BMJ Open

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051981
Article Type:	Protocol
Date Submitted by the Author:	01-Apr-2021
Complete List of Authors:	Moura da Silva, Patrícia; UFRN, Physical Therapy; Edmond and Lily Safra International Institute of Neuroscience of Natal Oliveira Bezerra, Ana Beatriz; Federal University of Rio Grande do Norte Araújo Farias, Luanna Barbara; Federal University of Rio Grande do Norte Ribeiro, Tatiana Souza; Federal University of Rio Grande do Norte; Federal University of Rio Grande do Norte Morya, Edgard; Edmond and Lily Safra International Institute of Neuroscience of Natal Cavalcanti, Fabrícia; Federal University of Rio Grande do Norte, Departament of Physical Therapy
Keywords:	DIABETES & ENDOCRINOLOGY, BIOTECHNOLOGY & BIOINFORMATICS, Neurophysiology < NEUROLOGY

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Patrícia Mayara Moura da Silva[1, 3, *], Ana Beatriz de Oliveira Bezerra[2], Luanna Barbara de Araújo Farias[2], Tatiana Souza Ribeiro[1,2], Edgard Morya[3], Fabrícia Azevêdo da Costa Cavalcanti[1,2]

[1]Graduate Program in Physiotherapy, Federal University of Rio Grande do Norte, Natal, RN, Brazil.

[2]Undergraduate Program in Physiotherapy, Federal University of Rio Grande do Norte, Natal, RN, Brazil.

[3]Neuroengineering Program, Edmond and Lily Safra International Institute of Neurosciences, Macaíba, RN, Brazil.

***Corresponding author:** Patrícia Mayara Moura da Silva at Physical Therapy Department, Federal University of Rio Grande do Norte, Natal, RN, 59072-970, Brazil, patriciamayarams@gmail.com

Word Count: 1842.

Abstract

Introduction: Diabetes can lead to gait abnormalities including a longer stance phase, shorter steps, and improper foot pressure distribution. Objective methods to evaluate gait pattern alterations can be decisive for preventing complications caused by diabetes. Besides, it can help predictive models to forecast complications and so develop early strategies to guide treatments. Therefore, the aims are to identify which predictive methods have been employed to assess the diabetic gait and verify which gait data input features are more used to implement a predictive model.

Methods and analysis: A systematic review of studies that evaluated gait in diabetic type 2 using a predictive model. Predictive models are mathematical equations that calculate the probability of an outcome developing in the future. Electronic searches will be performed in the Web of Science, PubMed/MEDLINE, IEEE Xplore Digital Library, Scopus, CINAHL, Embase, The Cochrane Library, and Google Scholar from inception to present. All published and unpublished studies, conference proceedings, or grey literature will also be searched without language restriction. Two independent reviewers will screen all titles, abstracts, and full texts. A third reviewer will be referred to solve any disagreements. This protocol will follow the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA-P) statement.

Ethics and dissemination: This systematic review will only use study-level data from public databases, so formal ethical approval is not required. The results will be disseminated in the form of a peer-reviewed journal and/or presentation at relevant conferences and media.

Trial registration number: PROSPERO (CDR 42020199495).

Article Summary

Strengths and limitations of this study:

- This systematic review protocol will be reported following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA-P) statement.
- This research will be the first systematic review to comprehensively approach the existing predictive models applied to gait analysis of type 2 diabetes patients. Besides, verify these predictive algorithms (e.g., machine learning approach, algorithm type) and input data features characteristics (e.g., data input format).
- Different input gait data features can improve algorithms learning to classify diabetic gait performance. However, heterogeneous database features used in predictive models can limit model generalizability to other settings as they are context-specific.
- This systematic review will focus on predictive model's performance (surrogate outcomes), rather than patient reported outcome measures.

INTRODUCTION

Diabetes is one of the worldwide health concerns, with 8.8% prevalence in 2017[1]. Diabetes type 2 is caused by pancreatic β-cell dysfunction to secret insulin, and insulin resistance in target organs fostered by unhealthy modern habits[1,2]. Patients can also present blood vessel degeneration[3,4] that can evolve into neuropathy and damage sensory and motor nerve fibers[3,4]. Diabetes alters physical function and mobility[5]. Both can lead to motor abnormalities such as slower gait speed and step length and longer stance time and cadence[6,7]. Also, change sensibility on the plantar face can worsen plantar pressure distribution, balance, and gait[4,6].

Boost insight into diabetic gait patterns alterations can be important for preventing complications caused by diabetes and developing strategies to guide treatments[5,8]. However, there is a high prevalence of observational methods[9,10], which may be inaccurate in assessing and diagnosing gait patterns. Subjective methods can impair decision-making for treatment due different interpretations[11-13]. Objective methods are quicker and cost-effective.

Objective gait analysis methods require data collected from patients wearing sensors or performing the gait in specific devices, such as optoelectronic systems or force platforms. The data can be analysed through various methods. One of them is the predictive analysis that can combine all the data collected and estimate probabilities that can aid clinicians and potentially influencing their decision to manage treatment to restore gait[14-17]. Predictive models are mathematical equations (from statistics or machine learning

approaches) that can combine information from a set of data resulting in a response forecasting the probability of a particular outcome[18,19]. Throughout the data collection, these models are trained to achieve an accurate response[14,15].

Emerging predictive methods include machine learning models for automatic gait recognition, opening new perspectives for early identification of gait disorders, and drawing personalized gait training[15]. As well as permit quantify the progress of gait treatment and follow-ups[20].

Based on these grounds, we raise an important question about the existence of predictive methods used to evaluate the gait of diabetic type 2 patients. Therefore, the purpose of this study is to conduct a systematic review of the literature to summarise the evidence of the existing predictive methods used in diabetics' gait patterns. Also, we intend to describe the characteristics of the studies identifying among a variety of gait data collected which input features are most common to implement a predictive method.

METHODS AND ANALYSES

Study design

This systematic review protocol was prepared using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P)[21].

Study registration

This protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO) no. CRD42020199495. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020199495

Eligibility criteria

Types of study

Articles will be eligible for the review when describing the development and/or validation of a predictive model to assess gait in human diabetic type 2. Further, all published and unpublished studies, conference proceedings, or grey literature that deal with diabetic gait analysis, independent of the parameters measured, will be included if they developed and/or validated a predictive model. There will be no geographical or language restriction.

Participants

We will include clinical data from adult participants (> 18 years old) who had type 2 diabetes diagnosed at any disease stage without lower limb amputations or use of gait assistive devices. There will be no restriction on sex and race.

Outcome measures

The primary outcome will comprise all predictive methods (e.g., machine learning models) applied to support gait analysis in diabetes type 2 patients.

The secondary outcome will include gait data input features (e.g., spatiotemporal and angular gait parameters, EMG data, force data, plantar pressure data) are most used to implement a predictive model.

Search strategy for identification of relevant studies

The search strategy will be guided the PRISMA extension for searching (PRISMA-S)[22]. We will search the following electronic bases from their inception to the present: Web of Science (Clarivate Analytics), MEDLINE (PubMed), Embase (Elsevier), IEEE Xplore Digital Library (IEEE), Scopus (Elsevier), CINAHL (EBSCOhost), Google Scholar (Google), and The Cochrane Library (Wiley). We will manually search the references of articles included in the review. All published and unpublished studies, conference proceedings, or grey literature will also be searched without language restriction, and limited to human participants.

The articles will be searched using the terminology registered in the Medical Subject Headings of the U.S. National Library of Medicine (MeSH). The keywords and their synonyms that will be used are related to diabetes terms (e.g., “Diabetes Mellitus, Type 2”), gait terms (e.g., “Gait”, “Gait Analysis”), and prediction-related terms (e.g., “Artificial Intelligence”, “Machine Learning”, "Statistical-learning", “Predictive Value of Tests”). The search strategy for MEDLINE will be adapted to suit the other databases (online supplementary appendix 1).

Screening of the studies

According to the previously described inclusion criteria, two independent reviewers from the group (PMMS, ABOB, LBAF, TSR, EM) will screen titles and abstracts identified during electronic and manual searches to determine the eligibility. Study record information including title and abstract from the searched online database will be imported into Rayyan systematic review software. This platform will guide authors to conduct the literature review process helping explore and filter searched studies. Duplicate studies will be removed, whether duplicity is not explicit, we will contact the authors to solve the problem. If the title or abstract does not provide enough information for inclusion, the full text will be obtained for a full review. The same two review authors independently will screen the full-text articles to identify studies for inclusion, and identify and record reasons for exclusion of the ineligible studies. Any disagreements that arise will be resolved initially by a discussion between the two authors, or if necessary with assistance from a third author (FACC).

All the reasons for the exclusion of ineligible studies will be recorded. The results of the screening process will be provided in detail using the PRISMA information flowchart (Figure 1).

Data extraction

Papers included will go forward to the data extraction and quality assessment stages of the review. Two independent reviewers from the group (PMMS, ABOB, LBAF, TSR, EM) will independently extract outcome data from included studies. A data extraction form was developed through discussion between all authors and adapted from the critical appraisal

and data extraction for systematic reviews of prediction modelling studies (CHARMS) checklist[23]. Disagreements in data extraction will be discussed between the two reviewers and judged by a third reviewer author (FACC), if necessary.

The data collection form will aim to extract the key features of the review. Hence, we will divide items within the data collection form into four blocks: (1) study information including publication year, author information, funding or sponsorship information, type of study, journal name, (exposure), control, population, intervention, and outcome (PICO elements); (2) database information including name, sample size, host organization, and sponsorship; (3) patient demographic information including gender, age, race, and disease severity; (4) predictive methodological information including the type of gait assessment, comparisons with gold-standard devices, type of predictive algorithm used including statistical or machine learning model name, the format of input feature, optimization algorithm, objective function, feature extraction methods, type of extraction feature and computational efficiency, and cost. An example of the data extraction form is presented in Table 1. These data will be presented in a 'Characteristics of included studies' table.

Types of missing data can be missing outcomes, missing summary data, missing individual results. We will consider the reason why the data are missing. Where possible, we will contact the original investigators for any missing data. However, in case of difficult contact, we will present the findings according to the statistical information available in each review, and this will be clearly stated in the final overview.

Table 1: An example of variables collected in the data extraction table

Study information	
Study year	Year of the study publication
Author information	Last name of the author, whether clinical practitioners participated in the study
Type of study	Source of data (e.g., cohort, case-control, randomised trial participants or registry data)
Journal name	Journal name
PICO* elements	PICO* elements in summary
Database information	
Database name	Name of the database used for modelling
Host organisation	Name of the hosting organisation of the database
Sponsorship	The funding or sponsorship information
Sample size	Sample size used for building the model
Source or data	From which source the database was used (e.g., electronic health records, clinical registry, administrative data, cohort study, clinical trial)
Patient demographic information	
Gender	Gender of adults (male, female, both)
Age	Age distribution
Country under study population	At which country the study population was based
Diabetes severity	Disease severity
Predictive methodological information	
Predictors	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation)
Tool used for gait assessment	Quantitative tool used to assess gait kinetic or kinematic (e.g., IMU, force platform, optoelectronic)
Used gold standard devices	Quantitative tool used to assess gait kinetic or kinematic was a device considered gold standards (e.g., force platform, optoelectronic)
Predictive method used	Type of predictive method used to assess gait (e.g., which machine learning techniques was used)
Model name	The name of the predictive model used. The underlying mathematical model used (e.g., linear regression, support vector machine)
Missing data	Number of participants with missing data for each predictor and the process handled with missing data

	(e.g., complete-case analysis, imputation, or other methods)
Format of input feature (predictor or variables)	Which input gait data was used (e.g., plantar pressure, frame, sequence or image)
Number of features	Number of features for building the model
Type of extracted feature	Which features the algorithm uses (e.g., pressure, gait velocity, cadence, step width, pixel feature, action unit, etc)
Selected features	The study reported the importance of selected features?
Model performance/ validation	Performance metrics and scores of how accurate the model used is predicting (e.g., accuracy, average errors, R-squared, confusion matrix etc)
Model evaluation	Method used for testing model performance: development dataset only (random split of data, resampling methods, e.g., bootstrap or cross-validation, none) or separate external validation (e.g., temporal, geographical, different setting, different investigators)
Computational efficiency and cost	Computational efficiency (speed, cloud space, etc) and cost related to the algorithm (e.g., require GPU resources, large cluster, etc)

*PICO: population, intervention (exposure), control, outcome; IMU: inertial measurement unit.

Risk of bias

The pre-selected articles will be evaluated and scored for methodological quality using the Prediction Model Risk of Bias Assessment Tool (PROBAST)[19] by two review authors from the group independently (PMMS, ABOB, LBAF, TSR, EM). The tool comprises a questionnaire of 20 items with four domains (participants, predictors, outcome and analysis) (online supplementary appendix 2). Based on the ratings of signaling questions, risk of bias for each domain will be ranked as low risk, high risk or too unclear for judgement.

PROBAST will be used to categorize the included studies regarding their methodological quality, but these studies will not be excluded based on this evaluation. The classification

of the selected studies will be performed by two independent reviewers from the group (PMMS, ABOB, LBAF, TSR, EM). In cases of a divergence of opinion, a third researcher (FACC) will decide the score.

Quality of evidence

The quality of the predictive model used on the eligible studies will be assessed based on Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) (online supplementary appendix 3) checklist[24]. The TRIPOD Statement is a checklist of 22 items for appropriate reporting of studies developing or validating multivariable prediction models[18]. Each item will be scored as 0, 1, and 2, respectively ranked as no report, inadequate report, and adequate report.

Strategy for data synthesis

A narrative synthesis will be conducted with the information presented in the text and tables to summarize and explain the characteristics and findings of the included studies. Data will be summarized using descriptive statistics and visual plots. Categorical data about the reporting, methodological conduct, and risks of bias will be described by numbers and percentages. The distribution of continuous data, such as sample size and the number of features, will be assessed and described using mean and standard deviation for normally distributed data and median and percentiles (25th and 75th) for non-normally distributed data.

