Dementia and Geriatric Cognitive Disorders Extra

Systematic Review

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744

Received: February 25, 2024
Accepted: June 7, 2024
Published online: June 10, 2024

# What's New in Dementia Risk Prediction Modelling? An Updated Systematic Review

Jacob Brain[a, b] Aysegul Humeyra Kafadar[a] Linda Errington[c] Rachael Kirkley[d] Eugene Y.H. Tang[d] Ralph K. Akyea[e] Manpreet Bains[f] Carol Brayne[g] Grazziela Figueredo[h] Leanne Greene[i] Jennie Louise[j] Catharine Morgan[k] Eduwin Pakpahan[l] David Reeves[m] Louise Robinson[d] Amy Salter[n] Mario Siervo[o, p] Phillip J. Tully[b, q] Deborah Turnbull[b] Nadeem Qureshi[e] Blossom C.M. Stephan[a, p]

[a]Institute of Mental Health, School of Medicine, University of Nottingham, Innovation Park, Jubilee Campus, Nottingham, UK; [b]Freemasons Foundation Centre for Men's Health, Discipline of Medicine, School of Psychology, The University of Adelaide, Adelaide, SA, Australia; [c]Walton Library, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; [d]Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK; [e]PRISM Group, Centre for Academic Primary Care, Lifespan and Population Health, School of Medicine, University of Nottingham, Nottingham, UK; [f]Nottingham Centre for Public Health and Epidemiology, Lifespan and Population Health, School of Medicine, University of Nottingham, Nottingham, UK; [g]Cambridge Public Health, University of Cambridge, Cambridge, UK; [h]School of Computer Science, University of Nottingham, Nottingham, UK; [i]Exeter Clinical Trials Unit, Department of Health and Community Sciences, University of Exeter Medical School, Exeter, UK; [j]Women's and Children's Hospital Research Centre and South Australian Health and Medical Research Institute, Adelaide, SA, Australia; [k]Division of Population Health, Health Services Research and Primary Care, University of Manchester, Manchester, UK; [l]Department of Mathematics, Physics and Electrical Engineering, Northumbria University, Newcastle upon Tyne, UK; [m]School for Health Sciences, University of Manchester, Manchester, UK; [n]School of Public Health, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, SA, Australia; [o]School of Population Health, Curtin University, Perth, WA, Australia; [p]Dementia Centre of Excellence, Curtin enAble Institute, Faculty of Health Sciences, Curtin University, Perth, WA, Australia; [q]Faculty of Medicine and Health, School of Psychology, University of New England, Armidale, NSW, Australia

**Keywords**
Risk prediction · Dementia · Alzheimer disease · Incidence · Statistical model

**Abstract**
***Introduction:*** Identifying individuals at high risk of dementia is critical to optimized clinical care, formulating effective preventative strategies, and determining eligibility for clinical trials. Since our previous systematic reviews in 2010 and 2015, there has been a surge in dementia risk prediction

karger@karger.com
www.karger.com/dee

Karger
OPEN ACCESS



Correspondence to:
Jacob Brain, jacob.brain@nottingham.ac.uk

modelling. The aim of this study was to update our previous reviews to explore, and critically review, new developments in dementia risk modelling. ***Methods:*** MEDLINE, Embase, Scopus, and Web of Science were searched from March 2014 to June 2022. Studies were included if they were population- or community-based cohorts (including electronic health record data), had developed a model for predicting late-life incident dementia, and included model performance indices such as discrimination, calibration, or external validation. ***Results:*** In total, 9,209 articles were identified from the electronic search, of which 74 met the inclusion criteria. We found a substantial increase in the number of new models published from 2014 (>50 new models), including an increase in the number of models developed using machine learning. Over 450 unique predictor (component) variables have been tested. Nineteen studies (26%) undertook external validation of newly developed or existing models, with mixed results. For the first time, models have also been developed in low- and middle-income countries (LMICs) and others validated in racial and ethnic minority groups. ***Conclusion:*** The literature on dementia risk prediction modelling is rapidly evolving with new analytical developments and testing in LMICs. However, it is still challenging to make recommendations about which one model is the most suitable for routine use in a clinical setting. There is an urgent need to develop a suitable, robust, validated risk prediction model in the general population that can be widely implemented in clinical practice to improve dementia prevention.



## Introduction

As strategies to tackle the global burden of disease associated with dementia move towards a greater focus on primary prevention, in particular risk reduction, the identification of individuals at high risk has become extremely important. Such an approach typically combines known risk factors into an algorithm to calculate the probability of an individual developing incident disease. Such strategies work well in areas such as cardiovascular disease [1, 2] and stroke [3, 4] prevention. Despite dementia being a leading cause of disability and death on a global scale [5], there is no recommended tool to identify individuals at increased risk who might benefit from early intervention to delay or prevent it.

While screening for dementia risk is not currently recommended [6, 7], numerous prediction models have been developed. Indeed, previous systematic reviews [8–10] have identified over 100 different models, largely developed in high-income White populations with variable discriminative performance (c-statistic range 0.49–0.89) and over a range of follow-up (1 year to >20 years). Where models have been externally validated the results are mixed [8, 9] and insufficiently reported [11]. New models are constantly being developed, including some that take advantage of more advanced computational techniques such as artificial intelligence (AI), in particular, machine learning (ML) algorithms.

Therefore, in this systematic review, we provide an update, covering all literature published after our 2015 review [9] on current developments in dementia risk prediction modelling. We aim to (i) synthesize the literature on all new dementia risk prediction models including reporting their discriminative accuracy and calibration where available, and external validation results; (ii) identify key risk factors for dementia incorporated into the different models; and (iii) by combining the results with our two previous reviews [8, 9] to make policy recommendations on dementia risk prediction. These recommendations are necessary to facilitate and coordinate ongoing and future dementia preventative work.

## Methods

The protocol was registered on PROSPERO (Reference CRD42022320630). The systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (online suppl. material 1; for all online suppl. material, see https://doi.org/10.1159/000539744) [12].

### *Search Strategy*

Embase (via Ovid), Medline (via Ovid), Scopus, and Web of Science were searched from the final date of our previous review (March 14, 2014) to the June 10, 2022. The following concepts were included in the search strategy and mapped to Medical Subject Headings (MeSH) where appropriate: dementia, Alzheimer disease, predict, develop, incident, sensitivity, specificity, ROC, area under the curve (AUC), and concordance statistic (online suppl. material 2). The electronic search results were transferred to Endnote software and de-duplicated [13]. The reference lists of all included articles were manually checked to ensure relevant studies were not missed in the electronic search.

### *Inclusion Criteria*

Studies were selected for inclusion based on three criteria. Firstly, the study sample had to be population- or community-based (including electronic health record datasets) and examine risk of incident dementia in later life (i.e., at ≥60 years of age). Cross-sectional, case-control, trials, and clinical-based studies were excluded. There was no restriction on baseline age or follow-up time, provided that the diagnosis of dementia was made after the age of 60 years. We therefore excluded publications focused

exclusively on early onset dementia as this is rare and typically has a different risk and disease profile compared to late-onset dementia [14]. Secondly, a risk model had to be reported along with performance indices (e.g., discriminative accuracy measured using the AUC or c-statistic [concordance statistic]). We used the following classification system to the interpret AUC/c-statistic: <0.70 = poor; ≤0.70 to <0.80 = fair; ≤0.80 to <0.90 = good; ≤0.90 to <1.00 = excellent [15, 16]. Thirdly, the outcome had to be dementia which included all-cause and its subtypes such as Alzheimer's disease and vascular dementia (VaD) assessed using standardised criteria (e.g., Diagnostic and Statistical Manual of Mental Disorders: DSM). Studies where the outcome was a combined cognitive group, e.g., mild cognitive impairment cases combined with dementia cases, were excluded. Unlike our previous reviews, no restriction was made based on language.

*Selection of Studies*

The selection of articles followed the same protocol as our previous reviews [8, 9]. Firstly, four authors (J.B., A.H.K., R.K., B.C.M.S.) independently screened titles and abstracts for relevant articles based on the inclusion criteria. This was conducted using Rayyan software [17]. Secondly, the full text of any identified articles was checked independently by four authors (J.B., A.H.K., R.K., B.C.M.S.). Where duplicate studies were identified, details of all novel risk models and their performance indices (AUC/c-statistic and calibration statistics) were extracted from each publication and reported separately. Disagreements were resolved through discussion.

*Risk of Bias*

Two authors (J.B., L.G.) assessed the risk of bias of included articles using a modified version of the Newcastle-Ottawa Scale for non-randomized studies [18], the same adapted tool used in our previous review [9]. Excluding items that describe non-intervention cohorts, studies could achieve a maximum of six stars (compared to the original nine) based on selection, comparability, and outcome criteria. All scores were cross-checked by another author (A.H.K.).

*Data Extraction*

An Excel data extraction sheet was created including the following information: author details, country, sample size, follow-up length, baseline age, outcome (e.g., all-cause and dementia subtypes; including number with dementia and diagnostic criteria used), the component (predictor) variables in each model, discriminative accuracy (AUC or c-statistic), calibration (e.g., Hosmer-Lemeshow test, Brier score, calibration slope, calibration intercept), and external validation metrics (e.g., external AUC, external calibration measures). Data were extracted independently by four authors (J.B., A.H.K., E.Y.H.T., B.C.M.S.). Any discrepancies were resolved via discussion.

*Data Synthesis*

As with our previous reviews [8, 9], a formal narrative synthesis was undertaken of all included studies. A table of all predictor variables included in each dementia risk prediction model was created to determine the key risk factors for incident dementia risk (all-cause and its subtypes). Using model performance indices (e.g., discriminative accuracy measured using the AUC and c-statistic), we created a plot for comparing the different AUC/c-statistic values of the different models, with the x-axis representing the AUC/c-statistic, and the y-axis representing the dementia risk model and study population. Different forms of c-statistic exist, and values are affected by factors such as population homogeneity and extent of censored observations, hence some caution is required when interpreting this plot. A meta-analysis was not possible due to little overlap in the nature of the predictor variables incorporated into the different risk models and large heterogeneity in populations studied, study design, and follow-up time.

## Results

In total, 9,209 papers were identified from the electronic search of which 4,376 were duplicates and therefore removed. After title/abstract screening of 4,833 records, 129 articles were identified for full-text review. From these, 51 were selected for inclusion. An additional 23 articles were identified from reviewing the reference list of included articles. Therefore, a total of 74 articles are included in the review (shown in Fig. 1).

Table 1 shows an overview of the included studies, with Online Supplementary Material 3 providing the complete data extraction. As shown, sample sizes ranged from 192 [19] to 930,395 [20], with baseline ages ranging from 18 [21] to 98 [22] years. Follow-up ranged from ~1 year [23] to >30 years [24, 25]. Most studies used data from high-income countries (HICs; $n$ = 67 studies), with a substantial proportion of these ($n$ = 18 studies) using cohorts from the USA. Six studies only drew on data from low- and middle-income countries (LMICs) including China, Cuba, the Dominican Republic, Mexico, Peru, Puerto Rico, and Venezuela [26–31]. Outcomes predominantly focused on all-cause dementia ($n$ = 61) or clinically diagnosed AD ($n$ = 19), followed by VaD ($n$ = 1) [32], mixed dementia ($n$ = 1) [33], Parkinson's disease dementia ($n$ = 1) [34], and Dementia with Lewy bodies ($n$ = 1) [23].

Of the 74 included studies, most used a cohort design ($n$ = 57), 12 used electronic health records, two used insurance claim data, and one used health survey data [55]. One study used a combination of administrative claims and electronic health records data [83], and another used cohort and electronic health records data [23]. Only 14 studies (19%) used population-representative data [27–29, 43, 45, 66, 69, 70, 82, 85, 90, 92, 99, 100]. As noted in online supplementary material 3, most studies ($n$ = 56) used traditional statistical modelling methods including Cox, logistic regression, and the Fine and Gray model. These different approaches estimate different outcomes, i.e., logistic regression estimates the odds of dementia as a binary outcome (yes/no), whilst

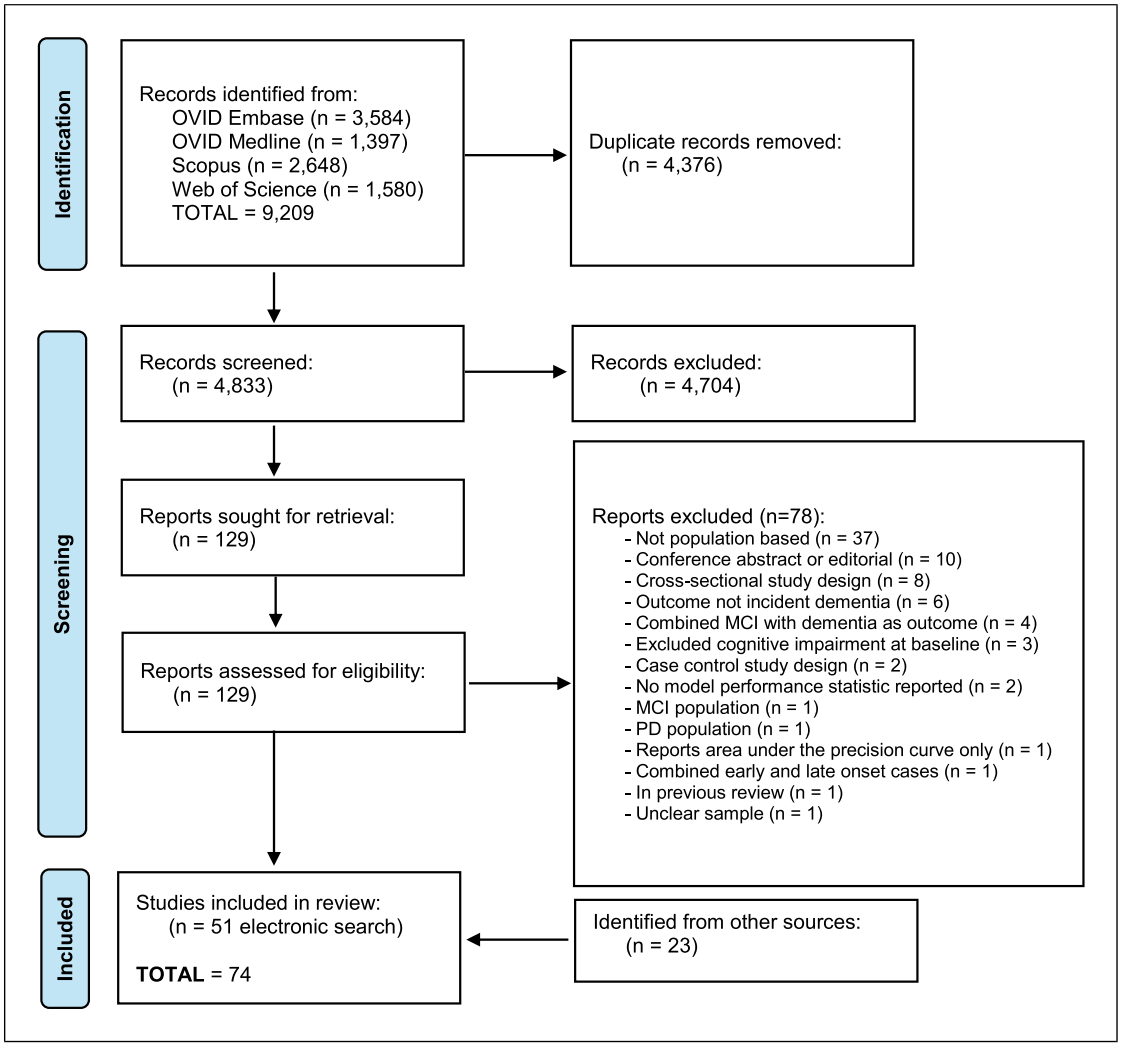



**Fig. 1.** PRISMA flow diagram showing the selection of studies for inclusion in the review study design and characteristics.

