

# Use of Immersive Virtual Reality in the Assessment and Treatment of Alzheimer's Disease: A Systematic Review

Felix Clay<sup>a,b,\*</sup>, David Howett<sup>c</sup>, James FitzGerald<sup>a,b</sup>, Paul Fletcher<sup>a,d,e</sup>, Dennis Chan<sup>c,f</sup> and Annabel Price<sup>a,b</sup>

<sup>a</sup>Cambridge and Peterborough Foundation Trust, Elizabeth House, Fulbourn Hospital Cambridge,

View metadata, citation and similar papers at [core.ac.uk](https://core.ac.uk)

brought to you by  CORE  
provided by Apollo

<sup>c</sup>University of Cambridge, Department of Clinical Neurosciences, Cambridge Biomedical Campus, Cambridge, UK

<sup>d</sup>Department of Psychiatry, University of Cambridge, Cambridge, UK

<sup>e</sup>Wellcome Trust MRC Institute of Metabolic Science, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

<sup>f</sup>Institute of Cognitive Neuroscience, University College London, Alexandra House, London, UK

Handling Associate Editor: Kerry Pike

Accepted 13 February 2020

## Abstract.

**Background:** Immersive virtual reality (iVR) allows seamless interaction with simulated environments and is becoming an established tool in clinical research. It is unclear whether iVR is acceptable to people with Alzheimer's disease (AD) dementia or useful in their care. We explore whether iVR is a viable research tool that may aid the detection and treatment of AD.

**Objectives:** This review examines the use of iVR in people with AD or mild cognitive impairment (MCI).

**Methods:** Medline, PsycINFO, Embase, CINAHL, and Web of Science databases were searched from inception. PRISMA guidelines were used with studies selected by at least two researchers.

**Results:** Nine studies were eligible for inclusion. None reported any issues with iVR tolerability in participants with MCI and AD on assessment or treatment tasks. One study demonstrated capability for detecting prodromal AD and correlated with neuroanatomical substrates. Two studies showed iVR to have high accuracy in differentiating participants with AD from controls but were not hypothesis driven or with adequate controls measures. In a small validation study and two longitudinal case studies, iVR cognitive training was positively rated but did not demonstrate reliable benefit.

**Conclusion:** iVR is emerging as a viable method of assessing older adults and people with AD. Strongest benefits were seen when closely integrated with theoretical models of neurodegeneration and existing screening methods. Further randomized controlled trials integrated with clinical populations are required. This will consolidate the power of iVR for assessment of MCI and clarify treatment efficacy beyond current applications in physical rehabilitation.

Keywords: Alzheimer's disease, cognitive remediation, mild cognitive impairment, spatial navigation, virtual reality

## INTRODUCTION

In this systematic review, we investigate the use of immersive virtual reality (iVR) technology in studies aiming to improve the diagnosis or treatment

\*Correspondence to: Dr. Felix Clay, Cambridge and Peterborough Foundation Trust, Elizabeth House, Fulbourn Hospital Cambridge, CB21 5EF, UK. Tel.: +44 1223 216 167; E-mail: felix.clay@cpft.nhs.uk.

of Alzheimer's disease (AD). New approaches are sought due to the high societal burden of AD combined with a lack of effective treatments. There is hope that supporting vulnerable people early may help prevent disease progression at or before the onset of mild cognitive impairment (MCI) [1]. iVR has already demonstrated advantages in the treatment of mental health conditions from anxiety disorders to psychosis [2]. iVR has been utilized in public awareness campaigns allowing users to take a 'walk through dementia' [3] and reminiscence activities allowing experiences to be shared with family [4]. These applications have an intuitive appeal but it remains unclear whether iVR can aid the diagnosis or treatment of AD dementia.

Virtual reality uses computer simulation to replace an individual's external sensory world with an artificial environment that updates corresponding to an individual's orientation and physical movement. Early virtual reality equipment was limited in its ability to create naturalistic interaction. Artificial environments were reliant on 2D displays and controllers such as joysticks and mouse devices which require practice for effective use (Picture 1). By contrast iVR uses head mounted displays (HMD) which adjust the input to each eye to create a true sense of perspective in 3D (Picture 2). Interaction is more naturalistic using head/hand movements or walking within a monitored field. Controllers are incorporated into the simulation, and joystick use is less common. We define 'immersive' virtual reality as systems that encompass the user's visual field and where virtual movement replicates actual head or bodily movement.

HMDs achieve immersion by excluding external visual input and updating the simulated visual environment in relation to head movements. This helps integrate vestibular, proprioceptive, and visual information. These sensory cues enable a more accurate representation of position [5] and are scarcely recruited during non-immersive VR paradigms. Collectively such cues contribute to an experience that is not only immersive but intuitive—participants do not need to be shown how to move, they just move. HMD systems are also rapidly dropping in price and are already affordable to individuals, clinics, and care homes.

iVR can also be achieved using Cave Automatic Virtual environments (CAVE systems, Picture 3). CAVE systems use a combination of video display walls, motion capture systems, and stereoscopic LCD shutter glasses to simulate a 3D world. An advantage

of CAVE systems is their ability to allow interaction with real world objects in a projected virtual world; however, these come at a greater cost than HMD-mediated iVR and require specialized and dedicated rooms.

Early iVR systems suffered lag in visual updating that led to mismatch between visual perception and vestibular feedback (e.g., head rotation or walking) that commonly induced nausea in participants [6]. Modern iVR technology overcomes these issues through a 1:1 correspondence of movement to visual perception; however, this has not yet been systematically appraised in a patient population. Other important technical issues in need of clarification include how well people with AD and MCI tolerate HMD equipment and whether individual differences in sensory impairment affects use of equipment.

iVR allows for naturalistic interaction with a simulated world which could potentially offer accurate discrimination of AD pathology on spatial cognitive tasks [7]. Unlike episodic memory, isolated impairments in spatial navigation are specific to early AD [8] and otherwise silent hippocampal strokes [9]. These impairments are not observed in aging, depression, or other neurodegenerative diseases. Tasks of spatial cognition may therefore offer earlier detection of AD ahead of current gold standard episodic memory measures. Immersive environments are critical for such spatial tasks that are dependent on medial temporal lobe structures; the primary site of neurodegeneration in AD [10–14]. iVR suits this purpose well through use of large landscapes that would not be possible with other techniques. While the utility of iVR is exemplified by spatial cognitive paradigms, this review examines the wider application of iVR to potential diagnostic and therapeutic applications that may benefit AD.

Previous reviews have looked at the use of virtual reality in diagnostic assessment of AD using a mixture of spatial cognition; task sequencing and episodic memory with no clear differential benefit noted compared to existing techniques [15–17]. However, these paradigms were not immersive and used deprecated technology. As a result, the validity, tolerability, and possible diagnostic utility of immersive tasks for AD remains unclear. It is also unclear whether iVR can be used for people with MCI to predict progression to AD or whether it might be sensitive to neurodegeneration before significant cognitive decline has occurred. In this review we address these issues.

Broader reviews of computer-based interventions for AD treatment have noted that non-immersive virtual experiences are enjoyed by participants, reduced apathy, and were preferred when compared to non-virtual experiences [18, 19]. Early reviews found very weak evidence that iVR was more effective than conventional virtual reality for cognitive training in people with a range of neurodegenerative conditions but did not use robust definitions of iVR or dementia subtypes meaning that it is harder to make conclusions about clinical applicability [20]. Kim et al.'s recent meta-analysis suggests modest effects of semi immersive VR treatment across a range of outcomes from physical fitness, cognition, and emotion but used an incomplete definition of iVR by suggesting use of a 3D display as sole criteria [21]. In both cases there was a risk of relevant studies being either missed or misclassified as immersive. Other recent reviews have looked much more broadly at treatment of neurocognitive disorders but did not look at specific applications in AD [22].

As a result, it remains unknown if iVR cognitive training or stimulation tasks lead to functional improvements in AD, benefit other important symptoms such as mood or wellbeing [23], or facilitate engagement in activities which provide other therapeutic benefit. Such activities might include physiotherapy or occupational therapy rehabilitation which could otherwise be difficult to access for patients with AD. No previous reviews have considered this area. In this review we included all treatment studies. While previous reviews of the field have produced a number of interesting observations, the current study sought to examine, extend, and refine these insights in two key ways. First, we emphasize the importance of a precise and consistent definition of iVR. Second, we focus entirely on AD rather than neurocognitive disorders more broadly.