The risk of bias assessment will be summarized and graphically presented for each PROBAST domain and the overall risk of bias judgment. Results will be stratified by prevalent predictive techniques and study design (development with internal validation and/or external validation). The quality of evidence based on TRIPOD will also be summarized and graphically presented for each included study and its respective score rank.

Analyses of subgroups or subsets

We plan to conduct subgroup analyses by predictive model types (e.g., regression models vs. classification models, neural networks vs. traditional machine learning models) and gait input parameters (e.g., kinematic vs. kinetic data features, IMUs vs. EMG data features). Also, subgroup participants' according to anthropometric characteristics (e.g., age, body index mass, and diabetes vitals). More exploratory subgroup analyses will be decided during the process of data extraction and analysis.

DISCUSSION, ETHICS AND DISSEMINATION

According to the best of our knowledge, this systematic review is the first that will investigate evidence regarding which type of predictive methods are used to assessing gait in type 2 diabetic patients. Predictive methods are increasingly being appraised and recommended for formal risk assessment in treatment decision-making and clinical guidelines. The proposed systematic review may contribute to support research and clinicians. For instance, it may help researchers design customizable prediction tools to

be used in diabetic care, and thus physiotherapists better conduct gait treatments in the diabetic type 2 population.

Since we will only be using secondary data sources ethical approval is not required for this systematic review study. Our findings will be disseminated through peer-reviewed publications, presentations at conferences, and through clinical and patient networks.

Acknowledgements: The authors would like to thank CAPES Foundation, an agency under the Ministry of Education (MEC) of Brazil for their assistance during the research.

Author Contributions: PMMS, ABOB, LBAF, FACC, TSR and EM wrote the main manuscript text. PMMS prepared all figures and tables. All authors reviewed the manuscript. All authors read and approved the final manuscript.

Funding statement: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1- Standl E, Khunti K, Hansen T, Schnell O. The global epidemics of diabetes in the 21st century: Current situation and perspectives. European Journal of Preventive Cardiology. 2019;26(2_suppl):7-14.

2- Chatterjee S, Khunti K, Davies M. Type 2 diabetes. The Lancet. 2017;389(10085):2239-2251.

3- Sawacha Z, Cristoferi G, Guarneri G, Corazza S, Donà G, Denti P et al. Characterizing multisegment foot kinematics during gait in diabetic foot patients. Journal of NeuroEngineering and Rehabilitation. 2009;6(1):37.

4- Huang C, Shivaswamy V, Thaisetthawatkul P, Mack L, Stergiou N, Siu K. An altered spatiotemporal gait adjustment during a virtual obstacle crossing task in patients with diabetic peripheral neuropathy. Journal of Diabetes and its Complications. 2019;33(2):182-188.

5- Ko S, Stenholm S, Chia C, Simonsick E, Ferrucci L. Gait pattern alterations in older adults associated with type 2 diabetes in the absence of peripheral neuropathy—Results from the Baltimore Longitudinal Study of Aging. Gait & Posture. 2011;34(4):548-552.

6- Crews R, Yalla S, Fleischer A, Wu S. A Growing Troubling Triad: Diabetes, Aging, and Falls. Journal of Aging Research. 2013.

7- Kirkwood R, Borém I, Sampaio R, Ferreira V, de Almeida J, Guimarães S et al. Frailty Status and Gait Parameters of Older Women With Type 2 Diabetes. Canadian Journal of Diabetes. 2019;43(2):121-127.

8- Wrobel J, Najafi B. Diabetic Foot Biomechanics and Gait Dysfunction. Journal of Diabetes Science and Technology. 2010;4(4):833-845.

9- Prakash C, Kumar R, Mittal N. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. Artificial Intelligence Review. 2016;49(1):1-40.

10- Mundt M, Thomsen W, Witter T, Koeppe A, David S, Bamer F et al. Prediction of lower limb joint angles and moments during gait using artificial neural networks. Medical & Biological Engineering & Computing. 2019;58(1):211-225.

11- Muro-de-la-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A. Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications. Sensors. 2014;14(2):3362-3394.

12- Gor-García-Fogeda M, Cano de la Cuerda R, Carratalá Tejada M, Alguacil-Diego I, Molina-Rueda F. Observational Gait Assessments in People With Neurological Disorders: A Systematic Review. Archives of Physical Medicine and Rehabilitation. 2016;97(1):131-140.

13- Roberts M, Mongeon D, Prince F. Biomechanical parameters for gait analysis: a systematic review of healthy human gait. Physical Therapy and Rehabilitation. 2017;4(1):6.

14- Begg R, Kamruzzaman J. A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data. Journal of Biomechanics. 2005;38(3):401-408.

15- Mannini A, Trojaniello D, Cereatti A, Sabatini A. A Machine Learning Framework for Gait Classification Using Inertial Sensors: Application to Elderly, Post-Stroke and Huntington's Disease Patients. Sensors. 2016;16(1):134.

16- König N, Singh N, Baumann C, Taylor W. Can Gait Signatures Provide Quantitative Measures for Aiding Clinical Decision-Making? A Systematic Meta-Analysis of Gait Variability Behavior in Patients with Parkinson's Disease. Frontiers in Human Neuroscience. 2016;10.

17- Caldas R, Mundt M, Potthast W, Buarque de Lima Neto F, Markert B. A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. Gait & Posture. 2017;57:204-210.

18- Moons K, Altman D, Reitsma J, Ioannidis J, Macaskill P, Steyerberg E et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Annals of Internal Medicine. 2015;162(1):W1.

19- Wolff R, Moons K, Riley R, Whiting P, Westwood M, Collins G et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Annals of Internal Medicine. 2019;170(1):51.

20- Figueiredo J, Santos C, Moreno J. Automatic recognition of gait patterns in human motor disorders using machine learning: A review. Medical Engineering & Physics. 2018;53:1-12.

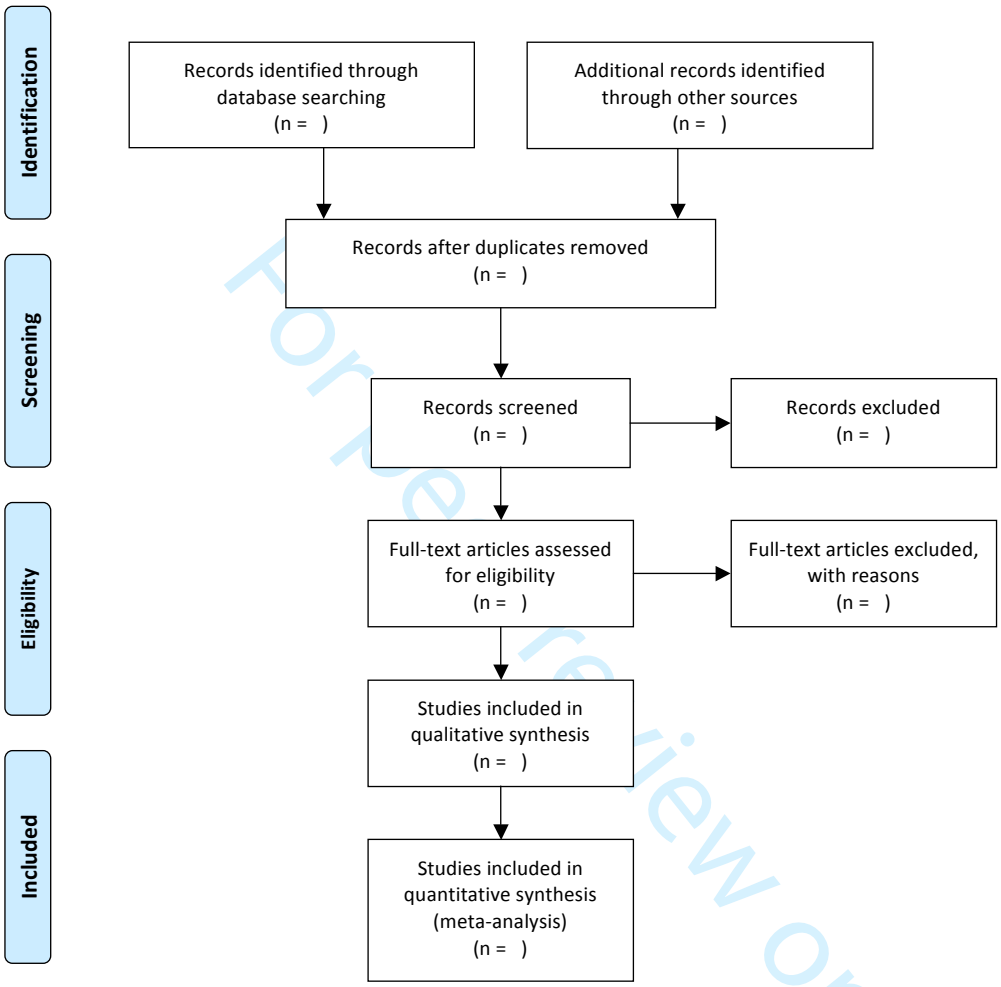
21- Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4:1.

22- Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, Koffel JB; PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev. 2021 Jan 26;10(1):39.

23- Moons K, de Groot J, Bouwmeester W, Vergouwe Y, Mallett S, Altman D et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Medicine. 2014;11(10):e1001744.

24- Collins G, Reitsma J, Altman D, Moons K. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Annals of Internal Medicine. 2015;162(1):55.

Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analysis flow diagram of the identification, screening, and eligibility of included articles.



Supplementary Appendix 1

MEDLINE (PubMed) – search date: March 29, 2021.

Table 1: Search strategy in MEDLINE (PubMed)

1	("Diabetes"[tiab]) OR ("Diabetes Mellitus/analysis"[Mesh] OR "Diabetes Mellitus/classification"[Mesh] OR "Diabetes Mellitus/rehabilitation"[Mesh] OR "Diabetes Mellitus/therapy"[Mesh] OR "Diabetes Mellitus, Type 2"[Mesh]) OR ("Diabetic"[tiab]) OR ("Diabetic Neuropathies/classification"[Mesh] OR "Diabetic Neuropathies/diagnosis"[Mesh] OR "Diabetic Neuropathies/diagnostic imaging"[Mesh] OR "Diabetic Neuropathies/physiopathology"[Mesh] OR "Diabetic Neuropathies/rehabilitation"[Mesh] OR "Diabetic Neuropathies/statistics AND numerical data"[Mesh] OR "Diabetes Complications"[Mesh])
2	("Gait"[tiab]) OR ("Gait Analysis"[tiab]) OR ("Gait/classification"[Mesh] OR "Gait/instrumentation"[Mesh] OR "Gait/methods"[Mesh] OR "Gait/organization and administration"[Mesh] OR "Gait/physiology"[Mesh] OR "Gait/standards"[Mesh] OR "Gait/statistics and numerical data"[Mesh] OR "Gait/trends"[Mesh]) OR ("Gait Disorders, Neurologic"[Mesh]) OR ("Walking Speed"[Mesh] OR "Walking"[tiab]) OR ("Locomotion"[tiab])
3	"Artificial Intelligence"[Mesh] OR Machine Learning[MeSH] OR Deep learning[MeSH] OR "Neural Networks, Computer"[Mesh] OR data mining[MeSH] OR machine[tiab] AND (learn* OR model*) OR (statistical[tiab] OR "statistical-learning"[tiab]) AND (strateg*[tiab]) OR multilayer perceptron*[tiab] OR random forest*[tiab] OR bayes* network*[tiab] OR support vector machine*[tiab] OR nearest neighbor*[tiab] OR k nearest neighbor*[tiab] OR elastic net[tiab] OR naive bayes*[tiab] OR (classification[tiab] OR regression[tiab] OR estimation[tiab] OR decision[tiab]) AND

	tree[tiab] OR ridge[tiab] OR kernel[tiab] OR ensemble[tiab] OR bagging[tiab] OR bagged[tiab] OR boosting[tiab] OR boosted[tiab] OR fuzzy[tiab] OR ("Predictive Value of Tests"[Mesh] OR "Probability Learning"[Mesh] OR "Forecasting"[Mesh] OR "Computing Methodologies"[Mesh] OR "Cluster Analysis"[Mesh]) OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination[tiab] OR discriminative[tiab] OR discriminatory[tiab]) AND (accuracy[tiab] OR ability[tiab] OR performance[tiab] OR value[tiab] OR model[tiab] OR models[tiab] OR power[tiab] OR efficiency[tiab])
4	#1 AND #2 AND #3
5	#4 NOT "review"[pt]

CINAHL search strategy – search date: March 29, 2021.

Table 1: Search strategy in CINAHL (EBSCOhost).

S1	("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR

	"Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications")
S2	("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion")
S3	"Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR machine AND (learn* OR model*) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency)
S4	(S1 AND S2 AND S3) NOT (review)

The Cochrane Library search strategy – search date: March 29, 2021.

Table 1: Search strategy in The Cochrane Library (Wiley).

1	("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications")
2	("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion")
3	"Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR machine AND (learn* OR model*) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND

(accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency)

4 #1 AND #2 AND #3

Embase search strategy – search date: March 29, 2021.

Table 1: Search strategy in Embase (Elsevier).

