



# Systematic Mapping Study of AI/Machine Learning in Healthcare and Future Directions

Gaurav Parashar<sup>1</sup> · Alka Chaudhary<sup>1</sup> · Ajay Rana<sup>1</sup>

Received: 4 August 2021 / Accepted: 1 September 2021 / Published online: 16 September 2021  
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd 2021

## Abstract

This study attempts to categorise research conducted in the area of: *use of machine learning in healthcare*, using a systematic mapping study methodology. In our attempt, we reviewed literature from top journals, articles, and conference papers by using the keywords *use of machine learning in healthcare*. We queried Google Scholar, resulted in 1400 papers, and then categorised the results on the basis of the objective of the study, the methodology adopted, type of problem attempted and disease studied. As a result we were able to categorize study in five different categories namely, interpretable ML, evaluation of medical images, processing of EHR, security/privacy framework, and transfer learning. In the study we also found that most of the authors have studied cancer, and one of the least studied disease was epilepsy, evaluation of medical images is the most researched and a new field of research, Interpretable ML/Explainable AI, is gaining momentum. Our basic intent is to provide a fair idea to future researchers about the field and future directions.

**Keywords** Machine learning (ML) · Transfer learning (TL) · Interpretable ML · Electronic health records (EHR) · Security framework · Privacy framework · Healthcare

## Introduction

Artificial intelligence (AI) can be defined as a field in which the machine demonstrates intelligence by learning itself. It can be done by deploying various techniques & algorithms to understand human intelligence but does not confine to—John McCarthy. Even though if we do not specifically program the machine and still it can automatically learn and improve itself this defines an intelligent behaviour of machine. Machine learning (ML) is a specific field of AI which relates to techniques that can automatically learn from experience.

The use of machine learning in healthcare had shown many promising solutions which had created confidence in the field. Researchers had used ICT tools with ML in developing solutions for increasing the effectiveness of the earlier methods or procedures. The field of healthcare had also shown tremendous improvement after the use of Big Data, ICT, and AI/machine learning (ML) in precision and speed. These tools have greatly helped physicians and healthcare professionals in their day-to-day working, research, testing the effect of biomedicine on humans using simulations. Every single detail of the patient gets recorded by the doctors with the other information like clinical notes, prescriptions, medical test results, diagnosis, X-rays, MRI scan, sonographic images, etc. This data becomes huge repository of information, which if churned, could give us better insights of treatment, fruitful suggestions and recommendations in diagnosis, progressive pattern of one disease could be correlated to another disease and may lead to new procedure for treatment of a disease and many more. There may be a chance that healthcare professional overlooked a symptom, which if not addressed early could lead to loss of life. Therefore, tools like AI/ML could help in better healthcare services.

---

This article is part of the topical collection “Intelligent Systems” guest edited by Geetha Ganesan, Lalit Garg, Renu Dhir, Vijay Kumar and Manik Sharma.

---

✉ Gaurav Parashar  
gauravparashar24@gmail.com  
Alka Chaudhary  
achaudhary4@amity.edu  
Ajay Rana  
ajayphdmba@gmail.com

<sup>1</sup> AIIT, AMITY University, Noida, Uttar Pradesh, India

The use of tools like IBM Watson<sup>1</sup> and Google DeepMind [1] have shown impressive results in healthcare. On top of these tools researchers and developers have designed applications, which harness the capabilities of these tools, to provide personalised patient care, better drug discovery, and improved healthcare organisational performance. According to Wired<sup>2</sup> Google DeepMind was used to identify protein structures associated with SARS-CoV-2 and understand how the virus functions. One of the oldest scientific puzzle of 'protein folding problem' was also solved and paved the way for faster development of drugs, better treatment by Google DeepMind. Other contributions in the field of healthcare are use of association rules (AR) which helped analyse malaria in Brazil [2], according to [3] diagnosing images of X-rays revealed respiratory condition of the patient and helped in better healthcare services.

ML can be applied to varied fields like defence, automation, finance, automobile, and manufacturing in performing tasks like classification, clustering, and forecasting. It can be categorised into three types supervised, unsupervised, and reinforcement learning.

