Naik, A, Nalepa, J, Wijata, AM, Mahon, J, Mistry, D, Knowles, AT, Dawson, EA, Lip, GYH, Olier-Caparroso, I and Ortega-Martorell, S

Artificial Intelligence and Digital Twins for the Personalised Prediction of Hypertension Risk

Computers in Biology and Medicine 196 (2025) 110718

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Artificial intelligence and digital twins for the personalised prediction of hypertension risk

Akhil Naik [a,b], Jakub Nalepa [c], Agata M. Wijata [d], Joseph Mahon [a,b], Dharmesh Mistry [a,b], Adam T. Knowles [a,b], Ellen A. Dawson [b,e], Gregory Y.H. Lip [b,f,g], Ivan Olier [a,b], Sandra Ortega-Martorell [a,b,*]

[a] Data Science Research Centre, Liverpool John Moores University, Liverpool, L3 3AF, UK
[b] Liverpool Centre for Cardiovascular Science at University of Liverpool, Liverpool John Moores University and Liverpool Heart & Chest Hospital, Liverpool, UK
[c] Department of Algorithmics and Software, Silesian University of Technology, Gliwice, Poland
[d] Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland
[e] Research Institute for Sport and Exercise Science, Liverpool John Moores University, Liverpool, L3 3AF, UK
[f] Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Aalborg, Denmark
[g] Medical University of Bialystok, Bialystok, Poland

ARTICLE INFO

Keywords:
Artificial intelligence
Digital twins
Personalised care
Precision medicine
Hypertension
Risk prediction
Prevention strategies
Dynamic models
Computational simulations
Explainable AI

ABSTRACT

Hypertension is a significant global health challenge, contributing substantially to morbidity and mortality through its association with various cardiovascular diseases. Traditional approaches to hypertension risk prediction, which rely on broad epidemiological data and common risk factors, often fail to account for individual variability, highlighting the need for advanced data-driven methodologies.

This review examines the role of Artificial Intelligence (AI) and Machine Learning (ML) in enhancing the prediction of hypertension risk by incorporating a range of data sources, including clinical, lifestyle, and genetic factors. Despite promising developments, challenges such as data standardisation, the need for high-quality datasets, model explainability, and class imbalance in medical data persist. The integration of wearable technologies, alongside the potential of emerging technologies in healthcare such as digital twins, presents significant opportunities in personalising care through the dynamic modelling of individual health profiles. This review synthesises current methodologies, identifies existing gaps, and highlights the transformative potential of AI-driven, personalised hypertension prevention and management, emphasising the importance of addressing issues of reproducibility and transparency to facilitate clinical adoption.

1. Introduction

Hypertension is a chronic condition where the pressure in the arteries is consistently high and can be defined in recent guidelines as an office systolic blood pressure of >140 mmHg (130 mmHg in some guidelines) and/or a diastolic blood pressure of more than 90 mmHg [1]. If left untreated, hypertension may lead to life-threatening cardiovascular diseases including myocardial infarction, stroke, kidney failure and vascular dementia [2]. Consequently, hypertension is one of the most prevalent causes of premature death globally, leading to an estimated 10.8 million avoidable deaths every year, and an annual burden of 235 million years of life lost or years lived with a disability [3].

Globally, it has been estimated that 1.28 billion adults aged 30–79 years have hypertension with 46 % of these unaware that they have the condition [4]. In the UK, approximately 30 % of adults are diagnosed with hypertension [5], whilst in the US it is almost half of the adult population (48.1 %) [6]. Given the large social and economic costs associated with hypertension [7,8], it is crucial to prioritise preventive measures and recognise risk factors early. By tailoring prevention efforts to individual needs and circumstances, it is expected that the incidence of hypertension could be reduced and management improved [9,10].

The surging availability of health data, together with technological advancements over the last decade in Artificial Intelligence (AI) and Machine Learning (ML, a subfield of AI), has brought a paradigm shift to

This article is part of a special issue entitled: Cardiovascular Digital Twins published in Computers in Biology and Medicine.

* Corresponding author. Data Science Research Centre, Liverpool John Moores University, Liverpool, L3 3AF, UK.
E-mail address: S.OrtegaMartorell@ljmu.ac.uk (S. Ortega-Martorell).

https://doi.org/10.1016/j.compbiomed.2025.110718
Received 9 December 2024; Received in revised form 7 May 2025; Accepted 2 July 2025
Available online 8 July 2025

healthcare [11]. In the context of hypertension, AI has played a significant role in applications such as the prediction of hypertension risk, estimation of blood pressure (BP), and prediction of treatment outcomes and long-term prognosis [12]. More recently, digital twin technology has made inroads into healthcare, promising new avenues for the personalisation of care [13,14]. These innovations promise to improve precision in both preventive and therapeutic strategies, paving the way for more personalised and effective hypertension care.

Traditional methods for predicting the risk of hypertension often rely on broad epidemiological data and common risk factors such as age, gender, family history, obesity, and lifestyle habits like smoking and alcohol consumption [15,16]. While these methods have been useful in identifying high-risk groups in the population, they often lack a personalised approach, failing to account for individual variability in genetic predispositions, comorbidities, and environmental influences. Consequently, the effectiveness of these methods in guiding targeted interventions and optimising preventive strategies may be limited, underscoring the need for more nuanced, data-driven approaches that incorporate individual-specific factors. The diverse causes and effects of hypertension, shaped by individual genetic and environmental factors, highlight the need for personalised prevention and management strategies [17].

This review article aims to synthesise the current state-of-the-art evidence regarding the application of AI methodologies for the personalised prediction of hypertension risk. Second, it seeks to critically examine how personalisation has been explicitly addressed in previous studies and to identify gaps in the current body of research, emphasising early identification of risk factors, enabling the prediction of disease onset and empowering individuals to implement timely interventions to prevent or delay its progression. Finally, this review explores the impact of AI and emerging technologies such as digital twins, assessing their transformative possibilities in enabling precision-driven prevention and fostering improved health outcomes.

2. Methods

2.1. Search strategy and selection process

We conducted a comprehensive literature search on personalised prediction of hypertension risk, prioritising studies that explicitly addressed personalisation and prediction of hypertension, using three online journal databases: PubMed, Scopus and IEEE. Fig. 1 presents the flowchart indicating the selection of the articles relevant to this study, including the search terms in order of execution and their location within the manuscript (i.e. title, abstract, keywords, etc.). Initial screening was conducted by title and abstract, followed by a full-text review of potentially relevant articles to ensure they met the inclusion criteria, as shown in Fig. 1.

3. Results

3.1. Search results

A total of n = 489 studies (Fig. 1) were obtained from the search process (date: March 13, 2024), with n = 373 unique studies between the three journal databases (after removing duplicates). Since our focus was on the singular prediction of regular arterial hypertension using AI/ML, the following studies were removed: i) studies not related to essential hypertension (n = 263); ii) studies related to pulmonary hypertension (n = 11); iii) studies of comorbidities which include hypertension (n = 9) where predicting risk of hypertension was not the focus; iv) studies related to BP monitoring, BP estimation, or and profiling of patients based on their BP (n = 22); and v) studies on hypertension prognosis, treatment or management (n = 25). Additionally, papers such as editorials (n = 11), reviews (n = 7), and letters or notes (n = 11) were also excluded. Furthermore, 2 studies were excluded due to lack of quality, i.e., unclear description of methodology and data used.

In total, 12 predictive modelling articles were read in full and initially selected for our review. However, one of the articles was later retracted by the publisher (date: August 19, 2024) and subsequently removed, resulting in a final inclusion of 11 studies [18–28].

3.2. Factors used to predict hypertension risk

Eleven of the reviewed articles [18–28], summarised in Table 1, investigated the development of predictive models for hypertension. Most of these studies have relied on demographic and clinical variables to predict hypertension. However, there have been other variables examined to elucidate the factors associated with hypertension, including lifestyle and genetics.

A key lifestyle factor associated with hypertension is physical activity [29]. Two studies [24,25], examined the relationship between hypertension and physical activity (accelerometer data). Yao and Wang [24] proposed a method to process and aggregate accelerometer data to characterise complex activities, which were further modelled using random forest [30] (AUC: 0.78) for classifying hypertensive and normotensive groups. Alternatively, Chiang et al. [25] used SHapley Additive exPlanations (SHAP) [31] as a feature selection strategy for their random forest model (MAE: 5.34, RMSE: 8.24, MAPE: 4.19, R^2: 0.51), to predict diastolic and systolic blood pressure using the previous 12-, 24- and 48-h readings and provide personalised activity-based recommendations aimed at reducing blood pressure and preventing hypertension.

In line with these studies that use wearable device data, Bernal et al. [19] presented a proof-of-concept web-based ecosystem with a wearable device to measure patient data such as blood pressure, heart rate, and step count; a mobile application to collect medication and lifestyle information and a web platform to organise the data. Multiple ML algorithms were used to model this data, with the optimal model obtained using random forest (R^2: 0.89) predicting real-time BP and assessing the risk of hypertensive crisis.

Abrar et al. [18] followed a similar framework of using wearable device data to predict blood pressure along with heart rate at 1–10 days using online infinite echo state Gaussian process (OIESGP) – a recurrent neural network based framework [32], with the optimal model outperforming traditional ML models. However, Abrar et al. did not include activity data, relying only on vital signs collected from wearable devices and other physiological parameters. Furthermore, they calculated a hypertension risk score for 1, 2, and 4 years using the Framingham Risk Score, but this scoring method may not be the most suitable for the Malaysian population studied in this paper due to ethnic, lifestyle, and environmental differences.

AI/ML systems integrated with wearable devices can leverage continuous real-time monitoring of individuals and environmental data to promote healthy choices and deliver personalised recommendations. This technology also has the potential to help prevent hypertension, as seen in earlier studies such as Chiang et al. [25] and Bernal et al. [19]. However, these studies were conducted on a limited number of patients, which makes the model performance evaluation unreliable and may not reflect broader data distribution, limiting its generalisability. Additionally, the adoption of wearable devices in clinical practice is limited due to frequent calibration needs, dependence on reliable measurements in specific body positions [33], and the need for simplicity, affordability, and consistency in design and operation.

