HAL
open science

Statistical methods for dementia risk prediction and recommendations for future work: A systematic review.

Jantje Goerdten, Iva Čukic, Samuel Danso, Isabelle Carrière, Graciela Terrera Muniz

To cite this version:

Jantje Goerdten, Iva Čukic, Samuel Danso, Isabelle Carrière, Graciela Terrera Muniz. Statistical methods for dementia risk prediction and recommendations for future work: A systematic review.. Alzheimer's & Dementia: Translational Research & Clinical Interventions, 2019, 5, pp.563-569. 10.1016/j.trci.2019.08.001 . inserm-02417286

HAL Id: inserm-02417286
https://inserm.hal.science/inserm-02417286

Submitted on 18 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Alzheimer's & Dementia: Translational Research & Clinical Interventions 5 (2019) 563-569

Alzheimer's & Dementia

Review Article

Statistical methods for dementia risk prediction and recommendations for future work: A systematic review

Jantje Goerdten[a,*], Iva Čukić[a], Samuel O. Danso[a], Isabelle Carrière[b], Graciela Muniz-Terrera[a]

[a]Edinburgh Dementia Prevention & Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK
[b]INSERM, Neuropsychiatrie, Recherche Epidemiologique et Clinique, Montpellier, France

Abstract

Introduction: Numerous dementia risk prediction models have been developed in the past decade. However, methodological limitations of the analytical tools used may hamper their ability to generate reliable dementia risk scores. We aim to review the used methodologies.

Methods: We systematically reviewed the literature from March 2014 to September 2018 for publications presenting a *dementia* risk prediction model. We critically discuss the analytical techniques used in the literature.

Results: In total 137 publications were included in the qualitative synthesis. Three techniques were identified as the most commonly used methodologies: machine learning, logistic regression, and Cox regression.

Discussion: We identified three major methodological weaknesses: (1) over-reliance on one data source, (2) poor verification of statistical assumptions of Cox and logistic regression, and (3) lack of validation. The use of larger and more diverse data sets is recommended. Assumptions should be tested thoroughly, and actions should be taken if deviations are detected.

© 2019 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Dementia risk models; Methodological review; Logistic regression; Cox models; Machine learning

1. Introduction

The prevalence of dementia is increasing globally, because of the rapid aging of the population. In 2015, 47 million people were affected by dementia worldwide, whereas dementia prevalence is predicted to almost triple by 2050 [1]. There is no cure for dementia yet; hence, the early identification of individuals at higher risk of developing dementia becomes critical, as this may provide a window of opportunity to adopt lifestyle changes to reduce dementia risk [1,2].

Numerous dementia risk prediction models to identify individuals at higher risk have been developed in the past decade. Three systematic reviews and meta-analyses summarizing dementia risk prediction models were published over the past years [3–5]. Stephan et al. [3] and Tang et al. [4] mainly focused on the critique of the variables selected for inclusion and the assessment of models' prognostic performance, whereas Hou et al. [5] reviewed published dementia risk models in terms of sensitivity, specificity, and area under the curve from receiving operating characteristic analysis. Stephan et al. [3] and Tang et al. [4] concluded that none of the published models could be recommended for dementia risk prediction, largely because of multiple methodological weaknesses of the models or study designs for their derivation. Methodological limitations of the models reviewed included the lack of discrimination of dementia type, lack of internal and external validations of the models, the long interval elapsed between assessments of individuals at risk, and notably, concerns about the analytical techniques used were also highlighted. Hou et al. [5] recommended four

Conflict of Interest: The authors have declared that no conflict of interest exists.

*Corresponding author. Tel: +44 131 537 6530.
E-mail address: jantje.goerdten@ed.ac.uk

https://doi.org/10.1016/j.trci.2019.08.001
2352-8737/ © 2019 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

risk prediction models for different populations (midlife, late-life, patients with diabetes, or mild cognitive impairment (MCI)) with acceptable predictive ability (area under the curve ≥0.74), but still concluded that the models showed methodological limitations, such as lack of external validation.

To date, there is no systematic literature review focusing solely on the methodological approaches used in the dementia risk literature. In the present study, we aim to identify and critically discuss the analytical techniques used in the dementia risk literature and provide suggestions for future prediction model developments, to increase model reliability and accuracy.

