Machine learning for subtype definition and risk prediction in heart failure, acute coronary syndromes and atrial fibrillation: systematic review of validity and clinical utility

Amitava Banerjee, Suliang Chen, Ghazaleh Fatemifar, Mohamad Zeina, R. Thomas Lumbers, Johanna Mielke, Simrat Gill, Dipak Kotecha, Daniel F. Freitag, Spiros Denaxas and Harry Hemingway

Abstract

Background: Machine learning (ML) is increasingly used in research for subtype definition and risk prediction, particularly in cardiovascular diseases. No existing ML models are routinely used for cardiovascular disease management, and their phase of clinical utility is unknown, partly due to a lack of clear criteria. We evaluated ML for subtype definition and risk prediction in heart failure (HF), acute coronary syndromes (ACS) and atrial fibrillation (AF).

Methods: For ML studies of subtype definition and risk prediction, we conducted a systematic review in HF, ACS and AF, using PubMed, MEDLINE and Web of Science from January 2000 until December 2019. By adapting published criteria for diagnostic and prognostic studies, we developed a seven-domain, ML-specific checklist.

Results: Of 5918 studies identified, 97 were included. Across studies for subtype definition (n = 40) and risk prediction (n = 57), there was variation in data source, population size (median 606 and median 6769), clinical setting (outpatient, inpatient, different departments), number of covariates (median 19 and median 48) and ML methods. All studies were single disease, most were North American (n = 61/97) and only 14 studies combined definition and risk prediction. Subtype definition and risk prediction studies respectively had limitations in development (e.g. 15.0% and 78.9% of studies related to patient benefit; 15.0% and 15.8% had low patient selection bias), validation (12.5% and 5.3% externally validated) and impact (32.5% and 91.2% improved outcome prediction; no effectiveness or cost-effectiveness evaluations).

(Continued on next page)
Conclusions: Studies of ML in HF, ACS and AF are limited by number and type of included covariates, ML methods, population size, country, clinical setting and focus on single diseases, not overlap or multimorbidity. Clinical utility and implementation rely on improvements in development, validation and impact, facilitated by simple checklists. We provide clear steps prior to safe implementation of machine learning in clinical practice for cardiovascular diseases and other disease areas.

Keywords: Cardiovascular disease, Machine learning, Subtype, Risk prediction, Informatics, Systematic review

Background
Disease definitions rely on expert consensus, informed by best available evidence [1–3]. Machine learning (ML), the use of algorithms to describe patterns in datasets with (supervised) or without (unsupervised) the need to define them a priori, is increasingly used in research for definition [4] and risk prediction [5]. However, established evaluation frameworks for clinical utility of such models [6, 7] have not been applied.

Better definitions for cardiovascular disease (CVD), the single greatest disease burden in the UK and globally [6], may improve prevention and treatment by characterising target populations, enabling comparability and generalisability across study designs and populations. Heart failure (HF), acute coronary syndromes (ACS) and atrial fibrillation (AF) are among the commonest CVDs globally [8, 9]. Despite frequent changes in disease definitions [3, 10, 11], continued diagnostic and prognostic uncertainties have led to ML research in all three diseases [12–17]. HF, ACS and AF frequently overlap, so phenotyping and risk prediction are relevant across diseases. To date, no existing ML models are routinely used for CVD management, and their phase of clinical utility is unknown, partly due to a lack of clear criteria.

Unlike drugs and devices [18, 19], the pathway to translation from research to routine use is unclear for ML. The AI-TREE criteria (Additional file_Web table 1) were developed for ML healthcare research [20], but consensus is lacking for best practice or regulatory requirements for implementation [21, 22]. Importance of such guidelines is emphasised by systematic reviews showing that ML performs no better than logistic regression or healthcare professionals in prediction across diseases, with high risk of bias and limited external validation in published studies [23, 24].

Evidence for ML in subtype definition and risk prediction in HF, ACS and AF has not been assimilated. Moreover, it is unclear how application of ML in CVD compares with other diseases. Our aims were to (i) develop a simple framework for clinical utility and validity of ML models in healthcare and (ii) conduct a systematic review to evaluate methods and results of ML for subtype definition or risk prediction in HF, AF and ACS.

Methods
We followed the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) statement, with a protocol agreed by all authors, without prospectively registering.

Systematic review
Identification of studies
We searched PubMed, MEDLINE and Web of Science databases from 1 January 2000 until 31 December 2019. Our search terms, agreed by co-authors, pertained to machine learning, clustering, cardiovascular disease, heart failure, atrial fibrillation, acute coronary syndromes, subtype and risk prediction. Reference lists and expert recommendations were taken into account, to identify grey literature, including conference reports and proceedings, guidelines, working papers and theses (Additional file_Search strategy).

Selection of studies
All abstracts were independently screened and then full text of selected abstracts were respectively assessed for eligibility by two reviewers (from AB, SC and MZ), and conflicts were resolved by a third reviewer (AB or SC).

Inclusion and exclusion criteria
Studies were eligible if the publication presented:

(i) ML models for disease subtype definition/clustering for HF, ACS and AF; or
(ii) Risk prediction for HF, ACS or AF

Studies were excluded if they:

(i) Were not original empirical data
(ii) Were not English language
(iii) Were not peer-reviewed (e.g. a published dissertation was not eligible)
(iv) Did not concern models developed for humans
(v) Did not have full text available

Data extraction
Tools for extraction were adapted and developed with consensus among co-authors from published
frameworks for new risk markers (AHA [6]), diagnostic accuracy (QUADAS-2 [25]), prognostic tools (CHARMS [26], PROGRESS [7], TRIPOD [27]) and ML (AI-TREE [20], Christodoulou [23]) (Additional file_Web table 2). The author, year and country of study, clinical setting, data source, outcome, comparator methods, ML method(s) and covariates were extracted. The final checklist was in the three stages of the translational pathway (“development”, “validation” and “impact”) as per AHA and PROGRESS statements [6, 7]. There were seven domains under the three main stages: clinical relevance, patients, algorithms (“development”), internal validation, external validation (“validation”), and clinical utility and effectiveness (“impact”). The questions for each domain were from published guidelines as above with similar extraction items for subtype definition and risk prediction studies. Data extraction was by two independent reviewers (from AB, MZ and SC), and disagreements were resolved by a third reviewer (AB or SC). Quantitative analysis was beyond the scope of this review and its aims.

Results
Systematic review of HF, AF and ACS
Of 5918 articles identified by our search, 97 met the inclusion criteria (Additional file_Figure S1).

Unsupervised ML for subtype definition
Of the 40 studies of unsupervised ML for subtype definition (included patients: median \( n = 606 \); min 117, max 251,451), there were 27 in HF, 9 in ACS and 4 in AF. All studies focused on a single disease, and only 6 (15%) studies included data regarding history of all three diseases. Twenty-three (57.5%) studies involved \(< 1000\) individuals (range 117–874) and 29 (72.5%) were based in North America with no analyses from low- or middle-income countries. Across diseases, 26 (65%) studies were in outpatients (8 inpatient and 6 mixed inpatient and outpatient) and 11 used trial data (10 prospective cohort, 5 retrospective cohort, 11 cross-sectional, 5 registries), with 7 using EHR data. The mean number of covariates was 31 (min 3, max 156), most commonly demographic and symptom variables (Table 1) [28–67].