Genetic factors also play a vital role in the development of hypertension. Two studies [21,23] incorporated genetic parameters, to further improve personalised approaches to hypertension prediction. Jusic et al. [23] explored the role of microRNAs in predicting essential hypertension, with their most effective model - Support Vector Machine [34] (AUC: 0.90), using a combination of two specific microRNAs - miR-361–3p and miR-501–5p, alongside common clinical parameters to differentiate between hypertensive and non-hypertensive individuals. In

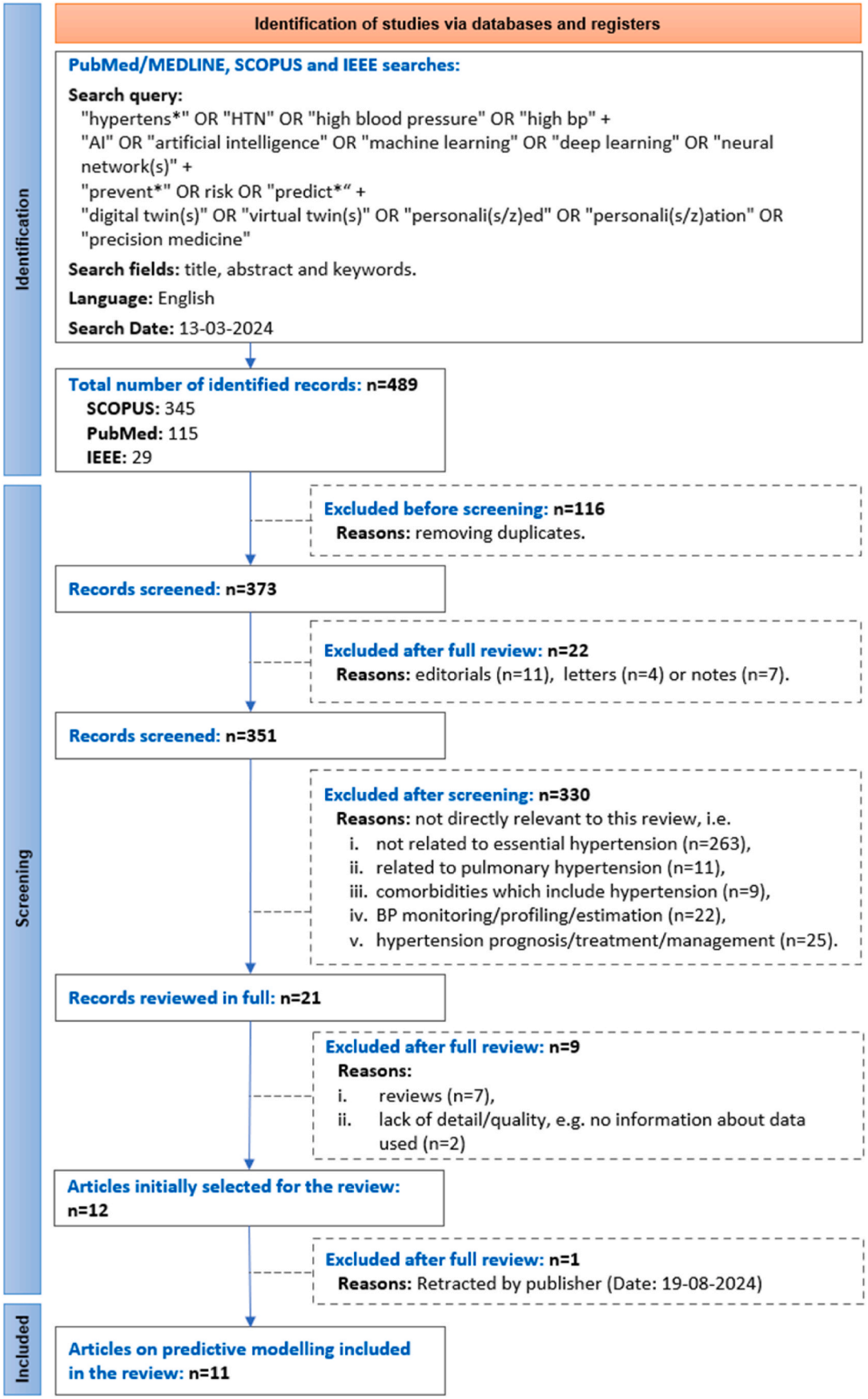

Fig. 1. Flowchart indicating the selection of articles relevant to this review.

Table 1
Summary of the 11 articles on predictive modelling that were fully reviewed. Abbreviations used are at the end of the table.

Authors and reference	Year	Downstream task	Type of task	Model (CML/DL)	Training	Features	Number of features	Dataset size	Data origin (modelling)	Evaluation	External validation	Metric	Reproducibility: Data	Reproducibility: Code
Abrar et al. [18]	2021	Prediction of the hypertension risk	Regression	CML, DL	Supervised	Blood pressure time series, physiological readings	9 + time series	4320 data records	Malaysia (UMMC)	SS (24 h of readings for training, 5 consecutive days for testing)	no	MAE, MSE, RMSE	no	no
Bernal et al. [19]	2021	Prediction of diastolic and systolic pressure	Regression	CML	Supervised	Medications, demographic and clinical parameters, weather quality, pollution, surveys (depression, social support, alcohol, stress, smoking, physical activity, diet)	>120	7 patients	Unknown[b] (volunteers)	SS (80–20)	no	RMSE, R^2	no	no
Davagdorj et al. [20]	2021	Identifying risk factors and prediction of hypertension	Classification	CML	Supervised	Demographic parameters, surveys (drinking behaviour, marital status, education, residential area), laboratory parameters	31	4094 patients	Korea (KNHANES)	CV (5 folds)	no	Pre, Re, F1-score, Acc	no	no
Li et al. [21]	2021	Classification of hypertension subtypes	Classification	CML	Supervised	Genetic and environmental parameters	up to 16227	2082 patients	USA (HyperGEN)	SS (80-10-10)	no	F1-score	no	no
Du et al. [22]	2023	Prediction of a hypertension risk	Classification	CML	Supervised	Age, gender, lifestyle, blood routine, biochemical examination parameters	57	1617 health check records	China (AHHNU)	SS (70–30)	no	Acc, Pre, Re, F1-score, macro-average ROC curve, micro-average ROC curve	no	no
Jusic et al. [23]	2023	Diagnosis of essential hypertension	Classification	CML, DL	Supervised	Clinical parameters and miRNAs	12	174 patients	Bosnia and Herzegovina (PMG)	LOOCV	no	AUC, balanced Acc, F1-score, Pre, Sen, Spe	no	no
Yao and Wang [24]	2023	Identification of the relationship between hypertension and physical exercise	Classification	CML, DL	Supervised	Gender, age, blood pressure readings, activity data	up to 3435	995 patients	USA (NHANES)	CV (50 folds)	no	Kappa, Spe, Pre, Re, F1-score, balanced Acc, ROC curve, AUC	no	no
Chiang et al. [25]	2021	Prediction of diastolic and systolic pressure	Regression	CML, DL	Supervised	Blood pressure time series and activity data	47	25 subjects	USA (ACTRI)	CV (5 folds)	no	MAE, RMSE, MAPE, R^2	no	no
Zhang et al. [26]	2023	Predicting risk of 9 chronic diseases	Classification	Survival Analysis	Supervised	clinical and demographic parameters	78	500000 participants	UK (UK Biobank)	CV (5 folds)	no	C-index, AUC, Spe, Re, Youden Index	no[a]	yes
Ismail et al. [27]	2020	Analysing health factors and prediction of 3 chronic diseases	Classification	DL	Supervised	Demographic, clinical and lifestyle parameters	20	10,806 participants	Korea (KNHANES)	CV (10 folds)	no	Acc	no	no
Nakamura et al. [28]	2023	Predict future onset of 11 chronic diseases	Classification	CML	Supervised	Physiological, biomarker, personal lifestyle, and socio-	25	3238 participants	Japan (IHPP)	CV (5 folds)	no	AUC, F1-score, Approached	no	no

(*continued on next page*)

Table 1 (*continued*)

Authors and reference	Year	Downstream task	Type of task	Model (CML/DL)	Training	Features	Number of features	Dataset size	Data origin (modelling)	Evaluation	External validation	Metric	Reproducibility: Data	Reproducibility: Code
		and set intervention goals for prevention				environmental, genetic parameters						Distance, Proportion		

Abbreviations: Acc: Accuracy; AI: Artificial Intelligence ARI: Adjusted Rand Index; AUC: Area Under [ROC] Curve; CV: Cross Validation; DCA: Decision Curve Analysis; DL: Deep Learning; LOOCV: Leave One Out Cross Validation; MAE: Mean Absolute Error; MAPE: Mean Absolute Percentage Error; ML: Machine Learning; MSE: Mean Squared Error; Pre: Precision; Re: Recall; RI: Rand Index; RMSE: Root Mean Squared Error; ROC: Receiver Operating Curve; Sen: Sensitivity; Spe: Specificity; SS: Single Split; UMMC: University Malaya Medical Centre; KNHANE: Korea National Health and Nutrition Examination Survey; HyperGEN: Hypertension Genetic Epidemiology Network; AHHNU: Affiliated Hospital of Hangzhou Normal University; PMG: Plava Medical Group; NHANES: National Health and Nutrition Examination Survey; ACTRI: Altman Clinical and Translational Research Institute, UC San Diego; IHPP: Iwaki Health Promotion Project.

Notes.

[a] The study uses UK Biobank data. The code provided in the supplementary information uses example data.

[b] While the specific origin of the volunteers is unknown, it is to be noted that the research is conducted by entities primarily based at the University of Murcia, Spain, and the University of Zurich, Switzerland.

contrast, Li et al. [21] highlighted the predictive value of single-nucleotide polymorphisms (SNPs) as genetic indicators of hypertension, with their optimised model - ScanMap [35] (F1: 0.63), employing ExAC-filtered SNPs to classify participants into non-hypertensive, mildly hypertensive, and severely hypertensive groups. While both studies demonstrated the relevance of genetic factors in hypertension, their findings were constrained by limited sample sizes. Additionally, challenges in genomic data collection and standardisation may impede the routine clinical use of these models.

