Document Version
Final published version

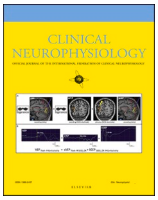
Citation (APA)
van der Geest, F. N., Tannemaat, M. R., Ye, F., Kononova, A. V., & van Rossum, I. A. (2025). Artificial intelligence for the analysis of head-upright tilt test data in syncope testing: a scoping review. *Clinical Neurophysiology*, *177*, Article 2110933. https://doi.org/10.1016/j.clinph.2025.2110933

Clinical Neurophysiology 177 (2025) 2110933

Contents lists available at ScienceDirect

Clinical Neurophysiology

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



Artificial intelligence for the analysis of head-upright tilt test data in syncope testing: a scoping review

F.N. van der Geest [a,b], M.R. Tannemaat [a], F. Ye [c], A.V. Kononova [c], I.A. van Rossum [a,*]

[a] Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands
[b] Educational Program Technical Medicine, Leiden University Medical Center, Delft University of Technology & Erasmus University Medical Center Rotterdam, the Netherlands
[c] Leiden Institute of Advanced Computer Science, Leiden, the Netherlands

ARTICLE INFO

Keywords:
Vasovagal syncope
Head-upright tilt test
Artificial intelligence
Machine learning

ABSTRACT

Objective: Syncope is defined as a sudden loss of consciousness due to cerebral hypoperfusion, with vasovagal syncope (VVS) being the most common form. The Head-Upright Tilt Test (HUTT) is the primary diagnostic tool but is time-consuming and has a suboptimal diagnostic yield. Machine Learning (ML) may improve early syncope prediction, thereby increasing diagnostic efficiency and reducing the burden on patients and healthcare professionals.

Methods: We searched PubMed for studies using ML on HUTT data for syncope testing. Extracted data included ML models, input features, performance metrics, preprocessing, and evaluation methods. Study quality was assessed using the STAR-ML checklist.

Results: Thirteen studies were included. Commonly used ML algorithms were support vector machines (SVM), neural networks, decision trees, k-nearest neighbor, and logistic regression. Features were derived from Electrocardiogram (ECG), continuous blood pressure (CBP), and transthoracic impedance (TIM). The highest-performing model used an SVM with features from ECG, CBP, and TIM.

Conclusions: ML integrated with HUTT signal analysis shows promise for improving diagnostic accuracy and efficiency. SVM models using multimodal features were particularly effective.

Significance: This review supports further development of ML-based tools to enhance diagnostic workflows in syncope care, especially for early VVS prediction.

1. Introduction

Syncope, a relatively common condition, is defined as a sudden loss of consciousness resulting from cerebral hypoperfusion. It is characterized by a rapid onset, short duration, and spontaneous complete recovery (Brignole et al., 2018a; Soteriades et al., 2002). The most common cause of syncope is vasovagal syncope (VVS). VVS can be triggered by specific circumstances such as fear, pain or prolonged standing. Typical symptoms that precede syncope are nausea, sweating and pallor (Brignole et al., 2018b).

The Head-Upright Tilt Test (HUTT) is the primary diagnostic tool for vasovagal syncope. The aim of HUTT is to elicit VVS. In subjects with VVS, blood pressure (BP) initially remains stable after tilt. At some point during the test, the reflex occurs, with both vasodepression and cardioinhibition leading to a sudden BP drop and syncope (van Dijk et al., 2021; van Rossum et al., 2023). Current HUTT protocols involve a 5–10 min stabilization phase in the supine position, followed by a passive phase of 10–20 min at a tilt angle of 60 degrees, then followed by another 10–20 min in the tilted position after sublingual administration of nitroglycerin (NTG) (Bartoletti et al., 2000; Benditt et al., 1996; Russo et al., 2023; Thijs et al., 2021). Using this protocol, sensitivity for reflex syncope is approximately 65 %, indicating that around one third of VVS patients will be tilted for 30–40 min and have a false negative result (Forleo et al., 2013).

Current analyses of HUTT data are often limited to the examination of BP, heartrate (HR), and sometimes stroke volume (SV) and total peripheral resistance (TPR). This approach disregards potentially crucial information within the dynamic interplay of these signals and significant temporal, structural, and spectral changes in physiological parameters that may occur before VVS becomes clinically apparent (Gilmore et al.,

---

* Corresponding author at: Leiden University Medical Center, Albinusdreef 2, 2300 RC Leiden, the Netherlands.
E-mail address: i.a.van_rossum@lumc.nl (I.A. van Rossum).

https://doi.org/10.1016/j.clinph.2025.2110933
Accepted 6 July 2025
Available online 9 July 2025

2021; Khodor et al., 2016). Exploring information possibly present within these parameters at early time points holds promise both for early prediction and for a deeper understanding of the pathophysiological mechanisms behind VVS.

In the era of digitization, the analysis of large amounts of complex data is becoming more common in healthcare (Rajpurkar et al., 2022; Wiens and Shenoy, 2018). Artificial intelligence (AI) including methods like machine learning (ML) are increasingly utilized successfully to improve the diagnostic yield of various tests (Ghorbani et al., 2020; Rajpurkar et al., 2022; Tannemaat et al., 2023). ML could give more insight into the underlying mechanism and prediction of syncope and optimize diagnostic yield.

This study aims to provide an overview of the different ways in which ML has been used to analyze HUTT data, compare different methods and features, evaluate the models' performance in the current literature and propose recommendations for further research.

2. Methods

2.1. Search strategy

We performed a Pubmed search on January 1, 2024. The search term covered two main subjects: the Head-Upright Tilt Test, including assessments for vasovagal syncope, and machine learning. For each of the elements, several terms were added. If available, specific terms provided by the NCBI MeSH database were applied. When MeSH terms were not available, general terms were combined with the term "[tiab]" to confine the search to titles and abstracts only. The full search term can be found in Appendix A. References from selected articles were also evaluated for eligibility.

2.2. Study selection

Articles underwent an initial screening based on their title and abstract, and their final eligibility for inclusion was determined through a comprehensive evaluation of the full text. Articles were selected based on the following inclusion criteria: (i) studies describing individuals who underwent HUTT with data measurements; (ii) studies using ML on the measured data; (iii) HUTT conducted explicitly for syncope testing.

Articles that met the following criteria were excluded: (i) articles without an English version; (ii) non-human evaluations (animal or model); (iii) studies involving only Postural Orthostatic Tachycardia Syndrome (POTS) patients; (iv) studies that only included children.

2.3. Quality assessment

The 'Screening Tool for Assessing Reporting of Machine Learning' (STAR-ML) checklist is specifically designed to evaluate the quality of ML reporting in research articles, focusing on aspects related to reproducibility and correctness (Koh et al., 2023). The STAR-ML version 2 checklist focuses on two key aspects: the quality of the data, assessed through four questions, and the quality of the algorithm, assessed through five questions. Each question in the checklist is assigned a score, in which a "yes" is notated as 1, and a "no" is notated as 0. The maximum achievable score is 9. The overall quality of the studies is categorized as low (0–4 points), medium (5–6 points), or high (7–9 points) based on the total score.

To assign scores to the questions, we made the following assumptions. If the features used were indicated, but it was unclear which features were ultimately used in the final model, a score of 0 was assigned. Additionally, if questions had at least one way of indicating consideration (for example, if one parameter was given while more parameters were used), a score of 1 was given. Regarding data distribution, it was assumed that if the distribution within the subject group was described beyond the number of positive or negative HUTT outcomes, a score of 1 was given.