It also remains unclear how well people with MCI or AD tolerate the experience of iVR. A previous review suggested people with AD suffer increased anxiety around virtual reality but this was not using an iVR application [24], whereas another study found that participants preferred an iVR-mediated cognitive assessment over pen-and-paper tests as it was more intuitive [25]. Evidence from managing behavioral disturbance in institutional settings has also demonstrated that even in these challenging settings iVR can be well tolerated and even enjoyed [24, 26]. In this review we explore this issue by appraising both quantitative and qualitative measures of acceptability.

In summary, iVR offers a novel and ecologically valid opportunity for diagnostic assessment and therapeutic intervention that may have unforeseen potential in AD. Given the prevailing emphasis on earlier diagnosis, cognitive-behavioral paradigms developed to target neuroanatomical sites compromised early in AD are vital. However, such assessments (e.g., navigation) have been overlooked by previous reviews, potentially owing to difficulty in application and replication. iVR paradigms overcome these limitations but the clinical benefit, acceptability, and engagement with this technology in AD and MCI has not been adequately reviewed. In this review we address these issues.

## METHOD

### *Systematic literature review*

A systematic literature review was conducted using the databases PubMed, Embase, Cinahl, PsycINFO, the Cochrane Library, and the Web of Science Core Collection. A protocol was developed based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement [27]. Details of the protocol for this systematic review were registered on the PROSPERO database [28] under registration number CRD42018115173.

### *Search strategies*

The databases were searched from inception up to July 26, 2019. Key phrases were determined in conjunction with an experienced research librarian (Isla Kuhn, IK) with the search terms optimized for each database. The search strategy was refined through comparison to the Cochrane review of virtual reality use with people who have suffered strokes [29].

Key terms are listed below and a copy of the full search strategy is available from the authors on request. Search items (also heading searches) included: 'Virtual Realit\*' OR 'VR' OR 'computer AND (simulation\* OR generat\* OR therap\* OR treatment\*)' AND 'Alzheimer\*'

### *Eligibility criteria*

Articles were accepted in all languages. No restrictions were set with regard to study type or the publication date.

*Inclusion:*

- Participants: Must include participants with AD or MCI at high risk of developing into AD (e.g., amnesic presentation/ biomarker positive).
- Intervention: Use of an immersive 3D virtual reality paradigm, e.g., 3D head mounted display or CAVE system
- Control: Nil specific criteria
- Outcome: Any study design including quantitative or qualitative outcomes based on individual data.

*Exclusion:*

- Participants: Study ONLY includes healthy controls/participants with other non-dementia or other dementia conditions, e.g., Lewy body dementia.
- Intervention: No virtual reality or non-immersive virtual reality used, e.g., 2D tablet/ conventional computer screen display.
- Control: Nil specific criteria.
- Outcome: Reviews/technology appraisals.

*Selection process*

- Step 1: Search results were uploaded into Rayyan Online Systematic Review Management Software [30] for initial screening. Duplicates were removed.
- Step 2: Title/abstract screening: The remaining titles and abstracts were screened independently by two team members (FC and JF) to identify studies that potentially met the inclusion criteria. If no abstract was available, additional information was searched manually on the internet to aid the selection process. Relevant review articles were retained for manual screening if they potentially referred to studies which met the eligibility criteria. A third author (AP) resolved disputes. Reasons for exclusion were noted.
- Step 3: Manual screening of review articles: The references and full text of retained review articles were searched independently by two members of the team (FC and JF) for additional relevant studies. Duplicates were removed.
- Step 4: Full text eligibility review: The full text of the remaining studies was assessed for eligibility independently by two team members (FC and DH). A third author (AP) resolved disputes. References of retained studies were searched for additional relevant studies which were then assessed for eligibility. Reasons for exclusion were noted.

- Step 5: Final screening during data extraction: Detailed analysis of papers for data extraction included final review of eligibility for inclusion.

*Data extraction and quality assessment*

A data extraction form for the included studies was developed in Windows Office Excel 2010, based on the Cochrane Handbook for Systematic Reviews [31]. The data extraction form used is available from the authors upon request: We included:

- Data about the publication: Authors, title of the article/journal, year of publication.
- Method
- Study design, setting, diagnosis and severity of dementia, comorbidities
- Description of the VR assessment/intervention and control, number of participants in the VR assessment/intervention and control/comparison group. For treatment studies: including how many were randomized and how many analyzed (with reasons for drop out), duration of the trial (incl. duration of the intervention and timing of follow up).
- Outcome measures, type of analysis
- Results
- Outcome data for patients with AD/high risk MCI was compared to controls including other neurodegenerative conditions.
- For assessment studies: between groups (VR versus control condition) effects: *p*-values & effect sizes (95% CI, Cohen's *d*).
- For treatment studies: baseline and follow up(s), between groups (VR versus control condition) effects: *p*-values & effect sizes (95% CI, Cohen's *d*).
- Summary of the study conclusion.

We noted whether the authors needed to be contacted for any additional information and if so, which specific information was required. Three reviewers (FC, JF, DH) extracted the data independently. Any disagreement between the reviewers was resolved through discussion with a fourth reviewer (AP). If relevant information could not be found in the included studies, the authors were requested to provide the missing information if possible.

Two reviewers independently (FC and DH) assessed the risk of bias of the included studies. This was based on the Cochrane Collaboration's tool for assessing risk of bias which includes selection bias, performance bias, detection bias, attrition bias,

reporting bias, and any other potential sources of bias [32]. We also used the ROBINS-I tool for assessing risk of bias in non-randomized trials where this was more suitable [33]. Any disagreements between the two review team members about the risk of bias assessment was resolved through discussion with a third review team member (AP) if necessary. DH was not involved with any data extraction or critical appraisal of the study by Howett et al. [34] due to his involvement with this study as primary author.

## RESULTS

### Study selection

The selection process is summarized according to PRISMA guidelines [27] in Fig. 1 as a flow chart. The initial search strategy yielded 7,805 results from the selected databases, comprising 2,869 records from Medline via OVID, 1,351 records from Embase via OVID, 736 records from PsycINFO via Ebsco, 136 records from CINAHL via Ebsco, 208 records from Cochrane, and 2,505 records from Web of Science.

After removal of duplicates, 5,823 records were screened by two authors based on title, abstract, and,

if applicable, additional information. Disagreements about in- or exclusion of 118 studies were resolved by a team of three authors. In this first stage, 5,731 records were excluded for the following primary reasons: 5,574 wrong intervention (i.e., not immersive VR study), 157 wrong participants (i.e., not including participants with probable or possible AD/MCI).

The full text was retrieved of the remaining articles ( $n=92$ ) and of relevant reviews ( $n=41$ ). References of these articles were screened. This led to 5 records being added for full text screening. Further details of 3 conference posters were not available but authors were contacted where possible to obtain details of subsequent publications.

The remaining 97 articles were read in full text and assessed against eligibility criteria by two authors. 88 were excluded leaving a total of 9 studies included in this review (5 assessment studies, 4 treatment studies). Meta-analysis was not possible due to heterogeneity of study designs, intervention contents, outcomes measured, and clinical groups sampled.

### Assessment studies

The five eligible studies are summarized below. Full details are shown in Table 1a:

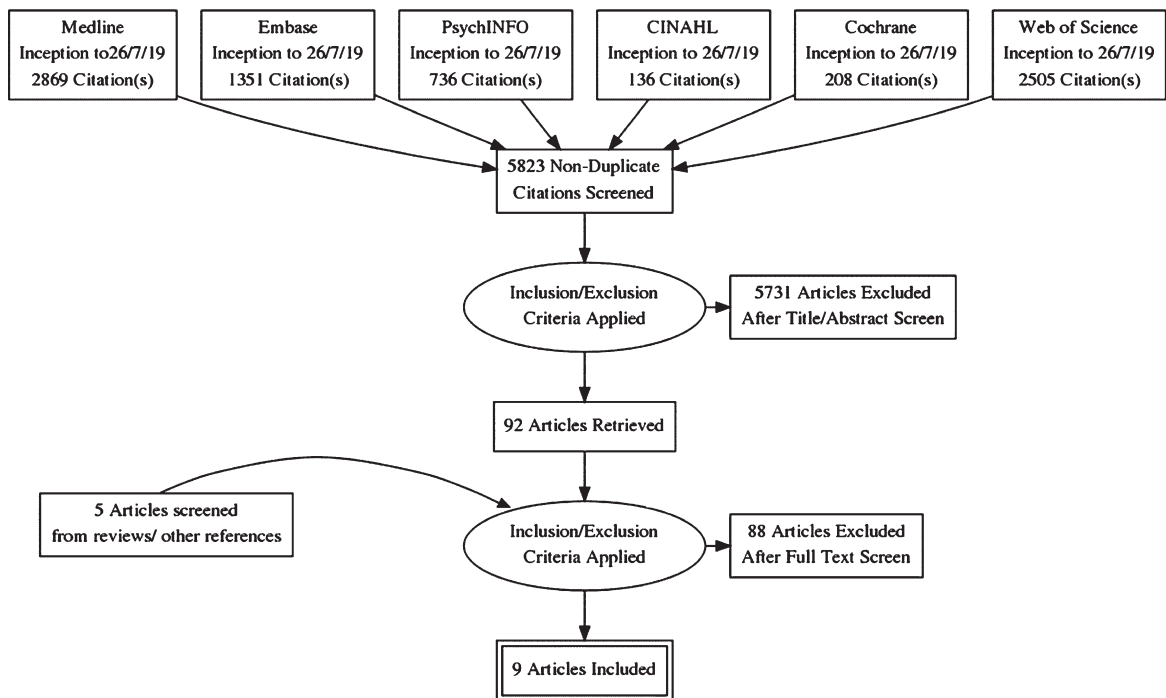


Fig. 1. PRISMA Diagram.