2. Methods

2.1. Search strategy

We searched MEDLINE, Embase, Scopus, and ISI Web of Science for articles published from March 1, 2014 to September 17, 2018 using combinations of the following terms: "dementia," "prediction," "development," "receiver operating characteristic," "sensitivity," "specificity," "area under the curve," and "concordance statistic." When possible, terms were mapped to Medical Subject Headings. We searched relevant systematic literature reviews for additional references. March 1, 2014 was chosen as earliest date for this review as it is the upper limit of Tang et al.'s [4] dementia risk review (see Supplementary Material 1 for an example of the search strategy). An updated search was performed from September 17, 2018 to June 12, 2019.

2.2. Selection of studies

First, two independent reviewers (I.Č. and J.G.) screened titles and abstracts for suitable articles. Next, full-text articles were screened for eligibility by one reviewer (J.G.). The following eligibility criteria was used to select the relevant publications: (1) the study has to use a population-based sample or a sample restricted to individuals with MCI; (2) the article provided a model to predict dementia (all-type dementia) risk; (3) the article described the statistical technique that was used for the model development; (4) the article was written in English. Conference abstracts and validation studies were excluded from the review. Any disagreements were resolved by consensus between two authors (I.Č. and J.G.), or if necessary, by a third author (G.M.T.) if the disagreement could not be resolved.

2.3. Data extraction

Data were extracted by three authors (I.Č., J.G., and S.D.) from each article. Information collected included data source, sample size, country, study population, dementia type, length of follow-up, statistical technique used for model development, tested assumptions, and validation method. Only information relevant to our review was extracted from the articles, that is, when studies investigated several aims, we only reported the statistical method and sample that were used for the prediction of dementia risk. In one case a reference is counted twice in the results, as it reports risk models developed from two separate techniques (see Supplementary Material 2 for tables describing publications extracted for review).

3. Results

A total of 2600 nonduplicated articles were identified from the database search and additional relevant references. During the title and abstract review phase 2328 articles were excluded. Full texts were screened for 272 articles, of which 137 were found to be eligible for inclusion in the quantitative synthesis. The most frequent reasons for exclusion were the article was a conference abstract, no prediction model for dementia risk was provided, outcomes other than dementia risk were predicted (e.g., combination of MCI and Alzheimer's disease (AD)), and nonpopulation-based samples or samples not consisting of MCI individuals were used (e.g., sample consisting of menopausal women) (see Fig. 1, for a flow chart of the review, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart diagram template was used [6]).

3.1. Outcomes and populations

Population-based samples were used in 31 (31/138, 22.5%) publications and 107 (107/138, 77.5%) publications used samples comprising MCI individuals for the development of a dementia risk prediction model. In total, 137 study populations were used for the development of the models, of which 74 are unique. In total 60 (60/138, 43.5%) samples were drawn from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The sample size of the studies reviewed ranged from 22 to 331,126 individuals, whereas 17 (17/138, 12.3%) studies had a sample size smaller than 100 participants. The follow-up time ranged from 1 to >30 years. In 103 (103/138, 74.6%) publications AD was the primary outcome. Other dementia types or a combination of dementia types with AD were regarded as the outcome in 35 (35/138, 25.4%) publications, including: dementia any type/not otherwise specified, vascular dementia, mixed dementia, frontotemporal dementia, Huntington disease, Lewy body dementia, multi-infarct type dementia, and Parkinson disease dementia. In 22 (22/138, 15.9%) publications, risk models were externally validated, whereas 46 (46/138, 33.3%) publications did not mention any validation procedure.

3.2. Analytical approaches

Machine learning (n = 55) was the most used technique for the development of dementia risk prediction models. In the publications selected for review, the support vector machine classifier (n = 17) was the most commonly used