2.4. Data extraction

From the included articles, information was extracted on: (i) the goal of using an ML model and which data was most useful; (ii) which ML models were used; (iii) which features were used; (iv) performance of the models as determined by their diagnostic yield (sensitivity, specificity, positive predictive value and negative predictive value). Secondary outcomes included: (i) pre-processing methods; (ii) evaluation methods.

Due to the wide range of features that can be utilized, we created an overview based on the predominant features and measurement methods used for data collection.

Diagnostic yield was determined based on the highest performance obtained per study. To make the performances comparable between the different studies, missing values were calculated based on the True Positive, True Negative, False Negative and True Negative.

3. Results

3.1. Article selection

A total of 31 articles were identified in the literature search: seventeen from the primary query and fourteen from the reference list of these articles After applying the predetermined inclusion and exclusion criteria, thirteen articles were included in this literature review based on the search strategy (Fig. 1).

3.2. Quality assessment

Completed START-ML questionnaires for all articles can be found in Appendix B. Table 1 shows an overview of the score of each study. Out of the thirteen articles, eight were categorized as high quality, two as medium quality, and three as low quality. Three articles (Ciliberti et al., 2018; Fortrat et al., 2007; Klemenc and Štrumbelj, 2015) received a low-quality score, because information was missing on both the selected data and the algorithm, negatively affecting reproducibility. Four articles (Couceiro et al., 2016; He et al., 2021; Schang et al., 2007, 2006) obtained the highest scores with a total of 8 points.

The absence of data normalization during pre-processing in six out of thirteen articles was notable.(Ciliberti et al., 2018; Fortrat et al., 2007; Shahadat Hussain et al., 2021; Khodor et al., 2014; Klemenc and Štrumbelj, 2015; Myrovali et al., 2021) The checklist contains several items pertaining to pre-processing steps, reflecting the significant impact normalization can have on model performance.(Koh et al., 2023) Another notable point is that only three articles (He et al., 2021; Schang et al., 2007, 2006) addressed the consideration of bias in their models.

Seven articles (Ciliberti et al., 2018; Fortrat et al., 2007; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014; Klemenc and Štrumbelj, 2015) did describe features and parameters used but were not complete in specifying the information used in the final model.

3.3. Study characteristics

3.3.1. Study population

The number of subjects used in the studies varied considerably between articles (Table 1). The median number of subjects was 86 [IQR: 129]. Notably, two studies (Shahadat Hussain et al., 2021; S. Hussain et al., 2021) included a very large number of subjects, with a total of 687 participants, consisting of 92 subjects with a positive HUTT result and 592 subjects with a negative HUTT result. The remaining eleven articles had a more balanced distribution, with approximately half of all subjects having a positive HUTT result. In addition, there was variability in the types of subjects who underwent HUTT testing across studies, with the most common focus on patients with suspected syncope or a history of syncope. One study (Myrovali et al., 2021) also included a comparison

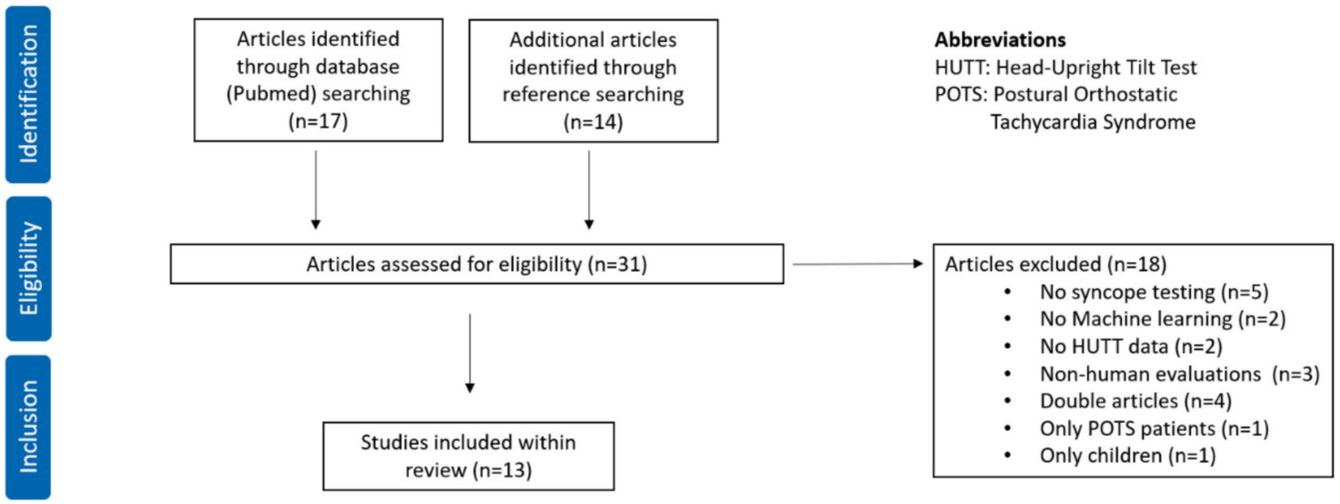


Fig. 1. Flow diagram of the study identification, exclusion and inclusion.

Table 1
Overview of the primary results and quality assessment of included articles.

First Author (Year)	Group size	Population	ML algorithm	Feature measurements	STAR-ML score
Classification in disease subgroups					
(Gilmore et al., 2021)	186 (97+, 89 −)	suspected syncope	RDF and K-means clustering	ECG and CBP	7
Prediction of syncope					
(Ciliberti et al., 2018)	26 (8 +, 18-)	suspected syncope	Logistic Regression with multivariable analysis	ECG	3
(Couceiro et al., 2016)	43 (21+,22–)	suspected syncope	Different distance metrics	ECG and CBP	8
(Fortrat et al., 2007)	86 (45 +, 41 −)	suspected syncope	NN	ECG, CBP, TIM, and other.	4
(He et al., 2021)	203(128+, 75-)	suspected syncope	SVR, LR, kNN, RF	ECG and CBP	8
(S. Hussain et al., 2021)	687 (92+, 592 −)	suspected syncope	DT, GNB, kNN, MNB, SVM and LR	ECG, CBP, and TIM	7
(Shahadat Hussain et al., 2021)	687 (92+, 592 −)	suspected syncope	SVM, kNN and SGD	ECG, CBP and TIM	6
(Khodor et al., 2014)	66 (35+, 31-)	healthy	KSVM	ECG and CBP	6
(Khodor et al., 2016)	57 (28+, 29-)	healthy	kNN and KSVM (Kernel: radial basis function)	ECG and CBP	7
(Klemenc and Štrumbelj, 2015)	92 (42 + 50-)	suspected syncope	Linear regression	ECG and CBP	5
(Myrovali et al., 2021)	36 (11 +, 15 −, 10c)	suspected syncope and healthy	MLP-NN	ECG and other	7
(Schang et al., 2006)	Retrospective:70 (33 −, 37 +), Prospective:59 (33 −, 26 +)	suspected syncope	NN and PCA	ECG, TIM and other	8
(Schang et al., 2007)	128 (65 −, 63 +),	suspected syncope	KSVM (kernels: linear kernel, polynomial kernel, sigmoidal neural network kernel and gaussian radial basis kernel)	ECG, TIM and other	8

+: number of subjects with a positive HUTT test, −: number of subjects with a negative HUTT test, c: healthy control.
NN: Neural Network, RDF: Random Decision Forest, DT: Decision Tree, SVR: Support vector regression, GNB: Gaussian Naïve Bayes, kNN: k-Nearest Neighbor, MNB: Multinomial Naïve Bayes, SVM: Support Vector Machine, KSVM: Kernel Support Vector Machine LR: Logistic Regression, SGD: Stochastic Gradient Descent, MLP-NN: Multilayer perceptron neural network, PCA: Principal Component Analysis.
ECG: Electrocardiogram, CBP: Continuous blood pressure measurement, TIM: Transthoracic impedance measurement,

with healthy subjects, providing a broader perspective on the evaluation of syncope.