Table 1a  
Design and Aims of Assessment Studies

Study	Design/aims	Setting/ participants	Diagnosis/ comorbidities	VR technology/ task	Outcome measures	Comparison task
Fernandez-Montenegro et al. [35]	Validation Study (Repeated Measures): To compare novel VR screening tests of AD to traditional pen and paper tests	Computer Sciences Department, Kingston University, UK <i>N</i> = 20 aged 23–82, healthy controls and patients (no details on proportion), 11 female, 11 high school education or higher. Nationalities included French, Spanish, Vietnamese and Greek, one participant was illiterate.	AD, No details given of criteria or number of patients/comorbidities	Oculus Rift DK2 HMD with Microsoft Kinect 2 depth sensor. Participant is sat in a simulated consulting room to complete tasks: Remembering object locations/noticing scene changes and sound location. No information on software used.	Four tasks measuring visual memory, auditory recognition, and discrimination	Dr. Oz Memory Quiz; Visual Association Test, Dichotic Listening Test
Gago et al. [36]	Case Control: To explore differences in provoked compensatory postural adjustments within an VR environment in patients with AD who had suffered falls in the past 12 months ('fallers') to 'non-fallers' and controls	Neurology Department, Guimaraes, Portugal. <i>N</i> = 21 patients recruited from local memory clinic. –11 'fallers' 9 female, mean age 71, average 1-4 years education, mean disease 3 years, mean MoCA score 9. –9 'non-fallers' 5 females, mean age 75, average 1-4 years education, mean disease duration 3 years, mean MoCA score 12 <i>N</i> = 19 healthy controls (from previous study). 11 females, mean age 71, average 1-4 years of education, mean MoCA score 24	AD, DSM-IV/NINCDS-ADRA, Comorbidities not noted. Exclusion: orthopedic, musculoskeletal, and vestibular disorder, significant auditory deficit, and alcohol abuse, and somatosensory deficit.	Oculus Rift HMD using Generic 'Tuscany scene'. Participants identify objects then stand still on a virtual stair case as unpredictable visual displacements are made.	Compensatory postural adjustment measured by vertical acceleration and time frequency distribution	Not used

(Continued)

Table 1a  
(Continued)

Study	Design/aims	Setting/ participants	Diagnosis/ comorbidities	VR technology/ task	Outcome measures	Comparison task
Seo et al. [37]	Validation study (Repeated Measures): To explore the validity and discriminative power of a virtual daily living test as a new diagnostic approach to assess MCI	Department of Neurology and Department of Engineering, Hanyang University Hospital, South Korea. <i>N</i> = 20 patients (8 female), recruited from tertiary medical center. <i>N</i> = 22 healthy controls (8 female) recruited from same medical center.	MCI, Albert (2011) criteria, Comorbidities not stated. Exclusion: psychiatric or neurological illness, drug/alcohol abuse within 4 weeks of the study.	CAVE system with 4'2.5'x2.5 m room, four rear projection screens, stereoscopic glasses to give 3D vision and Optitrack motion tracking. IADL tasks performed 'Withdraw money from ATM'; 'Take a bus'	Kinematic data collected during through motion tracking sensors (hands and head). Distance of sensors, time taken (whole task) and velocity of movement and number of errors: broken sequence, misremember information, e.g., pin code or incorrect bus information.	MMSE, IADL, FCSRT, Digit span forwards and backwards, 2*trail making tasks.
Zakzanis et al. [38]	Case Control: To examine age- and AD-related differences in route learning and memory using VR	Psychology Department, University of Toronto, Canada. <i>N</i> = 2 patients (mean age 70). <i>N</i> = 9 young controls (mean age = 26) <i>N</i> = 7 older controls (mean age 62). All highly educated, similar computer use difference gaming exp	Probable AD (NINCDS-ADRA). Comorbidities not stated.	Proview XL-50 HMD with LCD display and movement controlled using data glove. Navigation of specified routes around simulated town of Sunnybrook.	Completion time, distance travelled, wrong turns, 'falls off sidewalk', collisions with objects	Not used.
Howett et al. [34]	Case Control: To test the hypothesis that entorhinal cortex-based navigation is impaired in pre-dementia AD.	Cambridge University Hospitals Memory Clinic, Cambridge, UK <i>N</i> = 45 patients with MCI (26 AD biomarker tested: 12 positive, 14 negative) <i>N</i> = 41 controls recruited from 'Join Dementia Research' group. Age, gender, and education matched.	MCI (Petersen criteria. Exclusion: Patients with major medical, psychiatric, visual, mobility impairment, or history of alcohol dependency.	HTC Vive iVR equipment with external base stations used to monitor participants within a 3.5 m <sup>3</sup> space. Path integration task within a virtual open area. Participants walk a route and then return to their base.	Absolute distance error; proportional angular and linear errors.	ACE-R, NART, measures sensitive to transition from MCI to AD (FCSRT, Rey figure recall, trail Making Task, Digit Symbol test, 4MT)

4MT, 4 Mountains Test; ACE-R, Addenbrooke's Cognitive Examination Revised; AD, Alzheimer's disease; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders IV; FCSRT, Free and Cued Selective Reminding Test; HMD, head mounted display; IADL, instrumental activities of daily living; iVR, immersive virtual reality; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; NART, National Adult Reading Test; NINCDS-ADRA, National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association; VR, virtual reality.

Design/background; Table 2a: Outcome measures and acceptability; Table 3: Risk of bias.

*Fernandez-Montenegro et al. [35]*

This validation study compared four novel iVR tasks with existing AD screening measures with the aim of demonstrating potential alternative screening measures. A mixture of healthy controls and participants with AD completed static iVR object recall tasks in an HMD setting with joystick control (total  $n=20$ ). All iVR and existing screening tasks were able to discriminate participants with AD from controls ( $p < 0.001$  for all tasks). Results were potentially confounded by both age and education level. There was high correlation between three out of four of the novel VR tasks and existing measures. There was no comparison between groups and overall risk of bias was noted to be high. There was no specific data on acceptability. Some patients with 'advanced' AD did not complete testing.

*Gago et al. [36]*

This case control study used a generic scene from the Oculus Rift HMD catalogue to measure postural stability in an iVR task for participants with AD whom had suffered recent falls with those whom had not ( $n=11$ ,  $n=9$ ) and matched controls ( $n=19$ ). Non parametric statistics showed that AD fallers made higher frequency postural adjustments at all time periods ( $p < 0.05$ ) with both AD fallers and non-fallers also showing slower response times to visual perturbations compared to controls. Risk of bias around the design was high but results and analyses were less problematic. There was no specific data on acceptability. All participants completed all sessions.

*Seo et al. [37]*

This validation study compared participants with MCI ( $n=20$ ) to healthy controls ( $n=22$ ) during an iVR task in which participants completed instrumental activities of daily living (IADLs) in a CAVE system. Performance was compared to standard neuropsychology measures with the aim of finding tasks which contributed additional discriminative value between groups. Results showed significant differences between participants and controls of hand speed on a cash machine task ( $p < 0.005$ ) and head speed and errors on the catching a bus task ( $p < 0.005$ ,  $p < 0.001$ ). These corresponded to large effect sizes (1.1, 1.5). There were also significant between group differences on existing neuropsychology assessments. This study was noted to be of high

methodological quality with little risk of bias. There was no specific data on acceptability. All participants completed all sessions.