3.3. Explainability as an attempt to achieve personalisation

Different model-agnostic explainability methods have been used to deliver personalisation when determining future hypertension risk. For example, Li et al. [21] and Du et al. [22] incorporated SHAP analysis to identify key factors influencing predictions, and aid end-user understanding, by providing post-hoc explainability for their models. Similar to Li et al. [21], Du et al. [22] applied their best-performing model - LightGBM [36] (AUC: 0.82), for predicting hypertension subtypes (normotensive, pre-hypertensive, hypertensive). Additionally, Du et al. designed a visual risk assessment tool that provided individualised hypertension risk scores highlighting the contributing factors. However, it remains unclear whether these scores were derived from SHAP values or directly from the model itself, emphasising the need for greater transparency in the scoring process.

In contrast to these SHAP-based approaches, Davagdorj et al. [20] proposed a two-step method to enhance end-user interpretability of a hypertension prediction model using local interpretable model-agnostic explanations (LIME) [37]. Unlike the study from Jusic et al. [23], which used recursive feature elimination, Davagdorj et al. utilised multiple statistical tests (*t*-test, chi-square) to identify significant features, which were then fed into various models to identify the optimal classifier - XGBoost [38] (Recall: 0.90, F1: 0.83, Accuracy: 0.74), based on comprehensive performance metrics. LIME analysis was subsequently applied to explain individual predictions by highlighting key risk factors associated with hypertension and normotension. While model-agnostic methods like LIME and SHAP are valuable for understanding associations between individual predictors and outcomes, they cannot be used for personalisation due to the risk of misinterpretations (confounders), such as unwarranted causal inferences and misleading feature importance [39,40].

3.4. Leveraging shared knowledge

While the studies above primarily focused directly on hypertension, three additional studies [26–28] investigated the risk of multiple chronic diseases, including hypertension, leveraging shared knowledge across conditions.

Zhang et al. [26] harnessed concepts of multitask learning [41] and Cox proportional hazards model [42] to develop a multitask-Cox model for predicting the risk of nine common chronic diseases over time (1, 3, 5, and 7 years) using UK Biobank data, achieving better performance compared to single-task models. Ismail et al. [27] proposed a deep learning framework using convolutional neural network (CNN) [43] model enhanced by Pearsons correlation analysis and pattern mining to classify three chronic diseases. Nakamura et al. [28] on the other hand, proposed a methodology to assess individuals' health state in relation to the progression of 11 chronic diseases, aiming to predict disease onset within 1–3 years and establish individualised intervention targets for disease prevention. Their approach employs health-disease phase diagrams (HDPDs), which enable the visualisation of biomarker thresholds that delineate healthy states from disease-onset states, facilitating tailored early interventions for preventing non-communicable diseases, such as hypertension. The study is limited by single-site data with low onset records, and personalisation is attempted by visualising individual-specific biomarker thresholds that separate healthy states

from disease onset, as a mechanism to tailor prevention strategies.

4. Discussion

4.1. The impact of AI and emerging technologies in healthcare

AI is dedicated to developing systems capable of performing tasks traditionally requiring human intelligence, such as learning from experience, understanding natural language, recognising patterns and making decisions. These capabilities hold significant promise for advancing healthcare [44], particularly in the prediction and management of chronic conditions such as hypertension. This transformative potential is already reshaping the healthcare landscape, with AI applications proliferating across various domains.

Key advancements include (i) rule-based expert systems and ontological data models, which enhance clinical decision support [45,46]; (ii) natural language processing techniques that analyse unstructured clinical notes and other relevant patient information, which can aid in the identification of hypertension risk factors [45]; and (iii) AI/ML models capable of integrating diverse data sources to uncover complex interactions and associations, enabling risk predictions [47,48], enabling more accurate predictions of hypertension risk and other health outcomes.

Several review studies have addressed the increasing role of AI/ML and their transformative impact on hypertension care and cardiovascular events. A systematic review by Cai et al. [49] summarised the use of ML for the prediction, diagnosis and classification of hypertension, along with automated image assessment of heart failure; whilst Amaratunga et al. [50] focused on using AI to predict clinical outcomes in four studies that utilised large datasets, two of which related to predicting hypertension. In one of these two datasets, the authors explored the potential of employing wearable biosensors and portable devices for regular monitoring of patients at risk of hypertension. The other focused on developing a risk prediction model for incident essential hypertension within one year, predicting hypertension and finding the risk factors.

Layton [51] and Krittanawong et al. [52] also provided a comprehensive summary of currently utilised ML techniques for hypertension research applications, including random forest, decision trees, k-means, and support vector machines. Particularly, Layton [51] discussed the potential of the latest technologies such as large language models, suggesting a future where healthcare professionals will adapt and learn how to collaborate with AI. Additionally, Krittanawong et al. [52] and Chaikijurajai et al. [53] discussed future outlooks of using AI with multi-omics data, as well as socioeconomic, behavioural, and environmental factors. Their discussion highlighted the potential for the identification of novel risk factors and patient phenotypes, that will enable improvements in treatment outcomes, all aimed at reducing the global burden of hypertension.

Limitations and challenges have also been discussed in Refs. [48,51], including data compatibility issues arising from different versions of the International Coding of Diseases, as well as others related to AI/ML modelling, such as model generalisability (overfitting) and interpretability (black-box models). Dzau et al. [54] explored the potential and progress in "precision hypertension" with research in genomics, data, and AI to identify specific information related to data-driven approaches leading to precision diagnosis and therapy of hypertension. Their review also covered advances in biomedical therapeutics which can provide more precise therapies in specific populations compared to existing antihypertensive treatments. The study concentrated on treatment adherence for managing hypertension, discussing the potential application of omics data to generate a polygenic risk score indicating an individual's susceptibility to developing hypertension.

AI/ML can leverage data obtained from multiple data sources and in different formats and modalities, such as text (e.g., patient's past medical history), time series (e.g., vital signs, electrocardiograms), imaging (e.g. x-rays, computed tomography scans, echocardiograms) and cross-sectional tabular data (e.g., patient demographics) [55]. High-quality data are fundamental to the success of these models, as the accuracy of predictions, including hypertension risk, depends on robust training and validation datasets [56]. Large-scale data repositories such as the UK Biobank [57], BioBank Japan [58], Australian Genomics Health Alliance [59], and MIMIC [60], have played a significant role in the understanding of disease prediction, diagnosis, and therapeutic strategies, particularly when these databases are open-access [61]; changing the landscape of modern healthcare [56].

Emerging technologies in healthcare such as digital twins [13] further enhance the potential of AI/ML in precision medicine. A digital twin of a patient represents a virtual model of such a patient that integrates information such as medical history, demographic information, laboratory tests, vital measurements, and lifestyle factors. A defining feature of digital twins in healthcare is their capacity to model individual patients, setting them apart from traditional population-based models.

Digital twin technology enables the simulation of patient-specific responses to various treatments, allowing clinicians to explore the likely outcomes of different therapeutic strategies before applying them in practice. By continuously assimilating new data from the patient, digital twins have the potential to dynamically evolve and reflect changes in health status, thereby offering powerful means for testing interventions, predicting adverse events, and optimising treatment pathways in a personalised manner.

This innovative approach coupling digital twins and AI is currently being developed and applied in extensive research programmes [14,62] to improve diagnosis, risk prediction, peri-stroke management and post-stroke rehabilitation. Still, these digital twin technologies are yet to be used for predicting hypertension risk. By simulating disease progression and forecasting conditions like hypertension, digital twins offer the potential for proactive interventions and personalised treatment plans. As a tool for risk prediction and personalised care, patient digital twins could transform hypertension management and contribute to broader advancements in precision medicine.

4.2. Advantages of AI over traditional statistical techniques

Traditional statistical approaches have been widely used for hypertension-related studies. These approaches typically involve the use of clinically relevant variables to elucidate associations, aimed at helping to understand the underlying biological and pathophysiological mechanisms [63]. These clinical variables are usually limited, as they may not capture the full complexity of individual variability, genetic factors, environmental influences, or other non-clinical determinants that contribute to the development and progression of hypertension. Consequently, traditional methods may overlook important contributors to disease outcomes, leading to an incomplete understanding of the condition.

Several studies have employed conventional regression and survival analysis methods to predict hypertension risk. For example, logistic regression [64] has been used to develop a screening tool for hypertension [65], a hypertension risk calculator [66], and the prediction of incident hypertension at 8 years among women [67]. Linear and logistic regression [68] methods have been employed to evaluate the association of genetic risk scores with longitudinal changes in blood pressure and the incidence of hypertension [69]; while Cox proportional hazards models have been utilised to estimate the risk of developing hypertension [70]. Among the widely applied techniques for hypertension risk assessment is the Framingham Hypertension Risk Score [71]. As this score is based on a single measurement of blood pressure and other risk factors, and was originally derived from a predominantly Caucasian population [71], subsequent studies have aimed to recalibrate and validate it across more diverse population datasets [72,73].

These statistical approaches have offered the advantage of being

simple to understand and with fewer and less complex computations than AI/ML models. This simplicity comes with disadvantages; they are often poor at predicting risk in populations they were not validated on [67]. Moreover, the often 'one-off assessment' only gives long-term risk prediction, and does not consider that risk factors can change over time or in response to lifestyle modifications. Conversely, there is the argument that traditional AI/ML models are difficult to interpret [63]. However, recent advancements in explainable AI/ML have brought forward techniques like SHAP value analysis and LIME, among others, which have been utilised to help interpret and explain the decisions and predictions of complex AI/ML models [20–22].