Most studies used only one ML method (\( n = 22 \)). Non-hierarchical clustering was used in 12 studies (hierarchical: \( n = 25 \)). The most commonly used method was K-means (\( n = 11 \)), but other methods included Gaussian mixture modelling, latent class analysis and random forest (RF). All studies reported disease clusters. Most studies found 3 clusters (\( n = 17, 42.5\% \); median 3), and most clusters were based on symptom (\( n = 20 \)) or physical (\( n = 17 \)) variables. Clustering was usually case-wise (\( n = 37 \)), rather than variable-wise (\( n = 3 \)) and one study used both approaches.

At the development stage, there were limitations at “clinical relevance” (question related to patient benefit 15.0%, target condition applicability 62.5% and data suitable for clinical question 70%), “patient” (patient applicability 30.0%, low patient selection bias 15.0%) and “algorithm” (algorithm applicability 55.0%, low algorithm bias 40.0%) levels. Twelve studies did not validate or replicate findings; 28 (70.0%) had internal validation (most commonly using number of clusters or prediction of mortality/admissions) and only 5 (12.5%) externally validated findings. There were significant deficiencies under “clinical utility” (improved prediction of outcomes 32.5%, methods available 0%, clinically relevant metrics 72.5%, interpretable by clinicians 65%, clinically justified results 40%) and “effectiveness” (no studies showing effectiveness, real-world or cost) domains (Table 2) [28–67].

Supervised ML for risk prediction
Of the 57 studies of supervised ML for risk prediction (median \( n = 6769 \); min 28, max 2,994,837), 31 were in HF, 19 in ACS and 7 in AF. Ten of 57 studies involved \(< 1000\) individuals and most were from North America (\( n = 32 \), with one from a low- or middle-income country (\( n = 1 \)). Risk prediction studies focused on development of (i) HF, ACS or AF in healthy individuals or the general population by case-control, cross-sectional or cohort design (\( n = 25 \)) and (ii) outcomes in HF, ACS or AF (\( n = 32 \)) (Table 3) [13, 45, 53, 65–118].

Across diseases, 27 studies were in outpatients (23 inpatient and 7 mixed inpatient and outpatient) and 5 used trial data (6 prospective cohort, 30 retrospective cohort, 6 registry and 11 case-control). Thirty-one studies used EHR data. The mean number of covariates was 723 (median 48; min 6, max 15,815), most commonly demography, symptoms and comorbidities/drugs. The ML methods used were variable with neural networks (\( n = 19 \)), random forest (\( n = 23 \)) and support vector machine (\( n = 16 \)).

At the development stage, there were concerns at “clinical relevance” (question related to patient benefit 78.9%, target condition applicability 68.4% and data suitable for clinical question 68.4%), “patient” (patient applicability 43.9%, low patient selection bias 15.8%) and “algorithm” (algorithm applicability 68.4%, low algorithm bias 47.4%) levels. Only three studies (5.3%) had external validation in the clinical trial (\( n = 12,063 \), prospective cohort (\( n = 861 \)) and registry (\( n = 4759 \)) data. There was internal validation in 49/57 (86.0%) by “hold-out”, “leave one out” or k-fold cross-validation. The risk of bias was high in all studies with the commonest causes as patient selection (\( n = 47 \)), patient applicability to the clinical question (\( n = 35 \)) and bias in the algorithm(s) (\( n = 31 \)). Again, there were major limitations under “clinical relevance” (question related to patient benefit 15.0%, target condition applicability 62.5% and data suitable for clinical question 70%), “patient” (patient applicability 30.0%, low patient selection bias 15.0%) and “algorithm” (algorithm applicability 55.0%, low algorithm bias 40.0%) levels.
utility” (improved prediction of outcomes 91.2%, methods available 45.6%, clinically relevant metrics 49.1%, interpretable by clinicians 47.4%, clinically justified results 43.9%) and effectiveness (no studies showing effectiveness, real-world or cost) domains (Table 4) [13, 45, 53, 65–118].

Discussion

Our study highlighted three key findings with important impact on the use of ML in subtype definition and risk prediction for HF, ACS and AF. First, there is significant variation in methods, data and reporting of results. Second, we provide a pragmatic framework, based on published criteria, for assessing validity and clinical utility of ML studies. Third, there are major limitations at development, validation and impact stages in studies of ML in CVD, and our scoping review (Additional file_ Scoping review_Methods and Results, Additional file_ Web table 3 [119–129], Additional file_ Web table 4 [130–143]) [144, 145] suggests similar problems for non-CVDs. Addressing these issues has potential to improve future data science approaches and enhance value of these methods for clinical decision-making and better patient care.

There is increasing interest in guidelines for development and implementation of ML in healthcare from patients and public, academics and industry, partly due to current uncertainty regarding necessary approval and regulatory procedures [146, 147]. Whether in medicines or medical devices, the “development pipeline” (e.g.
### Table 2: Quality assessment of machine learning studies of subtype definition for heart failure, acute coronary syndromes and atrial fibrillation (n = 40)

| Author, Year (REF) | DEVELOPMENT | VALIDATION | IMPACT |
|--------------------|-------------|------------|--------|
|                    | Clinical relevance | Patients | Algorithm | Internal | External | Clinical utility | Effectiveness | Cost effectiveness |
| Heart failure      |              |           |          |          |          |                  |              |                  |
| Ieva, 2017 (28)    | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Ahmad, 2018 (15)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Ather, 2011 (29)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Kao, 2015 (30)     | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Sideris, 2015 (31) | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Ahmad, 2014 (32)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Panahizar, 2015 (33)| ● ● ● ●     | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Katz, 2017 (34)    | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Vellone, 2017 (35) | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Cikes, 2018 (46)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Moser, 2014 (36)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Jurgens, 2009 (37) | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Segar, 2019 (38)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Bose, 2018 (39)    | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Alonso-Betanzos, 2015 (40)| ● ● ● ● | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Song, 2010 (41)    | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Shah, 2015 (42)    | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Horiuchi, 2018 (43)| ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Lee, 2010 (44)     | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Huang, 2018 (45)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Przewlocka-Kosmala, 2019 (46)| ● ● ● ● | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Omar, 2017 (47)    | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Ahmad, 2016 (48)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Sanchez-Martinez, 2018 (49)| ● ● ● ● | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Hertog, 2010 (50)  | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Nowak, 2017 (51)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Harr, 2014 (52)    | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |

**Acute coronary syndromes**

| Sharma, 2019 (53) | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Riegel, 2010 (54) | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Ryan, 2007 (55)   | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Rosenfeld, 2015 (56)| ● ● ● ● | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| DeVon, 2010 (57)  | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| McSweeney, 2010 (58)| ● ● ● ● | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Lindgren, 2008 (59)| ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Abbott, 2010 (60) | ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Fukuda, 2007 (61)| ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |

**Atrial fibrillation**

| Inohara, 2018 (62)| ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Inohara, 2019 (63)| ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |
| Streuer, 2018 (64)| ● ● ● ●      | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●  | ● ● ● ●         | ● ● ● ●       | ● ● ● ●         |