In supervised learning, algorithms learn from the labeled datasets and prepare a model. After training, we give data to the model, which the model has not seen earlier and belongs to the same category so that it can correctly classify it. In unsupervised learning, algorithms themselves learn, by analysing data and the model then prepares a model, which can be used to correctly cluster the elements. Lastly, in reinforcement learning, machine learns itself from its mistakes or maximising rewards and by reducing penalties.

In this paper, we aim to categorise papers based on the healthcare and machine learning. With the use of AI and ML in healthcare, there have been significant changes in the life of healthcare professionals. The accuracy of medical diagnostic has increased, healthcare professionals have an assistant on which they can rely, they can predict diseases like pneumonia, cancer, heart diseases, Tumour, COVID-19, and many more with better accuracy & precision than before.

In this paper, we attempt to categorise research done in use of machine learning in healthcare. *According to best of our knowledge this type of categorisation has not been done earlier.* This attempt will become the basis of future research in the field. We attempt to categorise them on the basis of the objective of the study, methodology adopted, type of problem attempted, etc. We discussed in section "Research Methodology", in section "Literature Survey", in section "Results" and section "Conclusion".

<sup>1</sup> <https://www.healthcareglobal.com/technology-and-ai-3/four-ways-which-watson-transforming-healthcare-sector>

<sup>2</sup> <https://www.wired.co.uk/article/ai-healthcare-boom-deepmind>

## Research Methodology

This section describes the systematic mapping procedure adopted to study the use of AI/ML in the healthcare domain. This study was conducted using the keywords "Machine Learning" OR "Healthcare". The search was conducted on Google scholar and considered only results from Nature, Wiley periodicals, Elsevier, Taylor and Francis, IEEE transactions, ACM, SVN, IET, and ArXiv. Following steps were carried out: (1) Definition of research questions (2) conduct search for primary studies (3) screening of papers for inclusion and exclusion (4) keywording using abstracts (5) data extraction and mapping of studies. The above steps were proposed by [4].

## The Systematic Mapping Process

We have adopted the systematic mapping process from [4] and applied it to the study conducted on *use of ML in the healthcare*.

Systematic Mapping Process is a well defined, comprehensive overview study done on a particular research topic. According to [5] it helps researchers do verifiable, unbiased literature review, find research gaps by critical examination of research literature, helps collate evidence, reduce reviewer selection bias & publication bias with transparent inclusion and exclusion criteria.

The process is described here:

1. We first define research questions and scope of the study.
2. With respect to the questions framed from the previous step now search is conducted and literature is collected.
3. Proper screening is done to check whether the selected literature is related to the research question and scope of the research.
4. Abstract and keywords are scanned for critical survey of the content
5. In a spreadsheet, collected data is mapped with the RQs.

In the following (see Fig. 1) we had shown the process which we had implemented in the study.

## Definition of Research Questions

The main intent of the study is to find out the use of ML in the field of healthcare. To start with we have formulated three research questions (see Table 1) which are based on the topic of the study. Major goals of systematic mapping study are:-

1. Find review of the research area
2. Find the quantity, type of research, and result

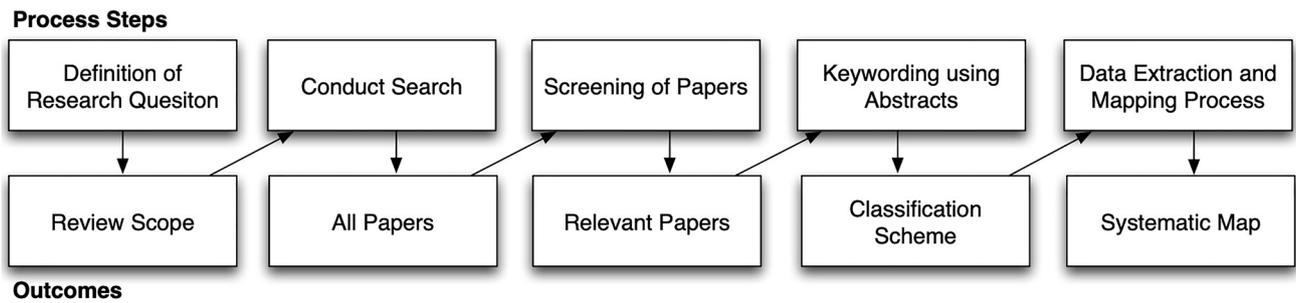


Fig. 1 The systematic mapping process

Table 1 Research questions for systematic map

Question number	Research question
RQ1:	What type of research was conducted in the area <i>use of AI/ML in Healthcare</i> ?
RQ2:	What are the broad categories of papers under the field?
RQ3:	What are the different diseases which have been studied with their paper count?

