Prediction Models for Conversion From Mild Cognitive Impairment to Alzheimer’s Disease: A Systematic Review and Meta-Analysis

Yanru Chen, Xiaoling Qian, Yuanyuan Zhang, Wenli Su, Yanan Huang, Xinyu Wang, Xiaoli Chen, Enhan Zhao, Lin Han and Yuxia Ma

1 Evidence-Based Nursing, School of Nursing, Lanzhou University, Lanzhou, China, 2 Department of Neurology, Second Hospital of Lanzhou University, Lanzhou, China, 3 Department of Nursing, Gansu Provincial Hospital, Lanzhou, China, 4 First School of Clinical Medicine, Lanzhou University, Lanzhou, China

Background and Purpose: Alzheimer’s disease (AD) is a devastating neurodegenerative disorder with no cure, and available treatments are only able to postpone the progression of the disease. Mild cognitive impairment (MCI) is considered to be a transitional stage preceding AD. Therefore, prediction models for conversion from MCI to AD are desperately required. These will allow early treatment of patients with MCI before they develop AD. This study performed a systematic review and meta-analysis to summarize the reported risk prediction models and identify the most prevalent factors for conversion from MCI to AD.

Methods: We systematically reviewed the studies from the databases of PubMed, CINAHL Plus, Web of Science, Embase, and Cochrane Library, which were searched through September 2021. Two reviewers independently identified eligible articles and extracted the data. We used the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies (CHARMS) checklist for the risk of bias assessment.

Results: In total, 18 articles describing the prediction models for conversion from MCI to AD were identified. The dementia conversion rate of elderly patients with MCI ranged from 14.49 to 87%. Models in 12 studies were developed using the data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI). C-index/area under the receiver operating characteristic curve (AUC) of development models were 0.67–0.98, and the validation models were 0.62–0.96. MRI, apolipoprotein E genotype 4 (APOE4), older age, Mini-Mental State Examination (MMSE) score, and Alzheimer’s Disease Assessment Scale cognitive (ADAS-cog) score were the most common and strongest predictors included in the models.

Conclusion: In this systematic review, many prediction models have been developed and have good predictive performance, but the lack of external validation of models limited the extensive application in the general population. In clinical practice, it is
recommended that medical professionals adopt a comprehensive forecasting method rather than a single predictive factor to screen patients with a high risk of MCI. Future research should pay attention to the improvement, calibration, and validation of existing models while considering new variables, new methods, and differences in risk profiles across populations.

**Keywords:** mild cognitive impairment, Alzheimer's disease, dementia, prediction models, systematic review

**INTRODUCTION**

Alzheimer's disease (often shortened to “Alzheimer’s” or “AD”) is the most common type of dementia occurring in older people and is defined as an irreversible, progressive neurodegenerative disorder characterized by abnormal accumulation of amyloid plaques and neurofibrillary tangles in the brain, causing a decline in thinking, memory, language, personality changes, and certain changes in the brain, that gradually get worse over time, eventually leading to a loss of ability to perform the simplest daily tasks (Ballard et al., 2011). As one of the greatest healthcare challenges of the twenty-first century, caring for patients with AD presents a heavy emotional and financial burden for families and society (Silva et al., 2019; Tiwari et al., 2019). Currently, AD affects more than 35 million people in the world, and its incidence is estimated to triple by 2050. In the United States alone, approximately 5.3 million people have AD, of which 5.1 million are aged 65 years or older (Alzheimer’s Association, 2015). The global socioeconomic costs for dementia were US$957.56 billion in 2015 and will reach US$2.54 trillion by 2030 and US$9.12 trillion by 2050 (Jia et al., 2018). Since current drug therapies can only postpone the progression of the disease and cannot directly prevent the progression of AD, more hope has been placed on the early prediction of AD.

Mild cognitive impairment (MCI) is considered to be a transitional stage between normal aging and AD and is a potential target for predicting individuals at risk of developing AD (Mariani et al., 2007). It can be defined as the presence of a memory complaint, objective memory impairment abnormal for age with relatively preserved general cognition, and essentially intact activities of daily living (no dementia) (Petersen, 2004). The annual conversion rate from MCI to AD has been reported as 10–15%. Approximately 80% of MCI patients will have converted to AD (Tábuas-Pereira et al., 2016), although some patients with MCI remain stable or convert back to normal (Lovell, 2009). Therefore, the predicted risk of conversion from MCI to AD and the early identification of patients with MCI with high risk are desperately required. These will facilitate the early treatment of patients with MCI before they convert to AD.

A systematic review of dementia prediction models was published in 2019, which reviewed the predictive performance and common predictors of dementia prediction models, but did not carry out an in-depth analysis of the prediction models and predictors of people for conversion from MCI to AD (Hou et al., 2019). In previous research, strenuous efforts had been made on the classification and prediction of MCI and AD based on the clinical, genetic, proteomic data and also on the AD imaging biomarkers. However, the predictive performance and clinical applicability of the current model still need to be further verified.

The purpose of this systematic review was to summarize the reported risk prediction models and identify the most prevalent factors for conversion from MCI to AD, so as to provide a theoretical basis for the construction, application, and optimization of the risk prediction model of dementia in patients with MCI and the early intervention.

**METHODS**

**Search Strategy**

We searched PubMed, CINAHL Plus, Web of Science, Cochrane Library, and Embase databases from the database inception through September 2021. The search strategies were performed through a combination of MeSH terms and free words. The following MeSH terms and free words were used: “cognitive dysfunctions,” “cognitive impairments,” “cognitive defect,” “mental disorders,” “dementia,” “dementias,” “Alzheimer's disease,” “Alzheimer Syndrome,” “dementia,” “prognostic model,” “prediction tool,” “prediction model,” “risk model,” and so on. Only articles published in English were considered for review. Additionally, the reference lists included in the identified articles were manually searched to identify additional relevant publications. The precise search strategies are shown in Figure 1.

