REVIEW ARTICLE   OPEN

Check for updates

Multimodal machine learning in precision health: A scoping review

Adrienne Kline[1], Hanyin Wang[1], Yikuan Li[1], Saya Dennis[1], Meghan Hutch[1], Zhenxing Xu[2], Fei Wang[2], Feixiong Cheng[3] and Yuan Luo[1]✉

Machine learning is frequently being leveraged to tackle problems in the health sector including utilization for clinical decision-support. Its use has historically been focused on single modal data. Attempts to improve prediction and mimic the multimodal nature of clinical expert decision-making has been met in the biomedical field of machine learning by fusing disparate data. This review was conducted to summarize the current studies in this field and identify topics ripe for future research. We conducted this review in accordance with the PRISMA extension for Scoping Reviews to characterize multi-modal data fusion in health. Search strings were established and used in databases: PubMed, Google Scholar, and IEEEXplore from 2011 to 2021. A final set of 128 articles were included in the analysis. The most common health areas utilizing multi-modal methods were neurology and oncology. Early fusion was the most common data merging strategy. Notably, there was an improvement in predictive performance when using data fusion. Lacking from the papers were clear clinical deployment strategies, FDA-approval, and analysis of how using multimodal approaches from diverse sub-populations may improve biases and healthcare disparities. These findings provide a summary on multimodal data fusion as applied to health diagnosis/prognosis problems. Few papers compared the outputs of a multimodal approach with a unimodal prediction. However, those that did achieved an average increase of 6.4% in predictive accuracy. Multi-modal machine learning, while more robust in its estimations over unimodal methods, has drawbacks in its scalability and the time-consuming nature of information concatenation.

INTRODUCTION

Clinical decision support has long been an aim for those implementing algorithms and machine learning in the health sphere[1–3]. Examples of algorithmic decision supports utilize lab test values, imaging protocols or clinical (physical exam scores) hallmarks[4,5]. Some health diagnoses can be made on a single lab value or a single threshold, such as in diabetes in older adults[6]. Other diagnoses are based on a constellation of the signs, symptoms, lab values and/or supportive imaging and are referred to as a clinical diagnosis. Oftentimes these clinical diagnoses are based on additive scoring systems that requires an admixture of positive and negative hallmarks prior to confirmatory labeling.

The modus operandi of a clinical diagnosis may fail to consider the relative weighting of these disparate data inputs and potentially non-linear relationships highlighting the limitations of human decision-making capacity. The strength of algorithmic decision-making support is that it can be used to offload such tasks, ideally yielding a more successful result. This is the promise of precision medicine. Precision medicine/health aims to create a medical model that customizes healthcare (decisions, treatments, practices etc.) that are tailored to either an individual or patient phenotype[7]. This includes tracking patients' health trajectories longitudinally[8], oftentimes incorporating genetics/epigenetics[9,10] and mathematical modeling[11] where diagnoses and treatments incorporate this unique information[12]. Contrast this with a one-drug-fits-all model, where there is a single treatment per disorder. Figure 1 illustrates the flow of information from hospitals/care centers that generate disparate data. It is through computational modeling and information fusion that outcomes of interest such

as drug and treatment targets ultimately facilitate better decision making at the patient level in those care centers. This phenomenon has sparked an interest in fusion studies using health care data.

Undertakings to characterize this literature have been performed by Huang et al.[13], who performed a systematic review of deep learning fusion of imaging and EHR data in health. However, it was limited to EHR and imaging data and deep learning applications. A follow-up review article included a commentary on omics and imaging data fusion[14]. The purpose of this study is to highlight the current scope of this research domain, summarize and offer suggestions to advance the field. The current study is more inclusive in the breadth of the types of machine learning protocols used and attempts to encompass all current modalities (information types/sources).

Data fusion is underpinned by information theory and is the mechanism by which disparate data sources are merged to create an information state based on the sources' complementarity[15,16] (Box 1). The expectation in machine learning is that data fusion efforts will result in an improvement in predictive power[17,18] and therefore provide more reliable results in potentially low validity settings[19]. Data fusion touts the advantage that the results of modeling become inherently more robust by relying on a multitude of informational factors rather than a single type. However, the methodology of combinatory information has drawbacks; it adds complexity to specifying the model and reduces the interpretability of results[19,20].

Data from different sources and file formats are rarely uniform, and this is especially the case with clinical data[21]. For example, data sets can have different naming conventions, units of

[1]Department of Preventive Medicine, Northwestern University, Chicago 60201 IL, USA. [2]Department of Population Health Sciences, Cornell University, New York 10065 NY, USA. [3]Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland 44195 OH, USA. ✉email: yuan.luo@northwestern.edu

measure, or represent different local population biases. Care must be taken to search and correct for systematic differences between datasets and assess their degree of inter-operability. For example, Colubri et al. aggregated computed tomography (CT) and PCR lab values, by performing an intra-site normalization. This ensured that the values were comparable across sites. In doing so they discarded several potentially informative clinical variables since they were not all available in all datasets[22].

A balance is required to allow information that is similar to work together (harmonization) and retain data purity (information correspondence)[23]. Successful fusion uses data harmonization techniques that assure both in the quality control of the

integration process. Clinical data harmonization requires multi-disciplinary research among medicine, biology, and computer science. The clinical area of heart failure with preserved ejection fraction (HFpEF) saw novel applications of multiple tensor factorization formulations to integrate the deep phenotypic and trans-omic information[24], and this extends to other areas of precision medicine[25]. To increase the portability of EHR-based phenotype algorithms, the Electronic Medical Records and Genomics (eMERGE) network has adopted common data models (CDMs) and standardized design patterns of the phenotype algorithm logic to integrate EHR data with genomic data and enable generalizability and scalability[26–29].

There are three main types of data fusion that are used in machine learning; early (data-level), intermediate (joint), and late (decision-level)[30]. In the case of early fusion, multiple data sources are converted to the same information space. This often results in vectorization or numerical conversion from an alternative state, such as that performed by Chen et al. via vectorized pathology reports[31]. Medical images possess characteristics that can undergo numerical conversion based on area, volume, and/or structural calculations[32]. These are then concatenated with additional measurements from structured data sources and fed into an individual classifier. Canonical correlation analysis[33], non-negative matrix factorization[34,35], Independent Component Analysis (ICA)

Box 1 Terms and Concepts

Multimodal machine learning: the area of machine learning concerned with bringing together disparate data sources to capitalize on the unique and complementary information in an algorithmic framework.
Data harmonization: using machine learning to unify different data sources to improve its quality and utilization.
Multiview machine learning: another term for multimodal machine learning.
Data fusion: the specific methodologies undertaken to perform data integration for multimodal/multiview machine learning; they come in three broad categories: early, intermediate/joint, and late.
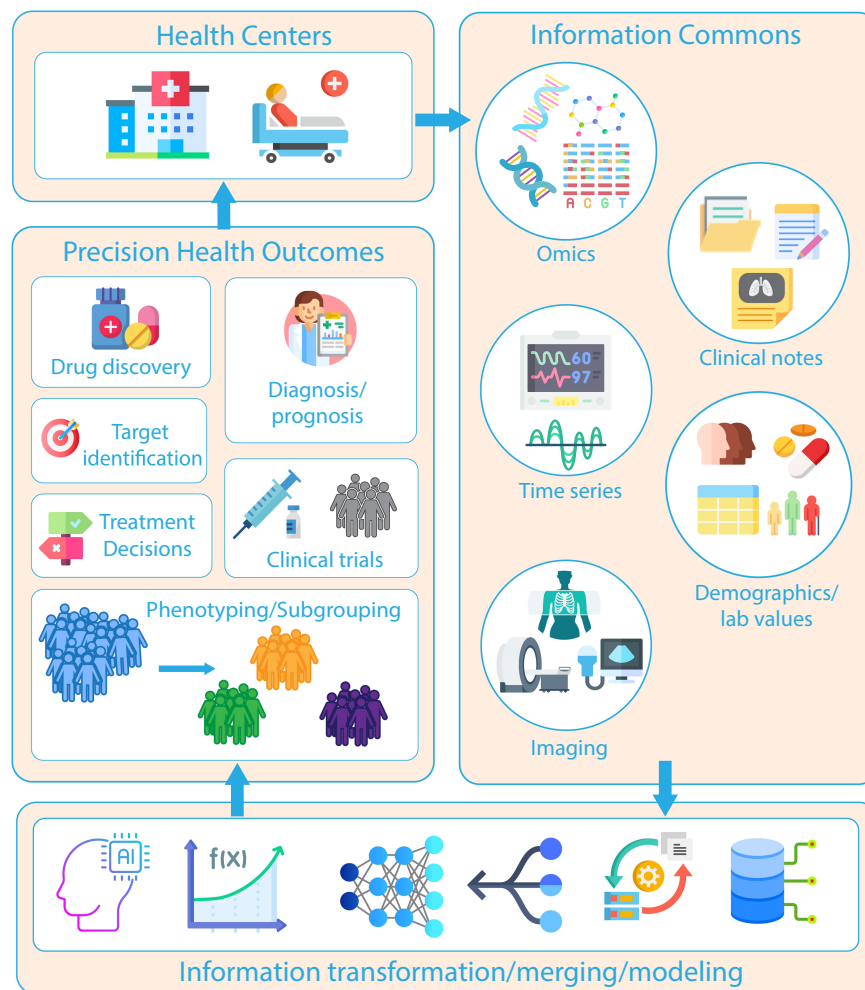

Fig. 1 Multimodal precision health; the flow of information. Information moves in a cyclical pattern from health centers to information commons, where it can be transformed and algorithmic modeling performed. These algorithms provide insight into many different health outcomes such as clinical trials, phenotyping, drug discovery, etc. These insights should return to health centers and practitioners to provide the most efficient, evidence-based medicine possible.
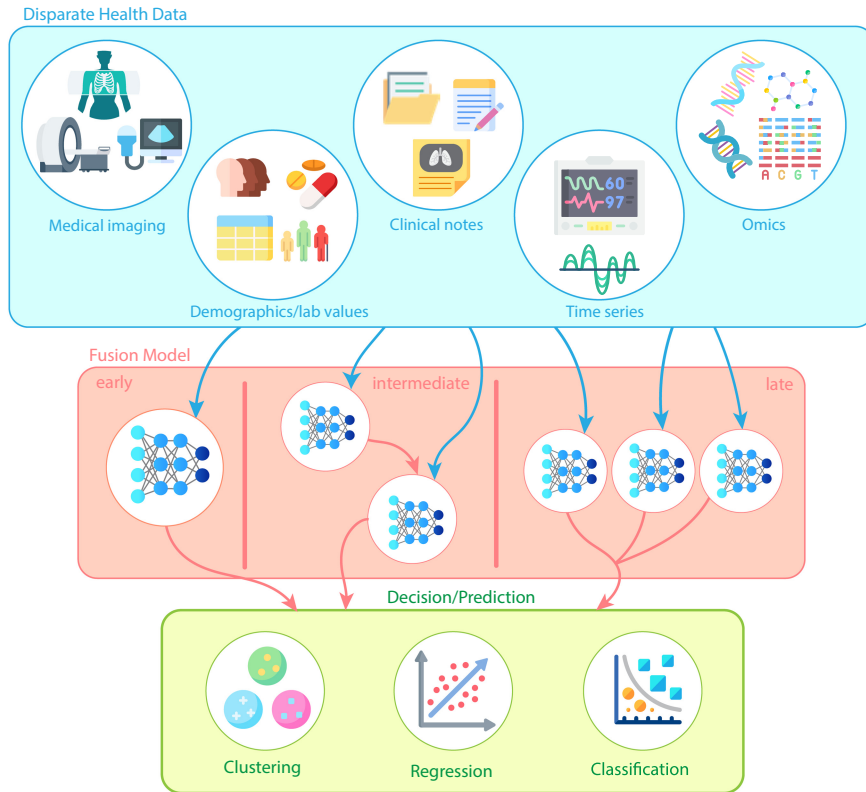
Fig. 2 Early, intermediate, and late fusion; flow of information from information commons to model structure to outcomes. Information fusion can occur in a myriad of ways. In machine learning, early, intermediate, and late fusion is typified by if all the information flows into a single model (early), a step-wise fashion where outputs from one model become inputs for another (intermediate), and lastly, where all unique data types undergo separate modelling after which ensembling and/or voting occurs (late).

and numerical feature conversion methodologies exist as common options to transform all data into the same feature space[36].