3.3.2. HUTT protocol

Syncope testing protocols across the studies shared a common starting point, initiating the examination in the supine position for several minutes. Subsequently, the patient was promptly tilted to an angle between 60–80 degrees, and the table was returned to the supine position either after a predetermined time or when the patient exhibited signs of pre-syncope or syncope. In most protocols, the patient remained in the supine position for intervals of 5–15 min. However, in one study (Fortrat et al., 2007), this position was maintained for 50 min. In addition, in four studies (Couceiro et al., 2016; Gilmore et al., 2021; Klemenc and Štrumbelj, 2015; Myrovali et al., 2021), the patient was administered NTG, isoproterenol or glyceryl trinitrate when they exhibited no (pre-)syncope symptoms after 20–30 min, and maintained in the tilted position for another 10–15 min. In the remaining studies, pharmacological agents were not mentioned, and the tilted position was maintained for 45 min or less. A schematic overview of the HUTT can be found in Fig. 2.

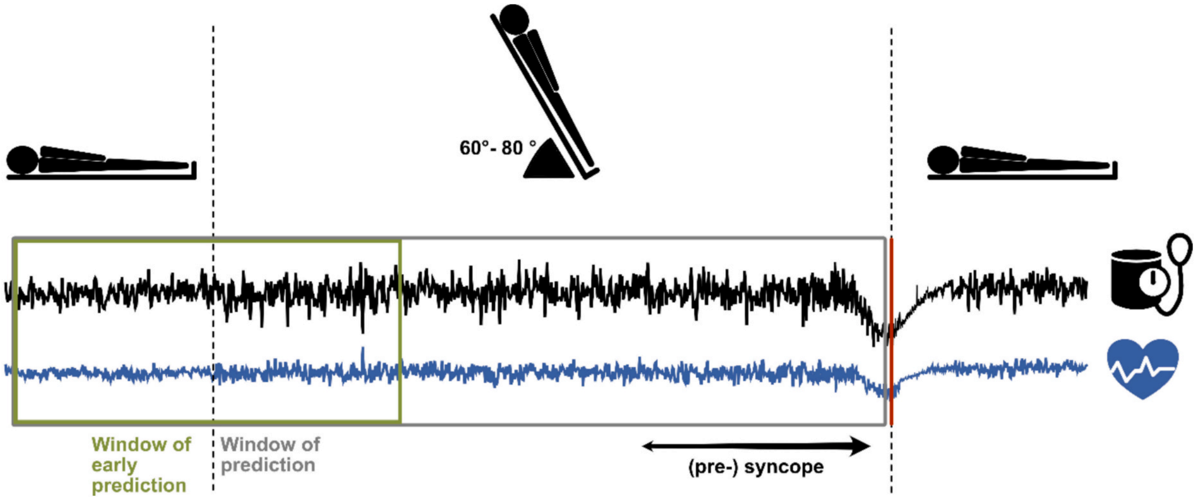


Fig. 2. Schematic overview of HUTT with start of (pre-) syncope and the (early) prediction window.

3.4. Primary results

3.4.1. The aim of ML and the data selection

ML was used for two distinct purposes in the analyzed studies: 1) to categorize patients into disease subgroups, including POTS, cardioinhibition, vasodepression, and mixed syncope(Gilmore et al., 2021) and 2) to predict early-stage syncope during the HUTT(Ciliberti et al., 2018; Couceiro et al., 2016; Fortrat et al., 2007; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014; Klemenc and Štrumbelj , 2015; Myrovali et al., 2021; Schang et al., 2007, 2006). A further distinction can be made between early syncope prediction and syncope prediction at a later time during the HUTT. It is notable that the majority of articles focused on shortening the duration of HUTT rather than classification in disease subgroups. Two studies (Khodor et al., 2016, 2014) compared different time intervals, including the first 5 supine minutes, the first 5 min after tilting, and the last 5 min to predict syncope. These articles showed no significant difference in sensitivity and specificity between these intervals. However, sensitivity increased from a maximum of 69.6 % to 87.5 % and the specificity from a maximum of 66.9 % to 93.8 % when the first 15 min after tilt were used compared to a 5-minute period (Khodor et al., 2016).

Another study (Gilmore et al., 2021) focused on optimizing data selection by using variable epochs of 4 min between tilt up and tilt down for classification into the correct disease subgroup.

3.4.2. Machine learning models

Both Neural Networks (NN) and classic ML methods have been used for the prediction and classification of syncope using HUTT data (Table 1).

Neural networks

NNs were studies in three articles (Fortrat et al., 2007; Myrovali et al., 2021; Schang et al., 2006). One of these articles used a multilayer perceptron neural network (MLP-NN), which is described as a nonlinear function with three layers, including an input layer that receives the inputs, a hidden layer that processes the input information, and an output layer that produces the classification results (Myrovali et al., 2021).

Classic machine learning models

Commonly used classic ML classifiers were Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR), and k-Nearest Neighbors (kNN).(Gilmore et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014; Schang et al., 2007) These classifiers were included in this paper due to their common usage or demonstrated good performance. While there were other algorithms used for comparison purposes, the following text focuses on these commonly used classifiers.

Support vector machine

The SVM is a classifier that performs classification by creating a hyperplane in a higher dimensional space. SVM can be divided into linear and nonlinear models and works by mapping data into a feature space in which each point in the feature space contains information about a data point. This feature space is created using a kernel function. The proper choice of a kernel can make inseparable classes separable by transforming data and creating the optimal decision boundary by maximizing the separation margin between classes.(Suthaharan, 2016) The most commonly used kernels are: linear kernel, polynomial kernel, sigmoid kernel, and Gaussian radial basis kernel.(Montesinos López et al., 2022; Schang et al., 2007) A SVM is the most commonly used method in the articles and was used in five of the six articles that used classic ML methods(Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014; Schang et al., 2007). One article (He et al., 2021) used Support Vector Regression (SVR), which applies the principle of SVM to perform regression tasks (Montesinos López et al., 2022).

Decision tree and random forest

The DT is a hierarchical classification technique that divides data into subsets, providing information to separate classes, and can be defined as a supervised learning model. It consists of internal nodes representing features, leaf nodes indicating outcomes, and branches dictating decision parameters. The computation of the decision parameters, including a feature and its split location, is determined by maximizing the information gain or other criteria like the Gini impurity. (Suthaharan, 2016) One article (S. Hussain et al., 2021) used a DT, and another article(Gilmore et al., 2021) used a combination of DTs, also known as Random Forest (RF). The advantage of RF is that a large number of DTs can be used, with each tree making a prediction and the algorithm selecting the best prediction through a majority voting (Gilmore et al., 2021).

k-Nearest neighbors

The kNN classifier assigns unlabeled data to the class of the most similar labeled examples. It relies on the distance to the nearest neighbor for classification, with the parameter k determining the number of neighboring data points included in the class prediction.(Zhang, 2016) kNN was used in three articles(Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016).

