org/10.1136/bmj.i1102>
- Hu, C., Yu, D., Sun, X., Zhang, M., Wang, L., & Qin, H. (2017). The prevalence and progression of mild cognitive impairment among clinic and community populations: A systematic review and meta-analysis. *International Psychogeriatrics*, 29(10), 1595–1608.
- Ienca, M., Vayena, E., & Blasimme, A. (2018). Big data and dementia: Charting the route ahead for research, ethics, and policy. *Frontiers in Medicine*, 5(13). <https://doi.org/10.3389/fmed.2018.00013>
- Iliffe, S., Manthorpe, J., & Eden, A. (2003). Sooner or later? Issues in the early diagnosis of dementia in general practice: A qualitative study. *Family Practice*, 20(4), 376–381.
- Islam, J., & Zhang, Y. (2017). A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data. *Brain Informatics—Lecture Notes in Computer Science*, 10654, 213–222. [https://doi.org/10.1007/978-3-319-70772-3\\_20](https://doi.org/10.1007/978-3-319-70772-3_20)
- Ithapu, V. K., Singh, V., Okonkwo, O. C., Chappell, R. J., Dowling, N. M., Johnson, S. C., & Alzheimer's Disease Neuroimaging Initiative. (2015). Imaging-based enrichment criteria using deep learning algorithms for efficient clinical trials in mild cognitive impairment. *Alzheimer's & Dementia*, 11(12), 1489–1499.
- Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., & Liu, E. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4), 535–562.
- James, C., Ranson, J. M., Everson, R., & Llewellyn, D. J. (2021). Performance of machine learning algorithms for predicting progression to dementia in memory clinic patients. *JAMA Network Open*, 4(12), e2136553. <https://doi.org/10.1001/jamanetworkopen.2021.36553>
- Jammeh, E. A., Camille, B. C., Stephen, W. P., Escudero, J., Anastasiou, A., Zhao, P., Chenore, T., Zajicek, J., & Ifeakor, E. (2018). Machine-learning based identification of undiagnosed dementia in primary care: A feasibility study. *BJGP Open*, 2(2), bjgpopen18X101589.
- Jha, A., Tabet, N., & Orrell, M. (2001). To tell or not to tell? Comparison of older patients' reaction to their diagnosis of dementia and depression. *International Journal of Geriatric Psychiatry*, 16(9), 879–885. <https://doi.org/10.1002/gps.412>
- Ju, R., Hu, C., Zhou, P., & Li, Q. (2019). Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 16(1), 244–257. <https://doi.org/10.1109/TCBB.2017.2776910>
- Karlawish, J. (2011). Addressing the ethical, policy, and social challenges of preclinical Alzheimer disease. *Neurology*, 77(15), 1487–1493. <https://doi.org/10.1212/WNL.0b013e318232ac1a>
- Kim, S. Y., Karlawish, J., & Berkman, B. E. (2015). Ethics of genetic and biomarker test disclosures in neurodegenerative disease prevention trials. *Neurology*, 84(14), 1488–1494.
- Knapp, M., Barlow, J., Comas-Herrera, A., Damant, J., Freddolino, P. P., Hamblin, K., Hu, B., Lorenz, K., Perkins, M., Rehill, A., & Wittenberg, R. (2016). *The case for investment in technology to manage the global costs of dementia*. [https://piru.ac.uk/assets/files/Dementia\\_IT\\_PIRU\\_publ\\_18.pdf](https://piru.ac.uk/assets/files/Dementia_IT_PIRU_publ_18.pdf)
- Kourtis, L. C., Regele, O. B., Wright, J. M., & Jones, G. B. (2019). Digital biomarkers for Alzheimer's disease: The mobile/wearable devices opportunity. *Npj Digital Medicine*, 2(1), 1–9.
- Koychev, I., Lawson, J., Chessell, T., Mackay, C., Gunn, R., Sahakian, B., Rowe, J. B., Thomas, A. J., Rochester, L., Chan, D., & Tom, B. (2019). Deep and frequent phenotyping study protocol: An observational study in prodromal Alzheimer's disease. *BMJ Open*, 9(3), e024498.



- Lang, L., Clifford, A., Wei, L., Zhang, D., Leung, D., Augustine, G., Danat, I. M., Zhou, W., Copeland, J. R., Anstey, K. J., & Chen, R. (2017). Prevalence and determinants of undetected dementia in the community: A systematic literature review and a meta-analysis. *BMJ Open*, 7(2), e011146.