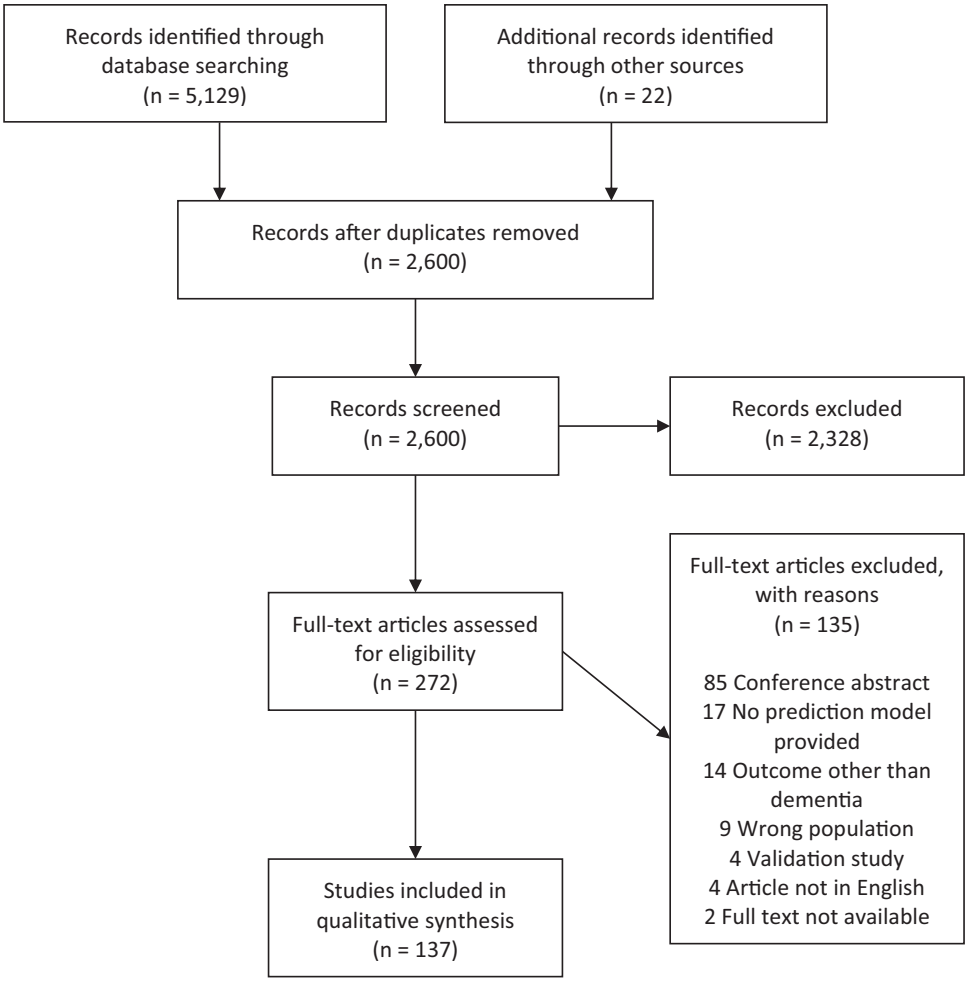

Fig. 1. Flow chart of review phases.

algorithm to predict dementia, followed by the disease state index ($n = 6$) and the random forest classifier ($n = 5$). Three studies used neural network algorithms to construct risk prediction models. Several different feature selection methods are used (including least absolute shrinkage and selection operator, recursive feature elimination, or correlation-based feature selection). In 48 (48/55, 87.3%) of the studies that used machine learning algorithms, prediction models were developed for individuals with MCI, whereas seven (7/55, 12.7%) studies developed models for individuals without clinically impaired cognition. Thirty-four (34/55, 61.8%) publications used the ADNI database. Twelve (12/55, 21.8%) models are externally validated with independent samples, of which 10 were MCI populations and two population-based, whereas 49 (49/55, 89.1%) models are internally validated using different cross-validation methods (e.g., 10-fold cross-validation), two (2/55, 3.6%) models are neither externally nor internally validated.

Logistic regression was used in 31 publications for the development of dementia risk prediction models. One study fitted a multinomial logistic regression including mortality as a third outcome and another study included follow-up time in the model. Eight (8/31, 25.8%) samples are drawn from the ADNI database. Five (5/31, 16.1%) studies checked for multicollinearity among the independent variables and two studies additionally checked the linearity assumption. Two (2/31, 6.5%) studies checked if the data were normally distributed. None of the dementia risk prediction models derived from logistic regression are externally validated. Eleven (11/31, 35.5%) models are internally validated using a cross-validation method, bootstrapping, or by splitting the sample into a testing and validation set, whereas 20 (20/31, 64.5%) models were not validated.

