Aalborg Universitet

Predicting falls in community-dwelling older adults

a systematic review of prognostic models

Gade, Gustav Valentin; Jørgensen, Martin Grønbech; Ryg, Jesper; Riis, Johannes; Thomsen, Katja; Masud, Tahir; Andersen, Stig

Link to publication from Aalborg University

BMJ Open

Predicting falls in community-dwelling older adults: a systematic review of prognostic models

Gustav Valentin Gade [ID] ,[1,2] Martin Grønbech Jørgensen,[1] Jesper Ryg [ID] ,[3,4] Johannes Riis [ID] ,[1,2] Katja Thomsen [ID] ,[3,4] Tahir Masud,[5] Stig Andersen [ID] [1,2]

¹Department of Geriatric Medicine, Aalborg University Hospital, Aalborg, Denmark
²Department of Clinical Medicine, Aalborg University, Aalborg, Denmark
³Department of Geriatric Medicine, Odense University Hospital, Odense, Denmark
⁴Department of Clinical Research, University of Southern Denmark, Odense, Syddanmark, Denmark
⁵Department of Healthcare for Older People, Nottingham University Hospitals NHS Trust, Nottingham, UK

Correspondence to
Dr Gustav Valentin Gade; gustavs@dcm.aau.dk

ABSTRACT

Objective To systematically review and critically appraise prognostic models for falls in community-dwelling older adults.

Eligibility criteria Prospective cohort studies with any follow-up period. Studies had to develop or validate multifactorial prognostic models for falls in community-dwelling older adults (60+ years). Models had to be applicable for screening in a general population setting.

Information source MEDLINE, EMBASE, CINAHL, The Cochrane Library, PsycINFO and Web of Science for studies published in English, Danish, Norwegian or Swedish until January 2020. Sources also included trial registries, clinical guidelines, reference lists of included papers, along with contacting clinical experts to locate published studies.

Data extraction and risk of bias Two authors performed all review stages independently. Data extraction followed the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies checklist. Risk of bias assessments on participants, predictors, outcomes and analysis methods followed Prediction study Risk Of Bias Assessment Tool.

Results After screening 11 789 studies, 30 were eligible for inclusion (n=86 369 participants). Median age of participants ranged from 67.5 to 83.0 years. Falls incidences varied from 5.9% to 59%. Included studies reported 69 developed and three validated prediction models. Most frequent falls predictors were prior falls, age, sex, measures of gait, balance and strength, along with vision and disability. The area under the curve was available for 40 (55.6%) models, ranging from 0.49 to 0.87. Validated models' The area under the curve ranged from 0.62 to 0.69. All models had a high risk of bias, mostly due to limitations in statistical methods, outcome assessments and restrictive eligibility criteria.

Conclusions An abundance of prognostic models on falls risk have been developed, but with a wide range in discriminatory performance. All models exhibited a high risk of bias rendering them unreliable for prediction in clinical practice. Future prognostic prediction models should comply with recent recommendations such as Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis.

PROSPERO registration number CRD42019124021.

Strengths and limitations of this study

► This systematic review is the first to summarise all prediction models on falls in community-dwelling older adults of the general population.
► The extensive search strategy supports identifying all available prospective cohort studies predicting falls in community-dwelling older adults (60+ years).
► Guidelines on prediction modelling reviews were strictly followed for search strings, data extraction (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies), risk of bias assessment (Prediction study Risk Of Bias Assessment Tool), along with development (Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA)-Protocol) and transparent reporting of the review (PRISMA).
► All review stages were performed independently and in duplicate.
► The exclusion of non-English language studies constitutes a risk of selection bias.

with one-third of community-dwelling adults ≥65 years and half of those ≥80 years falling annually.[1] Consequences of falls are considerable with loss of independence, increased morbidity and mortality.[2–4] Furthermore, the healthcare costs of falls increase substantially with age.[5] Therefore, as the prevalence of older fallers is predicted to increase with changes in demography, preventing falls is of utmost importance.[6]

Falls interventions have proven effective when aimed at older adults with a high risk of falling.[7] However, identifying these high-risk individuals is not straight-forward since falling is multifactorial. Prognostic models combine risk factors to estimate the individual's risk of a future outcome.[8] Thus, a prognostic model may be a valuable tool to discriminate between older adults at high versus low risk of falling. To prevent the consequences of falls, healthcare professionals could perform screening in the general population using prognostic models.[9] However, no systematic

BACKGROUND

The propensity to fall is a serious and common health issue among older adults

review has addressed prognostic models on falls for community-dwelling older adults.

This systematic review aims to provide an updated overview of available models to be used by healthcare professionals and for researchers to improve on. The primary objective was to describe the discriminatory performance of prognostic models for falls in prospective cohort studies on community-dwelling older adults. Secondary objectives were to describe the study and model characteristics of these models.

METHODS

A protocol was preregistered before commencing the review process[10] and is available in data supplements (online supplemental appendix 1). The review and its protocol followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (PRISMA)[11] and PRISMA-Protocols,[12] respectively. A completed PRISMA Checklist is available in data supplements (online supplemental appendix 2). During the review process, we found the protocol unclear in terms of eligibility criteria for the study designs, participants, models, outcomes and settings, for which reason we have further described these. Rationales for the changes are given in the protocol. Box 1 provides an introduction to commonly used prediction modelling terms.