Intermediate data fusion occurs as a stepwise set of models and offers the greatest latitude in model architecture. For example, a 3-stage deep neural learning and fusion model was proposed by Zhou et al.[37]. Stage 1 consists of feature selection by a soft-max classifier for independent modalities. Stages 2 and 3 constitute combining these selected features, establishing a further refined set of features, and feeding these into a Cox-nnet to perform joint latent feature representation for Alzheimer's diagnosis. In contrast to early fusion, intermediate fusion combines the features that distinguish each type of data to produce a new representation that is more expressive than the separate representations from which it arose.

In late fusion, typically multiple models are trained where each model corresponds to an incoming data source. This is akin to ensemble learning, which offers better performance over individual models[38]. Ensemble methods use multiple learning algorithms (typically applied to the same dataset) to obtain better predictive performance than could be obtained from any of the constituent learning algorithm alone. However, multimodal machine learning ensemble here can refer to ensemble learning within a data type or across data types. These take symbolic representations as sources and combine them to obtain a more accurate decision[39]. Bayesian's methods are typically employed at this level[40] to support a voting process between the set of models into a global decision. Within late fusion there has been headway made to perform multitask deep learning[41–47]. A schematic for the 3 subtypes of data fusion is presented in Fig. 2. Attributes in the fusion techniques are shown in Table 1.

Table 1. Comparison of fusion techniques.

Attribute	Early	Intermediate/joint	Late/ decision
Scalable	No	Yes	Yes
Multiple models needed	No	Yes	Yes
Improved accuracy	Yes	Yes	Yes
Voting of multiple models	No	Yes	Yes
Interaction effects across sources	Yes	Yes	No
Implemented in health	Yes	Yes	Yes

RESULTS
Topic Modeling
The topic modeling displayed in Fig. 3 showcases the category, specific health ailment under investigation, and the modality type for the studies included. These were subsequently mapped to the category of the combination of information that were merged to create models for prediction/classification/clustering (Table 2). This plot should serve as a resource to fellow researchers to identify areas that are less frequent, such as dermatology[48], hematology[49], medication/drug issues such as alcohol use disorder that may offer new research horizons[50]. Figure 4 identifies coding platforms, publishing trend and location over time, author locations and patient cohorts of the papers included in this review.
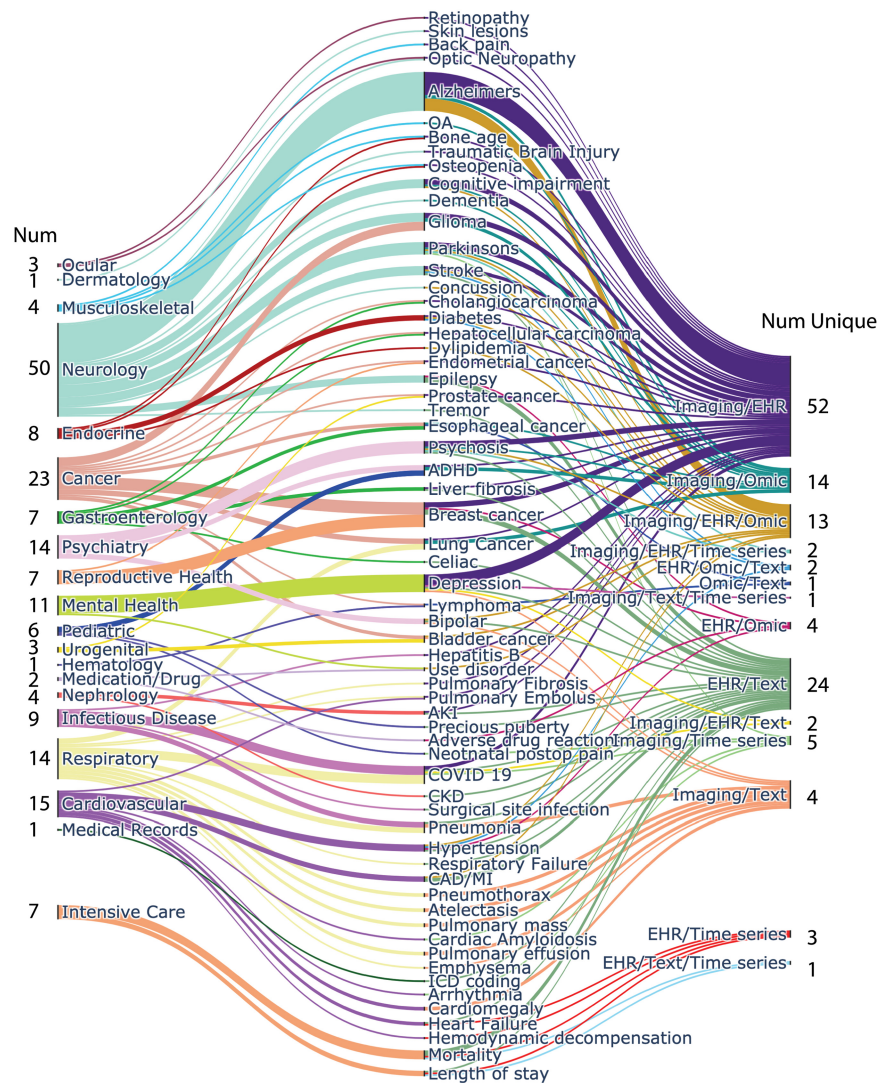
Fig. 3 Topic and Modality Modeling. Neurology, and in particular, Alzheimer's disease investigations accounted for the most papers published on this topic ($n = 22$). With the onset of the COVID-19 pandemic, several primary research articles were dedicated to this topic, which can be arrived at through the respiratory or infectious disease hierarchies. All papers noted in this review used either two or three disparate data sources when fusing their data, and specifically that of imaging and EHR ($n = 52$), was the most prevalent.

Model validation, techniques, and modalities used

Of the models used in the papers, 126/128 explicitly reported performing a validation procedure of them. The most common validation processes performed were N-fold cross validation (55)[51,52], train test split (51), leave one out cross validation (10), and external dataset (10). A cornucopia of machine learning techniques and methods were used within and across articles in this review. They have been summarized in Table 2, noting in which fusion umbrella subtype they were implemented.

Early fusion

Most papers were published using early fusion. Of those, most were published using medical imaging and EHR data[34,36,48,53,56,57,60,62–64,68,71,73,75,85–88,90,92–96,100,104,111]. Nearly all these papers performed numericalization of image features in essence converting them to structured data prior to processing, however, two performed matrix factorization[34,36]. A combination of EHR and text data was noted in 15 papers[31,54,69,72,79–81,91,99,102,106,112]. Meng et al. created a Bidirectional Representation Learning model used latent Dirichlet allocation (LDA) on clinical notes[112]. Cohen et al. used

unigrams and bigrams in conjunction with medication usage[54]. Zeng et al. used concept identifiers from text as input features[81]. Nine papers used early fusion with imaging, EHR and genomic data[32,50,51,55,61,65,83,89,108]. Doan et al. concatenated components derived from images with polygenic risk scores[83]. Lin et al. also created aggregated scores from MRI, cerebral spinal fluid, and genetic information and brought them together into a single cohesive extreme learning machine to predict mild cognitive impairment[55]. Tremblay et al. used a multivariate adaptive regression spline (MARS) after normalizing, removing highly correlated features[89]. Ten papers performed fusion using imaging and genomic data[33,52,70,76–78,82,84,97,110]. Three of these generated correlation matrices as features by vectorizing imaging parameters and correlating them with single nucleotide polymorphisms (SNPs) prior to feeding into the model[33,70,78]. Three papers in this category used EHR and time series[58,74,101]. Both Hernandez and Cannière et al. implemented their methods for purposes of cardiac rehabilitation and harnessed the power of support vector machines (SVMs). However, Hernandez preserved time series information by assembling ECG data into tensors that preserve the structural and temporal relationships inherent in the feature space[74], while Cannière performed dimensionality reduction of

Table 2. Fusion and machine learning methods included in this review.

Fusion type	Machine learning models/techniques implemented	References
Early	Support Vector Machine	31–34,36,50,51,53–81
	Random Forest, Decision Trees	31,33,34,36,67–71,73–76,78–80,82–91
	Gaussian model	36,55,92
	Bayesian models: Bayesian network, Naïve Bayes	31,54,71,73–75,80,93
	Regression: Logistic, Ridge, Cox, LASSO, MARS	31,51,67–69,77,79,80,87,89,90,94–99
	Multitask learning	62
	iMSF	100
	Boosted Models, Adaboost	68,71–75,79,90
	Deep Neural Networks: DNN, RNN, CNN, DUN, AutoEncoders	48,57,61,80,81,91,97,101–107
	Natural Language Processing, GPT, BERT	81,102,105
	Clustering: K-means, hierarchical, KNN	65,73,75,76,100,108
	Graph models	107–109
	Ensemble learning	65,88
	Artificial Neural Networks	57,64,69,73,78,81,88,91,101,104,107,110,111
	Markov Model	108
	Transfer learning	48,62,97,105,112
	Multitask learning	62
Intermediate	Regression: logistic, multivariate, support vector, LASSO	41,43,47,59,113–122
	Support Vector Machine	37,59,113–115,117,120,123–126
	Decision Trees, Random Forest	41,49,59,115–117,120,122,127,128
	Deep Learning: CNN, DNN, RNN, AutoEncoders	37,41–47,59,115–117,121–135
	Boosted Learning: gradient boosting, XGBoost, Adaboost	116,118,126
	Natural Language Processing	122,126,134,136
	Artificial Neural Network	43–45,49,116,117,121,126,135
	Naïve Bayes	59
	Ensemble Learning	120
	Transfer Learning	46
	Multitask learning	47,137
	Graph Learning	136
Late	Support Vector Machine	138–143
	Random Forest, Decision Trees	130,140–147
	Regression: LASSO, Logistic, Multivariate, GAM	141,143,147,148
	Deep Learning: CNN, DNN, RNN, AutoEncoders	39,130,140,142,144–146,149–153
	Ensemble learning	154
	Multitask learning	43,47,59,113,120
	XGBoost	143,147,153
	Artificial neural network	39,153,155
	Natural Language Processing	145,153,155
	Regression: GLM, Logistic	140–143,147,155,156
	Graph network	157
	Clustering: KNN, graphs	144,145,157
Mixed	Random Forest	158
	Artificial neural Network, ElasticNet	159

CNN convolutional neural network, *DUN* deep unified networks, *MARS* Multivariate adaptive regression splines, *MLP* multilayer perceptron, *GLM* Generalized Linear Model, *DNN* deep neural network, *GAM* generalized additive model, *iMSF* incomplete Multi-Source Feature, *RNN* Recurrent Neural Network.

the time series information using t-SNE plots[58]. Two papers comprised early fusion using imaging and time series[67,103]. There were two papers that leveraged EHR and genomic information[66,119]. Luo et al. implemented hybrid non-negative matrix factorization (HNMF) to find coherence between phenotypes and genotypes in those suffering from hypertension[119]. One paper leveraged early fusion using imaging and text data[105] and another used EHR, Genomics, Transcriptomics, and Insurance Claims[157].