Compared to traditional statistical models, AI/ML models offer flexibility and scalability and are less reliant on prior assumptions, such as normality, linearity and equality of variance [63]. Additionally, they are routinely validated with separate datasets during model development and can easily differentiate which set of variables is most relevant in risk prediction [49,63]. AI/ML has also outperformed traditional statistical methods in reducing bias, auto-managing missing variables, controlling for confounding and handling data imbalance, which are key factors in developing accurate models [74]. This is evident in several hypertension studies, including the work by Wu et al. [75], where the authors applied extreme gradient boosting (a method from the AI/ML family) for the prediction of clinical outcomes in young hypertensive patients aged 14–39, achieving a better concordance statistic score than statistical methods such as Cox proportion regression model and recalibrated Framingham risk. Similarly, a study in Japan to predict the new onset of hypertension employed an ensemble AI/ML approach which outperformed a regression-based classification model [76].

Using AI for the personalisation of care would require analysing increased volumes of data as it would need to use detailed information and many more potential predictors. AI/ML helps towards this as it excels in large data problems, analysing multiple types of data (e.g., imaging, demographic, and laboratory findings) and addressing complex interactions within numerous variables, where commonly used statistical approaches would struggle [63,74]. Furthermore, AI/ML algorithms have proven particularly effective for omics-based data analysis, facilitating the elucidation of complex relationships among variables and their impact on primary outcomes [63], a critical advancement for precision medicine. Additionally, AI/ML enables the integration of radiomics with manually curated, clinically relevant features, allowing for a more comprehensive and nuanced analysis of data [77].

4.3. Shortcomings of the existing approaches to predict hypertension risk

Our review revealed several shortcomings of the existing techniques for personalised hypertension risk prediction, particularly in terms of data sources and explainability. These limitations point to the need for more advanced methods that can handle complex, unstructured data. In this context, our findings highlight the prevalence of tree-based ensemble models across the majority of studies. While traditional AI/ML algorithms form the backbone of data-driven approaches, high-capacity deep learning models remain underrepresented. These models have the potential to automate representation learning, which could be particularly beneficial for analysing unstructured raw clinical data, a domain where hand-crafted feature engineering often proves inadequate. Leveraging deep learning for such tasks may unlock more nuanced insights and improve performance.

Genetic and physical activity data offer significant potential for advancing personalisation in healthcare, particularly in predicting the risk of hypertension. However, several shortcomings limit their current utility. Challenges related to data collection, standardisation, and processing hinder clinical applicability, and despite substantial advancements in genomics research, the integration of these findings into routine clinical practice remains slow and complex [78]. Furthermore, genetic information is not yet widely available in healthcare settings, limiting its use in studying conditions such as hypertension.

Deep learning, with its ability to process diverse and complex data types, is well-suited to analyse the intricate and heterogeneous nature of genomic data, thereby accelerating the development of precision medicine. However, its application to genomic data is constrained by the need for large, high-quality datasets, which are often unavailable. Similarly, whilst deep learning can accurately classify accelerometer wear-site and activity intensity directly from raw acceleration data, overcoming limitations associated with the specific placement of wearables [79], these methods are reliant on robust models and consistent input data. Variability in device usage and participant adherence further complicate efforts to utilise these approaches effectively in physical activity research, impacting their reliability for hypertension risk prediction.

Another critical challenge identified in our analysis is the issue of class imbalance in medical datasets, where normal cases far outnumber abnormal ones. This imbalance can significantly degrade model performance [80,81]. Data augmentation techniques such as SMOTE (Synthetic Minority Over-Sampling Technique) [82] have been employed to address this issue, as observed in one of the studies reviewed [22]. The application of similar strategies, including random oversampling and ADASYN (Adaptive Synthetic Sampling) [83], has also been reported in other healthcare-related studies confronting class imbalance—for instance, in machine learning-based analyses of early-onset hypertension risk factors [84] and, in body composition-based hypertension prediction models [85]. However, despite their intent to mitigate class imbalance, these conventional approaches exhibit notable limitations, including the risk of overfitting due to oversampling and the potential loss of information from under-sampling. Moreover, such methods may also result in poorly calibrated models, diminishing their clinical utility by producing inaccurate risk predictions [86,87].

In contrast, deep learning techniques such as autoencoders [88] and generative adversarial networks (GANs) [89] offer alternative approaches by generating synthetic data or learning robust representations [90,91]. Specialised strategies tailor-made for deep learning, such as Balanced-MixUp [92] - a data regularisation technique, may offer further alternatives to enhance model performance in federated learning settings, by generating synthetic data and ensuring a more uniform class distribution, while preserving privacy [93]. Additionally, transfer learning approaches could also be utilised to handle data scarcity or imbalance [94–96]. They can leverage knowledge from a model trained on a different but related dataset to improve performance, particularly for the underrepresented class of a skewed distribution. Transitioning to more such advanced methods could not only enhance predictive accuracy but also broaden the applicability of AI/ML techniques in clinical settings.

Furthermore, understanding causal relationships and identifying relevant risk factors is also crucial in healthcare. However, the black-box nature of many ML models limits transparency and trust. Interpretable models (e.g., decision trees) can offer a clear, human-understandable logic, supporting clinical trust and accountability, but may lack the accuracy of complex models. Explainability methods, such as SHAP, enable post hoc analysis, shedding light on the predictions from black-box models by highlighting influential features and helping identify potential risk factors. However, these methods primarily indicate correlations rather than causal relationships, and can be sensitive to model variations. To address these limitations, deep learning frameworks like graph neural networks (GNNs) [97] could offer a promising solution as they can model complex relationships between entities – such as patients, diseases, treatments, etc. By capturing both local and global contexts, GNNs can enhance predictive performance and support interpretability via node and edge analysis. This makes them well-suited for incorporating domain knowledge and understanding relational or causal structures in healthcare data. Additionally, they can form the backbone of digital twins, enabling the integration of multiscale

computational modelling with AI, as demonstrated by a recent study that used GNNs to forecast blood pressure and GANs for synthetic data generation, albeit with a focus on treatment rather than prevention or prediction of hypertension [98].

In addition to interpretability, issues of bias and generalisability remain major concerns in existing approaches. Notably, each of the studies [18–28] used data from a single country (see data origin column in Table 1), limiting demographic diversity and contributing to models being trained on datasets that may lead to biased predictions. As a result, many models may not generalise well to populations differing in ethnicity, sex, age, socioeconomic status, or geographical location. Such biases risk exacerbating health disparities if not properly addressed. Furthermore, none of these studies assessed their models using external datasets (see external validation column in Table 1). This is critical, as models that perform well on internal validation often fail to replicate their performance when tested on independent, external cohorts. To improve generalisability, models must be developed using larger, more representative datasets and validated rigorously across diverse populations. Efforts to mitigate bias are essential to ensure that AI models for hypertension risk prediction are equitable and clinically reliable across varied patient groups. Coupled with interpretable modelling, these practices can enhance both the trustworthiness and effectiveness of AI-based personalised medicine.

Building on the concept of leveraging shared knowledge to address disease interdependencies, the multitask learning approach described by Zhang et al. [26] demonstrated the ability to account for disease comorbidities by capturing interdependencies across conditions, although its reliance on the Cox proportional hazards model limits its ability to model complex, non-linear interactions. The study of Ismail et al. [27] emphasised the ability of deep learning models to capture more complex, non-linear, complex patterns, enhancing the classification of chronic diseases. In addition, the health-disease phase diagrams of Nakamura et al. [28] leveraged explainable AI techniques to tailor early prevention efforts by visualising biomarker thresholds, offering a more individualised (albeit not fully personalised) strategy for managing hypertension within the broader context of chronic disease prevention. However, despite their advancements, none of these approaches fully integrate individual-specific contexts, behaviours, or dynamic health trajectories into their predictive or intervention frameworks, and cannot be considered personalised.

The analysis conducted in this review also showed that most methods cannot be easily reproduced (see Table 1), as the utilised datasets and implementations are not publicly available. Additionally, these datasets substantially differ in terms of included predictors and underlying characteristics. Furthermore, the authors often follow inconsistent validation procedures and experimental regimes, defined as the training-test dataset splits and quality metrics used for quantifying the generalisation capabilities of AI/ML models, which makes direct comparisons between algorithms extremely challenging, even when addressing similar tasks. Given these issues, and in light of the reproducibility crisis affecting (not only) medical data analysis [99], addressing these gaps and ensuring full experimental reproducibility is crucial for accelerating the adoption of data-driven techniques in hypertension management. Standardising methodologies for developing and evaluating AI/ML approaches in hypertension prediction and management is essential to facilitate objective comparisons, ultimately helping to identify the most effective and generalisable models for integration into routine clinical practice.

As discussed by Royen et al. [100], there are five critical quality criteria which are pivotal to be met while introducing AI/ML solutions into the medical data analysis field. They include 1) reproducibility of AI/ML techniques, 2) their clear intended use, 3) rigorous validation of such methods with 4) appropriate sample size, as well as 5) openness of the data and software. Although we may indeed observe a promising trend in increasing the number of patients being included in the validation process of emerging AI/ML solutions in the context of hypertension management, our analysis revealed that the reported studies are extremely difficult or impossible to reproduce (as the code and data or the training/test dataset splits are not publicly available), hence the numerical results reported in such works cannot be compared directly; therefore, it is challenging to monitor the progress in the field of hypertension risk prediction and disease management.

Addressing the aforementioned issues and meeting the quality criteria are, in our view, of paramount importance. We emphasise that it is the responsibility of authors, researchers, and reviewers (readers) to ensure that clinical AI prediction modelling studies in cardiovascular health meet the five key quality criteria. These include comprehensive reporting, clearly defined model usage, rigorous validation, sufficient sample sizes, and transparency of code and software. Adhering to these standards will enhance the quality, clinical relevance, and impact of AI prediction studies in this field.

4.4. Precision medicine and personalisation – the likely future of hypertension prevention and management

The convergence of AI, digital twin technology, and precision medicine, is impacting healthcare by addressing complex problems in personalised care [56], offering transformative potential for the prediction and management of hypertension. In the context of hypertension, these technologies can integrate diverse data streams, such as wearable device outputs, environmental exposures, and behavioural patterns, to create dynamic models of an individual's health, which can be used to identify nuanced risk markers and simulate how different factors contribute to hypertension risk. This approach not only refines risk prediction but also enables the testing of potential interventions in the virtual space before applying them in real life, minimising trial-and-error approaches in patient care.