**Eligibility Criteria**

Two researchers (Chen and Su) did the literature search and extracted the data independently. Discrepancies were resolved by a third researcher (Ma).

The inclusion criteria of a study in the systematic review were as follows: (1) patients diagnosed with MCI; (2) the content of the study is the risk prediction model for conversion from MCI to AD; (3) the model has been built after internal or external validation; (4) study type is either a cohort study or a case-control study. The exclusion criteria were as follows: (1) Review articles; (2) reports; (3) commentary, abstracts, and presentations; (4) disease-specific dementia; (5) the process or method of model building is not described; (6) cannot access the original text or incomplete data; (7) republished literature.

**Data Extraction**

Key information of the included studies was extracted by two researchers working independently using a data extraction form. From each study, we abstracted the following: first author, publication year, country, research object and location, dementia...
Quality Evaluation

The quality of the included studies was evaluated independently by two investigators, using the CHARMS (Moons et al., 2014). Any disagreement regarding the quality of studies was resolved by a third investigator. It included bias risk and applicability and was assessed from 11 items; the main items of which included source of data, participants, outcome(s) to be predicted, candidate predictors, sample size, missing data, model development, model performance, model evaluation, results, interpretation, and discussion.

Statistical Analysis

We used descriptive analysis methods to sort out and summarize the basic characteristics of the included studies and models, development methods, verification methods, and prediction factors in the model. Stata 15.1 software was used for the meta-analysis of the predictive value of the predictors in the model. First, a Q-test was used to verify whether there was heterogeneity among the included models. The degree of heterogeneity was assessed using the $I^2$ statistic, with the $I^2$ values of 25, 50, and 75% being considered to indicate low, moderate, and high heterogeneity, respectively. If $I^2 > 50\%$, heterogeneity was considered larger, and the random effects model was used for analysis, otherwise the fixed effects model was used. The count data were represented by odds ratio (OR) and 95%CI, while the
measurement data were represented by weighted mean difference (WMD) and 95%CI.

RESULTS

Literature Search

Figure 1 shows the results of literature searching and selection. A total of 3,404 related articles were retrieved from the database, and 2,948 articles remained after deduplication. By reading the titles and abstracts, 2,904 articles were excluded from the study, and 28 articles were excluded from the study for reasons, such as failure to construct a risk prediction model, for the conversion from MCI to AD and repeated publication. In addition, two references were supplemented by consulting the included references. Ultimately, a total of 18 studies met the inclusion criteria and were utilized for the meta-analysis.

Characteristics of Eligible Studies

A total of 18 studies were included in this study. Notably, 12 of the studies' data were conducted using the data from the ADNI dataset. Of the 18 studies we included, 14 were retrospective cohort studies and four were prospective cohort studies. Among the diagnostic criteria for predicting the outcome of AD, five studies (Fleisher et al., 2008; Hansson et al., 2009; Korolev et al., 2016; Handels et al., 2017; Jang et al., 2018) used NINCDS-ADRDA diagnostic criteria, while other studies used other criteria, such as Diagnostic and Statistical Manual of Mental Disorders (DSM), Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale cognitive (ADAS-cog) scale, and Clinical diagnostic assessments. Shigemizu's study (Shigemizu et al., 2020) uses the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for AD (NIA) as the diagnosis of outcome indicators. The follow-up duration of the studies ranged from 0.5 to 7 years, and the dementia conversion rate of elderly patients with MCI ranged from 14.49% to 87%. The basic characteristics of the included studies are shown in Table 1.

Evaluation of Methodology Quality in Eligible Studies

All the articles included in this study were prospective or retrospective cohort studies, and the inclusion and exclusion criteria of the subjects were presented. Of note, eight studies (Prins et al., 2013; Korolev et al., 2016; Ding et al., 2017; Handels et al., 2017; Pereira et al., 2017; Jang et al., 2019; Lee et al., 2019; Shigemizu et al., 2020) reported missing data, and five of them processed missing data by mean interpolation (Pereira et al., 2017), excluded missing data cases (Korolev et al., 2016; Jang et al., 2019; Shigemizu et al., 2020), and multiple imputation (Handels et al., 2017); 11 studies (Fleisher et al., 2008; Hansson et al., 2009; Prins et al., 2013; Korolev et al., 2016; Handels et al., 2017; Pereira et al., 2017; Jang et al., 2018; Kauppi et al., 2018; Basaia et al., 2019; Huang et al., 2019; Jang et al., 2019; Lee et al., 2019; Sörensen et al., 2019; Shigemizu et al., 2020) only carried out internal verification, such as bootstrap method and cross-validation (CV), but did not conduct external verification. Only two studies (Pereira et al., 2017; Jang et al., 2018) assessed the method using internal and external model verification. In addition, 11 articles (Fleisher et al., 2008; Hansson et al., 2009; Prins et al., 2013; Korolev et al., 2016; Han et al., 2017; Handels et al., 2017; Jang et al., 2018; Kauppi et al., 2018; Jang et al., 2019; Sörensen et al., 2019; Shigemizu et al., 2020) presented complete regression equations, neural networks, nomograms, or Bayesian network models. The risk of bias assessment for included studies is shown in Figure 2.