3. Find journals of the published research topic

Therefore, on the basis of the above goals following research questions have been formed.

RQ1: What type of research has been conducted in the field of *use of AI/ML in healthcare*? *Rationale:* This question aims to find the type of research, which has been conducted under the field of the healthcare domain. We need to find out papers published under the topic.

RQ2: What are the broad categories of papers published under the topic? *Rationale:* The rationale for this question arises from the outcome of the RQ1. FQ1 gives research papers, then we need to find out the broad category under which the paper lie.

RQ3: What are the different diseases which have been studied and total number of total paper published under it? *Rationale:* After categorising the papers, we need to find out different diseases which are being studied in the research done by other researchers. The main intent is to find the least studied disease, which can become a starting point for new research.

Conduct Search for Primary Studies

To conduct the search we followed the steps:

1. Prepare the search string w.r.t to different databases(as described in Table 2). Since we have used only Google scholar therefore we had used a broad search string to cover all papers containing the keywords *healthcare* and *machine learning*.

2. Execute the search and collect the results (see Fig. 2).
3. Categorise the results by studying the papers and grouping them together on the basis of the disease studied (As mentioned in the Table 6) & intent of the paper (As mentioned in the Table 5).

We took 1400 search results from google scholar and transferred them to spreadsheet based on the query mentioned in Table 2. This data of around 1400 entries will be further drilled down in next section by excluding the entries which are not related to the study.

Screening of Papers for Inclusion and Exclusion (Relevant Papers)

In this step we exclude all the papers that are not relevant in the study. By this we also mean that the papers which are not related to the RQs (Refer Table 5), papers which do not from Nature, Wiley periodicals, Elsevier, Taylor and Francis, IEEE transactions, ACM, SVN, IET, and ArXiv are excluded from the final list.

Using the above criteria we retained those entries, which are based on Inclusion criteria (Refer Table 3). After using the above exclusion criteria we drilled down the entries which we finally considered were 42.

Table 2 Search string for Google scholar

Search string
“healthcare” OR ”machine learning”

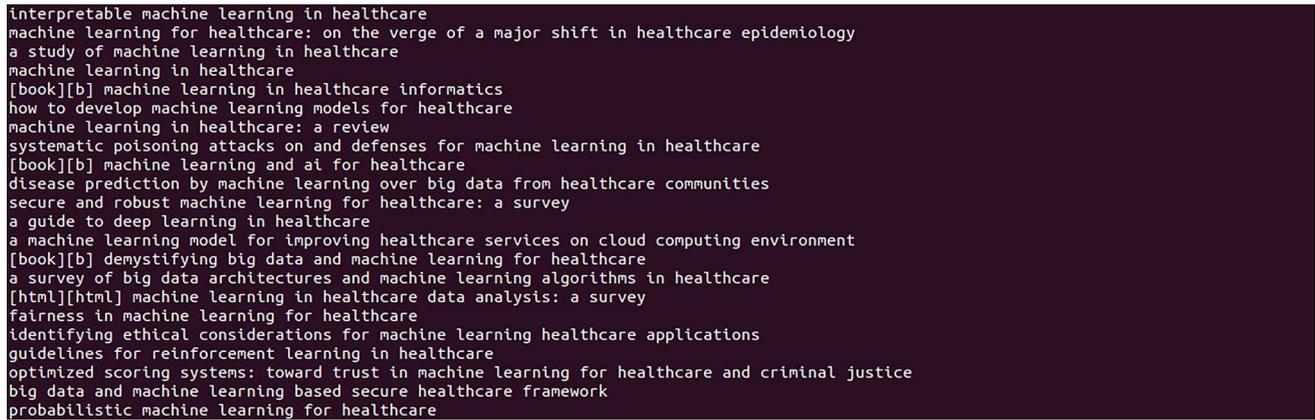


Fig. 2 Raw text from Google scholar results

**Keywording Using Abstracts**

For our study we followed the systematic process of classifying the results from Google Scholar. For Keywording we followed these steps:

**Data Extraction and Mapping of Studies**

We collected all the information in a spreadsheet with the

**Table 3** Inclusion and exclusion criteria

Use of machine learning in healthcare system map
Inclusion: books, papers, technical reports, white papers, periodicals from Nature, Wiley periodicals, Elsevier, Taylor and Francis, IEEE transactions, ACM, SVN, IET, and Arxiv, paper must include a disease, paper must be based on a technique
Exclusion: Papers other than the list in inclusion are excluded, papers not related to use of ML in healthcare, survey paper, research

**Table 4** Paper categories

S. No.	Context	Description
1.	Interpretable ML	It is a field of research which refers to methods and models that make the behaviour and predictions of ML systems understandable to human
2.	Evaluation of Medical Images	The field of research refers to evaluation of medical images like X-rays, computed tomography, magnetic resonance imaging, ultrasound, etc
3.	Processing of EHR	An Electronic Health Record are electronic record of patient medical history. It contains information like medical observations, prescriptions, medical tests and results, vital signs, past history, etc
4.	Security/Privacy Framework	Privacy framework focus on providing privacy and security to patient health records
5.	Transfer Learning	It is a machine learning method in which the model developed for one problem is used as starting point for another problem. This way we transfer learning of one domain to another

1. The result collected from the previous step are analysed by surveying abstract.
2. Abstract are surveyed for keywords and content. Then context of the study is evaluated.
3. Group the result on the basis of context and keywords (Refer Table 4).

information like s.no., paper title, abstract, keywords, year of publication, authors, name of publisher, name of periodical/journal/conference, major findings, major shortcomings. After that we mapped the RQs (see Table 1) to each entry.

**Table 5** Paper categories

S. No.	Category	Number of papers
1.	Interpretable ML	7
2.	Evaluation of Medical Images	17
3.	Processing of EHR	14
4.	Security/Privacy Framework	7
5.	Transfer Learning	4

**Table 6** Disease categories

S. No.	Disease	Frequency
1.	Cancer	38
2.	Heart	31
3.	Diabetes	27
4.	Mental	18
5.	Covid-19	13
6.	Lung	9
7.	Pneumonia	4
8.	Alzheimer	3
9.	Parkinson	2
10.	Epilepsy	2

## Literature Survey

### Interpretable Machine Learning

Interpretable models are those which explains itself. Interpretable models are linear regression, logistic regression and decision trees. For instance, if we use decision tree model then we can easily extract decision rules as explanations for the model.

In [6] authors referred to the use of ML in healthcare with an emphasis on Interpretability. Interpretable ML refers to models which can provide rationale on predictions made by the model. The basic impediment in the adoption of ML in healthcare is its BlackBox nature. Since we have to develop ML as a tool that can act as an assistant to physicians, therefore, we need to make its output more explainable. Mere providing metrics like AUC, recall, precision, F-Score may not suffice. We need to develop more interpretable models that themselves can provide explanations of their predictions. The authors [7] proposed a model which adds important value to features and make the output interpretable. Authors [8] developed reasoning through the use of visual indicators making the model interpretable. In [9] authors proposed an interpretable tree from a decision forest making understandable by humans. As proposed in [10, 11] interpretable ML models helps

develop a reasonable and data-driven decision support system that results in personalised decisions.

Authors [12] applied deep learning on medical data of patients for developing interpretable predictions for decision support.

### Evaluation of Medical Images

In this category, the authors discussed evaluation of medical images for better diagnosis using machine learning models.