*Zakzanis et al. [38]*

This case control study compared navigational abilities of young controls ( $n=9$ ) and older controls ( $n=7$ ) with a subset of two participants with AD on an iVR spatial navigation task using HMD and a joystick. There was no control task. Performance on navigation tasks was varied for the two participants with AD; one markedly impaired on all tasks, the other within 1 standard deviation of controls for all tasks. There was no between group comparison made which included participants with AD. There was no specific data on acceptability. All participants completed all sessions.

*Howett et al. [34]*

This case control study compared performance of participants with MCI ( $n=45$ ) to matched controls ( $n=41$ ) during an iVR path integration task designed to test entorhinal cortex function. A proportion of participants with MCI were tested for cerebrospinal fluid biomarkers of AD (amyloid- $\beta$  1-42, total tau, phosphorylated tau,  $n=26$ , 12 positive, 14 negative). A proportion of test subjects also completed MRI measures of entorhinal cortex volume (39 participants with MCI: 11 biomarker positive, 9 biomarker negative, 37 controls). Participants with MCI who had cerebrospinal fluid biomarkers positive for AD showed significantly impaired path integration performance compared to participants with MCI with negative biomarkers and controls whose performance did not significantly differ. The path integration task showed significantly higher diagnostic sensitivity and specificity for differentiating biomarker positive versus negative MCI compared to the best of current gold standard cognitive tests (area under the curve 0.90 versus 0.57). Entorhinal cortex volume on MRI was negatively associated with errors on the path integration task ( $p < 0.005$ ). This study was at low risk of bias with good methodological quality. There was no specific data on acceptability. All participants completed all sessions.

*Treatment studies*

The four eligible studies are summarized below. Full details are shown in Table 1b: Design/background; Table 2b: Outcome measures and acceptability; Table 3: Risk of bias.



Table 1b  
Design and Aims of Treatment Studies

Study	Design/aims	Setting/ participants	Diagnosis/ comorbidities	VR technology/ task	Outcome measures	Comparison task
Bourellier et al. [39]	Validation study (Repeated measures): To assess the use of a VR training environment to Engage Motor and Postural Abilities.	Department of Geriatrics, University of Bourgogne, France. <i>N</i> = 7 patients, average MMSE 23.2, Average timed up and go test 13.3 s <i>N</i> = 14, controls matched for age/gender, average MMSE 28.5, Average timed up and go test 9.5 secs	MCI, NINCDS-ADRDA criteria; comorbidities not mentioned, all normal/corrected hearing and vision	CAVE System with dimensions 3.40 m <sup>3</sup> , 3D vision via NVidia 3D Vision Pro stereoscopic glasses and ART DTrack 2motion tracking. Simulated fruit harvesting task designed to induce stretching: fruit in from different start positions are selected for ripeness and placed in a basket. 6* 3-min sessions with 1-min break between each session.	Number of appropriate movements: Ripe fruit picked on implicit VR task versus targeted movement on explicit exercise task.	Explicit exercise task: Arm pointing task with physiotherapist.
White and Moussavi. [40]	Case study: To determine whether an individual with AD could learn to navigate in a simple VR navigation environment	Department of Biomedical Engineering, University of Manitoba, Canada. <i>N</i> = 1, 74-year-old male, masters level education, MoCA 24/30	MCI, probable AD (no criteria given), comorbidities not mentioned.	Oculus Rift DK2 HMD: Participant is shown a room on the outside of a building and asked to navigate to it. 45-min training sessions, 3*week for 7 consecutive weeks (total 16 h)	Navigation errors, serial MoCA score.	Not used.
Optale et al. [41]	Case study: To explore whether cognitive processes could be restored via procedures practiced repetitively within an environment which contains functional real-world demands	Department of Neuro-Rehabilitation, Villa Salus Hospital, Venice, Italy. <i>N</i> = 1 Patient from memory clinic, female, aged 65, high school education.	Early onset AD (no criteria given). Full details of all comorbidities mentioned. No family history of dementia.	Virtual Research System HMD with Intersense motion tracker and Joystick control. Three listening experiences alternated with three different virtual experiences using Superscape VRT software: re-experience of childhood,	Subjective measures of improvement, Multiple Delayed Reaction Verbochronometry and Neuropsychology measures (MMSE, Wechsler Memory Scale, Digit span, Praxia Tests, Frontal Lobe measures	Not used.

(Continued)

Table 1b  
Continued)

Study	Design/aims	Setting/ participants	Diagnosis/ comorbidities	VR technology/ task	Outcome measures	Comparison task
				participating in a tournament, and walking the streets of a modern city. Tasks gradually more complex. 15 min, 3* week for 12 weeks, then reduced to 3*weekly sessions every two weeks for a further 12 weeks (13.5 h total).	(towers of Hanoi, Stroop test), Visual-spatial Tests Cubes, Global Deterioration Scale) at baseline, 5, 8, and 12 months	
Micarelli et al. [42]	Case Control: Explore the effect of a head motion-dependent VR racing game on vestibular rehabilitation in UVH patients encompassing MCI and cognitively unimpaired older adults.	ITER Center for Balance and Rehabilitation Research, Rome, Italy. $N = 24$ patients with MCI (mean age 74.4, 14 females, mean MMSE 25.5) $N = 23$ controls (mean age 75.6, 12 females, mean MMSE 28.1)	MCI diagnosis according to Albert criteria (2011). Chronic UVH diagnosed by a 25% reduction in vestibular response to bithermal water caloric irrigation on one side 3 months after the beginning of symptoms. Exclusion: Significant medical disorders.	VR Tech: 5.2" display of a Windows Phone placed in the HMD 'Revelation' VR Headset. Trackspeed 3D racing game 20 min per day: never ending racing race in which the car is steered from the cockpit by tilting the head	Otoneurological testing: Head impulse testing, Posturography. Self-report: Dizziness Handicap Inventory, activities-specific balance confidence scale, dynamic gait index. Simulator sickness Questionnaire, nausea oculomotor stress, and disorientation.	Vestibular rehabilitation only, twice daily 30–40 min in total.

UVH, unilateral vestibular hypofunction. See Table 1A for remaining abbreviations.

#### *Bourellier et al. [39]*

This validation study assessed the use of an iVR training environment for participants with MCI ( $n = 7$ ) versus matched controls ( $n = 14$ ) in comparison to a usual physiotherapy task. The iVR task used a CAVE system object sorting task which required large movements of the body, this was compared to a standard physiotherapy session practicing similar movements over a single session. Results showed a strong difference in appropriate movements between both MCI participants versus controls and iVR versus conventional physiotherapy (both  $p < 0.001$ ). There was also a strong interaction: participants with MCI were proportionately worse than controls on the iVR task in particular. There was a between groups differences on Mini-Mental State Examination (MMSE) and timed up and go task (TUG) which was con-

trolled for during statistical analysis. This study was at low risk of bias with good methodological quality. Acceptability of the technology was explored in detail through a 30-min semi structured interview: participants reported enjoying the iVR training more than the standard physiotherapy sessions which they found repetitive.

#### *White and Moussavi [40]*

This case study with a single participant looked at whether a patient with AD could engage with and learn from an iVR navigation program completed using HMD and innovative wheel chair control. No details were given about diagnostic criteria used. There was no control task. Training was completed three times a week over 7 weeks. Results showed

Table 2a  
Results of Assessment Studies

Study	Between groups differences	Outcome measures	Acceptability	Effect sizes	Subjective measures
Fernandez-Montenegro et al. [35]	No comparison made between groups based on demographic measures.	<p>Parametric statistics used to compare groups (Dementia versus controls, table 4): Significant difference noted between groups for all neuropsychology and novel measures (all <math>p &lt; 0.001</math>). Also, significant differences between age and DrOz test (<math>p = 0.033</math>), DLT (<math>p = 0.038</math>), and novel object recognition task (<math>p = 0.003</math>). Significant differences also between level of education and three of the novel tasks VOM (<math>p = 0.045</math>), AOR (<math>p = 0.004</math>), BDTT (<math>p = 0.0</math>).</p> <p>Correlation between novel and existing measures</p> <ul style="list-style-type: none"> <li>• VOM (<math>t = 0.59-0.83</math>, <math>p &lt; 0.01</math> for all measures)</li> <li>• AOR (<math>t = 0.53-0.74</math>, <math>p &lt; 0.01</math> for all measures except VAT <math>p = 0.017</math>)</li> <li>• VRS (<math>t = 0.49-0.72</math>, <math>p &lt; 0.01</math> for all measures except DrOz <math>p = 0.03</math>)</li> <li>• BDTT: Not significantly correlated with any existing measure.</li> </ul>	No specific data. Some patients with 'advanced' AD did not complete all sessions.	No data to allow this.	Positive feedback on VR headset, instructions and movement interaction, no statistics.
Gago et al. [36]	Nil significant differences demographic and anthropometric measures; significantly lower score on MoCA for both dementia groups versus controls (Kruskall Wallis Tests, $p < 0.001$ further details given).	<p><i>Vertical acceleration</i></p> <p>Significant difference between controls and AD fallers (higher mean, <math>U = 38.0</math> <math>p = 0.03</math>; higher average magnitude (<math>U = 53.0</math>, <math>p = 0.026</math>)).</p> <p>Movement power: Intergroup analysis</p> <ul style="list-style-type: none"> <li>• Significant difference between AD fallers to controls for all time periods (<math>p &lt; 0.05</math>)</li> <li>• No significant differences in low band (<math>&lt; 1.5</math> Hz) movements between other groups (<math>p</math> ranges 0.26-0.56)</li> <li>• Significantly higher value of high band (<math>&gt; 1.5</math> Hz) compensatory adjustments for AD fallers versus controls at all times (table 2, <math>p &lt; 0.05</math> for all comparisons).</li> </ul>	No specific data. All participants completed all sessions.	Not calculated.	Not used.