Basic Characteristics, Modeling, and Verification Methods

A total of 47 predictive models for the conversion from MCI to dementia were reported in the 18 articles included. In each study, we selected a model validated by the author or with the best performance in C-index/AUC for analysis and verification and statistical analysis. The first predictive model was published in 2008 (Fleisher et al., 2008), the latest in 2020 (Shigemizu et al., 2020), and 14 in the last 5 years. Models in 12 studies were developed using the data from the ADNI dataset. The sample size of method development in the studies is 98–1,327. Cox regression (n = 7), nomogram (n = 3), and neural networks (n = 3) are the most widely used modeling methods. Other modeling methods, such as Logistic regression (n = 2), support vector machine (SVM; n = 1), probabilistic multiple kernel learning (n = 1), Bayesian algorithm (n = 1), and supervised learning, approach based on time windows (n = 1). The sample size of method validation is 62–865, with eight studies using CV (Fleisher et al., 2008; Sabuncu, 2013; Korolev et al., 2016; Han et al., 2017; Kauppi et al., 2018; Basaia et al., 2019; Huang et al., 2019; Lee et al., 2019), three studies (Jang et al., 2018, 2019; Shigemizu et al., 2020) using a mix of CV and bootstrap validation, and only two studies (Pereira et al., 2017; Jang et al., 2018) using internal and external validation. Notably, nine of them reported the AUC value of the development model of 0.67–0.98, and 10 of them reported the AUC value of the validation model of 0.62–0.96.

Among the predictors included in the study models, a maximum of 10 predictors and a minimum of one predictor were reported. MRI (n = 12), apolipoprotein E genotype 4 (APOE4) (n = 10), older age (n = 7), female gender (n = 6), lower MMSE (n = 5), ADAS-cog scores (n = 5), cerebrospinal fluid (CSF) biomarkers (n = 5), and Functional Assessment Questionnaire (FAQ) scores (n = 4) were the most common predictors included in the models. Other than that, some other predictive factors, such as Fl8-fludeoxyglucose PET (FDG-PET; n = 3), education (n = 2), Rey Auditory Verbal Learning Test
Meta-Analysis for Predictive Factors

In this review, we analyzed the predictive value of the eight predictors of MRI, APOE4, age, gender, MMSE score, ADAS-cog, FAQ score, and FDG-PET with the highest frequency of entering the predictive model, on the predictive value of the risk of AD in patients from MCI. The results showed that the five influencing factors, such as MRI (OR = 1.419, 95%CI: 1.176–1.712), age (OR = 0.963, 95%CI: 0.935–0.991), gender (OR = 1.043, 95%CI: 1.001–1.086), APOE4 (OR = 1.877, 95%CI: 1.552–2.271), and MMSE score (OR = 0.997, 95%CI: 0.994–0.999), significantly predict the conversion from MCI to AD.

Table 1: Primary characteristics of the prediction model included in the review (n = 18).

| Source            | Country          | Type of study                        | Source of data | Subject                                                                 | AD assessment tool | Follow-up time | Incidence of AD |
|-------------------|------------------|---------------------------------------|----------------|-------------------------------------------------------------------------|-------------------|----------------|-----------------|
| Basaia et al. (2019) | United States/Italy | Retrospective cohort study           | ADNI/Milan dataset | AD, MCI and healthy controls                                             | Clinical assessment | 3 years        | 33%/54%         |
| Ding et al. (2017) | United States    | Retrospective cohort study           | ADNI dataset    | AD, MCI and healthy controls                                             | NR                | 1 year         | 14.49%          |
| Fleischer et al. (2008) | United States    | Retrospective cohort study           | ADCS MCI treatment trial | aMCI, NINCDS-ADRDA                                                        | 3 years            | 41.09%         |
| Han et al. (2017)  | United States    | Retrospective cohort study           | ADNI dataset    | MCI                                                                      | NR                | NR             | 60.18%          |
| Handels et al. (2017) | Multicountry, multicenter | Retrospective cohort study           | DESCRIPA multicenter study; LEARN multicenter study; Ljubljana University Medical Center and Karolinska University Hospital Huddinge memory clinic | MCI, DSN-IV-TR and NINCDS-ADRDA | 2 years | 87%          |
| Hansson et al. (2009) | Sweden           | Prospective cohort study            | Malmö University Hospital   | MCI, healthy controls                                                    | DSM-III and NINCDS-ADRDA | 4–6 years | 41%          |
| Huang et al. (2019) | United States    | Retrospective cohort study           | ADNI dataset    | MCI                                                                      | CDR scores        | At least 0.5 years | 26.21%         |
| Jang et al. (2018)  | South Korea      | Prospective cohort study            | Memory Disorder Clinic in Samsung Medical Center, Clinical Research Center for Dementia of South Korea | aMCI, DSN-IV and NINCDS-ADRDA | 3 years | 61.5%         |
| Jang et al. (2019)  | United States    | Retrospective cohort study           | ADNI dataset    | Aβ + MCI, Clinical diagnostic assessments                                 | 3 years            | 41.94%         |
| Kauppi et al. (2018) | United States    | Retrospective cohort study           | ADNI dataset    | MCI                                                                      | NR                | 3 years        | 54.17%          |
| Korolev et al. (2016) | United States    | Retrospective cohort study           | ADNI dataset    | MCI                                                                      | NINCDS-ADRDA      | 3 years        | 53.67%          |
| Lee et al. (2019)   | United States    | Retrospective cohort study           | ADNI dataset    | AD, MCI and healthy controls                                             | NR                | 2 years        | 35.49%         |
| Li et al. (2019)    | United States    | Retrospective cohort study           | ADNI dataset    | MCI                                                                      | ADAS-cog, RAVLT, FAQ and MMSE | 1 year       | 38.32%         |
| Pereira et al. (2017) | Portugal         | Prospective cohort study            | CCC             | MCI                                                                      | DSM-IV            | 5 years        | 36.00%         |
| Prins et al. (2013) | Netherlands      | Prospective cohort study            | Gal-Int-11      | MCI                                                                      | CDR scores        | 2 years        | 19.00%         |
| Sabuncu (2013)      | United States    | Retrospective cohort study           | ADNI dataset    | MCI                                                                      | NR                | 4.5 years      | 42.24%         |
| Shigemizu et al. (2020) | Japan           | Retrospective cohort study           | NCGG            | MCI                                                                      | NIA               | 0.5–7 years    | 42.10%         |
| Sörensen et al. (2019) | Germany         | Retrospective cohort study           | ADNI dataset    | MCI                                                                      | NR                | 3.8–4 years    | 33.27%         |