By incorporating digital twins, precision medicine will move beyond static datasets to continuous, adaptive insights. For hypertension, this would mean personalised prevention strategies can be modelled and optimised in real time. For instance, a digital twin might simulate the impact of dietary changes, medication adjustments, or stress-reduction techniques on blood pressure, offering tailored recommendations that are dynamically updated based on new data from wearables or routine check-ups.

From a data science and engineering perspective, building effective digital twins for the prediction and management of hypertension requires the integration of multimodal data through robust pipelines capable of real-time data ingestion, preprocessing, and feature extraction. Advanced AI/ML models, such as recurrent neural networks and graph-based architectures, can be leveraged to capture temporal and relational aspects of physiological and behavioural data. In addition, developing reliable patient-specific simulations demands sophisticated calibration and validation techniques, ensuring that digital twin outputs accurately reflect clinical realities. Engineering scalable frameworks for deploying digital twins will also be crucial, necessitating the development of secure, interoperable systems that maintain data privacy and support continuous model updating as new patient data becomes available.

Specifically for risk prediction, digital twins can integrate longitudinal datasets encompassing blood pressure trajectories, medication adherence patterns, genetic predisposition, and lifestyle factors to predict an individual's future risk of developing hypertension. Through counterfactual simulations, digital twins can also help assess how different modifiable factors, such as reductions in salt intake or increases in physical activity, may alter the risk profile over time. This predictive capacity enables early, highly targeted interventions aimed at delaying or preventing the onset of hypertension altogether.

The integration of AI with digital twins has the potential to enhance the precision and scalability of hypertension prevention and management. It will empower patients to engage actively with their health through actionable insights whilst supporting clinicians in making data-

driven, individualised decisions. This approach has the potential not only to improve prevention, diagnosis, and treatment but also to optimise resource utilisation, offering a pathway to reduce the burden of hypertension on healthcare systems and society.

4.5. Challenges to the clinical adoption of digital twin technology

Realising the full potential of digital twins in clinical settings requires addressing several critical challenges. As highlighted by the EDITH consortium [13], key barriers include the lack of robust, individual-level predictive models that can accurately represent the complexity and variability of human physiology. The integration of heterogeneous in-silico modelling approaches, ranging from data-driven to mechanistic methods, is also technically demanding and is further constrained by the scarcity of high-quality, representative datasets for model development and validation. Data accessibility is also constrained by stringent privacy regulations and a lack of consensus on data types, standards, and governance requirements.

In addition, regulatory uncertainty (particularly the absence of harmonised frameworks for evaluating the credibility and safety of in-silico technologies) impedes their incorporation into clinical and regulatory pathways. Broader implementation efforts are further hindered by limited stakeholder awareness, inadequate interoperability with electronic health record systems, and a shortage of professionals with expertise bridging clinical practice and computational science. These challenges are reflected in prior studies [101–103], which have similarly emphasised the importance of data quality, infrastructure, security, and validated methodologies.

Addressing these interrelated barriers will require coordinated, interdisciplinary collaboration to establish common standards, enhance data governance, and create supportive regulatory and commercial environments conducive to innovation and clinical translation.

5. Conclusion

The integration of AI/ML into personalised hypertension risk prediction is a rapidly advancing field, with emerging technologies such as digital twins offering promising avenues for enhanced precision in hypertension care. AI models, especially those that leverage diverse data sources, including genetic markers, wearable device outputs, and clinical variables, have demonstrated considerable potential in improving prediction accuracy and enabling more targeted, personalised prevention strategies. However, the field faces significant challenges, such as issues related to reproducibility, inconsistent methodological standards, and limited access to heterogeneous datasets. These barriers hinder the generalisability of findings and the broader adoption of AI-based solutions in clinical practice.

To realise the full potential of AI in hypertension prevention and management, it is essential to overcome these obstacles through the standardisation of methodologies, transparent reporting, and rigorous validation processes across diverse populations. A particularly promising innovation in this area is digital twin technology, which holds the potential to enable healthcare providers to both, tailor prevention measures and manage hypertension in real time, by simulating treatment outcomes and optimising personalised care strategies. These technologies have the potential not only to improve prevention, diagnosis, and treatment but also to optimise resource utilisation, offering a pathway to ultimately improve patient outcomes and reduce the global burden of hypertension. A coordinated, interdisciplinary collaboration will be required to realise the full potential of these novel technologies in clinical settings.

CRediT authorship contribution statement

Akhil Naik: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. **Jakub Nalepa:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Agata M. Wijata:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. **Joseph Mahon:** Writing – review & editing, Validation, Formal analysis, Data curation. **Dharmesh Mistry:** Writing – review & editing, Validation. **Adam T. Knowles:** Writing – review & editing, Validation. **Ellen A. Dawson:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis. **Gregory Y.H. Lip:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition. **Ivan Olier:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Sandra Ortega-Martorell:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization.

Ethics statement

This review paper is based on publicly available literature and does not involve primary data collection or research involving human participants or animals. The authors confirm that the work complies with the ethical standards for scholarly publishing, including:

- **Originality and Plagiarism**: The manuscript is original, has not been published elsewhere, and does not contain plagiarized material. All sources are appropriately cited and referenced.
- **Conflict of Interest**: There are no conflicts of interest that could influence the outcomes or interpretations presented in this review.
- **Acknowledgment of Sources**: All studies, data, and other materials referenced in this review are duly acknowledged to ensure transparency and credit to the original authors.
- **Compliance with Journal Guidelines**: The manuscript adheres to the ethical requirements and policies of the journal.

Declaration of competing interest

The below authors declare the following financial interests/personal relationships which may be considered as potential competing interests: **DM -** Team member of the TARGET project on health virtual twins for personalised management of atrial fibrillation and stroke (grant agreement no. 101136244).**ATK -** Team member of the TARGET project on health virtual twins for personalised management of atrial fibrillation and stroke (grant agreement no. 101136244).**EAD -** Team member of the TARGET project on health virtual twins for personalised management of atrial fibrillation and stroke (grant agreement no. 101136244).**GYHL** - Consultant and speaker for BMS/Pfizer, Boehringer Ingelheim, Daiichi-Sankyo, and Anthos. No fees are received personally. He is a National Institute for Health and Care Research (NIHR) Senior Investigator and co-PI of the AFFIRMO project on multimorbidity in AF (grant agreement no. 899871), TARGET project on health virtual twins for personalised management of atrial fibrillation and stroke (grant agreement no. 101136244) and ARISTOTELES project on artificial intelligence for the management of chronic long-term conditions (grant agreement no. 101080189), which are all funded by the EU's Horizon Europe Research & Innovation programme.**IO** - Methodological lead of the TARGET project on health virtual twins for personalised management of atrial fibrillation and stroke (grant agreement no. 101136244) and partner lead in the ARISTOTELES project on artificial intelligence for the management of chronic long-term conditions (grant agreement no. 101080189), both funded by the EU's Horizon Europe Research & Innovation programme. No fees are received personally.**SOM** - Principal Investigator of the TARGET project on health virtual twins for personalised management of atrial fibrillation and stroke (grant agreement no. 101136244) and senior investigator in the ARISTOTELES project on artificial intelligence for the management of chronic long-term conditions (grant agreement no. 101080189), both funded by the EU's

Horizon Europe Research & Innovation programme. She is also a member of the board of the ART (Ageing Research Translation) of Healthy Ageing Network funded by the Biotechnology and Biological Sciences Research Council (BBSRC). No fees are received personally.

Acknowledgements

This research was partially funded by Liverpool John Moores University via the Thematic Doctoral Programme: 'Misconceptions in the Treatment of Hypertension: MITH', supporting AN and JM. JN and AMW were supported by the Silesian University of Technology funds through the Excellence Initiative—Research University program (Grant 02/080/SDU/10-21-01), and by the Silesian University of Technology funds through the grant for maintaining and developing research potential. AMW was supported by the Just Transition Fund—project entitled "Development of the Silesian biomedical engineering potential in the face of the challenges of the digital and green economy (BioMeDiG)" (FESL.10.25-IZ.01-07G5/23). DM, EAD, GYH, IO and SOM are supported by the EU project TARGET on health virtual twins for personalised management of atrial fibrillation and stroke (grant agreement no. 101136244) funded by the Horizon Europe Research & Innovation programme.

References

[1] J.W. McEvoy, C.P. McCarthy, R.M. Bruno, S. Brouwers, M.D. Canavan, C. Ceconi, R.M. Christodorescu, S.S. Daskalopoulou, C.J. Ferro, E. Gerdts, H. Hanssen, J. Harris, L. Lauder, R.J. McManus, G.J. Molloy, K. Rahimi, V. Regitz-Zagrosek, G. P. Rossi, E.C. Sandset, B. Scheenaerts, J.A. Staessen, I. Uchmanowicz, M. Volterrani, R.M. Touyz, A. Abreu, M.H. Olsen, M. Ambrosetti, E. Androulakis, L.E. Bang, J.N. Bech, M.A. Borger, P. Boutouyrie, L. Bronze, S. Buccheri, R. Dalmau, M.C. De Pablo Zarzosa, C. Delles, M.M. Fiuza, R. Gabulova, B. O. Haugen, C. Heiss, B. Ibanez, S. James, V. Kapil, M. Kayikçioglu, L. Køber, K. C. Koskinas, E.T. Locati, S. MacDonald, A.S. Mihailidou, B. Mihaylova, R. Mindham, M.B. Mortensen, S. Nardai, L. Neubeck, J.C. Nielsen, P.M. Nilsson, A.A. Pasquet, M.M. Pedro, E. Prescott, A. Rakisheva, E. Rietzschel, B. Rocca, X. Rossello, J.-P. Schmid, E. Shantsila, I. Sudano, A.T. Timóteo, G. Tsivgoulis, A. Ungar, I. Vaartjes, F. Visseren, H. Voeller, C. Vrints, A. Witkowski, M.-C. Zennaro, K. Zeppenfeld, N. Shuka, N. Laredj, N. Pavo, U. Mirzoyev, P. van de Borne, Š. Sokolović, A. Postadzhiyan, J. Samardzic, P. Agathangelou, J. Widimsky, M.H. Olsen, W.M. El-Kilany, P. Pauklin, J.A. Laukkanen, R. Boulestreau, B. Tsinamdzgvrishvili, U. Kintscher, M. Marketou, D. Páll, þ. J. Hrafnkelsdóttir, E. Dolan, T. Wolak, G. Bilo, M.K. Tundybayeva, E. Mirrakhimov, K. Trusinskis, G. Kiwan, O. Msalem, J. Badarienė, C.-A. Banu, M. M. Balbi, A. Caraus, A. Boskovic, N. Mouine, T. Vromen, M. Bosevski, H. B. Midtbø, A. Doroszko, H. Dores, E. Badila, R. Bini, D.V. Simić, Z. Fras, P. Mazón, J. Spaak, T. Burkard, E. Barakat, S. Abdessalem, Y. Gunes, Y.M. Sirenko, A.J. B. Brady, G.A. Khamidullaeva, 2024 ESC guidelines for the management of elevated blood pressure and hypertension, Eur. Heart J. 45 (2024) 3912–4018, https://doi.org/10.1093/eurheartj/ehae178.