1	('Diabetes') OR ('Diabetes Mellitus/analysis' OR 'Diabetes Mellitus/classification' OR 'Diabetes Mellitus/rehabilitation' OR 'Diabetes Mellitus/therapy' OR 'Diabetes Mellitus, Type 2') OR ('Diabetic') OR ('Diabetic Neuropathies/classification' OR 'Diabetic Neuropathies/diagnosis' OR 'Diabetic Neuropathies/diagnostic imaging' OR 'Diabetic Neuropathies/physiopathology' OR 'Diabetic Neuropathies/rehabilitation' OR 'Diabetic Neuropathies/statistics AND numerical data' OR 'Diabetes Complications')
2	('Gait') OR ('Gait Analysis') OR ('Gait/classification' OR 'Gait/instrumentation' OR 'Gait/methods' OR 'Gait/organization and administration' OR 'Gait/physiology' OR Gait/standards' OR 'Gait/statistics and numerical data' OR 'Gait/trends') OR ('Gait Disorders, Neurologic') OR ('Walking Speed' OR 'Walking') OR ('Locomotion')
3	('artificial intelligence' OR machine) AND learning OR deep) AND learning OR 'neural networks, computer' OR data) AND mining OR machine) AND (learn* OR model*) OR statistical OR 'statistical-learning') AND strateg* OR multilayer) AND perceptron* OR random) AND forest* OR bayes*) AND network* OR support) AND vector AND machine* OR nearest) AND neighbor* OR k) AND nearest AND neighbor* OR elastic) AND net OR naive) AND bayes* OR classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR 'predictive value of tests' OR 'probability learning' OR 'forecasting' OR 'computing

	methodologies' OR 'cluster analysis' OR validat* OR predict* OR rule* OR (predict* AND outcome*) OR risk* OR model* OR history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) AND (predict* OR model* OR decision* OR identif* OR prognos*) OR (decision* AND model*) OR clinical* OR (prognostic AND history) OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) AND (predict* OR model* OR decision* OR identif* OR prognos*) OR (decision* AND model*) OR clinical* OR (prognostic AND history) OR variable* OR criteria OR OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency)
4	#1 AND #2 AND #3
5	#4 NOT 'review'

Google Scholar – search date: March 29, 2021.

Table 1: Search strategy in Google Scholar.

Diabetes AND Gait "Artificial Intelligence" OR machine OR AND OR learning OR deep OR AND OR learning OR "Neural Networks" OR data OR AND OR mining OR "Predictive Value of Tests" OR "Cluster Analysis" -review

IEEE Xplore Digital Library – search date: March 29, 2021.

Table 1: Search strategy in IEEE Xplore Digital Library (IEEE).

((("All Metadata":"Diabetes") OR ("All Metadata":"Diabetic Neuropathies") OR ("All Metadata":"Diabetes Complications")) AND (("All Metadata":Gait) OR ("All Metadata":"Gait Analysis") OR ("All Metadata":"Walking") OR ("All Metadata":"Locomotion")) AND (("All Metadata":"Artificial Intelligence") OR ("All

Metadata":"Machine Learning") OR ("All Metadata":"Deep learning") OR ("All Metadata":"Neural Networks") OR ("All Metadata":"Predictive Value of Tests")))

Scopus – search date: March 29, 2021.

Table 1: Search strategy in Scopus (Elsevier).

1	ALL(("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications"))
2	ALL(("Gait") OR ("Gait Analysis") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion"))
3	ALL ("Artificial Intelligence" OR machine AND learning OR deep AND learning OR "Neural Networks" OR data AND mining OR "Predictive Value of Tests" OR "Cluster Analysis") AND NOT "review"
4	ALL(("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications")) AND ALL(("Gait") OR ("Gait Analysis") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion")) AND ALL ("Artificial Intelligence" OR machine AND learning

OR deep AND learning OR "Neural Networks" OR data AND mining OR "Predictive Value of Tests" OR "Cluster Analysis") AND NOT "review"

Web of Science – search date: March 29, 2021.

Table 1: Search strategy in Web of Science (Clarivate Analytics).

1	TS=(("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics and numerical data" OR "Diabetes Complications")) *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years*
2	TS=(("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion")) *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years*
3	TS=("Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural N etworks, Computer" OR data mining OR (machine AND (learn* OR model*)) OR (statistical OR ”statistical-learning”) AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive b ayes* OR (classification OR regression OR estimation OR decision) AND tree OR

ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR “Computing Methodologies” OR “Cluster Analysis”) OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency))

Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years

4 #3 AND #2 AND #1

Refined by: DOCUMENT TYPES: (ARTICLE OR MEETING ABSTRACT OR EDITORIAL MATERIAL OR PROCEEDINGS PAPER OR LETTER OR EARLY ACCESS)

Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years

Supplementary Appendix 2

PROBAST

(Prediction model study Risk Of Bias Assessment Tool)

Published in Annals of Internal Medicine (freely available):

1. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies
2. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration

What does PROBAST assess?

PROBAST assesses both the *risk of bias* and *concerns regarding applicability* of a study that evaluates (develops, validates or updates) a multivariable diagnostic or prognostic prediction model. It is designed to assess primary studies included in a systematic review.

Bias occurs if systematic flaws or limitations in the design, conduct or analysis of a primary study distort the results. For the purpose of prediction modelling studies, we have defined *risk of bias* to occur when shortcomings in the study design, conduct or analysis lead to systematically distorted estimates of a model's predictive performance or to an inadequate model to address the research question. Model predictive performance is typically evaluated using calibration, discrimination and sometimes classification measures, and these are likely inaccurately estimated in studies with high risk of bias. *Applicability* refers to the extent to which the prediction model from the primary study matches your systematic review question, for example in terms of the participants, predictors or outcome of interest.

A primary study may include the development and/or validation or update of more than one prediction model. A PROBAST assessment should be completed for each distinct model that is developed, validated or updated (extended) for making individualised predictions. Where a publication assesses multiple prediction models, only complete a PROBAST assessment for those models that meet the inclusion criteria for your systematic review. Please note that subsequent use of the term "model" includes derivatives of models, such as simplified risk scores, nomograms, or recalibrations of models.

PROBAST is not designed for all multivariable diagnostic or prognostic studies. For example, studies using multivariable models to identify predictors associated with an outcome but not attempting to develop a model for making individualised predictions are not covered by PROBAST.

PROBAST includes four steps.

Step	Task	When to complete
1	Specify your systematic review question(s)	Once per systematic review
2	Classify the type of prediction model evaluation	Once for each model of interest in each publication being assessed, for each relevant outcome
3	Assess risk of bias and applicability	Once for each development and validation of each distinct prediction model in a publication
4	Overall judgment	Once for each development and validation of each distinct prediction model in a publication

If this is your first time using PROBAST, we strongly recommend reading the detailed explanation and elaboration (E&E, see link above) paper and to check the examples on www.probast.org

Step 1: Specify your systematic review question

State your systematic review question to facilitate the assessment of the applicability of the evaluated models to your question. *The following table should be completed once per systematic review.*

Criteria	Specify your systematic review question
Intended use of model:	
Participants *including selection criteria and setting:*	
Predictors *(used in prediction modelling), including types of predictors (e.g. history, clinical examination, biochemical markers, imaging tests), time of measurement, specific measurement issues (e.g., any requirements/ prohibitions for specialized equipment):*	
Outcome to be predicted:	

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development		Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation		Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation		External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	
Models of interest	
Outcome of interest	

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain.

The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.

Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection:			
		Dev	Val
1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?			
1.2 Were all inclusions and exclusions of participants appropriate?			
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*		
Rationale of bias rating:			
B. Applicability			
Describe included participants, setting and dates:			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*		
Rationale of applicability rating:			

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment:			
		Dev	Val
2.1 Were predictors defined and assessed in a similar way for all participants?			
2.2 Were predictor assessments made without knowledge of outcome data?			
2.3 Are all predictors available at the time the model is intended to be used?			
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*		
Rationale of bias rating:			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*		
Rationale of applicability rating:			

DOMAIN 3: Outcome			
A. Risk of Bias			
Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:			
		Dev	Val
3.1 Was the outcome determined appropriately?			
3.2 Was a pre-specified or standard outcome definition used?			
3.3 Were predictors excluded from the outcome definition?			
3.4 Was the outcome defined and determined in a similar way for all participants?			
3.5 Was the outcome determined without knowledge of predictor information?			
3.6 Was the time interval between predictor assessment and outcome determination appropriate?			
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*		
Rationale of bias rating:			
B. Applicability			
At what time point was the outcome determined: *If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:*			
Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*		
Rationale of applicability rating:			

DOMAIN 4: Analysis			
Risk of Bias			
Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor:			
Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition):			
Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants):			
Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism:			
Describe any participants who were excluded from the analysis:			
Describe missing data on predictors and outcomes as well as methods used for missing data:			
		Dev	Val
4.1 Were there a reasonable number of participants with the outcome?			
4.2 Were continuous and categorical predictors handled appropriately?			
4.3 Were all enrolled participants included in the analysis?			
4.4 Were participants with missing data handled appropriately?			
4.5 Was selection of predictors based on univariable analysis avoided?			
4.6 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?			
4.7 Were relevant model performance measures evaluated appropriately?			
4.8 Were model overfitting and optimism in model performance accounted for?			
4.9 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?			
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*		
Rationale of bias rating:			

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	
Summary of sources of potential bias:		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	
Summary of applicability concerns:		

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	
	5b	D;V	Describe eligibility criteria for participants.	
	5c	D;V	Give details of treatments received, if relevant.	
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	
Sample size	8	D;V	Explain how the study size was arrived at.	
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	
	10c	V	For validation, describe how the predictions were calculated.	
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	
	15b	D	Explain how to the use the prediction model.	
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Section and topic	Item No	Checklist item	Reported on page #
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	14
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	
Support:			
Sources	5a	Indicate sources of financial or other support for the review	14
Sponsor	5b	Provide name for the review funder and/or sponsor	14
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	14
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	4
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	6
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	6 and 7
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	6 and 7
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	7
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	7
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening,	8

		eligibility and inclusion in meta-analysis)	
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	8
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	8
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	6
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	11 and 12
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	12
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	13
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	13
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	13
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	11
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	12

***It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051981.R1
Article Type:	Protocol
Date Submitted by the Author:	13-Dec-2021
Complete List of Authors:	Moura da Silva, Patrícia; UFRN, Physical Therapy; Edmond and Lily Safra International Institute of Neuroscience of Natal Oliveira Bezerra, Ana Beatriz; Federal University of Rio Grande do Norte Araújo Farias, Luanna Barbara; Federal University of Rio Grande do Norte Ribeiro, Tatiana Souza; Federal University of Rio Grande do Norte; Federal University of Rio Grande do Norte Morya, Edgard; Edmond and Lily Safra International Institute of Neuroscience of Natal Cavalcanti, Fabrícia; Federal University of Rio Grande do Norte, Departament of Physical Therapy
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diabetes and endocrinology, Research methods, Health informatics, Neurology
Keywords:	DIABETES & ENDOCRINOLOGY, BIOTECHNOLOGY & BIOINFORMATICS, Neurophysiology < NEUROLOGY

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Patrícia Mayara Moura da Silva[1, 3, *], Ana Beatriz de Oliveira Bezerra[2], Luanna Barbara de Araújo Farias[2], Tatiana Souza Ribeiro[1,2], Edgard Morya[3], Fabrícia Azevêdo da Costa Cavalcanti[1,2]

[1]Department of Physical Therapy, Graduate Program, Federal University of Rio Grande do Norte, Natal, RN, Brazil.

[2]Department of Physical Therapy, Undergraduate Program, Federal University of Rio Grande do Norte, Natal, RN, Brazil.

[3]Edmond and Lily Safra International Institute of Neurosciences, Neuroengineering Program, Santos Dumont Institute, Macaíba, RN, Brazil.

***Corresponding author:** Patrícia Mayara Moura da Silva at Physical Therapy Department, Federal University of Rio Grande do Norte, Natal, RN, 59072-970, Brazil, patriciamayarams@gmail.com

Word Count: 2,031

ABSTRACT

Introduction: Type 2 diabetes can lead to gait abnormalities, including a longer stance phase, shorter steps, and improper foot pressure distribution. Quantitative data from objective methods for evaluating gait patterns are accurate and cost-effective. In addition, it can also help predictive methods to forecast complications and develop early strategies to guide treatments. To date, no research has systematically summarised the predictive methods used to assess type 2 diabetic gait. Therefore, this protocol aims to identify which predictive methods have been employed to assess the diabetic gait.

Methods and analysis: This protocol will follow the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) statement. Electronic searches will be performed in the Web of Science, MEDLINE, Embase, IEEE Xplore Digital Library, Scopus, CINAHL, Google Scholar, and the Cochrane Library, as well as in reference lists from key articles and grey literature without language restrictions, from May 2021 to 31 January 2022. We will include studies that examined the development and/or validation of predictive methods to assess type 2 diabetic gait in adults aged > 18 years without amputations, use of assistive devices, ulcers, or neuropathic pain. Two independent reviewers will screen the search results and extract the data using a customised charting form from the included articles. A third reviewer will resolve any disagreements. A narrative synthesis will be performed for the included studies. Risk of bias and quality of evidence will be assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD).

Ethics and dissemination: Ethical approval is not required because only available secondary published data will be analysed. The findings will be disseminated through peer-reviewed journals and/or presentations at relevant conferences and other media platforms.

Trial registration number: PROSPERO (CDR 42020199495).

ARTICLE SUMMARY

Strengths and limitations of this study:

- This study will be the first systematic review to comprehensively analyse the existing predictive methods for gait analysis in patients with type 2 diabetes.
- This systematic review will focus on the predictive method's performance (surrogate outcomes) rather than on patient-reported outcome measures.
- A broad search strategy and robust quality assessment criteria will be used to appraise and examine the existing literature.
- Two independent reviewers will be responsible for conducting the study selection, data extraction, and quality assessment.
- A limitation could be the potential lack of studies that meet the established inclusion criteria.

INTRODUCTION

Diabetes mellitus is a worldwide health concern, with a prevalence of 8.8% in 2017[1]. With type 2 diabetes mellitus being the most common[2], this condition is related to a dysfunction either in the pancreatic β-cells' ability to secrete insulin, insulin resistance in target organs, or a combination of both, resulting in hyperglycaemia[1,3]. Patients can also present with blood vessel degeneration[4,5] that can evolve into neuropathy and damage sensory and motor nerve fibres[4,5]. Diabetes also alters physical function and mobility[6]. Both can lead to motor abnormalities such as longer stance time (i.e., greater support base) and shorter steps, which may exhibit as slower gait speeds and increased cadence[7,8]. In addition, changes in the sensibility of the plantar surface of the foot can worsen plantar pressure distribution, balance, and gait[5,7].