ADNI, Alzheimer’s Disease Neuroimaging Initiative; AD, Alzheimer’s disease; MCI, mild cognitive impairment; ADCS, Alzheimer’s Disease Cooperative Study; aMCI, amnestic mild cognitive impairment; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders-Stroke/Alzheimer’s Disease and Related Disorders Association; DSM, Psychiatry Diagnostic & Statistical Manual of Mental Disorders; Aβ, amyloid-beta positivity; MRI, magnetic resonance imaging; APOE, apolipoprotein E; MMSE, Mini-Mental State Examination; ADAS-cog, Alzheimer’s Disease Assessment Scale-cognitive subscale; FAQ, Functional Assessment Questionnaire; RAVLT, Rey Auditory Verbal Learning Test; NCGG, the National Center for Geriatrics and Gerontology; NIA, the criteria of the National Institute on Aging Alzheimer’s Association workgroups.
FIGURE 2 | Risk of bias assessment for included studies.

\[ P = 0.023 \], MMSE score \( (WMD = 0.877, 95\% CI: 0.573–1.182, P = 0.000) \), and ADAS-cog score \( (WMD = 4.211, 95\% CI: 3.488–4.934, P = 0.000) \), were statistically significant \( (P < 0.05) \). Among them, older age \( (I^2 = 75.5\%) \) and MMSE score \( (I^2 = 68.5\%) \) showed large heterogeneity \( (I^2 > 50\%) \). The results are shown in Table 3.

## DISCUSSION

We conducted a systematic review of prediction models for conversion from MCI to AD and a meta-analysis of the performance of prediction factors. The 18 studies included in this systematic review are carefully designed in the process of model development and evaluation. The included studies are both prospective cohort studies and retrospective cohort studies, and they all define the source and inclusion criteria of the study objects, which effectively reduce the selection bias. Among the 47 predictive models reported, the AUC of 40 models in the model population were all \( > 0.7 \), indicating that the constructed model can accurately identify the risk of dementia with MCI. However, there are some shortcomings in the model construction process.

For example, of the studies we analyzed, 12 were based on the ADNI, with fewer studies from other independent cohorts, which will limit the generalizability of those models. During the validation of the model, most studies only use CV or bootstrap internal validation, only two studies (Pereira et al., 2017; Jang et al., 2018) used external validation, while it is well known that external validation is critical in assessing a model’s capability and applicability (Hou et al., 2019). In addition, the candidate refers to the predictors chosen to be studied for their predictive performance and is not restricted to those included in the multivariable analysis (Moons et al., 2012). The number of candidate predictors analyzed in the primary studies is highly important. However, a few studies (Fleisher et al., 2008; Sabuncu, 2013; Ding et al., 2017; Pereira et al., 2017; Basaia et al., 2019) have not reported the number and process of analyzing candidate predictors before developing models, which may lead to the overfitting of final prediction model (Moons et al., 2014). In future studies, it is recommended to conduct a multicountry, multicenter prospective cohort study to explore more generalized and applicable predictive tools and factors. At the same time, researchers should pay attention to improvement, calibration, and validation of the models and also to the screening of candidate prediction factors to improve the reliability of research results.

### Demographic Data (Apolipoprotein E Genotype \( \varepsilon 4 \), Age, and Gender)

The results of this study showed that APOE4 was identified as one of the strongest predictors of the transition from MCI to dementia, consistent with previous studies (Lambert et al., 2013; Kunkle et al., 2019). In the meta-analysis of prediction factors, APOE4 also showed high homogeneity and prediction performance. In 1993, APOE4, as an important risk marker of AD, was first reported by Corder et al. (1993) and has been validated in major cohorts around the globe in recent years (Rasmussen et al., 2015; Livingston et al., 2020). Previous studies have shown that APOE4 has several effects on AD. First, APOE4 interferes with Aβ clearance from the brain and is also processed into neurotoxic fragments (Mahley and Huang, 2012). Furthermore, APOE4 causes the disinhibition of the
TABLE 2 | Modeling, verification methods, and predictive factors for included studies (n = 18).