[2] British Heart Foundation, High blood pressure, (n.d.). https://www.bhf.org.uk/informationsupport/risk-factors/high-blood-pressure (accessed April 19, 2024).

[3] World Health Organization, Global Report on Hypertension: the Race Against a Silent Killer - World | ReliefWeb, (n.d.). https://reliefweb.int/report/world/global-report-hypertension-race-against-silent-killer (accessed April 23, 2024).

[4] World Health Organization, Hypertension, (n.d.). https://www.who.int/news-room/fact-sheets/detail/hypertension (accessed April 19, 2024).

[5] Adults' health: Hypertension - NHS England Digital, (n.d.). https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2021-part-2/adult-health-hypertension (accessed July 6, 2024).

[6] C.D. Fryar, B. Kit, M.D. Carroll, J. Afful, Key findings data from the national health and nutrition examination survey. https://www.cdc.gov/nchs/products/index.htm, 2021.

[7] K.T. Mills, A. Stefanescu, J. He, The global epidemiology of hypertension, Nat. Rev. Nephrol. 16 (2020) 223–237, https://doi.org/10.1038/s41581-019-0244-2.

[8] E. Wierzejska, B. Giernaś, A. Lipiak, M. Karasiewicz, M. Cofta, R. Staszewski, A global perspective on the costs of hypertension: a systematic review, Arch. Med. Sci. 16 (2020) 1078–1091, https://doi.org/10.5114/aoms.2020.92689.

[9] M. Volpe, G. Tocci, A. de la Sierra, R. Kreutz, S. Laurent, A.J. Manolis, K. Tsioufis, Personalised single-pill combination therapy in hypertensive patients: an update of a practical treatment platform, High Blood Pres. Cardiovasc. Prev. 24 (2017) 463–472, https://doi.org/10.1007/s40292-017-0239-7.

[10] E.K. Oikonomou, E.S. Spatz, M.A. Suchard, R. Khera, Individualising intensive systolic blood pressure reduction in hypertension using computational trial phenomaps and machine learning: a post-hoc analysis of randomised clinical trials, Lancet Digit. Health 4 (2022), https://doi.org/10.1016/S2589-7500(22)00170-4.

[11] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, Y. Wang, Artificial intelligence in healthcare: past, present and future, Stroke Vasc Neurol 2 (2017), https://doi.org/10.1136/svn-2017-000101.

[12] V. Visco, C. Izzo, C. Mancusi, A. Rispoli, M. Tedeschi, N. Virtuoso, A. Giano, R. Gioia, A. Melfi, B. Serio, M.R. Rusciano, P. Di Pietro, A. Bramanti, G. Galasso, G. D'Angelo, A. Carrizzo, C. Vecchione, M. Ciccarelli, Artificial intelligence in hypertension management: an ace up your sleeve, J. Cardiovasc. Dev. Dis. 10 (2023), https://doi.org/10.3390/jcdd10020074.

[13] M. Viceconti, M. De Vos, S. Mellone, L. Geris, Position paper from the digital twins in healthcare to the virtual human twin: a moon-shot project for digital health research, IEEE J. Biomed. Health Inform. (2023), https://doi.org/10.1109/JBHI.2023.3323688. PP.

[14] S. Ortega-Martorell, I. Olier, M. Ohlsson, G.Y.H. Lip, T. Consortium, TARGET: a major European project aiming to advance the personalised management of atrial fibrillation-related stroke via the development of health virtual twins technology and artificial intelligence, Thromb. Haemost. (2024), https://doi.org/10.1055/a-2438-5671.

[15] H. Zhao, X. Zhang, Y. Xu, L. Gao, Z. Ma, Y. Sun, W. Wang, Predicting the risk of hypertension based on several easy-to-collect risk factors: a machine learning method, Front. Public Health 9 (2021), https://doi.org/10.3389/fpubh.2021.619429.

[16] S.M.S. Islam, A. Talukder, MdA. Awal, MdM.U. Siddiqui, MdM. Ahamad, B. Ahammed, L.B. Rawal, R. Alizadehsani, J. Abawajy, L. Laranjo, C.K. Chow, R. Maddison, Machine learning approaches for predicting hypertension and its associated factors using population-level data from three South Asian countries, Front. Cardiovasc. Med. 9 (2022), https://doi.org/10.3389/fcvm.2022.839379.

[17] S.T. Turner, G.L. Schwartz, E. Boerwinkle, Personalized medicine for high blood pressure, Hypertension 50 (2007) 1–5, https://doi.org/10.1161/HYPERTENSIONAHA.107.087049.

[18] S. Abrar, C.K. Loo, N. Kubota, A multi-agent approach for personalized hypertension risk prediction, IEEE Access 9 (2021) 75090–75106, https://doi.org/10.1109/ACCESS.2021.3074791.

[19] S. López Bernal, J. Martínez Valverde, A.H. Celdrán, G. Martínez Pérez, SENIOR: an intelligent web-based ecosystem to predict high blood pressure adverse events using biomarkers and environmental data. https://doi.org/10.3390/app1106, 2021.

[20] K. Davagdorj, M. Li, K.H. Ryu, Local interpretable model-agnostic explanations of predictive models for hypertension, in: Smart Innovation, Systems and Technologies, Springer Science and Business Media Deutschland GmbH, 2021, pp. 426–433, https://doi.org/10.1007/978-981-33-6757-9_53.

[21] Y. Li, S.J. Shah, D. Arnett, R. Irvin, Y. Luo, SNPs filtered by allele frequency improve the prediction of hypertension subtypes, in: Proceedings - 2021 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2021, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 2796–2802, https://doi.org/10.1109/BIBM52615.2021.9669758.

[22] J. Du, X. Chang, C. Ye, Y. Zeng, S. Yang, S. Wu, L. Li, Developing a hypertension visualization risk prediction system utilizing machine learning and health check-up data, Sci. Rep. 13 (2023), https://doi.org/10.1038/s41598-023-46281-y.

[23] A. Jusic, I. Junuzovic, A. Hujdurovic, L. Zhang, M. Vausort, Y. Devaux, A machine learning model based on microRNAs for the diagnosis of essential hypertension, Noncoding RNA 9 (2023), https://doi.org/10.3390/ncrna9060064.

[24] X. Yao, Q. Wang, A Classification Approach for the Sports Behavior Data with Random Forest, SPIE-Intl Soc Optical Eng, 2023, p. 20, https://doi.org/10.1117/12.2687417.

[25] P.-H. Chiang, M. Wong, S. Dey, Using wearables and machine learning to enable personalized lifestyle recommendations to improve blood pressure, IEEE J. Transl. Eng. Health Med. 9 (2021) 1–13, https://doi.org/10.1109/JTEHM.2021.3098173.

[26] S. Zhang, F. Yang, L. Wang, S. Si, J. Zhang, F. Xue, Personalized prediction for multiple chronic diseases by developing the multi-task cox learning model, PLoS Comput. Biol. 19 (2023), https://doi.org/10.1371/journal.pcbi.1011396.

[27] W.N. Ismail, M.M. Hassan, H.A. Alsalamah, G. Fortino, CNN-based health model for regular health factors analysis in internet-of-medical things environment, IEEE Access 8 (2020) 52541–52549, https://doi.org/10.1109/ACCESS.2020.2980938.

[28] K. Nakamura, E. Uchino, N. Sato, A. Araki, K. Terayama, R. Kojima, K. Murashita, K. Itoh, T. Mikami, Y. Tamada, Y. Okuno, Individual health-disease phase diagrams for disease prevention based on machine learning, J. Biomed. Inf. 144 (2023) 104448, https://doi.org/10.1016/j.jbi.2023.104448.

[29] S.M. Hegde, S.D. Solomon, Influence of physical activity on hypertension and cardiac structure and function, Curr. Hypertens. Rep. 17 (2015), https://doi.org/10.1007/s11906-015-0588-3.

[30] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[31] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, in: Adv Neural Inf Process Syst, 2017.

[32] H. Soh, Y. Demiris, Iterative temporal learning and prediction with the sparse online echo state gaussian process, in: The 2012 International Joint Conference on Neural Networks (IJCNN), IEEE, 2012, pp. 1–8, https://doi.org/10.1109/IJCNN.2012.6252504.

[33] D. Konstantinidis, P. Iliakis, F. Tatakis, K. Thomopoulos, K. Dimitriadis, D. Tousoulis, K. Tsioufis, Wearable blood pressure measurement devices and new approaches in hypertension management: the digital era, J. Hum. Hypertens. 36 (2022) 945–951, https://doi.org/10.1038/s41371-022-00675-z.

[34] M.A. Hearst, S.T. Dumais, E. Osuna, J. Platt, B. Scholkopf, Support vector machines, IEEE Intell. Syst. Their Appl. 13 (1998) 18–28, https://doi.org/10.1109/5254.708428.

[35] Y. Luo, C. Mao, ScanMap: supervised confounding aware non-negative matrix factorization for polygenic risk modeling, in: Proc Mach Learn Res, 2020.

[36] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.Y. Liu, LightGBM: a highly efficient gradient boosting decision tree, in: Adv Neural Inf Process Syst, 2017.