Eligibility criteria
Participants and setting
All participants had to be community dwelling, 60 years of age or older, and be recruited from a general population setting. For that reason, we excluded models intended for hospitals, general practitioners and nursing homes. Studies restricted to participants with prespecified diseases, conditions or symptoms such as Parkinson's disease or stroke were excluded to raise external validity. However, we included studies that excluded certain types of community-dwelling older adults, such as those with known neurological, spinal or cognitive disorders.

Index (model)
Studies had to present a final multifactorial prognostic model defined by the inclusion of two prognostic factors or more. This definition was chosen as causes of falls are multifactorial and coexisting.[1] Thus, prognostic factor studies investigating the association between predictors and prospective falls were excluded. Both development studies and validation studies with and without model updating were included.

Outcome
We included studies defining falls as 'an unexpected event in which the participants come to rest on the ground, floor or lower level'.[13] However, studies without an outcome definition were also included since this would not rule out the definition mentioned above. We excluded studies using fall definitions excluding certain

---

Box 1 Commonly used prediction modelling terms with examples related to falls

Prognostic factor
A prognostic factor, also called a predictor, is any measure that, among people with a given health condition, is associated with a subsequent clinical outcome such as falls.[57]

Prognostic prediction model
A prognostic prediction model is a statistical combination of multiple predictors from which risks of a longitudinal outcome, for example, falls, can be calculated for individuals.[8]

Development and validation studies
A prediction model development study aims to develop a prediction model by combining essential predictors from a data set into a model and testing its predictive performance within the same development data set.[54] A model validation study aims to assess the predictive performance of a developed prediction model using new data not used in the development of the model.[54]

Model performance, overfitting and internal validation
A model's predictive performance is termed model performance. This term encompasses several measures with the two most important being discrimination and calibration.[17] Estimates of model performance derived directly from a data set used for developing the model is termed the apparent performance.[54] Since the model is fitted explicitly to the development data set, predictions on new data, that is, new older adults with different characteristics, may yield poorer model performance estimates, that is, poor generalisability. Hence, clinicians would typically find the apparent performance optimistic in terms of predicting a fall in their population which has not been used for developing the model. In consequence, fall preventive interventions could end up being provided to those not needing it and not offered to those actually in need hereof. The optimism in apparent performance is due to the model fitting too well to its data, a term known as overfitting. In such situations, predictions would be biased when the model is used on older adults with different characteristics, that is, frequency distributions of predictors.[17] Estimating the amount of optimism in the development study's model can be done using internal validation techniques such as bootstrap validation. However, since the population of older adults is heterogeneous, generalising a model's performance to the entire population would be more clinically relevant. Here, internal validation procedures fail, and the model should instead be tested in a validation study.

Model discrimination
Model discrimination is a performance measure referring to the models' ability to correctly predict if an individual will experience a fall or not. Therefore, as an example, it can be used by healthcare professionals to assess how confident a model assigns individuals to a high-risk group and guides the clinician when allocating fall preventive interventions.[58] A perfectly discriminating model assigns a higher risk of falling to all older adults experiencing a fall. Likewise, a lower risk is appointed to those not suffering a fall. Usually, discrimination is reported as a concordance index (c-index) or an area under the curve, but other measures are also available. Here, a value of 1 equals perfect discrimination, and 0.5 indicates that the model discriminates no better than chance. If a model shows poor discriminative performance, it could predict low-risk older adults to fall and high-risk older adults to not fall.

Model calibration
Model calibration is a performance measure used to examine whether a model over- or underestimates the predicted risks in a sample. More

Continued

Box 1   Continued

specifically, it is the agreement between predictions made by the model and the frequency of the outcome to be predicted.[54] Healthcare professionals can use this information to assess how confident the model predicts the specific risk of having a fall for the individual. In brief, it is crucial when counselling older adults on their fall risk that the risk estimate is as accurate as possible.[58] If the model predicts a person to have a 10% risk of falling within 1 year, the observed frequency of people falling with such a predicted risk should be 10 out of 100 for the model to have good calibration. However, should the frequency be only 5 of 100 people, the model overestimates the risk. Calibration is typically assessed graphically using calibration plots. In development studies, models are usually calibrated well to the data from which they are developed and therefore yield limited information.[4] Thus, it is more relevant how well the model is calibrated when introduced to a new sample of older adults used to validate the model. This information would enable healthcare professionals to evaluate whether the model over- or underestimates the risk of falling when used in their population of community-dwelling older adults. If the model is not correctly calibrated, it could predict low-risk older adults to have a higher risk and vice versa, or systematically overestimate or underestimate all predictions.

types of falls presumed to be due to a specific cause, for example, external forces or acute medical events. This approach was chosen since postfall classification methods may introduce recall bias.[14] Finally, we did not include studies predicting only injurious falls in older adults since risk factors for these are different from those experiencing non-injurious falls.[15] No restrictions were made on method or timing of outcome assessment other than it had to be prospectively recorded.

Study designs

We only included prospective cohort studies since this study design allows optimal control when measuring predictors and outcomes. Thus, it is the recommended study design for prognostic modelling studies.[16] We excluded randomised controlled trials since these can have different limitations incorporated within their design. Typically, strict eligibility criteria are used that generate a highly selected sample of participants. This narrows predictors' distribution and hence reduces the discriminatory performance in the prognostic models.[17] Also, strict criteria may compromise generalisability to the target population.[18] Lastly, interventions in the study may also influence the discriminatory performance of the models.[18] Retrospective cohort studies were also excluded due to issues of missing data and restrictions on which predictors to apply since data are already collected.[18]

Timing

No restrictions on follow-up or predictive horizon were made since we found it clinically relevant to include models both able to predict falls within short and long periods ahead in time.