| Source           | Model development | Model validation | Risk factors in final model                                                                 | Sample size | AUC              | Accuracy | Sensitivity | Specificity |
|------------------|-------------------|------------------|------------------------------------------------------------------------------------------------|-------------|------------------|----------|-------------|-------------|
|                  |                   |                  |                                                                                                | Development| Validation      | Development| Validation |
| Basaia et al. (2019) | CNN               | 10-fold cross-validation | Brain structural MRI scan                                                                     | 1,327       | 147              | NR       | NR          | 74.90%      |
| Ding et al. (2017)    | SVM               | NR               | MRI (cortical thickness), FDG-PET (cerebellum and the whole white matter)                      | 214         | NR               | 0.67     | NR          | 72.78%      |
| Fleisher et al. (2008) | LR                | 10-fold cross-validation | ADAS-Cog, MMSE, FAQ                                                                            | 129         | 129              | NR       | 0.89        | 78.80%      |
| Han et al. (2017)     | CR                | 10-fold cross-validation | APOE ε4 allele, Neurological disorder (other than AD), CDR-SB, ADAS-Cog 13, MMSE, FAQ         | 658         | 74               | NR       | 0.84        | 77%         |
| Handels et al. (2017) | LR                | Bootstrap validation | Female gender, MMSE, MCI stage on MRI and CSF biomarkers (Aβ1–42, t-tau and p-tau)           | 250         | 250              | NR       | 0.85        | NR          |
| Han et al. (2017)     | CR                | NR               | Older age, female gender, APOE ε4 carrier, rCBF and CSF biomarkers (t-tau, p-tau and Aβ1–42) | 167         | NR               | 0.78     | NR          | 77.2%       |
| Huang et al. (2019)   | Nomogram          | 10-fold cross-validation | Radiomics signature (MRI (cortical features)), FAQ scores and Aβ1–42 CSF concentrations     | 191         | 99               | 0.98     | 0.96        | NR          |
| Jang et al. (2018)    | Nomogram          | Bootstrapping, 10-fold cross-validation + External validation | Older age, APOE4 and neuropsychological features (modality, severity, and multiplicity)       | 167         | 75               | 0.80     | 0.75,0.82   | NR          |
| Jang et al. (2019)    | Nomogram          | Bootstrapping, 10-fold cross-validation | APOE4, MCI stage, MRI (hippocampal volume), FDG-PET SUVR, CSF (t-tau and P-tau)             | 124         | 62               | 0.93     | 0.91        | NR          |
| Kauppi et al. (2018)  | CR                | Cross-validated  | APOE ε4 alleles, MMSE, MRI (brain atrophy score), and PHS                                      | 336         | 336              | 0.84     | NR          | 78.9%       |
| Korolev et al. (2016) | pMKL              | 10-fold cross-validation | Cognitive scores (ADAS-Cog and RAVLT), functional assessments (FAQ) and MRI measures (volume/cortical thickness of left hippocampus, middle temporal gyrus, and inferior parietal cortex) | 259         | 259              | NR       | 0.87        | 79.9%       |
| Lee et al. (2019)     | RNN               | 5-fold cross-validation | Cognitive score (executive functioning and memory), MRI (hippocampal volume and entorhinal cortical thickness), CSF biomarker (Aβ1–42, t-tau and p-tau), demographic data, APOE ε4 | 865         | 865              | NR       | 0.86        | 81%         |
| Li et al. (2019)      | RNN, CR           | NR               | Demographic data, APOE ε4, and hippocampal MRI measures, Cognitive measures (ADAS-Cog13, RAVLT immediate, RAVLT learning, FAQ, and MMSE) | 822         | 439              | 0.90     | NR          | NR          |
| Pereira et al. (2017) | A supervised learning approach based on time windows | 5-fold cross-validation and External validation | Neuropsychological data                                                                       | 719         | 604,115          | 0.88     | 0.76        | 0.88        |
| Prins et al. (2013)   | CR                | NR               | Age, gender, MTA scores on MRI, ADAS-cog/MCI and Delayed recall test                            | 426         | NR               | NR       | NR          | NR          |
| Sabuncu (2013)        | BN                | 5-fold cross-validation | Structural MRI (cortical thickness)                                                              | 393         | 315              | NR       | 0.62        | NR          |
| Shigenzi et al. (2020) | CR               | 3-fold cross-validation, Bootstrap resampling | 24 miR-eQTLs, Older age, and APOE ε4                                                           | 98          | 99               | 0.72     | 0.70        | NR          |
| Sörensen et al. (2019) | CR                | Proportional hazard assumption | APOE ε4 alleles, MMSE, FDG-PET                                                                  | 272         | 272              | NR       | NR          | NR          |

CNN, Convolutional neural networks; AUC, area under the receiver operating characteristic curve; SVM, Support Vector Machine; FDG-PET, Fluorodeoxyglucose (FDG) positron emission tomography (PET); LR, Logistic regression; APOE, apolipoprotein E genotype; ADAS-cog, Alzheimer’s Disease Assessment Scale–Cognitive Subscale; CR, COX regression; CDR-SB, cognitive dementia rating scale sum of boxes; MMSE, mini-mental state examination score; FAQ, Functional Assessment Questionnaire; MTA, medial temporal lobe atrophy; CSF, cerebrospinal fluid; Aβ1–42, amyloid-β 1–42; t-tau, total tau; P-tau, phosphorylated tau; rCBF, regional cerebral blood flow; PHS, polygenic hazard score; pMKL, probabilistic multiple kernel learning; RNN, recurrent neural network; RAVLT, Rey Auditory Verbal Learning Test; BN, Bayesian algorithm; miR-eQTLs, microRNA expression quantitative trait loci.
TABLE 3 | Results of meta-analysis for predictive factors.

| Predictors        | Number of studies | OR/WMD | 95% CI | Z    | P  | I² (%) | P  |
|-------------------|-------------------|--------|--------|------|----|-------|----|
| MRI               | 12                | 1.419  | 1.176–1.712 | 3.65 | 0.000 | 0.0%  | 0.557 |
| APOE4             | 10                | 1.877  | 1.552–2.271 | 6.48 | 0.000 | 38.4% | 0.112 |
| Age               | 7                 | 1.073* | 1.010–1.140 | 2.28 | 0.023 | 75.5% | 0.003 |
| Gender            | 6                 | 1.235  | 0.960–1.588 | 1.64 | 0.101 | 50.0% | 0.075 |
| MMSE              | 5                 | 0.877* | 0.573–1.182 | 5.65 | 0.000 | 68.5% | 0.042 |
| ADAS-cog          | 5                 | 4.211* | 3.488–4.934 | 11.41 | 0.000 | 25.6% | 0.246 |
| FAQ score         | 4                 | 7.648* | -0.919–6.210 | 1.75 | 0.080 | 98.7% | 0.000 |
| FDG-PET           | 3                 | 0.748  | 0.280–1.995 | 0.58 | 0.561 | 93.6% | 0.000 |

*WMD.

Chen et al. Systematic Review of Prediction Models

Cyclophilin A signaling mechanism in the pericytes of the brain blood vessels, leading to a degeneration of these vessels, leakage of the blood-brain barrier, and tau-induced neurodegeneration and atrophy (Bell et al., 2012; Serrano-Pozo et al., 2021). However, no therapies directed at APOE are currently available. Although several therapeutic approaches have been successful in mouse models expressing human APOE alleles, including increasing or decreasing APOE levels, enhancing its lipidation, blocking the interactions between APOE and Aβ peptide, and genetically switching APOE4 to APOE3 or APOE2 isoforms, translation to human clinical trials has proven challenging (Serrano-Pozo et al., 2021).