(Continued)

Table 2a  
(Continued)

Study	Between groups differences	Outcome measures	Acceptability	Effect sizes	Subjective measures
		<p>Movement power: Intragroup analysis</p> <ul style="list-style-type: none"> <li>• Controls increased power immediately after perturbation (0-4 s) versus before (-4-0 s, <math>p=0.002</math>) and at delay (4-8 s, <math>p=0.01</math>)</li> <li>• Both AD faller and non-faller groups showed increased power after a delayed period (0-4 versus 4-8 s, <math>p=0.005</math> and <math>p=0.20</math>)</li> </ul>			
Seo et al. [37]	<p>Nil significant difference between groups based on gender or demographics/ depression scale (Table 1). There were significant differences between groups on some of the cognitive assessments on MANOVA</p> <ul style="list-style-type: none"> <li>• MMSE, <math>F=7.9</math>, <math>p=0.007</math>,</li> <li>• Immediate recall, <math>F=22.1</math>, <math>p&lt;0.001</math></li> <li>• Delayed recall, <math>F=7.4</math>, <math>p=0.009</math></li> <li>• Digit span backwards, <math>F=7.7</math>, <math>p=0.009</math></li> <li>• Trail making task A time, <math>F=4.4</math>, <math>p=0.043</math></li> </ul>	<p>VDLT compared between MCI and controls MANOVA (<math>F(8,33)=5.6</math>, <math>p&lt;0.001</math>) univariate analysis with Bonferroni correction showed significant results for Hand speed on task 1 (<math>p=0.001</math>), slower head speed and more errors on task 2, <math>p=0.002</math>, <math>p&lt;0.001</math>).</p> <p>Correlation with neuropsychology tests Hand speed in task 1 and head speed in task 2 did not correlate with neuro psych measures but number of errors was inversely correlated.</p> <p>Discriminating performance Compared using neuropsychology measure (immediate free recall) and two non-correlated but discriminating kinematic measures (hand/ head speed); Wilks Lambda showed significant changes as each was added with eventual accuracy 92.9%, sensitivity 90%, specificity 95.5%</p>	<p>No specific data. All participants completed all sessions.</p>	<p>Hand speed on task 1 : 1 Head speed on task 2 : 1 Number of errors on task 2 : 1.5</p>	<p>Not used.</p>
Zakzanis et al. [38]	<p>No data on patients. Young versus old controls:</p> <ul style="list-style-type: none"> <li>• No difference in education level/years of computer use (<math>p=0.09</math>, <math>p=0.89</math>).</li> </ul>	<p>No comparative statistics for the two patients versus controls.</p> <p>Neuropsychology measures</p> <ul style="list-style-type: none"> <li>• Patient JH was marked impairment on verbal tasks versus performance;</li> <li>• Patient SD more impaired on recall tasks.</li> </ul> <p>Spatial navigation in VR</p>	<p>Patients with AD did not complete the cybersickness questionnaire. No specific data. All participants completed all sessions.</p>	<p>Not applicable</p>	<p>No data on patients. Cybersickness measures for controls showed significant increase during versus after immersion (<math>r=0.65</math>, <math>p=0.01</math>)</p>

(Continued)

Table 2a  
(Continued)

Study	Between groups differences	Outcome measures	Acceptability	Effect sizes	Subjective measures
	<ul style="list-style-type: none"> <li>• Significant difference in years of experience with video games/age (<math>p &lt; 0.05</math>, <math>p &lt; 0.01</math>).</li> </ul>	<ul style="list-style-type: none"> <li>• JH: performance was within one standard deviation of older controls for all tasks</li> <li>• SD: more than 2-3 standard deviations worse for all tasks.</li> </ul>			
Howett et al. [34]	<ul style="list-style-type: none"> <li>• Nil significant differences age, gender, educational level.</li> <li>• Significant difference in ACE-R (<math>p &lt; 0.05</math>) and other cognitive assessments between controls and patients with MCI (<math>p &lt; 0.05</math>)</li> </ul>	<p>Absolute distance error increased for patients with MCI compared to controls (57.33 cm, <math>t(1, 107)=3.24</math>, <math>p &lt; 0.01</math>). Increased errors for biomarker positive patients versus biomarker negative patients (97.56 cm, <math>t, 163, 4.69</math>, <math>p &lt; 0.001</math>). Lower ACE-R scores correlated with greater distance errors (<math>t(1,85)=2.89</math>, <math>p &lt; 0.01</math>).</p> <p>Significant negative associations between absolute distance error and total entorhinal cortex volume on MRI (<math>F(1,64)=9.60</math>, <math>p &lt; 0.005</math>). Area under the curve for absolute distance error MCI biomarker+ve versus -ve 0.9 (CI 0.5-1) compared to ACE-R (0.53, CI 0.24-0.73) and other cognitive assessments (e.g., Trail making task B, 0.57, 0.22-0.69, 4MT 0.56, CI 0.2-0.72).</p>	No specific data. All participants completed all sessions.	Not applicable	Not used.

4MT, 4 Mountains Test; ACE-R, Addenbrooke's Cognitive Examination Revised; AD, Alzheimer's disease; AOR, Abnormal Objects Recognition Test; BDTT, Bot Doctor Turing Test; DLT, Dichotic Listening Test; MANOVA, multivariate analysis of variance; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; MRI, magnetic resonance imaging; VAT, Visual Association Test; VDLT, Virtual Daily Living Test; VOM, Virtual Objects Memorization; VR, virtual reality; VRS, Virtual vs. Real Sounds

that neuropsychology scores remained stable over a 28-week period following treatment. Their participant showed sustained learning on the navigation task. Subjective improvements were noted including a reduction in self-deprecating comments. There was no specific data on acceptability.

#### *Optale et al. [41]*

This case study with a single participant with AD monitored the effect of practice of functional real-world tasks in iVR using HMD controlled by joystick. No details were given about diagnostic criteria used. There was no control task. Training was

completed three times a week over 12 weeks. Results showed that neuropsychology scores over a 28-week period following treatment and subjective improvement. There was no specific data on acceptability.

#### *Micarelli et al. [42]*

This case control study examined the effect of a VR racing game, steered via head movement, on vestibular rehabilitation. Participants with unilateral vestibular hypoactivity (UVH,  $n=47$ ) comprising older adults and MCI were further divided into treatment (older adults  $n=12$ , MCI  $n=11$ ) and control groups (older adults and MCI  $n=12$ ). All groups

Table 2b  
Results of Treatment Studies

Study	Between groups differences	Outcome measures	Acceptability	Effect sizes	Subjective measure	Validation measures
Bourellier et al. [39]	<ul style="list-style-type: none"> <li>No difference in age/educational level.</li> <li>Significant difference in MMSE (<math>p &lt; 0.001</math>) and TUG test (<math>p = 0.005</math>).</li> </ul>	<p>2*2 ANOVA Group (MCI, Control) and Session (Implicit/VR versus Explicit/Physio)</p> <ul style="list-style-type: none"> <li>Session: F (1,20)=30.70, <math>p &lt; 0.001</math></li> <li>Group: F (1, 20)=16.47, <math>p &lt; 0.001</math></li> <li>Interaction: F (1, 20)=6.67, <math>p = 0.01</math>, MCI subjects sig fewer appropriate movements in implicit session (<math>p &lt; 0.001</math>), no difference in explicit session.</li> </ul>	Semi-structured interview of 30 min in a quiet room. Patients reported enjoying iVR more than the explicit physiotherapy session which they found repetitive.	Not calculated.	Not assessed quantitatively, subjects reported enjoying VR tasks but some found 3D glasses unwieldy and preferred therapist presence (explicit task)	Some content validity, construct validity not calculated.
White and Moussavi. [40]	Not applicable to case study	Average number of navigation errors and wall errors both improved over time (R2 0.463 and 0.458, respectively), MoCA scores stable (23-26 over 28-week period). No other statistics calculated.	Subjective reduction in self-deprecating comments, no other data reported.	Not applicable	Patient and his wife reported more confidence driving and improved mood	Not used.
Optale et al. [41]	Not applicable to case study	No statistics calculated. Neuropsychology measures stable over 28-week period. No changes reported in serial EEG, CT, MRI, laboratory tests (no further data on which tests or interval).	No specific data.	Not applicable	Patient reports improvements in word finding, sleep.	Not used.