Cox proportional hazards regression (Cox regression) models were used for the development of 25 dementia risk prediction models. Of these, one study used time-dependent covariates in the Cox model, one study included death as a competing risk, one further study used a penalized Cox regression, and another study used age as the time axis. Seven (7/25, 28%) models are developed based on data from the ADNI database. Eight (8/25, 32%) studies verified the proportional hazard assumption

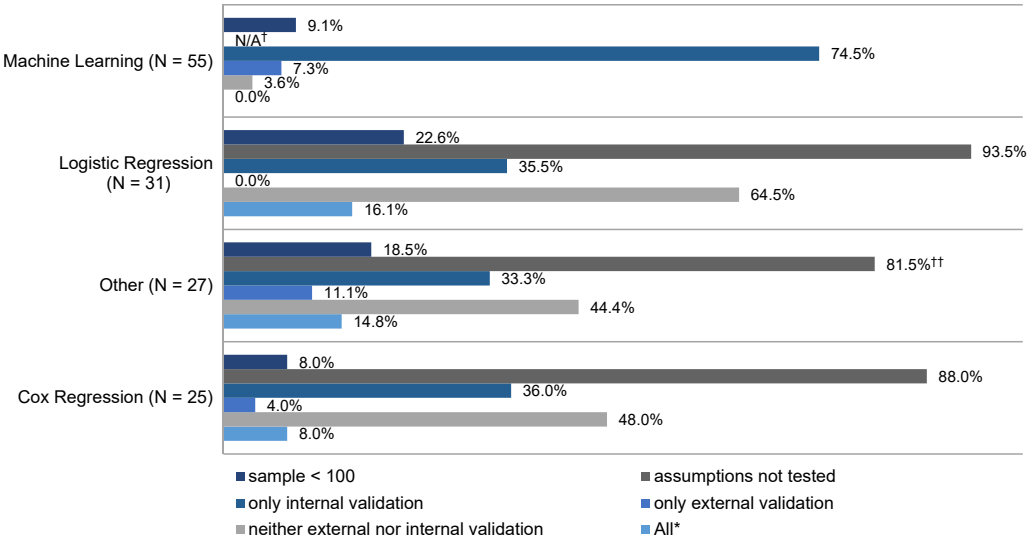

Fig. 2. Limitations of included studies stratified by technique used. *Study used a sample <100 individuals, did not test the assumptions (if applicable), and did not validate the results internally or externally. †Not applicable (N/A). ††Percentage calculated for whole group (22/27); however, only 22 studies needed to test assumptions (22/22, 100%).

and three studies additionally checked the linearity assumption. Four (4/25, 16%) risk models were externally validated with independent samples, of which one was an MCI population and three were population-based. Twelve (12/25, 48%) risk models are validated internally, using cross-validation methods or bootstrapping, whereas 12 (12/25, 48%) risk models are neither externally nor internally validated.

Five studies used a combination of a machine learning approach and a regression analysis (e.g., disease state index and Cox regression), four a combination of two regressions (e.g., logistic and Cox regression), two a joint longitudinal survival model, two an analysis of variance, two a bilinear regression, and two a receiver operating characteristic curve analysis to develop a dementia risk prediction model.

Less frequently used techniques were linear regression ($n = 1$), polynomial regression ($n = 1$), χ^2 test and Kruskal-Wallis test ($n = 1$), power of the t-sum score ($n = 1$), Poisson regression ($n = 1$), illness-death model ($n = 1$), multivariate ordinal regression ($n = 1$), event-based probabilistic model ($n = 1$), mixed linear model ($n = 1$), and general linear model ($n = 1$). An overview of the limitations found in the studies is provided in Fig. 2.

4. Discussion

Our review of analytical approaches in dementia risk prediction identified three techniques as the most commonly used methodologies: machine learning, Cox regression, and logistic regression models.