Older age was also strongly predictive for patients with MCI to AD. The reason may be that aging acts through various biological mechanisms at the cellular or tissue level which lead to multisystem loss of reserve and function and affect the health of the body's blood vessels (Fabbri et al., 2015). Besides, amyloid and tau pathologies as well as brain atrophy increase with age (Vemuri et al., 2017). Of all the models we analyzed, six used female gender as a predictor, but the meta-analysis in this review found that the heterogeneity among the analyzed studies was large and not statistically significant, which is not consistent with the results of previous studies (Vermunt et al., 2019; Yoo et al., 2020). It may be related to the population selection (e.g., the sample size for the development and validation of the model in Shigemizu et al., 2020, study was less than 100) and the follow-up time (e.g., Li et al., 2019, study was only followed for 1 year) of the included models, which needs to be further verified in future research. It is suggested that future prediction models should be further validated by increasing sample size or longer follow-up time.

Cognitive Scores (Mini-Mental State Examination, Alzheimer’s Disease Assessment Scale Cognitive, and Functional Assessment Questionnaire)

Notably, 28% of models we selected considered MMSE as a high factor for AD. Other cognitive screening scales, such as ADAS-cog score and FAQ score, also show good prediction performance. This is consistent with the findings of McEvoy et al. (2009). But Mitchell (2015) pointed out that the MMSE when used alone is not a good tool for prediction of future decline in people with MCI. A combination of predictors would be more accurate in predicting progression from MCI to dementia.

Both FAQ and ADAS-cog are the frequently used indices of cognitive decline in AD. In recent years, they have also been seen as strong predictors of the conversion from MCI to AD. In this review, five studies selected ADAS-cog score as a predictor, and all had good predictive performance; it suggests that baseline impairment in multiple cognitive domains is predictive of future progression to dementia (Korolev et al., 2016). Patients with MCI with both memory and non-memory deficits have a greater risk of progression to AD than those with isolated memory deficits (Bozoki et al., 2001). Impairment in multiple cognitive domains, as measured by the performance on the ADAS-cog, can be viewed as reflecting a more advanced MCI stage. The selection of FAQ scores as predictors for conversion from MCI to AD indicates that a subtle but reliable impairment in functional status precedes the development of overt dementia in patients with MCI (Korolev et al., 2016). In addition, longitudinal glucose metabolism decline was associated with concurrent ADAS-cog and FAQ decline, which has a value in predicting the future cognitive decline of patients with MCI (Landau et al., 2011). In general, the cognitive scores of MMSE, ADAS-cog, and FAQ are cost-effective. Neuropsychological test performance is relatively easy to gather, compared to the more expensive MRI scan, PET scan, and CSF biomarkers which require invasive lumbar puncture.

Imaging Biomarkers (Magnetic Resonance Imaging and F18-Fluorodeoxyglucose Positron Emission Tomography)

The MRI is one of the most widely studied imaging techniques because it is completely non-invasive, highly available, and inexpensive compared to PET and has an excellent contrast between different soft tissues (Johnson et al., 2012). The
results of this review showed that MRI was identified as the strongest predictor of the transition from MCI to AD. In our review, 67% of models we selected considered MRI as a high factor for conversion from MCI to AD, and all had a good predictive performance. The AUCs ranged from 0.62 to 0.98. In general, models that combined, such as demographic data, cognitive scores, fluid biomarkers, and other clinical markers, have better predictive performance than a single MRI model. The results of our study showed that many MRI markers, such as the whole brain, hippocampal, entorhinal cortex atrophy, and medial temporal lobe atrophy (MTA) on MRI score, have significant predictive value. However, the systematic studies of advantages/disadvantages of various MRI markers have been limited so far, and the existing studies do not allow to make definite conclusions. There is still no preferred MRI representation for AD-conversion prediction (Gómez-Sancho et al., 2018).

Previous research has shown that FDG-PET is also one of the effective predictors for predicting the conversion from MCI to AD. However, since there were only three cases of meta-analysis in this review, and the model sample size of some studies was small, the results showed a large heterogeneity, which needs to be further verified in future research. In addition, FDG-PET is relatively expensive and, similar to all PET techniques, has more limited availability. It requires intravenous access and involves exposure to radioactivity. Brain FDG retention is a non-specific indicator of metabolism that can be deranged for a variety of reasons (e.g., ischemia or inflammation) and may, in certain individuals, be irrelevant or only indirectly related to any AD-related process.

**CONCLUSION**

In this systematic review, many prediction models have been developed and have good predictive performance, but the lack of external validation of models limited the extensive application in the general population. In addition, 67% of models were based on the ADNI dataset, with fewer studies from other independent cohorts, which will limit the generalizability of those models. MRI, APOE4, older age, MMSE score, and ADAS-cog score were the most common and strongest predictors included in the models. However, a few of the predictors are still highly heterogeneous. In clinical practice, it is recommended that medical professionals adopt a comprehensive forecasting method rather than a single predictive factor to screen patients with a high risk of MCI. Furthermore, future research should pay attention to improvement, calibration, and validation of existing models while considering new variables, new methods, and differences in risk profiles across populations.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

**AUTHOR CONTRIBUTIONS**

YC, XQ, and YM contributed to the study design. WS, YH, XW, and EZ contributed to the data collection. YZ, YC, and XC performed statistical analyses and interpretation of results. YC, XQ, and LH drafted the manuscript and edited the language. All authors participated in the critical revisions and approved the final version of the manuscript.