- Negative/No
- Positive/Yes
phases I to IV of drug development) and guidelines are much clearer, ultimately promoting clinical effectiveness and patient safety. The wide variations in methods, datasets and reporting are not surprising, given the current lack of such guidelines with respect to the use of ML in subtype definition and risk prediction [147]. On the basis of existing checklists, we have developed a straightforward 7-domain checklist (16 points) to capture development, validation and impact. Although there are multiple efforts to standardise reporting guidelines for AI in healthcare [27, 148], the relationship with the “translational pathway” in terms of development, validation and impact has not been emphasised. We show that no published studies have results which are ready to be implemented, again likely to be related to lack of consensus regarding what is required in research and clinical practice to ensure clinical effectiveness and safety of ML.

| Table 3 Machine learning risk prediction studies in heart failure, acute coronary syndromes and atrial fibrillation (n = 57) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Source | Type | Design | Setting and Population | Outcome | Comparison method | Machine learning methods | Example |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Banerjee et al. BMC Medicine (2021) 19:85 | 10 | 10 | United States | Major adverse cardiovascular events | Logistic regression | Random forest | Impact was demonstrated on clinical outcomes. |
| Achenbach et al. JAMA Network Open (2020) 3:6 | 10 | 10 | United States | Major adverse cardiovascular events | Logistic regression | Random forest | Impact was demonstrated on clinical outcomes. |
| Akritidis et al. JAMA Network Open (2020) 3:3 | 10 | 10 | United States | Major adverse cardiovascular events | Logistic regression | Random forest | Impact was demonstrated on clinical outcomes. |
| Alsi et al. JAMA Network Open (2020) 3:1 | 10 | 10 | United States | Major adverse cardiovascular events | Logistic regression | Random forest | Impact was demonstrated on clinical outcomes. |
| Anderson et al. JAMA Internal Medicine (2019) 179:20 | 10 | 10 | United States | Major adverse cardiovascular events | Logistic regression | Random forest | Impact was demonstrated on clinical outcomes. |
| Banerjee et al. BMC Medicine (2021) 19:85 | 10 | 10 | United States | Major adverse cardiovascular events | Logistic regression | Random forest | Impact was demonstrated on clinical outcomes. |

ACS, acute coronary syndrome; AF, atrial fibrillation; Atherosclerosis Risk in Communities Study; CHARGE Cohorts for Heart and Aging Research in Genomic Epidemiology; CHA2DS2-VASc congestive heart failure, hypertension, age > 75, diabetes mellitus, stroke, vascular disease, sex category; CVD, cardiovascular disease; ECG, electrocardiogram; EHR, electronic health records; GRACE, Global Registry of Acute Coronary Events; HF, heart failure; HFrEF, heart failure with preserved ejection fraction; IP, hospital inpatient; LV, left ventricular; OP, hospital outpatient; RCT, randomised controlled trial; TIMI, thrombolysis in myocardial infarction; UK, United Kingdom; US, United States

*Australia, Austria, Brazil, Canada, China, Denmark, Korea, Finland, France, Germany, Italy, Japan, Mexico, Norway, Poland, Spain, Sweden, Netherlands, and UK

**Negative/no for all columns (except in the “Baseline population” column, where it denotes “unreported”)**

● Positive/yes
### Table 4 Quality assessment of machine learning studies of risk prediction for heart failure, acute coronary syndromes and atrial fibrillation (n = 57)

| Author, Year (REF) | DEVELOPMENT | VALIDATION | IMPACT |
|-------------------|-------------|------------|--------|
|                   | Clinical relevance | Patients | Algorithm | Internal | External | Clinical utility | Effectiveness |
| Heart failure     | Question relates to patient benefit? | Target condition replicability? | Baseline replicability? | New patient selection bias? | Algorithm replicability? | Internal validation? | Extern validation? | Clinical utility | Effectiveness |
| Rasmy, 2018 (60) | ● ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Banerjee, 2019 (66) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Chol, 2017 (67) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Ng, 2016 (68) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Wang, 2015 (69) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Yn, 2019 (70) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Lindahl, 2018 (71) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Nagar, 2019 (72) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Blecker, 2016 (73) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Blecker, 2017 (74) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Austin, 2017 (75) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Nascimento, 2018 (76) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Tabassian, 2018 (77) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Masereika, 2016 (78) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Van, 2018 (79) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Fritzr, 2017 (80) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Miller, 2019 (81) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Gelles, 2018 (82) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Awan, 2018 (83) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Khon, 2019 (84) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Anisias, 2019 (85) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Tenamiketakina, 2016 (86) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Panahizar, 2015 (87) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Wang, 2018 (88) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Martaç, 2018 (89) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Herki, 2018 (90) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Shimer, 2017 (91) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Mahjhan, 2015 (92) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Gleeson, 2017 (93) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Valenza, 2018 (94) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Hernán, 2015 (95) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |

**Acute coronary syndromes**

| Author, Year (REF) | DEVELOPMENT | VALIDATION | IMPACT |
|-------------------|-------------|------------|--------|
| VanKooten, 2014 (96) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Ikemoto, 2014 (97) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| A/An, 2019 (98) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Noi, 2019 (99) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Laviole, 2009 (100) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Wu, 2019 (101) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Borko, 2016 (102) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| A/An, 2019 (103) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Declin, 2018 (104) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Hu, 2019 (105) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Gibson, 2020 (106) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Sharma, 2019 (107) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |

**Atrial fibrillation**

| Author, Year (REF) | DEVELOPMENT | VALIDATION | IMPACT |
|-------------------|-------------|------------|--------|
| Hil, 2019 (111) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Twark, 2020 (112) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Bundy, 2020 (113) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Gofill, 2016 (114) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Goto, 2020 (115) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Tog, 2016 (116) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |
| Lai, 2017 (117) | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● | ● ● |

* Negative/No
● Positive/Yes
Problems at the development stage

Studies to date of ML in HF, ACS and AF have been limited by number and type of covariates, population size, geographic location (mainly North American) and clinical setting. As with epidemiologic studies and trials, the generalisability of data and reproducibility of methods [6, 7] are crucial to make findings interpretable and translatable to clinical care. However, the majority of studies to date have not fully considered these factors, resulting in high risk of bias in all studies.

EHR data and advanced data analytics, which have been either under-used or under-reported in studies to date, offer research opportunities across diseases, but ML studies have focused on single diseases, when diseases often co-exist, as for HF, ACS and AF. Outside CVD, perhaps the most promising ML studies to date have been in either large imaging datasets or settings where there is linkage of multimodal clinical data. Whereas the popular concept of ML suggests the use of a wide range of covariates, particularly in clustering, the number and type of variables have been limited. Availability of covariates and quality of recording in EHR determine these limitations rather than research or clinical need. In the future, as more complete and larger EHR datasets become available for research, studies of ML can focus on improved use of limited data in clinical practice, or use of more comprehensive lists of covariates and coexisting diseases for the discovery of new factors in disease definition or prediction. For example, the ESC classification lists 89 causes of HF, which have not been studied together in a single population-based study in EHR or otherwise [3].

The majority of studies have had positive results suggesting publication bias (not formally assessed here), which likely overestimates potential healthcare impact. Lessons must be learned from biomarkers and genomics [149, 150], where lack of standardisation and biased reporting have contributed to lack of translation from science to practice. It is concerning that a significant proportion of studies ML are being developed with questions unrelated to patient benefit, or with data unsuitable to answer them. “Data-driven care” is important in both personalised and precision medicine [151], both of which can benefit from advances in the use of ML. However, a “data-centred” or “data-driven” agenda must remain “patient-centred” rather than “technology-centred”
(or in this case, “ML-centred”) [152]. Our framework can be used to plan studies, facilitating clinical relevance and reducing bias.