Language and publication year

Due to the composition of the study group, we only included published studies reported in English, Danish, Norwegian or Swedish languages. No restrictions on publication year were made.

Information sources

We searched electronic databases, trial registries and clinical guidelines. Furthermore, we consulted with additional clinical experts. Lastly, we screened conference abstracts along with reference lists of both the included studies and systematic reviews found during the search. Databases included Pubmed.gov (PubMed interface), EMBASE (Embase.com), CINAHL (EBSCOhost interface), The Cochrane Library (Wiley interface), PsycINFO (APA PsycNET interface), and Web of Science (Web of Science Core Collection). All databases were searched from inception dates to the 3 January 2020. Trial registries included PROSPERO, ClinicalTrials.gov, WHO International Clinical Trials Registry Platform and Open Grey. Guidelines included Guidelines International Network, the National Institute for Health and Care Excellence, Centre for Reviews and Dissemination and to Health Technology Assessments and Scottish Intercollegiate Guidelines Network. Conference abstracts and studies in trial registries were used to obtain full-text papers through contact with authors. Letters to the editor were excluded.

Search

We used a validated search string for prediction models.[19] With the help of a research librarian in health science, we added the following terms to the search string: independent living, aged and accidental falls. Details on the search string are available online supplemental appendix 3. No search filters were applied. We included 'Aged' as a search term in the search string. Since this would restrict the number of search hits and thus the sensitivity of the search string, we pretested the search string without 'Aged' in all databases before commencing the review. From this, the first 3000 hits were screened independently and in duplicate, and we did not find studies not identified by the final search string. Thus, we believe this had a limited influence on the sensitivity of the search string.

Study selection

Duplicates were removed using EndNote (EndNote X9, Clarivate Analytics, Philadelphia, USA). Two reviewers independently screened titles and abstracts (GVG and JRi) and full-text papers (GVG and KT) according to the inclusion criteria. We contacted authors for clarification when information on this review's eligibility criteria was missing. Disagreement among reviewers was resolved by consensus for one study by including a third reviewer (MGJ). For screening of titles and abstracts along with full-text reading, we used Covidence (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org). Exclusion of studies after full-text reading was performed using a

prioritised list of reasons (online supplemental appendix 4). Reviewers were not blinded to author names, institutions or journal titles.

Data collection process

We developed a standardised data collection form using Research Electronic Data Capture (REDCap),[20] a research electronic data capture software, following the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies checklist.[18] Data extraction was performed in duplicate and independently by two reviewers (GVG and JRi). Independence between reviewers was ensured using a double data entry module in REDCap, thereby denying access to each other's responses. Disagreements among the reviewers were discussed, and the third reviewer (MGJ) was not consulted during data collection since consensus was reached in all studies. We contacted all study authors for retrieval of information on data items not reported. None of the included studies were published more than once.

Data items

We extracted data on the following items: country, publication year, authors, inclusion criteria, exclusion criteria, age, outcome definition, number of falls and fallers, candidate predictors, missing data, choice of statistical analysis, C-statistic and area under the curve (AUC), internal and external validation procedures, final model presentation and sources of funding. If available, 51 data items were extracted from each paper as detailed in online supplemental appendix 5.

Risk of bias and reporting transparency

To follow current recommendations,[21] the Prediction study Risk Of Bias Assessment Tool[22] was used for the risk of bias assessment in individual studies. The tool comprises 20 signalling questions in four domains: participants, predictors, outcomes, and analysis. The tool also includes an evaluation of each model's applicability for the intended population, predictors and outcome of the review. Reporting transparency was assessed using the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis, TRIPOD) adherence assessment form.[23] The bias, applicability and reporting assessments were performed in duplicate and independently by two reviewers (GVG and JRi). Independence between reviewers was ensured using a double data entry module in REDCap, thereby denying access to each other's responses. Disagreements among the reviewers were discussed, and consensus was reached for all studies. Thus, a third reviewer was not consulted for a final decision. The reviewers were not blinded to study authors, institutions or journal titles. The results of the risk of bias assessments of all included studies were incorporated into the qualitative synthesis. We sought to investigate outcome reporting bias by comparing the study papers to their pertaining protocols

to examine whether outcomes were prespecified and not differing from the published paper.

Summary measures and planned method of analysis

The principal summary measure of this systematic review was the discriminatory performance measured either in a C-index or AUC. In the prespecified protocol, we decided not to perform meta-analyses due to the presumed heterogeneity of the prognostic models. This assumption was confirmed after the review was complete. Furthermore, we summarised the study and model characteristics using ranges and percentage proportions when appropriate. When data were available, we summarised continuous measures using medians and IQRs.

Patient public involvement

We did not involve patients or the public in the research.