**FUNDING**

This study was supported by the National Nature Science Foundation of China (grant 71704071), Natural Science Foundation of Gansu Province (20JR10RA603), the Fundamental Research Funds for the Central Universities (Iuzjkby-2020-10 and Iuzjkby-2021-33), National Research Training Program of Gansu Provincial Hospital (19SYPYA-4), the Research Funds for the School of Nursing of Lanzhou University (LZUSON202002), Gansu Province Health and Health Industry Scientific Research Program Project (GSWSHL2021-016), and The Second Hospital of Lanzhou University Cuiying Science and Technology Innovation Clinical Nursing Research Project (CY2021-HL-05).

**ACKNOWLEDGMENTS**

We are grateful to Lanzhou University Library for the assistance during the literature search. Furthermore, we sincerely appreciate the hard work from the distinguished editor and reviewers.

**REFERENCES**

Alzheimer’s Association. (2015). 2015 Alzheimer’s disease facts and figures. *Alzheimer’s Dement.* 11, 332–384. doi: 10.1016/j.jalz.2015.02.003

Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., and Jones, E. (2011). Alzheimer’s disease. *Lancet.* 377, 1019–1031. doi: 10.1016/S0140-6736(10)61349-9

Bassia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., et al. (2019). Alzheimer’s Disease Neuroimaging Initiative. Automated classification of Alzheimer’s disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin.* 21:101645. doi: 10.1016/j.nicl.2018.101645

Bell, R. D., Winkler, E. A., Singh, I., Sagare, A. P., Deane, R., Wu, Z., et al. (2012). Apolipoprotein E controls cerebrovascular integrity via cyclophilin A. *Nature* 485, 512–516. doi: 10.1038/nature11087

Bozoki, A., Giordani, B., Heidebrink, J. L., Berent, S., and Foster, N. L. (2001). Mild cognitive impairments predict dementia in nondemented elderly patients with memory loss. *Arch Neurol.* 58, 411–416. doi: 10.1001/archneur.58.3.411

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., et al. (1993). Gene dose of apolipoprotein E type 4 allele...
and the risk of Alzheimer’s disease in late onset families. Science 261, 921–923. doi: 10.1126/science.8346443

Ding, J. W., Huang, Q., and Ieee. (2017). Prediction of MCI to AD Conversion Using Laplace Eigenmaps Learned from FDG and MRI Images of AD Patients and Healthy Controls. 2017 2nd International Conference on Image, Vision and Computing (ICIVC).(New Jersey, NJ: IEEE) 660–664. doi: 10.1109/ICIVC.2017. 7984638

Fabbri, E., Zoli, M., Gonzalez-Freire, M., Salive, M. E., Studenski, S. A., and Ferrucci, L. (2015). Aging and Multimorbidity: New Tasks. Priorities, and Frontiers for Integrated Gerontological and Clinical Research. J. Am. Med. Dir. Assoc. 16, 640–647. doi: 10.1016/j.jamda.2015.03.013

Flsher, A. S., Sun, S., Taylor, C., Ward, C. P., Gamst, A. C., Petersen, R. C., et al. (2008). Volumetric MRI vs clinical predictors of Alzheimer disease in mild cognitive impairment. Neurology. 70, 191–199. doi: 10.1212/01.wnl.0000270911.57376.65

Gómez-Sancho, M., Tohka, J., Gómez-Verdejo, V., Alzheimer’s Disease Neuroimaging Initiative. (2018). Comparison of feature representations in MRI-based MCI-to-AD conversion prediction. Magn. Reson. Imaging 50, 84–95. doi: 10.1016/j.mri.2018.03.003

Han, X., Zhang, Y., and Shao, Y. (2017). Application of concordance probability estimate to predict conversion from mild cognitive impairment to Alzheimer’s disease. Biostat. Epidemiol. 1, 105–118. doi: 10.20514/bse.2017.134 2187

Handels, R. L. H., Vogt, S. B., Krambmer, M. G., Jelic, V., Blennow, K., van Buchem, M., et al. (2017). Predicting progression to dementia in persons with mild cognitive impairment using cerebrospinal fluid markers. Alzheimer’s Dementia. 13, 903–912. doi: 10.1016/j.jalz.2016.12.015

Hansson, O., Buchhave, P., Zetterberg, H., Blennow, K., Minthon, L., and Warkentin, S. (2009). Combined rCBF and CSF biomarkers predict progression from mild cognitive impairment to Alzheimer's disease. Neurobiol. Aging. 30, 165–173. doi: 10.1016/j.neurobiolaging.2007.06.009

Hou, X., Feng, L., Zhang, C., Cao, X. P., Tan, L., and Yu, J. T. (2019). Models for predicting risk of dementia: a systematic review. J. Neurol. Neurosurg. Psychiatry 90, 373–379. doi: 10.1136/jnnp-2018-318212

Huang, K., Lin, Y., Yang, L., Wang, Y., Cai, S., Pang, L., et al. (2019). A multipredictor model to predict the conversion of mild cognitive impairment to Alzheimer’s disease by using a predictive nomogram. Neuropsychopharmacology. 45, 358–366. doi: 10.1038/s41386-019-0551-0

Jang, H., Park, J., Woo, S., Kim, S., Kim, H. J., Na, D. L., et al. (2019). Prediction of fast decline in amyloid positive mild cognitive impairment patients using multimodal biomarkers. NeuroImage Clin. 24:101941. doi: 10.1016/j.nicl.2019.10.1941

Jang, H., Ye, B. S., Woo, S., Kim, S. W., Chin, J., Choi, S. H., et al. (2018). Prediction Model of Conversion to Dementia Risk in Subjects with Amnestic Mild Cognitive Impairment: A Longitudinal Multi-Center Clinic-Based Study. J. Alzheimers Dis. 60, 1579–1587. doi: 10.3233/JAD-170507

Jia, J., Wei, C., Chen, S., Li, F., Tang, Y., Qin, W., et al. (2018). The cost of Alzheimer’s disease in China and re-estimation of costs worldwide. Alzheimer’s Dement 14, 483–491. doi: 10.1016/j.jalz.2017.12.006