[37] M.T. Ribeiro, S. Singh, C. Guestrin, Why should I trust you?, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM, New York, NY, USA, 2016, pp. 1135–1144, https://doi.org/10.1145/2939672.2939778.

[38] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.

[39] C. Molnar, G. König, J. Herbinger, T. Freiesleben, S. Dandl, C.A. Scholbeck, G. Casalicchio, M. Grosse-Wentrup, B. Bischl, General Pitfalls of Model-Agnostic Interpretation Methods for Machine Learning Models, 2022, pp. 39–68, https://doi.org/10.1007/978-3-031-04083-2_4.

[40] X. Huang, J. Marques-Silva, On the failings of shapley values for explainability, Int. J. Approx. Reason. 171 (2024), https://doi.org/10.1016/j.ijar.2023.109112.

[41] Y. Zhang, Q. Yang, A survey on multi-task learning, IEEE Trans. Knowl. Data Eng. 34 (2022) 5586–5609, https://doi.org/10.1109/TKDE.2021.3070203.

[42] S.V. Deo, V. Deo, V. Sundaram, Survival Analysis—part 2: cox proportional hazards model, Indian J. Thorac. Cardiovasc. Surg. 37 (2021) 229, https://doi.org/10.1007/S12055-020-01108-7.

[43] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, https://doi.org/10.1109/CVPR.2015.7298594.

[44] P. Kumar, S. Chauhan, L.K. Awasthi, Artificial intelligence in healthcare: review, ethics, trust challenges & future research directions, Eng. Appl. Artif. Intell. 120 (2023) 105894, https://doi.org/10.1016/j.engappai.2023.105894.

[45] T. Davenport, R. Kalakota, The potential for artificial intelligence in healthcare, Future Healthc J 6 (2019), https://doi.org/10.7861/futurehosp.6-2-94.

[46] S. Rajapaksha, W.J.A. Abhayarathne, S.G.K. Kumari, M.V.L.U. De Silva, W.M.S. M. Wijesuriya, A Mobile application to predict and manage high blood pressure and personalized recommendations, in: 2019 International Conference on Advancements in Computing, ICAC 2019, 2019, https://doi.org/10.1109/ICAC49085.2019.9103337.

[47] S. Padmanabhan, T.Q.B. Tran, A.F. Dominiczak, Artificial intelligence in hypertension: seeing through a glass darkly, Circ. Res. 128 (2021) 1100–1118, https://doi.org/10.1161/CIRCRESAHA.121.318106.

[48] V. Visco, C. Izzo, C. Mancusi, A. Rispoli, M. Tedeschi, N. Virtuoso, A. Giano, R. Gioia, A. Melfi, B. Serio, M.R. Rusciano, P. Di Pietro, A. Bramanti, G. Galasso, G. D'Angelo, A. Carrizzo, C. Vecchione, M. Ciccarelli, Artificial intelligence in hypertension management: an ace up your sleeve, J. Cardiovasc. Dev. Dis. 10 (2023), https://doi.org/10.3390/jcdd10020074.

[49] A. Cai, Y. Zhu, S.A. Clarkson, Y. Feng, The use of machine learning for the care of hypertension and heart failure, JACC (J. Am. Coll. Cardiol.): Asia 1 (2021) 162–172, https://doi.org/10.1016/j.jacasi.2021.07.005.

[50] D. Amaratunga, J. Cabrera, D. Sargsyan, J.B. Kostis, S. Zinonos, W.J. Kostis, Uses and opportunities for machine learning in hypertension research, Int. J. Cardiol. Hypertens 5 (2020), https://doi.org/10.1016/j.ijchy.2020.100027.

[51] A.T. Layton, AI, machine learning, and ChatGPT in hypertension, Hypertension 81 (2024) 709–716, https://doi.org/10.1161/HYPERTENSIONAHA.124.19468.

[52] C. Krittanawong, A.S. Bomback, U. Baber, S. Bangalore, F.H. Messerli, W. H. Wilson Tang, Future direction for using artificial intelligence to predict and manage hypertension, Curr. Hypertens. Rep. 20 (2018), https://doi.org/10.1007/s11906-018-0875-x.

[53] T. Chaikijurajai, L.J. Laffin, W.H. Wilson Tang, Artificial intelligence and hypertension: recent advances and future outlook, Am. J. Hypertens. 33 (2020) 967–974, https://doi.org/10.1093/ajh/hpaa102.

[54] V.J. Dzau, C.P. Hodgkinson, Precision hypertension, Hypertension (2023), https://doi.org/10.1161/hypertensionaha.123.21710.

[55] F. Krones, U. Marikkar, G. Parsons, A. Szmul, A. Mahdi, Review of multimodal machine learning approaches in healthcare, Inf. Fusion 114 (2025) 102690, https://doi.org/10.1016/j.inffus.2024.102690.

[56] K.B. Johnson, W.Q. Wei, D. Weeraratne, M.E. Frisse, K. Misulis, K. Rhee, J. Zhao, J.L. Snowdon, Precision medicine, AI, and the future of personalized health care, Clin. Transl. Sci. 14 (2021), https://doi.org/10.1111/cts.12884.

[57] C. Bycroft, C. Freeman, D. Petkova, G. Band, L.T. Elliott, K. Sharp, A. Motyer, D. Vukcevic, O. Delaneau, J. O'Connell, A. Cortes, S. Welsh, A. Young, M. Effingham, G. McVean, S. Leslie, N. Allen, P. Donnelly, J. Marchini, The UK biobank resource with deep phenotyping and genomic data, Nature 562 (2018), https://doi.org/10.1038/s41586-018-0579-z.

[58] A. Nagai, M. Hirata, Y. Kamatani, K. Muto, K. Matsuda, Y. Kiyohara, T. Ninomiya, A. Tamakoshi, Z. Yamagata, T. Mushiroda, Y. Murakami, K. Yuji, Y. Furukawa, H. Zembutsu, T. Tanaka, Y. Ohnishi, Y. Nakamura, M. Kubo, M. Shiono, K. Misumi, R. Kaieda, H. Harada, S. Minami, M. Emi, N. Emoto, H. Daida, K. Miyauchi, A. Murakami, S. Asai, M. Moriyama, Y. Takahashi, T. Fujioka, W. Obara, S. Mori, H. Ito, S. Nagayama, Y. Miki, A. Masumoto, A. Yamada, Y. Nishizawa, K. Kodama, H. Kutsumi, Y. Sugimoto, Y. Koretsune, H. Kusuoka, H. Yanai, Overview of the BioBank Japan project: study design and profile, J. Epidemiol. 27 (2017), https://doi.org/10.1016/j.je.2016.12.005.

[59] Z. Stark, T. Boughtwood, P. Phillips, J. Christodoulou, D.P. Hansen, J. Braithwaite, A.J. Newson, C.L. Gaff, A.H. Sinclair, K.N. North, Australian genomics: a federated model for integrating genomics into healthcare, Am. J. Hum. Genet. 105 (2019), https://doi.org/10.1016/j.ajhg.2019.06.003.

[60] A.E.W. Johnson, L. Bulgarelli, L. Shen, A. Gayles, A. Shammout, S. Horng, T. J. Pollard, S. Hao, B. Moody, B. Gow, L.-W.H. Lehman, L.A. Celi, R.G. Mark, MIMIC-IV, a freely accessible electronic health record dataset, Sci. Data 10 (2023) 1, https://doi.org/10.1038/s41597-022-01899-x.

[61] I. Olier, S. Ortega-Martorell, M. Pieroni, G.Y.H. Lip, How machine learning is impacting research in atrial fibrillation: implications for risk prediction and future management, Cardiovasc. Res. 117 (2021) 1700–1717, https://doi.org/10.1093/cvr/cvab169.

[62] S. Ortega-Martorell, I. Olier, G.Y.H. LipOhlsson, G.Y.H. Lip, TARGET: a major European project aiming to advance the personalised management of atrial fibrillation-related stroke via the development of health virtual twins technology and artificial intelligence, Thromb. Haemost. (2024), https://doi.org/10.1055/a-2438-5671.

[63] H.S.R. Rajula, G. Verlato, M. Manchia, N. Antonucci, V. Fanos, Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment, Medicina (Lithuania) 56 (2020), https://doi.org/10.3390/medicina56090455.

[64] S. Sperandei, Understanding logistic regression analysis, Biochem. Med. 24 (2014), https://doi.org/10.11613/BM.2014.003.

[65] Q. Jiang, D. Gong, H. Li, D. Zhang, S. Hu, Q. Xia, H. Yuan, P. Zhou, Y. Zhang, X. Liu, M. Sun, J. Lv, C. Li, Development and validation of a risk score screening tool to identify people at risk for hypertension in shanghai, China, Risk Manag. Healthc. Policy 15 (2022) 553–562, https://doi.org/10.2147/RMHP.S354057.

[66] A.V. Kshirsagar, Y. lin Chiu, A.S. Bomback, P.A. August, A.J. Viera, R. E. Colindres, H. Bang, A hypertension risk score for middle-aged and older adults, J. Clin. Hypertens. 12 (2010), https://doi.org/10.1111/j.1751-7176.2010.00343.x.

[67] N.P. Paynter, N.R. Cook, B.M. Everett, H.D. Sesso, J.E. Buring, P.M. Ridker, Prediction of incident hypertension risk in women with currently normal blood pressure, Am. J. Med. 122 (2009), https://doi.org/10.1016/j.amjmed.2008.10.034.

[68] G. Tripepi, K.J. Jager, F.W. Dekker, C. Zoccali, Linear and logistic regression analysis, Kidney Int. 73 (2008), https://doi.org/10.1038/sj.ki.5002787.

[69] C. Fava, M. Sjögren, M. Montagnana, E. Danese, P. Almgren, G. Engström, P. Nilsson, B. Hedblad, G.C. Guidi, P. Minuz, O. Melander, Prediction of blood pressure changes over time and incidence of hypertension by a genetic risk score in swedes, Hypertension 61 (2013), https://doi.org/10.1161/HYPERTENSIONAHA.112.202655.