Intermediate fusion

Intermediate fusion had the second highest number of papers published. 14 used imaging and EHR data[43,59,113,114,118,121,123,125,126,129,131–133,135,137]. Zihni et al. merged the output from a Multilayer Perceptron (MLP) for modeling clinical data and convolutional neural network (CNN) for modeling imaging data into a single full connected final layer to predict stroke[135]. A very similar approach
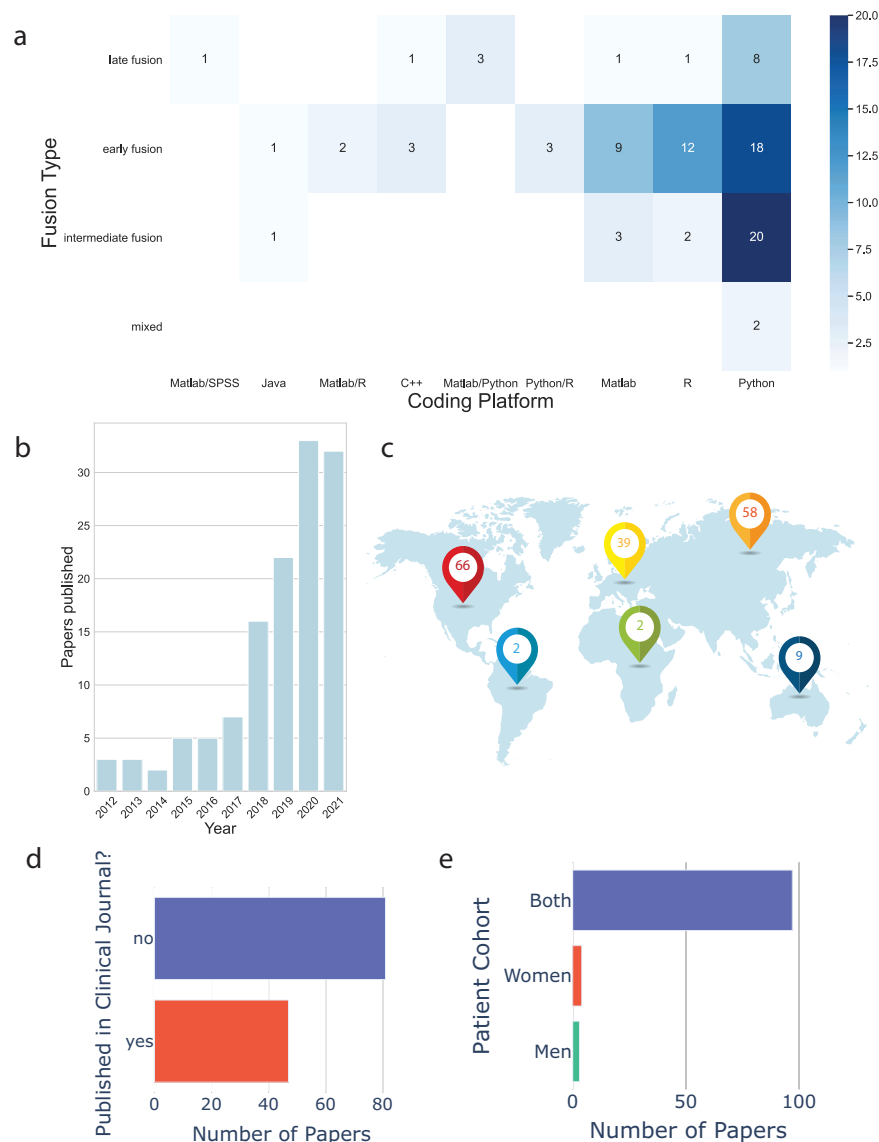
Fig. 4 Meta-data from the review process. a Heat map of fusion type broken down into the coding platforms papers used by summing over paper counts (those that mentioned platform used), the most popular being the Python platform and early fusion. Of note, 37 of the papers did not explicitly mention a platform. **b** Total number of original research papers published in this sphere in the last 10 years. **c** Continental breakdown of author contributions (note some papers have authors from multiple continents). **d** Breakdown of publication type (clinical/non-clinical journal). Less than half (37.6%) of the papers were published in a journal intended for a clinical audience. **e** Sex breakdown of populations studied. Both men and women were represented in the papers, however, the degree of representation varied within an individual studies.

was taken by Tang et al. who used three-dimensional CNNs and merged the layers in the last layer[113]. EHR and text data were fused together in 11 papers[41,44,80,107,109,116,122,126,134,136,142]. Of these, six[41,44,80,122,134,142] used long term short term (LSTM) networks, CNNs, or knowledge-guided CNNs[160] in their fusion of EHR and clinical notes. Chowdhury et al. used graph neural networks and autoencoders to learn meta-embeddings from structured lab test results and clinical notes[107,109]. Pivovarov et al. learned probabilistic phenotypes from clinical notes and medication/lab orders (EHR) data[136]. Two models each employing LDA where data type was treated as a bag of elements and to bring coherence between the two models to identify unique phenotypes. Ye et al. and Shin et al. used concept identifiers via NLP and bag-of-words techniques, respectively, prior to testing a multitude of secondary models[116,126]. In general, clinical notes can provide complementary information to structured EHR data, where natural language processing (NLP) is

often needed to extract such information[161–163]. A few studies were published using imaging and genomic[37,117,120]. Here radiogenomics were used to diagnose attention-deficit/hyperactivity disorder (ADHD), glioblastoma survival, and dementia respectively. Polygenic risk scores were combined with MRI by Yoo et al. who used an ensemble of random forests for ADHD diagnosis[120]. Zhou et al. fused SNPs information together with MRI and positron emission tomography (PET) for dementia diagnosis by learning latent representations (i.e., high-level features) for each modality independently. Subsequently learning joint latent feature representations for each pair of modality combination and then learning the diagnostic labels by fusing the learned joint latent feature representations from the second stage was carried out[37]. Wijethilake used MRI and gene expression profiling, performing recursive feature elimination prior to merging into multiple models SVM, linear regression, and artificial neural network (ANN). The linear

Table 3. Research questions as outlined in Methods.

RQ1	• The literature published in this area as displayed and characterized in the Results' section is one that is of growing and global interest.
	• Fueled by a desire to improve predictive capabilities, relying on complementary and correlative (reinforcing) data. This was found to be the case in the papers surveyed and included in this review, with an increase in 6.4% accuracy.
	• Most common health topics were neurology and cancer. This is likely fostered by curated databases that lend themselves to multi-modality predictions such as Alzheimer's Disease Neuroimaging Initiative[165] and The Cancer Genome Atlas Program[166].
	• Dominance of early data fusion methods likely owe their pervasiveness for three reasons:
	○ 2 modalities over 3 means less work overall in model building and deployment.
	○ EHR and image data do not require extensive digital conversion for models as does text.
	○ Early fusion is built on a single model with a multitude of feature inputs and is typically less computationally complex than is intermediate or late fusion.
	• Seldomly did articles perform comparisons of machine learning findings against their human clinician counterparts.
	• Several did perform comparisons between uni-modal and multi-modal predictions, with the majority having found a consistent improvement in classification accuracy, sensitivity, and specificity[53,113,159] when leveraging multi-modal data.
	• Performance benefits seemingly not limited to a particular subtype of multi-modal strategy that was detectable in our metadata.
	• Genera recommendation that multi-modal data integration be attempted to improve performance and better mirror a human expert by creating a higher validity environment from which to make clinical decisions.
RQ2	• The analysis techniques are varied and currently do not showcase a gold standard machine learning method in the field. This is likely linked to it being a relatively new and emerging field.
	• The varied techniques implemented are highlighted in Table 3.
	• N-cross fold validation was the most common and a robust estimator in the face of bias within a dataset.
	• Strength of generalizability stems from either the dataset set containing multi-site/location patient data to begin with, or using an external dataset from a remote location[167].
RQ3	• Health contexts predominantly impacted by this include Neurology and Cancer.
	• No domain/method laid claim to building translation models via FDA (or equivalent) approval for use in clinical circumstances.
	• Compare models more readily to physician decision makers[168–171]. This will guide the validity of the environments suitable to machine learning and increase adoption and permit FDA approval of these tools[172].

regression model outperformed the other two merged models and any single modality[117]. Wang et al. and Zhang et al. showcased their work in merging imaging and text information[45,46]. Both used LSTM for language modeling a CNN to generate embeddings that were joined together in a dual-attention model. This is achieved by computing a context vector with attended information preserved for each modality resulting in joint learning. Seldom were articles published using: Imaging/EHR/Text[115], Genomic/Text[49], Imaging/Time series[127], Imaging/Text/Time series[47], Imaging/EHR/Genomic[130], Imaging/EHR/Time series[124], EHR/Genomic[128], EHR/Text/Time series[42].

Late fusion

A much smaller number (*n* = 20) of papers used late fusion. Seven of those used imaging and EHR data types[138,139,144,150,151,154,164]. Both Xiong et al. and Yin et al. fed outputs into a CNN to provide a final weighting and decision[150,151]. Three papers were published using a trimodal approach: imaging, EHR and genomic[130,147,148]. Xu et al. and Faris et al. published papers using EHR and text data[146,155]. Faris et al. processed clinical notes using TF-IDF, hashing vectorizer and document embeddings in conjunction with binarized clinical data[155]. Logistic Regression (LR), Random Forest (RF), Stochastic Gradient Descent Classifier (SGD Classifier), and a Multilayer Perceptron (MLP) were applied to both sets of data independently and final outputs of the two models were combined using different schemes: ranking, summation, and multiplication. Two articles were published using imaging and time series[149,152] both of which employed CNNs, one in video information of neonates[149] and the other in chest x-rays[152]. However, they differed in their processing of the time series data. Salekin used a bidirectional CNN and Nishimori used a one-dimensional CNN. Far fewer papers were published using Imaging/EHR/Text[153], EHR/Genomic/Text[145], imaging/EHR, time series/[141], Imaging/Genomic[156], EHR/Genomic[140], and Imaging/Text[39].