Cox estimates hazard ratios, and Fine-Gray models (competing-risk) estimate sub-distribution or cause-specific hazard ratios. Models using supervised ML and AI methods have emerged, having been absent in both our previous reviews. This includes generalised linear models like Elastic Net [108], tree-based models utilising decision trees, random forests, and XGBoost [43, 44], non-linear models, for instance, long short-term memory networks [66, 97], as well as combined approaches ML/AI [19, 24, 28, 34, 37, 56, 60, 76, 77, 83, 86, 90, 96].

*Risk of Bias*

In total, 50 articles scored a maximum of six stars on the modified Newcastle-Ottawa Scale, whilst 19 articles scored five stars, four articles scored four stars, and one article scored three stars (online suppl. material 3). From this, it can be inferred that most studies included in this review have moderate or high methodological quality, with minimal bias observed in study reporting.

*Adherence to Reporting Guidelines*

Of the 74 included studies, only 10 studies adhered to established reporting guidelines such as the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement [110] and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [111]. Specifically, nine studies followed the TRIPOD statement [22, 28–30, 40, 51, 55, 72, 86] while one study adhered to STROBE guidelines [108]. This indicates that

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744

**Table 1.** Overview of included studies

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Achterberg [35] (2014) | RS | NA | Netherlands | 511 (52; 41 with AD) | 10 | Hippocampal volume, hippocampal shape, and hippocampal volume + shape | All-cause and AD: hippocampal volume + shape | All-cause: 0.72–0.77<br>AD: 0.66–0.72 | All-cause: 0.72–0.76<br>AD: [a] | NU |
| Achterberg [36] (2019) | RS | NA | Netherlands | 510 (52) | 11 | Hippocampal volume, hippocampal shape, hippocampal texture, volume + shape, volume + texture, shape + texture, and volume + shape + texture | Hippocampal volume + shape + texture | All-cause: 0.63–0.79 | All-cause: 0.70–0.79 | NU |
| Acosta [26] (2018) | 10/66 | NA | Mexico | 1,355 (129) | 3 | Age, location (rural), MCI, diabetes, illiteracy and ≥2 NPS | Sum of age 80+, live in rural area, MCI, diabetes, illiteracy, and 2 or more NPSs | All-cause: 0.72–0.75 | NU | NU |
| Beebee-Wang [37] (2021) | ROS/MAP | NA | USA | 1,597 (521) | 3 | Age, sex, education, APOE, MCI and cognitive tests including categorical fluency (semantic memory), symbol digit modality test, word list (immediate and delayed), MMSE, logical memory II, category fluency (fruits, animals), Boston Naming, logical memory I, number comparison, East Boston Test (immediate and delayed), Stroop Colour Naming Task and line orientation (perceptual orientation) | All features model | NR | All-cause: 0.81 (SE = 0.0005)–0.92 (SE = 0.0044) | NU |
| Ben-Hassen [38] (2021) | PAQUID | 3C | France | PAQUID = 2,880 (813)<br>3C = 3,953 (602) | 5 and 10 | Age, sex, education, IST, MMSE, BVRT and IADL4 (NOTE: value and slope included for cognitive and ADL variables) | Age, sex, education, MMSE, IST, and IADL4 | NR | All-cause: 0.90–0.96 | All-cause: 0.85–0.96 |
| Calvin [39] (2019) | UK Biobank | NA | UK | 81,823 to 145,068 (1,051; *n* = 352 AD & *n* = 169 VaD) | 3–8 | Age, sex, education, cognitive tests (reaction time, visual memory, verbal-numerical, and prospective memory), family history of dementia, depression, and APOE e4 | Fully adjusted model: age, sex, education, visual memory, verbal-numerical reasoning, reaction time, family history of AD, depression, and APOE e4 | All-cause: 0.76 (0.73–0.80)–0.86 (0.81–0.90) | NU | NU |
| Capuano [40] (2022) | ROS/MAP | MARS | USA | MAP = 1,872 (NR)<br>ROS = 1,308 (NR)<br>MARS = 999 (NR) | 3 | RADaR: age, memory complaints, handling of finances, month recall, room recall, and three-word delayed recall<br>BDSI: age, education, BMI, diabetes, stroke, IADL (needs help with money or medications), and depression | RADaR | RADaR (all-cause): 0.81 (0.78–0.85) | RADaR (all-cause): 0.82 (0.78–0.86) | RADaR (all-cause): 0.85 (0.80-0.90)<br><br>BDSI (all-cause): 0.72 (0.67-0.77) |
| Casanova [19] (2016) | NA | BLSA and AGES-RS | USA and Iceland | BLSA = 192 (93)<br>AGES-RS = 200 (100) | 5 | Mapstone's (2014) 10-metabolite panel model [41] | Mapstone's (2014) 10-metabolite panel model [41] | NA | NA | All-cause: 0.40–0.64 |
| Chouraki [42] (2016) | NA | FHS, 3C, CHS, AGES, ROS/MAP, ACT, and WHICAP | Finland<br>France<br>Netherlands<br>USA | 19,687 (2,782) | 5, 6, 7, and 8 (results reported for 7-year follow-up) | Age, sex, education, APOE e4 status, and GRS | Age, sex, education, APOE e4 status, and GRS (Lambert et al.) | AD: 0.69–0.80 | NU | NU |
| Dallora [43] (2020) | NA | SNAC | Sweden | 726 (91) | 10 | See original paper[b] | Age, single leg standing test with left leg, past smoking, present smoking frequency, diabetes type 2, alcohol consumption, number of medications taken regularly, hand grip strength, BMI, and backwards digit span | NR | All-cause: 0.74 | NU |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744



**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Danso [44] (2021) | Share | Prevent | 20 European countries | 84,856 (4,157) | 14 | Age, gender, education, marital status (widowed, divorced, married living with spouse, married not living with spouse, never married, registered partnership), no children, daily sport (vigorous and moderate), emotional disorders, BMI, hypertension, cholesterol, diabetes, heart attack, smoking, lung disease, peptic ulcer, osteoarthritis, and Parkinson's disease | XGBoost modelling | NR | All-cause: 0.94–0.96 | All-cause: 0.63 |
| de la Fuente [45] (2020) | ELSA | NA | UK | 2,255 (NR) | 10 | Self-reported sensory functioning scale and cognitive measures | Sensory-cognitive difficulties Latent score (including visual, hearing, and cognitive functioning) | All-cause: 0.50 (0.37–0.63)–0.80 (0.75–0.86) | NU | NU |
| Deckers [46] (2020) | NA | CAIDE study cohort | Finland | 1,024 (84) | Mean: 20.9 and 28.9 | LIBRA: age, education, coronary heart disease, diabetes, hypercholesterolaemia, hypertension, depression, obesity, smoking, physical inactivity, renal disease, alcohol consumption, cognitive activity, healthy diet, and APOE | LIBRA + age, education, and diet: age, education, coronary heart disease, diabetes, hypercholesterolaemia, hypertension, depression, obesity, smoking, physical inactivity, renal disease, alcohol consumption, cognitive activity, and healthy diet | NA | NA | All-cause: 0.50–0.75 |
| Devanand [47] (2015) | WHICAP | NA | USA | 757 (101) | 4 | Age, sex, education, functional impairment, odour identification (UPSIT), and SRT-TR | UPSIT + SRT-TR | AD: 0.65–0.77 | NU | NU |
| Ding [48] (2020) | Shanghai Aging Study | NA | China | 947 (75) | Mean: 4.9 | Age, sex, BMI, height, education, smoking, drinking, coronary artery disease, hypertension, diabetes, stroke, APOE, MMSE, and olfactory identification test (including 12 smells) | Full model: age, sex, BMI, height, education, smoking, drinking, coronary artery disease, hypertension, diabetes, stroke, APOE, MMSE, and olfactory identification test (including 12 smells) | NR | All-cause: 0.90–0.92 | NU |
| Downer [27] (2016) | MHAS | MHAS | Mexico | 3,002 (251) | 11 | MHAS late-life risk index: age, sex, education, hypertension, diabetes, stroke, high depressive symptoms, IADL impaired, ADL impaired, Fall in past 2 years, and fair-poor vision BDSI [49]: age, education, stroke, type 2 diabetes, BMI, needing assistance with finances/ medication, and depressive symptoms | MHAS late-life risk index | NR | MHAS (all-cause): 0.70 (0.64–0.73)–0.74 (0.70–0.77) | BDSI (all-cause): 0.72 (0.69–0.76) |
| Downer [50] (2016) | H-EPESE | NA | USA (Mexican Americans from Texas, New Mexico, Colorado, Arizona and California) | 1,739 (229) | NR | MADeN: age, gender, education, not having friends to count on, not attending community events, diabetes, feeling the blues, pain, IADL impaired, and being unable to walk a half-mile | MADeN | NR | MADeN (all-cause): 0.74 (0.70–0.78) | NU |

**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Fayosse [51] (2020) | NA | Whitehall II | UK | 7,553 (318) | Mean: 23.5 | CAIDE [52]: age, education, sex, physical activity, SBP, BMI, and total cholesterol<br>FINDRISC [53]: age, sex, family history of diabetes, daily fruit/vegetable intake, physical activity, antihypertension medication, history of high blood glucose, BMI, and waist circumference<br>FRS [54]: age, sex, smoking status, diabetes, SBP, HDL, and total cholesterol | CAIDE and FRS | NA | NA | CAIDE (all-cause): 0.50 (0.46–0.54)–0.72 (0.69–0.75)<br>FINDRISC (all-cause): 0.52 (0.48–0.56)–0.63 (0.60–0.66)<br>FRS (all-cause): 0.53 (0.49-0.57)–0.72 (0.69-0.75) |
| Fisher [55] (2021) | Canadian Community Health Survey | NA | Canada | 75,460 (8,448) | 5 | DemPoRT: see original paper[b] | DemPoRT (full model) | DemPoRT (all-cause): 0.82 (0.80–0.83)–0.82 (0.81–0.83) | DemPoRT (all-cause): 0.83 (0.81–0.85) | NU |
| Fukunishi [56] (2020) | Health insurance claim and long-term care insurance data | NA | Japan | 48,123 (3,095) | 31 months | See original paper[b] | Sparse logistic regression with L1 regularisation | AD: 0.65–0.65–0.65)–0.73 (0.72–0.74) | AD: 0.65 (0.65–0.65)–0.66 (NR) | NU |
| Graves [57] (2019) | NA | Intermountain healthcare hospitals | USA | 74,081 (449) with atrial fibrillation | 10 | $CHA_2DS_2$-VASc [58]: age, congestive heart failure, hypertension, diabetes, stroke/TIA, vascular disease, and sex<br>IMRS [59]: Complete blood count and basic metabolic panel | $CHA_2DS_2$-VASc score with moderate Intermountain Mortality Risk Score | NA | NA | $CHA_2DS_2$-VASc (all-cause): 0.27 (0.16–0.39)–0.77 (0.66–0.87) |
| Hall [60] (2019) | Vantaa +85 cohort | NA | Finland | 245 (97) | Mean: 5.6 | APOE genotype, APOE e2 carrier, APOE e4 carrier, all genotypes, sociodemographics, social class, and functioning | Overall model | NR | AD: 0.57 (0.51–0.62)–0.68 (0.61–0.76) | NU |
| Honda [61] (2021) | Hisayama Study | NA | Japan | 795 (364) | 24 (calculated in 10-year probabilities) | Age, sex, education, hypertension, hypercholesterolaemia, diabetes, BMI, stroke, electrocardiogram abnormalities, cancer, current smoking, current alcohol consumption, no regular exercise, and sedentariness | Simplified risk score: age, sex, education, hypertension, diabetes, BMI, stroke, current smoking, and sedentariness | All-cause: 0.72 (0.69–0.75)–0.76 (0.72–0.79) | All-cause: 0.70 | NU |
| Hu [62] (2019) | NA | NHIRD | Taiwan | 387,595 (NR) heart failure patients | Mean: 2.9 | $CHA_2DS_2$-VASc [58]: atrial fibrillation, hyperlipidaemia, COPD, hyperthyroidism, sleep disorder, gout, chronic kidney disease, anaemia, head injury, depression, and alcoholism-related disease<br>AHEAD [63]: gender, CVA or TIA, vascular disease, hypertension, COPD, hyperlipidaemia, hyperthyroidism, sleep disorder, gout, head injury, depression, and alcoholism-related disease | $CHA_2DS2$-VASc | NA | NA | $CHA_2DS2$-VASc (all-cause): 0.61 (0.60–0.61)<br>AHEAD (all-cause): 0.55 (0.54–0.55) |
| Hu [28] (2021) | CLHLS | NA | China | 6,718 (991) | 3 | See original paper[b] | Age, IADL, baseline MMSE, and marital status | NR | All-cause: 0.77–0.82 | NU |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744



**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Hu [29] (2022) | CLHLS | CLHLS | China | 10,053 (1,750) | 6 | Age, sex, education, marital status, economic status, smoking status, smoking index, drinking status, drinking index, activity status, activity duration, playing cards or mah-jongg, watching TV or listening to the radio, doing garden work, taking part in some social activities, sleep quality, sleep duration, hypertension, heart disease, diabetes, stroke or cardiovascular diseases, BMI, mood, SBP, and DBP | Final model: age, sex, education, marital status, activity duration, playing cards or mah-jongg, watching television or listening to the radio, presence of stroke, or cardiovascular disease | NR | All-cause: 0.71 (0.70–0.72)–0.71 (0.70–0.73) | All-cause: 0.67 (0.65–0.69)–0.69 (0.67–0.70) |
| Ibarrondo [64] (2022) | EPIC-Spain | NA | Spain | 24,507 (755) | 5–25 | Age, sex, education, physical activity, hypertension, waist circumference, hyperlipidaemia, diabetes, smoker, BMI, and Mediterranean diet | Age, sex, education, physical activity, hypertension, waist circumference, hyperlipidaemia, diabetes, smoker, BMI, and Mediterranean diet | NR | All-cause: 0.55 (0.48–0.61)–0.81 (0.76–0.86) | NU |
| Jacqmin-Gadda [65] (2014) | PAQUID | NA | France | Sample A: 1,313 (121) Sample B: 2,795 (265) | 5 or 10 | Age, sex, education, forgetfulness in daily living, 4-IADL score, Isaacs Set Test score, DSST score, and MMSE | Age, sex, education, forgetfulness in daily living, 4-IADL score, Isaacs Set Test score, DSST score, and MMSE | All-cause: 0.75 (SE = 0.016)–0.85 (SE = 0.030) | All-cause: 0.75–0.84 | NU |
| Kim [66] (2019) | NHIS-HEALS | NA | South Korea | 43,648 (21,824) | 10 | Age, sex, BMI, SBP, DBP, fasting plasma glucose, total cholesterol, smoking, no exercise, cardiovascular disease, diabetes, hypertension, psychiatric disorder, and neurological disorder | Deep learning model with repeated measurements | NR | All-cause: 0.84 (0.83–0.85)–0.90 (0.90–0.90)–AD: 0.87 (0.86–0.88)–0.91 (0.91–0.91) | NU |
| Kochan [67] (2016) | Sydney-MAS | NA | Australia | 861 (48) | 4 | Age, education, complex mean RT, simple mean RT, complex IIVRT, simple IIVRT, combined RT measures, logical memory (delayed), RAVLT (delayed), category fluency (animals), coding, block design, Benton Visual Retention, trail making test, Boston Naming Test, trail making test A, letter fluency, and combined neuropsych tests | Age, sex, education, and combined neuropsychological measures | All-cause: 0.56 (0.48–0.64)–0.89 (0.83–0.94) | NU | NU |
| Li [68] (2018) | FHS | NA | USA | 2,383 (778) | 5, 10, and 20 | Age, sex, marital status, BMI, smoking status, alcohol consumption, daily consumption of tea, daily consumption of coffee, low salt diet, hypertension, stroke, diabetes, ischaemic attack, cancer, and vascular disease of the brain | Age, marital status, BMI, stroke, diabetes, ischaemic attacks, and cancer | All-cause: 0.72 | NU | NU |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744