(Continued)

Table 2b  
(Continued)

Study	Between groups differences	Outcome measures	Acceptability	Effect sizes	Subjective measure	Validation measures
Micarelli et al. [42]	Lower mean MMSE for MCI versus controls (25.8 versus 28). Age, gender, educational level noted but not formally compared	VOR MCI HMD versus Control ( $p=0.0008$ ), Otoneurological measures Power Spectra low frequency (most affected by vestibular issues) effect sizes eyes closed: 0.34, 0.27, eyes open 1.14, 0.97 Positive correlation between MMSE & pre-post differences in low frequency power spectra ( $r=0.72$ ) and DGI ( $r=0.76$ )	Not measured directly. Subjective measures of nausea and disorientation were reduced in the treatment group over time and compared to controls.	Not calculated:	Cohens d calculated for MCI HMD versus control Dizziness Handicap Inventory 1.78, DGI 0.29, ABC 0.25	Nothing formal, construct and face validity evident.

ABC, Activities-specific Balance Confidence Scale; ANOVA, analysis of variance; CT, computerized tomography; DGI, Dynamic Gait Index; EEG, electroencephalogram; HMD, head mounted display; iVR, Immersive virtual reality; TUG, timed up and go task; VOR, vestibulo-ocular reflex. See Table 2a for remaining abbreviations.

received the same vestibular rehabilitation program for 40 min daily but the treatment group additionally spent 20 min per day playing the VR racing game. Participants in the treatment condition (including those with MCI) exhibited greater improvement in vestibular function than those in the non-treatment condition. There was no alternative treatment to VR for the control group which led to a high risk of performance bias. Acceptability was not measured directly but subjective measures of nausea and disorientation were reduced in the treatment group compared to controls.

## DISCUSSION

The breadth of studies included in this review demonstrate the rich opportunity available to researchers by iVR technology. These range from simulating IADLs such as catching a bus and cash machine withdrawals [37], to balance retraining [42] and assessment of AD-specific navigation deficits [34]. iVR offers distinct advantages for research and therapy reliant on movement and engagement in physical space. Where they were asked, patient groups reported enjoying iVR and this may contribute

to therapeutic advantage [39, 42]. iVR also has the potential advantage of rich and meaningful outcome measures such as IADLs and navigation.

iVR equipment continues to improve and become more affordable. Several of the studies included use off the shelf commercial iVR equipment [36, 42]. Earlier acceptability concerns around motion sickness and related side effects were not problematic in any of the studies. Even when tasks included vestibular rehabilitation all groups tolerated iVR without incident [42]. Some studies did mention the need to adjust equipment according to hair/visual acuity and mobility [37, 42]. This must be considered with any future technology developments if these are to be used with older people. Other potential limiting factors on the clinical utility of iVR are the tendency of HMD equipment to be relatively cumbersome and heavy, lack of standardized programs and difficulty implementing equipment in a healthcare setting.

*iVR can be used to discriminate people with AD from controls/other dementia processes*

Howett et al. [34] designed a navigation task to test the function of the entorhinal cortex. This task differentiated participants with biomarker positive MCI

Table 3  
Risk of Bias

Author	Selection Bias		Performance Bias		Detection Bias		Attrition Bias		Reporting Bias		Other Bias		Total Score /6
	Risk	Information	Risk	Information	Risk	Information	Risk	Information	Risk	Information	Risk	Information	
<b>Assessment Studies</b>													
Fernandez-Montenegro et al. [35]	High	Convenience sampling not explained further	High	Same tasks for all participants but order not explained	High	No blinding of assessors attempted	High	More unwell patients were unable to complete the task	High	Selective reporting of significant/interesting results	High	Limited matching of controls/confounders	6
Gago et al. [36]	Low	Random clinic patients	High	Unblinded researcher triggers postural disturbance	High	No blinding of assessors attempted	Low	No attrition over the session	Low	Outcomes relevant to initial question: unclear AD criteria, incomplete statistics	High	Non-random triggering of translation, arbitrary definition of fallers versus non-fallers	3
Seo et al. [37]	Low	Random clinic patients	Low	Same tasks for all participants in a random order.	High	No blinding of assessors attempted	Low	No attrition over the session	Low	All outcomes are reported	Low		1
Zakzamis et al. [38]	High	Convenience sampling not explained further	High	Controls & patients completed different neuropsych measures	High	No blinding of assessors attempted	Low	No attrition over the session	High	Very limited reporting of results for patient group	High	Small number of patients with lower average level of education versus controls.	5
Howett et al. [34]	Low	Random clinic patients	Low	Researchers blinded	Low	Researchers blinded	Low	No attrition over the session	Low	All anticipated outcomes are reported	Low		0
<b>Treatment Studies</b>													
Bourellier et al. [39]	High	Convenience sampling not explained further	Low	Same tasks for all participants in a random order.	High	No blinding of assessors attempted	Low	No attrition over the session	Low	All outcomes are reported	Low		2
Miracelli et al. [42]	Low	Group assignment was random balanced for lesion side, age, sex,	High	No additive control group. VR participants received more treatment	Low	Oto neurologist/physiotherapist blinded	Low	No attrition over the sessions	Low	Stats and control for multiple comparisons (Bonferroni) reported, although not the Bonferroni adjusted alpha	Low		1



from those with biomarker negative MCI and controls. The task was critically dependent on iVR. The integration of multisensory cues generated through physical movement and head rotation was pivotal to the success of this task. Other studies were able to discriminate participants from controls but used complex tasks which were not hypothesis driven. In these studies, it is difficult to determine whether iVR or other aspects of the tasks were important. For example, the two tasks studied by Seo et al. [37] were very well suited to iVR and all participants were able to interact in a meaningful way with the virtual environment (boarding the correct bus or use a cash machine with a specific pin code). Spatial and episodic memory-based elements of the task were not separately measured. As a result, it is unclear which aspects of the task discriminated participants with AD from controls.

Similar issues affected other studies where iVR tasks were designed as proofs of concept [35, 36, 38]. This led to tasks not being assessed as independent screening measures [38] or the use of iVR simply making tasks more confusing for patient participants [35]. Gago et al. used the same task to distinguish participants with AD [36] and Parkinson's disease [43] from controls. This measure was well tolerated by participants but may simply select out any patient at high risk of falls. The common problem for these studies was lack of focus in the design on AD-specific neuropathology. To aid the diagnosis and treatment of AD, iVR tasks should adopt a hypothesis driven approach targeting specific cognitive processes that are dependent upon neuroanatomical structures compromised in early AD pathology.

*There is limited evidence that iVR can contribute to improvements in real-world function*

Despite lack of significant therapeutic benefit, participants with AD and MCI reported enjoying treatment tasks. These ranged from fruit picking as an alternative to standard physiotherapy [39] to car racing for vestibular rehabilitation [42] and simulated navigation tasks [40, 41]. Increased engagement should be considered a positive factor in designing training system for participants with cognitive deficits. On some tasks, overall performance and learning were also improved compared to controls tasks [42]. This suggests that iVR has the potential to engage vulnerable groups meaningful in rehabilitation tasks which are otherwise very difficult to sustain.

There was a lack of longitudinal studies. Case studies focused on cognitive training in AD [40, 41] noted high patient engagement and subjective benefit for the participants with cognitive measures remaining stable over a significant period (28 weeks for both) and improved performance shown on a spatial cognition task [40]. It will be important to complete larger longitudinal studies to address longer term response to treatment. Methodological issues which should be anticipated include ensuring consistent supervision of iVR sessions and measures to reduce risk of attrition/assess reasons for this which might relate to tolerability.