Mixed fusion

Two papers performed multiple data fusion architectures[158,159]. Huang et al. created seven different fusion architectures. These included, early, joint, and late fusion. The architecture that performed the best was the late elastic average fusion for the diagnosis of pulmonary embolism using computed tomography and EHR data[159]. Their Late Elastic Average Fusion leveraged an ElasticNet (linear regression with combined L1 and L2 priors that act as regularizers) for EHR variables. El-Sappagh et al. performed early and late fusion to create an interpretable Alzheimer's diagnosis and progression detection model[158]. Their best performing model was one that implemented instance-based explanations of the random forest classifier by using the SHapley Additive exPlanations (SHAP) feature attribution. Despite using clinical, genomic, and imaging data, the most influential feature was found to be the Mini-Mental State Examination.

Clinical relevance

Data fusion may help address sex representation and increase population diversity issues (including minority populations) in health modeling by creating a more representative dataset if one datatype contained more of one sex and another datatype contained more of the other. This reciprocal compensation ability of employing various data sets would also hold true for racial or ethnic diversities.

Less than half (37.6%) of the papers were published in a journal intended for a clinical audience. None of the papers included in the final cohort of studies had created tools for clinical use that had FDA approval. Based on the rising number of papers in this field there is a growing and global need and interest to characterize these findings.

DISCUSSION

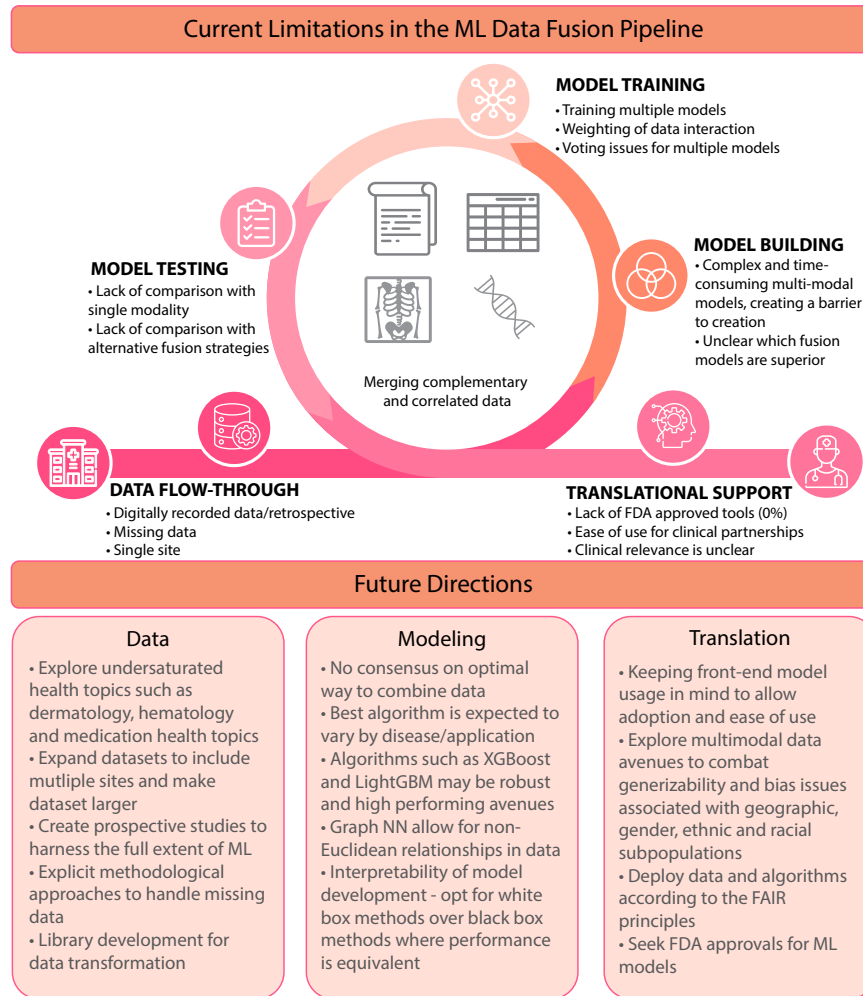Returning to our research questions, we outlined from the inception of this work, we arrive at Table 3.

Fig. 5 Limitations to multimodal fusion in health and proposed future directions of the fields. Limitations to multimodal fusion implementation are stratified by their location in the workflow. These include issues associated with the underlying data, the modeling that arises from that data, and finally how these are ported back to health systems to provide translational decision support.

Many issues were raised in the papers included in this review. The most common reported limitations were cohorts from a single site, small sample sizes, retrospective data, imbalanced samples, handling of missing data, feature engineering, controlling for confounding factors, and interpretation of the models employed. Samples were most often built from a single hospital or academic medical center[148]. Small sample sizes often lead to poor model fitting and generalizability. The median number of unique patients reported across the studies was 658 with a standard deviation of 42,600. This suggests that while some studies were able to leverage large and multi-center cohorts, a great many were not able to do so[70,82,120,131].

Seldom were machine learning investigations on prospective data, an issue endemic in the field[84]. Sample imbalances were often ignored, which results in biased models and misleading performance metrics[75,151]. Missing data were usually ignored by dropping data or imputing, if not dealt with appropriately can skew the results[68,106,173]. More studies need to discuss frequencies and types of missing data[174–177]. Comparison of different imputation methods on the final results should be part of the reporting process[178]. When performing statistical analysis, researchers usually ignored possible confounding factors such as age or gender. Doing so may have major effects on the impact of results[153]. Such possible confounding effects should either be

taken into consideration by the model[179,180] or adjusted for first, prior to reporting model results. Reasonable interpretations of the model and outputs must be presented so that clinicians find the results credible and then use them to provide guidance for treatments. However, most authors did not take the time to interpret the models for clinical audiences. Additionally, how the results may function as a clinical decision support tool. Different types of models warrant different explanations[129,130]. These limitations are highlighted where they occur in the data processing and modeling building pipeline in Fig. 5.

To expedite and facilitate this field, we have outlined several gaps for future research in this field. These are listed in Fig. 5 and explored. Medication/drug topics present an underrepresented area, with only two papers being published in this field[50,66]. Awareness of drug interaction effects is a difficult and growing issue[181–184], particularly in geriatrics, which gave rise to Beer's criteria[185]. Performing multimodal machine learning may offer an earlier detection of adverse events associated with medication misuse that is a result of iatrogenic error, non-compliance, or addiction. Similar justifications as outlined above could be applied to other areas seen as 'under saturated' such as hematology with only one paper[49] and nephrology having just three[41,87,99].

Augmenting clinical decision-making with ML to improve clinical research and outcomes offers positive impacts that have

economic, ethical, and moral ramifications, as it can reduce suffering and save human lives. Multiple studies have now pointed out that if the data an ML model is trained on is biased this often yields bias in the predictions[186,187]. Ensuring multisite, representative data will limit model biases. We also advocate for the creation of open access pipelines/libraries to speed up data conversion to make the technology more widely available[188,189]. Improving accuracy at the expense of complex and time-consuming data transformations may mean the predictive power gained from a multimodal approach is offset by this front-end bottleneck, meaning predictions are no longer temporally relevant or useful.

While incorporating disparate data does lend itself to seemingly better predictions[139], as knowledge around certain diseases accumulates, data fusion in healthcare is an evolving target that warrants proactively adapting to the dynamic landscape[190]. There is no single ML model with ubiquitous applicability. For example, it has been shown in protein-protein interactions that utilization of the XGBoost ensemble algorithm reduces noisy features, maintains the significant raw features, and prevents overfitting[122]. Similarly, LightGBM[191] has the advantages of faster training speed, higher efficiency, lower memory usage, better accuracy[192], and has been consistently outperforming other models[193,194]. Graph neural networks can synthesize new connections leading to drug discovery/targets[122].

In the same vein, models that permit interpretability should always be considered. For example, the Perotte et al.[99] model was not compared with conventional simpler machine learning classifiers, and collective matrix factorization becomes inherently difficult to interpret[79]. Contrast this with the work of Fraccaro et al. whose study of macular degeneration noted their white box performed as well as black box methods implementions[68].

As this field and the datasets associated mature there is work needed to address the tenets of data management: Findability, Accessibility, Interoperability, and Reuse of digital datasets (FAIR)[195]. This entails having metadata that are unique/de-identified and searchable, with open or federated access points (Findability/Accessibility), data that are shared broadly (Interoperable), and finally data that contain accurate and relevant attributes under a clear data usage agreement/license (Reusable). It is imperative there exist a clear definition of outcomes, assessment of biases and interpretability/transparency of results, and limitations inherent in its predictions[196].

Of crucial importance for uptake is that predictions be patient-specific and actionable at a granular level[197]. For example, a 30-day readmission prediction algorithm[106], if implemented, may inform resource management and prompt additional research that may decrease the number of patients re-admitted. Linden et al. developed Deep personalized LOngitudinal convolutional RIsk model (DeepLORI) capable of creating predictions that can be interpreted on the level of individual patients[122]. Leveraging both and clinical and empirically driven information to create meaningful and usable recommendations[136] may improve clinician/end-user under understanding by relating to existing frameworks. Resources such as CRISP-ML provide a framework for moving use cases into more practical applications[198], while efforts to vie for Food and drug administration (FDA) approvals as a tool for use are encouraged to increased adoption.

Deployment of models with user interfaces annotating limitations inherent on those predictions[196] will allow clinical decision makers to interface and implement change accordingly. Follow-through on the aforementioned tasks will push individual fields to create recommendations for subsequent real-world implementations that are relevant, actionable, and transcend regional/subpopulation differences. Limitations of this scoping review include that it is not a systematic review. Therefore, it is possible that some titles that should have been included were missed. As

the primary purpose of this study was to perform scientific paper profiling on multimodal machine learning in health, a critical appraisal of individual methodological quality of the included studies was not performed. However, commentary is provided on the methodological limitations that could have affected their results and impacted their claims. This review offers comprehensive meta-data and evaluation across health domains, immaterial to the type of machine learning or the data used. This work serves as both a summary and steppingstone for future research in this field. Data fusion in health is a growing field of global interest. The topic areas of health that have high frequency relative to others were neurology and cancer, which serve to highlight opportunities for further exploration in understudied topics (hematology, dermatology). Unimodal machine learning is inherently in contrast to current routine clinical practice in which imaging, clinical or genomic data are interpreted in unison to inform accurate diagnosis and warrants further work for ease of use and implementation. Overall, it appears justified to claim that multimodal data fusion increases predictive performance over unimodal approaches (6.4% mean improvement in AUC) and is warranted where applicable. Multimodal machine learning may be a tool leveraged in precision medicine to further subgroup patients' and their unique health fingerprint. Furthermore, as no papers in our review sought FDA approval, we advocate for more efforts into model translation and explore necessities that facilitate that end.