**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Li [69] (2018) | NHIRD and NDCMP | NA | Taiwan | 27,540 type 2 diabetes patients ($n$ = 18,360 derivation dataset) (1,228 (853 in derivation dataset)) | Mean: 8.1 | Age, sex, smoking habit, alcohol consumption, age of diabetes onset, duration of type 2 diabetes, BMI, obesity, diabetes related factor and biomarkers, hypertension, stroke, coronary artery disease, peripheral artery disease, peripheral neuropathy, diabetes retinopathy, disease of the peripheral circulatory disturbance, hypoglycaemia, traumatic amputation, ketoacidosis, postural hypotension, arterial embolism and thrombosis, hyperlipidaemia, anti-diabetic medication, antihypertensive medication, and cardiovascular medications | Age, sex, diabetes duration, BMI, variation, fasting plasma glucose, variation HBa1c, stroke, hypoglycaemia, postural hypertension, coronary artery disease, and anti-diabetes medications | All-cause: 0.76 (0.75–0.77)–0.82 (0.80–0.84) | All-cause: 0.75 (0.73–0.77)–0.84 (0.80–0.88) | NU |
| Li [24] (2018) | Framingham Offspring Study | NA | USA | 2,461 (227) | 30+ | Age, sex, marital status, BMI, alcohol, smoke, salt intake, sleep, physical activity | Age, sex, marital status, BMI, alcohol, smoke, salt intake, sleep, physical activity | All-cause: 0.79–0.89 | NU | NU |
| Liao [70] (2015) | NA | NHIRD | Taiwan | 332,665 AF patients (29,012) | NR | $CHADS_2$ [71]: age, hypertension, diabetes mellitus, heart failure, and stroke/TIA<br>$CHA_2DS_2$-VASc [58]: age, history of hypertension, diabetes, recent cardiac failure, vascular disease (myocardial infarction or peripheral artery disease), sex, and stroke/TIA | $CHA_2DS_2$-VASc | NA | NA | $CHADS_2$ (all-cause: 0.59 (0.59–0.59)<br>$CHA_2DS_2$-VASc (all-cause): 0.61 (0.61–0.61) |
| Licher [72] (2018) | NA | RS | Netherlands | 6,667 (867) | 2, 5, 10, and 15 | CAIDE [52]: age, sex, obesity, hypertension, education, hypercholesterolaemia, and psychical activity<br><br>BDSI [49]: age, obesity, education, type 2 diabetes, stroke, depressive symptoms, and needs help with money/medications<br><br>ANU-ADRI [73]: age, sex, BMI, education, alcohol, smoking, hypercholesterolaemia, diabetes, TBI, depressive symptoms, social network, and fish consumption<br><br>DRS [20]: age, gender, BMI, hypertension, alcohol, smoking, diabetes, stroke, atrial fibrillation, depressive symptoms, social deprivation, anxiety, aspirin use, NSAID use, and calendar year | BDSI and DRS | NA | NA | CAIDE (all-cause): 0.49 (0.42–0.56)–0.55 (0.53–0.57)<br>CAIDE (AD): 0.51 (0.43–0.59)–0.55 (0.53–0.57)<br>BDSI (all-cause): 0.59 (0.57–0.61)–0.83 (0.75–0.90)<br><br>BDSI (AD): 0.78 (0.75–0.80)–0.83 (0.75–0.90)<br>ANU-ADRI (All-cause): 0.51 (0.47–0.55)–0.81 (0.77–0.86)<br>ANU-ADRI (AD): 0.69 (0.68–0.71)–0.80 (0.75–0.85)<br>DRS (all-cause): 0.55 (0.53–0.57)–0.84 (0.77–0.92)<br>DRS (AD): 0.80 (0.78–0.82)–0.84 (0.77–0.92) |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744



**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/ c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Licher [74] (2019) | RS | EPOZ | Netherlands | 2,710 (181) | 10 | BDRM: age, age-squared, sex, education level, systolic blood pressure, current smoking, history of diabetes, history of symptomatic stroke, depressive symptoms, parental history of dementia, presence of subjective memory decline, need for assistance with finances or medication, letter digit substitution test, the delayed word learning test, APOE-ε4 carrier status, and all imaging markers except brain infarcts | BDRM basic model + letter digit substitution test, the delayed word learning test, APOE-ε4, and all imaging markers except brain infarcts | All-cause: 0.79 (0.76–0.83) | All-cause: 0.78 (0.75–0.81)–0.86 (0.83–0.88)–AD: 0.77 (0.75–0.81)–0.86 (0.83–0.88) | All-cause: 0.75 (0.67-0.82)–0.81 (0.74-0.88) |
| Lin [75] (2018) | JAGES | NA | Japan | 72,127 (6,656) | Mean: 3 | Age, sex, and 25 Kihon Checklist items | Age-sex categories and Kihon Checklist items | All-cause: 0.73 (0.73–0.74)–0.79 (0.79–0.80) | All-cause: 0.73 (0.73–0.74)–0.79 (0.79–0.80) | NU |
| Lin [76] (2022) | Framingham Offspring Study | NA | USA | 1,642 (243) | Mean: 12 | Age, sex, survival time, APOE e4 status, and 38 blood biomarkers | Age, sex, survival time, APOE e4 status, and nine most informative biomarkers | NR | All-cause: 0.69 (SE±0.01)–0.76 (SE±0.01) | NU |
| Makino [77] (2021) | Sub cohort of the NCGG-SCG | NA | Japan | 4,298 (93) | 2 | STAD: Do you forget where you have left things more than you used to? Do other people find you forgetful? Do you find yourself not knowing today's date? Have you dropped many of your activities and interests? Do you often get bored? Do you feel helpless? Do you prefer to stay at home, rather than going out and doing new things? In the last 2 weeks have you felt tired without a reason? Do you go out less frequently compared to last year? Do you engage in low levels of physical exercise aimed at health at least five times a week? Do you use maps to go to unfamiliar places? and Do you engage in cognitive stimulation, such as board games and learning | STAD | NR | All-cause: 0.65–0.70 | NU |
| McCoy [21] (2020) | Academic medical centres for adults data | NA | USA | 267,855 (6,516) | 8 | Age, sex, race, Charlson comorbidity index, cognitive symptom burden score (at discharge from narrative hospital notes) | Age, sex, race, Charlson comorbidity index, cognitive symptom burden score (at discharge from narrative hospital notes) | All-cause: 0.58 (0.58–0.59)–0.77 (0.73–0.81) | NU | NU |
| Mehta [78] (2016) | CPRD | NA | UK | Participants with diabetes and hypertension: 133,176 (NR) Participants with diabetes and no hypertension: 16,677 (validation sample) | 9 | RxDx: myocardial infarction, congestive heart failure, coronary and peripheral vascular disease, cerebrovascular disease, chronic pulmonary disease, rheumatologic disease, peptic ulcer disease, renal disease/end stage renal disease, mild liver disease/moderate or severe liver disease, any malignancy, including lymphoma | RxDx (categorical) | All-cause: 0.78 (0.78–0.79)–0.81 (0.80–0.81) | All-cause: 0.81 (0.80–0.81)–0.86 (0.84–0.87) | NU |

**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/ c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | and leukaemia/metastatic solid tumour, epilepsy, hyperlipidaemia, Parkinson's disease, cardiac disease ASCVD, glaucoma, transplantation, thyroid disorder, gout, Crohn's and ulcerative disease, pain and inflammation/ pain, depression, psychotic illness, bipolar disorders, anxiety and tension), Charlson comorbidity score, chronic disease score | | | | |
| Muller [79] (2020) | SNAC-K (MRI sub-sample) | NA | Sweden | 212 (16) | 6 | Age, sex, education, cognition, APOE, Brain volume, white matter hyperintensities, measures of microstructural white matter integrity including mean diffusivity, and fractional anisotropy | Age, sex, education, total brain tissue volume to global cognition, and APOE | All-cause: 0.55–0.92 | NU | NU |
| Mura [80] (2017) | 3C | NA | France | 2,558 (173; 116AD) | 3 and 5 | Free and Cued Selective Reminding Test (various subtests) | Free and Cued Selective Reminding Test – Free Recall | All-cause: 0.65 (0.58–0.72)–0.85 (0.79–0.90)–<br>AD: 0.64 (0.56–0.73)–0.88 (0.81–0.95) | All-cause: 0.71 (0.60–0.82)–0.89 (0.83–0.95)–<br>AD: 0.69 (0.53–0.85)–0.92 (0.82–1.00) | NU |
| Nagata [81] (2019) | The Hisayama Study | The Hisayama Study | Japan | 1,635 (377) | Median = 10.2 | Age, sex, education level, systolic blood pressure, use of antihypertensive agents, use of heart disease agents, diabetes mellitus, hypercholesterolaemia, BMI, estimated glomerular filtration rate, atrial fibrillation, history of stroke, smoking habit, alcohol intake, regular exercise, and serum high-sensitivity C-reactive protein, and log serum NT-proBNP levels<br>CAIDE [52]: age, education status, sex, systolic blood pressure, BMI, total cholesterol, and physical activity + serum NT-proBNP levels<br>FRS [54]: age, marital status, BMI, stroke, diabetes, and cancer + serum NT-proBNP levels | Basic model + serum NT-proBNP levels: age, sex, education level, systolic blood pressure, use of antihypertensive agents, use of heart disease agents, diabetes mellitus, hypercholesterolaemia, body mass index, estimated glomerular filtration rate, atrial fibrillation, history of stroke, smoking habit, alcohol intake, regular exercise, and serum high-sensitivity C-reactive protein | NR | All-cause: 0.78 (0.76–0.80)–0.79 (0.77–0.81) | FRS (all-cause): 0.75–0.77<br>CAIDE (all-cause): 0.57–0.69 |
| Nakazawa [82] (2022) | The Hisayama Study | NA | Japan | 1,158 (113) | 5 | Age, sex, education status, systolic blood pressure, antihypertensive medication, diabetes mellitus, serum total cholesterol, body mass index, ECG abnormalities, cerebrovascular lesions on MRI, smoking habits, alcohol intake, regular exercise, hippocampal atrophy, multiple-region grey matter atrophy, medial temporal GMV/TBV, insular GMV/TBV, hippocampal GMV/TBV, amygdala GMV/TBV | Basic model +accumulated numbers of grey matter atrophy: Age, sex, education status, systolic blood pressure, antihypertensive medication, diabetes mellitus, serum total cholesterol, body mass index, ECG abnormalities, cerebrovascular lesions on MRI, smoking habits, alcohol intake, regular exercise, and the total brain volume-to-intracranial volume ratio | All-cause: 0.75—0.78<br>AD: 0.77—0.80 | NU | NU |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744



**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/ c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Nori [83] (2019) | Electronic health record dataset from the OptumLabs Data Warehouse | NA | USA | 215,196 (179,330 controls) (35,866) | 4––5 | See original paper[b] | Label learning models | NR | All-cause: 0.64––0.90 | NU |
| Ostovaneh [84] (2020) | MESA | NA | USA | 4,563 (223) | 10 | See original paper[b] | Multivariable mode l+ baseline and change in log-NT-proBNP | All-cause: 0.85 (0.83–0.88)–0.87 (0.84–0.89) | NU | NU |
| Park [85] (2019) | NHIS-HEALS | NA | South Korea | 331,126 (15,501) | Mean: 10 | Age, sex, BMI, hypertension, cardiovascular disease, diabetes, psychiatric disorder, neurological disorder, smoking, physical activity, and blood pressure | Age, sex, BMI, hypertension, cardiovascular disease, diabetes, psychiatric disorder, neurological disorder, smoking, physical activity, and blood pressure | All-cause: 0.81 (0.80–0.82)–0.81 (0.81–0.82) | All-cause: 0.81 (0.81–0.82) | NU |
| Park [86] (2020) | NHIS-HEALS | NA | South Korea | 40,736 (614 definite-AD, 2,026 probable-AD) | 1–4 | See original paper[b] | All feature model | AD: 0.56–0.89 | AD: 0.61–0.90 | NU |
| Paul [87] (2021) | Faroese Septuagenarian Cohort | NA | Denmark | 713 (65) | 10 | Sex and various neuropsychological variables (see original paper[b]) | Sex, delayed recall, incidental memory, attention speed, and visuospatial perception | NR | All-cause: 0.77 | NU |
| Payton [88] (2018) | SNAC-K | NA | Sweden | 418 (28) | 6 | Age, sex, education, episodic memory, semantic memory, verbal fluency, perceptual speed, executive function, APOE, MRI (total tissue volume, total grey matter volume, hippocampal volume), and general knowledge | Age, sex, education, APOE, word recall, and general knowledge | All-cause: 0.56–0.92 | NU | NU |
| Payton [89] (2020) | SNAC-K | NA | Sweden | 2,357 (246) | 12 | Age, sex, education, episodic memory, semantic memory, verbal fluency, perceptual speed, and executive function | Age, sex, education, category fluency, word recall, and pattern comparison | All-cause: 0.64–0.91 | NU | NU |
| Phongpreecha [34] (2020) | Pacific Udall Center | NA | USA | 827 (NR) | Mean: 4 | Age, education, sex, disease duration (time since initial Onset of PD motor symptoms), total levodopa equivalent daily dose, depression, MoCA, HVLT-R, Letter-Number Sequencing, Digit Symbol, verbal fluency, Benton Judgment of Line Orientation, trail making test B-A, APOE, MAPT, and GBA | Cognitive test score model | NR | PDD: 0.77–unclear[a] | NU |
| Reinke [90] (2022) | German health insurance company "Allgemeine Ortskrankenkasse" | NA | Germany | 117,895 (27,651) | 10 | See original paper[b] | Logistic regression model | NR | All-cause: 0.64 (0.63–0.64)–0.71 (0.70–0.72) | NU |
| Schiepers [91] (2018) | MAAS | NA | Netherlands | 949 (61) | 16 | Age, sex, education, LIBRA Score (alcohol, CHD, physical activity, renal dysfunction, diabetes, high cholesterol, smoking, obesity, hypertension, depression, and cognitive activity) | LIBRA + age, sex, and education | All-cause: 0.59 (0.52–0.66)–0.75 (0.69–0.80) | NU | NU |
| Song [92] (2014) | CSHA | NA | Canada | 7,239 (607) | 10 | 42 variables encompassing demographic, health, lifestyle, and environmental factors[b] | All-factor index | NR | All-cause: 0.66 (±0.03)–0.67 (±0.03) | NU |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744