- Largent, E. A., Harkins, K., van Dyck, C. H., Hachey, S., Sankar, P., & Karlawish, J. (2020). Cognitively unimpaired adults' reactions to disclosure of amyloid PET scan results. *PLoS One*, 15(2), e0229137.
- Laver, K., Dyer, S., Whitehead, C., Clemson, L., & Crotty, M. (2016). Interventions to delay functional decline in people with dementia: A systematic review of systematic reviews. *BMJ Open*, 6(4), e010767. <https://doi.org/10.1136/bmjopen-2015-010767>
- Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., ... Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, 396(10248), 413–446. [https://doi.org/10.1016/S0140-6736\(20\)30367-6](https://doi.org/10.1016/S0140-6736(20)30367-6)
- Lussier, M., Lavoie, M., Giroux, S., Consel, C., Guay, M., Macoir, J., Hudon, C., Lorrain, D., Talbot, L., Langlois, F., Pigot, H., & Bier, N. (2019). Early detection of mild cognitive impairment with In-home monitoring sensor technologies using functional measures: A systematic review. *IEEE Journal of Biomedical and Health Informatics*, 23(2), 838–847. <https://doi.org/10.1109/JBHI.2018.2834317>
- Luz, S., Haider, F., De La Fuente Garcia, S., Fromm, D., & MacWhinney, B. (2021). Alzheimer's dementia recognition through spontaneous speech. *Frontiers in Computer Science*, 3, 780169.
- Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2007). Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19(9), 1498–1507.
- Martin, S., Kelly, S., Khan, A., Cullum, S., Denning, T., Rait, G., Fox, C., Katona, C., Cosco, T., Brayne, C., & LaFortune, L. (2015). Attitudes and preferences towards screening for dementia: A systematic review of the literature. *BMC Geriatrics*, 15(1), 1–12.
- Marzanski, M. (2000). Would you like to know what is wrong with you? On telling the truth to patients with dementia. *Journal of Medical Ethics*, 26(2), 108–113. <https://doi.org/10.1136/jme.26.2.108>
- Mattap, S. M., Mohan, D., McGrattan, A. M., Allotey, P., Stephan, B. C., Reidpath, D. D., Siervo, M., Robinson, L., & Chaiyakunapruk, N. (2022). The economic burden of dementia in low-and middle-income countries (LMICs): A systematic review. *BMJ Global Health*, 7(4), e007409.
- Matthews, F. E., Stephan, B. C., Robinson, L., Jagger, C., Barnes, L. E., Arthur, A., & Brayne, C. (2016). A two decade dementia incidence comparison from the cognitive function and ageing studies I and II. *Nature Communications*, 7, 11398.
- Mayeda, E. R., Glymour, M. M., Quesenberry, C. P., & Whitmer, R. A. (2016). Inequalities in dementia incidence between six racial and ethnic groups over 14 years. *Alzheimer's & Dementia*, 12(3), 216–224.
- McCoy, T. H., Han, L., Pellegrini, A. M., Tanzi, R. E., Berretta, S., & Perlis, R. H. (2019). Stratifying risk for dementia onset using large-scale electronic health record data: A retrospective cohort study. *Alzheimer's & Dementia*, 16, 531–540. <https://doi.org/10.1016/j.jalz.2019.09.084>
- Meier, I. B., Buegler, M., Harms, R., Seixas, A., Çöltekin, A., & Tarnanas, I. (2021). Using a digital neuro signature to measure longitudinal individual-level change in Alzheimer's disease: The Altoidea large cohort study. *Npj Digital Medicine*, 4(1), 1–9.
- Meyers, E., & Carrillo, M. C. (2020). Establishing a trial ready cohort to accelerate Alzheimer's clinical trial enrollment and treatments. *The Journal of Prevention of Alzheimer's Disease*, 7(4), 202–203.
- Milne, R., Diaz, A., Badger, S., Bunnik, E., Fauria, K., & Wells, K. (2018). At, with and beyond risk: Expectations of living with the possibility of future dementia. *Sociology of Health & Illness*, 40(6), 969–987.