Problems at the validation stage
In order to improve data-driven characterisation of CVD and influence clinical decision-making, ML studies for subtyping and prediction should be larger-scale, across diseases, with standardised reporting and proven external validity. To date, the degree of external validation across subtype definition and risk prediction ML studies has been disappointing, making ML difficult to implement in routine care. External validation will help to understand which clustering and prediction tools are of greatest use, and greater availability of electronic health record data should facilitate this step in the pathway to implementation of ML.

Problems at the impact stage
There are major gaps at the phases of clinical utility, particularly open availability of methods and interpretability of methods and results by clinicians. A limitation of our study is that we did not conduct meta-analysis for studies of either subtype classification or risk prediction because it was out of scope for our review. However, variation and lack of standardisation in methods and reporting make meta-analysis challenging and potentially unrepresentative. These issues can be addressed at design and implementation phases, but have been relatively neglected. Importantly, in applications of ML in CVD to date, studies of effectiveness and cost-effectiveness are lacking. Without these evaluations, ML cannot be implemented safely or effectively.

Figure 1 shows deficiencies at development, validation and impact stages and that ML research is not necessarily being conducted in a “sequential” manner, e.g. to ensure that the evaluation of impact is in validated ML models, perhaps reflecting lack or under-use of consensus guidelines. Particularly in CVD, with high disease burden globally and in low- and middle-income countries, the impact of ML has to be considered through a global lens. The fact that the majority of the literature regarding ML in CVD is from the North American context does raise concern that AI could broaden research and clinical inequalities both within and across countries, beyond current debates about the inequalities which may be inherent in algorithms. If individuals are not represented in the data feeding into the algorithms, then the AI cannot benefit them. These inequalities still exist in other domains, such as pharmaceutical trials and genomics, but the situation is improving, e.g. the proportion of RCTs recruiting in low- to middle-income countries has increased in recent years. As EHR and digital healthcare become global phenomena, there is scope to use ML in diverse data and settings. Validation of ML applications in clustering and risk prediction across countries and settings, even between high-income countries, are urgently needed to advance this agenda. We have proposed steps to improve the development and validation of ML in clinical datasets (Additional file_ Web table 5).

Conclusions
Studies to date of ML in HF, ACS and AF have been limited by number and type of covariates, population size, geographic location (mainly North American), clinical setting, focus on single diseases (where overlap and multimorbidity are common) and ML methods used. Moreover, flaws at stages of development, validation and impact reduce the clinical utility and likelihood of implementation of ML in routine healthcare. To improve the generalisability, applicability and clinical utility of ML in CVD and other diseases, and to influence clinical decision-making internationally, we provide a simple checklist to foster standardised reporting and validation.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12916-021-01940-7.

**Additional file 1.** Search terms and search strategy.

**Additional file 2 Figure S1.** PRISMA flow diagram.

**Additional file 3 Web Table 1.** AI-TREE checklist.

**Additional file 4 Web Table 2.** Data extraction for included studies.

**Additional file 5.** Methods and Results for scoping review.

**Additional file 6 Web Table 3.** Subtype classification studies in other disease areas.

**Additional file 7 Web Table 4.** Risk prediction studies in other disease areas.

**Additional file 8 Web Table 5.** Proposed methods for development and validation of ML.

**Additional file 9.** PRISMA checklist.

Acknowledgements
Not applicable.

Authors’ contributions
Research question: AB
Funding: AB and HH
Study design and analysis plan: AB and HH
Data extraction: AB, SC and MZ
Drafting initial versions of manuscript: AB
Drafting final versions of manuscript: AB and HH
Critical review of early and final versions of manuscript: all authors
The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. AB is the guarantor. The authors read and approved the final manuscript.

Funding
AB is supported by research funding from NIHR, British Medical Association, Astra-Zeneca and UK Research and Innovation. HH is funded by NIHR (Senior Investigator and University College London Hospitals Biomedical Research Centre). HH is supported by Health Data Research UK (grant No. LOND1), which is funded by the UK Medical Research Council, Engineering and
Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome Trust. HH, AB, SD, SC, DK, DFF, JM, GZ, SC and RTL are supported by the BigData@Heart consortium, funded by the Innovative Medicines Initiative-2 Joint Undertaking under grant agreement No. 116074. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and EFPIA; it is chaired by DE Grobbee and SD Anker partnering with 20 academic and industry partners and ESC. GF is funded by AHA Institutional Data Fellowship programme. SD is supported by an Alan Turing Fellowship. This paper represents independent research [part] funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at University College London Hospital (UCLH BRC). DK reports grants from the National Institute of Health Research (NIHR CDF-2015-08-074 and NIHR HTA-130280), the British Heart Foundation (PG/17/55/33087 and AA/18/2/34218) and IRCSS San Raffaele/Menarini (Beta-blockers in Heart Failure Collaborative Group NCT00832444), in addition to personal fees from Bayer (Advisory Board), AtriCure (Speaker fees), Amomed (Advisory Board), Menarini (Research Grant) and Myokardia (Advisory Board), all outside the submitted work.

Availability of data and materials

The data underlying this article are available in the article and in its online supplementary material.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

AB has received research funding from Astra-Zeneca. DF and JM are employees of Bayer. All other authors declare none.

Author details

1Institute of Health Informatics, University College London, 222 Euston Road, London NW1 2DA, UK. 2Health Data Research UK, University College London, 5. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA. 2018;319(13):1317–8.