**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Stephan [93] (2015) | 3C | NA | France | 1,721 (119) | 10 | Age, sex, education, ADL impairment, MMSE, Benton Visual Retention Test, Digit Span, smoking, alcohol, diabetes, cardiovascular disease, SBP, APOE, MRI (white matter lesion volume, whole brain volume, hippocampal volume) | Model 1 + hippocampal volume/ all MRI variables | NR | All-cause: 0.77 (0.71–0.82)–0.79 (0.74–0.84) AD: 0.77 (0.71–0.83)–0.81 (0.75–0.86) | NU |
| Stephan [30] (2020) | NA | 10/66 | China, Cuba, the Dominican Republic, Mexico, Peru, PR, and Venezuela | 11,143 (1,069) | 3–5 | CAIDE [52]: age, sex, education, systolic blood pressure, BMI, total cholesterol, and physical activity<br>AgeCoDe [94]: age, subjective memory impairment, verbal fluency, delayed recall, MMSE, and IADL<br>ANU-ADRI [73]: age group (by sex), education, BMI, diabetes, symptoms of depression, total cholesterol, traumatic brain injury, smoking, alcohol use, social engagement, physical activity, and fish intake<br>BDSI [49]: age, education, BMI, diabetes, stroke, IADL (needs help with money or medications), and depressive symptoms<br>BDRM [74]: age, stroke, subjective memory decline, IADL (needs help with finances or medication) | BDRM | NA | NA | CAIDE (all-cause): 0.52 (0.47–0.57)–0.58 (0.50–0.61)<br>AgeCoDe (all-cause): 0.57 (0.52–0.62)–0.74 (0.69–0.79)<br>ANU-ADRI (all-cause): 0.66 (0.61–0.71)–0.78 (0.68–0.88)<br>BDSI (all-cause): 0.62 (0.57–0.68) too 0.78 (0.71–0.86)<br>BDRM (all-cause): 0.66 (0.62–0.70)–0.79 (0.74–0.83) |
| Stocker [95] (2020) | ESTHER | NA | Germany | 770 (150) | 8–14 | Age, sex, education, physical activity, obesity, diabetes, hypertension, cardiovascular events, depression, APOE, and Aβmisfolding | Model including APOE and Aβmisfolding | AD: 0.59 (0.48–0.69)–0.85 (0.79–0.90) | NU | NU |
| Stocker [32] (2021) | ESTHER | NA | Germany | 5,203 (631) | 17 | Age, sex, education, PRS, APOE | Age, sex, education + PRS + APOE | All-cause: 0.78–0.79<br>AD: 0.77–0.81<br>VaD: 0.78–0.78<br>Mixed dementia: 0.84–0.85 | NU | NU |
| Stocker [33] (2021) | ESTHER | NA | Germany | 728 (147) | 14 | Age, sex, education, Aβmisfolding, PRS, APOE | Age, sex, education + APOE and Aβ misfolding | AD: 0.55 (0.46–0.63)–0.86 (0.80–0.91) | NU | NU |
| Tjandra [96] (2020) | Michigan-ADRC and University of Michigan's Research Data Warehouse | NA | USA | 8,474 (350) | 10 | Laboratory tests, procedures, healthcare utilisation, diagnosis, vital signs, age features, medications, sociodemographic | All categories of EHR data combined | NR | AD: 0.54 (0.47–0.61)–0.70 (0.63–0.77) | NU |
| Tjandra [97] (2021) | VA | MM | USA | 5,488 (VA development); 1,372 (VA validation); 1,201 (external MM) (133 [VA development]; 33 [VA validation]; 30 [external MM]) | 5 | Age, sex, race, blood pressure trajectories and summary statistics | Blood pressure trajectories | NR | AD: 0.59 (0.49–0.68)–0.64 (0.54–0.73) | AD: 0.63 (0.55–0.72)–0.66 (0.55–0.76) |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744



**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/ c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Tomata [98] (2017) | Ohsaki Cohort 2006 Study | NA | Japan | 13,974 (1,229) | 5.7 | Age, sex, Kihon Checklist-Cognitive Function | Age, sex, Kihon Checklist-Cognitive Function | All-cause: 0.63 (0.61–0.66)–0.65 (0.63–0.66) | NU | NU |
| Tynkkynen [99] (2015) | FINRISK | NA | Finland | 7,158 (220) | Median = 13.8 | Age, sex, HDL, BMI, prevalent ischaemic heart disease, prevalent diabetes, and NT-proBNP | Age, sex, HDL, BMI, prevalent ischaemic heart disease, prevalent diabetes + NT-proBNP | NR | All-cause: 0.86–0.87 | NU |
| Tynkkynen [100] (2017) | FINRISK | NA | Finland | 7,114 (407) | 18 | Sex, total cholesterol, SBP, BMI, education, APOE, hs-TNI, NT-proBNP | All-cause: model + hs-TNI<br><br>AD: model + NT-proBNP | NR | All-cause: 0.64–0.70<br>AD: 0.66–0.72 | NU |
| Vonk [22] (2021) | NA | AGES-RS | Iceland | 5,343 (all-cause: 1,099; AD: 492) | 6 and 10 | ANU-ADRI [73]: age, sex, education, diabetes, head trauma, alcohol, smoking, fish intake, physical activity, mental leisure activity, social leisure activity, and depressive symptoms<br>Barnes et al. (2009) model [101]: age, coronary bypass surgery, APOE, white matter disease, ventricular enlargement, carotid intima-media thickness, BMI, MMSE, DSST, difficulty to dress, and alcohol consumption<br>BDSI [49]: age, education, stroke, diabetes, BMI, difficulty managing money, and depressive symptoms<br>Hogan et al. (2000) model [102]: age, MMSE and subjective memory concerns<br>Li et al. (2018) model [68]: age, stroke, diabetes, TIA, cancer, BMI, and marital status<br><br>BDRM [74]: age, stroke, subjective memory concerns, and difficulty managing money<br><br>MADeN [50]: age, sex, education, diabetes, pain walking/standing, difficulty walking, IADL, social leisure activity, not having friends, and depressive symptoms<br>Tierney et al. (2010) model [103]: age, sex, education, delayed memory recall, and DSST<br>Mura et al. (2017) model [80]: age, sex, education, and delayed memory recall<br>Tierney et al. (2005) model [104]: age, education, and delayed memory recall<br>Verhaaren et al. (2013) model [105]: age, sex, and APOE | All-cause: Hogan et al. (2000) and Barnes et al. (2009) models<br>AD: Mura et al. (2017) model | NA | NA | ANU–ADRI (all–cause): 0.71 0.69–0.74)–0.73 (0.71–0.76)<br><br>Barnes et al. (all–cause): 0.80 (0.78–0.82)<br><br>BDSI (all–cause): 0.72 (0.70–0.75)<br><br>Hogan et al. (all–cause): 0.80 (0.78–0.82)<br>Li et al. (all–cause): 0.70 (0.68–0.72)–0.71 (0.69–0.73)<br>BDRM (all–cause): 0.74 (0.72–0.76)–0.75 (0.73–0.77)<br>MADeN (all–cause): 0.72 (0.71–0.74)<br><br>Tierney et al. (2010; all–cause): 0.77 (0.76–0.79)<br>ANU–ADRI (AD): 0.67 (0.63–0.70)–0.68 (0.64–0.71)<br>Mura et al. (AD): 0.81 (0.78–0.84)<br><br>Tierney et al. (2005; AD): 0.76 (0.74–0.78)<br>Verhaaren et al. (AD): 0.70 (0.68–0.72)–0.73 (0.71–0.75) |

Dement Geriatr Cogn Disord Extra 2024;14:49–74
DOI: 10.1159/000539744

**Table 1** (continued)

| Author | Development and internal validation cohort(s) | External validation cohort(s) | Country | Baseline sample followed up (incident cases) | Follow-up, years | Considered predictor variables | Best performing model(s) | Development AUC/c-statistic range (95% CI/SE) | Internal validation AUC/c-statistic (95% CI/SE) | External validation AUC/c-statistic (95% CI/SE) |
|---|---|---|---|---|---|---|---|---|---|---|
| Vos [106] (2017) | NA | DESCRIPA | Sweden, Italy, Netherlands, and France | 9,387 (1,120) | 16 | LIBRA [107]: depression, diabetes, physical inactivity, hypertension, obesity, smoking, hypercholesterolaemia, coronary heart disease, mild/moderate alcohol use | LIBRA in midlife | NA | NA | LIBRA (all-cause): 0.50 (SE = 0.01)–0.57 (SE = 0.03) |
| Vuoksimaa [25] (2016) | FTC | NA | Finland | 2,602 (265) | 20–40 Mean: 28 | Age, sex, follow-up time, cognitive reserve score (telephone assessment for dementia, and telephone interview for cognitive status) | Age, sex, follow-up time, cognitive reserve score | All-cause: 0.61 (0.52–0.69)–0.82 (0.72–0.91) | NU | NU |
| Walker [108] (2021) | ARIC | NA | USA | 4,110 (428) | 5 | Age, sex, education, race, APOE, proteins (top 50), BMI, diabetes, hypertension, smoking status, eGFR-creatinine | Demographic characteristics + clinical variables + proteins | AD: 0.77 (0.74–0.79)–0.82 (0.80–0.84) | AD: 0.74 (0.71–0.77)–0.78 (0.75–0.81) | NU |
| Walters [20] (2016) | THIN | THIN | UK | Development cohort: 930,395 (All-cause: 6,017; AD: 1,831; VaD: 1,308; unspecified/MD: 2,878) Validation cohort 60–79 years: 226,140 (all-cause: 1,699; AD: 528; VaD: 384; unspecified/MD: 787) Validation cohort 80–95 years: 38,084 (all-cause: 923; AD: 408; VaD: 364; unspecified/MD: 1,151) | 5 | DRS: age, sex, calendar year, BMI, current antihypertensive use, SBP, lipid ratio, smoking status, hazardous/harmful alcohol drinking, current depression/anti-depressants, current anxiety/anxiolytics, current aspirin use, current other NSAID use, and history of diabetes, stroke, or TIA and atrial fibrillation | DRS (60–79 age group): age, sex, calendar year, deprivation, BMI, antihypertensive use, smoking, alcohol, diabetes, depression/antidepressant use, stroke/TIA, atrial fibrillation, and aspirin use | NR | DRS (all-cause): 0.84 (0.81–0.87) | DRS (all-cause): 0.56 (0.55–0.58)–0.84 (0.81–0.87) |
| Wang [31] (2017) | Beijing Longitudinal Study of Aging | NA | China | 2,788 (351) | 7 | Frailty Index: age, sex, education, stroke, hypertension, MMSE, diabetes, CHD, and 40+ other health deficits[b] | Frailty Index | All-cause: 0.74 (0.69–0.78) AD: 0.77 (0.74–0.79) | NU | NU |
| Wyman-Chick [23] (2022) | NACC-UDS | NA | USA | 464 (116) | 1 year prior–DLB diagnosis | CCF, MMSE, MoCA, NPS other than psychosis | Two or more CCF, or one CCF with low MMSE/MoCA scores, or one CCF with one or more NPS | DLB: 0.56 (0.52–0.59)–0.90 (0.86–0.93) | NU | NU |
| Zhou [109] (2022) | Clinical Data Analysis and Reporting System | NA | Hong Kong | 74,855 (3,550) | Median: 9.3 | Age, prior hypertension, high DBP, high SBP, prior cardiovascular, respiratory and gastrointestinal diseases, hypertension, diabetes, stroke, and laboratory examinations | Scoring system considering BP variabilities | NR | All-cause: 0.82–0.91 | NU |

Multiple studies did not state any internal validation methods, so we assume it is the development AUC/c-statistic. AD, Alzheimer's disease; ADL, activities of daily living; AgeCoDe, Study on Aging, Cognition and Dementia; AHEAD, Age, Health, Education, APOE genotype, and Depression; ANU-ADRI, Australian National University-Alzheimer Disease Risk Index; APOE, apolipoprotein E; ASCVD, atherosclerotic cardiovascular disease; AUC, area under the curve; BDRM, basic dementia risk model; BDSI, brief dementia screening indicator; BMI, body mass index; BP, blood pressure; BVRT, Benton visual retention test; CAIDE, cardiovascular risk factors, aging and dementia; CCF, clinical core features; CHD, coronary heart disease; CI, confidence interval; CIND, cognitive impairment, no dementia; COPD, chronic obstructive pulmonary disease; DBP, diastolic blood pressure; DemPoRT, dementia population risk tool; DLB, dementia with Lewy bodies; DRS, dementia risk score; DSI, dementia screening indicator; DSST, digit symbol substitution test; ECG, electrocardiogram; FINDRISC, Finnish diabetes risk score; FRS, Framingham risk score; GBM, Gradient boosting machines; GMV, grey matter volume; GRS, genetic risk score; HBP, high blood pressure; HDL, high-density lipoprotein; HVLT-R, Hopkins verbal learning test-revised; IADL, instrumental activities of daily living; IIVRT, intraindividual variability of reaction time; IMRS, intermountain mortality risk score; IST, Isaacs set test; LIBRA, lifestyle for brain health; MADeN, Mexican American Dementia Nomogram; MCI, mild cognitive impairment; MD, mixed dementia; MMSE, mini-mental state exam; MoCA, Montreal Cognitive Assessment; MRI, magnetic resonance imaging; NPS, neuropsychological symptoms; NR, not reported; NA, not applicable; NR, not reported; NSAID, non-steroidal anti-inflammatory drug; NU, not undertaken; PD, Parkinson's disease; PDD, Parkinson's disease dementia; PRS, polygenic risk score; RADaR, rapid assessment of dementia risk; RF, random forest; RT, reaction time; RVLT, Rey's auditory verbal learning test; SBP, systolic blood pressure; SE, standard error; SRT, selective reminding test; STAD, simplified telephone assessment for dementia risk; TBI, traumatic brain injury; TBV, total brain volume; TIA, transient ischaemic attack; UPSIT, University of Pennsylvania Smell Identification Test; VaD, vascular dementia; 3C, three-city study; ACT, adult changes in thought; AGES-RS, age, gene/environment susceptibility-Reykjavik study; ARIC, atherosclerosis risk in communities cohort; BLSA, Baltimore longitudinal study of aging; CASPER, Cognition and Affect After Stroke: Prospective Evaluation of Risks; CHS, cardiovascular health study; CLHLS, Chinese longitudinal healthy longevity survey; CPRD, clinical practice research datalink; CSHA, Canadian study of health and aging; CU-STRIDE, Chinese University of Hong Kong-Stroke Registry Investigating Cognitive Decline; DESCRIPA, development of screening guidelines and clinical criteria for pre-dementia AD study; ELSA, English longitudinal study of ageing; EPIC, European prospective investigation into cancer; EpiUSA, epidemiologic study of the risk of dementia after stroke; EPOZ, epidemiological prevention study of Zoetermeer; ESTHER, Epidemiologische Studie zu Chancen der Verhütung, Früherkennung und optimierten THerapie chronischer ERkrankungen in der älteren Bevölkerung study; FHS, Framingham heart study; FINRISK, National FINRISK 1997 study; FTC, Finnish twin cohort; H-EPESE, Hispanic Established Populations for the Epidemiologic Study of the Elderly; JAGES, Japan Gerontological Evaluation Study; MAAS, Maastricht Ageing Study; MAP, memory and aging project; MARS, Minority Aging Research Study; MESA, multi-ethnic study of atherosclerosis; MHAS, Mexican health and aging study; Michigan-ADRC, Michigan Alzheimer's Disease Research Centre; MM, Michigan Medicine's Epic EHR instance; NACC-UDS, National Alzheimer's Coordinating Centre Uniform Data Set; NCGG-SCG, National Centre for Geriatrics and Gerontology-Study of Geriatric Syndromes; NDCMP, Taiwan National Diabetes Care Management Program; NHIRD, National Health Insurance Research Database; NHIS-HEALS, National Health Insurance Service-Health Screening Cohort; PAQUID, Personnes Agées QUID; ROS, religious orders study; RS, Rotterdam study; SHARE, the Survey of Health, Ageing and Retirement in Europe; SNAC, Swedish national study on aging and care; SNAC-K, Swedish National Study on Aging and Care in Kungsholmen; STROKDEM, study of factors influencing post-stroke dementia; Sydney-MAS, Sydney Memory and Ageing Study; THIN, The health improvement network database; VA, Veterans affairs; VISN, Veterans integrated service network; WHICAP, Washington heights-Inwood community aging project. [a]Unable to ascertain specific values due to the presentation of results. [b]Due to the high number of component variables, it was not pragmatic to include them all in this table.

most studies did not utilise standardised reporting frameworks, which may impact the transparency and reproducibility of their findings.