Logistic regression

A LR model is a statistical model estimating the probability of an event based on individual characteristics. It is a supervised learning technique that utilizes a logistic function or sigmoid function to map predictions to probabilities.(Sperandei, 2014) This method was used in one article(S. Hussain et al., 2021).

3.4.3. Features

Measurements

Features can be derived from parameters obtained through different types of measurements. These measurements can be divided into the Electrocardiogram (ECG), Continuous Blood Pressure (CBP), Transthoracic Impedance Measurement (TIM), and other measurements. An overview of the measurements per article can be found in Table 1. It is notable that most studies extracted parameters from at least two different measurement methods.

The most frequently used parameters related to these measurements are discussed below. Table 2 provides an overview of the different types of parameters that were used in at least two articles.

Electrocardiogram

ECG was consistently used in all included studies, measuring parameters related to cardiac activity. These include heart rate (HR), heart rate variability (HRV), and various intervals such as RR-interval, QT-interval, PR-segment, and QRS-duration. In addition, heart rate can also be used for other calculations such as cardiac output (CO).

Continuous blood pressure

The CBP measurement is the second most prevalent measurement tool and was used in nine other articles.(Couceiro et al., 2016; Fortrat et al., 2007; Gilmore et al., 2021; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014; Klemenc and Štrumbelj, 2015) There are several devices that obtain this non-invasive measurement using a finger sensor. Systolic BP, diastolic BP, mean BP, and Pulse pressure can be derived from this measurement.

Table 2

An overview of the frequently measured parameters.

Parameters	Measurement
HR	ECG(Couceiro et al., 2016; Fortrat et al., 2007; Gilmore et al., 2021; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Schang et al., 2007, 2006)
RR-interval (HRV)	ECG(Ciliberti et al., 2018; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014; Klemenc and Štrumbelj, 2015; Myrovali et al., 2021)
Systolic BP	CBP(Fortrat et al., 2007; Gilmore et al., 2021; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014)
Diastolic BP	CBP(Fortrat et al., 2007; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021)
Mean BP	CBP(Fortrat et al., 2007; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021)
PTT	ECG(Couceiro et al., 2016; Khodor et al., 2016), CBP(Couceiro et al., 2016; Khodor et al., 2016)
Baroreflex sensitivity	ECG(Khodor et al., 2014; Klemenc and Štrumbelj, 2015), CBP(Khodor et al., 2014; Klemenc and Štrumbelj, 2015)
Cardiac output	CBP(He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021) ECG(Shahadat Hussain et al., 2021; S. Hussain et al., 2021), TIM(Shahadat Hussain et al., 2021; S. Hussain et al., 2021)
Contractability index	TIM(Fortrat et al., 2007; Shahadat Hussain et al., 2021; S. Hussain et al., 2021)
Stroke volume	CBP(He et al., 2021), TIM(Fortrat et al., 2007; Shahadat Hussain et al., 2021; S. Hussain et al., 2021)
TPR	CBP,(He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021) TIM(Shahadat Hussain et al., 2021; S. Hussain et al., 2021)
LVET	CBP(Couceiro et al., 2016; He et al., 2021), TIM (Fortrat et al., 2007; Schang et al., 2007, 2006)
Impedance waveform (IW)	TIM(Fortrat et al., 2007; Schang et al., 2007, 2006)
IW AUC	TIM(Schang et al., 2007, 2006)
IW Slopes	TIM(Schang et al., 2007, 2006)
Age	Other(Fortrat et al., 2007; Schang et al., 2007, 2006)
Sex	Other(Fortrat et al., 2007; Schang et al., 2007, 2006)

ECG: Electrocardiogram, TIM: Transthoracic impedance measurement, CBP: Continuous blood pressure measurementHR: Heart rate, HRV: Heart rate variability, BP: Blood pressure, PTT: Pulse transit time, TPR: total peripheral resistance, LVET: Left ventricular ejection time, AUC: Area under the curve.

Transthoracic impedance measurement

TIM includes impedance cardiography (ICG) and transthoracic impedance (TTI) measurements indicating the resistance in the thorax through the flow of current. Parameters include the impedance waveform slope and area under the curve, but also hemodynamic parameters like stroke volume (SV), cardiac output (CO), and contractility index (CI). TIM was used in five articles (Fortrat et al., 2007; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Schang et al., 2007, 2006).

Other measurement

Other measurements include the age, sex, and sometimes blood values (e.g., hematocrit and hemoglobin) and body measurements (e.g., fat mass and body surface area) as parameters.

Calculations and Transformations

These parameters can be used as features themselves, but they can also be used to derive additional secondary features. These secondary features include different minima and maxima, averages, intervals, amplitudes, and derivatives. In addition, multiple parameters can be used to calculate a new variable such as the stroke index (SI), calculated by dividing the SV by the body surface area (BSA) and multiplying it by *1000 (Shahadat Hussain et al., 2021). Parameters and their intervals can also be converted to the frequency-time domain using the wavelet function (Khodor et al., 2016; Myrovali et al., 2021) or to the frequency domain (Shahadat Hussain et al., 2021; Khodor et al., 2014). In the frequency domain, the power spectral density (PSD) can also be calculated, and frequencies can be divided into low and high frequencies. Examples are the high-frequency RR-interval or the low-frequency diastolic blood pressure (Shahadat Hussain et al., 2021).

Features can be determined automatically in a neural network that extracts the needed information (Fortrat et al., 2007). However, when using a classic ML classifier, feature selection methods are used to select relevant features.

Feature selection methods

The inclusion of characteristics can be based on statistical significance, in which features are included only if a significant difference can be verified or when it improves the classification accuracy rate (ACC) (Myrovali et al., 2021).

In one study (Khodor et al., 2016), three different feature selection methods were compared. The first method is the Relief method in which feature weights are iteratively computed based on their ability to distinguish between neighboring models. The second method is the Sequential Forward Selection (SFS) in which features are included one at a time in a candidate set until the inclusion of additional features no longer improves the specific termination criterion. The third method is the Probe feature algorithm in which classification is based on the correlation coefficient. The selection of relevant features is performed using a stopping criterion, which sets a threshold for the inclusion of features, ensuring that each added variable contributes more than a random variable. The model that achieved the highest performance in the final analysis used the SFS feature selection method (Khodor et al., 2016).

3.4.4. Performance

A summary of the performance and the method used to obtain those values is shown in Table 3. This table does not include the single study that dealt with a different research question: the accurate classification of the disease subgroup, for which the study reached a minimum value of 93 % correct classification per disease group (Gilmore et al., 2021). Another study was not included because it provided only the accuracy of the model (Klemenc and Štrumbelj, 2015).

3.5. Secondary results

3.5.1. Pre-processing

The reviewed articles offer limited insight into the use of pre-

Table 3
The sensitivity, specificity, positive predictive value, negative predictive value and ML model of the included articles testing the predictability of syncope during HUTT.