References

1. Otsubo F, Joner M, Prat F, Virmani R, Narula J. Clinical classification of plaque morphology in coronary disease. Nat Rev Cardiol. 2014;11(7):379–89.
2. Dunlay SM, Roger VL, Redfield MM. Epidemiology of heart failure with preserved ejection fraction. Nat Rev Cardiol. 2017;14(10):591–602.
3. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J. 2016;37(27):2129–220.
4. Hwang T, Atluri G, Xie M, Dey S, Hong C, Kumar V, et al. Co-clustering phenotype-genome for phenotype classification and disease gene discovery. Nucleic Acids Res. 2012;40(19):e146.
5. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA. 2018;319(13):1317–8.
6. Hatky MA, Greenland P, Ametek DK, Ballantyne CM, Criqui MH, Eblin MS, Go AS, Harrell FE Jr, Hong Y, Howard BV, Howard VL, Hsue PY, Kramer CM, McConnell JP, Normand SL, O’Donnell CJ, Smith SC Jr, Wilson PW, American Heart Association Expert Panel on Subclinical Atherosclerotic Diseases and Emerging Risk Factors and the Stroke Council. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. Circulation. 2009;119(7):2408–16. https://doi.org/10.1161/CIRCULATIONAHA.109.102278.
7. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG, PROGRESS Group. Prognosis Research Strategy (PROGRESS): 3: prognostic model research. PLoS Med. 2013;10(2):e1001381. https://doi.org/10.1371/journal.pmed.1001381.
8. Newton JN, Briggs AD, Murray CJ, Dicker D, Foreman KJ, Wang H, et al. Changes in health in England, with analysis by English regions and areas of deprivation, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;386(10010):2257–74.
9. Joseph P, Leong D, McKee M, Anand SS, Schwalmb JD, Teo K, et al. Reducing the global burden of cardiovascular disease, part 1: the epidemiology and risk factors. Circ Res. 2017;121(6):77–94.
10. Kirchoff P, Curtis AB, Skanes AC, Gillis AM, Samuel Wann L, John CA. Atrial fibrillation guidelines across the Atlantic: a comparison of the current recommendations of the European Society of Cardiology/European Heart Rhythm Association/European Association of Cardiothoracic Surgeons, the American College of Cardiology Foundation/American Heart Association/Heart Rhythm Society, and the Canadian Cardiovascular Society. Eur Heart J. 2019;34(20):1471–4.
11. Thygesen K, Alpert JS, Jaffe AS, Chairman BR, Box JI, Morrow DA, et al. Fourth universal definition of myocardial infarction (2018). Circulation. 2018;138(20):e618–e51.
12. AwO SE, Sohle F, Sanfilippo FM, Binnamouns M, Dwored G. Machine learning in heart failure: ready for prime time. Curr Opin Cardiol. 2018;33(2):190–5.
13. Myers PD, Scirica BM, Stultz CM. Machine learning improves risk stratification after acute coronary syndrome. Sci Rep. 2017;7(1):12692.
14. Sreer M, Ractiffe SJ, Callans D, Shoemaker MB, Riegel B. Atrial fibrillation symptom clusters and associated clinical characteristics and outcomes: a cross-sectional secondary data analysis. Eur J Cardiovasc Nurs. 2018;17(8):707–16.
15. Ahmad T, Lund LH, Rao P, Ghosh R, Water P, Vaccaro B, et al. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. J Am Heart Assoc. 2018;7(8):008801.
16. Gikes M, Sanchez-Martinez S, Cloggert B, Duchateau N, Piella G, Butkoff C, et al. Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. Eur J Heart Fail. 2019;21(11):174–85.
17. Alonso-Betanzos A, Bolon-Canedo V. Big data-analysis, cluster analysis, and machine-learning approaches. Adv Exp Med Biol. 2018;1065:607–26.
18. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c343.
19. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2013;346:f3529.
20. Vollmer S, Mateen BA, Bohner G, Kiraly FI, Ghani R, Jonsson P, Cumber S, Jonas A, McCAllister KSL, Myles P, Granger D, Bixie M, Bransons R, Moons KG, Collins GS, Ioannidis JPA, Holmes C, Hemingway H. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ. 2020;368:l6297.
21. Department of Health and Social Care. Code of conduct for data-driven health and care technology. 2018.
22. Wiens J, Sara S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, Jung K, Heller K, Kale D, Saeed M, Ossorio PN, Thadaney-Israni S, Goldenberg A. Do no harm: a roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337–40.
23. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbalik JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019;110:12–22.
24. Lui X, Paes L, Kale AU, Wagner SK, Fu DJ, Brunysseels A, Mahendiran T, Moraes G, Shamdas M, Kern C, Ledsam JR, Smith MC, Balasak K, Topol EJ, Bachmann LM, Keane PA, Denniston AK. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digital Health. 2019;1(6):271–97.
25. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36.

26. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med. 2014;11(10):e1001744.

27. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g794. https://doi.org/10.1136/bmj.g794.

28. Ieva F, Pagani AM, Pietrabissa T. Dynamic clustering of hazard functions: an application to disease progression in chronic heart failure. Health Care Manag Sci. 2017;20(3):353–64. https://doi.org/10.1007/s10729-016-9537-3. Epub 2016 Feb 4.

29. Ather S, Peterson LE, Divakaran VG, Deswal A, Ramasubbu K, Giorgberidze I, Pandey A. Phenomapping of patients with heart failure with preserved ejection fraction. Circ Cardiovasc Nurs. 2017;32(2):180–1. https://doi.org/10.1111/ccn.12423.

30. Wang J, Zhong P, Li X, Duan H, Liu J. A regularized deep learning approach for clinical risk prediction of acute coronary syndrome using electronic health records. IEEE Trans Biomed Eng. 2018;65(5):956–68. https://doi.org/10.1109/TBME.2017.2731158. Epub 2017 Jul 24.

31. Sideris C, Alisharaf M, Pourhamzayoon M, Shahimhammadfard M, Samy L, Sarrafzadegan M. A data-driven feature extraction framework for predicting the severity of condition of congestive heart failure patients. Conf Proc IEEE Eng Med Biol Soc. 2015;2015:2534–7. https://doi.org/10.1109/IEMBS.2015.7318008.

32. Ahmad T, Desai N, Wilson F, Schulte P, Dunning A, Jacoby D, Allen L, Fizazi M, Rogers J, Felker GM, O'Connor C, Patel CB. Clinical implications of cluster analysis-based classification of acute decompensated heart failure and correlation with bedside hemodynamic profiles. PLoS One. 2016(11(6):e0156881.

33. Sanchez-Martinez S, Duchateau N, Erdi E, Kunig K, Aakhus S, Degiovanni A, Marino P, Cautiero E, Piella G, Frazier AG, Bijnens BH. Machine learning analysis of left ventricular function to characterize heart failure with preserved ejection fraction. Circ Cardiovasc Imaging. 2018;11(4):e007138. https://doi.org/10.1161/CIRCIMAGING.117.007138.

34. Hertzog MA, Poezel B, Duncan K. Cluster analysis of symptom occurrence to identify subgroups of heart failure patients: a pilot study. J Cardiovasc Nurs. 2010;25(2):273–83. https://doi.org/10.1097/JCN.0b013e318181fbc6.

35. Nowak RM, Reed BP, DiSomma S, Nanayakkara P, Moyer M, Mills S, Levy P. Presenting symptoms of acute heart failure patients in the ED: identification and implications. Am J Emerg Med. 2017;35(4):536–42. https://doi.org/10.1016/j.ajem.2016.12.003. Epub 2016 Dec 8.

36. Herr JK, Salyer J, Flattery M, Goodloe L, Lyon DE, Kabbab CS, Clement DG. Heart failure symptom clusters and functional status - a cross-sectional study. J Adv Nurs. 2015;71(6):1274–87. https://doi.org/10.1111/jan.12596. Epub 2014 Dec 23.

37. Sharma A, Sun JL, Lokhnyngyin Y, Roe MT, Ahmad T, Desai NR, Blazing MA. Patient phenotypes, cardiovascular risk, and ezetimibe treatment in patients after acute coronary syndromes (from IMPROVE-IT). Am J Cardiol. 2019;123(8):1193–201.

38. Riegel B, Hanlon AL, McKinley S, Moser DK, Meischke H, Doering LV, et al. Symptom clusters in men and women with heart failure and their impact on cardiac event-free survival. J Cardiovasc Nurs. 2010;25(4):263–72. https://doi.org/10.1097/JCN.0b013e3181c5bb88.

39. Huang Z, Dong W, Duan H, Liu J. A regularized deep learning approach for clinical risk prediction of acute coronary syndrome using electronic health records. IEEE Trans Biomed Eng. 2018;65(5):956–68. https://doi.org/10.1109/TBME.2017.2731158. Epub 2017 Jul 24.

40. Przewlocka-Kosmala M, Marwick TH, Dabrowski A, Kosmala W. Contribution of cardiovascular reserve to prognostic categories of heart failure with preserved ejection fraction: a classification based on machine learning. J Am Soc Echocardiogr. 2019;32(5):604–615.e6. https://doi.org/10.1016/j.echo.2018.12.002. Epub 2019 Feb 1.