*Component Variables*

Across the 74 articles, over 450 unique predictor variables have been incorporated within the context of dementia risk modelling, including demographic ($n$ = 8), cognitive ($n$ = 63), health (self-reported or objectively measured health status; $n$ = 187), lifestyle ($n$ = 13), imaging (e.g., magnetic resonance imaging; $n$ = 35), genetic ($n$ = 23), blood and protein biomarkers ($n$ = 58), physical functioning ($n$ = 25), and others ($n$ = 55) including, e.g., multilingualism and healthcare utilization. Like our previous reviews, models range from incorporating single predictors [80, 98] to a maximum of 10,363 variables within a single model [83]. Figure 2 shows the 34 most used model variables; defined as being incorporated into a prediction model a minimum of four times across the 74 included studies. All factors previously identified in the Lancet Commission on Dementia Prevention, Intervention, and Care [112] were frequently used as model variables, except for air pollution. Most variables are modifiable except for age, genotypic sex, ethnicity/race, and genes (e.g., apolipoprotein E e4).

*Model Development and Internal Validation*

Overall, 63 studies developed new dementia risk models since our previous reviews [8, 9], with 21 studies developing models without undergoing any validation procedures, 15 reporting both development and internal validation metrics, and 27 reporting internal validation results only (Table 1). This is a notable increase in the number of studies that have undertaken internal validation, with cross-validation resampling most frequently used.

Discriminative accuracy of models for all-cause dementia ranged from an AUC/c-statistic of 0.50 (logistic regression model using latent scores of hearing difficulties; follow-up = 10 years) [45] to 0.96 (discrete Bayesian networks developed model incorporating age, mini-mental state examination score, and fish olfactory identification; mean follow-up = 4.9 years) [48]. For AD, discriminatory accuracy ranged from 0.54 (ML incorporating numerous laboratory tests, healthcare data, and sociodemographic features; follow-up = 10 years) [96] to 0.92 (logistic regression using a Free and Cued Selective Reminding Test Risk Score; follow-up = 3 years) [113]. Developing a model exclusively for Dementia with Lewy bodies, one study found that the number of core clinical features, neuropsychiatric symptoms, and cognitive score affects discriminative accuracy (range from 0.56 to 0.90; follow-up = 1 year) [23]. One study incorporating polygenic risk scores into basic models reported fair discriminative accuracy in predicting VaD (AUC/c-statistic = 0.78; follow-up = 17 years) and mixed dementia (AUC/c-statistic = 0.84–0.85; follow-up = 17 years) [32], whilst another study utilising cognitive-based variables reported fair discriminative performance in predicting Parkinson's disease dementia (AUC/c-statistic = 0.77+; follow-up = mean of 4 years) [34]. As expected, internal validation was often fair (AUC/c-statistic between 0.70 and 0.79) or excellent (AUC/c-statistic ≥0.8), although four ML-based models reported poor internal validation results (AUC/c-statistic <0.7) [60, 83, 96, 97].

*External Validation*

In total, 27 dementia risk prediction models were externally validated (Fig. 3). This included models specifically aimed at developing a score for dementia risk, e.g., the Brief Dementia Screening Indicator (BDSI) [49], the Cardiovascular Risk Factors, Ageing and Dementia risk score (CAIDE) [52], the Australian National University AD Risk Index (ANU-ADRI) [73], Study on Aging, Cognition and Dementia (AgeCoDe) [94], and Lifestyle for Brain Health (LIBRA) [91, 107], or models that were developed to index risk in cardiovascular disease, such as the $CHADS_2$ [71], $CHA_2DS_2$-VASc [58], and AHEAD [63]. As shown in Figure 3, model discriminative accuracy and calibration varied greatly across the external validation cohorts, but overall was generally poor. Models that transported well include the Basic Dementia Risk Model (BDRM) [74], BDSI [49], and Chouraki et al.'s [42] model. Other models such as Ben-Hassen et al.'s [38] and the Rapid Risk Assessment of Dementia (RADaR) [40] reported high discriminative performances but need further external validation to assess their transportability. Discriminative performance across models not specifically aimed at dementia risk was generally poor.

*Model Fit*

Only 20 studies reported calibration results. Studies that developed new models generally reported good or excellent calibration [50, 61, 65, 69, 85, 96, 99], with few studies reporting mixed results with calibration ranging from poor to good [55, 90], or poor [37, 78]. Where previously developed models were tested outside the context in which they were developed (i.e., externally validated), the results varied from having poor [51, 72] to mixed/good calibration [20, 30, 22].

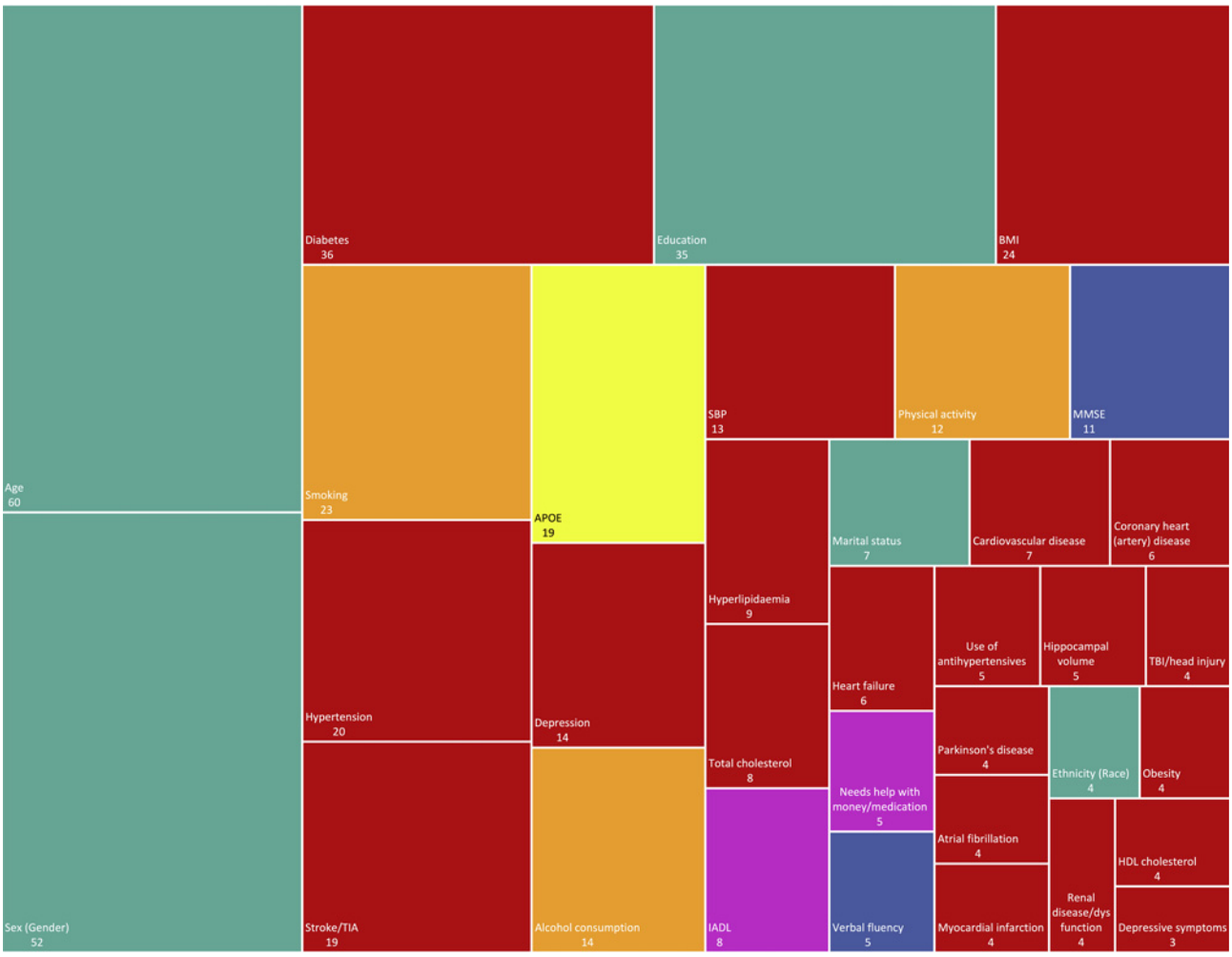
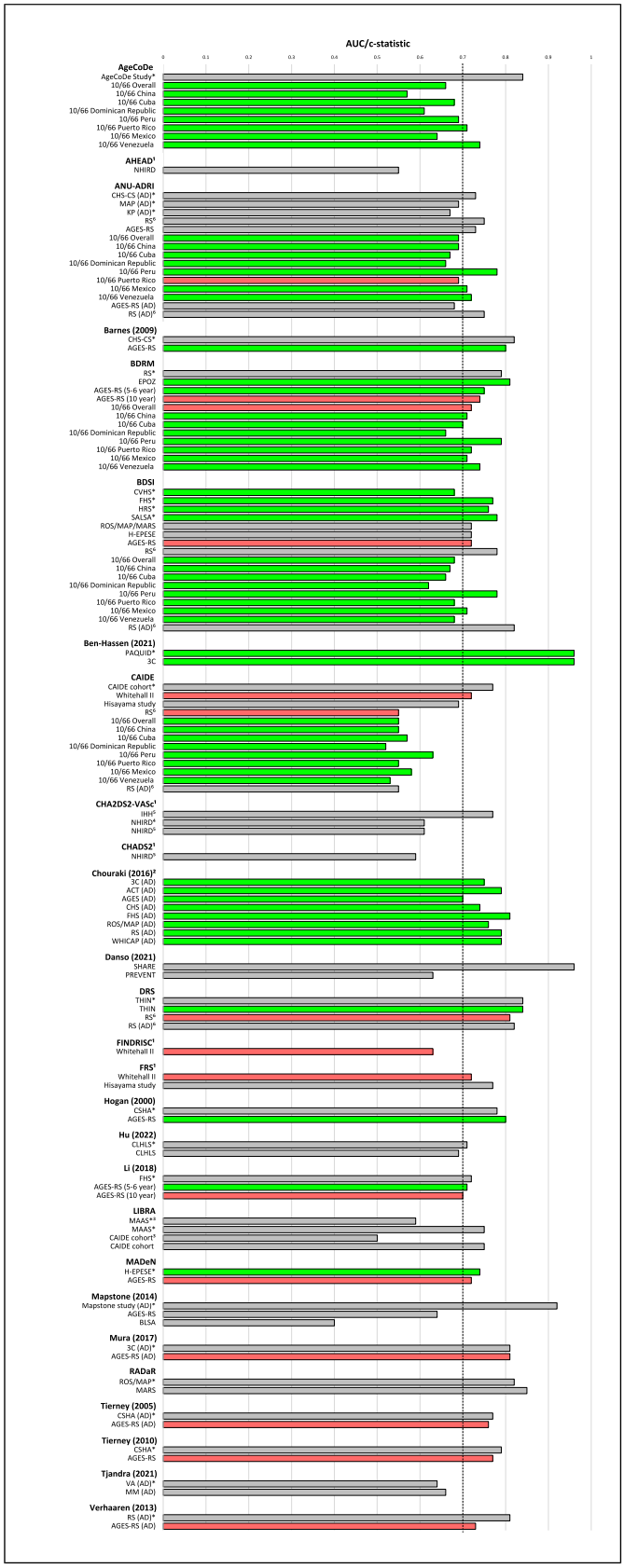



**Fig. 2.** Treemap of the most commonly used variables, with size and colour representing frequency and variable category. Key: types of variables represented: green = demographic; red = health; orange = lifestyle; pink = functioning; blue = cognitive; yellow = genetic. BMI, body mass index; HDL, high-density lipoprotein; IADL, instrumental activities of daily living; MMSE, Mini-Mental State Examination; TBI, traumatic brain injury.

*Ethnicity and Race*

Of the 74 included studies, 16 developed risk models in Asian populations using data from Japan ($n = 7$), China ($n = 3$, South Korea ($n = 3$), Taiwan ($n = 1$), and Hong Kong ($n = 1$). Only five studies developed models in LMICs, using data from Mexico [26, 27] and China [28, 48, 31].

Ethnic variability in predictive accuracy of risk scores has been investigated within African American, Caribbean, Asian, and Hispanic populations residing in the USA [40, 42, 47, 50, 84]. Developed in a predominantly White cohort, the RADaR model was externally validated in the Minority Aging Research Study (MARS; 100% Black cohort) [40]. With a prediction horizon of 3 years, the RADaR displayed excellent discriminative accuracy (AUC = 0.85) that was higher than the development cohort (AUC = 0.81), with baseline characteristics varying greatly, i.e., MARS participants having, on average, higher prevalence of diabetes and increased body mass index. Four studies explored dementia risk in Mexican [26, 27, 30] and American-Mexican [50] (participants identifying as Mexican and residing in the USA) cohorts. Using data from the Mexico 10/66 cohort, one study developed a model with "fair" discriminative performance (AUC range: 0.72–0.75), utilising variables such as neuropsychiatric symptoms, rurality, illiteracy, mild cognitive




impairment, and diabetes, with participants followed for 3 years and aged >80 years [26]. To note, this does not mean it is suitable for applications to unselected populations in LMICs. As shown in Figure 3, another study externally validated various dementia risk models using data from the 10/66 study over a 10-year follow-up [30], reporting that the BDRM, BDSI, and ANU-ADRI transported well into Mexican populations (c-statistic = 0.71), whilst transportability of the AgeCoDe and CAIDE was poor (c-statistic = 0.64 and 0.58, respectively). Another study developed the Mexican American Dementia Nomogram (MADeN) risk index that aims to predict dementia in Mexican Americans over a 10 years period [50]. The MADeN risk index included education, diabetes, age, and other novel measures such as pain and social support/engagement. Overall, the MADeN exhibited fair predictive accuracy (AUC = 0.74) with good calibration.

### *Risk Model Development and Testing in Disease-Specific Groups*

Six studies have developed and tested models in disease-specific groups including: type 2 diabetes ($n$ = 2) [69, 78], cardiovascular disease ($n$ = 2) [57, 62, 70], and stroke ($n$ = 1) [21] populations.

#### Diabetes

One study developed a model incorporating; age, sex, diabetes duration, body mass index, variation (%) fasting plasma glucose, HBa1c, stroke, hypoglycaemia, postural hypertension, coronary artery disease, and anti-diabetes medications in a sample of $n$ = 27,540 type 2 diabetic patients in Taiwan, aged 50–94 years, with a mean follow-up of 8 years [69]. The model had fair/good discriminative accuracy in the derivation (AUC range: 0.76–0.82) and internal validation (AUC range: 0.75–0.84) datasets; with shorter follow-up times having higher predictive accuracy scores (i.e., 3-year vs. 5-year and 10-year follow-up). Another study developed and validated the RxDx dementia risk index, incorporating demographics, the Charlson comorbidity index, and a chronic disease score specifically for older (≥60 years) individuals with type 2 diabetes and comorbid hypertension [78]. The RxDx risk model demonstrated excellent discriminative accuracy in both the training (c-statistic = 0.81) and validation (c-statistic = 0.86; type 2 diabetic patients without hypertension) datasets. However, calibration was poor.