RESULTS

Study selection

The search yielded 19 612 publications with 11 789 remaining after removal of duplicates. Screening titles and abstracts led to the exclusion of 11 611 publications leaving 178 for full-text reading. Of these, 148 were excluded due to: wrong outcome (n=45), wrong study design (n=45), not being a prediction model (n=25), no full-text paper published (n=14), wrong population (n=8), not multifactorial (n=8) or wrong setting (n=3). Thirty studies met the eligibility criteria and were included.[24–53] Figure 1 displays the PRISMA flow diagram
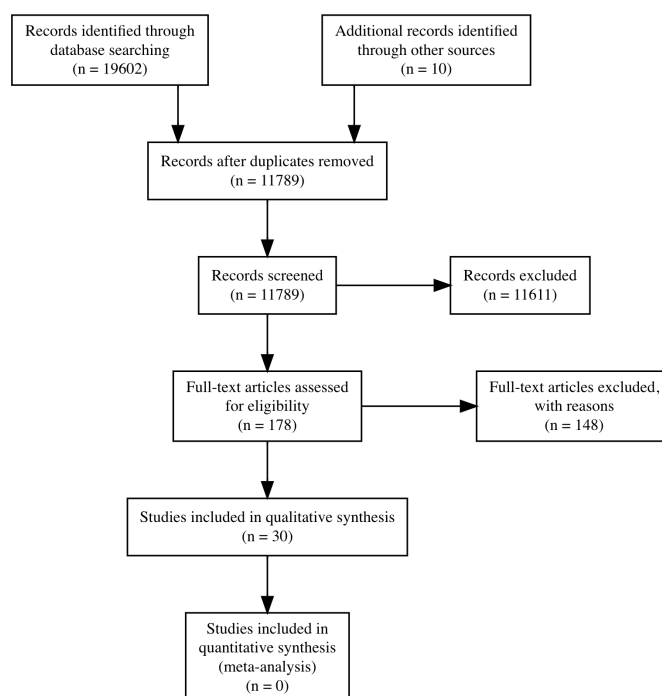

Figure 1 PRISMA diagram of the study selection process. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

for the study selection. Details on excluded papers can be found as online supplemental appendices 6–8.

Characteristics of included studies

A summary of included studies along with models' performance can be found in online supplemental table 1 (online supplemental appendix 9). All studies were published in English from 1994 to 2019. Seventy-two prognostic models were reported, of which 69 models were developed, and three were validated.

Participants

Studies were conducted in Australia, Belgium, Canada, France, Germany, Israel, Italy, Japan, the Netherlands, Malaysia, Spain, the UK and the USA. Sample sizes ranged from 65 to 23 417 participants with median ages from 67.5 to 83 years. Studies used primarily a probability sampling method (n=16), followed by convenience sampling (n=8) and consecutive sampling (n=2). Four studies did not report their sampling methods.

Index/model

The median (IQR) number of predictors in the final models were five (3–9) and ranged from two to 96 predictors. Figure 2 shows the number of studies including a specific predictor. The most frequently applied predictors were prior falls, age, sex, measures of gait, balance and strength, along with vision and disability. Predictors were measured in homes (n=19), research centres (n=19), or both (n=12). Locations for measuring predictors were not reported for 22 models.

Outcomes and timing

The percentage of fallers ranged from 5.9% to 59%, and the number of recurrent fallers (≥2 falls) ranged from 6.3%–54.1%. Models primarily predicted any falls (n=34), that is, single and recurrent falls, and recurrent falls only (n=34). Two models predicted first-time falls[31 45] and two predicted time to a fall.[34 46] Participants were followed for a median (IQR) time of 12 (9.75–12) months. Individual study data items extracted are available in online supplemental appendix 10.

Model performance

Discriminatory measures were reported for 40 (55.6%) models. AUCs were 0.49–0.87 and 0.62–0.69 for developed (n=37) and validated (n=3) models, respectively. Corresponding CIs were reported for 27 (37.5%) models. Calibration measures were available for seven (9.7%) models. For validated models (n=1), calibration was imperfect due to over- and underestimated predicted risks of falling for high-risk and low-risk participants, respectively.[39] Regarding developed models (n=6), calibration was found acceptable, but studies did not assess model calibration in new participants.[41 44 47 49]