Johnson, K. A., Fox, N. C., Sperling, R. A., and Klunk, W. E. (2012). Brain imaging predicting risk loci and implicates Aβ in injury triggers neuropathology. Neuron 76, 871–885. doi: 10.1016/j.neuron.2012.11.020

Mariani, E., Monastero, R., and Mecocci, P. (2007). Mild cognitive impairment: a systematic review. J. Alzheimers Dis. 22, 33–35. doi: 10.3233/jad-2007-1 2104

McEvoy, L. K., Fennema-Notestine, C., Roddey, J. C., Hagler, D. J. Jr., Holland, D., Karow, D. S., et al. (2009). Alzheimer’s Disease Neuroimaging Initiative. Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. Radiology. 251, 195–205. doi: 10.1148/radiol.251080924

Mitchell A. J. (2015). Can the MMSE help clinicians predict progression from mild cognitive impairment to dementia? Biophys Adv. 21, 363–366. doi: 10.1192/apt. 21.6.363

Moons, K. G., de Groot, J. A., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D. G., et al. (2014). Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 11:e1001744. doi: 10.1371/journal.pmed.1001744

Moons, K. G., Kengne, A. P., Woodward, M., Royston, P., Vergouwe, Y., Altman, D. G., et al. (2012). Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 98, 683–690. doi: 10.1136/heartjnl-2011-301246

Pereira, T., Lemos, L., Cardoso, S., Silva, D., Rodrigues, A., Santana, I., et al. (2017). Predicting progression of mild cognitive impairment to dementia using neuropsychological data: a supervised learning approach using time windows. BMC Med. Inform. Decision Making. 17:110. doi: 10.1186/s12911-017-0497-2

Peterson, R. C. (2004). Mild cognitive impairment as a diagnostic entity. J. Intern. Med. 256, 183–194. 2004.01388. x doi: 10.1111/j.1365-2796

Prins, N. D., Von Der Flier, W. M., Brashear, H. R., Barkhof, F., and Scheltens, P. (2013). Predictors of progression from mild cognitive impairment to Alzheimer’s disease in the placebo arm of a clinical trial population. J. Nut. Health Aging. 36, 79–85. doi: 10.3233/JAD-122233

Rasmussen, K. L., Tybjerg-Hansen, A., Nordestgaard, B. G., and Frikke-Schmidt, R. (2015). Plasma levels of apolipoprotein E and risk of dementia in the general population. Ann. Neurol. 77, 301–311. doi: 10.1002/ana.24326

Sabuncu, M. R. (2013). “A Bayesian Algorithm for Image Time-to-Event Prediction”. In: G. Wu, D. Zhang, Shen,P. Yan, K. Suzuki, F. Wang (eds) Machine Learning in Medical Imaging, MLMI 2013. Lecture Notes in Computer Science, 8184. (Switzerland: Springer) 74–81. doi: 10.1007/978-3-319-0226 7-3_10

Serrano-Pozo, A., Das, S., and Hyman, B. T. (2021). APOE and Alzheimer’s disease: advances in genetics, pathophysiology, and therapeutic approaches. Lancet Neurol. 20, 68–80. doi: 10.1016/S1474-4422(20)30412-9

Shigemizu, D., Akiyama, S., Higaki, S., Sugimoto, T., Sakurai, T., Boroevich, K. A., et al. (2020). prognosis prediction model for conversion from mild cognitive impairment to Alzheimer’s disease created by integrative analysis of
multi-omics data. *Alzheimer's Res. Ther.* 12:145. doi: 10.1186/s13195-020-0716-0

Silva, M. V. F., Loures, C. M. G., Alves, L. C. V., de Souza, L. C., Borges, K. B. G., and Carvalho, M. D. G. (2019). Alzheimer's disease: risk factors and potentially protective measures. *J. Biomed. Sci.* 26:33. doi: 10.1186/s12929-019-0524-y

Sörensen, A., Blazhenets, G., Rücker, G., Schiller, F., Meyer, P. T., and Frings, L. (2019). Prognosis of conversion of mild cognitive impairment to Alzheimer's dementia by voxel-wise Cox regression based on FDG PET data. *Neuroimage Clin.* 21:101637. doi: 10.1016/j.nicl.2018.101637

Tábuas-Pereira, M., Baldeiras, I., Duro, D., Santiago, B., Ribeiro, M. H., Leitão, M. J., et al. (2016). Prognosis of Early-Onset vs. Late-Onset Mild Cognitive Impairment: Comparison of Conversion Rates and Its Predictors. *Geriatrics* 1:11. doi: 10.3390/geriatrics1020011

Tiwari, S., Atturi, V., Kaushik, A., Yndart, A., and Nair, M. (2019). Alzheimer's disease: pathogenesis, diagnostics, and therapeutics. *Int. J. Nanomed.* 14, 5541–5554. S200490 doi: 10.2147/ijn

Vemuri, P., Lesnick, T. G., Przybelski, S. A., Knopman, D. S., Lowe, V. I., Graff-Radford, J., et al. (2017). Age, vascular health, and Alzheimer disease biomarkers in an elderly sample. *Ann. Neurol.* 82, 706–718. doi: 10.1002/ana.25071

Vermunt, I., Sikkes, S. A. M., van den Hout, A., Handels, R., Bos, I., van der Flier, W. M., et al. (2019). Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. *Alzheimers Dement.* 15, 888–898. doi: 10.1016/j.jalz.2019.04.001

Yoo, J. E., Shin, D. W., Han, K., Kim, D., Won, H. S., Lee, J., et al. (2020). Female reproductive factors and the risk of dementia: a nationwide cohort study. *Eur. J. Neurol.* 27, 1448–1458. doi: 10.1111/ene.14315

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Chen, Qian, Zhang, Su, Huang, Wang, Chen, Zhao, Han and Ma.** This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.