4.1. Machine learning

A growing number of dementia risk prediction models have been developed using machine learning algorithms [7,8]. Machine learning consists of computational methods, which are able to find meaningful patterns in the data [9], while using experience to improve and make predictions [10]. This means machine learning techniques can explore the structure of the data, in terms of associations between the variables, without having a theory of how the structure looks like. This might make them better suited to detect associations between variables than logistic or Cox regression [11]. However, as discussed by Pellegrini et al. [8], in a published systematic literature and meta-analyses of machine learning techniques in neuroimaging for cognitive impairment and dementia, studies using machine learning algorithms also show limitations. Generalizability of results generated from the application of these techniques and their transfer to clinical use are likely to be constrained because of their over-reliance on one data source, the fact that they commonly use data from populations with greater proportions of cases (i.e., individuals with the diseases) and lower proportions of control subjects, they are usually derived using only one machine learning method and the application of varying validation methods [7,8]. Furthermore, although machine learning methods perform well (accuracy $\geq$0.8) in differentiating healthy control subjects from individuals with dementia, their performance

when identifying individuals at high risk of developing dementia is poorer (accuracy from 0.5 to 0.85) [8]. Similar methodological limitations were found in the present review: more than half the studies used the same data source (34/55, 61.8%), only six (7/55, 12.7%) studies investigated the prediction abilities of a machine learning method in a non-MCI population and only 12 of 55 (21.8%) studies externally validated their model. Relying mainly on one data source and focusing on individuals already at a higher risk of developing dementia results in limitations of clinical relevance and generalizability. Furthermore, without additional studies externally validating these prediction models, it is not clear if these models are overfitted and further limits the generalizability of findings.

4.2. Logistic and Cox regression

Cox and logistic regression, two traditional statistical techniques, are used frequently in dementia risk prediction [4].

Despite their popularity, several features of logistic and traditional Cox regression need to be reflected on when using these methods in dementia risk prediction (for a comparison of these approaches in general settings see Ingram and Kleinman [12] and Peduzzi et al. [13], but it is also worth remembering that although logistic regression aims at the estimation of odds ratios, Cox modeling aims to estimate hazard ratios over time). First, an aspect of both approaches that is relevant to note is that they both generate static risk predictions as they are based on a designated time 0 and on data (baseline covariates) collected at a single time point (time 0 or before). Extensions to time-dependent Cox models exist that are appropriate for use when risk factors themselves change over time [14]. Although not implemented yet in dementia risk prediction, these extended models are likely to be informative in the context of dementia risk prediction as change in predictors over time is likely to be more informative than a single value. However, if prediction is short term, models with time-dependent variables may not be necessary.

Second, although our review identified only one publication where a Cox model based on age was used, the choice of the time axis in Cox modeling is a methodological aspect that also needs consideration as different choices hamper the comparison of results across (and within) studies. When age is used as time axis, the analysis needs to be corrected for delayed entry. A discussion of this issue in the methodological literature and empirical demonstrations showing high sensitivity of results to different choices can be found in Pencina et al. [15].

Third, both methods assume a data structure that may not be adequately fulfilled when used in dementia prediction. Both techniques assume linear relationships between the independent and the dependent variables, that is, a linear relationship between the log of the odds (logistic regression) or the relative risk (Cox regression) and the covariates is assumed. Yet, this assumption is likely not to hold for critical variables used as input in the model (e.g., biomarker) [16]. Notably, our review identified only five of 56 (8.9%) studies that explicitly tested the linearity assumption. Cox regression additionally assumes proportional hazards, which postulates that the impact of a prognostic factor on dementia remains constant over the entire follow-up. Only a third (8/25, 32%) of the identified studies in our review that used Cox regression tested the proportional hazards assumption. Violations of the underlying assumptions in logistic and Cox regression result in biased estimates [16,17]. Furthermore, merely five (5/56, 8.9%) studies incorporated interactions in their regression model. Although an interaction makes it harder to interpret the estimates, it is still relevant and potentially informative to test these.

4.3. Validation, sample size, and data source

External and internal validations are crucial steps when developing a reliable prediction model. Although internal validation ensures the robustness of the findings, that is, there are no alternative explanations for the findings, external validation provides information to which extent results can be generalized, that is, the model can be applied to a wider population than the one from which it was developed [18,19]. Although, the validation phase is highly recommended in prediction models [20], a third of the studies did not perform internal or external validation (46/138, 33.3%). Of the 138 prediction models reviewed, only 14 (10.1%) studies validated their model internally and externally. Too many studies did not perform any validation, whereas too few studies performed both internal and external validations. This is a poor state and the field would benefit from a change in practice.