A dashboard resource published in conjunction with this review article is available at: https://multimodal-ml-health.herokuapp.com/. This dashboard was created as an interactive infographic-based display of the major findings presented in this paper. To foster future work, a drop-down menu was created to help researchers filter the underlying data file of titles based on the specific overarching health topic by selection. This will facilitate the location of relevant papers.

METHODS

Search strategy and selection criteria

Inclusion requirements were: (a) original research article; (b) published within the last 10 years (encompassing years 2011–2021); (c) published in English; and (d) on the topic of multi-modal or multi-view using machine learning in health for diagnostic or prognostication applications.'Multi-modal' or'multi-view' for our context means the multiple data sources were not of the same type. For example, while a paper using CT and MRI may be considered multi-modal imaging; however, under our criteria it would be considered uni-modal (i.e., only included imaging). Exclusions for the purposes of this review were: (a) scientific articles not published in English; (b) commentaries or editorials; or (c) other review articles. Papers were also excluded if the data were not human-derived. We also excluded papers where the fusion already occurred at the data generation stage, such as spatial transcriptomics producing integrated tissue imaging and transcriptomics data[199–201]. All papers underwent a 2-person verification for inclusion in the manuscript.

Search strings were established via literature searches and domain expertize. Additional keywords were identified based on key word co-occurrence matrices established from the abstracts of the previously included articles. Figure 6a displays the search strings, where an individual string would include one keyword from each column, this was performed for all combinations of search strings. An overview of the inclusion/exclusion process is noted in Fig. 6b and follows the standard set by PRISMA extension for scoping reviews[202].
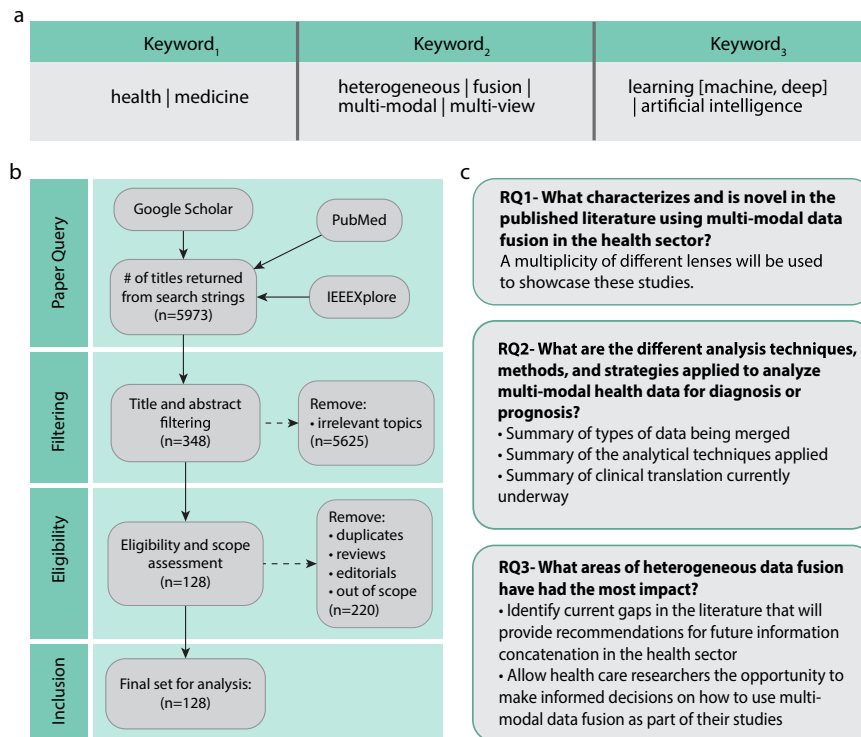
Fig. 6 Overview of our PRIMSA-SCR process. a Health-related keyword, Multimodal-related keyword, machine learning-related keywords, |: or. For example, "health + heterogeneous data + machine learning" would be one of the search strings. **b** Overview of study inclusion process. **c** Research questions posed.

Data extracted

Information garnered from the articles included title, year published, FDA approval of the tool, whether published in a clinical journal, author affiliations, number of authors, locations (continents), and abstract. Health topic(s) addressed were extracted, as well as the broader medical topic(s) that encompass the disease. For example, lung cancer would be the specific disease in question. It arises from the topics of Cancer and Respiratory according to our classification. Health topic classification was overseen and reviewed by a medical doctor to ensure accuracy. As multiple health topics often encompassed a single health disease addressed in each paper, several papers are counted twice. This is true when being mapped from the right side of the Sankey plot to the specific health disease in the middle.

We recorded and extracted the number of different modalities and the divisions (i.e., text/image vs EHR/genomic/time series) used. The objective of each paper was extracted in a 1–2 sentence summary along with the keyword (if available). Patient characterization in the studies was performed by ascertaining the number of unique patients in the cohort and patient sex (i.e., Men/women/both or not mentioned).

Computational information extracted included: (a) the coding interface(s) used in data processing/analysis, (b) machine learning type, (c) data merging technique (early, intermediate, late), and (d) types of machine learning algorithms used. Whether validation was performed (yes/no), the statistical tests run, the nature of the validation, and outcomes measures were all recorded for each paper. The significance, impact, and limitations of each paper were extracted by reviewing the primary findings and limitations as noted in the papers.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

REFERENCES

1. Nils, B. H. & Sabine, S. The ethics of machine learning-based clinical decision support: an analysis through the lens of professionalisation theory. *BMC Med. Ethics* **22**, 112 (2021).
2. Sanchez-Pinto, L. N., Luo, Y. & Churpek, M. M. Big data and data science in critical care. *Chest* **154**, 1239–1248 (2018).
3. Miotto, R., Wang, F., Wang, S., Jiang, X. & Dudley, J. T. Deep learning for healthcare: Review, opportunities, and challenges. *Brief Bioinform.* **19**, 1236–1246 (2017).
4. Timothy, J. W. D. et al. Machine learning of threedimensional right ventricular motion enables outcome prediction in pulmonary hypertension: A cardiac MR imaging study. *Radiology* **283**, 381–390 (2017).
5. Gigi, F. S., Gregory, R. H., Bradley, J. N. & Jun, D. Predicting breast cancer risk using personal health data and machine learning models. *PLoS One* **14**, e0226765 (2019).
6. Michael, F., Justin, B. E.-T. & Elizabeth, S. Clinical and public health implications of 2019 endocrine society guidelines for diagnosis of diabetes in older adults. *Diabetes Care* **43**, 1456–1461 (2020).
7. Gambhir, S. S., Ge, T. J., Vermesh, O. & Spitler, R. Toward achieving precision health. *Sci. Transl. Med.* **10**, eaao3612 (2018).
8. Schüssler-Fiorenza Rose, S. M. et al. A longitudinal big data approach for precision health. *Nat. Med.* **25**, 792–804 (2019).
9. Feero, W. G. Introducing "genomics and precision health". *JAMA* **317**, 1842–1843 (2017).
10. Kellogg, R. A., Dunn, J. & Snyder, M. P. Personal omics for precision health. *Circulation Res.* **122**, 1169–1171 (2018).
11. Thapa, C. & Camtepe, S. Precision health data: Requirements, challenges and existing techniques for data security and privacy. *Comput. Biol. Med.* **129**, 104130 (2021).
12. Pranata, S. et al. Precision health care elements, definitions, and strategies for patients with diabetes: A literature review. *Int. J. Environ. Res. Public Health* **18**, 6535 (2021).
13. Shih Cheng, H., Anuj, P., Saeed, S., Imon, B. & Matthew, P. L. Fusion of medical imaging and electronic health records using deep learning: A systematic review and implementation guidelines. *npj Digital Med.* **3**, 136 (2020).

14. Weixian, H., Kaiwen, T., Jinlong, H., Ziye, Z. & Shoubin, D. A review of fusion methods for omics and imaging data. In *IEEE/ACM Trans Comput Biol Bioinform* (IEEE, 2022).

15. Federico, C. A review of data fusion techniques. *Scientific World J.* **2013**, 704504 (2013).

16. Alan, N. S., Christopher, L. B. & Franklin, E. W. Revisions to the JDL data fusion model. *Sens. Fusion.: Architectures, Algorithms, Appl. III* **3719**, 430 (1999). publisher = SPIE.

17. Erik, M.-M.-R., Antonio, A. A., Ramon, F. B. & Enrique, G.-C. Improved accuracy in predicting the best sensor fusion architecture for multiple domains. *Sensors* **21**, 7007 (2021).

18. Ahmad, F. S., Luo, Y., Wehbe, R. M., Thomas, J. D. & Shah, S. J. Advances in machine learning approaches to heart failure with preserved ejection fraction. *Heart Fail Clin.* **18**, 287–300 (2022).

19. Erik, B. et al. Machine learning/artificial intelligence for sensor data fusion-opportunities and challenges. In *IEEE Aerospace and Electronic Systems Magazines* Vol. 36, 80–93 (IEEE, 2021).

20. Li, Y., Wu, X., Yang, P., Jiang, G. & Luo, Y. Machine learning applications in diagnosis, treatment, and prognosis of lung cancer. Preprint at https://arxiv.org/abs/2203.02794 (2022).

21. Kohane, I. S. et al. What every reader should know about studies using electronic health record data but may be afraid to ask. *J. Med. Internet Res.* **23**, e22219 (2021).

22. Andres, C. et al. Machine-learning Prognostic Models from the 2014–16 Ebola Outbreak: Data-harmonization challenges, validation strategies, and mHealth applications. *EClinicalMedicine* **11**, 54–64 (2019).

23. Afshin, J., Jean Pierre, P. & Johanne, M.-P. Machine-learning-based patient-specific prediction models for knee osteoarthritis. *Nat. Rev. Rheumatol.* **15**, 49–60 (2019).

24. Luo, Y., Ahmad, F. S. & Shah, S. J. Tensor factorization for precision medicine in heart failure with preserved ejection fraction. *J. Cardiovasc. Transl. Res.* **10**, 305–312 (2017).

25. Luo, Y., Wang, F. & Szolovits, P. Tensor factorization toward precision medicine. *Briefings Bioinform.* **18**, 511–514 (2016).

26. Rasmussen L. et al. Considerations for improving the portability of electronic health record-based phenotype algorithms. In *Proceedings of 2019 AMIA Annual Symposium* 2019 (2019).

27. Zhong, Y. et al. Characterizing design patterns of EHR-driven phenotype extraction algorithms. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* 1143–1146 (IEEE, 2018).

28. Rasmussen, L. V. et al. Solutions for unexpected challenges encountered when integrating research genomics results into the EHR. *ACI Open* **4**, e132–e5 (2020).

29. Shang, N. et al. Making work visible for electronic phenotype implementation: Lessons learned from the eMERGE network. *J. Biomed. Inf.* **99**, 103293 (2019).

30. Dana, L., Tulay, A. & Christian, J. Multimodal data fusion: An overview of methods, challenges, and prospects. *Proc. IEEE* **103**, 1449–1477 (2015).