First author (year)	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Method
(Ciliberti et al., 2018)	87.5 %	72.2 %	75 %	89 %	LR with multivariable analysis, using ECG and CBP features from the supine position
(Couceiro et al., 2016)	95.2 %	95.4 %	90.9 %	89 %	Minkowski distance with ECG and CBP features from 2 min after tilt as reference and the time after during the tilted position, LOOCV results
(Fortrat et al., 2007)	76 %	81 %	78 %	80 %	NN, using a combination of ECG and other measurement features from different moments during supine position,
(He et al., 2021)	86 %	82 %	89 %	78 %	SVR, using ECG and CBP features from 3 min supine position, 15 min tilted position
(S. Hussain et al., 2021)	97.61 %	98.74 %	92.23 %	99.63 %	LR, using ECG, CBP and TIM features from the whole test
(Shahadat Hussain et al., 2021)	99,10 %	99.73 %	98.23 %,	99.86 %	SVM, using ECG, CBP and TIM features from the whole test
(Khodor et al., 2014)	88.5 %	80.6 %	83.7 %	86.1 %	SVM, using ECG and CBP features from first 15 min of the tilting position, LOOCV results
(Khodor et al., 2016)	87.5 %	93.8 %.	93.2 %	88.6 %	SVM, using feature selection method SFS on ECG and CBP features from the first 15 min of tilting position
(Myrovali et al., 2021)	82.39 %	91.78 %	92.30 %	92.57 %.	MLP-NN, using ECG and other measurement features from the 5 min before tilt and 5 min after tilt
(Schang et al., 2006)	88 %	64 %	66 %	88 %	NN, using ECG, TIM and other measurement features from the 10 min supine position tested on prospective subjects
(Schang et al., 2007)	94 %	79 %	81 %	93 %	SVM, gaussian kernel, using ECG, TIM and other measurement features from the 10 min supine position

NN: Neural Network, SVM: Support Vector Machine, SVR: Support Vector Regression, LR: Logistic Regression, MLP-NN: Multilayer perceptron neural network, SFS: Sequential feature selection, min: minutes, LOOCV: Leave One Out Cross-Validation
ECG: Electrocardiogram, TIM: Transthoracic impedance measurement, CBP: Continuous blood pressure measurement.

processing methods. However, two articles (Couceiro et al., 2016; Khodor et al., 2016) specifically mentioned their use of a Butterworth filter, such as a high-pass 0.01 Hz filter. The main practice across the other studies involves normalization and re-sampling the data, coupled with the application of various algorithms for feature extraction. This includes the Pan-Tompkins algorithm and the Morlet wavelet function (Myrovali et al., 2021).

Additionally, in one study missing values were replaced with the average of the respective indicators for all patients, accompanied by min–max normalization (S. Hussain et al., 2021).

Furthermore, two studies (Shahadat Hussain et al., 2021; S. Hussain et al., 2021) encountered a data imbalance. Out of the 687 subjects, only 92 subjects had a positive HUTT result. To address this data imbalance, Synthetic Minority Oversampling Techniques (SMOTE) was used. SMOTE creates artificial instances of the minority class to address this data imbalance (Khushi et al., 2021).

3.5.2. *Evaluation method*

The most commonly used evaluation method was K-fold cross-validation. Six articles (Gilmore et al., 2021; He et al., 2021; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2014; Myrovali et al., 2021) used this method for validation with different values of K. In some cases, K was determined by experimentation (Shahadat Hussain et al., 2021), while in other cases, the number was determined by the approximation of the classification rate (Gilmore et al., 2021). Another K-fold cross-validation method used was K-1-fold cross-validation, in which the remaining set is used as the evaluation set (S. Hussain et al., 2021). This is different from the leave-one-out cross-validation, in which the ML model is trained on all but one data point, iteratively leaving out each point to evaluate the model's generalization performance, as was used in four other instances (Couceiro et al., 2016; Gilmore et al., 2021; Khodor et al., 2014; Klemenc and Štrumbelj, 2015).

Another way to divide the data into a train and a test set is the train-test-split method, in which a predetermined percentage of the number of subjects is used as the training set, and the remaining percentage of subjects is then used as the test set for evaluation. This method was used in three articles (Shahadat Hussain et al., 2021; Khodor et al., 2016; Schang et al., 2007).

One study (Schang et al., 2006) used a prospective validation group to determine the performance of the algorithm.

4. Discussion

In this review, we provide an overview of the current literature on ML for HUTT, with a specific emphasis on data selection, feature extraction and accuracy. ML has been used in multiple studies for the classification of syncope using HUTT test data, with the main focus on early prediction of syncope. The optimal selection of data varied among the studies. Longer epochs (i.e. 15 min instead of 5) appear to improve predictive accuracy (Khodor et al., 2016, 2014). Many different ML approaches were used, with SVM emerging as the most widely used method. The most frequently used parameters were extracted from ECG and CBP, which were often used to derive features such as means and maxima or values from the frequency spectrum. The study with the highest performance used an SVM model with a diverse set of features that could be measured through ECG, CBP and TIM (Shahadat Hussain et al., 2021).

Data selection

The integration of ML with signal analysis of HUTT data opens up new possibilities for the early prediction and accurate classification of syncope by combining traditional medical assessment with ML classification techniques. Data selection differed between studies: longer epochs appear to improve prediction. However, this does defeat the purpose of shortening the required duration of HUTT. Furthermore, it was unclear whether subjects experiencing syncope during the 15 min in the tilted position were excluded. If these patients were not excluded, the obtained ML model should be considered a diagnostic rather than a predictive model, and this may explain the better performance of the 15 min compared to the 5 min data that was selected (Klemenc and Štrumbelj, 2015). Although 15 min after tilt has been defined several times as early prediction, true early prediction can only occur in the first minutes before or after tilt, well before the occurrence of syncope. The comparison of data between the supine and tilted positions within a subject may also provide extra information, as significant differences in the autonomic response between the two positions can be found early after tilting.(Vybiral et al., 1989) This is consistent with the results found by (Khodor et al., 2014) and warrants further investigation.

Machine learning algorithms

Various ML algorithms have been used to enable syncope classification, each with its advantages and disadvantages. SVM was used most frequently because of several advantages: The risk of overfitting is low, it performs well with clear separation classes, and outliers have little influence, it works In the nonlinear case via the usage of appropriate kernels, and it can operate with high-dimensional data. SVM was indeed used in the two best-performing models(Shahadat Hussain et al., 2021; S. Hussain et al., 2021), although both articles used the same imbalanced subject group. Investigation of the different ways of dealing with this imbalance in the data would have been useful additional information to be investigated in relation to the number of variations (Kong et al., 2020). However, SMOTE was the method of choice for this study. While SMOTE is commonly effective in addressing data imbalances, its application in both articles might have contributed to the observed high performance, given the similarity in datasets.(Khushi et al., 2021) Nonetheless, SVM appears to be a promising approach, as it has not only been extensively used but has also outperformed various other methods in comparative studies (Shahadat Hussain et al., 2021; Khodor et al., 2016).

Although SVM was the most frequently reviewed approach in the studies, only three studies included a neural network-based approach. However, limited details were provided on the architectures used, making it difficult to assess the quality of their implementation or compare performance effectively. Furthermore, the included studies did not explore other deep learning models such as CNNs or LSTMs, which could potentially model HUTT data more effectively. This absence may be due to the small sample sizes typically required for the effective training of deep neural networks. The lack of deep learning models reflects the relatively early stage at which deep learning methods are being applied in this specific clinical context. Future studies with larger datasets could help determine whether deep learning approaches offer significant advantages over traditional ML techniques for syncope classification.