#### Cardiovascular Conditions

One study investigated risk of incident dementia using the $CHA_2DS_2$-VASc and AHEAD scores in a large national database in Taiwan of $n$ = 387,595 participants with heart failure [62]. Although high $CHA_2DS_2$-VASc and AHEAD

scores were strongly associated with an increased risk of dementia over a mean of 2.9-year follow-up, model discriminative accuracy was poor (AUC range: 0.55–0.61). In the same dataset, but focused on participants with atrial fibrillation ($n$ = 332,665), $CHADS_2$ and $CHA_2DS_2$-VASc scores again showed poor discriminative accuracy (c-statistic range: 0.59–0.61) for predicting incident dementia [70]. Another study looking at dementia risk in patients with atrial fibrillation ($n$ = 74,081; followed for 10 years) in the USA, reported a higher predictive accuracy of the $CHA_2DS_2$-VASc (AUC range: 0.65–0.73), compared to the Intermountain Mortality Risk Score (IMRS) (AUC range: 0.68–0.69) [57].

Stroke

A dementia risk model incorporating demographics, the Charlson comorbidity index, and cognitive symptom burden showed poor discriminative accuracy (c-statistic = 0.59) for incident dementia in stroke patients ($n$ = 18,681) over 5-year follow-up [21].

*Sex Stratification*

Research stratifying sex frequently showed analogous discriminative performance. Over a 10-year observation period using common demographic, health, and lifestyle variables, Park et al. [85] reported consistent discriminative performance across male, female, and aggregated participants in both development and internal validation cohorts ($n$ = 331,126 using Cox regression; AUC = 0.81) with excellent calibration, as reported by both calibration slope and intercept. Two studies used Fine-Gray sub-distribution hazard modelling methods, reporting fair or good discriminative ability across sexes over 5 years ($n$ = 75,460; c-statistic = 0.83 for both men and women) [55] and 10 years ($n$ = 2,710; c-statistic ranging from 0.79 to 0.87 for men, 0.78–0.85 for women) [74]. One study reported poor discriminative accuracy across both sexes in their 10 years, 42 health and functioning variable model ($n$ = 7,239; AUC = 0.67 for men 0.66 for women) [92]. One study using the $CHA_2DS_2$-VASc and IMRS models to predict 10-year dementia risk in atrial fibrillation patients ($n$ = 74,081) found discriminative performance varied more in men (AUC range: 0.27–0.68) compared to women (AUC range: 0.59–0.77) [57].

**Fig. 3.** AUC/c-statistic indices comparing development and external validation of dementia risk models. Key: The dashed line represents the lowest cut-off value (AUC/c-statistic = 0.70) for acceptable discriminative accuracy and clinical significance. See Table 1 for a comprehensive list of abbreviations. Grey = calibration not reported; green = good calibration; red = poor calibration (underestimated/overestimated risk). *Model development cohort. [1]Risk model not initially developed for dementia risk prediction. [2]Taken from the total population at 7-year follow-up due to having calibration results. [3]LIBRA basic model. [4]Heart failure population. [5]Atrial fibrillation population. [6]Taken from 10-year follow-up due to having calibration results.

## Discussion

In this review, 74 studies were identified, adding to the 46 studies already published. Since our previous reviews in 2010 [8] and 2015 [9], there has been an increase in the development of novel dementia risk prediction models. While traditional analytical methods are still being used, more advanced computational methods (i.e., AI and ML), population-specific models targeting groups with higher risk of dementia based on disease comorbidity, such as individuals with type 2 diabetes, atrial fibrillation and stroke, and a greater focus on validation (both internal and external) is now taking place. Further, for the first time, models have been developed and tested in LMIC settings. However, model performance indices, captured in the AUC/c-statistic, remain variable (0.50–0.96) and not all models show good transportability/generalisability, particularly when external validation is tested. Moreover, calibration was rarely undertaken, making it difficult to ensure the predicted probabilities of developing dementia align closely with the observed probabilities, and whether risk models may be under- or overestimating risk. Although some models show acceptable transportability, it is still not possible to recommend one model for use in clinical or population-based settings without further robust validation, as highlighted by other reviews [9, 114]. However, there are several models that show promise, such as the BDSI, BDRM, and RADaR that could be the focus of further investment, as the proliferation of developed risk models may further fragment the evidence base.

The inconsistent adherence to established reporting guidelines such as TRIPOD and STROBE raises concerns about the completeness and transparency of the reported methods and results. Adherence to guidelines like TRIPOD and STROBE is crucial for ensuring that studies are reported with sufficient detail to allow for critical appraisal and replication. Future research in this area should aim to adhere to these guidelines to enhance the quality and reliability of prognostic models. Below, Table 2 outlines standards risk prediction should be held at, in addition to pre-existing guidelines.

*Risk Score Components and Risk Calculation*

Demographic, health, and lifestyle variables are often incorporated into dementia risk prediction models. This resulted in substantial overlap between model components, e.g., the 12 factors outlined in the Lancet Commission on

**Table 2.** An overview of standards for utilising dementia risk prediction models at the individual and population-based level (information adapted from the PHG Foundation's dementia risk prediction models report)

| | |
|---|---|
| **Model characteristics and reporting** | |
| Accuracy and uncertainty | Models must have sufficient accuracy, validity, and acceptability for the purpose of their use. In particular, rates of false positives (patients classed as at high risk who will never develop dementia) and false negatives (patients classed as low risk who go on the develop dementia) must be at acceptably low levels, which may vary depending upon the clinical application |
| Reporting | Models should be reported with standardized performance metrics, detailing vital information about the target population. Adhering to guidelines, such as the TRIPOD statement, can be a useful reference for this standardisation to ensure transparency and reproducibility |
| **Model implementation** | |
| Context-specific considerations | Models should be relevant to the context in which they are to be applied, taking account of the patient population (age, gender, socio-economic group, ethnicity, country, etc.), the clinical context (population screening, primary care consultations, geriatric assessments, etc.), and the data available. As such they may need context-specific calibration |
| Promote health and equality | If implemented, models should not be applied in ways that exacerbate existing health disparities across and within populations |
| Timeframe of risk | Models should estimate risk for clinically relevant timeframes (e.g., short-medium-long-term risk), dependent upon the context of application |
| Education | Prior to implementation, all stakeholders (e.g., patients, healthcare providers) must be provided with accessible information – and training where appropriate – on dementia risk (and protective factors), the interpretation of risk scores, and the steps that individuals can take to reduce their risk |
| Pre- and post-risk identification | Before risk identification, pre-risk counselling should be offered to individuals to understand the implications of receiving risk scores. This counselling should align with best practices from genetic counselling environments, ensuring informed consent, especially if invasive investigations inform the model's calculation<br>For post-risk identification, clarity is needed on how risk will be communicated, who calculates it, the subsequent referral pathways, and transparency in risk score use, such as potential implications on health insurance |
| Ethical implications | Issues relating to potential bias, fairness, and discrimination must be considered in the model development and resultant risk scores |
| Cost-effectiveness | Evidence should be sought that implementing a system for dementia risk assessment, and incorporating the model, will produce value for money |
| Public health impact | Use of such models must lead to improved health outcomes that manifest for individuals and populations, i.e., reduction in burden of disease and costs associated with dementia |

Dementia Prevention and Care including: early life low educational attainment, midlife obesity, alcohol consumption, and hypertension, and later life diabetes and smoking [112]. Importantly, the risk modelling literature expands on these 12 factors and includes over 450 predictor variables such as cerebrovascular diseases (e.g., stroke), cardiometabolic conditions and their risk factors (e.g., hyperlipidaemia, heart failure, atrial fibrillation, and coronary heart disease), genetics, and cognitive function (e.g., verbal fluency).

The use of ML/AI methods in large datasets has allowed for more intensive testing of a larger number of variables. Indeed, in one study over 10,000 predictor variables were used in the training model matrix with 50 predictors incorporated into the final model [83]. This exploration of the number of predictors tested could potentially help identify new mechanisms of disease and inform the development of novel risk reduction and prevention strategies. However, it is crucial to evaluate whether ML/AI models outperform those models developed using traditional methods and more restricted risk predictors. Specifically, the advantages of a greater number of predictor variables used in ML/AI models are uncertain when balanced against the time and cost of obtaining data, overfitting and the consequent lack of transportability, and model predictive accuracy. Indeed, our review highlights that most AI/ML models did not outperform those developed using traditional statistical methods.

As in our previous reviews [8, 9], some models include time and cost-intensive variables such as blood-based biomarkers, genetics, cerebrospinal fluid analysis, and brain imaging. However, as highlighted in one study such information may only add limited predictive value at huge infrastructure and personnel costs [93]. It is also possible that different risk factors and the way they are combined (or scored) should consider the setting in which risk modelling is being undertaken. Indeed, an attempt to transport models developed in high-income country settings to LMICs in one study showed poor external validation for most models tested with the exception of the ANU-ADRI, BDSI, and BDRM [30]. Different risk factors may require testing for LMICs (e.g., illiteracy, food insecurity, and infectious disease profiles) which may be less applicable to high-income settings.

*Risk Modelling within Populations*

In studies that have considered the impact of ethnicity; namely African American, Caribbean, Asian, and Hispanic populations in the USA [40, 42, 47, 50, 84], it has been observed that olfactory variables, may require adaptation to be culturally appropriate for the studied population [47]. In contrast, the RADaR model, incorporating common demographic, health, cognitive, and genetic, and functioning variables, showed similar performance in White versus Black populations [40]. Risk models either developed or validated in Mexican populations (also residing in the USA) often showed fair predictive discriminative performance, such as the MADeN model that was exclusively developed for use in Mexican populations [27, 50], as well as the ANU-ADRI, BDSI, and BDRM [30]. Although very few studies explored sex differences in dementia risk, the majority that undertook this demonstrated comparable results, with both sexes presenting a similar level of discriminative performance suggesting robust reliability and validity across both sexes [55, 74, 85, 92].

*Disease-Specific Modelling*

There has been an increase in the number of models developed in disease-specific populations including individuals with stroke, atrial fibrillation, heart failure, and type 2 diabetes. These typically incorporate risk factors identified in the whole population as well as disease-specific variables such as anti-diabetic medication, hypoglycaemia, and renal parameters. Parallel to whole population findings, the accuracy of disease-specific models varies. Generally higher accuracy is observed in cardiometabolic-specific versus stroke-specific models. Difficulties with predicting post-stroke dementia could be due heterogeneity in stroke subtype, impacted brain region, pre-stroke cognition or other yet undetermined risk factors specific to this population, all of which could differentially influence risk models.

One study tested whether models developed in the general population transported well to stroke populations. The results showed poor performance across all tested risk models, indicating that future work is needed to develop dementia risk prediction models specifically in stroke populations [115]. Two studies investigated whether scores originally developed to predict long-term mortality in European heart failure patients or stroke risk in atrial fibrillation [62, 70] were predictive of incident dementia in heart failure and atrial fibrillation populations. The models also showed poor discriminative performance. Overall, the results highlight the need to develop dementia risk models tailored to specific disease populations [116, 117].

*Strengths and Limitations*

The review has several strengths. Firstly, the systematic approach, using the same protocol as our previous reviews, ensures that all relevant literature was captured, and the results are comparable across publications. Secondly, no language restrictions were applied thereby reducing the bias associated with using articles only published in the English language. There are also some limitations. Due to considerable heterogeneity across studies, e.g., in the predictor variables selected, follow-up times, dementia outcome use, data used, analytical methods, and sample demographics, only a narrative synthesis of the results was possible. Further, few studies used population-representative data thereby limiting generalisability of the results. Limited testing across settings (high-income country vs. LMIC) and subgroups within society (e.g., stratified by ethnicity, sex and socio-economic status) also raises questions around transportability/applicability. Unfortunately, we were unable to report some model performance metrics where authors presented the data graphically, preventing determination of specific figures. It is also important to note that while some studies reported good calibration, calibration assessments based solely on the Hosmer-Lemeshow test may be misleading. As such, calibration should also be evaluated using calibration plots or tables comparing predicted versus observed outcome frequencies, as the Hosmer-Lemeshow test alone has limited suitability for assessing calibration, particularly in large datasets [118].

*Implications and Conclusion*

While the growing number of dementia risk models reflect the increasing interest and investment into the area, our systematic review shows that there is still no one model suitable for use in routine clinical practice or the

general population. It is essential to now validate a limited number of high-quality, context-specific models to determine whether they are useful if implemented in diverse populations. Once the clinical validity and effectiveness of risk prediction models are confirmed, further analyses on cost-effectiveness and clinical utility should be undertaken. This includes evaluating potential harms, patient, and public acceptability of such tools, the implications of any such effort for health systems to address questions around who would be responsible for undertaking scoring and risk communication, and how model outcomes could inform "personalised medicine" plans [119]. Moreover, population-based strategies, in tandem with personalised medicine approaches, may provide a more holistic strategy to preventing dementia by not only having the potential to significantly reduce the overall prevalence of dementia but may also help address health disparities by ensuring that dementia prevention efforts are built in to society, and those that are individually orientated are appropriate and accessible [120, 121]. Finally, there needs to be a consensus on the most well-developed models and focussing research resources on improving the clinical validity and utility of these models as a clinical risk prediction tool for application across diverse communities in the general population.

## Statement of Ethics

A statement of ethics is not applicable as this study is based exclusively on published literature.

## Conflict of Interest Statement

The authors declare no conflict of interest.


## Funding Sources

This work was supported by the NIHR (Three Schools: Dementia Research Program, Improving the lives of people living with dementia and carers 2021–2024; l Grant reference #DP06) and UKRI (Research Grant: MR/X005437/1). The authors declare no conflict of interest.


## Author Contributions

R.K.A., M.B., C.B., G.F., L.R., M.S., E.P., C.M., D.R., N.Q., and B.C.M.S. conceptualised and helped in the funding acquisition. L.E. refined the search strategy. J.B., A.H.K., R.K., and B.C.M.S. undertook title, abstract, and full-text screening. J.B., E.T., and B.C.M.S. developed the methodology. J.B., E.Y.H.T., and R.K. performed data extraction. J.B. and L.G. undertook the risk of bias assessment. J.B. curated the data, undertook project administration and data synthesis, and prepared the original draft. B.C.M.S, J.L., A.S., P.J.T, and D.T. supervised the project. All authors reviewed and edited the final manuscript.

## Data Availability Statement

All data supporting the findings of this study are available within the article and its supplementary materials. Further enquiries can be directed to the corresponding author.

## References


1 Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;357:j2099. https://doi.org/10.1136/bmj.j2099

2 Hageman S, Pennells L, Ojeda F, Kaptoge S, Kuulasmaa K, De Vries T, et al. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur Heart J. 2021;42(25):2439–54. https://doi.org/10.1093/eurheartj/ehab309

3 Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. BMJ. 2013;346(may02 1):f2573–73. https://doi.org/10.1136/bmj.f2573

4 Parmar P, Krishnamurthi R, Ikram MA, Hofman A, Mirza SS, Varakin Y, et al. The stroke Riskometer™ app: validation of a data collection tool and stroke risk predictor. Int J Stroke. 2015;10(2):231–44. https://doi.org/10.1111/ijs.12411

5 The Lancet Public Health. Reinvigorating the public health response to dementia. Lancet Public Health. 2021;6(10):e696. https://doi.org/10.1016/s2468-2667(21)00215-2

6 America GSo. The Gerontological Society of America workgroup on cognitive impairment detection and earlier diagnosis: report and recommendations; 2015.