- Modarres, M. H., Kalafatis, C., Apostolou, P., Marefat, H., Khanbagi, M., Karimi, H., Vahabi, Z., Aarsland, D., & Khaligh-Razavi, S. M. (2021). A remote digital tool for diagnosis and monitoring of Alzheimer's disease. *Alzheimer's & Dementia*, 17, e054123.
- Molinuevo, J. L., Cami, J., Carné, X., Carrillo, M. C., Georges, J., Isaac, M. B., Khachaturian, Z., Kim, S. Y. H., Morris, J. C., Pasquier, F., Ritchie, C., Sperling, R., & Karlawish, J. (2016). Ethical challenges in preclinical Alzheimer's disease observational studies and trials: Results of the Barcelona summit. *Alzheimer's & Dementia*, 12(5), 614–622. <https://doi.org/10.1016/j.jalz.2016.01.009>
- Müller, K., Fröhlich, S., Germano, A. M. C., Kondragunta, J., Agoitia Hurtado, M. F. D. C., Rudisch, J., Schmidt, D., Hirtz, G., Stollmann, P., & Voelcker-Rehage, C. (2020). Sensor-based systems for early detection of dementia (SENDA): A study protocol for a prospective cohort sequential study. *BMC Neurology*, 20(1), 1–15. <https://doi.org/10.1186/S12883-020-01666-8>
- Murugan, S., Venkatesan, C., Sumithra, M. G., Gao, X. Z., Elakkiya, B., Akila, M., & Manoharan, S. (2021). DEMNET: A deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images. *IEEE Access*, 9, 90319–90329. <https://doi.org/10.1109/ACCESS.2021.3090474>
- Muurling, M., deBoer, C., Kozak, R., Religa, D., Koychev, I., Verheij, H., Nies, V. J., Duyndam, A., Sood, M., Fröhlich, H., & Hannesdottir, K. (2021). Remote monitoring technologies in Alzheimer's disease: Design of the RADAR-AD study. *Alzheimer's Research & Therapy*, 13(1), 1–13.
- National Institute for Health and Care Excellence (NICE). (2018). *Dementia: Assessment, management and support for people living with dementia and their carers*. <https://www.nice.org.uk/guidance/ng97>
- Neuropathology Group, & Medical Research Council Cognitive Function and Ageing Study. (2001). Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *Lancet*, 357(9251), 169–175. [https://doi.org/10.1016/s0140-6736\(00\)03589-3](https://doi.org/10.1016/s0140-6736(00)03589-3)
- NHS Digital. (2022). *Recorded dementia diagnoses, May 2022*. <https://digital.nhs.uk/data-and-information/publications/statistical/recorded-dementia-diagnoses/may-2022>
- NHS England. (2017). *Dementia diagnosis rate workbooks*. <https://www.england.nhs.uk/publication/dementia-diagnosis-rate-workbook/>



- Nolin, P., Banville, F., Cloutier, J., & Allain, P. (2013). Virtual reality as a new approach to assess cognitive decline in the elderly. *Academic Journal of Interdisciplinary Studies*, 2(8), 612–616. <https://doi.org/10.5901/AJIS.2013.V2N8P612>
- Nori, V. S., Hane, C. A., Sun, Y., Crown, W. H., & Bleicher, P. A. (2020). Deep neural network models for identifying incident dementia using claims and EHR datasets. *PLoS One*, 15(9), e0236400.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- Odusami, M., Maskeliūnas, R., Damaševičius, R., & Krilavičius, T. (2021). Analysis of features of Alzheimer's disease: Detection of early stage from functional brain changes in magnetic resonance images using a Finetuned ResNet18 network. *Diagnostics*, 11(6), 1071. <https://doi.org/10.3390/diagnostics11061071>
- Ott, B. R., Pelosi, M. A., Tremont, G., & Snyder, P. J. (2016). A survey of knowledge and views concerning genetic and amyloid positron emission tomography status disclosure. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 2(1), 23–29.
- Parlevliet, J. L., Uysal-Bozkir, Ö., Goudsmit, M., van Campen, J. P., Kok, R. M., ter Riet, G., Schmand, B., & de Rooij, S. E. (2016). Prevalence of mild cognitive impairment and dementia in older non-western immigrants in the Netherlands: A cross-sectional study. *International Journal of Geriatric Psychiatry*, 31(9), 1040–1049.