Features

Each article used different combinations of features derived from the measured parameters, making direct comparisons between their relative value challenging. Nevertheless, several comparisons were made between the types of features within the articles. For instance, one article (Fortrat et al., 2007) found that incorporating both cardiovascular and body composition parameters produced superior results compared to using either alone. Notably, most articles used a combination of features from at least two measurement methods to achieve the highest scores, showing that combining different types of features is likely to contribute to higher scores. Relying on a single measurement method may not be enough to determine the interaction between the different reflex components (Khodor et al., 2016).

Pre-processing and evaluation

The documentation of pre-processing including the steps for the handling of missing data was not consistently provided in every report, which limits their reproducibility and comparability across studies. This lack of standardization complicates the interpretation of performance metrics. Nonetheless, some articles provided a general description of the steps involved. One article (S. Hussain et al., 2021) showed some of the pre-processing steps encompassing commonly used methods, including the cleaning of the data such as handling missing values, and transformation of the data including normalization (Frye et al., 2021). Across the included studies, various methods were used to evaluate the performance. One study (Schang et al., 2006) used a prospective validation group, which provides a more realistic reflection of how the algorithm would perform in a clinical setting compared to studies that only used retrospective data. The remaining studies used both a retrospective train set and a test set, with more than half of the articles using k-fold cross-validation. One of the studies (Khodor et al., 2014) specifically mentioned LOOCV, a technique that can be used to prevent overfitting in a smaller subject group. However, the diverse and missing information regarding preprocessing and validation approaches in different studies highlights the need for more standardized methodology and reporting practices. Future research should focus on the adoption and detailed documentation of consistent preprocessing methods to improve the reliability and clinical relevance of machine-learning applications in syncope testing.

Study characteristics

The variation in subject group sizes is an important aspect to consider when evaluating the performance of the various models and features. Large group sizes are likely to improve performance, but the studies involving the largest group of subjects (n = 687) also showed the greatest data imbalance (Shahadat Hussain et al., 2021; S. Hussain et al., 2021), which may contribute to an overestimation of diagnostic accuracy. Most other studies in this review involved data from fewer than 100 subjects, which is likely on the low side for the application of advanced ML techniques(Fortrat et al., 2007; Khodor et al., 2016, 2014; Myrovali et al., 2021; Schang et al., 2006). Going forward, we recommend the use of larger and more balanced groups.

An additional consideration is the diversity in included subjects. Some research used healthy individuals undergoing HUTT (Khodor et al., 2016, 2014). Physiological responses and the perception of presyncopal symptoms in subjects with a clinical suspicion or history of syncope may be different compared to completely healthy subjects. The impact on syncope classification of these patient characteristics requires further investigation.

The final aspect to consider is the variation between the HUTT protocols across studies. Generally, the same sequence from the supine position to a tilt position between 60–80 degrees was maintained. Nonetheless, there were notable differences in how long a particular position was held, ranging from 10 to 50 min. Similarly, the duration a subject remained in a tilted position varied from 20 to 45 min, or until (pre-)syncope occurred. Shorter tilt times are likely to lead to a higher number of subjects with negative HUTT outcomes.

Limitations.

Five articles had a score below 7 points on the quality assessment (Ciliberti et al., 2018; Fortrat et al., 2007; S. Hussain et al., 2021; Khodor et al., 2014; Klemenc and Štrumbelj, 2015). The STAR-ML guide recommends the exclusion of articles with a score is below 7. These articles were included nonetheless to get a broader picture of the ML methods and features that can be used for the analysis of HUTT data. In future research, it is important that the used data and algorithm are properly recorded so that the quality of the articles increases in accordance with the STAR-ML guide.

Another limitation of this review is the fact that only one database (Pubmed) was searched. This may have led to the omission of relevant studies. We aimed to mitigate this limitation by also searching the references of selected articles. We included multiple pairs from the same research groups, in some cases even describing the same study population, which may have caused some bias.(Fortrat et al., 2007; Shahadat Hussain et al., 2021; S. Hussain et al., 2021; Khodor et al., 2016, 2014; Schang et al., 2007, 2006).

Clinical implications

Symptom recognition during tilt-induced VVS remains crucial for accurate diagnosis, and current HUTT protocols, based on real-time physician observation, continue to provide indispensable clinical value. The use of ML data cannot fully replace the current tilt protocols for VVS. However, further exploration of the use of ML for early prediction of syncope during tilt may aid in improving syncope care in three ways. Firstly, in approximately a third of subjects with a clinical diagnosis of VVS, HUTT does not lead to syncope, even after administration of NTG. Identifying those subjects during the first minutes of HUTT can save time for both patients and health care professionals. Secondly, the use of ML for classification of syncope during tilt may improve the understanding of hemodynamic changes prior to the actual syncope and thereby help to further unravel the pathophysiology of VVS. Thirdly, prediction of VVS during the early stages of HUTT can increase diagnostic accuracy in patients whose test must be stopped prematurely (e.

g., due to feeling unwell). Nonetheless, the potential benefits must be considered alongside the limitations of the integration of ML models.

Integration challenges

Integrating ML into healthcare can present various challenges that must be addressed for this technology to be successfully implemented (Assadi et al., 2022). Several factors must also be considered when integrating an ML model into HUTT workflows to ensure a smooth workflow. Firstly, variability in patient responses during tilt testing significantly limits the generalizability of models. Various factors, such as cardiovascular health, emotional state, hydration status and environmental conditions, could influence physiological responses during testing (Chu et al., 2015; Russo et al., 2017; Shinohara et al., 2014). These differences, both between and within patients, complicate the training and validation of ML models designed for real-time prediction.

Secondly, it is important to distinguish between ML applications that aim to predict long-term outcomes or treatment responses, and those that attempt to predict short-term physiological events, as in the case during HUTT. The latter requires a more nuanced understanding of individual patient responses in real time, which, as described above, can vary widely. The clinical application of ML for predicting responses during HUTT will need to be rigorously validated in order to account for context sensitive interpretation of the data.

Thirdly, implementation of these models in practice requires acceptance by healthcare professionals, as well as seamless integration into existing clinical workflows (Chen et al., 2022). Healthcare professionals must be able to interpret ML predictions and act on them without creating new diagnostic uncertainties or additional work. Until this is possible and the models have been proven reliable for all patient subgroups, ML-based predictions will be considered an adjunct to current practice rather than a replacement.

Overall, clinical decision-making in syncope diagnosis still requires nuanced interpretation of physiological signals in the right context. Therefore, future studies should take into account these different factors and create models that can interpret real-time inputs in the appropriate context.

5. Conclusion and recommendations for future research

The integration of ML with signal analysis of HUTT data offers new possibilities for syncope prediction. The use of ML can help to gain insight in the pathophysiology of tilt induced VVS and may eventually lead to a decreased test duration and improved accuracy of HUTT in a selection of subjects. SVM appears to be an effective prediction method based on the performance of multiple studies. Using features of at least two measurement methods better reflects the interaction between different physiological parameters corresponding to the complex mechanism of syncope. However, challenges such as variations in subject group size, patient characteristics, and HUTT protocols require careful interpretation of the results. Standardization of methodologies, improvement of predictive accuracy and further validation in prospective studies are crucial next steps before adopting ML for syncope prediction in a clinical setting.