The data used for the development of a prediction model are as important as the technique used for the derivation of the model. Several studies (17/138, 12.3%) used a sample smaller than 100 participants, which likely is a limitation of these studies. There are no recommendations for a specific sample size, as it dependents on various factors (e.g., which technique is used, number of cases, and number of predictors). Nevertheless, the sample size should be considered when planning a study. The studies reviewed here used 74 unique study populations. The ADNI database was used frequently: 60 (60/138, 43.5%) models were derived using information from subsamples drawn from the ADNI database. Although this overlap makes the results more comparable, it renders at the same time generalizability and might inflate accuracy. For instance, predictions for individuals in ethnic minorities are unlikely to be accurate if models are derived from the ADNI database, where ethnic minorities are largely under-represented, as inferences will be based on a low number of cases and the studies underpowered. Furthermore, the generalizability and replication of findings generated from populations of different sociodemographic

characteristics (age distribution, for instance) and study design (years of follow-up) are likely to be hampered as left censoring will almost certainly operate differently.

4.4. Recommendations

In this review, we identified three major methodological weaknesses, which we encourage researchers to address in future dementia risk prediction work: (1) over-reliance on one data source, (2) the limited evaluation of analytical assumptions of the models used (Cox and logistic regression), and (3) poor internal and external validations of the prediction models. Hence, we suggest the following recommendations to improve the reliability and accuracy of dementia risk prediction models and provide researches with some guidance:

1. A broader selection of data sources should be considered when developing dementia prediction models, including more diverse samples. Although we acknowledge challenges for differentiation between the dementia types, the discrimination of individuals by dementia type will facilitate the identification of risk factors specific to each dementia type. Data sets with different lengths of follow-up time will permit the evaluation of risk progression over different time frames. We encourage researchers to perform where possible, subgroup analyses to evaluate consistency of results in subgroups of similar features.
2. When using regression analyses for dementia risk prediction model development the assumptions need to be tested thoroughly. If deviations from the (linearity) assumptions are detected, appropriate actions need to be taken. There are a number of more flexible nonparametric extensions for regression analyses, through which the linearity assumption can be relaxed: polynomials or restricted cubic splines can be added to the regression model or the predictor can be (log-) transformed [21]. Similarly, the proportional hazard assumption for Cox regression can be relaxed by implementing alternative formulations of the models (i.e., adding splines).
3. Internal and external validations are key steps during the development and implementation of a new prediction model. Internal validation provides insight to which extent the model is overfitted and whether the predictive ability is too optimistic, whereas external validation proves the ability of the prediction model to perform similarly well in a comparable population. There are different internal validation methods, such as splitting the data into two subsets (a development and a validation sample), leave one-out cross-validation or bootstrapping. Bootstrapping is a recommended internal validation method, also when a large number of predictors are used [18]. However, the method might be limited when used in a small sample. For external validation data with a similar but different population to its development population are needed. As mentioned in recommendation 1, more and easily accessible data are required to enable fast and uncomplicated external validation.
4. We encourage researchers to adopt innovative methodologies such as dynamic risk prediction models [22], as the incorporation of within person change in markers of disease progression is likely to be more informative of risk than data collected at a single point in time while also being more likely to reflect clinical practice.
5. We strongly suggest the adoption of Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis guidelines [20] when developing and validating risk prediction models.

4.5. Strength and limitations

The strength of this review is the broad inclusion, allowing a good representation of all possible methodological approaches used in the dementia risk literature. However, we only included published articles and excluded conference abstracts. We also included only population-based samples or samples consisting of MCI individuals, prediction models for other populations (e.g., individuals with Parkinson) might have been developed with different methodological approaches. Furthermore, only dementia was used as a search term, whereby studies looking at specific types of dementia could have been missed.