31. Wei, C., Yungui, H., Brendan, B. & Simon, L. The utility of including pathology reports in improving the computational identification of patients. *J. Pathol. Inform.* **7**, 46 (2016).

32. Yubraj, G., Ramesh Kumar, L. & Goo Rak, K. Prediction and classification of Alzheimer's disease based on combined features from apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers. *Front. Comput. Neurosci.* **13**, 72 (2019).

33. Xia An, B., Xi, H., Hao, W. & Yang, W. Multimodal data analysis of Alzheimer's disease based on clustering evolutionary random forest. *IEEE J. Biomed. Health Inform.* **24**, 2973–2983 (2020).

34. Ariana, A. et al. Non-negative matrix factorization of multimodal MRI, fMRI and phenotypic data reveals differential changes in default mode subnetworks in ADHD. *Neuroimage* **102**, 207–219 (2014).

35. Chao, G., Luo, Y. & Ding, W. Recent advances in supervised dimension reduction: A survey. *Mach. Learn. Knowl. Extraction* **1**, 341–358 (2019).

36. Pillai, P. S. L. T. Y. Fusing heterogeneous data for Alzheimer's disease classification. *Stud. Health Technol. Inf.* **216**, 731–735 (2015).

37. Tao, Z., Kim Han, T., Xiaofeng, Z. & Dinggang, S. Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Hum. Brain Mapp.* **40**, 1001–1016 (2019).

38. Robi, P. Ensemble based systems in decision making. *IEEE Circuits Syst. Mag.* **6**, 21–44 (2006).

39. Francesco, C. & Marwa, M. Multimodal temporal machine learning for bipolar disorder and depression recognition. *Pattern Analysis Appl.* **25**, 493–504 (2021).

40. Durrant-Whyte, H. F. Sensor models and multisensor integration. *Int. J. Robot. Res.* **7**, 97–113 (1988).

41. Xu, Z. et al. Identification of predictive sub-phenotypes of acute kidney injury using structured and unstructured electronic health record data with memory networks. *J. Biomed. Inform.* **102**, 103361 (2019).

42. Dongdong, Z., Changchang, Y., Jucheng, Z., Xiaohui, Y. & Ping, Z. Combining structured and unstructured data for predictive models: A deep learning approach. *BMC Med. Inform. Decis. Mak.* **20**, 280 (2020).

43. Shaker, E.-S., Tamer, A., Islam, S. M. R. & Kyung Sup, K. Multimodal multitask deep learning model for Alzheimer's disease progression detection based on time series data. *Neurocomputing* **412**, 197–215 (2020).

44. Haiyang, Y., Li, K. & Feng Qiang, X. Multimodal temporal-clinical note network for mortality prediction. *J. Biomed. Semantics* **12**, 3 (2021).

45. Zizhao, Z., Pingjun, C., Manish, S. & Lin, Y. TandemNet: Distilling knowledge from medical images using diagnostic reports as optional semantic references. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, **10435**, 320–328 (Springer, Cham, 2017).

46. Xiaosong, W., Yifan, P., Le, L., Zhiyong, L. & Ronald, M. S. TieNet: Text-Image embedding network for common thorax disease classification and reporting in chest X-rays. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 9049–9058 (IEEE Computer Society, 2018).

47. Syed Arbaaz, Q., Sriparna, S., Mohammed, H., Gael, D. & Erik, C. Multitask representation learning for multimodal estimation of depression level. *IEEE Intell. Syst.* **34**, 45–52 (2019).

48. Jordan, Y., William, Y. & Philipp, T. Multimodal skin lesion classification using deep learning. *Exp. Dermatol.* **27**, 1261–1267 (2018).

49. Kai, Z et al. MLMDA: A machine learning approach to predict and validate MicroRNA-disease associations by integrating of heterogenous information sources. *J. Transl. Med.* **17**, 260 (2019).

50. Sivan, K. et al. Predicting risk for Alcohol Use Disorder using longitudinal data with multimodal biomarkers and family history: A machine learning study. *Mol. Psychiatry* **26**, 1133–1141 (2021).

51. Mara Ten, K et al. MRI predictors of amyloid pathology: Results from the EMIF-AD Multimodal Biomarker Discovery study. *Alzheimer's Res. Ther.* **10**, 100 (2018).

52. Isamu, H. et al. Radiogenomics predicts the expression of microRNA-1246 in the serum of esophageal cancer patients. *Sci. Rep.* **10**, 2532 (2020).

53. Jesus, J. C., Jianhua, Y. & Daniel, J. M. Enhancing image analytic tools by fusing quantitative physiological values with image features. *J. Digit Imaging* **25**, 550–557 (2012).

54. Kevin Bretonnel, C. et al. Methodological issues in predicting pediatric epilepsy surgery candidates through natural language processing and machine learning. *Biomed. Inform. Insights* **8**, BII.S38308 (2016).

55. Weiming, L. et al. Predicting Alzheimer's disease conversion from mild cognitive impairment using an extreme learning machine-based grading method with multimodal data. *Front. Aging Neurosci.* **12**, 77 (2020).

56. Micah, C. et al. Predicting rehospitalization within 2 years of initial patient admission for a major depressive episode: A multimodal machine learning approach. *Transl. Psychiatry* **9**, 285 (2019).

57. Jongin, K. & Boreom, L. Identification of Alzheimer's disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine. *Hum. Brain Mapp.* **39**, 3728–3741 (2018).

58. Hélène De, C. et al. Wearable monitoring and interpretable machine learning can objectively track progression in patients during cardiac rehabilitation. *Sensors (Switz.)* **20**, 1–15 (2020).

59. Tamer, A., Shaker, E.-S. & Jose, M. A. Robust hybrid deep learning models for Alzheimer's progression detection. *Knowledge-Based Syst.* **213**, 106688 (2021).

60. Jeungchan, L. et al. Machine learning-based prediction of clinical pain using multimodal neuroimaging and autonomic metrics. *Pain* **160**, 550–560 (2019).

61. Uttam, K., Goo Rak, K. & Horacio, R.-G. An efficient combination among sMRI, CSF, cognitive score, and APOE ϵ 4 biomarkers for classification of AD and MCI using extreme learning machine. *Comput. Intell. Neurosci.* **2020**, 8015156 (2020).

62. Bo, C., Mingxia, L., Heung, I. S., Dinggang, S. & Daoqiang, Z. Multimodal manifold-regularized transfer learning for MCI conversion prediction. *Brain Imaging Behav.* **9**, 913–926 (2015).

63. Kevin, H., Ulrike, L., Markus, M. & Katja, B-B. Separating generalized anxiety disorder from major depression using clinical, hormonal, and structural MRI data: A multimodal machine learning study. *Brain Behavior* **7**, e00633 (2017).

64. Fayao, L., Luping, Z., Chunhua, S. & Jianping, Y. Multiple kernel learning in the primal for multimodal alzheimer's disease classification. *IEEE J. Biomed. Health Inform.* **18**, 984–990 (2014).

65. Diego, C.-B. et al. Robust ensemble classification methodology for I123-Ioflupane SPECT images and multiple heterogeneous biomarkers in the diagnosis of Parkinson's disease. *Front. Neuroinform.* **12**, 53 (2018).

66. Yi, Z. et al. Predicting adverse drug reactions of combined medication from heterogeneous pharmacologic databases. *BMC Bioinform.* **19**, 517 (2018).

67. Chin Po, C., Susan Shur Fen, G. & Chi Chun, L. Toward differential diagnosis of autism spectrum disorder using multimodal behavior descriptors and executive functions. *Comput. Speech Lang.* **56**, 17–35 (2019).

68. Paolo, F. et al. Combining macula clinical signs and patient characteristics for age-related macular degeneration diagnosis: A machine learning approach Retina. *BMC Ophthalmol.* **15**, 10 (2015).

69. Benjamin, D. W. et al. Early identification of epilepsy surgery candidates: A multicenter, machine learning study. *Acta Neurol. Scand.* **144**, 41–50 (2021).

70. Xia An, B., Wenyan, Z., Lou, L. & Zhaoxu, X. Detecting risk gene and pathogenic brain region in EMCI using a novel GERF algorithm based on brain imaging and genetic data. *IEEE J. Biomed. Health Inform.* **25**, 3019–3028 (2021).

71. Prashanth, R., Sumantra Dutta, R., Pravat, K. M. & Shantanu, G. High-accuracy detection of early Parkinson's disease through multimodal features and machine learning. *Int. J. Med. Inf.* **90**, 13–21 (2016).

72. Ali, A.-M. et al. Machine learning for localizing epileptogenic-zone in the temporal lobe: Quantifying the value of multimodal clinical-semiology and imaging concordance. *Front. Digital Health* **3**, 559103 (2021).

73. Baiying, L. et al. Assessment of liver fibrosis in chronic hepatitis B via multimodal data. *Neurocomputing* **253**, 169–176 (2017).

74. Larry, H. et al. Multimodal tensor-based method for integrative and continuous patient monitoring during postoperative cardiac care. *Artif. Intell. Med.* **113**, 102032 (2021).

75. Ivo, D. D. et al. Predictive big data analytics: A study of Parkinson's disease using large, complex, heterogeneous, incongruent, multi-source and incomplete observations. *PLoS One* **11**, e0157077 (2016).

76. Eleftherios, T., Ioannis, S., Apostolos, H. K. & Kostas, M. Deep radiotranscriptomics of non-small cell lung carcinoma for assessing molecular and histology subtypes with a data-driven analysis. *Diagnostics* **11**, 2383 (2021).

77. Hua, W. et al. Identifying disease sensitive and quantitative trait-relevant biomarkers from multidimensional heterogeneous imaging genetics data via sparse multimodal multitask learning. *Bioinformatics* **28**, i127–36 (2012).

78. Xia, B., Xi, H., Yiming, X. & Hao, W. A novel CERNNE approach for predicting Parkinson's Disease-associated genes and brain regions based on multimodal imaging genetics data. *Med. Image Anal.* **67**, 101830 (2021).

79. Vijay, H. et al. Predicting complications in critical care using heterogeneous clinical data. *IEEE Access* **4**, 7988–8001 (2016).

80. Wang, H., Li, Y., Khan, S. A. & Luo, Y. Prediction of breast cancer distant recurrence using natural language processing and knowledge-guided convolutional neural network. *Artif. Intell. Med.* **110**, 101977 (2020).

81. Zeng, Z. et al. Identifying breast cancer distant recurrences from electronic health records using machine learning. *J. Healthcare Inform. Res.* **3**, 283–299 (2019).

82. Kautzky, A. et al. Machine learning classification of ADHD and HC by multimodal serotonergic data. *Transl. Psychiatry* **10**, 104 (2020).

83. Nhat Trung, D. et al. Distinct multivariate brain morphological patterns and their added predictive value with cognitive and polygenic risk scores in mental disorders. *NeuroImage: Clin.* **15**, 719–731 (2017).

84. Niha, B. et al. Radiogenomic analysis of hypoxia pathway is predictive of overall survival in Glioblastoma. *Sci. Rep.* **8**, 7 (2018).