Appendix A. Search term

((“Syncope, Vasovagal”[Mesh]) OR (Vasovagal Syncope[tiab]) OR (“Tilt-Table Test”[Mesh]) OR (Head-Up Tilt Test[tiab]) OR (Tilt Test[tiab])).

AND

((“Machine Learning”[Mesh]) OR (Machine Learning[tiab]) OR (Supervised Machine Learning[tiab]) OR (Supervised Learning[tiab]) OR (Deep Learning[tiab]) OR (Unsupervised Machine Learning[tiab]) OR (Unsupervised Learning[tiab]) OR (Support Vector Machine[tiab]) OR (“Artificial Intelligence”[Mesh])).

Appendix B. Quality assessment scoring

First Author (Year)	**Q1:** Did the study report the data used?	**Q2:** Did the study report data quality or data pre-processing?	**Q3:** Did the study report data distribution and if imbalanced did they handle it	**Q4:** Did the study report data normalization	**Q5:** Did the study report any rationale behind their choice of algorithm	**Q6:** Focusing on modeling, did the study report any measures to address their model bias?	**Q7:** Focusing on reproducibility, did the study report their model parameters?	**Q8:** Focusing on model training, did the study report validation technique(s) for the model?	**Q9:** Focusing on model test/validation, did the study report any performance evaluation metric of the used algorithm?
Ciliberti et al. (2018)	1	0	1	0	0	0	0	0	1
Couceiro et al. (2016)	1	1	1	1	1	0	1	1	1
Fortrat et al., 2007	1	0	1	0	0	0	0	1	1
Gilmore et al. (2021)	1	1	1	1	0	0	1	1	1
He et al. (2021)	1	1	1	1	0	1	1	1	1
Shahadat Hussain et al. (2021)	0	1	1	1	1	0	1	1	1
Hussain et al. (2021)	0	1	1	0	1	0	1	1	1
Khodor et al. (2014)	1	1	1	0	1	0	0	1	1
Khodor et al. (2016)	1	1	1	1	1	0	0	1	1
Klemenc and Štrumbelj (2015)	1	1	1	0	0	0	0	1	1
Myrovali et al. (2021)	1	1	1	0	1	0	1	1	1
Schang et al. (2006)	1	1	1	1	0	1	1	1	1
Schang et al. (2007)	1	1	1	1	0	1	1	1	1

References

Assadi, A., Laussen, P.C., Goodwin, A.J., Goodfellow, S., Dixon, W., Greer, R.W., et al., 2022. An integration engineering framework for machine learning in healthcare. Front Digit Health 4, 932411. https://doi.org/10.3389/FDGTH.2022.932411/BIBTEX.

Bartoletti, A., Alboni, P., Ammirati, F., Brignole, M., Del Rosso, A., Foglia Manzillo, G., et al., 2000. "The Italian Protocol": a simplified head-up tilt testing potentiated with oral nitroglycerin to assess patients with unexplained syncope. Europace 2, 339–342. https://doi.org/10.1053/EUPC.2000.0125.

Benditt, D.G., Ferguson, D.W., Grubb, B.P., Kapoor, W.N., Kugler, J., Lerman, B.B., et al., 1996. Tilt table testing for assessing syncope. J. Am. Coll. Cardiol. 28, 263–275. https://doi.org/10.1016/0735-1097(96)00236-7.

Brignole, M., Moya, A., De Lange, F.J., Deharo, J.C., Elliott, P.M., Fanciulli, A., et al., 2018a. 2018 ESC guidelines for the diagnosis and management of syncope. Eur. Heart J. 39, 1883–1948. https://doi.org/10.1093/EURHEARTJ/EHY037.

Brignole, M., Moya, A., De Lange, F.J., Deharo, J.C., Elliott, P.M., Fanciulli, A., et al., 2018b. Practical Instructions for the 2018 ESC guidelines for the diagnosis and management of syncope. Eur. Heart J. 39, e43–e80. https://doi.org/10.1093/EURHEARTJ/EHY071.

Chen, M., Zhang, B., Cai, Z., Seery, S., Gonzalez, M.J., Ali, N.M., et al., 2022. Acceptance of clinical artificial intelligence among physicians and medical students: a systematic review with cross-sectional survey. Front Med (Lausanne) 9, 990604. https://doi.org/10.3389/FMED.2022.990604/FULL.

Chu, W., Wang, C., Wu, L., Lin, P., Li, F., Zou, R., 2015. Oral rehydration salts: an effective choice for the treatment of children with vasovagal syncope. Pediatr. Cardiol. 36, 867–872. https://doi.org/10.1007/S00246-015-1097-5.

Ciliberti, M.A.P., Santoro, F., Di Martino, L.F.M., Rinaldi, A.C., Salvemini, G., Cipriani, F., et al., 2018. Predictive value of very low frequency at spectral analysis among patients with unexplained syncope assessed by head-up tilt testing. Arch. Cardiovasc. Dis. 111, 95–100. https://doi.org/10.1016/J.ACVD.2017.04.006.

Couceiro, R., Carvalho, P., Paiva, R.P., Muehlsteff, J., Henriques, J., Eickholt, C., et al., 2016. Real-time prediction of neurally mediated syncope. IEEE J. Biomed. Health Inform. 20, 508–520. https://doi.org/10.1109/JBHI.2015.2408994.

van Dijk, J.G., van Rossum, I.A., Thijs, R.D., 2021. The pathophysiology of vasovagal syncope: novel insights. Auton. Neurosci. 236. https://doi.org/10.1016/J.AUTNEU.2021.102899.

Forleo, C., Guida, P., Iacoviello, M., Resta, M., Monitillo, F., Sorrentino, S., et al., 2013. Head-up tilt testing for diagnosing vasovagal syncope: a meta-analysis. Int. J. Cardiol. 168, 27–35. https://doi.org/10.1016/J.IJCARD.2012.09.023.

Fortrat, J.O., Schang, D., Bellard, E., Victor, J., Lefthériotis, G., Investigator, F., 2007. Cardiovascular variables do not predict head-up tilt test outcome better than body composition. Clin. Auton. Res. 17, 206–210. https://doi.org/10.1007/S10286-007-0423-2.

Frye, M., Mohren, J., Schmitt, R.H., 2021. Benchmarking of data preprocessing methods for machine learning-applications in production. Procedia CIRP 104, 50–55. https://doi.org/10.1016/J.PROCIR.2021.11.009.

Ghorbani, A., Ouyang, D., Abid, A., He, B., Chen, J.H., Harrington, R.A., et al., 2020. Deep learning interpretation of echocardiograms. NPJ Digit Med 3. https://doi.org/10.1038/S41746-019-0216-8.

Gilmore, S., Hart, J., Geddes, J., Olsen, C.H., Mehlsen, J., Gremaud, P., et al., 2021. Classification of orthostatic intolerance through data analytics. Med. Biol. Eng. Comput. 59, 621–632. https://doi.org/10.1007/S11517-021-02314-0.

He, Z., Du, L., Du, S., Wu, B., Fan, Z., Xin, B., et al., 2021. Machine learning for the early prediction of head-up tilt testing outcome. Biomed. Signal Process. Control 69, 102904. https://doi.org/10.1016/J.BSPC.2021.102904.