- Pearson, C., Watson, N., Brunner, R., Cullingworth, J., Hameed, S., Scherer, N., & Shakespeare, T. (2022). Covid-19 and the crisis in social care: Exploring the experiences of disabled people in the pandemic. *Social Policy and Society*, 1–16, 1–16. <https://doi.org/10.1017/S1474746422000112>
- Pellegrini, E., Ballerini, L., Hernandez, M. d. C. V., Chappell, F. M., González-Castro, V., Anblagan, D., Danso, S., Muñoz-Maniega, S., Job, D., Pernet, C., Mair, G., MacGillivray, T. J., Trucco, E., & Wardlaw, J. M. (2018). Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10, 519–535. <https://doi.org/10.1016/j.dadm.2018.07.004>
- Perakslis, E., & Ginsburg, G. S. (2021). Digital health—The need to assess benefits, risks, and value. *JAMA*, 325(2), 127–128. <https://doi.org/10.1001/jama.2020.22919>
- Piau, A., Wild, K., Mattek, N., & Kaye, J. (2019). Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: Systematic review. *Journal of Medical Internet Research*, 21(8), e12785. <https://doi.org/10.2196/12785>
- Plancher, G., Tirard, A., Gyselinck, V., Nicolas, S., & Piolino, P. (2012). Using virtual reality to characterize episodic memory profiles in amnesic mild cognitive impairment and Alzheimer's disease: Influence of active and passive encoding. *Neuropsychologia*, 50(5), 592–602. <https://doi.org/10.1016/j.neuropsychologia.2011.12.013>
- Post, S. G. (1996). On not jumping the gun: Ethical aspects of APOE gene testing for Alzheimer's disease. *Annals of the New York Academy of Sciences*, 802(1), 111–119. <https://doi.org/10.1111/j.1749-6632.1996.tb32604.x>
- Prados, M. J., Liu, Y., Jun, H., Lam, J., & Mattke, S. (2022). Projecting the long-term societal value of a disease-modifying treatment for Alzheimer's disease in the United States. *Alzheimer's & Dementia*, 18(1), 142–151.
- Prince, M., Albanese, E., Guerchet, M., & Prina, M. (2014). *World Alzheimer report 2014 dementia and risk reduction: An analysis of protective and modifiable factors*. World Alzheimer Reports <https://www.alzint.org/u/WorldAlzheimerReport2014.pdf>
- Prince, M., Bryce, R., & Ferr, C. (2011). *Alzheimer's Disease International World Alzheimer Report 2011: The benefits of early diagnosis and intervention*. Institute of Psychiatry. King's College London. Alzheimer's Disease International <https://www.alzint.org/u/WorldAlzheimerReport2011.pdf>
- Qiu, S., Miller, M. I., Joshi, P. S., Lee, J. C., Xue, C., Ni, Y., Wang, Y., de Anda-Duran, I., Hwang, P. H., Cramer, J. A., Dwyer, B. C., Hao, H., Kaku, M. C., Kedar, S., Lee, P. H., Mian, A. Z., Murman, D. L., O'Shea, S., Paul, A. B., ... Kolachalama, V. B. (2022). Multi-modal deep learning for Alzheimer's disease dementia assessment. *Nature Communications*, 13(1), 3404. <https://doi.org/10.1038/s41467-022-31037-5>
- Quinn, C., Toms, G., Jones, C., Brand, A., Edwards, R. T., Sanders, F., & Clare, L. (2016). A pilot randomized controlled trial of a self-management group intervention for people with early-stage dementia (the SMART study). *International Psychogeriatrics*, 28(5), 787–800. <https://doi.org/10.1017/S1041610215002094>
- Rawtaer, I., Mahendran, R., Kua, E. H., Tan, H. P., Tan, H. X., Lee, T.-S., & Ng, T. P. (2020). Early detection of mild cognitive impairment with in-home sensors to monitor behavior patterns in community-dwelling senior citizens in Singapore: Cross-sectional feasibility study. *Journal of Medical Internet Research*, 22(5), e16854. <https://doi.org/10.2196/16854>
- Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, 27(3), 491–497.
- Reinke, C., Doblhammer, G., Schmid, M., & Welchowski, T. (2022). Dementia risk predictions from German claims data using methods of machine learning. *Alzheimer's & Dementia*. <https://doi.org/10.1002/alz.12663>
- Richard, E., den Brok, M. G., & van Gool, W. A. (2021). Bayes analysis supports null hypothesis of anti-amyloid beta therapy in Alzheimer's disease. *Alzheimer's & Dementia*, 17(6), 1051–1055.