41. Omar AMS, Narula S, Abdel Rahman MA, Pedrzizzetti G, Raslan H, Riafei O, Narula J. Sengupta PP. Precision phenotyping in heart failure and pattern clustering of ultrasound data for the assessment of diastolic dysfunction. JACC Cardiovasc Imaging. 2017;10(1):291–303. https://doi.org/10.1016/j.jcmg.2016.10.012. Epub 2016 Jan 18.

42. Ahmad T, Desai N, Wilson F, Schulte P, Dunning A, Jacoby D, Allen L, Fizazi M, Rogers J, Felker GM, O'Connor C, Patel CB. Clinical implications of cluster analysis-based classification of acute decompensated heart failure and correlation with bedside hemodynamic profiles. PLoS One. 2016(11(6):e0156881.

43. Horiuchi Y, Tanimoto S, Latif AHMM, Urayama KY, Aoki J, Yahagi K, Okuno T, Blaustein A, Wehrens XH, Mann DL, Bozkurt B. Digoxin treatment in heart failure patients with preserved ejection fraction: possible implications for clinical and therapeutic use. J Electrocardiol. 2011;44(5):368–74. https://doi.org/10.1016/j.jelectrocard.2010.04.021. Epub 2010 May 15.

44. Kar S, Selvaraj S, Burke M, Yancy CW, Gheorghide M, Bonow RO, Huang CC, Deo RC. Phenomapping for novel classification of heart failure with preserved ejection fraction. Circulation. 2015;131(3):269–79. https://doi.org/10.1161/CIRCULATIONAHA.114.016367. Epub 2014 Nov 14.

45. Horuchi Y, Tanimoto S, Latif AHMM, Urayama KY, Aoki J, Yahagi K, Okuno T, Sato Y, Tanaka T, Koseki K, Komiya Y, Nakajima H, Hara K, Tanabe K. Identifying novel phenotypes of acute heart failure using cluster analysis of clinical variables. Int J Cardiovasc Sci. 2018;26:525–73. https://doi.org/10.1016/j.jcardiovasc.2018.03.098. Epub 2018 Mar 29.

46. Lee KS, Song EK, Lennie TA, Frazier SK, Chung ML, Heo S, Wu JR, Rayens MK, Riegel B, Moser DK. Symptom clusters in men and women with heart failure and their impact on cardiac event-free survival. J Cardiovasc Nurs. 2010;25(4):263–72. https://doi.org/10.1097/JCN.0b013e3181c5bb88.

47. Omar AMS, Narula S, Abdel Rahman MA, Pedrzizzetti G, Raslan H, Riafei O, Narula J. Sengupta PP. Precision phenotyping in heart failure and pattern clustering of ultrasound data for the assessment of diastolic dysfunction. JACC Cardiovasc Imaging. 2017;10(1):291–303. https://doi.org/10.1016/j.jcmg.2016.10.012. Epub 2016 Jan 18.

48. Ahmad T, Desai N, Wilson F, Schulte P, Dunning A, Jacoby D, Allen L, Fizazi M, Rogers J, Felker GM, O'Connor C, Patel CB. Clinical implications of cluster analysis-based classification of acute decompensated heart failure and correlation with bedside hemodynamic profiles. PLoS One. 2016(11(6):e0156881.
60. Abbott AA, Barnason S, Zimmerman L. Symptom burden clusters and their impact on psychosocial functioning following coronary artery bypass surgery. J Cardiovasc Nurs. 2010;25(4):301–10. https://doi.org/10.1097/JCN.0b013e3181eb4e46.

61. Abbas FR, Haddad S, Sood ES, Lyall BD, Gupta M, Stuebe AM, Macon WG, deSouza CB, deFilippi CR, W的眼神,59–62. https://doi.org/10.1001/jama.2019.10189

62. Alves SE, Bennamoun M, Soel, S., Filippini FM, D. Wildevs G. Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. ESC Heart Fail. 2019;6(2):428–35. https://doi.org/10.1002/ehf2.12419

63. Al-Abed M, Massanari A, Banerjee A, Cadeiras M, Alaa AM, van der Schaar M. A cluster analysis of the Japanese multicenter outpatient registry of patients with atrial fibrillation. Am J Cardiol. 2019;124(6):871–8. https://doi.org/10.1016/j.amjcard.2019.05.071

64. Alward K, Fraser AG, Duffield M, et al. Predicting heart failure using whole-slide images of H&E tissue. PLoS One. 2018;13(4):e0197318. https://doi.org/10.1038/EMBC.2018.303

65. Al-Bayati M, Kojima G, Felsted J, Kallarakal S, Kvedar J, Jethwani K. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. BMC Med Inform Decis Mak. 2018;18(1):44. https://doi.org/10.1186/s12911-018-0620-2

66. Allan CI, Thielke R, Zhou Z, et al. Predictive modeling of hospital readmission rates using electronic health record data. J Biomed Inform. 2015;55:106–14. https://doi.org/10.1016/j.jbi.2015.04.003

67. Allan R, Moomey S, O’Mahony S, et al. Using methods from the data-mining and machine-learning literature for disease classification and outcomes: a multicenter registry study. JAMA Cardiol. 2018;3(1):54–63. https://doi.org/10.1001/jama cardiol.2017.4665

68. Alkontiainen M, Parkkinen J, Aaltonen M, et al. Analysis of the Japanese multicenter outpatient registry of patients with atrial fibrillation. Am J Cardiol. 2019;124(6):871–8. https://doi.org/10.1016/j.amjcard.2019.05.071

69. Al-Khaismeh T, Tyrer P, Bell S, et al. Symptom clusters in adults with chronic atrial fibrillation. J Cardiovasc Nurs. 2017;32(3):296–303. https://doi.org/10.1097/JCN.0000000000000344.

70. Al-Khateeb N, Al-Khateeb N. A comparative study of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2019;4(2):e182723.

71. Al-Khateeb N, Al-Khateeb N. A comparative study of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2019;4(2):e182723.

72. Al-Khateeb N, Al-Khateeb N. A comparative study of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2019;4(2):e182723.

73. Al-Khateeb N, Al-Khateeb N. A comparative study of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2019;4(2):e182723.

74. Al-Khateeb N, Al-Khateeb N. A comparative study of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2019;4(2):e182723.

75. Al-Khateeb N, Al-Khateeb N. A comparative study of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2019;4(2):e182723.

76. Al-Khateeb N, Al-Khateeb N. A comparative study of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2019;4(2):e182723.
94. Valenza G, Wendt H, Kyone K, Hayano J, Watanabe E, Yamamoto Y, Aby P, Barbieri R. Mortality prediction in severe congestive heart failure patients with multifractal point-process modeling of heartbeat dynamics. IEEE Trans Eng Med Biol. 2016;35(10):2545–54. https://doi.org/10.1109/TBME.2016.2797158 ePub 2018 Jan 23.

95. Henriquez J, Canvalho P, Paredes S, Rocha T, Habetha J, Antunes M, Morais J. Prediction of heart failure decompensation events by trend analysis of telemonitoring data. IEEE J Biomed Health Inform. 2015;19(5):1757–69. https://doi.org/10.1109/JBHI.2014.2358715 ePub 2014 Sep 17.

96. Vanhouten JP, Stammer JM, Lorenzi NM, Maron DJ, Lasko TA. Machine learning for risk prediction of acute coronary syndrome. AMIA Ann Symp Proc. 2014;2014:1946–9 eCollection 2014.