In [13–17] authors used deep learning models, Neural Network to classify different diseases, organ segmentation and compared it with the diagnosis of health care professionals for diagnostic accuracy. In [18] authors proposed a novel colour deconvolution for stain separation and colour normalisation of images. In [19] authors performed a comparison of five colour normalisation algorithms and found stain colour normalisation algorithms performed better, which had high stain segmentation accuracy and low computational complexity. In their review paper authors [20] did a comparison of different image segmentation methods and related studies with medical imaging for AI-assisted diagnosis of COVID-19. In [21] authors explained AI, ML, DL, and CNN and the use of these techniques in imaging. [22] discussed image enhancements method with noise suppression by enhancing low light regions.

### Processing of Electronic Health Record (EHR)

In this category, we had compiled papers that had processed electronic health records of patients.

In the paper [17] authors proposed diagnostic of pneumonia in a resource-constrained environment. The authors of [23–25] discussed the processing of electronic health records and used ML algorithms to categorise disease. The authors [26] trained their proposed model on large dataset and performed regression and classification to check their effectiveness and accuracy. In [27], a medical recommendation system was proposed using Fast Fourier transformation coupled with a machine learning ensemble model. The model uses this model for disease risk prediction to provide medical recommendations like medical tests and other recommendations for chronic heart disease patients. In [28] authors proposed the use of graphical structure of electronic health records and find hidden structure from it. In [29] proposed a model that provides help to physicians to evaluate the quality of evidence for better decision making. Authors used risk of bias assessment in textual data using Natural Language Processing.

## Security/Privacy Framework

Under this category, we will summarised papers related to the security and privacy framework for safeguarding health records transferred over network or internet.

Authors of [30] researched on novel design of smart and secure healthcare information system by adopting machine learning. It also employed advanced security mechanism to handle the big data of the healthcare industry. This framework used many security tools to secure the data like encryption, monitoring the activity, access control, and many other mechanism. This paper [31] discussed the privacy-preserving collaborative model using ML tools for medical diagnosis systems.

Most of the privacy protection methods are centralized. There is a need for a decentralized system that can help in mitigating several challenges like single-point-of-failure, modification of records, privacy preservation, improper information exchange that may lead to risk of patient's life. To protect, many researchers have proposed different algorithms [32–35]. Models like VERTIGO, GLORE, and WebDISCO were designed for privacy preservation and predictive modelling. These models aimed to preserve privacy by sending partially-trained machine learning models rather than patient data. This way the information is preserved, and develop trust between different parties.

Many other distributed privacy-preserving models were developed those were based on Blockchain technology. They use the technology to update models as in Blockchain like ModelChain, EXPLORER, Distributed Autonomous Online Learning sequentially.

Secure multiparty computation(SMC) for privacy preservation that do computations on encrypted data with personally identifiable information had opened a new dimension. Data is a very precious commodity, therefore techniques like privacy preserving scoring of Tree Ensembles [36] are designed to provide a framework that provides cryptographic protocols for sending data securely.

## Transfer Learning

In this category, we summarised research papers related to transfer learning. Transfer learning is a technique in which we gain knowledge from one problem and use the same knowledge to solve different but related problem. In [37] authors proposed a technique for handling missing data using transfer learning perspective. The proposed classifier learn weights and then complete portion of the dataset and then transfer it to the target domain. In [38] authors used transfer learning approach to predict breast cancer using model trained on a task some other task. A model trained on ImageNet databases containing 1.2 million images is used as feature extractor. The model is combined with other

components to perform classification. [39, 40] uses data generated by different wearable devices using federated learning, and then builds machine learning model by transfer learning. The study was applied to diagnose Parkinson's disease.

## Results

After the systematic mapping process we got categories of research literature as mentioned in Table 5. This table describes category and total number of papers under the category.

From Table 5 we can clearly observe that most research is being done in Processing of Medical Images this might be due to the availability of the dataset for research purpose. In case of Processing of EHR, which is second most researched category, might be again due to availability of the dataset. In case of Interpretable ML, since it is a new field and slowly gaining momentum therefore, researchers are taking interest because this gives a rationale to the outcome of the model result. This is a very important attribute, when it comes to certain domains where high stakes are at risk. Like for example in healthcare, defence, and finance. Lastly, In case of Transfer Learning, it is a field which talks about using the domain knowledge of one domain and use it in another related domain. So according to us researchers use this technique to apply for testing the results. Therefore it has a very limited number of research.