Risk of bias, reporting transparency and applicability within studies

Risk of bias

Table 1 summarises ratings on risk of bias, applicability and reporting transparency for the individual studies. All
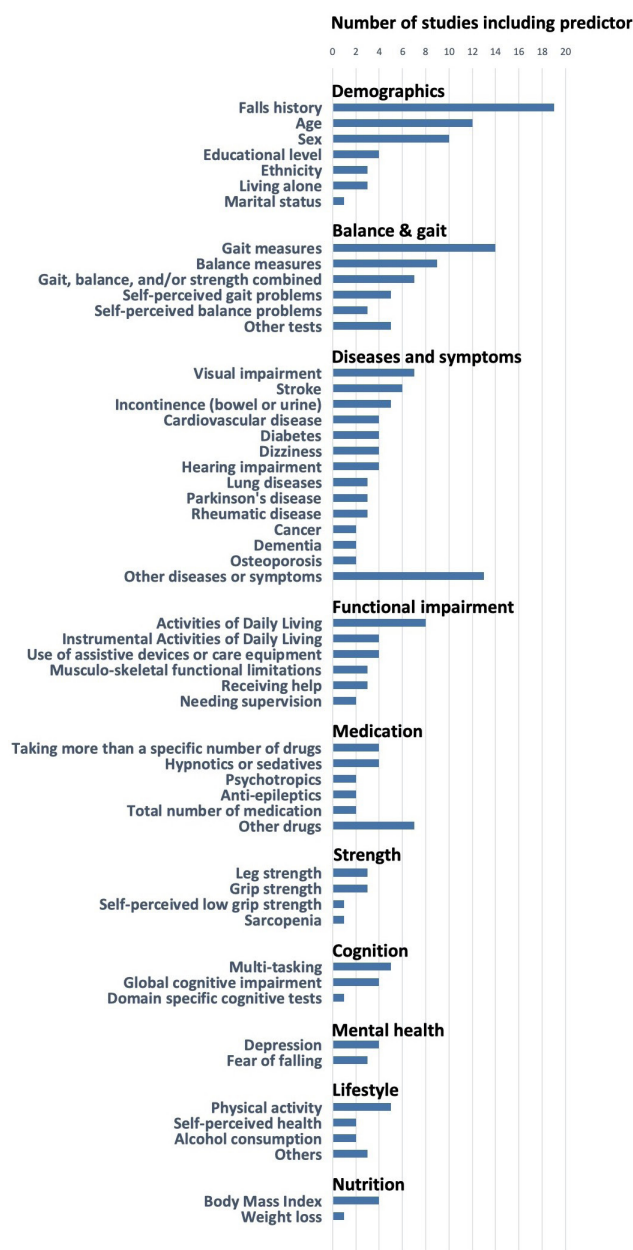

Figure 2 Number of studies using a specific predictor.

studies had a high risk of bias mainly due to methods of analysis and outcome assessment along with restrictive eligibility criteria. Regarding analysis methods, missing data were excluded in 13 out of 30 studies, and no internal validation methods were applied. As to the outcome, only four studies recorded falls daily with monthly notifications.[25 27 32 35] Also, the majority of studies did not report the outcome definition used or whether outcomes assessors were blinded. Eligibility criteria were found restrictive for the majority of studies due to the exclusion of individuals with falls-risk-increasing conditions. These selective criteria limit the usability of models for the target population of community-dwellers. Overall, risk of bias and applicability assessments were complicated by studies only reporting, on average, 50% of all items recommended in reporting guidelines. Furthermore, this

Table 1 Reporting, risk of bias and applicability ratings of individual studies

Study	Reporting TRIPOD, %	Risk of bias				Applicability			Overall	
		Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	ROB	Applicability
Yamashita et al[28]	34	–	+	–	–	–	+	?	–	–
Weiss et al[32]	36	+	+	–	–	–	+	+	–	–
Muhaidat et al[35]	36	+	+	?	–	+	+	?	–	?
Brauer et al[25]	37	–	+	?	–	–	–	+	–	–
Panzer et al[30]	37	–	+	–	–	–	+	?	–	?
Maki et al[24]	41	+	+	?	–	+	+	+	–	+
Boulgarides et al[52]	41	+	+	–	–	+	+	?	–	?
Viccaro et al[31]	41	+	+	–	–	?	+	+	–	?
Lindemann et al[51]	44	–	+	–	–	–	+	+	–	–
Singh et al[44]	44	–	+	?	–	–	+	?	–	–
Nandy et al[53]	45	+	+	–	–	?	+	+	–	?
Okochi et al[50]	45	+	+	–	–	?	+	+	–	?
Rodriguez-Molinero et al[41]	46	–	+	–	–	+	+	+	–	+
Gadkaree et al[37]	48	+	+	–	–	+	+	+	–	+
de Vries et al[34]	50	+	+	–	–	+	+	+	–	+
Hnizdo et al[33]	52	+	+	?	–	–	+	+	–	–
Lohman et al[42]	52	+	+	–	–	+	+	–	–	–
Gillain et al[45]	54	–	+	–	–	–	–	+	–	–
Delbaere et al[27]	55	+	+	–	–	+	+	–	–	–
Cawthon et al[38]	56	+	+	–	–	–	+	+	–	–
Kim et al[43]	56	–	+	–	–	–	+	+	–	–
Tromp et al[36]	57	+	+	–	–	+	+	+	–	+
Stel et al[48]	57	+	+	–	–	+	+	+	–	+
Lamb et al[26]	61	+	+	–	–	–	+	?	–	–
Yamashita et al[29]	61	–	+	–	–	+	+	+	–	–
Stalenhoef et al[47]	62	?	+	–	–	+	+	?	–	?
Palumbo et al[40]	68	+	+	?	–	?	+	+	–	–
Bongue et al[46]	70	–	+	–	–	–	+	?	–	–
Palumbo et al[39]	73	+	+	–	–	+	+	?	–	?
Pluijm et al[49]	76	+	+	–	–	+	+	+	–	+

Note: + and green colour indicates low risk of bias /low concern regarding applicability; – and red colour indicates high risk of bias/high concern regarding applicability; ? and blue colour indicates unclear risk of bias/unclear concern regarding applicability; Bright colours pertain to the PROBAST domains, and dark colours pertain to the overall PROBAST assessments; ROB indicates risk of bias; TRIPOD indicates the percentage proportion of reported items recommended in reporting guidelines.
ROB, risk of bias; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis.

was complicated by a low response rate with four out of 30 study authors responding when contacted for clarification on study characteristics and data extraction items. Finally, outcome reporting bias assessments were not possible due to studies not referring to a preregistered protocol for their prognostic modelling study.