85. Jan, C. P. et al. Combining multimodal imaging and treatment features improves machine learning-based prognostic assessment in patients with glioblastoma multiforme. *Cancer Med.* **8**, 128–136 (2019).

86. Hao, Z. et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. *J. Neurooncol* **142**, 299–307 (2019).

87. Namyong, P. et al. Predicting acute kidney injury in cancer patients using heterogeneous and irregular data. *PLoS One* **13**, e0199839 (2018).

88. Wei Liang, T., Chee Kong, C., Sim Heng, O. & Alvin Choong Meng, N. Ensemble-based regression analysis of multimodal medical data for osteopenia diagnosis. *Expert Syst. Appl.* **40**, 811–819 (2013).

89. Sébastien, T., Yasser, I.-M., José María, M.-P., Alan, C. E. & Louis De, B. Defining a multimodal signature of remote sports concussions. *Eur. J. Neurosci.* **46**, 1956–1967 (2017).

90. Gianluca, B. et al. Multimodal predictive modeling of endovascular treatment outcome for acute ischemic stroke using machine-learning. *Stroke* **51**, 3541–3551 (2020).

91. Yiming, X. et al. Explainable dynamic multimodal variational autoencoder for the prediction of patients with suspected central precocious puberty. *IEEE J. Biomed. Health Inform.* **26**, 1362–1373 (2021).

92. Alan, D. K. et al. Mixture model framework for traumatic brain injury prognosis using heterogeneous clinical and outcome data. *IEEE J. Biomed. Health Inform.* **26**, 1285–1296 (2021).

93. Casper, R. et al. Preoperative risk stratification in endometrial cancer (ENDORISK) by a Bayesian network model: A development and validation study. *PLoS Med.* **17**, e1003111 (2020)

94. Tommaso, G. et al. SARS-COV-2 comorbidity network and outcome in hospitalized patients in Crema, Italy. *PLoS One* **16**, e0248498 (2021).

95. Huan, Q. et al. Machine-learning radiomics to predict early recurrence in perihilar cholangiocarcinoma after curative resection. *Liver Int.* **41**, 837–850 (2021).

96. Ramon, C. et al. Alzheimer's disease risk assessment using large-scale machine learning methods. *PLoS One* **8**, e77949 (2013).

97. Aleksei, T. et al. Multimodal machine learning-based knee osteoarthritis progression prediction from plain radiographs and clinical data. *Sci. Rep.* **9**, 20038 (2019).

98. Michael, J. D. et al. Development and validation of a novel automated Gleason grade and molecular profile that define a highly predictive prostate cancer progression algorithm-based test. *Prostate Cancer Prostatic Dis.* **21**, 594–603 (2018).

99. Perotte, A., Ranganath, R., Hirsch, J. S., Blei, D. & Elhadad, N. Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. *J. Am. Med. Inf. Assoc.* **22**, 872–880 (2015).

100. Lei, Y., Yalin, W., Paul, M. T., Vaibhav, A. N. & Jieping, Y. Multi-source feature learning for joint analysis of incomplete multiple heterogeneous neuroimaging data. *Neuroimage* **61**, 622–632 (2012).

101. Yanbo, X., Siddharth, B., Shriprasad, R. D., Kevin, O. M. & Jimeng, S. RAIM: Recurrent attentive and intensive model of multimodal patient monitoring data. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2565–2573 (Association for Computing Machinery, 2018).

102. Yixue, H., Mohd, U., Jun, Y., Hossain, M. S. & Ahmed, G. Recurrent convolutional neural network based multimodal disease risk prediction. *Future Gener. Computer Syst.* **92**, 76–83 (2019).

103. Shinichi, G., et al. Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. *Nat. Commun.* **12**, 2726 (2021).

104. Rui, Y. et al. Richer fusion network for breast cancer classification based on multimodal data. *BMC Med. Inform. Decis. Mak.* **21**, 134 (2021).

105. Li, Y., Wang, H. & Luo, Y. A comparison of pre-trained vision-and-language models for multimodal representation learning across medical images and reports. *2020 IEEE Int. Conf. Bioinform. Biomedicine (BIBM)* 2020 1999–2004 (2020).

106. Sara Bersche, G. et al. 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.* **18**, 44 (2018).

107. Chowdhury, S., Zhang, C., Yu, P. S. & Luo, Y. Mixed pooling multi-view attention autoencoder for representation learning in healthcare. Preprint at https://arxiv.org/abs/1910.06456 (2019).

108. Ilan, S. et al. An unsupervised learning approach to identify novel signatures of health and disease from multimodal data. *Genome Med.* **12**, 7 (2020).

109. Chowdhury, S., Zhang, C., Yu, P. S. & Luo, Y. Med2Meta: Learning representations of medical concepts with meta-embeddings. *HEALTHINF* **2020**, 369–376 (2020).

110. Subramanian V, Do MN, Syeda-Mahmood T. Multimodal fusion of imaging and genomics for lung cancer recurrence prediction, *IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 804–808, (2020).

111. Michele, D. et al. Combining heterogeneous data sources for neuroimaging based diagnosis: Re-weighting and selecting what is important. *Neuroimage* **195**, 215–231 (2019).

112. Yiwen, M., William, S., Michael, K. O. & Corey, W. A. Bidirectional representation learning from transformers using multimodal electronic health record data to predict depression. *IEEE J. Biomed. Health Inform.* **25**, 3121–3129 (2021).

113. Xing, T. et al. Elaboration of a multimodal MRI-based radiomics signature for the preoperative prediction of the histological subtype in patients with non-small-cell lung cancer. *BioMed. Engineering Online* **19**, 5 (2020).

114. Kathleen, C. F., Kristina Lundholm, F., Marie, E., Fredrik, Ö & Dimitrios, K. Predicting MCI status from multimodal language data using cascaded classifiers. *Front. Aging Neurosci.* **11**, 205 (2019).

115. Amir Hossein, Y. et al. Multimodal mental health analysis in social media. *PLoS One* **15**, e0226248 (2020).

116. Ye, J., Yao, L., Shen, J., Janarthanam, R. & Luo, Y. Predicting mortality in critically ill patients with diabetes using machine learning and clinical notes. *BMC Med Inf. Decis. Mak.* **20**, 1–7 (2020).

117. Navodini, W., Mobarakol, I. & Hongliang, R. Radiogenomics model for overall survival prediction of glioblastoma. *Med Biol. Eng. Comput* **58**, 1767–1777 (2020).

118. Solale, T. et al. A distributed multitask multimodal approach for the prediction of Alzheimer's disease in a longitudinal study. *Neuroimage* **206**, 116317 (2020).

119. Luo, Y. et al. Integrating hypertension phenotype and genotype with hybrid non-negative matrix factorization. *Bioinformatics* **35**, 1395–403 (2019).

120. Jae Hyun, Y., Johanna Inhyang, K., Bung Nyun, K. & Bumseok, J. Exploring characteristic features of attention-deficit/hyperactivity disorder: Findings from multi-modal MRI and candidate genetic data. *Brain Imaging Behav.* **14**, 2132–2147 (2020).

121. Chao, T., Baoyu, L., Jun, L. & Zhigao, Z. A Deep automated skeletal bone age assessment model with heterogeneous features learning. *J. Med. Syst.* **42**, 249 (2018).

122. Cheng, C. et al. Improving protein–protein interactions prediction accuracy using XGBoost feature selection and stacked ensemble classifier. *Comput. Biol. Med.* **123** (2020).

123. Xueyi, Z. et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat. Commun.* **11**, 1236 (2020).

124. Juan Camilo, V.-C. et al. Multimodal assessment of Parkinson's disease: A deep learning approach. *IEEE J. Biomed. Health Inform.* **23**, 1618–1630 (2019).

125. Ping, Z. et al. Deep-learning radiomics for discrimination conversion of Alzheimer's disease in patients with mild cognitive impairment: A study based on 18F-FDG PET imaging. *Front. Aging Neurosci.* **13**, (2021).

126. Shin, J., Li, Y. & Luo, Y. Early prediction of mortality in critical care setting in sepsis patients using structured features and unstructured clinical notes. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* 2885–2890 (IEEE, 2021).

127. Jayachitra, V. P., Nivetha, S., Nivetha, R. & Harini, R. A cognitive IoT-based framework for effective diagnosis of COVID-19 using multimodal data. *Biomed. Signal Processing Control* **70**, 102960 (2021).

128. Thomas, L. et al. An explainable multimodal neural network architecture for predicting epilepsy comorbidities based on administrative claims data. *Front. Artificial Intelligence* **4**, 610197 (2021).

129. Alan Baronio, M., Carla Diniz Lopes, B. & Silvio Cesar, C. Computer-aided diagnosis of hepatocellular carcinoma fusing imaging and structured health data. *Health Inform. Sci. Syst.* **9**, 20 (2021)

130. Janani, V., Li, T., Hamid Reza, H. & May, D. W. Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci. Rep.* **11**, 3254 (2021).

131. Cam Hao, H. et al. Bimodal learning via trilogy of skip-connection deep networks for diabetic retinopathy risk progression identification. *Int. J. Med. Inform.* **132**, 103926 (2019).

132. Yucheng, T. et al. Prediction of type II diabetes onset with computed tomography and electronic medical records. *Lect. Notes Comput. Sci.* **12445**, 13–23 (2020).

133. Rui, Y. et al. Integration of multimodal data for breast cancer classification using a hybrid deep learning method. *Lect. Notes Computer Sci.* **11643**, 460–469 (2019).

134. Batuhan, B. & Mehmet, T. Improving clinical outcome predictions using convolution over medical entities with multimodal learning. *Artif. Intell. Med.* **117**, 102112 (2021).

135. Esra, Z. et al. Multimodal fusion strategies for outcome prediction in Stroke. In *HEALTHINF 2020 - 13th International Conference on Health Informatics, Proceedings; Part of 13th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2020* 421–428 (2020).

136. Rimma, P. et al. Learning probabilistic phenotypes from heterogeneous EHR data. *J. Biomed. Inform.* **58**, 156–165 (2015).

137. Leon, M. A. et al. Modeling longitudinal imaging biomarkers with parametric Bayesian multi-task learning. *Hum. Brain Mapp.* **40**, 3982–4000 (2019).

138. Paris Alexandros, L. et al. Heterogeneity and classification of recent onset psychosis and depression: A multimodal machine learning approach. *Schizophr. Bull.* **47**, 1130–1140 (2021).

139. Nikolaos, K. et al. Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: A multimodal, multisite machine learning analysis. *JAMA Psychiatry* **75**, 1156–1172 (2018).

140. Dongdong, S., Minghui, W. & Ao, L. A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data. *IEEE/ACM Trans. Comput Biol. Bioinform.* **16**, 841–850 (2019).

141. Karen, S. A. et al. A machine-learning framework for robust and reliable prediction of short- and long-term treatment response in initially antipsychotic-naïve schizophrenia patients based on multimodal neuropsychiatric data. *Translational Psychiatry* **10**, 276 (2020).

142. Sun, M. et al. Early prediction of acute kidney injury in critical care setting using clinical notes and structured multivariate physiological measurements. *Stud. Health Technol. Inf.* **264**, 368–372 (2019).