### *Strengths*

The present review is the first to focus specifically on iVR in AD. We have developed clear inclusion and exclusion criteria not utilized in previous work [21]. This makes our definitions of technology reliable and reproducible. We used a rigorous and systematic search strategy which was supported by a specialist research librarian (IK). This leads to very low risk of missing relevant studies or reviews. Our authors are from clinical and clinical research backgrounds. As a result, we are able to offer a meaningful appraisal of the uses and challenges of iVR in real clinical practice.

### *Limitations*

The search strategy for the current review only searched major databases. Any non-academic/games industry/grey literature sources will have been omitted. This is a very new and rapidly developing area of research and it is possible that new terms/journals and technologies will have developed which are not indexed by major databases.

We were unable to answer more detailed questions around the evidence base for iVR in AD treatment due to sparsity of evidence. Small studies and variable outcome measures suggest a field in its infancy and limit the generalizability of conclusions. These issues also preclude the meaningful use of meta-analysis and increases the risk of bias.

### *Focus for future research*

Future tasks implemented in iVR should target cognitive tasks known to be affected by neurodegenerative processes. Association of biomarker positive MCI with navigation deficits [34] suggests

several outstanding research questions. Firstly, do participants identified at high risk of AD actually progress to AD diagnoses? Secondly, do participants with positive AD biomarkers but without cognitive deficits demonstrate navigation deficits specific to the entorhinal cortex? If deficits are noted, this offers the potential for identifying targets for therapy at earlier stages than currently possible. It may also be feasible to complete functional imaging during task performance [44] allowing more direct comparison with animal models. Indeed, there is already a literature emerging in which human and animal iVR performance is compared in this manner which may contribute to increased power to distinguish useful new drug treatments at an animal stage of trials [45].

By linking iVR training studies to existing clinical cohorts it will be possible to determine much more precisely if these techniques have additional benefit beyond conventional measures. The PROTECT program has already demonstrated that practice on abstract reasoning tasks can lead to differential improvement in overall cognitive function for older participants [46]. These generalized effects appear stronger in individuals with cognitive impairment and we might therefore expect iVR spatial navigation training to have particular benefits for less mobile elderly people with MCI/AD who suffer impairments both in their physical ability to mobilize and the cognitive systems which control these processes. Additional benefit could also be seen from errorless learning and empowerment techniques [47].

An ethical concern around the use of iVR in AD is whether users may become confused about whether they are inhabiting the real or a virtual world. This may present a challenge to any users given sufficiently sophisticated technology. People with dementia would be expected to be at earlier risk. This could potentially lead to increased distress or even, conversely, reluctance to leave pleasant or safe therapeutic iVR environments for the confusion and stress of reality. There is little evidence yet in the field of these tensions. Reasons why participants were unable to participate in studies should be explored in greater detail [35]. Limited duration of iVR use and rigorous testing of software on cognitively intact participants before use in patient groups will help. Clear standards over the therapeutic focus of technology will also reduce risk, but this is an area in need of regular review as opportunities within the field expand.

As experience with iVR develops, it may become possible to augment specific memories or skills in the

real world through simulated practice. This could be in a similar manner to Occupational Therapy (OT) practice in a hospital kitchen. OT sessions clarify deficits and helps with function when the patient returns home. Already this area is being explored with retraining cooking skills in iVR [48, 49]. Augmented and mixed reality technology will allow more seamless integration. We may be also able to tailor systems to individual users' impairments by integrating familiar people/places into virtual worlds to reduce cognitive stress and reducing risk of behavioral and psychological symptoms of dementia in a similar manner to shared reminiscence therapy [50]. Early work shows such treatment is tolerated by participants but benefits are not yet clear [26].

Other preliminary research suggests residential home residents with dementia reported enjoying shared iVR sessions and that these also reduced apathy [51]. Further controlled trials are underway. Finally, it is also feasible to imagine iVR being used for assessment of high-risk tasks such as driving [52] and to help maintain interest in spatial activities when driving is no longer safe.

### Summary

This systematic review shows the field of iVR for the assessment and treatment of AD is starting to step out its infancy. Participants found iVR tasks demanding but were able to participate and reported enjoying the tasks and interfaces [35, 39]. The intuitive nature of the technology makes it very appealing even to those with a significant burden of disease. Incorporating measures of spatial cognition within iVR assessment has already demonstrated remarkable discriminative potential for detecting preclinical AD [34].

It is less clear whether iVR can offer any specific cognitive benefits, but potential has been demonstrated to use iVR to facilitate broader physical and vestibular rehabilitation [39, 42]. As patients with AD often find it difficult to engage with repetitive activities as part of physical therapy, it is useful to note that iVR can facilitate such treatment. It is important to ensure that future use of this technology is theoretically grounded and controlled to avoid potential harm to vulnerable patient groups. Collaboration between clinical researchers and software engineering teams is crucial to build on the early potential demonstrated.

## ACKNOWLEDGMENTS

Our thanks to Cambridge University Medical School Librarian Isla Kuhn for crucial support with the search strategy. This is a summary of research funded by the National Institute for Health Research (NIHR) Applied Research Collaboration East of England. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Authors' disclosures available online (<https://www.j-alz.com/manuscript-disclosures/19-1218r1>).

## REFERENCES

- [1] Prince M, Bryce R, Ferri C (2011) *World Alzheimer Report 2011: The benefits of early diagnosis and intervention*. Alzheimer's Disease International, London, UK.
- [2] Cogné M, Taillade M, N'Kaoua B, Tarruella A, Klinger E, Larue F, Sauzéon H, Joseph PA, Sorita E (2017) The contribution of virtual reality to the diagnosis of spatial navigation disorders and to the study of the role of navigational aids: A systematic literature review. *Ann Phys Rehabil Med* **60**, 164–176.
- [3] Alzheimer's Research UK (2019) *A Walk Through Dementia*. <https://www.alzheimersresearchuk.org/campaigns/awtd/>. Accessed 28 May 2019.
- [4] Johnson L (2018) For people with dementia, virtual reality can be life-changing. *Wired*, <https://www.wired.co.uk/article/virtual-reality-dementia-technology>. Posted 21 May 2018, Accessed 9 July 2019.
- [5] McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser MB (2006) Path integration and the neural basis of the "cognitive map." *Nat Rev Neurosci* **7**, 663-678.
- [6] Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, Slater M (2017) Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychol Med* **47**, 2393-2400.
- [7] Slobounov SM, Ray W, Johnson B, Slobounov E, Newell KM (2015) Modulation of cortical activity in 2D versus 3D virtual reality environments: an EEG study. *Int J Psychophysiol* **95**, 254-260.
- [8] Yew B, Alladi S, Shailaja M, Hodges JR, Hornberger M (2013) Lost and forgotten? Orientation versus memory in Alzheimer's disease and frontotemporal dementia. *J Alzheimers Dis* **33**, 473-481.
- [9] Faraji J, Soltanpour N, Moeeni R, Roudaki S, Soltanpour N, Abdollahi AA, Metz GAS (2014) Topographical disorientation after ischemic mini infarct in the dorsal hippocampus: Whispers in silence. *Front Behav Neurosci* **8**, 216.
- [10] Widmann CN, Beinhoff U, Riepe MW (2012) Everyday memory deficits in very mild Alzheimer's disease. *Neurobiol Aging* **33**, 297-303.
- [11] Allison SL, Fagan AM, Morris JC, Head D (2016) Spatial navigation in preclinical Alzheimer's disease. *J Alzheimers Dis* **52**, 77-90.
- [12] Moodley K, Minati L, Contarino V, Prioni S, Wood R, Cooper R, D'Incerti L, Tagliavini F, Chan D (2015) Diagnostic differentiation of mild cognitive impairment due to Alzheimer's disease using a hippocampus-dependent test of spatial memory. *Hippocampus* **25**, 939-951.
- [13] Mokrisova I, Laczó J, Andel R, Gazova I, Vyhnalek M, Nedelska Z, Levcik D, Cerman J, Vlcek K, Hort J (2016) Real-space path integration is impaired in Alzheimer's disease and mild cognitive impairment. *Behav Brain Res* **307**, 150-158.
- [14] Vlček K, Laczó J (2014) Neural correlates of spatial navigation changes in mild cognitive impairment and Alzheimer's disease. *Front Behav Neurosci* **8**, 89.
- [15] Dejos M, Sauzeon H, N'Kaoua B (2012) Virtual reality for clinical assessment of elderly people: Early screening for dementia. *Rev Neurol (Paris)* **168**, 404-414.
- [16] Abichou K, La Corte V, Piolino P (2017) Does virtual reality have a future for the study of episodic memory in aging? *Geriatr Psychol Neuropsychiatr Vieil* **15**, 65-74.
- [17] García-Betances RI, Arredondo Waldmeyer MT, Fico G, Cabrera-Umpiérrez MF (2015) A succinct overview of virtual reality technology use in Alzheimer's disease. *Front Aging Neurosci* **7**, 80.
- [18] García-Casal JA, Loizeau A, Csipke E, Franco-Martín M, Perea-Bartolomé MV, Orrell M (2017) Computer-based cognitive interventions for people living with dementia: a systematic literature review and meta-analysis. *Aging Ment Health* **21**, 454-467.
- [19] D'Cunha NM, Nguyen D, Naumovski N, McKune AJ, Kellest J, Georgouspoulou EN, Frost J, Isabel S (2019) A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. *Gerontology* **65**, 430-440.
- [20] Hill NTM, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A (2017) Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. *Am J Psychiatry* **174**, 329-340.
- [21] Kim O, Pang Y, Kim JH (2019) The effectiveness of virtual reality for people with mild cognitive impairment or dementia: A meta-analysis. *BMC Psychiatry* **19**, 219.
- [22] Moreno A, Wall KJ, Thangavelu K, Craven L, Ward E, Dissanayaka NN (2019) A systematic review of the use of virtual reality and its effects on cognition in individuals with neurocognitive disorders. *Alzheimers Dement (N Y)* **5**, 834-850.
- [23] Woods B, O'Philbin L, Farrell EM, Spector AE, Orrell M (2018) Reminiscence therapy for dementia. *Cochrane Database Syst Rev* **3**, CD001120
- [24] Rose V, Stewart I, Jenkins KG, Ang CS, Matsangidou M (2018) A scoping review exploring the feasibility of virtual reality technology use with individuals living with dementia. *ICAT-EGVE 2018*, pp. 131-139.
- [25] Manera V, Chapoulie E, Bourgeois J, Guerchouche R, David R, Ondrej J, Drettakis G, Robert P (2016) A feasibility study with image-based rendered virtual reality in patients with mild cognitive impairment and dementia. *PLoS One* **11**, e0151487.
- [26] Rose V, Stewart I, Jenkins KG, Tabbaa L, Ang CS, Matsangidou M (2019) Bringing the outside in: The feasibility of virtual reality with people with dementia in an inpatient psychiatric care setting. *Dementia (London)*, doi: 10.1177/1471301219868036.
- [27] Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Med* **6**, e1000097.
- [28] Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, Stewart L (2012) The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev* **1**, 2.