7 Pottie K, Rahal R, Jaramillo A, Birtwhistle R, Thombs BD, Singh H, et al. Recommendations on screening for cognitive impairment in older adults. CMAJ. 2016;188(1):37–46. https://doi.org/10.1503/cmaj.141165

8 Stephan BCM, Kurth T, Matthews FE, Brayne C, Dufouil C. Dementia risk prediction in the population: are screening models accurate? Nat Rev Neurol. 2010;6(6):318–26. https://doi.org/10.1038/nrneurol.2010.54

9 Tang EY, Harrison SL, Errington L, Gordon MF, Visser PJ, Novak G, et al. Current developments in dementia risk prediction modelling: an updated systematic review. PLoS One. 2015;10(9):e0136181. https://doi.org/10.1371/journal.pone.0136181

10 Hou X-H, Feng L, Zhang C, Cao X-P, Tan L, Yu J-T. Models for predicting risk of dementia: a systematic review. J Neurol Neurosurg Psychiatry. 2019;90(4):373–9. https://doi.org/10.1136/jnnp-2018-318212

11 John LH, Kors JA, Fridgeirsson EA, Reps JM, Rijnbeek PR. External validation of existing dementia prediction models on observational health data. BMC Med Res Methodol. 2022;22(1):311. https://doi.org/10.1186/s12874-022-01793-5

12 Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg. 2021;88:105906. https://doi.org/10.1016/j.ijsu.2021.105906

13 Hupe M. EndNote X9. J Electron Resour Med Libraries. 2019;16(3–4):117–9. https://doi.org/10.1080/15424065.2019.1691963
14 Gumus M, Multani N, Mack ML, Tartaglia MC. Progression of neuropsychiatric symptoms in young-onset versus late-onset Alzheimer's disease. GeroScience. 2021;43(1):213–23. https://doi.org/10.1007/s11357-020-00304-y
15 Hajian-Tilaki K. Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med. 2013;4(2):627.
16 Corbacioglu Ş, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: a guide to interpreting the area under the curve value. Turk J Emerg Med. 2023;23(4):195. https://doi.org/10.4103/tjem.tjem_182_23
17 Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan: a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. https://doi.org/10.1186/s13643-016-0384-4
18 Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Oxford; 2000.
19 Casanova R, Varma S, Simpson B, Kim M, An Y, Saldana S, et al. Blood metabolite markers of preclinical Alzheimer's disease in two longitudinally followed cohorts of older individuals. Alzheimers Dement. 2016;12(7):815–22. https://doi.org/10.1016/j.jalz.2015.12.008
20 Walters K, Hardoon S, Petersen I, Iliffe S, Omar RZ, Nazareth I, et al. Predicting dementia risk in primary care: development and validation of the Dementia Risk Score using routinely collected data. BMC Med. 2016;14(1):6. https://doi.org/10.1186/s12916-016-0549-y
21 McCoy TH, Han L, Pellegrini AM, Tanzi RE, Berretta S, Perlis RH. Stratifying risk for dementia onset using large-scale electronic health record data: a retrospective cohort study. Alzheimers Dement. 2020;16(3):531–40. https://doi.org/10.1016/j.jalz.2019.09.084
22 Vonk JMJ, Greving JP, Gudnason V, Launer LJ, Geerlings MI. Dementia risk in the general population: large-scale external validation of prediction models in the AGES-Reykjavik study. Eur J Epidemiol. 2021;36(10):1025–41. https://doi.org/10.1007/s10654-021-00785-x
23 Wyman-Chick KA, O'Keefe LR, Weintraub D, Armstrong MJ, Rosenbloom M, Martin PK, et al. Prodromal dementia with Lewy bodies: evolution of symptoms and predictors of dementia onset. J Geriatr Psychiatry Neurol. 2022;35(4):527–34. https://doi.org/10.1177/08919887211023586
24 Li J, Ogrodnik M, Kolachalama VB, Lin H, Au R. Assessment of the mid-life demographic and lifestyle risk factors of dementia using data from the framingham heart study offspring cohort. J Alzheimers Dis. 2018;63(3):1119–27. https://doi.org/10.3233/jad-170917
25 Vuoksimaa E, Rinne JO, Lindgren N, Heikkilä K, Koskenvuo M, Kaprio J. Middle age self-report risk score predicts cognitive functioning and dementia in 20–40 years. Alzheimers Dement (Amst). 2016;4(1):118–25. https://doi.org/10.1016/j.dadm.2016.08.003
26 Acosta I, Borges G, Aguirre-Hernandez R, Sosa AL, Prince M. Neuropsychiatric symptoms as risk factors of dementia in a Mexican population: a 10/66 Dementia Research Group study. Alzheimers Dement. 2018;14(3):271–9. https://doi.org/10.1016/j.jalz.2017.08.015
27 Downer B, Veeranki SP, Wong R. A late life risk index for severe cognitive impairment in Mexico. J Alzheimers Dis. 2016;52(1):191–203. https://doi.org/10.3233/jad-150702
28 Hu M, Shu X, Yu G, Wu X, Välimäki M, Feng H. A risk prediction model based on machine learning for cognitive impairment among Chinese community-dwelling elderly people with normal cognition: development and validation study. J Med Internet Res. 2021;23(2):e20298. https://doi.org/10.2196/20298
29 Hu M, Gao Y, Kwok TCY, Shao Z, Xiao LD, Feng H. Derivation and validation of the cognitive impairment prediction model in older adults: a national cohort study. Front Aging Neurosci. 2022;14:755005. https://doi.org/10.3389/fnagi.2022.755005
30 Stephan BCM, Pakpahan E, Siervo M, Licher S, Muniz-Terrera G, Mohan D, et al. Prediction of dementia risk in low-income and middle-income countries (the 10/66 Study): an independent external validation of existing models. Lancet Glob Health. 2020;8(4):e524–35. https://doi.org/10.1016/s2214-109x(20)30062-0
31 Wang C, Ji X, Wu X, Tang Z, Zhang X, Guan S, et al. Frailty in relation to the risk of Alzheimer's disease, dementia, and death in older Chinese adults: a seven-year prospective study. J Nutr Health Aging. 2017;21(6):648–54. https://doi.org/10.1007/s12603-016-0798-7
32 Stocker H, Perna L, Weigl K, Möllers T, Schöttker B, Thomsen H, et al. Prediction of clinical diagnosis of Alzheimer's disease, vascular, mixed, and all-cause dementia by a polygenic risk score and APOE status in a community-based cohort prospectively followed over 17 years. Mol Psychiatry. 2021;26(10):5812–22. https://doi.org/10.1038/s41380-020-0764-y
33 Stocker H, Nabers A, Perna L, Möllers T, Rujescu D, Hartmann AM, et al. Genetic predisposition, Aβ misfolding in blood plasma, and Alzheimer's disease. Transl Psychiatry. 2021;11(1):261. https://doi.org/10.1038/s41398-021-01380-0
34 Phongpreecha T, Cholerton B, Mata IF, Zabetian CP, Poston KL, Aghaeepour N, et al. Multivariate prediction of dementia in Parkinson's disease. Npj Parkinsons Dis. 2020;6(1):20. https://doi.org/10.1038/s41531-020-00121-2
35 Achterberg HC, Van Der Lijn F, Den Heijer T, Vernooij MW, Ikram MA, Niessen WJ, et al. Hippocampal shape is predictive for the development of dementia in a normal, elderly population. Hum Brain Mapp. 2014;35(5):2359–71. https://doi.org/10.1002/hbm.22333
36 Achterberg HC, Sørensen L, Wolters FJ, Niessen WJ, Vernooij MW, Ikram MA, et al. The value of hippocampal volume, shape, and texture for 11-year prediction of dementia: a population-based study. Neurobiol Aging. 2019;81:58–66. https://doi.org/10.1016/j.neurobiolaging.2019.05.007
37 Beebe-Wang N, Okeson A, Althoff T, Lee SI. Efficient and explainable risk assessments for imminent dementia in an aging cohort study. IEEE J Biomed Health Inform. 2021;25(7):2409–20. https://doi.org/10.1109/jbhi.2021.3059563
38 Ben-Hassen C, Helmer C, Berr C, Jacqmin-Gadda H. Five-year dynamic prediction of dementia using repeated measures of cognitive tests and a dependency scale. Am J Epidemiol. 2021;191(3):453–64. https://doi.org/10.1093/aje/kwab269
39 Calvin CM, Wilkinson T, Starr JM, Sudlow C, Hagenaars SP, Harris SE, et al. Predicting incident dementia 3-8 years after brief cognitive tests in the UK Biobank prospective study of 500,000 people. Alzheimers Dement. 2019;15(12):1546–57. https://doi.org/10.1016/j.jalz.2019.07.014
40 Capuano AW, Shah RC, Blanche P, Wilson RS, Barnes LL, Bennett DA, et al. Derivation and validation of the Rapid Assessment of Dementia Risk (RADaR) for older adults. PLoS One. 2022;17(3):e0265379. https://doi.org/10.1371/journal.pone.0265379
41 Mapstone M, Cheema AK, Fiandaca MS, Zhong X, Mhyre TR, MacArthur LH, et al. Plasma phospholipids identify antecedent memory impairment in older adults. Nat Med. 2014;20(4):415–8. https://doi.org/10.1038/nm.3466
42 Chouraki V, Reitz C, Maury F, Bis JC, Bellenguez C, Yu L, et al. Evaluation of a genetic risk score to improve risk prediction for Alzheimer's disease. J Alzheimers Dis. 2016;53(3):921–32. https://doi.org/10.3233/jad-150749
43 Dallora AL, Minku L, Mendes E, Rennemark M, Anderberg P, Sanmartin Berglund J. Multifactorial 10-year prior diagnosis prediction model of dementia. Int J Environ Res Publ Health. 2020;17(18):6674. https://doi.org/10.3390/ijerph17186674
44 Danso SO, Zeng Z, Muniz-Terrera G, Ritchie CW. Developing an explainable machine learning-based personalised dementia risk prediction model: a transfer learning approach with ensemble learning algorithms. Front Big Data. 2021;4:613047. https://doi.org/10.3389/fdata.2021.613047

45 de la Fuente J, Moreno-Agostino D, de la Torre-Luque A, Prina AM, Haro JM, Caballero FF, et al. Development of a combined sensory-cognitive measure based on the common cause hypothesis: heterogeneous trajectories and associated risk factors. Gerontologist. 2020;60(5):e357–66. https://doi.org/10.1093/geront/gnz066
46 Deckers K, Barbera M, Köhler S, Ngandu T, van Boxtel M, Rusanen M, et al. Long-term dementia risk prediction by the LIBRA score: A 30-year follow-up of the CAIDE study. Int J Geriatr Psychiatry. 2020;35(2):195–203. https://doi.org/10.1002/gps.5235.
47 Devanand D, Lee S, Manly J, Andrews H, Schupf N, Doty RL, et al. Olfactory deficits predict cognitive decline and Alzheimer dementia in an urban community. Neurology. 2015;84(2):182–9. https://doi.org/10.1212/wnl.0000000000001132
48 Ding D, Liang X, Xiao Z, Wu W, Zhao Q, Cao Y. Can dementia be predicted using olfactory identification test in the elderly? A Bayesian network analysis. Brain Behav. 2020;10(11):e01822. https://doi.org/10.1002/brb3.1822
49 Barnes DE, Beiser AS, Lee A, Langa KM, Koyama A, Preis SR, et al. Development and validation of a brief dementia screening indicator for primary care. Alzheimers Dement. 2014;10(6):656–65.e1. https://doi.org/10.1016/j.jalz.2013.11.006
50 Downer B, Kumar A, Veeranki SP, Mehta HB, Raji M, Markides KS. Mexican-American dementia Nomogram: development of a dementia risk index for Mexican-American older adults. J Am Geriatr Soc. 2016;64(12):e265–69. https://doi.org/10.1111/jgs.14531
51 Fayosse A, Nguyen D-P, Dugravot A, Dumurgier J, Tabak AG, Kivimäki M, et al. Risk prediction models for dementia: role of age and cardiometabolic risk factors. BMC Med. 2020;18(1):107. https://doi.org/10.1186/s12916-020-01578-x
52 Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. Lancet Neurol. 2006;5(9):735–41. https://doi.org/10.1016/s1474-4422(06)70537-3
53 Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care. 2003;26(3):725–31. https://doi.org/10.2337/diacare.26.3.725
54 D'Agostino RB Sr, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117(6):743–53. https://doi.org/10.1161/circulationaha.107.699579
55 Fisher S, Manuel DG, Hsu AT, Bennett C, Tuna M, Bader Eddeen A, et al. Development and validation of a predictive algorithm for risk of dementia in the community setting. J Epidemiol Community Health. 2021;75(9):843–53. https://doi.org/10.1136/jech-2020-214797
56 Fukunishi H, Nishiyama M, Luo Y, Kubo M, Kobayashi Y. Alzheimer-type dementia prediction by sparse logistic regression using claim data. Comput Methods Programs Biomed. 2020;196:105582. https://doi.org/10.1016/j.cmpb.2020.105582
57 Graves KG, May HT, Jacobs V, Knowlton KU, Muhlestein JB, Lappe DL, et al. CHA(2)DS(2)-VASc scores and Intermountain Mortality Risk Scores for the joint risk stratification of dementia among patients with atrial fibrillation. Heart Rhythm. 2019;16(1):3–9. https://doi.org/10.1016/j.hrthm.2018.10.018
58 Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest. 2010;137(2):263–72. https://doi.org/10.1378/chest.09-1584
59 Horne BD, May HT, Kfoury AG, Renlund DG, Muhlestein JB, Lappé DL, et al. The Intermountain Risk Score (including the red cell distribution width) predicts heart failure and other morbidity endpoints. Eur J Heart Fail. 2010;12(11):1203–13. https://doi.org/10.1093/eurjhf/hfq115
60 Hall A, Pekkala T, Polvikoski T, van Gils M, Kivipelto M, Lötjönen J, et al. Prediction models for dementia and neuropathology in the oldest old: the Vantaa 85+ cohort study. Alzheimers Res Ther. 2019;11(1):11. https://doi.org/10.1186/s13195-018-0450-3
61 Honda T, Ohara T, Yoshida D, Shibata M, Ishida Y, Furuta Y, et al. Development of a dementia prediction model for primary care: the Hisayama Study. Alzheimers Dement (Amst). 2021;13(1):e12221. https://doi.org/10.1002/dad2.12221
62 Hu W-S, Lin C-L. Comparison of CHA2DS2-VASc and AHEAD scores for the prediction of incident dementia in patients hospitalized for heart failure: a nationwide cohort study. Intern Emerg Med. 2019;14(3):395–402. https://doi.org/10.1007/s11739-018-1961-4
63 Spinar J, Jarkovsky J, Spinarova L, Mebazaa A, Gayat E, Vitovec J, et al. AHEAD score: long-term risk classification in acute heart failure. Int J Cardiol. 2016;202:21–6. https://doi.org/10.1016/j.ijcard.2015.08.187
64 Ibarrondo O, Huerta JM, Amiano P, Andreu-Reinón ME, Mokoroa O, Ardanaz E, et al. Dementia risk score for a population in southern Europe calculated using competing risk models. J Alzheimers Dis. 2022;86(4):1751–62. https://doi.org/10.3233/jad-215211
65 Jacqmin-Gadda H, Blanche P, Chary E, Loubere L, Amieva H, Dartigues J-F. Prognostic score for predicting risk of dementia over 10 years while accounting for competing risk of death. Am J Epidemiol. 2014;180(8):790–8. https://doi.org/10.1093/aje/kwu202
66 Kim WJ, Sung JM, Sung D, Chae MH, An SK, Namkoong K, et al. Cox proportional hazard regression versus a deep learning algorithm in the prediction of dementia: an analysis based on periodic health examination. JMIR Med Inform. 2019;7(3):e13139. https://doi.org/10.2196/13139
67 Kochan NA, Bunce D, Pont S, Crawford JD, Brodaty H, Sachdev PS. Reaction time measures predict incident dementia in community-living older adults: the Sydney memory and ageing study. Am J Geriatr Psychiatry. 2016;24(3):221–31. https://doi.org/10.1016/j.jagp.2015.12.005
68 Li J, Ogrodnik M, Devine S, Auerbach S, Wolf PA, Au R. Practical risk score for 5-10- and 20-year prediction of dementia in elderly persons: framingham Heart Study. Alzheimers Dement. 2018;14(1):35–42. https://doi.org/10.1016/j.jalz.2017.04.013
69 Li C-I, Li T-C, Liu C-S, Liao L-N, Lin W-Y, Lin C-H, et al. Risk score prediction model for dementia in patients with type 2 diabetes. Euro J Neurol. 2018;25(7):976–83. https://doi.org/10.1111/ene.13642
70 Liao J-N, Chao T-F, Liu C-J, Wang K-L, Chen S-J, Tuan T-C, et al. Risk and prediction of dementia in patients with atrial fibrillation: a nationwide population-based cohort study. Int J Cardiol. 2015;199:25–30. https://doi.org/10.1016/j.ijcard.2015.06.170
71 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting StrokeResults from the national Registry of atrial fibrillation. JAMA. 2001;285(22):2864–70. https://doi.org/10.1001/jama.285.22.2864
72 Licher S, Yilmaz P, Leening MJ, Wolters FJ, Vernooij MW, Stephan BC, et al. External validation of four dementia prediction models for use in the general community-dwelling population: a comparative analysis from the Rotterdam Study. Eur J Epidemiol. 2018;33(7):645–55. https://doi.org/10.1007/s10654-018-0403-y
73 Anstey KJ, Cherbuin N, Herath PM, Qiu C, Kuller LH, Lopez OL, et al. A self-report risk index to predict occurrence of dementia in three independent cohorts of older adults: the ANU-ADRI. PLoS One. 2014;9(1):e86141. https://doi.org/10.1371/journal.pone.0086141
74 Licher S, Leening MJ, Yilmaz P, Wolters FJ, Heeringa J, Bindels PJ, et al. Development and validation of a dementia risk prediction model in the general population: an analysis of three longitudinal studies. Am J Psychiatr. 2019;176(7):543–51. https://doi.org/10.1176/appi.ajp.2018.18050566
75 Lin HR, Tsuji T, Kondo K, Imanaka Y. Development of a risk score for the prediction of incident dementia in older adults using a frailty index and health checkup data: the JAGES longitudinal study. Prev Med. 2018;112:88–96. https://doi.org/10.1016/j.ypmed.2018.04.004