[70] Y. Chen, C. Wang, Y. Liu, Z. Yuan, W. Zhang, X. Li, Y. Yang, X. Sun, F. Xue, C. Zhang, Incident hypertension and its prediction model in a prospective northern urban han Chinese cohort study, J. Hum. Hypertens. 30 (2016), https://doi.org/10.1038/jhh.2016.23.

[71] N.I. Parikh, M.J. Pencina, T.J. Wang, E.J. Benjamin, K.J. Lanier, D. Levy, R. B. D'Agostino, W.B. Kannel, R.S. Vasan, A risk score for predicting near-term incidence of hypertension: the framingham heart study, Ann. Intern. Med. 148 (2008) 102–110, https://doi.org/10.7326/0003-4819-148-2-200801150-00005.

[72] M. Kivimäki, G.D. Batty, A. Singh-Manoux, J.E. Ferrie, A.G. Tabak, M. Jokela, M. G. Marmot, G.D. Smith, M.J. Shipley, Validating the framingham hypertension risk score: results from the whitehall II study, Hypertension 54 (2009), https://doi.org/10.1161/HYPERTENSIONAHA.109.132373.

[73] F. Koohi, E.W. Steyerberg, L. Cheraghi, A. Abdshah, F. Azizi, D. Khalili, Validation of the framingham hypertension risk score in a middle eastern population: tehran lipid and glucose study (TLGS), BMC Public Health 21 (2021), https://doi.org/10.1186/s12889-021-10760-6.

[74] S.M.S. Islam, A. Talukder, M.A. Awal, M.M.U. Siddiqui, M.M. Ahamad, B. Ahammed, L.B. Rawal, R. Alizadehsani, J. Abawajy, L. Laranjo, C.K. Chow, R. Maddison, Machine learning approaches for predicting hypertension and its associated factors using population-level data from three South Asian countries, Front. Cardiovasc. Med. 9 (2022), https://doi.org/10.3389/fcvm.2022.839379.

[75] X. Wu, X. Yuan, W. Wang, K. Liu, Y. Qin, X. Sun, W. Ma, Y. Zou, H. Zhang, X. Zhou, H. Wu, X. Jiang, J. Cai, W. Chang, S. Zhou, L. Song, Value of a machine learning approach for predicting clinical outcomes in young patients with hypertension, Hypertension 75 (2020) 1271–1278, https://doi.org/10.1161/HYPERTENSIONAHA.119.13404.

[76] H. Kanegae, K. Suzuki, K. Fukatani, T. Ito, N. Harada, K. Kario, Highly precise risk prediction model for new-onset hypertension using artificial intelligence techniques, J. Clin. Hypertens. 22 (2020), https://doi.org/10.1111/jch.13759.

[77] K. Kotowski, D. Kucharski, B. Machura, S. Adamski, B. Gutierrez Becker, A. Krason, L. Zarudzki, J. Tessier, J. Nalepa, Detecting liver cirrhosis in computed tomography scans using clinically-inspired and radiomic features, Comput. Biol. Med. 152 (2023) 106378, https://doi.org/10.1016/j.compbiomed.2022.106378.

[78] L. Koumakis, Deep learning models in genomics; are we there yet? Comput. Struct. Biotechnol. J. 18 (2020) 1466–1473, https://doi.org/10.1016/j.csbj.2020.06.017.

[79] J.Y.Y. Ng, J.H. Zhang, S.S. Hui, G. Jiang, F. Yau, J. Cheng, A.S. Ha, Development of a multi-wear-site, deep learning-based physical activity intensity classification algorithm using raw acceleration data, PLoS One 19 (2024) e0299295, https://doi.org/10.1371/journal.pone.0299295.

[80] Haibo He, E.A. Garcia, Learning from imbalanced data, IEEE Trans. Knowl. Data Eng. 21 (2009) 1263–1284, https://doi.org/10.1109/TKDE.2008.239.

[81] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, G. Bing, Learning from class-imbalanced data: review of methods and applications, Expert Syst. Appl. 73 (2017) 220–239, https://doi.org/10.1016/j.eswa.2016.12.035.
[82] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (2002), https://doi.org/10.1613/jair.953.
[83] Haibo He, Yang Bai, E.A. Garcia, Shutao Li, ADASYN: adaptive synthetic sampling approach for imbalanced learning, in: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), IEEE, 2008, pp. 1322–1328, https://doi.org/10.1109/IJCNN.2008.4633969.
[84] M. Martínez-García, G.O. Gutiérrez-Esparza, M.F. Márquez, L.M. Amezcua-Guerra, E. Hernández-Lemus, Machine learning analysis of emerging risk factors for early-onset hypertension in the Tlalpan 2020 cohort, Front. Cardiovasc. Med. 11 (2025), https://doi.org/10.3389/fcvm.2024.1434418.
[85] M.A. Nematollahi, S. Jahangiri, A. Asadollahi, M. Salimi, A. Dehghan, M. Mashayekh, M. Roshanzamir, G. Gholamabbas, R. Alizadehsani, M. Bazrafshan, H. Bazrafshan, H. Bazrafshan drissi, S.M. Shariful Islam, Body composition predicts hypertension using machine learning methods: a cohort study, Sci. Rep. 13 (2023) 6885, https://doi.org/10.1038/s41598-023-34127-6.
[86] R. van den Goorbergh, M. van Smeden, D. Timmerman, B. Van Calster, The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression, J. Am. Med. Inf. Assoc. 29 (2022) 1525–1534, https://doi.org/10.1093/jamia/ocac093.
[87] C. Yang, E.A. Fridgeirsson, J.A. Kors, J.M. Reps, P.R. Rijnbeek, Impact of random oversampling and random undersampling on the performance of prediction models developed using observational health data, J. Big Data 11 (2024) 7, https://doi.org/10.1186/s40537-023-00857-7.
[88] K. Berahmand, F. Daneshfar, E.S. Salehi, Y. Li, Y. Xu, Autoencoders and their applications in machine learning: a survey, Artif. Intell. Rev. 57 (2024) 28, https://doi.org/10.1007/s10462-023-10662-6.
[89] T. Chakraborty, U. Reddy K S, S.M. Naik, M. Panja, B. Manvitha, Ten years of generative adversarial nets (GANs): a survey of the state-of-the-art, Mach. Learn Sci. Technol. 5 (2024) 011001, https://doi.org/10.1088/2632-2153/ad1f77.
[90] V.-I. Tomescu, G. Czibula, Ş. Niţică, A study on using deep autoencoders for imbalanced binary classification, Procedia Comput. Sci. 192 (2021) 119–128, https://doi.org/10.1016/j.procs.2021.08.013.
[91] X. Xia, X. Pan, N. Li, X. He, L. Ma, X. Zhang, N. Ding, GAN-based anomaly detection: a review, Neurocomputing 493 (2022) 497–535, https://doi.org/10.1016/j.neucom.2021.12.093.
[92] A. Galdran, G. Carneiro, M.A. González Ballester, Balanced-MixUp for Highly Imbalanced Medical Image Classification, 2021, pp. 323–333, https://doi.org/10.1007/978-3-030-87240-3_31.
[93] G. Tsoumplekas, I. Siniosoglou, V. Argyriou, I.D. Moscholios, P. Sarigiannidis, Enhancing Performance for Highly Imbalanced Medical Data via Data Regularization in a Federated Learning Setting, 2024, pp. 302–315, https://doi.org/10.1007/978-3-031-67285-9_22.
[94] K. Araujo-Moura, L. Souza, T.A. de Oliveira, M.S. Rocha, A.C.F. De Moraes, A. Chiavegatto Filho, Prediction of hypertension in the pediatric population using machine learning and transfer learning: a multicentric analysis of the SAYCARE Study, Int. J. Publ. Health 70 (2025), https://doi.org/10.3389/ijph.2025.1607944.
[95] A.B. Bada, A.B. Garko, D. Gabi, Hypertension prediction using deep learning with transfer learning techniques, Fudma J. SCI. 8 (2024) 257–263, https://doi.org/10.33003/fjs-2024-0806-2855.
[96] Y. Liu, G. Yang, S. Qiao, M. Liu, L. Qu, N. Han, T. Wu, G. Yuan, T. Wu, Y. Peng, Imbalanced data classification: using transfer learning and active sampling, Eng. Appl. Artif. Intell. 117 (2023) 105621, https://doi.org/10.1016/j.engappai.2022.105621.
[97] B. Khemani, S. Patil, K. Kotecha, S. Tanwar, A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions, J. Big Data 11 (2024) 18, https://doi.org/10.1186/s40537-023-00876-4.
[98] P. Barbiero, R. Viñas Torné, P. Lió, Graph representation forecasting of patient's medical conditions: toward a digital twin, Front. Genet. 12 (2021), https://doi.org/10.3389/fgene.2021.652907.
[99] S. Kapoor, A. Narayanan, Leakage and the reproducibility crisis in machine-learning-based science, Patterns 4 (2023) 100804, https://doi.org/10.1016/j.patter.2023.100804.
[100] F.S. van Royen, F.W. Asselbergs, F. Alfonso, P. Vardas, M. van Smeden, Five critical quality criteria for artificial intelligence-based prediction models, Eur. Heart J. 44 (2023) 4831–4834, https://doi.org/10.1093/eurheartj/ehad727.
[101] N. Mohamed, J. Al-Jaroodi, I. Jawhar, N. Kesserwan, Leveraging digital twins for healthcare systems engineering, IEEE Access 11 (2023) 69841–69853, https://doi.org/10.1109/ACCESS.2023.3292119.
[102] MdD. Xames, T.G. Topcu, A systematic literature review of digital twin research for healthcare systems: research trends, gaps, and realization challenges, IEEE Access 12 (2024) 4099–4126, https://doi.org/10.1109/ACCESS.2023.3349379.
[103] S.M. Zayed, G.M. Attiya, A. El-Sayed, E.E.-D. Hemdan, A review study on digital twins with artificial intelligence and internet of things: concepts, opportunities, challenges, tools and future scope, Multimed. Tool. Appl. 82 (2023) 47081–47107, https://doi.org/10.1007/s11042-023-15611-7.