- Ritchie, C. W., Terrera, G. M., & Quinn, T. J. (2015). Dementia trials and dementia tribulations: Methodological and analytical challenges in dementia research. *Alzheimer's Research & Therapy*, 7(1), 31. <https://doi.org/10.1186/s13195-015-0113-6>
- Roshanzamir, A., Aghajan, H., & Soleymani Baghshah, M. (2021). Transformer-based deep neural network language models for Alzheimer's disease risk assessment from targeted speech. *BMC Medical Informatics and Decision Making*, 21(1), 92. <https://doi.org/10.1186/s12911-021-01456-3>

- Royal College of General Practitioners. (2019). *RCGP position statement on screening by organisations which have not been approved by the UK National Screening Committee*. <https://www.rcgp.org.uk/policy/rcgp-policy-areas/screening.aspx>
- Sabbagh, M. N., Boada, M., Borson, S., Doraiswamy, P. M., Dubois, B., Ingram, J., Iwata, A., Porsteinsson, A. P., Possin, K. L., Rabinovici, G. D., & Vellas, B. (2020). Early detection of mild cognitive impairment (MCI) in an at-home setting. *The Journal of Prevention of Alzheimer's Disease*, 7(3), 171–178.
- Sarraf, S., & Tofighi, G. (2016). *Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks*. arXiv pre-print arXiv:1603.08631.
- Saunders, S., Ritchie, C. W., Russ, T. C., Muniz-Terrera, G., & Milne, R. (2022). Assessing and disclosing test results for 'mild cognitive impairment': The perspective of old age psychiatrists in Scotland. *BMC Geriatrics*, 22(1), 1–12.
- Schmutte, T., Olsson, M., Maust, D. T., Xie, M., & Marcus, S. C. (2022). Suicide risk in first year after dementia diagnosis in older adults. *Alzheimer's & Dementia*, 18(2), 262–271. <https://doi.org/10.1002/ALZ.12390>
- Schweda, M., Kögel, A., Bartels, C., Wiltfang, J., Schneider, A., & Schicktz, S. (2018). Prediction and early detection of Alzheimer's dementia: Professional disclosure practices and ethical attitudes. *Journal of Alzheimer's Disease*, 62(1), 145–155.
- Shao, Y., Zeng, Q. T., Chen, K. K., Shutes-David, A., Thielke, S. M., & Tsuang, D. W. (2019). Detection of probable dementia cases in undiagnosed patients using structured and unstructured electronic health records. *BMC Medical Informatics and Decision Making*, 19(1), 128. <https://doi.org/10.1186/s12911-019-0846-4>
- Shore, J., Kalafatis, C., Modarres, M. H., & Khaligh-Razavi, S.-M. (2022). Health economic analysis of the integrated cognitive assessment tool to aid dementia diagnosis in the United Kingdom. *medRxiv*, 2022.2005.2009.22274765.
- Silverstein, N. M. (2008). *Alzheimer's disease and fitness to drive*. Paper presented at the AAA Foundation for Traffic Safety conference AAA Foundation for Traffic Safety, Washington, DC.
- Slot, R. E., Sikkes, S. A., Berkhof, J., Brodaty, H., Buckley, R., Cavado, E., Dardiotis, E., Guillo-Benarous, F., Hampel, H., Kochan, N. A., & Lista, S. (2019). Subjective cognitive decline and rates of incident Alzheimer's disease and non-Alzheimer's disease dementia. *Alzheimer's & Dementia*, 15(3), 465–476.
- So, A., Hooshyar, D., Park, K. W., & Lim, H. S. (2017). Early diagnosis of dementia from clinical data by machine learning techniques. *Applied Sciences*, 7(7), 651. <https://www.mdpi.com/2076-3417/7/7/651>
- Solomonides, A. E., Koski, E., Atabaki, S. M., Weinberg, S., McGreevey, J. D., III, Kannry, J. L., Petersen, C., & Lehmann, C. U. (2022). Defining AMIA's artificial intelligence principles. *Journal of the American Medical Informatics Association*, 29(4), 585–591.