Applicability

Seven (23%) studies, with a total of 21 models, had low applicability concerns for the review question. Regarding participants, 17 (56.7%) studies were rated as having high or unclear applicability concerns for the review question. This concern was primarily due to restrictive eligibility criteria impeding generalisation to the general population of community dwellers. Restrictions were made by excluding participants with specific diseases or conditions that could increase the risk of falling, such as disability or impaired mobility. Furthermore, studies rated as having unclear applicability concerns did not sufficiently report whether the participants were community-dwellers or whether the setting was the general population rather than, for example, primary or secondary care. Regarding predictors, 28 (93.3%) studies had no concerns. The remaining two studies used specific laboratory measures which may be challenging to apply in a general population setting.[25 45] Regarding applicability concerns for the outcome, 18 (60%) studies had no concerns since they reported using the falls definition of the review or similar.

DISCUSSION

The current systematic review found 72 prognostic models on falls risk with the area under the curve ranging from 0.49 to 0.87. All models had a high risk of bias mostly due to limitations in statistical methods and outcome assessments, combined with restrictive eligibility criteria. Thus, using the models in clinical practice would entail unreliable predictions. This review provides an extensive overview of prediction models for falls and information for future study methodology.

Strengths

The current review followed guidelines on prediction modelling reviews strictly for search strings,[19] data extraction,[18] risk of bias assessment,[22] along with development[12 18] and transparent reporting of the review.[11]

Limitations
Review level

We excluded potentially eligible studies without full text (n=14) or published in other languages (n=7) during screening of titles and abstracts. These studies are listed in the data supplement. Furthermore, we excluded randomised controlled trials and retrospective cohort studies. Consequently, we were only able to include 0.25% (30/11 789) of studies screened even though other models, based on other study designs, may had been available. As prespecified in the study protocol, this exclusion

criterion was chosen due to limitations with generalisability and missing data when developing or validating prediction models using these designs. Thus, this systematic review only provides an overview of models based on a specific study design, but we consider this exclusion of the other studies to be justified.

Study level

Limitations were found in the studies with a high risk of bias, poor quality of reporting, and finally, a low response rate when contacting authors for retrieval of missing data extraction items.

Risk of bias

We found a high risk of bias within all studies. Hence, the predictive performance may be low, and predictions unreliable when models are used in clinical practice. The bias ratings were primarily based on eligibility criteria, methods of outcome assessments, and statistical analysis. Building prediction models on selected subgroups of the target population can yield biased performance estimates when used in clinical practice on a different population.[17] Thus, study eligibility criteria should be aligned with the research questions, that is, broad and with as few exclusion criteria as possible. In terms of outcome assessments, we found the definition of falls missing for one-third of studies along with varying falls recording methods. These findings are similar to results of a previous review on methodology in falls prevention trials, where only half of the studies provided a falls definition and recording methods varied highly.[14] The problem with not defining a fall is that the notion of falls is taken for granted. As seen in our review, the prevalence of falls differed markedly between studies, which could be due to different understandings of the fall's definition. Consequently, falls become harder to predict[17] while at the same time, comparing and combining studies in systematic reviews with meta-analyses becomes complicated. To address these issues, a common outcome data set on falls trials is available along with a falls definition and recommendations for falls recording methods.[13] Finally, statistical analysis methods raised concerns for risk of bias. Primarily, this was due to the handling of missing data with most of the studies applying a complete-case analysis method. Significant limitations can arise from the exclusion of participants due to missing data, for example, on a single predictor among many, since otherwise useful predictors on each participant are lost. Consequently, this can lead to low sample sizes and biased model performances. In such cases, imputation methods have proved useful when dealing with missing data.[17] Furthermore, for the majority of developed models, no internal validation procedures were applied. This shortcoming typically causes models' predictive performance to be optimistic.[54] Finally, the critical appraisal was compromised due to incomplete reporting. We believe that future studies and systematic reviews would benefit from adhering to reporting guidelines for prediction modelling studies.[16]

Implications for clinical practice

Only seven studies could address the review question appropriately, and all of these had a high risk of bias. Consequently, the evidence available to inform healthcare professionals is limited and, as mentioned, possibly biased. Thus, no model can currently be recommended for clinical practice.

Implications for research

We recognise that most studies (n=23/30) were conducted before the publishing of prediction modelling guidelines.[16 22] Thus, with the benefit of hindsight, studies would be expected to have different shortcomings within their methods and reporting. On the other hand, this also supports the reason for publishing guidelines in the first place. Despite this, the included studies provide valuable information on future candidate predictors. Thus, selecting predictors for prediction models on non-statistical grounds, that is, based on literature and clinical knowledge, is commonly used to avoid predictor selection bias.[55] Therefore, future development studies may include the most frequently applied predictors found in this review. Lastly, it is essential to test the generalisability of developed models by performing validation studies to determine which models provide stable predictions across different populations.[56]

CONCLUSIONS

There are several studies on falls prognostic models intended for a general population setting, but only a few are fully applicable to the heterogeneous population of community-dwelling older adults. Thus, the evidence available to address this is limited. From all included studies, we found an abundance of falls prognostic models available. However, the discriminatory performance of these varied and was only reported for half of the models. Each model had concerns regarding risk of bias mainly due to restrictive eligibility criteria along with methods of statistical analysis and outcome assessments. Consequently, this could give rise to unreliable predictions should the models be used in clinical practice. Future prognostic prediction models should comply with TRIPOD.