76 Lin H, Himali JJ, Satizabal CL, Beiser AS, Levy D, Benjamin EJ, et al. Identifying blood biomarkers for dementia using machine learning methods in the framingham heart study. Cells. 2022;11(9):1506. https://doi.org/10.3390/cells11091506

77 Makino K, Lee S, Bae S, Chiba I, Harada K, Katayama O, et al. Development and validation of new screening tool for predicting dementia risk in community-dwelling older Japanese adults. J Transl Med. 2021;19(1):448. https://doi.org/10.1186/s12967-021-03121-9

78 Mehta HB, Mehta V, Tsai C-L, Chen H, Aparasu RR, Johnson ML. Development and validation of the RxDx-dementia risk index to predict dementia in patients with type 2 diabetes and hypertension. J Alzheimers Dis. 2015;49:423–32. https://doi.org/10.3233/jad-150466

79 Müller T, Payton NM, Kalpouzos G, Jessen F, Grande G, Bäckman L, et al. Cognitive, genetic, brain volume, and diffusion tensor imaging markers as early indicators of dementia. J Alzheimers Dis. 2020;77(4):1443–53. https://doi.org/10.3233/jad-200445

80 Mura T, Baramova M, Gabelle A, Artero S, Dartigues J-F, Amieva H, et al. Predicting dementia using socio-demographic characteristics and the Free and Cued Selective Reminding Test in the general population. Alzheimers Res Ther. 2017;9(1):21. https://doi.org/10.1186/s13195-016-0230-x

81 Nagata T, Ohara T, Hata J, Sakata S, Furuta Y, Yoshida D, et al. NT-proBNP and risk of dementia in a general Japanese elderly population: the hisayama study. J Am Heart Assoc. 2019;8(17):e011652. https://doi.org/10.1161/jaha.118.011652

82 Nakazawa T, Ohara T, Hirabayashi N, Furuta Y, Hata J, Shibata M, et al. Multiple-region grey matter atrophy as a predictor for the development of dementia in a community: the Hisayama Study. J Neurol Neurosurg Psychiatry. 2022;93(3):263–71. https://doi.org/10.1136/jnnp-2021-326611

83 Nori VS, Hane CA, Martin DC, Kravetz AD, Sanghavi DM. Identifying incident dementia by applying machine learning to a very large administrative claims dataset. PLoS One. 2019;14(7):e0203246. https://doi.org/10.1371/journal.pone.0203246

84 Ostovaneh MR, Moazzami K, Yoneyama K, A Venkatesh B, Heckbert SR, Wu CO, et al. Change in NT-proBNP (N-terminal pro-B-type natriuretic peptide) level and risk of dementia in Multi-Ethnic Study of Atherosclerosis (MESA). Hypertension. 2020;75(2):316–23. https://doi.org/10.1161/hypertensionaha.119.13952

85 Park KM, Sung JM, Kim WJ, An SK, Namkoong K, Lee E, et al. Population-based dementia prediction model using Korean public health examination data: a cohort study. PLoS One. 2019;14(2):e0211957. https://doi.org/10.1371/journal.pone.0211957

86 Park JH, Cho HE, Kim JH, Wall MM, Stern Y, Lim H, et al. Machine learning prediction of incidence of Alzheimer's disease using large-scale administrative health data. NPJ Digit Med. 2020;3(1):46. https://doi.org/10.1038/s41746-020-0256-0

87 Paul KC, Debes F, Eliasen E, Weihe P, Petersen MS. Incidence, gender influence, and neuropsychological predictors of all cause dementia in the Faroe Islands: the Faroese Septuagenarian cohort. Aging Clin Exp Res. 2021;33(1):105–14. https://doi.org/10.1007/s40520-020-01520-4

88 Payton NM, Kalpouzos G, Rizzuto D, Fratiglioni L, Kivipelto M, Bäckman L, et al. Combining cognitive, genetic, and structural neuroimaging markers to identify individuals with increased dementia risk. J Alzheimers Dis. 2018;64(2):533–42. https://doi.org/10.3233/jad-180199

89 Payton NM, Rizzuto D, Fratiglioni L, Kivipelto M, Bäckman L, Laukka EJ. Combining cognitive markers to identify individuals at increased dementia risk: influence of modifying factors and time to diagnosis. J Int Neuropsychol Soc. 2020;26(8):785–97. https://doi.org/10.1017/s1355617720000272

90 Reinke C, Doblhammer G, Schmid M, Welchowski T. Dementia risk predictions from German claims data using methods of machine learning. Alzheimers Dement. 2023;19(2):477–86. https://doi.org/10.1002/alz.12663

91 Schiepers OJG, Köhler S, Deckers K, Irving K, O'Donnell CA, van den Akker M, et al. Lifestyle for Brain Health (LIBRA): a new model for dementia prevention. Int J Geriat Psychiatry. 2018;33(1):167–75. https://doi.org/10.1002/gps.4700

92 Song X, Mitnitski A, Rockwood K. Age-related deficit accumulation and the risk of late-life dementia. Alzheimers Res Ther. 2014;6(5–8):54. https://doi.org/10.1186/s13195-014-0054-5

93 Stephan BCM, Tzourio C, Auriacombe S, Amieva H, Dufouil C, Alpérovitch A, et al. Usefulness of data from magnetic resonance imaging to improve prediction of dementia: population based cohort study. BMJ Br Med J. 2015;350(jun22 8):h2863. https://doi.org/10.1136/bmj.h2863

94 Jessen F, Wiese B, Bickel H, Eiffländer-Gorfer S, Fuchs A, Kaduszkiewicz H, et al. Prediction of dementia in primary care patients. PLoS One. 2011;6(2):e16852. https://doi.org/10.1371/journal.pone.0016852

95 Stocker H, Nabers A, Perna L, Möllers T, Rujescu D, Hartmann A, et al. Prediction of Alzheimer's disease diagnosis within 14 years through Aβ misfolding in blood plasma compared to *APOE4* status, and other risk factors. Alzheimers Dement. 2020;16(2):283–91. https://doi.org/10.1016/j.jalz.2019.08.189

96 Tjandra D, Migrino RQ, Giordani B, Wiens J. Cohort discovery and risk stratification for Alzheimer's disease: an electronic health record-based approach. AD Transl Res Clin Interv. 2020;6(1):e12035. https://doi.org/10.1002/trc2.12035

97 Tjandra D, Migrino RQ, Giordani B, Wiens J. Use of blood pressure measurements extracted from the electronic health record in predicting Alzheimer's disease: a retrospective cohort study at two medical centers. Alzheimers Dement. 2022;18(11):2368–72. https://doi.org/10.1002/alz.12676

98 Tomata Y, Sugiyama K, Kaiho Y, Sugawara Y, Hozawa A, Tsuji I. Predictive ability of a simple subjective memory complaints scale for incident dementia: evaluation of Japan's national checklist, the "Kihon Checklist". Geriatr Gerontol Int. 2017;17(9):1300–5. https://doi.org/10.1111/ggi.12864

99 Tynkkynen J, Laatikainen T, Salomaa V, Havulinna AS, Blankenberg S, Zeller T, et al. NT-proBNP and the risk of dementia: a prospective cohort study with 14 Years of follow-up. J Alzheimers Dis. 2015;44(3):1007–13. https://doi.org/10.3233/jad-141809

100 Tynkkynen J, Hernesniemi JA, Laatikainen T, Havulinna AS, Salo P, Blankenberg S, et al. High-sensitivity cardiac troponin I and NT-proBNP as predictors of incident dementia and Alzheimer's disease: the FINRISK Study. J Neurol. 2017;264(3):503–11. https://doi.org/10.1007/s00415-016-8378-7

101 Barnes D, Covinsky K, Whitmer R, Kuller L, Lopez O, Yaffe K. Predicting risk of dementia in older adults: the late-life dementia risk index. Neurology. 2009;73(3):173–9. https://doi.org/10.1212/wnl.0b013e3181a81636

102 Hogan DB, Ebly EM. Predicting who will develop dementia in a cohort of Canadian seniors. Can J Neurol Sci. 2000;27(1):18–24. https://doi.org/10.1017/s0317167100051921

103 Tierney MC, Moineddin R, McDowell I. Prediction of all-cause dementia using neuropsychological tests within 10 and 5 years of diagnosis in a community-based sample. J Alzheimers Dis. 2011;22(4):1231–40. https://doi.org/10.3233/jad-2010-100516

104 Tierney MC, Yao C, Kiss A, McDowell I. Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. Neurology. 2005;64(11):1853–9. https://doi.org/10.1212/01.wnl.0000163773.21794.0b

105 Verhaaren BF, Vernooij MW, Koudstaal PJ, Uitterlinden AG, van Duijn CM, Hofman A, et al. Alzheimer's disease genes and cognition in the nondemented general population. Biol Psychiatr. 2013;73(5):429–34. https://doi.org/10.1016/j.biopsych.2012.04.009

106 Vos SJB, van Boxtel MPJ, Schiepers OJG, Deckers K, de Vugt M, Carrière I, et al. Modifiable risk factors for prevention of dementia in midlife, late life and the oldest-old: validation of the LIBRA index. J Alzheimers Dis. 2017;58(2):537–47. https://doi.org/10.3233/jad-161208

107 Deckers K, van Boxtel MP, Schiepers OJ, de Vugt M, Muñoz Sánchez JL, Anstey KJ, et al. Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. Int J Geriatr Psychiatry. 2015;30(3):234–46. https://doi.org/10.1002/gps.4245
108 Walker KA, Chen J, Zhang J, Fornage M, Yang Y, Zhou L, et al. Large-scale plasma proteomic analysis identifies proteins and pathways associated with dementia risk. Nat Aging. 2021;1(5):473–89. https://doi.org/10.1038/s43587-021-00064-0
109 Lee S, Zhou J, Liu T, Zhang Q, Tse G, Lee TTL. Gender-specific clinical risk scores incorporating blood pressure variability for predicting incident dementia. J Am Med Inf Assoc. 2022;29(10):1825–6. https://doi.org/10.1093/jamia/ocac117
110 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Med. 2015;67(6):1142–51. https://doi.org/10.1016/j.eururo.2014.11.025
111 Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61(4):344–9. https://doi.org/10.1016/j.jclinepi.2007.11.008
112 Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet. 2020;396(10248):413–46. https://doi.org/10.1016/s0140-6736(20)30367-6
113 Kapogiannis D, Mustapic M, Shardell MD, Berkowitz ST, Diehl TC, Spangler RD, et al. Association of extracellular vesicle biomarkers with Alzheimer disease in the Baltimore longitudinal study of aging. JAMA Neurol. 2019;76(11):1340–51. https://doi.org/10.1001/jamaneurol.2019.2462
114 Mohanannair Geethadevi G, Quinn TJ, George J, Anstey KJ, Bell JS, Sarwar MR, et al. Multi-domain prognostic models used in middle-aged adults without known cognitive impairment for predicting subsequent dementia. Cochrane Database Syst Rev. 2023;2023(6):CD014885. https://doi.org/10.1002/14651858.cd014885.pub2
115 Tang EYH, Price CI, Robinson L, Exley C, Desmond DW, Köhler S, et al. Assessing the predictive validity of simple dementia risk models in harmonized stroke cohorts. Stroke. 2020;51(7):2095–102. https://doi.org/10.1161/strokeaha.120.027473
116 Brain J, Tully PJ, Turnbull D, Tang E, Greene L, Beach S, et al. Risk factors for dementia in the context of cardiovascular disease: a protocol of an overview of reviews. PLoS One. 2022;17(7):e0271611. https://doi.org/10.1371/journal.pone.0271611
117 Brain J, Greene L, Tang EYH, Louise J, Salter A, Beach S, et al. Cardiovascular disease, associated risk factors, and risk of dementia: an umbrella review of meta-analyses. Front Epidemiol. 2023;3:1095236. https://doi.org/10.3389/fepid.2023.1095236
118 Moons KG, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med. 2019;170(1):W1–33. https://doi.org/10.7326/m18-1377
119 Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. Lancet Neurol. 2011;10(9):819–28. https://doi.org/10.1016/s1474-4422(11)70072-2
120 Walsh S, Govia I, Peters R, Richard E, Stephan BCM, Wilson NA, et al. What would a population-level approach to dementia risk reduction look like, and how would it work? Alzheimers Dement. 2023;19(7):3203–9. https://doi.org/10.1002/alz.12985
121 Stephan BCM, Cochrane L, Kafadar AH, Brain J, Burton E, Myers B, et al. Population attributable fractions of modifiable risk factors for dementia: a systematic review and meta-analysis. Lancet Healthy Longevity. 2024;5(6):e406–21. https://doi.org/10.1016/s2666-7568(24)00061-8