- [29] Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M (2017) Virtual reality for stroke rehabilitation. *Cochrane Database Syst Rev* **11**, CD008349.
- [30] Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A (2016) Rayyan—a web and mobile app for systematic reviews. *Syst Rev* **5**, 210.
- [31] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (2019) *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0. Cochrane, <http://www.training.cochrane.org/handbook>.
- [32] Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JAC (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**, d5928.
- [33] Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, Carpenter JR, Chan A-W, Churchill R, Deeks JJ, Hróbjartsson A, Kirkham J, Jüni P, Loke YK, Pigott TD, Ramsay CR, Regidor D, Rothstein HR, Sandhu L, Santaguida PL, Schünemann HJ, Shea B, Shrier I, Tugwell P, Turner L, Valentine JC, Waddington H, Waters E, Wells GA, Whiting PF, Higgins JP (2016) ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* **355**, i4919.
- [34] Howett D, Castegnaro A, Krzywicka K, Hagman J, Marchment D, Henson R, Rio M, King JA, Burgess N, Chan D (2019) Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. *Brain* **142**, 1751-1766.
- [35] Fernandez Montenegro JM, Argyriou V (2017) Cognitive evaluation for the diagnosis of Alzheimer's disease based on Turing Test and virtual environments. *Physiol Behav* **173**, 42-51.
- [36] Gago MF, Yelshyna D, Bicho E, Silva HDHD, Rocha LL, Lurdes Rodrigues M, Sousa N (2016) Compensatory postural adjustments in an oculus virtual reality environment and the risk of falling in Alzheimer's disease. *Dement Geriatr Cogn Dis Extra* **6**, 252-267.
- [37] Seo K, Kim J-K, Oh DH, Ryu H, Choi H (2017) Virtual daily living test to screen for mild cognitive impairment using kinematic movement analysis. *PLoS One* **12**, e0181883.
- [38] Zakzanis KK, Quintin G, Graham SJ, Mraz R (2009) Age and dementia related differences in spatial navigation within an immersive virtual environment. *Med Sci Monit* **15**, CR140-150.
- [39] Bourrelier J, Ryard J, Dion M, Merienne F, Manckoundia P, Mourey F (2016) Use of a virtual environment to engage motor and postural abilities in elderly subjects with and without mild cognitive impairment (MAAMI Project). *IRBM* **37**, 75-80.
- [40] White PJF, Moussavi Z (2016) Neuro-cognitive treatment for an Alzheimer's patient using a virtual reality navigational environment. *J Exp Neurosci* **10**, 129-135.
- [41] Optale G, Capodiecici S, Pinelli P, Zara D, Gamberini L, Riva G (2001) Music-enhanced immersive virtual reality in the rehabilitation of memory-related cognitive processes and functional abilities: A case report. *Presence Teleoperators Virtual Environ* **10**, 450-462.
- [42] Micarelli A, Viziano A, Micarelli B, Augimeri I, Alessandrini M (2019) Vestibular rehabilitation in older adults with and without mild cognitive impairment: Effects of virtual reality using a head-mounted display. *Arch Gerontol Geriatr* **83**, 246-256.
- [43] Yelshyna D, Gago MF, Bicho E, Fernandes V, Gago NF, Costa L, Silva H, Rodrigues ML, Rocha L, Sousa N (2016) Compensatory postural adjustments in Parkinson's disease assessed via a virtual reality environment. *Behav Brain Res* **296**, 384-392.
- [44] Reggente N, Essoe JK-YKY, Aghajan ZM, Tavakoli AV, McGuire JF, Suthana NA, Rissman J (2018) Enhancing the ecological validity of fMRI memory research using virtual reality. *Front Neurosci* **12**, 1-9.
- [45] Dolins FL1, Klimowicz C, Kelley J, Menzel CR (2016) Cognitive abilities in chimpanzees and humans. *Am J Primatol* **76**, 496-513.
- [46] Corbett A, Owen A, Hampshire A, Grahn J, Stenton R, Dajani S, Burns A, Howard R, Williams N, Williams G, Ballard C (2015) The effect of an online cognitive training package in healthy older adults: an online randomized controlled trial. *J Am Med Dir Assoc* **16**, 990-997.
- [47] Perez-Marcos D, Bieler-Aeschlimann M, Serino A (2018) Virtual reality as a vehicle to empower motor-cognitive neurorehabilitation. *Front Psychol* **9**, 2120.
- [48] Follope DA, Richard P, Yamaguchi T, Etcharry-Bouyx F, Allain P (2018) The potential of virtual reality-based training to enhance the functional autonomy of Alzheimer's disease patients in cooking activities: A single case study. *Neuropsychol Rehabil* **28**, 709-733.
- [49] Davis R, Ohman JM, Weisbeck C (2017) Salient cues and wayfinding in Alzheimer's disease within a virtual senior residence. *Environ Behav* **49**, 1038-1065.
- [50] Burdea G, Polistico K, Krishnamoorthy A, House G, Rethage D, Hundal J, Damiani F, Pollack S (2015) Feasibility study of the BrightBRAIN integrative cognitive rehabilitation system for elderly with dementia. *Disabil Rehabil Assist Technol* **10**, 421-432.
- [51] Brimelow RE, Dawe B, Dissanayaka N (2020) Preliminary research: Virtual reality in residential aged care to reduce apathy and improve mood. *Cyberpsychol Behav Soc Netw* **23**, 165-170.
- [52] Bennett CR, Corey RR, Giudice U, Giudice NA (2016) Immersive virtual reality simulation as a tool for aging and driving research. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, pp. 377-385.

**ADDITIONAL MATERIALS**

*Picture 1. 2D display of virtual environments*



Copyright free image from Humanities Commons (<http://hcommons.org>)

*Picture 2. HMD iVR*



By ESA, image is for free distribution under Creative Commons Share Alike Licence (<https://creativecommons.org/licenses/by-sa/3.0-igo>)

*Picture 3. CAVE System iVR*



Copyright free image from Wikimedia Commons (<http://commons.wikimedia.org>)