- Stephan, B. C., Pakpahan, E., Siervo, M., Licher, S., Muniz-Terrera, G., Mohan, D., Acosta, D., Pichardo, G. R., Sosa, A. L., Acosta, I., & Llibre-Rodriguez, J. J. (2020). Prediction of dementia risk in low-income and middle-income countries (the 10/66 Study): An independent external validation of existing models. *The Lancet Global Health*, 8(4), e524–e535.
- Sturdy, S., Miller, F., Hogarth, S., Armstrong, N., Chakraborty, P., Cressman, C., Dobrow, M., Flitcroft, K., Grossman, D., Harris, R., Hoebee, B., Holloway, K., Kinsinger, L., Krag, M., Löblová, O., Löwy, I., Mackie, A., Marshall, J., O'Hallahan, J., ... Zappa, M. (2020). Half a century of Wilson & Jungner: Reflections on the governance of population screening [version 2; peer review: 3 approved]. *Wellcome Open Research*, 5(158). <https://doi.org/10.12688/wellcomeopenres.16057.2>
- Swallow, J. (2020). Constructing classification boundaries in the memory clinic: Negotiating risk and uncertainty in constituting mild cognitive impairment. *Sociology of Health & Illness*, 42, 99–113.
- Teipel, S., Drzezga, A., Grothe, M. J., Barthel, H., Chételat, G., Schuff, N., Skudlarski, P., Cavado, E., Frisoni, G. B., Hoffmann, W., Thyrian, J. R., Fox, C., Minoshima, S., Sabri, O., & Fellgiebel, A. (2015). Multimodal imaging in Alzheimer's disease: Validity and usefulness for early detection. *The Lancet Neurology*, 14(10), 1037–1053. [https://doi.org/10.1016/S1474-4422\(15\)00093-9](https://doi.org/10.1016/S1474-4422(15)00093-9)
- Thabtah, F., Mampusti, E., Peebles, D., Herradura, R., & Varghese, J. (2020). A mobile-based screening system for data analyses of early dementia traits detection. *Journal of Medical Systems*, 44(1), 24. <https://doi.org/10.1007/s10916-019-1469-0>
- Thabtah, F., Peebles, D., Retzler, J., & Hathurusingha, C. (2020). Dementia medical screening using mobile applications: A systematic review with a new mapping model. *Journal of Biomedical Informatics*, 111, 103573. <https://doi.org/10.1016/j.jbi.2020.103573>
- The PHG Foundation. (2019). *Dementia risk prediction models: What do policy makers need to know?* <https://www.phgfoundation.org/documents/dementia-risk-prediction-models.pdf>
- Thyrian, J. R., Eichler, T., Pooch, A., Albuene, K., Dreier, A., Michalowsky, B., Wucherer, D., & Hoffmann, W. (2016). Systematic, early identification of dementia and dementia care management are highly appreciated by general physicians in primary care; results within a cluster-randomized-controlled trial (DelpHi). *Journal of Multidisciplinary Healthcare*, 9, 183–190. <https://doi.org/10.2147/JMDH.S96055>
- Thyrian, J. R., Hoffmann, W., & Eichler, T. (2018). Early recognition of dementia in primary care-current issues and concepts. *Current Alzheimer Research*, 15(1), 2–4.
- Tong, T., & Chignell, M. (2014). *Developing a serious game for cognitive assessment: Choosing settings and measuring performance*. Paper presented at the Proceedings of the Second International Symposium of Chinese CHI, Toronto, Ontario, Canada. <https://doi.org/10.1145/2592235.2592246>
- UK Government. (2021). *The pros and cons of screening*. <https://www.gov.uk/guidance/the-pros-and-cons-of-screening>
- US Federal Trade Commission. (2016). *Lumosity to pay \$2 million to settle FTC deceptive advertising charges for its "brain training" program*. <https://www.ftc.gov/news-events/news/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges-its-brain-training-program>

- Valladares-Rodriguez, S., Fernández-Iglesias, M. J., Anido-Rifón, L., Facal, D., Rivas-Costa, C., & Pérez-Rodríguez, R. (2019). Touchscreen games to detect cognitive impairment in senior adults. A user-interaction pilot study. *International Journal of Medical Informatics*, 127, 52–62. <https://doi.org/10.1016/J.IJMEDINF.2019.04.012>
- Valladares-Rodriguez, S., Perez-Rodriguez, R., Facal, D., Fernandez-Iglesias, M. J., Anido-Rifon, L., & Mouriño-Garcia, M. (2017). Design process and preliminary psychometric study of a video game to detect cognitive impairment in senior adults. *PeerJ*, 5, e3508. <https://doi.org/10.7717/peerj.3508>
- van der Flier, W. M., & Scheltens, P. (2005). Epidemiology and risk factors of dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, 76-(Suppl 5), v2–v7. <https://doi.org/10.1136/jnnp.2005.082867>
- Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific Reports*, 11(1), 3254. <https://doi.org/10.1038/s41598-020-74399-w>
- Walsh, C. G., Chaudhry, B., Dua, P., Goodman, K. W., Kaplan, B., Kavuluru, R., Solomonides, A., & Subbian, V. (2020). Stigma, biomarkers, and algorithmic bias: Recommendations for precision behavioral health with artificial intelligence. *JAMIA Open*, 3(1), 9–15. <https://doi.org/10.1093/jamiaopen/ooz054>
- Weatherby, T. J., & Agius, M. (2018). Ethical and organisational considerations in screening for dementia. *Psychiatry Danubina*, 30(Suppl 7), 463–468.