From Table 6 it is clearly evident that most researched disease is Cancer, which is 38 and Pneumonia with frequency as 4, Alzheimer as 3, Parkinson as 2 and Epilepsy as 2 are the least researched diseases. These results have been extracted from 1400 papers downloaded from Google Scholar.

## Conclusion

In this paper, we have provided a brief overview of the directions of research in the healthcare domain using Machine Learning. As described earlier, these papers can show researchers path where they can work. The result is based on literature review done on around 1400 papers and filtered down to 42 papers. As described in section “[Literature Survey](#)”, we categorised the research into 5 broad areas and found most of the research is done in the field of Evaluation of Medical Images in which authors researched many diseases like cancer, heart disease, COVID-19, Parkinson, etc. Authors used different kinds of dataset like images, voice, and electronic health records. Using these dataset they predicted these diseases using machine learning/AI. As described in section “[Processing of Electronic](#)

**Health Record (EHR)**” second major contribution is done in this category. We would like to conclude that in section “**Interpretable Machine Learning**” very little research has been done so this area can be chosen for further research.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Powles J, Hodson H. Google DeepMind and healthcare in an age of algorithms. *Health Technol.* 2017;7(4):351.
- Baroni L, Salles R, Salles S, Guedes G, Porto F, Bezerra E, Barcellos C, Pedrosa M, Ogasawara E. An analysis of malaria in the Brazilian Legal Amazon using divergent association rules. *J Biomed Inf.* 2020;108:103512.
- Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, et al. 2017. [arXiv:1711.05225](https://arxiv.org/abs/1711.05225).
- Petersen K, Feldt R, Mujtaba S, Mattsson M. In: 12th international conference on evaluation and assessment in software engineering (EASE) 12; 2008. pp. 1–10.
- Haddaway NR, Westgate MJ. Predicting the time needed for environmental systematic reviews and systematic maps. *Conserv Biol.* 2019;33(2):434.
- Ahmad MA, Eckert C, Teredesai A. In: Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics; 2018. pp. 559–60.
- Lundberg S, Lee SI. 2017. [arXiv:1705.07874](https://arxiv.org/abs/1705.07874).
- Yu F, Ip HH. Semantic content analysis and annotation of histological images. *Comput Biol Med.* 2008;38(6):635.
- Sagi O, Rokach L. Explainable decision forest: transforming a decision forest into an interpretable tree. *Inf Fusion.* 2020;61:124.
- Stiglic G, Kocbek P, Fijacko N, Zitnik M, Verbert K, Cilar L. Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdiscipl Rev Data Min Knowl Disc.* 2020;10(5):e1379.
- Ribeiro MT, Singh S, Guestrin C. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining; 2016. pp. 1135–44.
- Rebane J, Samsten I, Papapetrou P. Exploiting complex medical data with interpretable deep learning for adverse drug event prediction. *Artif Intell Med.* 2020;109:101942.
- Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, Mahendiran T, Moraes G, Shamdas M, Kern C, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Dig Health.* 2019;1(6):e271.
- Panayides AS, Amini A, Filipovic ND, Sharma A, Tsafaris SA, Young A, Foran D, Do N, Golemati S, Kurc T, et al. AI in medical imaging informatics: current challenges and future directions. *IEEE J Biomed Health Inf.* 2020;24(7):1837.
- Oktay O, Ferrante E, Kamnitsas K, Heinrich M, Bai W, Caballero J, Cook SA, De Marvao A, Dawes T, Oregan DP, et al. Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. *IEEE Trans Med Imaging.* 2017;37(2):384.
- Jeyaraj PR, Nadar ERS. Deep Boltzmann machine algorithm for accurate medical image analysis for classification of cancerous region. *Cogn Comput Syst.* 2019;1(3):85.
- Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, Yang D, Myronenko A, Anderson V, Amalou A, et al. A systematic review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. *Nature Commun.* 2020;11(1):1.
- Zheng Y, Jiang Z, Zhang H, Xie F, Shi J, Xue C. Adaptive color deconvolution for histological WSI normalization. *Comput Methods Progr Biomed.* 2019;170:107.
- Hoffman RA, Kothari S, Wang MD. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society, IEEE; 2014. pp. 194–7.
- Feng S, et al. 2020. [arXiv:2004.02731](https://arxiv.org/abs/2004.02731).
- Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine learning and deep learning in medical imaging: intelligent imaging. *J Med Imaging Radiat Sci.* 2019;50(4):477.
- Xia W, Chen EC, Peters T. Endoscopic image enhancement with noise suppression. *Healthcare Technol Lett.* 2018;5(5):154.
- Capotorti A. Probabilistic inconsistency correction for misclassification in statistical matching, with an example in health care. *Int J Gener Syst.* 2020;49(1):32.
- Li JP, Haq AU, Din SU, Khan J, Khan A, Saboor A. Heart disease identification method using machine learning classification in e-healthcare. *IEEE Access.* 2020;8:107562.
- Naydenova E, Tsanas A, Casals-Pascual C, De Vos M. In: 2015 IEEE global humanitarian technology conference (GHTC), IEEE; 2015. pp. 377–84.
- Haq AU, Li JP, Memon MH, Malik A, Ahmad T, Ali A, Nazir S, Ahad I, Shahid M, et al. Feature selection based on L1-norm support vector machine and effective recognition system for Parkinson disease using voice recordings. *IEEE Access.* 2019;7:37718.
- Zhang J, Lafta RL, Tao X, Li Y, Chen F, Luo Y, Zhu X. Coupling a fast fourier transformation with a machine learning ensemble model to support recommendations for heart disease patients in a telehealth environment. *IEEE Access.* 2017;5:10674.
- Choi E, Xu Z, Li Y, Dusenberry M, Flores G, Xue E, Dai A. In: Proceedings of the AAAI conference on artificial intelligence, vol. 34; 2020. pp. 606–13.
- Pereira RG, Castro GZ, Azevedo P, Tôrres L, Zuppo I, Rocha T, Júnior AAG. In: 2020 IEEE 33rd international symposium on computer-based medical systems (CBMS), IEEE; 2020. pp. 1–6.
- Kaur P, Sharma M, Mittal M. Big data and machine learning based secure healthcare framework. *Procedia Comput Sci.* 2018;132:1049.
- Wang F, Zhu H, Liu X, Lu R, Hua J, Li H, Li H. Privacy-preserving collaborative model learning scheme for E-healthcare. *IEEE Access.* 2019;7:166054.
- Wu Y, Jiang X, Kim J, Ohno-Machado L. Grid Binary LOGistic REGression (GLORE): building shared models without sharing data. *J Am Med Inf Assoc.* 2012;19(5):758.
- Kuo TT, Ohno-Machado L. Modelchain: decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks. 2018.
- Wang S, Jiang X, Wu Y, Cui L, Cheng S, Ohno-Machado L. Expectation propagation logistic regression (explorer): distributed privacy-preserving online model learning. *J Biomed Inf.* 2013;46(3):480.
- Li Y, Jiang X, Wang S, Xiong H, Ohno-Machado L. Vertical grid logistic regression (vertigo). *J Am Med Inf Assoc.* 2016;23(3):570.
- Fritchman K, Saminathan K, Dowsley R, Hughes T, De Cock M, Nascimento A, Teredesai A. In: 2018 IEEE international conference on big data (Big Data); 2018. pp. 2413–22. <https://doi.org/10.1109/BigData.2018.8622627>.

37. Wang G, Lu J, Choi KS, Zhang G. A transfer-based additive LS-SVM classifier for handling missing data. *IEEE Trans Cybern.* 2018;50(2):739.
38. Dey N, Das H, Naik B, Behera HS. *Big data analytics for intelligent healthcare management.* Cambridge: Academic Press; 2019.
39. Chen Y, Qin X, Wang J, Yu C, Gao W, Chen Y, Qin X, Wang J, Yu C, Gao W. Fedhealth: a federated transfer learning framework for wearable healthcare. *IEEE Intell Syst.* 2020;35(4):83.
40. Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, Yang D, Myronenko A, Anderson V, Amalou A, et al. A systematic

review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. *Nature Commun.* 2020;11(1):1.

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