97. Beunza JJ, Puertas E, Garcia-Ovejero E, Villalba G, Condes E, Koleva G, Hurtado C, Landecho MF. Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease). J Biomed Inform. 2019;97:103257. https://doi.org/10.1016/j.jbi.2019.103257 ePub 2019 Jul 30.

98. Al'Aref SJ, Singh G, van Rosendael AR, Kolli KK, Ma X, Maliakal G, Pandey M, Banerjee K, Aszczynski R, Burchardt P. Machine-learned models using electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease on coronary computed tomography angiography: analysis from the CONFIRM registry. Eur Heart J. 2020;41(13):359–67. https://doi.org/10.1093/eurheartj/ehz265.

99. Noh YK, Park JY, Choi BG, Kim KE, Rha SW. A machine learning-based approach for predicting major adverse cardiac events in acute coronary syndromes from electronic health record data. PLoS One. 2018;13(8):e0202344. https://doi.org/10.1371/journal.pone.0202344 eCollection 2018.

100. Noh YK, Park JY, Choi BG, Kim KE, Rha SW. Determinants of in-hospital mortality after percutaneous coronary revascularization. J Med Syst. 2019;43(8):253. https://doi.org/10.1007/s10916-019-1359-5.

101. Lavesson N, Halling A, Freitag M, Odeberg J, Odeberg H, Davidsson P. Machine learning for risk prediction of major adverse cardiac events after acute coronary syndrome: a pooled randomized clinical trial analysis. J Thromb Thrombolysis. 2020;49(1):1–10. https://doi.org/10.1007/s11239-019-01940-8.

102. Hsu CC, Hsu MD, Islam MM, Poly TN, Yang HC, Nguyen PA, Wang YC, Li YJ. An artificial intelligence approach to early predict non-ST-elevation myocardial infarction patients with chest pain. Comput Methods Prog Biomed. 2019;173:105507. https://doi.org/10.1016/j.cmpb.2019.01.013.

103. Berikol GB, Yildiz O, Ozcan IT. Diagnosis of acute coronary syndrome with a support vector machine. J Med Syst. 2016;40(4):84. https://doi.org/10.1007/s10916-016-0432-6 ePub 2016 Jan 27.

104. Al’Aref SJ, Singh G, van Rosendael AR, Kolli KK, Ma X, Malakal G, Pandey M, Lee BC, Wang J, Xu Z, Zhang Y, Min JK, Wong SC, Minutello RM. Determinants of in-hospital mortality after percutaneous coronary intervention: a machine learning approach. J Am Heart Assoc. 2019;8(5):e011160. https://doi.org/10.1161/JAHA.118.011160.

105. Steele JJ, Denaxas S, Shah AD, Hemingway H, Luscombe NM. Machine learning versus traditional risk stratification methods in acute coronary syndromes using electronic health records. IEEE J Biomed Health Inform. 2019;23(5):2117–26. https://doi.org/10.1109/JBHI.2018.2882518 ePub 2018 Nov 20.

106. Shouval R, Hadanny A, Shlomo N, Iakobishvili Z, Unger R, Zahger D, Alcalai R, Astar S, Gotsfel A, Matesky S, Gofen Y, Beigel R. Machine learning for prediction of 30-day mortality after ST elevation myocardial infarction: an Acute Coronary Syndrome Israei Survey data mining study. Int J Cardiol. 2017;2467–13. https://doi.org/10.1016/j.ijcard.2017.05.067.

107. Hill NR, Ayoubkhani D, McEwan P, Sugrue DM, Faroqiu U, Lister S, Lumley M, Balhain A, Cohen AT, O'Neill M, Clifton D, Gordon J. Predicting atrial fibrillation in primary care using machine learning. PLoS One. 2019;14(11):e0224582. https://doi.org/10.1371/journal.pone.0224582 eCollection 2019.

108. Tiwari P, Colborn KL, Smith DE, Xing F, Ghosh D, Rosenberg MA. Assessment of a machine learning model applied to harmonized electronic health record data for the prediction of incident atrial fibrillation. JAMA Netw Open. 2020;3(1):e191936.

109. Bundy JD, Heckbert SR, Chen LY, Lloyd-Jones DM, Greenland P. Evaluation of risk prediction models of atrial fibrillation (from the Multi-Ethnic Study of Atherosclerosis (MESA)). Am J Cardiol. 2020;125(1):55–62. https://doi.org/10.1016/j.amjcard.2019.09.032 ePub 2019 Oct 10.

110. Goodfellow SD, Goodwin A, Green R, Lausen PC, Mazwi M, Eytan D. Atrial fibrillation classification using step-by-step machine learning. Biomed Phys Eng Expos. 2018;4045005.

111. Goto S, Goto S, Pieper KS, Bassand JP, Camm AJ, Fitzmaurice DA, Goldhaber SZ, Haas S, Parente S, Otto M, Misselwitz F, Turpie AGG, Vermeulen AT, Vrijland WA, Kerkhof CP, Kakkad AK, GARFIELD-AF Investigators. New AI prediction model using serial PT-INR measurements in AF patients on VKAs: GARFIELD-AF. Eur Heart J Cardiovasc Pharmacother. 2019;pvx076. https://doi.org/10.1093/ehjcvp/pvx076 eCollection 2019.

112. Han L, Askari M, Altman RB, Schmitt SK, Fan J, Bentley JP, Narayan SM, Turakhia MP. Atrial fibrillation burden signature and near-term prediction of stroke: a machine learning analysis. Circ Cardiovasc Qual Outcomes. 2019;12(1):e005595.

113. Huang Z, Dong W, Lu X, Duan H, He K, Huang Z. Evidential MACE prediction after acute coronary syndrome. J Biomed Inform. 2019;93:103297. https://doi.org/10.1016/j.jbi.2016.10.007 Epub 2016 Oct 12.

114. Lynch CM, van Berkel VH, Friebos HB. Application of unsupervised analysis techniques to lung cancer patient data. PLoS One. 2017;12(9):e0184370. https://doi.org/10.1371/journal.pone.0184370 eCollection 2017.

115. Schuler A, Liu Y, Van W, Callahan A, Udelil M, Stark DE, Shah NH. Discovering patient phenotypes using generalized low rank models. Pac Symp Biocomput. 2016:124–155.

116. Conrad DJ, Bailey BA. Multidimensional clinical phenotyping of an adult cystic fibrosis patient population. PLoS One. 2015;10(3):e0122705 eCollection 2015.

117. Niehaus KE, Uhlig HH, Clifton DA. Phenotypic characterisation of Crohn’s disease severity. Conf Proc IEEE Eng Med Biol Soc. 2015;2015:7502–3. https://doi.org/10.1109/EMBC.2015.7320006.

118. Beaulieu-Jones BK, Greene CS, Pooled Resource Open-Access ALS Clinical Trials Consortium. Semi-supervised learning of the electronic health record data for phenotype stratification. J Biomed Inform. 2016;64:168–78. https://doi.org/10.1016/j.jbi.2016.10.007 ePub 2016 Oct 12.

119. Cho SB, Kim SC, Chung MG. Identification of novel population clusters with different susceptibilities to type 2 diabetes and their impact on the prediction of diabetes. Sci Rep. 2019;9(1):1–10. https://doi.org/10.1038/s41598-019-40058-y.