Boosting insight into diabetic gait pattern alterations can be important for preventing complications caused by diabetes and developing strategies to guide treatments[6,9]. In clinical practice, while there is a high prevalence of observational methods[10,11], this may be unreliable in assessing and diagnosing gait patterns. Observational methods are subjective and can generate inaccuracies during the assessment and diagnosis of the patient's movements due to different interpretations between examiners. In addition, these differences can impair decision making to address a specific treatment[12-14]. On the contrary, objective methods are reliable, accurate, quicker, and cost-effective owing to quantitative metric results[8].

Objective gait analysis methods require data collected from patients wearing sensors or performing gait in specific devices, such as Inertial Motion Units (IMUs), electromyography (EMG), optoelectronic systems, or force platforms[10]. Data from quantitative gait measures can be analysed using various methods. One of these is the use of predictive analytics that combine the collected data and estimate probabilities that can assist clinicians and potentially influence their decision to manage treatments to restore gait[15-18]. Predictive methods are mathematical equations (from statistics or machine learning approaches) that can combine information from a set of data, resulting in a response forecasting the probability of a particular outcome[19,20].

Emerging predictive methods include machine learning (ML) algorithms. ML can be used for automatic gait recognition to predict possible complications such as the risk of falls and pressure ulcers[21]. Newer methods have opened new perspectives for the early diagnosis of gait disorders. This is essential in preventing potential future complications and to draw on personalised gait training[15] by quantifying the treatment progress and follow-ups[22].

To our knowledge, no research has systematically summarised predictive algorithms used to assess gait in patients with type 2 diabetes. Based on this, we raise an important question about the existence of predictive methods used to evaluate the gait of patients with type 2 diabetes. Therefore, the purpose of this study is to conduct a systematic review of the literature to summarise the evidence regarding existing predictive methods used in the gait patterns of patients with diabetes. In addition, we intend to describe the

characteristics of the studies identified among the variety of gait data collected regarding which input features are the most commonly used to implement a predictive method.

METHODS AND ANALYSES

Study design

This systematic review protocol was prepared using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P)[23].

Study registration

This protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO) no. CRD42020199495. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020199495

Eligibility criteria

Types of study

Articles will be eligible for review when they describe the development and/or validation of a predictive method to assess gait in human type 2 diabetes. Furthermore, all published and unpublished studies (e.g., dissertations and theses), conference proceedings that deal with diabetic gait analysis, independent of the parameters measured, will be included if developed and/or validated as a predictive method. There will be no geographical or language restrictions. Wherever necessary, relevant articles will be arranged for translation.

Participants

We will include clinical data from adult participants (> 18 years old) who had type 2 diabetes diagnosed at any disease stage without lower limb amputations or the use of gait assistive devices. In addition, data with participants with ulcers or neuropathic pain (that could have interfered in the gait execution) will be excluded. There will be no restrictions on sex or race.

Outcome measures

The primary outcome will comprise all predictive methods (e.g., machine learning models) applied to analyse gait in patients with type 2 diabetes. The secondary outcome will include gait data input features (e.g., spatiotemporal, angular gait parameters, EMG data, force data, and plantar pressure data) most commonly used to implement a predictive model.

Search strategy for identification of relevant studies

The search strategy will be guided by the PRISMA extension for searching (PRISMA-S)[24]. The following electronic databases will be searched: Web of Science (Clarivate Analytics), MEDLINE (PubMed), Embase (Elsevier), IEEE Xplore Digital Library (IEEE), Scopus (Elsevier), CINAHL (EBSCOhost), Google Scholar (Google), and the Cochrane Library (Wiley) from May 2021 to 31 January 2022. The time range of the published studies was from inception to January 2022. We will manually search the reference list of the studies included in the review. Grey literature involving published and unpublished

studies (e.g., dissertations and theses) and conference proceedings will also be searched without language restrictions, but this must be limited to human participants.

The articles will be searched using a combination of free keywords and the terminology registered in the Medical Subject Headings (MeSH) of the U.S. National Library of Medicine. The terms that will be used are related to diabetes (e.g., “Type 2 Diabetes”, “Diabetes, Type 2”, “Diabetes Mellitus, Type 2”), gait (e.g., “Gait”, “Gait Analysis”, “Kinematic”, “Kinetic”, “Range of Motion”), and prediction-related (e.g., “Artificial Intelligence”, “Machine Learning”, "Statistical-learning", “Predictive Value of Tests”, “Support Vector Machine”, “Neural Networks, Computer”). The search strategy was pilot tested and finalised in MEDLINE (PubMed) before being translated for use in other databases. Details of the search strategies are provided in Online Supplementary Appendix 1.

Screening of the studies

Based on the previously described inclusion criteria, two independent reviewers (PMMS, ABOB) will screen titles and abstracts identified during electronic and manual searches to determine its eligibility. Study record information, including title and abstract from the searched online database, will be imported into the Rayyan systematic review software[25]. This platform will guide the reviewers in conducting the literature review process through its ability to explore and filter searched studies. Duplicate studies will be removed. If the title or abstract does not provide enough information for inclusion, the full text will be obtained for a full review. The same two reviewers (PMMS, ABOB) will independently

screen the full-text articles to identify studies for inclusion and record the reasons for exclusion for ineligible studies. Any disagreements that arise will be resolved initially by a discussion between the two reviewers, or, if necessary, with assistance from a third reviewer (FACC).

All reasons for the exclusion of ineligible studies will be recorded. The results of the screening process will be provided in detail using the PRISMA information flowchart (Figure 1).

Data extraction

A data extraction form was developed through a discussion among all authors and adapted from the critical appraisal and data extraction for systematic reviews of prediction modelling studies (CHARMS) checklist[26]. The included studies will go forward to the data extraction and quality assessment stages of the review. Two independent reviewers (PMMS, ABOB) will extract the outcome data from the included studies. If necessary, disagreements in data extraction will be discussed between the two reviewers and judged by a third reviewer (FACC).

The data collection form will aim to extract the key features of the review. Hence, we will divide the items within the data collection form into four blocks: (1) study information including publication year, author information, funding or sponsorship information, type of study, journal name, control, population, intervention, and outcome (PICO elements); (2) database information including name, sample size, host organisation, and sponsorship;

(3) patient demographic information including sex, age, race, and disease severity; and (4) predictive methodological information including the type of gait assessment, comparisons with gold standard devices, type of predictive algorithm used (including its statistical or machine learning model name), format of input feature, optimisation algorithm, objective function, feature extraction methods, type of extraction feature and computational efficiency, and cost. Table 1 presents an example of the data extraction form. These data will be presented in the 'Characteristics of included studies' table.

Missing data may include missing outcomes, missing summary data, or missing individual results. The authors will consider the reasons for the missing data. Where possible, we will contact the original investigators to obtain any missing data. However, in the case of contact difficulty, we will present the findings according to the statistical information available in each review, and this will be clearly stated in the final overview.

Table 1: Example of the data extraction form for all included studies.

Study information	
Study year	Year of the study publication
Author information	Last name of the author, whether clinical practitioners participated in the study
Type of study	Source of data (e.g., cohort, case-control, randomised trial participants or registry data)
Journal name	Journal name
PICO* elements	PICO* elements in summary
Database information	
Database name	Name of the database used for modelling
Host organisation	Name of the hosting organisation of the database
Sponsorship	The funding or sponsorship information
Sample size	Sample size used for building the model

Source or data	From which source the database was used (e.g., electronic health records, clinical registry, administrative data, cohort study, clinical trial)
Patient demographic information	
Sex	Sex of adults (male, female, both)
Age	Age distribution
Country under study population	At which country the study population was based
Diabetes severity	Disease severity
Predictive methodological information	
Predictors	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation)
Tool used for gait assessment	Quantitative tool used to assess gait kinetic or kinematic (e.g., IMU, force platform, optoelectronic, EMG)
Used gold standard devices	Quantitative tool used to assess gait kinetic or kinematic was a device considered gold standards (e.g., force platform, optoelectronic)
Predictive method used	Type of predictive method used to assess gait (e.g., which machine learning techniques was used)
Model name	The name of the predictive model used. The underlying mathematical model used (e.g., linear regression, support vector machine)
Missing data	Number of participants with missing data for each predictor and the process handled with missing data (e.g., complete-case analysis, imputation, or other methods)
Format of input feature (predictor or variables)	Which input gait data was used (e.g., plantar pressure, frame, sequence or image)
Number of features	Number of features for building the model
Type of extracted feature	Which features the algorithm uses (e.g., pressure, gait velocity, cadence, step width, pixel feature, action unit, etc.)
Selected features	The study reported the importance of selected features?
Model performance/ validation	Performance metrics and scores of how accurate the model used is predicting (e.g., accuracy, average errors, R-squared, confusion matrix, etc.)
Model evaluation	Method used for testing model performance: development dataset only (random split of data, resampling methods, e.g., bootstrap or cross-validation) or separate external validation (e.g., temporal, geographical, different setting, different investigators)
Computational efficiency and cost	Computational efficiency (speed, cloud space, etc.) and cost related to the algorithm (e.g., require GPU resources, large cluster, etc.)

*PICO: population, intervention (exposure), control, outcome; IMU: inertial measurement unit; EMG: electromyography; GPU: graphics processing unit.

Risk of bias

The pre-selected articles will be evaluated and scored for methodological quality using the Prediction Model Risk of Bias Assessment Tool (PROBAST)[20] by two independent reviewers (PMMS, ABOB). In cases of opinion divergence, a third reviewer (FACC) will provide the judgment. The questionnaire is comprised of 20 items with four domains (participants, predictors, outcome, and analysis). Based on the questionnaire ratings, the risk of bias for each domain will be ranked as 'low risk', 'high risk', or 'too unclear for judgment'.

PROBAST will be used to categorise the included studies regarding their methodological quality, but these studies will not be excluded based solely on this evaluation. The classification of the selected studies will be performed by two independent reviewers (PMMS, ABOB). In cases of opinion divergence, a third reviewer (FACC) will decide the score.

Quality of evidence

The quality of the predictive model used on the eligible studies will be assessed based on Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) checklist[27]. The TRIPOD Statement is a checklist of 22 items for the appropriate reporting of studies developing or validating multivariable prediction

models[19]. Each item will be scored as 0, 1, and 2, ranked as 'no report', 'inadequate report', and 'adequate report', respectively.

Strategy for data synthesis

A narrative synthesis will be conducted with the information presented in the text and tables to summarise and explain the characteristics and findings of the included studies. Data will be summarised using descriptive statistics and visual plots. Categorical data about the reporting, methodological conduct, and risks of bias will be described using numbers and percentages. The distribution of continuous data, such as sample sizes and the number of features, will be described using measures of central tendency such as mean and standard deviation for normally distributed data and median and percentiles (25th and 75th) for non-normally distributed data.

The risk of bias assessment will be summarised and graphically presented for each PROBAST domain and the overall risk of bias judgment. The results will be stratified by prevalent predictive techniques and study design (development with internal validation and/or external validation). The quality of evidence based on TRIPOD will also be summarised and graphically presented for each included study and its respective score rank.

Analyses of subgroups or subsets

We plan to conduct subgroup analyses using predictive model types (e.g., regression models vs. classification models, neural networks vs. traditional machine learning

models) and gait input parameters (e.g., kinematic vs. kinetic data features, IMUs vs. EMG data features). In addition, we plan to classify participants according to their anthropometric characteristic subgroup (e.g., age, body index mass, and diabetes vitals). More exploratory subgroup analyses will be decided during the data extraction and analysis process.

ETHICS AND DISSEMINATION

To the best of our knowledge, this systematic review is the first that will synthesize existing evidence regarding the types of predictive methods used to assess gait in patients with type 2 diabetes. Predictive methods are increasingly being appraised and recommended for formal risk assessment in treatment decision making and clinical guidelines. The proposed systematic review may inform future research and clinicians. For instance, it may help researchers in designing customisable prediction tools to be used in diabetic care, and thus allow physiotherapists to better conduct rehabilitative gait treatments in the patients with type 2 diabetes.

Because we will be using secondary data sources, ethical approval is not required for this systematic review study. Our findings will be disseminated through peer-reviewed publications, presentations at conferences, and clinical and patient networks.

Acknowledgements: The authors would like to thank CAPES Foundation, an agency under the Ministry of Education (MEC) of Brazil for their assistance during the research.

Authors' Contributions: PMMS, ABOB, LBAF, TSR, EM, and FACC wrote the main manuscript text. PMMS prepared all figures and tables. All authors reviewed the manuscript. All authors read and approved the final manuscript.

Funding statement: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Patient and Public Involvement: Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

References

1- Standl E, Khunti K, Hansen T, et al. The global epidemics of diabetes in the 21st century: Current situation and perspectives. Eur J Prev Cardiol 2019;26(2_suppl):7-14.

2- Skyler JS, Bakris GL, Bonifacio E, et al. Differentiation of diabetes by pathophysiology, natural history, and prognosis. Diabetes 2017;*66*(2):241-255.

3- Chatterjee S, Khunti K, Davies M. Type 2 diabetes. Lancet 2017;389(10085):2239-2251.

4- Sawacha Z, Cristoferi G, Guarneri G, et al. Characterizing multisegment foot kinematics during gait in diabetic foot patients. J Neuroeng Rehabil 2009;6(1):37.

5- Huang C, Shivaswamy V, Thaisetthawatkul P, et al. An altered spatiotemporal gait adjustment during a virtual obstacle crossing task in patients with diabetic peripheral neuropathy. J Diabetes Complications 2019;33(2):182-188.

6- Ko S, Stenholm S, Chia C, et al. Gait pattern alterations in older adults associated with type 2 diabetes in the absence of peripheral neuropathy—Results from the Baltimore Longitudinal Study of Aging. Gait Posture 2011;34(4):548-552.

7- Crews R, Yalla S, Fleischer A, et al. A Growing Troubling Triad: Diabetes, Aging, and Falls. J Aging Res 2013.

8- Kirkwood R, Borém I, Sampaio R, et al. Frailty Status and Gait Parameters of Older Women With Type 2 Diabetes. Can J Diabetes 2019;43(2):121-127.

9- Wrobel J, Najafi B. Diabetic Foot Biomechanics and Gait Dysfunction. J Diabetes Sci Technol 2010;4(4):833-845.

10- Prakash C, Kumar R, Mittal N. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. Artif Intell Rev 2018;49(1):1-40.