ORCID iDs
Gustav Valentin Gade http://orcid.org/0000-0002-7944-9913
Jesper Ryg http://orcid.org/0000-0002-8641-3062
Johannes Riis http://orcid.org/0000-0001-8398-5420
Katja Thomsen http://orcid.org/0000-0002-0454-836X
Stig Andersen http://orcid.org/0000-0003-3632-5213

REFERENCES

1 Masud T, Morris RO. Epidemiology of falls. *Age Ageing* 2001;30 Suppl 4:3–7.
2 Juel K, Sørensen J, Brønnum-Hansen H. Supplement: risk factors and public health in Denmark. *Scand J Public Health* 2008;36:1–227.
3 Rubenstein LZ. Falls in older people: epidemiology, risk factors and strategies for prevention. *Age Ageing* 2006;35 Suppl 2:ii37–41.
4 Cummings-Vaughn LA, Gammack JK, Falls GJK. Falls, osteoporosis, and hip fractures. *Med Clin North Am* 2011;95:495–506.
5 Scuffham P, Chaplin S, Legood R. Incidence and costs of unintentional falls in older people in the United Kingdom. *J Epidemiol Community Health* 2003;57:740–4.
6 United Nations, Department of Economic and Social Affairs PD. *World population prospects 2019: highlights (ST/ESA/SER.A/423*, 2019.
7 Panel on Prevention of Falls in Older Persons, American Geriatrics Society and British Geriatrics Society. Summary of the updated American geriatrics Society/British geriatrics Society clinical practice guideline for prevention of falls in older persons. *J Am Geriatr Soc* 2011;59:148–57.
8 Steyerberg EW, Moons KGM, van der Windt DA, *et al*. Prognosis research strategy (progress) 3: prognostic model research. *PLoS Med* 2013;10:e1001381.
9 Sattin RW. Falls among older persons: a public health perspective. *Annu Rev Public Health* 1992;13:489–508.
10 Sørensen GVB, Jorgensen MG, Ryg J. Systematic review of prognostic models for predicting falls in community-dwelling older adults (protocol). *Int Prospect Regist Syst Rev* 2019.
11 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535–336.
12 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
13 Lamb SE, Jørstad-Stein EC, Hauer K, *et al*. Development of a common outcome data set for fall injury prevention trials: the

prevention of falls network Europe consensus. *J Am Geriatr Soc* 2005;53:1618–22.

14 Hauer K, Lamb SE, Jorstad EC, *et al*. Systematic review of definitions and methods of measuring falls in randomised controlled fall prevention trials. *Age Ageing* 2006;35:5–10.

15 Tinetti ME, Doucette J, Claus E, *et al*. Risk factors for serious injury during falls by older persons in the community. *J Am Geriatr Soc* 1995;43:1214–21.

16 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med* 2015;13:1.

17 Moons KGM, Wolff RF, Riley RD, *et al*. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1–33.

18 Moons KGM, de Groot JAH, Bouwmeester W, *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the charms checklist. *PLoS Med* 2014;11:e1001744.

19 Geersing G-J, Bouwmeester W, Zuithoff P, *et al*. Search filters for finding prognostic and diagnostic prediction studies in MEDLINE to enhance systematic reviews. *PLoS One* 2012;7:e32844.

20 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.

21 Debray TPA, Damen JAAG, Snell KIE, *et al*. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;356:i6460.

22 Wolff RF, Moons KGM, Riley RD, *et al*. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:51–8.

23 Heus P, Damen JAAG, Pajouheshnia R, *et al*. Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies. *BMJ Open* 2019;9:25611.

24 Maki BE, Holliday PJ, Topper AK. A prospective study of postural balance and risk of falling in an ambulatory and independent elderly population. *J Gerontol* 1994;49:M72–84.

25 Brauer SG, Burns YR, Galley P. A prospective study of laboratory and clinical measures of postural stability to predict community-dwelling fallers. *J Gerontol A Biol Sci Med Sci* 2000;55:M469–76.

26 Lamb SE, McCabe C, Becker C, *et al*. The Optimal Sequence and Selection of Screening Test Items to Predict Fall Risk in Older Disabled Women: The Women's Health and Aging Study. *J Gerontol A Biol Sci Med Sci* 2008;63:1082–8.

27 Delbaere K, Close JCT, Heim J, *et al*. A multifactorial approach to understanding fall risk in older people. *J Am Geriatr Soc* 2010;58:1679–85.

28 Yamashita T, Jeon H, Bailer AJ, *et al*. Fall risk factors in community-dwelling elderly who receive Medicaid-supported home- and community-based care services. *J Aging Health* 2011;23:682–703.

29 Yamashita T, Noe DA, Bailer AJ. Risk factors of falls in community-dwelling older adults: logistic regression tree analysis. *Gerontologist* 2012;52:822–32.

30 Panzer VP, Wakefield DB, Hall CB, *et al*. Mobility assessment: sensitivity and specificity of measurement sets in older adults. *Arch Phys Med Rehabil* 2011;92:905–12.

31 Viccaro LJ, Perera S, Studenski SA. Is timed up and go better than gait speed in predicting health, function, and falls in older adults? *J Am Geriatr Soc* 2011;59:887–92.