120. Ahkvist E, Storm P, Kirjàmäki A, Martineill M, Dorkhan M, Carlsson A, Vikman P, Prasad RB, Aly DM, Almgren P, Wessman Y, Shaat N, Spégel P, Mulder H, Lindholm E, Melander O, Hansson O, Malqvist U, Lennmark Å, Lahtt K, Forsén T, Tuomain T, Rosenberg AH, Groop L. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol. 2018;6(5):361–9. https://doi.org/10.1016/S2213-8588(18)30051-2 ePub 2018 Mar 5.

121. Ceret de Langavant L, Bayen E, Yaffe K. Unsupervised machine learning to identify high risk variables in population-based surveys: development and validation study. J Med Int Res. 2018;20(7):10493. https://doi.org/10.2196/10493.
127. Seymour CW, Kennedy JN, Wang S, Chang CH, Elliott CF, Xu Z, Berry S, Clermont G, Cooper G, Gomez H, Huang DT, Kellum JA, Mi Q, Opal SM, Talisa V, van der Poll T, Visserwan Saran, Vodovotz Y, Weiss JC, Yealy DM, Yende S, Angus DC. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. JAMA. 2019;321(20): 2003–17. https://doi.org/10.1001/jama.2019.5791.

128. Pikoulas M, Quint JK, Nissen F, Hemingway H, Smeeth L, Denaxas S. Identifying clinically important COPD sub-types using data-driven approaches in primary care population based electronic health records. BMC Med Inform Decis Mak. 2019;19(1):86. https://doi.org/10.1186/s12911-019-0805-0.

129. Wojnarski CM, Roselli EE, Idrees JJ, Zhu Y, Carnes TA, Lowry AM, Collier PH, Tomaš Taylor RA, Pare JR, Venkatesh AK, Mowafi H, Melnick ER, Fleischman W, Hall MK. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK Biobank participants. PLoS One. 2019;14(5):e0216533.

130. Weng SF, Reps J, Kaj J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One. 2017;12(4): e0174944. https://doi.org/10.1371/journal.pone.0174944.eCollection 2017.

131. Zhao J, Feng Q, Wu P, Lupta RA, Willke RA, Wells QS, Denny JC, Wei WQ. Learning from longitudinal data in electronic health record and genetic data to improve cardiovascular event prediction. Sci Rep. 2019;9(1):1717. https://doi.org/10.1038/s41598-018-36745-6.

132. Zambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folston AR, Shea S, Guallar E, Blumenta DM, Lima JC. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. Circ Res. 2017;121(10):1092–101. https://doi.org/10.1161/CIRCRESAHA.117.311312.

133. Farani B, Channanath AM, Behbehani K, Thanaraj TA. Predictive models to assess risk of type 2 diabetes, hypertension and comorbidity: machine-learning algorithms and validation using national health data from Kuwait—a cohort study. BMJ Open. 2013;3(5):e002457. https://doi.org/10.1136/bmjopen-2012-002457.

134. Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT, Hamzah H, Ho J, Lee XQ, Hsu W, Lee ML, Musonda I, Chandran M, Chipato-Mutapi G, Muma M, Tan GS, Siwapradas D, Menon G, Wong TF, Ting DSH. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. Lancet Digital Health. 2019;1(1):35–44.

135. Bertimas D, Dunn J, Pawlowski C, Silberholz J, Weinstein A, Zhuo YD, Chen E, Effkay AA. Applied informatics decision support tool for mortality predictions in patients with cancer. JCO Clin Cancer Inform. 2018;2:1–11. https://doi.org/10.1200/CCI.18.00003.

136. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining techniques. Artif Intell Med. 2003;29(2):89–127.

137. Chevray CJ, Liu N, Wong TH, Sim YE, Abdulrah HR. Utilizing machine learning methods for preoperative prediction of postsurgical mortality and intensive care unit admission. Ann Surg. 2019; https://doi.org/10.1097/SLA.0000000000003297.

138. De Fauw J, Ledsmar JR, Romero-Paredes B, Nikolov S, Tomasiev N, Blackwell S, Askham H, Giroir B, O’Donoghue B, Visentin D, van den Driessche G, Lakshminarayanan B, Meyer C, Mackinder F, Boustou S, Ayoub K, Chorpa R, King D, Karthikasalangam A, Hughes CO, Rainé R, Hughes J, Sim DA, Egan C, Tufal A, Montgomery H, Hassabis D, Rees G, Back T, Khaw PT, Suleyman M, Cornebise J, Keane PA, Ronneberger O. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med. 2018;24(9):1342–50. https://doi.org/10.1038/s41591-018-0107-6 Epub 2018 Aug 13.

139. Goldstein BA, Pomann GM, Winkelmayer WC, Perciavilla MJ. A comparison of risk prediction methods using repeated observations: an application to electronic health records for hemodialysis. Stat Med. 2017;36(17):2750–63. https://doi.org/10.1002/sim.7308 Epub 2017 May 2.

140. Taylor RA, Pare JR, Venkatesh AK, Mowafi H, Melnick ER, Fleischman W, Hall MK. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach. Acad Emerg Med. 2016;23(3):269–78. https://doi.org/10.1111/acem.13276 Epub 2016 Feb 13.

141. Tomasiev N, Giroir B, Rae MW, Ziekmink M, Askham H, Saravia A, Mottram A, Meyer C, Rauri S, Protsyuk I, Connell A, Hughes CO, Karthikasalangam A, Cornebise J, Montgomery H, Rees G, Laing C, Baker CR, Petersen K, Reeves R, Hassabis D, King D, Suleyman M, Back T, Nelson C, Ledsmar JR, Mohamed S. A clinically applicable approach to continuous prediction of future acute kidney injury. Nature. 2019;572(7767):16–9. https://doi.org/10.1038/s41586-019-1390-1 Epub 2019 Jul 31.

142. Tibble H, Tanas A, Home E, Home R, Muzani M, Simpson CR, Sheikh A. Predicting asthma attacks in primary care: protocol for developing a machine learning-based prediction model. BMJ Open. 2019;9(7):e028375. https://doi.org/10.1136/bmjopen-2018-028375.

143. Moher D, Stewart L, Shamse B. All in the family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. Syst Rev. 2015;4:183.

144. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc. 2015;13(3):141–6.

145. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. Lancet. 2019;393(10181):1577–9.

146. Coiera E, Ammenwerth E, Georgiou A, Magrabi F. Does health informatics have a replication crisis? J Am Med Inform Assoc. 2018;25(8):963–8.

147. Hernandez-Boussard T, Bokurt S, Ioannidis JP, Shah NH. MINIMAR (MINimum Information for Medical AI Reporting): developing reporting standards for artificial intelligence in health care. JAMIA. 2020; Published online 28 June 2020. https://academic.oup.com/jamia/article/doi/10.1093/jamia/ocaa080/5864179.

148. Sugrue LP, Desikam RS. What are polygenic scores and why are they important? JAMIA. 2019;32(18):1820–1.

149. Gudusanì D, Barros I, Zeggini E, Sandhu MS. Genomics of disease risk in globally diverse populations. Nat Rev Genet. 2019;20(9):520–35.

150. Schussler-Fiorenza Rose SM, Contraposi K, Moneghetti KJ, Zhou W, Mishra T, Mataraso S, et al. A longitudinal big data approach for precision health. Nat Med. 2019;25(5):792–804.

151. Panch T, Matte H, Gell LA. The “inconvenient truth” about AI in healthcare. NPJ Digit Med. 2019;2:77.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.