11- Mundt M, Thomsen W, Witter T, et al. Prediction of lower limb joint angles and moments during gait using artificial neural networks. Med Biol Eng Comput 2019;58(1):211-225.

12- Muro-de-la-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A. Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications. Sens 2014;14(2):3362-3394.

13- Gor-García-Fogeda M, Cano de la Cuerda R, Carratalá Tejada M, et al. Observational Gait Assessments in People With Neurological Disorders: A Systematic Review. Arch Phys Med Rehabil 2016;97(1):131-140.

14- Roberts M, Mongeon D, Prince F. Biomechanical parameters for gait analysis: a systematic review of healthy human gait. Phys Ther Rehabil 2017;4(1):6.

15- Begg R, Kamruzzaman J. A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data. J Biomech 2005;38(3):401-408.

16- Mannini A, Trojaniello D, Cereatti A, et al. A Machine Learning Framework for Gait Classification Using Inertial Sensors: Application to Elderly, Post-Stroke and Huntington's Disease Patients. Sens 2016;16(1):134.

17- König N, Singh N, Baumann C, et al. Can Gait Signatures Provide Quantitative Measures for Aiding Clinical Decision-Making? A Systematic Meta-Analysis of Gait

Variability Behavior in Patients with Parkinson's Disease. Front Hum Neurosci 2016;10.

18- Caldas R, Mundt M, Potthast W, et al. A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. Gait Posture 2017;57:204-210.

19- Moons K, Altman D, Reitsma J, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 2015;162(1):W1.

20- Wolff R, Moons K, Riley R, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med 2019;170(1):51.

21- Khera P, Kumar N. Role of machine learning in gait analysis: a review. J Med Eng Technol, Taylor & Francis 2020;44(8):441-467.

22- Figueiredo J, Santos C, Moreno J. Automatic recognition of gait patterns in human motor disorders using machine learning: A review. Med Eng Phys 2018;53:1-12.

23- 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.

24- Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev 2021;10(1):39.

25- Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile APP for systematic reviews. Syst Rev 2016;5:210

26- Moons K, de Groot J, Bouwmeester W, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Med 2014;11(10):e1001744.

27-Collins G, Reitsma J, Altman D, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med 2015;162(1):55.

Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analysis flow diagram of the identification, screening, and eligibility of included articles.

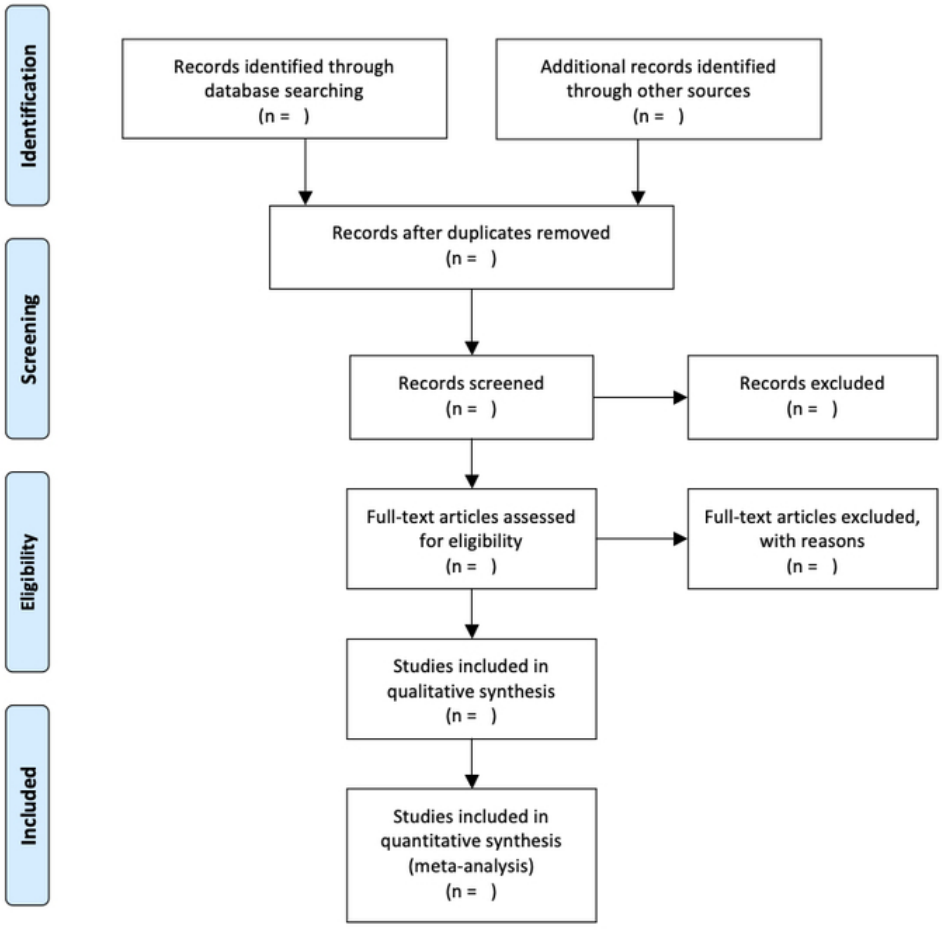


Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analysis flow diagram of the identification, screening, and eligibility of included articles.

60x57mm (300 x 300 DPI)

Supplementary Appendix 1. The search terms across databases

Database	Search terms
MEDLINE (PubMed)	("Diabetes"[tiab]) OR ("Diabetes Mellitus/analysis"[Mesh] OR "Diabetes Mellitus/classification"[Mesh] OR "Diabetes Mellitus/rehabilitation"[Mesh] OR "Diabetes Mellitus/therapy"[Mesh] OR "Diabetes Mellitus, Type 2"[Mesh]) OR ("Diabetic"[tiab]) OR ("Diabetic Neuropathies/classification"[Mesh] OR "Diabetic Neuropathies/diagnosis"[Mesh] OR "Diabetic Neuropathies/diagnostic imaging"[Mesh] OR "Diabetic Neuropathies/physiopathology"[Mesh] OR "Diabetic Neuropathies/rehabilitation"[Mesh] OR "Diabetic Neuropathies/statistics AND numerical data"[Mesh] OR "Diabetes Complications"[Mesh]) OR "Type 2 Diabetes" [tw] OR "Diabetes, Type 2" [tw] AND ("Gait"[tiab]) OR ("Gait Analysis"[tiab]) OR ("Gait/classification"[Mesh] OR "Gait/instrumentation"[Mesh] OR "Gait/methods"[Mesh] OR "Gait/organization and administration"[Mesh] OR "Gait/physiology"[Mesh] OR "Gait/standards"[Mesh] OR "Gait/statistics and numerical data"[Mesh] OR "Gait/trends"[Mesh]) OR ("Gait Disorders, Neurologic"[Mesh]) OR ("Walking Speed"[Mesh] OR "Walking"[tiab]) OR ("Locomotion"[tiab]) OR "Locomotion"[tiab] OR "Gait Kinetic*"[tw] OR "Gait Kinematic*"[tw] OR "range of motion"[tw] AND "Artificial Intelligence"[Mesh] OR Machine Learning[MeSH] OR Deep learning[MeSH] OR "Neural Networks, Computer"[Mesh] OR data mining[MeSH] OR machine[tiab] AND (learn* OR model*) OR (statistical[tiab] OR "statistical-learning"[tiab]) AND (strateg*[tiab]) OR multilayer perceptron*[tiab] OR random forest*[tiab] OR bayes* network*[tiab] OR support vector machine*[tiab] OR nearest neighbor*[tiab] OR k nearest neighbor*[tiab] OR elastic net[tiab] OR naive bayes*[tiab] OR (classification[tiab] OR regression[tiab] OR estimation[tiab] OR decision[tiab]) AND tree[tiab] OR ridge[tiab] OR kernel[tiab] OR ensemble[tiab] OR bagging[tiab] OR bagged[tiab] OR boosting[tiab] OR boosted[tiab] OR fuzzy[tiab] OR ("Predictive Value of Tests"[Mesh] OR "Probability Learning"[Mesh] OR "Forecasting"[Mesh] OR "Computing Methodologies"[Mesh] OR "Cluster Analysis"[Mesh]) OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination[tiab] OR discriminative[tiab] OR discriminatory[tiab]) AND (accuracy[tiab] OR ability[tiab] OR performance[tiab] OR value[tiab] OR model[tiab] OR models[tiab] OR power[tiab] OR efficiency[tiab]) OR "Generalized linear models"[tw] NOT "review"[pt]
Web of Science	(((TS=(("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes

(Clarivate Analytics)	Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics and numerical data" OR "Diabetes Complications" OR "Type 2 Diabetes" OR "Diabetes, Type 2"))) AND TS=(("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion")) AND TS=("Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR (machine AND (learn* OR model*)) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency) OR "Generalized linear models")) NOT TS=(Review)
Embase (Elsevier)	('diabetes' OR 'diabetes mellitus/analysis' OR 'diabetes mellitus/classification' OR 'diabetes mellitus/rehabilitation' OR 'diabetes mellitus/therapy' OR 'diabetes mellitus, type 2' OR 'diabetic' OR (('diabetic neuropathies/classification':ti,ab,kw OR 'diabetic neuropathies/diagnosis':ti,ab,kw OR 'diabetic neuropathies/diagnostic imaging':ti,ab,kw OR 'diabetic neuropathies/physiopathology':ti,ab,kw OR 'diabetic neuropathies/rehabilitation':ti,ab,kw OR 'diabetic neuropathies/statistics':ti,ab,kw) AND 'numerical data':ti,ab,kw) OR 'diabetic complication':ti,ab,kw OR 'non insulin dependent diabetes mellitus':ti,ab,kw) AND ('gait':ti,ab,kw OR 'gait analysis':ti,ab,kw OR 'gait/classification':ti,ab,kw OR 'gait/instrumentation':ti,ab,kw OR 'gait/methods':ti,ab,kw OR 'gait/organization':ti,ab,kw) AND administration:ti,ab,kw OR 'gait/physiology':ti,ab,kw OR 'gait/standards':ti,ab,kw OR 'gait/statistics':ti,ab,kw) AND 'numerical data':ti,ab,kw OR 'gait/trends':ti,ab,kw OR 'gait disorders, neurologic':ti,ab,kw OR 'walking speed':ti,ab,kw OR 'walking':ti,ab,kw OR 'locomotion':ti,ab,kw OR 'gait kinetic*':ti,ab,kw OR 'gait kinematic*':ti,ab,kw OR 'range of motion':ti,ab,kw) AND ('artificial intelligence' OR 'machine learning' OR 'deep learning' OR 'neural networks, computer' OR 'data mining' OR 'machine') AND ('learn*' OR 'model*') OR 'statistical' OR 'statistical-learning') AND 'strateg*' OR 'multilayer perceptron*' OR 'random

	forest*' OR 'bayes* network*' OR 'support vector machine*' OR 'nearest neighbor*' OR 'k nearest neighbor*' OR 'elastic net' OR 'naive bayes*' OR 'classification' OR 'regression' OR 'estimation' OR 'decision') AND 'tree' OR 'ridge' OR 'kernel' OR 'ensemble' OR 'bagging' OR 'bagged' OR 'boosting' OR 'boosted' OR 'fuzzy' OR 'predictive value of tests' OR 'probability learning' OR 'forecasting' OR 'computing methodologies' OR 'cluster analysis' OR 'validat*' OR 'predict*' OR 'rule*' OR ('predict*' AND 'outcome*') OR 'risk*' OR 'model*' OR 'history' OR 'variable*' OR 'criteria' OR 'scor*' OR 'characteristic*' OR 'finding*' OR 'factor*') AND ('predict*' OR 'model*'OR 'decision*' OR 'identif*' OR 'prognos*') OR ('decision*' AND 'model*') OR 'clinical*' OR ('prognostic' AND 'history') OR 'variable*' OR 'criteria' OR 'scor*' OR 'charcteristic*'OR 'finding*' OR 'factor*' OR 'model*' OR 'discrimination' OR 'discriminative' OR 'discriminatory') AND ('accuracy' OR 'ability' OR 'performance' OR 'value' OR 'model' OR 'models' OR 'power' OR 'efficiency') OR 'generalized linear models') NOT review:ab,ti
IEEE Xplore Digital Library (IEEE)	("All Metadata":"Type 2 Diabetes" OR "All Metadata":"Diabetes" OR "All Metadata":"Diabetic Neuropathies") AND ("All Metadata":"Gait" OR "All Metadata":"Gait Analysis" OR "All Metadata":"Range of Motion" OR "All Metadata":"Walking" OR "All Metadata":"Locomotion") AND ("All Metadata":"Artificial Intelligence" OR "All Metadata":"Machine learning" OR "All Metadata":"Predictive Value of Tests")
Scopus (Elsevier)	ALL (("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications" OR "Type 2 Diabetes")) AND ALL (("Gait") OR ("Gait Analysis") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion") AND ALL ("Artificial Intelligence" OR machine AND learning OR deep AND learning OR "Neural Networks" OR data AND mining OR "predictive value off tests" OR "Classification" OR "Cluster Analysis" OR "support vector machine" OR "Random Forest" OR "Naive Bayes" OR "Generalized linear models" OR "nearest neighbor*" OR "k nearest neighbor*") AND NOT "Review"
CINAHL (EBSCOhost)	TI (("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications" OR "Type 2 Diabetes" OR "Diabetes, Type 2")) AND TI (("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and

	administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion") AND ("Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR machine AND (learn* OR model*) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency) OR "Generalized linear models") NOT review
Google Scholar (Google)	Diabetes AND Gait "Artificial Intelligence" OR machine OR AND OR learning OR deep OR AND OR learning OR "Neural Networks" OR data OR AND OR mining OR "Predictive Value of Tests" OR "Cluster Analysis" -review
The Cochrane Library (Wiley)	("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications") OR "Type 2 Diabetes" OR "Diabetes, Type 2" in Title Abstract Keyword AND ("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion" in Title Abstract Keyword AND "Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR machine AND (learn* OR model*) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome*

	OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency) OR "Generalized linear models" OR "Random Forest" in Title Abstract Keyword NOT "review"

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Section and topic	Item No	Checklist item	Reported on page #
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	N/A
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	3
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	15
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	N/A
Support:			
Sources	5a	Indicate sources of financial or other support for the review	15
Sponsor	5b	Provide name for the review funder and/or sponsor	15
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	15
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	6
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	6
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	7
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	8
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	9
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	9
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening,	9

		eligibility and inclusion in meta-analysis)	
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	10
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	10
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	7
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	13
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	14
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	14
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	14
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	14
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	13
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	13

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051981.R2
Article Type:	Protocol
Date Submitted by the Author:	17-Jan-2022
Complete List of Authors:	Moura da Silva, Patrícia; UFRN, Physical Therapy; Edmond and Lily Safra International Institute of Neuroscience of Natal Oliveira Bezerra, Ana Beatriz; Federal University of Rio Grande do Norte Araújo Farias, Luanna Barbara; Federal University of Rio Grande do Norte Ribeiro, Tatiana Souza; Federal University of Rio Grande do Norte; Federal University of Rio Grande do Norte Morya, Edgard; Edmond and Lily Safra International Institute of Neuroscience of Natal Cavalcanti, Fabrícia; Federal University of Rio Grande do Norte, Departament of Physical Therapy
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diabetes and endocrinology, Research methods, Health informatics, Neurology
Keywords:	DIABETES & ENDOCRINOLOGY, BIOTECHNOLOGY & BIOINFORMATICS, Neurophysiology < NEUROLOGY

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Patrícia Mayara Moura da Silva[1, 3, *], Ana Beatriz de Oliveira Bezerra[2], Luanna Barbara de Araújo Farias[2], Tatiana Souza Ribeiro[1,2], Edgard Morya[3], Fabrícia Azevêdo da Costa Cavalcanti[1,2]

[1]Department of Physical Therapy, Graduate Program, Federal University of Rio Grande do Norte, Natal, RN, Brazil.