143. Dennis, S. R., Simuni, T. & Luo, Y. A predictive model for Parkinson's disease reveals candidate gene sets for progression subtype. *2020 IEEE Int. Conf. Bioinforma. Biomedicine (BIBM) 2020*, 417–420 (2020).

144. Ming, X. et al. Accurately differentiating COVID-19, other viral infection, and healthy individuals using multimodal features via late fusion learning. *J. Med. Internet Res.* **23** (2021).

145. Min, C., Yixue, H., Kai, H., Lin, W. & Lu, W. Disease prediction by machine learning over big data from healthcare communities. *IEEE Access* **5**, 8869–8879 (2017).

146. Keyang, X. et al. Multimodal machine learning for automated ICD coding. *Proc. Mach. Learn. Res.* **106**, 1–17 (2019).

147. Liuqing, Y. et al. Deep learning based multimodal progression modeling for Alzheimer's disease. *Stat. Biopharma. Res.* **13**, 337–343 (2021).

148. Peng, L. et al. A radiogenomics signature for predicting the clinical outcome of bladder urothelial carcinoma. *Eur. Radio.* **30**, 547–557 (2020).

149. Md Sirajus, S. et al. Multimodal spatio-temporal deep learning approach for neonatal postoperative pain assessment. *Comput. Biol. Med.* **129**, 104150 (2021).

150. Jian, X. et al. Multimodal machine learning using visual fields and peripapillary circular OCT scans in detection of glaucomatous optic neuropathy. *Ophthalmology* **129**, 171–180 (2021).

151. Dai, Y., Yiqi, Z., Yang, W., Wenpu, Z. & Xiaoming, H. Auxiliary diagnosis of heterogeneous data of Parkinson's disease based on improved convolution neural network. *Multimed. Tools Appl.* **79**, 24199–24224 (2020).

152. Makoto, N. et al. Accessory pathway analysis using a multimodal deep learning model. *Sci. Rep.* **11**, 8045 (2021).

153. Wenhuan, Z., Anupam, G. & Daniel, H. H. On the application of advanced machine learning methods to analyze enhanced, multimodal data from persons infected with covid-19. *Computation* **9**, 1–15 (2021).

154. Jeremy, A. T., Kit, M. L. & Marta, I. G. Multi-dimensional predictions of psychotic symptoms via machine learning. *Hum. Brain Mapp.* **41**, 5151–5163 (2020).

155. Hossam, F., Maria, H., Mohammad, F., Haya, E. & Alaa, A. An intelligent multimodal medical diagnosis system based on patients' medical questions and structured symptoms for telemedicine. *Inform. Med. Unlocked* **23**, 100513 (2021).

156. Md Ashad, A. et al. A kernel machine method for detecting higher order interactions in multimodal datasets: Application to schizophrenia. *J. Neurosci. Methods* **309**, 161–174 (2018).

157. Luo, Y. et al. A multidimensional precision medicine approach identifies an autism subtype characterized by dyslipidemia. *Nat. Med.* **26**, 1375–1379 (2020).

158. Shaker, E.-S., Jose, M. A., Islam, S. M. R., Ahmad, M. S., Kyung Sup, K. A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Sci. Rep.* **11**, 2660 (2021).

159. Shih Cheng, H., Anuj, P., Roham, Z., Imon, B. & Matthew, P. L. Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: A case-study in pulmonary embolism detection. *Sci. Rep.* **10**, 22147 (2020).

160. Yao, L., Mao, C. & Luo, Y. Clinical text classification with rule-based features and knowledge-guided convolutional neural networks. *BMC Med. Inf. Decis. Mak.* **19**, 71 (2019).

161. Velupillai, S. et al. Using clinical Natural Language Processing for health outcomes research: Overview and actionable suggestions for future advances. *J. Biomed. Inf.* **88**, 11–19 (2018).

162. Luo, Y., Uzuner, Ö. & Szolovits, P. Bridging semantics and syntax with graph algorithms—state-of-the-art of extracting biomedical relations. *Brief. Bioinform.* **18**, 160–178 (2016).

163. Zeng, Z., Deng, Y., Li, X., Naumann, T. & Luo, Y. Natural language processing for EHR-based computational phenotyping. *IEEE/ACM Trans. Comput Biol. Bioinform.* **16**, 139–153 (2018).

164. Nikolaos, K. et al. Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. *JAMA Psychiatry* **78**, 195–209 (2021).

165. Petersen, R. C. et al. Alzheimer's disease neuroimaging initiative (ADNI): Clinical characterization. *Neurology* **74**, 201–209 (2010).

166. Weinstein, J. N. et al. The cancer genome atlas pan-cancer analysis project. *Nat. Genet.* **45**, 1113 (2013).

167. Christopher, J. K., Alan, K., Mustafa, S., Greg, C. & Dominic, K. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* **17**, 195 (2019).

168. Michael, A. M. et al. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. *J. Am. Acad. Dermatol.* **78**, 270–7.e1 (2018).

169. Jannin, P. G. C. & Gibaud, B. Medical Applications of NDT Data Fusion (Springer, 2001).

170. Guang, Y., Qinghao, Y. & Jun, X. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. *Inf. Fusion* **77**, 29–52 (2022).

171. Zhang, Z., & Sejdić, E., Radiological images and machine learning: trends, perspectives, and prospects. *Comput. Biol. Med.* **108**, 354–370 (2019).

172. David, L., Enrico, C., Jessica, C., Parina, S. & Farah, M. How machine learning is embedded to support clinician decision making: An analysis of FDA-approved medical devices. *BMJ Health Care Inform.* **28**, e100301 (2021).

173. Luo, Y., Szolovits, P., Dighe, A. S. & Baron, J. M. Using machine learning to predict laboratory test results. *Am. J. Clin. Pathol.* **145**, 778–788 (2016).

174. Thakur, S., Choudhary, J. & Singh, D. P. *Systems* 435–443 (Springer, 2021).

175. Luo, Y., Szolovits, P., Dighe, A. S. & Baron, J. M. 3D-MICE: Integration of cross-sectional and longitudinal imputation for multi-analyte longitudinal clinical data. *J. Am. Med. Inform. Assoc. (JAMIA)* **25**, 645–653 (2017).

176. Xue, Y., Klabjan, D. & Luo, Y. Mixture-based multiple imputation model for clinical data with a temporal dimension. *2019 IEEE International Conference on Big Data (Big Data) 2019* 245–252 (IEEE, 2019).

177. Cao, W. et al. Brits: Bidirectional Recurrent Rmputation for Time Series. *NeurIPS,* **31** (2018).

178. Luo, Y. Evaluating the state of the art in missing data imputation for clinical data. *Brief. Bioinform.* **23**, bbab489 (2022).

179. Zhao, Q., Adeli, E. & Pohl, K. M. Training confounder-free deep learning models for medical applications. *Nat. Commun.* **11**, 1–9 (2020).

180. Luo, Y. & Mao, C. ScanMap: Supervised confounding aware non-negative matrix factorization for polygenic risk modeling. In *Machine Learning for Healthcare Conference; 2020: PMLR* 27–45 (2020).

181. Kalavathy, R. & Suresh, R. M. Pharmacovigilance from electronic medical records to report adverse events. *J. Chem. Pharm. Sci.* **2015**, 188–191 (2015).

182. Luo, Y. et al. Natural language processing for EHR-based pharmacovigilance: A structured review. *Drug Saf.* https://doi.org/10.1007/s40264-017-0558-6 (2017).

183. Segura Bedmar, I., Martínez, P. & Herrero Zazo, M. Semeval-2013 task 9: Extraction of drug–drug interactions from biomedical texts (ddiextraction 2013). *2013: Association for Computational Linguistics* (2013).

184. Hammann, F. & Drewe, J. Data mining for potential adverse drug–drug interactions. *Expert Opin. Drug Metab. Toxicol.* **10**, 665–671 (2014).

185. Donna, M. F. et al. American geriatrics society 2019 updated AGS beers criteria for potentially inappropriate medication use in older adults. *J. Am. Geriatr. Soc.* **67**, 674–694 (2019).

186. Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447–53. (2019).

187. Wang, H. et al. Using machine learning to integrate sociobehavioral factors in predicting cardiovascular-related mortality risk. *Stud. Health Technol. Inf.* **264**, 433–437 (2019).

188. Christof, E. Open source software in industry. *IEEE Softw.* **25**, 52–53 (2008).

189. Robert, M. S. Why develop open-source software? The role of non-pecuniary benefits, monetary rewards, and open-source licence type. *Oxf. Rev. Economic Policy* **23**, 605–619 (2007).

190. Luo, Y., Wunderink, R. G. & Lloyd-Jones, D. Proactive vs reactive machine learning in health care: Lessons from the COVID-19 pandemic. *JAMA* **327**, 623–624 (2022).

191. Ke, G. et al. LightGBM: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process Syst.* **30** (2017).

192. Derara Duba, R., Taye Girma, D., Achim, I. & Worku Gachena, N. Diagnosis of diabetes mellitus using gradient boosting machine (Lightgbm). *Diagnostics* **11**, 1714 (2021).

193. Xiaolei, S., Mingxi, L. & Zeqian, S. A novel cryptocurrency price trend forecasting model based on LightGBM. *Financ. Res. Lett.* **32**, 101084 (2020).

194. Dongzi, J., Yiqin, L., Jiancheng, Q., Zhe, C. & Zhongshu, M. SwiftIDS: Real-time intrusion detection system based on LightGBM and parallel intrusion detection mechanism. *Comput. Secur.* **97**, 101984 (2020).

195. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. *Nature* **3**, 160018 (2016).

196. Weissler, E. H., et al. The role of machine learning in clinical research: Transforming the future of evidence generation. *Trials* **22**, 537 (2021).

197. Pratik, S. et al. Artificial intelligence and machine learning in clinical development: A translational perspective. *npj Digital Med.* **2**, 69 (2019).

198. Inna, K. & Simeon, S. Interpretability of machine learning solutions in public healthcare: The CRISP-ML approach. *Front. Big Data* **4**, 660206 (2021).

199. Marx, V. Method of the Year: Spatially resolved transcriptomics. *Nat. Methods* **18**, 9–14 (2021).

200. Zeng, Z., Li, Y., Li, Y. & Luo, Y. Statistical and machine learning methods for spatially resolved transcriptomics data analysis. *Genome Biol.* **23**, 1–23. (2022).

201. Longo, S. K., Guo, M. G., Ji, A. L. & Khavari, P. A. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* **22**, 627–644 (2021).

202. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann. Intern. Med.* **169** 467–473 (2018).

AUTHOR CONTRIBUTIONS
A.K. led the review process, performed data extraction, performed the computation analysis, figure generation, writing, and dashboard creation. H.W., Y.L., S.D., and M.H. performed data extraction. Z.X. synthesized limitations of the studies. F.W. and F.C. performed proof reading and content curation. Y.L. conceived the review, oversaw the review process, and provided necessary feedback, proof reading, and content curation.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-022-00712-8.

Correspondence and requests for materials should be addressed to Yuan Luo.

Reprints and permission information is available at http://www.nature.com/reprints

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