32 Weiss A, Brozgol M, Dorfman M, *et al*. Does the evaluation of gait quality during daily life provide insight into fall risk? a novel approach using 3-day accelerometer recordings. *Neurorehabil Neural Repair* 2013;27:742–52.

33 Hnizdo S, Archuleta RA, Taylor B, *et al*. Validity and reliability of the modified John Hopkins fall risk assessment tool for elderly patients in home health care. *Geriatr Nurs* 2013;34:423–7.

34 de Vries OJ, Peeters GMEE, Lips P, *et al*. Does frailty predict increased risk of falls and fractures? A prospective population-based study. *Osteoporos Int* 2013;24:2397–403.

35 Muhaidat J, Kerr A, Evans JJ, *et al*. Validity of simple gait-related Dual-task tests in predicting falls in community-dwelling older adults. *Arch Phys Med Rehabil* 2014;95:58–64.

36 Tromp AM, Pluijm SM, Smit JH, *et al*. Fall-risk screening test: a prospective study on predictors for falls in community-dwelling elderly. *J Clin Epidemiol* 2001;54:837–44.

37 Gadkaree SK, Sun DQ, Huang J, *et al*. Comparison of simple vs. performance-based fall prediction models: data from the National health and aging trends study. *Gerontol Geriatr Med* 2015;1. doi:10.1177/2333721415584850. [Epub ahead of print: 11 May 2015].

38 Cawthon PM, Blackwell TL, Cauley J, *et al*. Evaluation of the usefulness of consensus definitions of sarcopenia in older men: results from the observational osteoporotic fractures in men cohort study. *J Am Geriatr Soc* 2015;63:2247–59.

39 Palumbo P, Klenk J, Cattelani L, *et al*. Predictive performance of a fall risk assessment tool for community-dwelling older people (FRAT-up) in 4 European cohorts. *J Am Med Dir Assoc* 2016;17:1106–13.

40 Palumbo P, Becker C, Bandinelli S, *et al*. Simulating the effects of a clinical guidelines screening algorithm for fall risk in community dwelling older adults. *Aging Clin Exp Res* 2019;31:1069–76.

41 Rodríguez-Molinero A, Gálvez-Barrón C, Narvaiza L, Mi A, Ruiz J, Valldosera E, *et al*. A two-question tool to assess the risk of repeated falls in the elderly. *PLoS One* 2017;12:e0176703.

42 Lohman MC, Crow RS, DiMilia PR, *et al*. Operationalisation and validation of the stopping elderly accidents, deaths, and injuries (STEADI) fall risk algorithm in a nationally representative sample. *J Epidemiol Community Health* 2017;71:jech-2017-209769–-1197.

43 Kim DH, Schneeweiss S, Glynn RJ, *et al*. Measuring frailty in Medicare data: development and validation of a Claims-Based frailty index. *J Gerontol A Biol Sci Med Sci* 2018;73:980–7.

44 Singh DKA, Samah ZA, Shahar S, *et al*. Sociodemographic characteristics as determinants and physical performance measures as correlates of falls among Malaysian community-dwelling older adults. *Top Geriatr Rehabil* 2019;35:113–8.

45 Gillain S, Boutaayamou M, Schwartz C, *et al*. Using supervised learning machine algorithm to identify future fallers based on gait patterns: a two-year longitudinal study. *Exp Gerontol* 2019;127:110730.

46 Bongue B, Dupré C, Beauchet O, *et al*. A screening tool with five risk factors was developed for fall-risk prediction in community-dwelling elderly. *J Clin Epidemiol* 2011;64:1152–60.

47 Stalenhoef PA, Diederiks JPM, Knottnerus JA, *et al*. A risk model for the prediction of recurrent falls in community-dwelling elderly: a prospective cohort study. *J Clin Epidemiol* 2002;55:1088–94.

48 Stel VS, Pluijm SMF, Deeg DJH, *et al*. A classification tree for predicting recurrent falling in community-dwelling older persons. *J Am Geriatr Soc* 2003;51:1356–64.

49 Pluijm SMF, Smit JH, Tromp EAM, *et al*. A risk profile for identifying community-dwelling elderly with a high risk of recurrent falling: results of a 3-year prospective study. *Osteoporos Int* 2006;17:417–25.

50 Okochi J, Toba K, Takahashi T, *et al*. Simple screening test for risk of falls in the elderly. *Geriatr Gerontol Int* 2006;6:223–7.

51 Lindemann U, Lundin-Olsson L, Hauer K, *et al*. Maximum step length as a potential screening tool for falls in non-disabled older adults living in the community. *Aging Clin Exp Res* 2008;20:394–9.

52 Boulgarides LK, McGinty SM, Willett JA, *et al*. Use of clinical and impairment-based tests to predict falls by community-dwelling older adults. *Phys Ther* 2003;83:328–39.

53 Nandy S, Parsons S, Cryer C, *et al*. Development and preliminary examination of the predictive validity of the falls risk assessment tool (Frat) for use in primary care. *J Public Health* 2004;26:138–43.

54 Moons KGM, Altman DG, Reitsma JB, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.

55 Moons KGM, Kengne AP, Woodward M, *et al*. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012;98:683–90.

56 Altman DG, Vergouwe Y, Royston P, *et al*. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605–1435.

57 Riley RD, Hayden JA, Steyerberg EW, *et al*. Prognosis research strategy (progress) 2: prognostic factor research. *PLoS Med* 2013;10:e1001380.

58 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130:515–24.