[2]Department of Physical Therapy, Undergraduate Program, Federal University of Rio Grande do Norte, Natal, RN, Brazil.

[3]Edmond and Lily Safra International Institute of Neurosciences, Neuroengineering Program, Santos Dumont Institute, Macaíba, RN, Brazil.

***Corresponding author:** Patrícia Mayara Moura da Silva at Physical Therapy Department, Federal University of Rio Grande do Norte, Natal, RN, 59072-970, Brazil, patriciamayarams@gmail.com

Word Count: 2,009

ABSTRACT

INTRODUCTION: Type 2 diabetes can lead to gait abnormalities, including a longer stance phase, shorter steps, and improper foot pressure distribution. Quantitative data from objective methods for evaluating gait patterns are accurate and cost-effective. In addition, it can also help predictive methods to forecast complications and develop early strategies to guide treatments. To date, no research has systematically summarised the predictive methods used to assess type 2 diabetic gait. Therefore, this protocol aims to identify which predictive methods have been employed to assess the diabetic gait.

METHODS AND ANALYSIS: This protocol will follow the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) statement. Electronic searches of articles from inception to January 2022 will be performed, from May 2021 to 31 January 2022, in the Web of Science, MEDLINE, Embase, IEEE Xplore Digital Library, Scopus, CINAHL, Google Scholar, APA PsycInfo, the Cochrane Library, and in references of key articles and grey literature without language restrictions. We will include studies that examined the development and/or validation of predictive methods to assess type 2 diabetic gait in adults aged > 18 years without amputations, use of assistive devices, ulcers, or neuropathic pain. Two independent reviewers will screen the included studies and extract the data using a customised charting form. A third reviewer will resolve any disagreements. A narrative synthesis will be performed for the included studies. Risk of bias and quality of evidence will be assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD).

ETHICS AND DISSEMINATION: Ethical approval is not required because only available secondary published data will be analysed. The findings will be disseminated through peer-reviewed journals and/or presentations at relevant conferences and other media platforms.

TRIAL REGISTRATION NUMBER: PROSPERO (CDR 42020199495).

ARTICLE SUMMARY

Strengths and limitations of this study:

- This study will be the first systematic review to comprehensively analyse the existing predictive methods for gait analysis in patients with type 2 diabetes.
- This systematic review will focus on the predictive method's performance (surrogate outcomes) rather than on patient-reported outcome measures.
- A broad search strategy and robust quality assessment criteria (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) will be used to appraise and examine the existing literature.
- Two independent reviewers will be responsible for conducting the study selection, data extraction, and quality assessment.
- A limitation could be the potential lack of studies that meet the established inclusion criteria.

INTRODUCTION

Diabetes mellitus is a worldwide health concern, with a prevalence of 8.8% in 2017[1]. With type 2 diabetes mellitus being the most common[2], this condition is related to a dysfunction either in the pancreatic β-cells' ability to secrete insulin, insulin resistance in target organs, or a combination of both, resulting in hyperglycaemia[1,3]. Patients can also present with blood vessel degeneration[4,5] that can evolve into neuropathy and damage sensory and motor nerve fibres[4,5]. Diabetes also alters physical function and mobility[6]. Both can lead to motor abnormalities such as longer stance time (i.e., greater support base) and shorter steps, which may exhibit as slower gait speeds and increased cadence[7,8]. In addition, changes in the sensibility of the plantar surface of the foot can worsen plantar pressure distribution, balance, and gait[5,7].

Boosting insight into diabetic gait pattern alterations can be important for preventing complications caused by diabetes and developing strategies to guide treatments[6,9]. In clinical practice, while there is a high prevalence of observational methods[10,11], this may be unreliable in assessing and diagnosing gait patterns. Observational methods are subjective and can generate inaccuracies during the assessment and diagnosis of the patient's movements due to different interpretations between examiners. In addition, these differences can impair decision making to address a specific treatment[12-14]. On the contrary, objective methods are reliable, accurate, quicker, and cost-effective owing to quantitative metric results[8].

Objective gait analysis methods require data collected from patients wearing sensors or performing gait in specific devices, such as Inertial Motion Units (IMUs), electromyography (EMG), optoelectronic systems, or force platforms[10]. Data from quantitative gait measures can be analysed using various methods. One of these is the use of predictive analytics that combine the collected data and estimate probabilities that can assist clinicians and potentially influence their decision to manage treatments to restore gait[15-18]. Predictive methods are mathematical equations (from statistics or machine learning approaches) that can combine information from a set of data, resulting in a response forecasting the probability of a particular outcome[19,20].

Emerging predictive methods include machine learning (ML) algorithms. ML can be used for automatic gait recognition to predict possible complications such as the risk of falls and pressure ulcers[21]. Newer methods have opened new perspectives for the early diagnosis of gait disorders. This is essential in preventing potential future complications and to draw on personalised gait training[15] by quantifying the treatment progress and follow-ups[22].

To our knowledge, no research has systematically summarised predictive algorithms used to assess gait in patients with type 2 diabetes. Based on this, we raise an important question about the existence of predictive methods used to evaluate the gait of patients with type 2 diabetes. Therefore, the purpose of this study is to conduct a systematic review of the literature to summarise the evidence regarding existing predictive methods used in the gait patterns of patients with diabetes. In addition, we intend to describe the

characteristics of the studies identified among the variety of gait data collected regarding which input features are the most commonly used to implement a predictive method.

METHODS AND ANALYSES

Study design

This systematic review protocol was prepared using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P)[23].

Study registration

This protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO) no. CRD42020199495. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020199495

Eligibility criteria

Types of study

Articles will be eligible for review when they describe the development and/or validation of a predictive method to assess gait in human type 2 diabetes. Furthermore, all published and unpublished studies (e.g., dissertations and theses), conference proceedings that deal with diabetic gait analysis, independent of the parameters measured, will be included if developed and/or validated as a predictive method. There will be no geographical or language restrictions. Wherever necessary, relevant articles will be arranged for translation.

Participants

We will include clinical data from adult participants (> 18 years old) who had type 2 diabetes diagnosed at any disease stage without lower limb amputations or the use of gait assistive devices. In addition, data with participants with ulcers or neuropathic pain (that could have interfered in the gait execution) will be excluded. There will be no restrictions on sex or race.

Outcome measures

The primary outcome will comprise all predictive methods (e.g., machine learning models) applied to analyse gait in patients with type 2 diabetes. The secondary outcome will include gait data input features (e.g., spatiotemporal, angular gait parameters, EMG data, force data, and plantar pressure data) most commonly used to implement a predictive model.

Search strategy for identification of relevant studies

The search strategy will be guided by the PRISMA extension for searching (PRISMA-S)[24]. The following electronic databases will be searched: Web of Science (Clarivate Analytics), MEDLINE (PubMed), Embase (Elsevier), IEEE Xplore Digital Library (IEEE), Scopus (Elsevier), CINAHL (EBSCOhost), Google Scholar (Google), APA PsycInfo (APA PsycNet), and the Cochrane Library (Wiley) from May 2021 to 31 January 2022. The time range of the published studies was from inception to January 2022. We will manually search the reference list of the studies included in the review. Grey literature involving published and unpublished studies (e.g., dissertations and theses) and conference

proceedings will also be searched without language restrictions, but this must be limited to human participants.

The articles will be searched using a combination of free keywords and the terminology registered in the Medical Subject Headings (MeSH) of the U.S. National Library of Medicine. The terms that will be used are related to diabetes (e.g., “Type 2 Diabetes”, “Diabetes, Type 2”, “Diabetes Mellitus, Type 2”), gait (e.g., “Gait”, “Gait Analysis”, “Kinematic”, “Kinetic”, “Range of Motion”), and prediction-related (e.g., “Artificial Intelligence”, “Machine Learning”, "Statistical-learning", “Predictive Value of Tests”, “Support Vector Machine”, “Neural Networks, Computer”). The search strategy was pilot tested and finalised in MEDLINE (PubMed) before being translated for use in other databases. Details of the search strategies are provided in Online Supplementary Appendix 1.

Screening of the studies

Based on the previously described inclusion criteria, two independent reviewers (PMMS, ABOB) will screen titles and abstracts identified during electronic and manual searches to determine its eligibility. Study record information, including title and abstract from the searched online database, will be imported into the Rayyan systematic review software[25]. This platform will guide the reviewers in conducting the literature review process through its ability to explore and filter searched studies. Duplicate studies will be removed. If the title or abstract does not provide enough information for inclusion, the full text will be obtained for a full review. The same two reviewers (PMMS, ABOB) will independently

screen the full-text articles to identify studies for inclusion and record the reasons for exclusion for ineligible studies. Any disagreements that arise will be resolved initially by a discussion between the two reviewers, or, if necessary, with assistance from a third reviewer (FACC).

All reasons for the exclusion of ineligible studies will be recorded. The results of the screening process will be provided in detail using the PRISMA information flowchart (Figure 1).

Data extraction

A data extraction form was developed through a discussion among all authors and adapted from the critical appraisal and data extraction for systematic reviews of prediction modelling studies (CHARMS) checklist[26]. The included studies will go forward to the data extraction and quality assessment stages of the review. Two independent reviewers (PMMS, ABOB) will extract the outcome data from the included studies. If necessary, disagreements in data extraction will be discussed between the two reviewers and judged by a third reviewer (FACC).

The data collection form will aim to extract the key features of the review. Hence, we will divide the items within the data collection form into four blocks: (1) study information including publication year, author information, funding or sponsorship information, type of study, journal name, control, population, intervention, and outcome (PICO elements); (2) database information including name, sample size, host organisation, and sponsorship;

(3) patient demographic information including sex, age, race, and disease severity; and (4) predictive methodological information including the type of gait assessment, comparisons with gold standard devices, type of predictive algorithm used (including its statistical or machine learning model name), format of input feature, optimisation algorithm, objective function, feature extraction methods, type of extraction feature and computational efficiency, and cost. Table 1 presents an example of the data extraction form. These data will be presented in the 'Characteristics of included studies' table.

Missing data may include missing outcomes, missing summary data, or missing individual results. The authors will consider the reasons for the missing data. Where possible, we will contact the original investigators to obtain any missing data. However, in the case of contact difficulty, we will present the findings according to the statistical information available in each review, and this will be clearly stated in the final overview.

Table 1: Example of the data extraction form for all included studies.

Study information	
Study year	Year of the study publication
Author information	Last name of the author, whether clinical practitioners participated in the study
Type of study	Source of data (e.g., cohort, case-control, randomised trial participants or registry data)
Journal name	Journal name
PICO* elements	PICO* elements in summary
Database information	
Database name	Name of the database used for modelling
Host organisation	Name of the hosting organisation of the database
Sponsorship	The funding or sponsorship information
Sample size	Sample size used for building the model

Source or data	From which source the database was used (e.g., electronic health records, clinical registry, administrative data, cohort study, clinical trial)
Patient demographic information	
Sex	Sex of adults (male, female, alternative gender)
Age	Age and/or year of birth
Country	Country or countries in which study was based
Diabetes severity	Disease severity
Predictive methodological information	
Predictors	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation)
Number of features	Number of features for building the model
Selected features	The study reported the importance of selected features?
Type of extracted feature	Which features the algorithm uses (e.g., pressure, gait velocity, cadence, step width, pixel feature, action unit, etc.)
Tool used for gait assessment	Quantitative tool used to assess gait kinetic or kinematic (e.g., IMU, force platform, optoelectronic, EMG)
Used highly rated standard devices	Quantitative tool used to assess gait kinetic or kinematic was a device considered gold standards (e.g., force platform, optoelectronic)
Predictive method used	Type of predictive method used to assess gait (e.g., which machine learning techniques was used)
Model name	The name of the predictive model used. The underlying mathematical model used (e.g., linear regression, support vector machine)
Missing data	Number of participants with missing data for each predictor and the process handled with missing data (e.g., complete-case analysis, imputation, or other methods)
Format of input feature (predictor or variables)	Which input gait data was used (e.g., plantar pressure, frame, sequence or image)
Model performance/ validation	Performance metrics and scores of how accurate the model used is predicting (e.g., accuracy, average errors, R-squared, confusion matrix, etc.)
Model evaluation	Method used for testing model performance: development dataset only (random split of data, resampling methods, e.g., bootstrap or cross-validation) or separate external validation (e.g., temporal, geographical, different setting, different investigators)
Computational efficiency and cost	Computational efficiency (speed, cloud space, etc.) and cost related to the algorithm (e.g., require GPU resources, large cluster, etc.)