- Wells, C. E., & Smith, S. J. (2016). Diagnostic care pathways in dementia: A review of the involvement of primary care in practice and innovation. *Journal of Primary Care & Community Health*, 8(2), 103–111. <https://doi.org/10.1177/2150131916678715>
- Wikler, E. M., Blendon, R. J., & Benson, J. M. (2013). Would you want to know? Public attitudes on early diagnostic testing for Alzheimer's disease. *Alzheimer's Research & Therapy*, 5(5), 1–11.
- Wilson, J. M. G., & Jungner, G. (1968). *Principles and practice of screening for disease*. [http://apps.who.int/iris/bitstream/handle/10665/37650/WHO\\_PHP\\_34.pdf?sequence=17](http://apps.who.int/iris/bitstream/handle/10665/37650/WHO_PHP_34.pdf?sequence=17)
- Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., Prina, A. M., Winblad, B., Jönsson, L., Liu, Z., & Prince, M. (2017). The worldwide costs of dementia 2015 and comparisons with 2010. *Alzheimers Dement*, 13(1), 1–7. <https://doi.org/10.1016/j.jalz.2016.07.150>
- Wittenberg, R., Hu, B., Barraza-Araiza, L., & Rehill, A. (2019). *Projections of older people with dementia and costs of dementia care in the United Kingdom, 2019–2040* (CPEC Working Paper 5). [https://www.alzheimers.org.uk/sites/default/files/2019-11/cpec\\_report\\_november\\_2019.pdf](https://www.alzheimers.org.uk/sites/default/files/2019-11/cpec_report_november_2019.pdf)
- World Health Organisation. (2017). *Mental health atlas 2017*. Mental Health Atlas <https://www.who.int/publications/i/item/9789241514019>
- World Health Organisation. (2019). *Risk reduction of cognitive decline and dementia: WHO guidelines*. <https://www.who.int/publications/i/item/9789241550543>
- World Health Organisation. (2020). *Screening programmes: A short guide. Increase effectiveness, maximize benefits and minimize harm*. <https://apps.who.int/iris/bitstream/handle/10665/330829/9789289054782-eng.pdf>
- World Health Organisation. (2021). *Ethics and governance of artificial intelligence for health*. <https://www.who.int/publications/i/item/9789240029200>
- Zhang, D., & Shen, D. (2012). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS One*, 7(3), e33182. <https://doi.org/10.1371/journal.pone.0033182>
- Zhou, T., Thung, K.-H., Zhu, X., & Shen, D. (2019). Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Human Brain Mapping*, 40(3), 1001–1016. <https://doi.org/10.1002/hbm.24428>
- Zhu, F., Li, X., Tang, H., He, Z., Zhang, C., Hung, G.-U., Chiu, P. Y., & Zhou, W. (2020). Machine learning for the preliminary diagnosis of dementia. *Scientific Programming*, 2020, 5629090. <https://doi.org/10.1155/2020/5629090>

**How to cite this article:** Ford, E., Milne, R., & Curlewis, K. (2023). Ethical issues when using digital biomarkers and artificial intelligence for the early detection of dementia. *WIREs Data Mining and Knowledge Discovery*, e1492. <https://doi.org/10.1002/widm.1492>