Shahadat, H., Raza, Z., Giacomini, G., Goswami, N., 2021. Support vector machine-based classification of vasovagal syncope using head-up tilt test. Biology (Basel) 10. https://doi.org/10.3390/BIOLOGY10101029.

Hussain, S., Raza, Z., Kumar, T., Goswami, N., Investigator, F., Investigator, F., 2021. Diagnosing neurally mediated syncope using classification techniques. J. Clin. Med. 10. https://doi.org/10.3390/JCM10215016.

Khodor, N., Carrault, G., Matelot, D., Amoud, H., Khalil, M., Boullay, T.D., N,, et al., 2016. Early syncope detection during head up tilt test by analyzing interactions between cardio-vascular signals. Digit. Signal Process. 49, 86–94. https://doi.org/10.1016/J.DSP.2015.11.005.

Khodor, N., Matelot, D., Carrault, G., Amoud, H., Khalil, M., Ville, N., et al., 2014. Kernel based support vector machine for the early detection of syncope during head-up tilt test. Physiol. Meas. 35, 2119–2134. https://doi.org/10.1088/0967-3334/35/10/2119.

Khushi, M., Shaukat, K., Alam, T.M., Hameed, I.A., Uddin, S., Luo, S., et al., 2021. A comparative performance analysis of data resampling methods on imbalance medical data. IEEE Access 9, 109960–109975. https://doi.org/10.1109/ACCESS.2021.3102399.

Klemenc, M., Štrumbelj, E., 2015a. Predicting the outcome of head-up tilt test using heart rate variability and baroreflex sensitivity parameters in patients with vasovagal syncope. Clin. Auton. Res. 25, 391–398. https://doi.org/10.1007/S10286-015-0318-6.

Koh, R.G.L., Khan, M.A., Rashidiani, S., Hassan, S., Tucci, V., Liu, T., et al., 2023. Check it before you wreck it: a guide to STAR-ML for screening machine learning reporting in research. IEEE Access 11, 101567–101579. https://doi.org/10.1109/ACCESS.2023.3316019.

Kong J, Rios T, Kowalczyk W, Menzel S, Bäck T. On the Performance of Oversampling Techniques for Class Imbalance Problems. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12085 LNAI, 2020. Doi: 10.1007/978-3-030-47436-2_7.

Montesinos López, O.A., Montesinos López, A., Crossa, J., 2022. Support vector machines and support vector regression. Multivariate Statistical Machine Learning Methods for Genomic Prediction 337–378. https://doi.org/10.1007/978-3-030-89010-0_9.

Myrovali, E., Fragakis, N., Vassilikos, V., Hadjileontiadis, L.J., 2021. Efficient syncope prediction from resting state clinical data using wavelet bispectrum and multilayer perceptron neural network. Med. Biol. Eng. Comput. 59, 1311–1324. https://doi.org/10.1007/S11517-021-02353-7.

Rajpurkar, P., Chen, E., Banerjee, O., Topol, E.J., 2022. AI in health and medicine. Nat. Med. 28, 31–38. https://doi.org/10.1038/S41591-021-01614-0.

van Rossum, I.A., de Lange, F.J., Benditt, D.G., van Zwet, E.W., van Houwelingen, M., Thijs, R.D., et al., 2023. Variability of cardioinhibition in vasovagal syncope: differences between subgroups during cardioinhibition and beyond. Clin. Auton. Res. 33, 749–755. https://doi.org/10.1007/S10286-023-00991-5.

Russo, V., AlTurki, A., Rago, A., Proietti, R., Chaussé, G., Monteleone, A.M., et al., 2017. Temperament and character personality dimensions in nitrate-tilt-induced vasovagal syncope patients. Hellenic J. Cardiol. 58, 411–416. https://doi.org/10.1016/J.HJC.2017.01.026.

Russo, V., Parente, E., Tomaino, M., Comune, A., Sabatini, A., Laezza, N., et al., 2023. Short-duration head-up tilt test potentiated with sublingual nitroglycerin in suspected vasovagal syncope: the fast Italian protocol. Eur. Heart J. 44, 2473–2479. https://doi.org/10.1093/EURHEARTJ/EHAD322.

Schang, D., Bellard, E., Plantier, G., Dupuis, J.M., Victor, J., Leftheriotis, G., 2006. Comparison of computational algorithms applied on transthoracic impedance waveforms to predict head-up tilt table testing outcome. Comput. Biol. Med. 36, 225–240. https://doi.org/10.1016/J.COMPBIOMED.2004.09.004.

Schang, D., Feuilloy, M., Plantier, G., Fortrat, J.O., Nicolas, P., 2007. Early prediction of unexplained syncope by support vector machines. Physiol. Meas. 28, 185–197. https://doi.org/10.1088/0967-3334/28/2/007.

Shinohara, T., Ebata, Y., Ayabe, R., Fukui, A., Okada, N., Yufu, K., et al., 2014. Cardiac autonomic dysfunction in patients with head-up tilt test-induced vasovagal syncope. PACE - Pacing and Clinical Electrophysiology 37, 1694–1701. https://doi.org/10.1111/PACE.12484;JOURNAL:JOURNAL:15408159;WGROUP:STRING:PUBLICATION.

Soteriades, E.S., Evans, J.C., Larson, M.G., Chen, M.H., Chen, L., Benjamin, E.J., et al., 2002. Incidence and prognosis of syncope. N. Engl. J. Med. 347, 878–885. https://doi.org/10.1056/NEJMOA012407.

Sperandei S. Understanding logistic regression analysis. Biochem Med (Zagreb) 2014;24:12–8. Doi: 10.11613/BM.2014.003.

Suthaharan, S., 2016. Machine Learning Models and Algorithms for Big Data Classification. Integrated Series in Information Systems. vol. 36. Springer, Boston, MA; 2016. Doi: 10.1007/978-1-4899-7641-3_10.

Tannemaat, M.R., Kefalas, M., Geraedts, V.J., Remijn-Nelissen, L., Verschuuren, A.J.M., Koch, M., et al., 2023. Distinguishing normal, neuropathic and myopathic EMG with an automated machine learning approach. Clin. Neurophysiol. 146, 49–54. https://doi.org/10.1016/J.CLINPH.2022.11.019.

Thijs, R.D., Brignole, M., Falup-Pecurariu, C., Fanciulli, A., Freeman, R., Guaraldi, P., et al., 2021. Recommendations for tilt table testing and other provocative cardiovascular autonomic tests in conditions that may cause transient loss of consciousness : Consensus statement of the European Federation of Autonomic Societies (EFAS) endorsed by the American Autonomic Society (AAS) and the European Academy of Neurology (EAN). Clin. Auton. Res. 31, 369–384. https://doi.org/10.1007/S10286-020-00738-6.

Vybiral, T., Bryg, R.J., Maddens, M.E., Boden, W.E., 1989. Effect of passive tilt on sympathetic and parasympathetic components of heart rate variability in normal subjects. Am. J. Cardiol. 63, 1117–1120. https://doi.org/10.1016/0002-9149(89)90089-1.

Wiens, J., Shenoy, E.S., 2018. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. Clin. Infect. Dis. 66, 149. https://doi.org/10.1093/CID/CIX731.

Zhang, Z., 2016. Introduction to machine learning: k-nearest neighbors. Ann Transl Med 4. https://doi.org/10.21037/ATM.2016.03.37.