*PICO: population, intervention (exposure), control, outcome; IMU: inertial measurement unit; EMG: electromyography; GPU: graphics processing unit.

Quality of evidence

The quality of the predictive model used on the eligible studies will be assessed based on Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) checklist[27]. The TRIPOD Statement is a checklist of 22 items for the appropriate reporting of studies developing or validating multivariable prediction models[19]. Each item will be scored as 0, 1, and 2, ranked as 'no report', 'inadequate report', and 'adequate report', respectively.

Risk of bias

The pre-selected articles will be evaluated and scored for methodological quality using the Prediction Model Risk of Bias Assessment Tool (PROBAST)[20] by two independent reviewers (PMMS, ABOB). In cases of opinion divergence, a third reviewer (FACC) will decide the score. The questionnaire consists of 20 items with four domains (participants, predictors, outcome, and analysis). Based on the questionnaire ratings, the risk of bias for each domain will be ranked as 'low risk', 'high risk', or 'too unclear for judgment'. PROBAST will be used to categorise the included studies regarding their methodological quality, but these studies will not be excluded based solely on this evaluation.

Strategy for data synthesis

A narrative synthesis will be conducted with the information presented in the text and tables to summarise and explain the characteristics and findings of the included studies.

Data will be summarised using descriptive statistics and visual plots. Categorical data about the reporting, methodological conduct, and risks of bias will be described using numbers and percentages. The distribution of continuous data, such as sample sizes and the number of features, will be described using measures of central tendency such as mean and standard deviation for normally distributed data and median and percentiles (25th and 75th) for non-normally distributed data.

The risk of bias assessment will be summarised and graphically presented for each PROBAST domain and the overall risk of bias judgment. The results will be stratified by prevalent predictive techniques and study design (development with internal validation and/or external validation). The quality of evidence based on TRIPOD will also be summarised and graphically presented for each included study and its respective score rank.

Analyses of subgroups or subsets

We plan to conduct subgroup analyses using predictive model types (e.g., regression models vs. classification models, neural networks vs. traditional machine learning models) and gait input parameters (e.g., kinematic vs. kinetic data features, IMUs vs. EMG data features). In addition, we plan to classify participants according to their anthropometric subgroup (e.g., age, body index mass, height, weight, gait measurements, and diabetes vitals). More exploratory subgroup analyses will be decided during the data extraction and analysis process.

ETHICS AND DISSEMINATION

To the best of our knowledge, this systematic review is the first that will synthesize existing evidence regarding the types of predictive methods used to assess gait in patients with type 2 diabetes. Predictive methods are increasingly being appraised and recommended for formal risk assessment in treatment decision making and clinical guidelines. The proposed systematic review may inform future research and clinicians. For instance, it may help researchers in designing customisable prediction tools to be used in diabetic care, and thus allow physiotherapists to better conduct rehabilitative gait treatments in the patients with type 2 diabetes.

Because we will be using secondary data sources, ethical approval is not required for this systematic review study. Our findings will be disseminated through peer-reviewed publications, presentations at conferences, and clinical and patient networks.

Acknowledgements: The authors would like to thank CAPES Foundation, an agency under the Ministry of Education (MEC) of Brazil for their assistance during the research.

Authors' Contributions: PMMS, ABOB, LBAF, TSR, EM, and FACC wrote the main manuscript text. PMMS prepared all figures and tables. All authors reviewed the manuscript. All authors read and approved the final manuscript.

Funding statement: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Patient and Public Involvement: Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

References

1- Standl E, Khunti K, Hansen T, et al. The global epidemics of diabetes in the 21st century: Current situation and perspectives. Eur J Prev Cardiol 2019;26(2_suppl):7-14.

2- Skyler JS, Bakris GL, Bonifacio E, et al. Differentiation of diabetes by pathophysiology, natural history, and prognosis. Diabetes 2017;*66*(2):241-255.

3- Chatterjee S, Khunti K, Davies M. Type 2 diabetes. Lancet 2017;389(10085):2239-2251.

4- Sawacha Z, Cristoferi G, Guarneri G, et al. Characterizing multisegment foot kinematics during gait in diabetic foot patients. J Neuroeng Rehabil 2009;6(1):37.

5- Huang C, Shivaswamy V, Thaisetthawatkul P, et al. An altered spatiotemporal gait adjustment during a virtual obstacle crossing task in patients with diabetic peripheral neuropathy. J Diabetes Complications 2019;33(2):182-188.

6- Ko S, Stenholm S, Chia C, et al. Gait pattern alterations in older adults associated with type 2 diabetes in the absence of peripheral neuropathy—Results from the Baltimore Longitudinal Study of Aging. Gait Posture 2011;34(4):548-552.

7- Crews R, Yalla S, Fleischer A, et al. A Growing Troubling Triad: Diabetes, Aging, and Falls. J Aging Res 2013.

8- Kirkwood R, Borém I, Sampaio R, et al. Frailty Status and Gait Parameters of Older Women With Type 2 Diabetes. Can J Diabetes 2019;43(2):121-127.

9- Wrobel J, Najafi B. Diabetic Foot Biomechanics and Gait Dysfunction. J Diabetes Sci Technol 2010;4(4):833-845.

10- Prakash C, Kumar R, Mittal N. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. Artif Intell Rev 2018;49(1):1-40.

11- Mundt M, Thomsen W, Witter T, et al. Prediction of lower limb joint angles and moments during gait using artificial neural networks. Med Biol Eng Comput 2019;58(1):211-225.

12- Muro-de-la-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A. Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications. Sens 2014;14(2):3362-3394.

13- Gor-García-Fogeda M, Cano de la Cuerda R, Carratalá Tejada M, et al. Observational Gait Assessments in People With Neurological Disorders: A Systematic Review. Arch Phys Med Rehabil 2016;97(1):131-140.

14- Roberts M, Mongeon D, Prince F. Biomechanical parameters for gait analysis: a systematic review of healthy human gait. Phys Ther Rehabil 2017;4(1):6.

15- Begg R, Kamruzzaman J. A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data. J Biomech 2005;38(3):401-408.

16- Mannini A, Trojaniello D, Cereatti A, et al. A Machine Learning Framework for Gait Classification Using Inertial Sensors: Application to Elderly, Post-Stroke and Huntington's Disease Patients. Sens 2016;16(1):134.

17- König N, Singh N, Baumann C, et al. Can Gait Signatures Provide Quantitative Measures for Aiding Clinical Decision-Making? A Systematic Meta-Analysis of Gait

Variability Behavior in Patients with Parkinson's Disease. Front Hum Neurosci 2016;10.

18- Caldas R, Mundt M, Potthast W, et al. A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. Gait Posture 2017;57:204-210.

19- Moons K, Altman D, Reitsma J, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 2015;162(1):W1.

20- Wolff R, Moons K, Riley R, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med 2019;170(1):51.

21- Khera P, Kumar N. Role of machine learning in gait analysis: a review. J Med Eng Technol, Taylor & Francis 2020;44(8):441-467.

22- Figueiredo J, Santos C, Moreno J. Automatic recognition of gait patterns in human motor disorders using machine learning: A review. Med Eng Phys 2018;53:1-12.

23- 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.

24- Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev 2021;10(1):39.

25- Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile APP for systematic reviews. Syst Rev 2016;5:210

26- Moons K, de Groot J, Bouwmeester W, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Med 2014;11(10):e1001744.

27-Collins G, Reitsma J, Altman D, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med 2015;162(1):55.

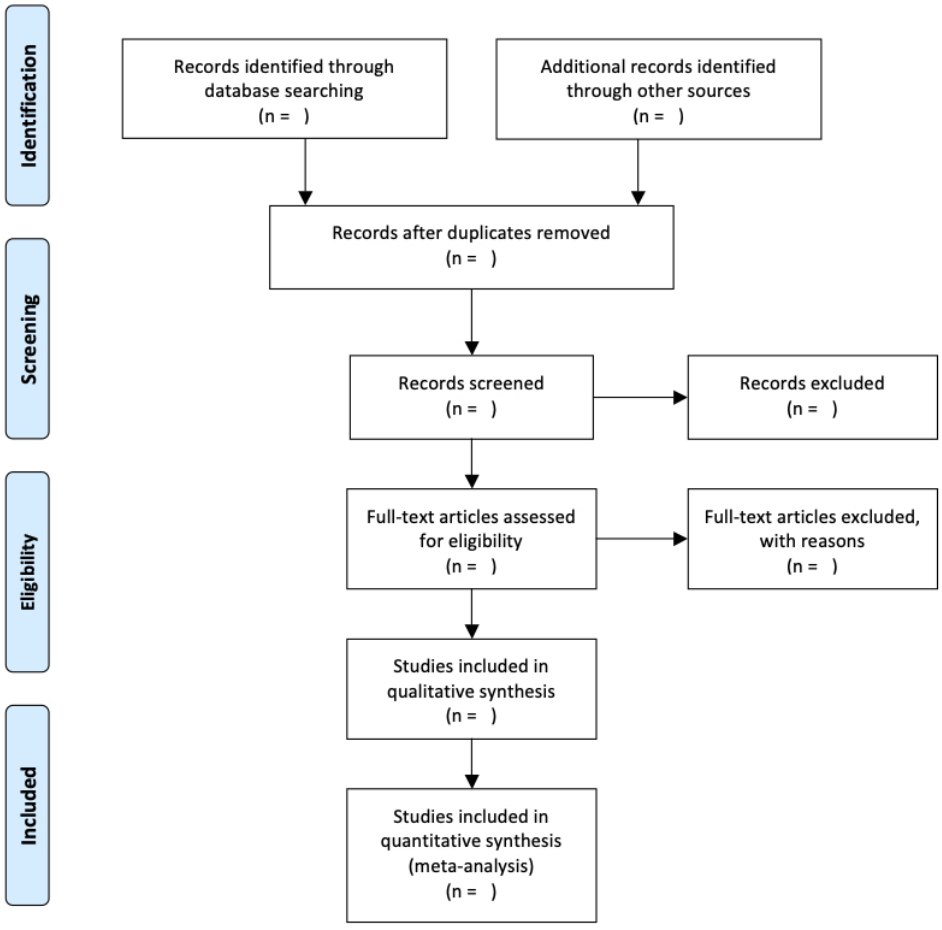


Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analysis flow diagram of the identification, screening, and eligibility of included articles.

60x57mm (350 x 350 DPI)

Supplementary Appendix 1. The search terms across databases

Database	Search terms
MEDLINE (PubMed)	("Diabetes"[tiab]) OR ("Diabetes Mellitus/analysis"[Mesh] OR "Diabetes Mellitus/classification"[Mesh] OR "Diabetes Mellitus/rehabilitation"[Mesh] OR "Diabetes Mellitus/therapy"[Mesh] OR "Diabetes Mellitus, Type 2"[Mesh]) OR ("Diabetic"[tiab]) OR ("Diabetic Neuropathies/classification"[Mesh] OR "Diabetic Neuropathies/diagnosis"[Mesh] OR "Diabetic Neuropathies/diagnostic imaging"[Mesh] OR "Diabetic Neuropathies/physiopathology"[Mesh] OR "Diabetic Neuropathies/rehabilitation"[Mesh] OR "Diabetic Neuropathies/statistics AND numerical data"[Mesh] OR "Diabetes Complications"[Mesh]) OR "Type 2 Diabetes" [tw] OR "Diabetes, Type 2" [tw] AND ("Gait"[tiab]) OR ("Gait Analysis"[tiab]) OR ("Gait/classification"[Mesh] OR "Gait/instrumentation"[Mesh] OR "Gait/methods"[Mesh] OR "Gait/organization and administration"[Mesh] OR "Gait/physiology"[Mesh] OR "Gait/standards"[Mesh] OR "Gait/statistics and numerical data"[Mesh] OR "Gait/trends"[Mesh]) OR ("Gait Disorders, Neurologic"[Mesh]) OR ("Walking Speed"[Mesh] OR "Walking"[tiab]) OR ("Locomotion"[tiab]) OR "Locomotion"[tiab] OR "Gait Kinetic*"[tw] OR "Gait Kinematic*"[tw] OR "range of motion"[tw] AND "Artificial Intelligence"[Mesh] OR Machine Learning[MeSH] OR Deep learning[MeSH] OR "Neural Networks, Computer"[Mesh] OR data mining[MeSH] OR machine[tiab] AND (learn* OR model*) OR (statistical[tiab] OR "statistical-learning"[tiab]) AND (strateg*[tiab]) OR multilayer perceptron*[tiab] OR random forest*[tiab] OR bayes* network*[tiab] OR support vector machine*[tiab] OR nearest neighbor*[tiab] OR k nearest neighbor*[tiab] OR elastic net[tiab] OR naive bayes*[tiab] OR (classification[tiab] OR regression[tiab] OR estimation[tiab] OR decision[tiab]) AND tree[tiab] OR ridge[tiab] OR kernel[tiab] OR ensemble[tiab] OR bagging[tiab] OR bagged[tiab] OR boosting[tiab] OR boosted[tiab] OR fuzzy[tiab] OR ("Predictive Value of Tests"[Mesh] OR "Probability Learning"[Mesh] OR "Forecasting"[Mesh] OR "Computing Methodologies"[Mesh] OR "Cluster Analysis"[Mesh]) OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination[tiab] OR discriminative[tiab] OR discriminatory[tiab]) AND (accuracy[tiab] OR ability[tiab] OR performance[tiab] OR value[tiab] OR model[tiab] OR models[tiab] OR power[tiab] OR efficiency[tiab]) OR "Generalized linear models"[tw] NOT "review"[pt]
CINAHL (EBSCOhost)	TI (("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis"

	OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications" OR "Type 2 Diabetes" OR "Diabetes, Type 2")) AND TI (("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion") AND ("Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR machine AND (learn* OR model*) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency) OR "Generalized linear models") NOT review
The Cochrane Library (Wiley)	("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications") OR "Type 2 Diabetes" OR "Diabetes, Type 2" in Title Abstract Keyword AND ("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion" in Title Abstract Keyword AND "Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR machine AND (learn* OR model*) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability

	Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency) OR "Generalized linear models" OR "Random Forest" in Title Abstract Keyword NOT "review"
Embase (Elsevier)	('diabetes' OR 'diabetes mellitus/analysis' OR 'diabetes mellitus/classification' OR 'diabetes mellitus/rehabilitation' OR 'diabetes mellitus/therapy' OR 'diabetes mellitus, type 2' OR 'diabetic' OR (('diabetic neuropathies/classification':ti,ab,kw OR 'diabetic neuropathies/diagnosis':ti,ab,kw OR 'diabetic neuropathies/diagnostic imaging':ti,ab,kw OR 'diabetic neuropathies/physiopathology':ti,ab,kw OR 'diabetic neuropathies/rehabilitation':ti,ab,kw OR 'diabetic neuropathies/statistics':ti,ab,kw) AND 'numerical data':ti,ab,kw) OR 'diabetic complication':ti,ab,kw OR 'non insulin dependent diabetes mellitus':ti,ab,kw) AND ('gait':ti,ab,kw OR 'gait analysis':ti,ab,kw OR 'gait/classification':ti,ab,kw OR 'gait/instrumentation':ti,ab,kw OR 'gait/methods':ti,ab,kw OR 'gait/organization':ti,ab,kw AND administration:ti,ab,kw OR 'gait/physiology':ti,ab,kw OR 'gait/standards':ti,ab,kw OR 'gait/statistics':ti,ab,kw) AND ('numerical data':ti,ab,kw OR 'gait/trends':ti,ab,kw OR 'gait disorders, neurologic':ti,ab,kw OR 'walking speed':ti,ab,kw OR 'walking':ti,ab,kw OR 'locomotion':ti,ab,kw OR 'gait kinetic*':ti,ab,kw OR 'gait kinematic*':ti,ab,kw OR 'range of motion':ti,ab,kw) AND ('artificial intelligence' OR 'machine learning' OR 'deep learning' OR 'neural networks, computer' OR 'data mining' OR 'machine') AND ('learn*' OR 'model*') OR 'statistical' OR 'statistical-learning' AND 'strateg*' OR 'multilayer perceptron*' OR 'random forest*' OR 'bayes* network*' OR 'support vector machine*' OR 'nearest neighbor*' OR 'k nearest neighbor*' OR 'elastic net' OR 'naive bayes*' OR 'classification' OR 'regression' OR 'estimation' OR 'decision' AND 'tree' OR 'ridge' OR 'kernel' OR 'ensemble' OR 'bagging' OR 'bagged' OR 'boosting' OR 'boosted' OR 'fuzzy' OR 'predictive value of tests' OR 'probability learning' OR 'forecasting' OR 'computing methodologies' OR 'cluster analysis' OR 'validat*' OR 'predict*' OR 'rule*' OR ('predict*' AND 'outcome*') OR 'risk*' OR 'model*' OR 'history' OR 'variable*' OR 'criteria' OR 'scor*' OR 'characteristic*' OR 'finding*' OR 'factor*') AND ('predict*' OR 'model*'OR 'decision*' OR 'identif*' OR 'prognos*') OR ('decision*' AND 'model*') OR 'clinical*' OR ('prognostic' AND 'history') OR 'variable*' OR 'criteria' OR 'scor*' OR 'charcteristic*'OR 'finding*' OR 'factor*' OR 'model*' OR 'discrimination' OR 'discriminative' OR 'discriminatory') AND ('accuracy' OR 'ability' OR 'performance' OR 'value' OR 'model' OR 'models' OR 'power' OR 'efficiency') OR 'generalized linear models') NOT review:ab,ti
APA PsycInfo (APA PsycNet)	(Any Field: "Diabetes" [tiab]) *OR* (Any Field: "Diabetes Mellitus/analysis" [Mesh] *OR* Any Field: "Diabetes Mellitus/classification" [Mesh] *OR* Any Field: "Diabetes

	Mellitus/rehabilitation" [Mesh] *OR* Any Field: "Diabetes Mellitus/therapy" [Mesh] *OR* Any Field: "Diabetes Mellitus, Type 2" [Mesh]) *OR* (Any Field: "Diabetic" [tiab]) *OR* (Any Field: "Diabetic Neuropathies/classification" [Mesh] *OR* Any Field: "Diabetic Neuropathies/diagnosis" [Mesh] *OR* Any Field: "Diabetic Neuropathies/diagnostic imaging" [Mesh] *OR* Any Field: "Diabetic Neuropathies/physiopathology" [Mesh] *OR* Any Field: "Diabetic Neuropathies/rehabilitation" [Mesh] *OR* Any Field: "Diabetic Neuropathies/statistics AND numerical data" [Mesh] *OR* Any Field: "Diabetes Complications" [Mesh]) *OR* Any Field: "Type 2 Diabetes" [tw] *OR* Any Field: "Diabetes, Type 2" [tw]*AND* (Any Field: "Gait" [tiab]) *OR* (Any Field: "Gait Analysis" [tiab]) *OR* (Any Field: "Gait/classification" [Mesh] *OR* Any Field: "Gait/instrumentation" [Mesh] *OR*Any Field: "Gait/methods" [Mesh] *OR* Any Field: "Gait/organization and administration" [Mesh] *OR* Any Field: "Gait/physiology" [Mesh] *OR* Any Field: "Gait/standards" [Mesh] *OR* Any Field: "Gait/statistics and numerical data" [Mesh] *OR* Any Field: "Gait/trends" [Mesh]) *OR* (Any Field: "Gait Disorders, Neurologic" [Mesh]) *OR* (Any Field: "Walking Speed" [Mesh] *OR* Any Field: "Walking" [tiab]) *OR* (Any Field: "Locomotion" [tiab]) *OR* Any Field: "Locomotion" [tiab]*OR* Any Field: "Gait Kinetic*" [tw] *OR* Any Field: "Gait Kinematic*" [tw] *OR* Any Field: "range of motion" [tw]*AND* Any Field: "Artificial Intelligence" [Mesh] *OR* Any Field: Machine Learning[MeSH] *OR* Any Field: Deep learning[MeSH] *OR* Any Field: "Neural Networks, Computer" [Mesh] *OR* Any Field: data mining[MeSH]*OR* Any Field: machine[tiab] *AND* (Any Field: learn**OR* Any Field: model*) *OR* (Any Field: statistical[tiab]*OR* Any Field: "statistical-learning" [tiab]) *AND* (Any Field: strateg*[tiab]) *OR* Any Field: multilayer perceptron*[tiab] *OR* Any Field: random forest*[tiab]*OR* Any Field: bayes* network*[tiab] *OR* Any Field: support vector machine*[tiab] *OR* Any Field: nearest neighbor*[tiab] *OR* Any Field: k nearest neighbor*[tiab]*OR* Any Field: elastic net[tiab] *OR* Any Field: naive bayes*[tiab] *OR* (Any Field: classification[tiab] *OR* Any Field: regression[tiab] *OR* Any Field: estimation[tiab]*OR* Any Field: decision[tiab]) *AND* Any Field: tree[tiab]*OR* Any Field: ridge[tiab] *OR* Any Field: kernel[tiab] *OR*Any Field: ensemble[tiab] *OR* Any Field: bagging[tiab]*OR* Any Field: bagged[tiab] *OR* Any Field: boosting[tiab] *OR* Any Field: boosted[tiab] *OR* Any Field: fuzzy[tiab] *OR* (Any Field: "Predictive Value of Tests" [Mesh] *OR* Any Field: "Probability Learning" [Mesh] *OR* Any Field: "Forecasting" [Mesh] *OR* Any Field: "Computing Methodologies" [Mesh] *OR* Any Field: "Cluster Analysis" [Mesh]) *OR* (Any Field: Validat* *OR* Any Field: Predict* *OR* Any Field: Rule*) *OR* (Any Field: Predict* *AND* Any Field: Outcome* *OR*Any Field: Risk* *OR* Any Field: Model*) *OR* (Any Field: History *OR* Any Field: Variable* *OR* Any Field: Criteria *OR* Any Field: Scor* *OR* Any Field: Characteristic* *OR* Any Field: Finding* *OR* Any Field: Factor*) *AND* (Any Field: Predict* *OR* Any Field: Model*

	OR Any Field: Decision* *OR* Any Field: Identif**OR* Any Field: Prognos*) *OR* (Any Field: Decision**AND* Any Field: Model* *OR* Any Field: Clinical*) *OR*(Any Field: Prognostic *AND* Any Field: History *OR* Any Field: Variable* *OR* Any Field: Criteria *OR* Any Field: Scor* *OR* Any Field: Charcteristic* *OR* Any Field: Finding* *OR* Any Field: Factor* *OR* Any Field: Model*) *OR* (Any Field: discrimination[tiab] *OR* Any Field: discriminative[tiab] *OR* Any Field: discriminatory[tiab]) *AND* (Any Field: accuracy[tiab] *OR* Any Field: ability[tiab] *OR* Any Field: performance[tiab] *OR* Any Field: value[tiab] *OR* Any Field: model[tiab] *OR* Any Field: models[tiab] *OR* Any Field: power[tiab] *OR* Any Field: efficiency[tiab]) *OR* Any Field: "Generalized linear models" [tw] NOT "review" [pt]
Google Scholar	Diabetes AND Gait "Artificial Intelligence" OR machine OR AND OR learning OR deep OR AND OR learning OR "Neural Networks" OR data OR AND OR mining OR "Predictive Value of Tests" OR "Cluster Analysis" -review
IEEE Xplore Digital Library (IEEE)	("All Metadata":"Type 2 Diabetes" OR "All Metadata":"Diabetes" OR "All Metadata":"Diabetic Neuropathies") AND ("All Metadata":"Gait" OR "All Metadata":"Gait Analysis" OR "All Metadata":"Range of Motion" OR "All Metadata":"Walking" OR "All Metadata":"Locomotion") AND ("All Metadata":"Artificial Intelligence" OR "All Metadata":"Machine learning" OR "All Metadata":"Predictive Value of Tests")
Scopus (Elsevier)	ALL (("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics AND numerical data" OR "Diabetes Complications" OR "Type 2 Diabetes")) AND ALL (("Gait") OR ("Gait Analysis") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion") AND ALL ("Artificial Intelligence" OR machine AND learning OR deep AND learning OR "Neural Networks" OR data AND mining OR "predictive value off tests" OR "Classification" OR "Cluster Analysis" OR "support vector machine" OR "Random Forest" OR "Naive Bayes" OR "Generalized linear models" OR "nearest neighbor*" OR "k nearest neighbor*") AND NOT "Review"
Web of Science (Clarivate Analytics)	(((TS=(("Diabetes") OR ("Diabetes Mellitus/analysis" OR "Diabetes Mellitus/classification" OR "Diabetes Mellitus/rehabilitation" OR "Diabetes Mellitus/therapy" OR "Diabetes Mellitus, Type 2") OR ("Diabetic") OR ("Diabetic Neuropathies/classification" OR "Diabetic Neuropathies/diagnosis" OR "Diabetic Neuropathies/diagnostic imaging" OR "Diabetic Neuropathies/physiopathology" OR "Diabetic Neuropathies/rehabilitation" OR "Diabetic Neuropathies/statistics and numerical data" OR "Diabetes Complications" OR "Type 2 Diabetes" OR "Diabetes, Type 2"))) AND

	TS=(("Gait") OR ("Gait Analysis") OR ("Gait/classification" OR "Gait/instrumentation" OR "Gait/methods" OR "Gait/organization and administration" OR "Gait/physiology" OR "Gait/standards" OR "Gait/statistics and numerical data" OR "Gait/trends") OR ("Gait Disorders, Neurologic") OR ("Walking Speed" OR "Walking") OR ("Locomotion") OR "Gait Kinetic*" OR "Gait Kinematic*" OR "range of motion")) AND TS=("Artificial Intelligence" OR Machine Learning OR Deep learning OR "Neural Networks, Computer" OR data mining OR (machine AND (learn* OR model*)) OR (statistical OR "statistical-learning") AND (strateg*) OR multilayer perceptron* OR random forest* OR bayes* network* OR support vector machine* OR nearest neighbor* OR k nearest neighbor* OR elastic net OR naive bayes* OR (classification OR regression OR estimation OR decision) AND tree OR ridge OR kernel OR ensemble OR bagging OR bagged OR boosting OR boosted OR fuzzy OR ("Predictive Value of Tests" OR "Probability Learning" OR "Forecasting" OR "Computing Methodologies" OR "Cluster Analysis") OR (Validat* OR Predict* OR Rule*) OR (Predict* AND Outcome* OR Risk* OR Model*) OR (History OR Variable* OR Criteria OR Scor* OR Characteristic* OR Finding* OR Factor*) AND (Predict* OR Model* OR Decision* OR Identif* OR Prognos*) OR (Decision* AND Model* OR Clinical*) OR (Prognostic AND History OR Variable* OR Criteria OR Scor* OR Charcteristic* OR Finding* OR Factor* OR Model*) OR (discrimination OR discriminative OR discriminatory) AND (accuracy OR ability OR performance OR value OR model OR models OR power OR efficiency) OR "Generalized linear models")) NOT TS=(Review)

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Existing predictive methods applied to gait analysis of diabetic patients: study protocol for a systematic review

Section and topic	Item No	Checklist item	Reported on page #
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	N/A
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	3
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	15
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	N/A
Support:			
Sources	5a	Indicate sources of financial or other support for the review	15
Sponsor	5b	Provide name for the review funder and/or sponsor	15
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	15
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	6
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	6
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	7
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	8
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	9
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	9
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening,	9

		eligibility and inclusion in meta-analysis)	
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	10
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	10
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	7
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	13
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	13
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	13
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	14
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	14
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	13
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	13

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.