5. Conclusion

Dementia is one of the leading causes of disability and dependence in late-life [2]. There is a great need to identify individuals at high risk of developing dementia early on. Therefore, the large reliance on one data source, poor validation of results, and limited verification of model assumptions when developing dementia risk prediction models are of concern. It has been shown by Abrahamowicz et al. [16] and Exalto et al. [23] that an application of a more accurate or different analytical technique can result in altered risk prediction. An inaccurate representation of the true relationship of a predictor variable with the outcome might cause false identification of high-risk groups and biased prognosis. To ensure valid conclusions and accurate risk prediction, prognostic studies should rely on statistical methods that correctly represent the actual structure of empirical data and the true complexity of the biological processes under study. Improved practice in data analysis and innovative data designs may advance derivation of dementia risk scores. Machine learning approaches are frequently used for dementia risk prediction model development. As machine learning approaches still need to

improve prediction abilities, regression analyses are robust techniques for prediction model development when applied correctly. Compared with machine learning methods, regression analyses are cost effective and require less computational time.

Advanced and innovative dynamic methods already adopted in other research and clinical areas are likely to be the best choice for future dementia risk prediction developments for now. The community will also benefit from the adoption of new data collection modes to advance knowledge in the short term.

Acknowledgments

Funding for this study was provided by Merck Sharp and Dohme Ltd [grant number 2726958 RB0538].

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.trci.2019.08.001.

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed the literature using traditional sources and references from previous publications.
2. Interpretation: Our findings identified several methodological limitations in the existing literature on dementia risk prediction.
3. Future directions: Future research about dementia risk prediction should use a more thorough methodological approach and devote efforts to ensure fulfilment of assumptions, explore interactions, and validation of results.

References

[1] Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017;390:2673–734.
[2] Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. BMJ 2015;350:h3029.
[3] Stephan BC, Kurth T, Matthews FE, Brayne C, Dufouil C. Dementia risk prediction in the population: are screening models accurate? Nat Rev Neurol 2010;6:318–26.
[4] 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:e0136181.
[5] Hou XH, Feng L, Zhang C, Cao XP, Tan L, Yu JT. Models for predicting risk of dementia: a systematic review. J Neurol Neurosurg Psychiatry 2019;90:373–9.
[6] Moher D, Liberati A, Tetzlaff J, Altman DG, The PG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
[7] Dallora AL, Eivazzadeh S, Mendes E, Berglund J, Anderberg P. Machine learning and microsimulation techniques on the prognosis of dementia: a systematic literature review. PLoS One 2017; 12:e0179804.
[8] Pellegrini E, Ballerini L, Hernandez M, Chappell FM, Gonzalez-Castro V, Anblagan D, et al. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: a systematic review. Alzheimers Dement (Amst) 2018;10:519–35.
[9] Shalev-Shwartz S. In: Ben-David S, ed. Understanding machine learning: from theory to algorithms. Cambridge: Cambridge University Press; 2014.
[10] Mohri M, Rostamizadeh A, Talwalkar A. Foundations of machine learning. MIT Press; 2012:480.
[11] Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? Heart 2018;104:1156–64.
[12] Ingram DD, Kleinman JC. Empirical comparisons of proportional hazards and logistic regression models. Stat Med 1989;8:525–38.
[13] Peduzzi P, Holford T, Detre K, Chan Y-K. Comparison of the logistic and Cox regression models when outcome is determined in all patients after a fixed period of time. J Chronic Dis 1987;40:761–7.
[14] Fisher LD, Lin DY. Time-dependent covariates in the Cox proportional-hazards regression model. Annu Rev Public Health 1999;20:145–57.
[15] Pencina MJ, Larson MG, D'Agostino RB. Choice of time scale and its effect on significance of predictors in longitudinal studies. Stat Med 2007;26:1343–59.
[16] Abrahamowicz M, du Berger R, Grover SA. Flexible modeling of the effects of serum cholesterol on coronary heart disease mortality. Am J Epidemiol 1997;145:714–29.
[17] Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.
[18] Moons KGM, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart 2012;98:683.
[19] Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98:691.
[20] 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; 13:1.
[21] Nick TG, Campbell KM. Logistic regression. In: Ambrosius WT, ed. Topics in biostatistics. Totowa, NJ: Humana Press; 2007. p. 273–301.
[22] Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. Biometrics 2011; 67:819–29.
[23] Exalto LG, Quesenberry CP, Barnes D, Kivipelto M, Biessels GJ, Whitmer RA. Midlife risk score for the prediction of dementia four decades later. Alzheimers Dement 2014;10:562–70.