Heart Failure Diagnosis, Readmission, and Mortality Prediction Using Machine Learning and Artificial Intelligence Models

Aixia Guo 1 · Michael Pasque 2 · Francis Loh 1 · Douglas L. Mann 3 · Philip R. O. Payne 1

Accepted: 23 October 2020 / Published online: 31 October 2020 © The Author(s) 2020

Abstract

Purpose of Review One in five people will develop heart failure (HF), and 50% of HF patients die in 5 years. The HF diagnosis, readmission, and mortality prediction are essential to develop personalized prevention and treatment plans. This review summarizes recent findings and approaches of machine learning models for HF diagnostic and outcome prediction using electronic health record (EHR) data.

Recent Findings A set of machine learning models have been developed for HF diagnostic and outcome prediction using diverse variables derived from EHR data, including demographic, medical note, laboratory, and image data, and achieved expert-comparable prediction results.

Summary Machine learning models can facilitate the identification of HF patients, as well as accurate patient-specific assessment of their risk for readmission and mortality. Additionally, novel machine learning techniques for integration of diverse data and improvement of model predictive accuracy in imbalanced data sets are critical for further development of these promising modeling methodologies.

Keywords Heart failure (HF) · Machine learning · Deep learning · Artificial intelligence · Readmission · Mortality

Introduction

Cardiovascular diseases (CVDs), which cause over 18.9 million deaths globally each year, are the number 1 cause of death, responsible for approximately 31% of all health-related deaths worldwide [1, 2]. Heart failure (HF) accounts for a large portion of this CVD morbidity and mortality, as well as an equally large portion of related healthcare expense [2]. One in five people will develop HF in their lifetime, and about 50% of these HF patients will die within 5 years [3]. In the management of this expanding HF patient population, the accurate prediction of HF outcomes is critical to effective prevention and treatment, as well as to the reduction of the burdensome expenditure of related healthcare dollars. The importance of accurate outcome prediction is accentuated by the impact of HF readmissions, which will cost Medicare approximately 17 billion dollars expended on the approximately 20% of patients who are readmitted within 30 days of HF discharge [4].

Expansive implementation of the EHR has led to a revolution in the introduction of demographic characteristics, genetic profiles, medical treatments, diagnostic notes, laboratory results, and image data of individual patients into an electronic format that facilitates access and use in “big data” research investigations. The feasibility of truly personalized and precision medicine is dependent upon the management of the vast quantities of EHR data that has become available and is critical to the development of these models. The sheer volume and heterogeneous nature of EHR data have raised new challenges regarding the integration and analysis of this data. Therefore, machine learning computational data integration and analysis models using EHR data are critical for developing personalized and precision prevention and treatment, with improved HF patient outcomes.
Machine learning models, such as random forests [5], decision trees [6], logistic regression [7], and support vector machines [8], have been successfully and widely used in many prediction and classification tasks. Moreover, deep learning models, like deep belief neural networks [9], deep convolutional networks [10], and long short-term memory [11] models, as well as more complicated deep learning models, have for the most part demonstrated stronger prediction and classification capability than traditional machine learning models. In HF-related studies, machine learning and deep learning models have been developed using the variables derived from the complex and diverse EHR data of HF patients. Critical to this rapidly developing area is a functional knowledge of the application of machine learning and deep learning models to the EHR, wearable sensor, genetic, and proteomic data associated with HF diagnosis, hospitalization, readmission, and mortality prediction. Our goal is to summarize state-of-the-art machine learning approaches to HF risk prediction.

Several challenges must be overcome for machine learning models to be applied on a personalized and precision basis for diagnostic and predictive management (diagnosis, prevention, risk stratification, and treatment) of HF patients. For example, algorithms must be developed to allow the full integration of the widely diverse data available in the EHR, ranging from textual medical reports, a wide variety of imaging data formats, and such developing fields as personalized genetic profiles. Furthermore, the application of machine learning to prediction in rare disease patient populations will mandate further enrichment of techniques for managing unbalanced dataset effects and as the identification of stable, clinically applicable, and informative risk factors to make the models interpretable and actionable.

Methods

We conducted a comprehensive review of available literature between January 2015 and August 2020 by a search of the PubMed library database for relevant papers. The keywords searched included “machine learning heart failure” and “deep learning heart failure.” By searching “machine learning” and “heart failure” in the PubMed database, 353 papers were obtained. The search for “deep learning” and “heart failure” revealed 69 papers. Removing the common articles from the above two searches, 374 unique papers were obtained. Among them were 335 relevant papers published after 2015. We reviewed and selected a subset of the most applicable publications (Table 1).

Statistics of Study Articles

Several trends are apparent in the applications of machine learning and deep learning in HF subpopulations (Fig. 1). Figure 1a shows the publishing trends by plotting the quantity of publications related to the machine learning and deep learning in HF from 2015 to 2020. A stable growth in the number of publications can be observed after 2015, especially after 2018, suggesting a progressive clinical recognition of the value of machine learning and deep learning algorithms applied in HF. Figure 1b shows the top 20 journals in which the collected 335 papers were published. These journals contained 35.2% of the papers published in the past 5 years. Most of them are influential journals in the research fields of HF (e.g., European Journal of Heart Failure, JACC Heart Failure, and Circulation Heart Failure), health and medical informatics (e.g., JMIR Medical Informatics, IEEE Journal Biomedical Informatics, and BMC Medical Informatics and Decision Making), and image and biotechnology (e.g., Medical Image Analysis, Computer Biology Medicine).

Moreover, we collected all the paper titles of the 335 published papers and generated a word cloud to capture the most studied topics in the application of machine learning and deep learning algorithms in HF patients. Figure 1c shows the word cloud of all the paper titles resulting from the use of the Natural Language Toolkit (NLTK) tool [32] to lemmatize each word. Figure 1c illustrates the specific techniques used in studying these HF patients: machine learning, deep learning, neural network, artificial intelligence; and medical outcomes such as readmission, mortality, and detection. Other typical words included “patient,” “prediction,” and “risk model,” suggesting that a primary focus was HF patient risk stratification.

Machine Learning and Deep Learning for Heart Failure Risk Prediction

HF outcome prediction is critical to the accurate application of many available therapeutic options, ranging from pharmacologic to highly invasive mechanical ventricular assistance and cardiac transplantation. Recent HF outcome prediction investigations have focused upon EHR, echocardiographic, proteomic, and wearable sensor data. In one investigation, using quantitative features derived from echocardiography images, domain expert selected features and data-driven selected features were integrated in machine learning models, including decision tree models. Data-driven feature selection had much better prediction accuracy than expert-driven feature selection [12]. In another study, the timing and amplitude of left ventricular (LV) images were analyzed to obtain the myocardial motion and deformation information in the cardiac cycle during rest and stress. Their results suggested that LV images can provide informative features for HF with preserved ejection fraction (HFpEF) prediction [15].

In addition to imaging data, EHR data are also informative in HF risk prediction. In one investigation, structured and unstructured EHR data were used to evaluate four approaches of HF hospitalization prediction [14•]. The results indicated...
Table 1 Summary of selected articles related to heart failure readmission and mortality prediction

| Title of articles                                                                 | Study design                                                                 | Sample size | Location | Main findings                                                                                       |
|----------------------------------------------------------------------------------|------------------------------------------------------------------------------|-------------|----------|---------------------------------------------------------------------------------------------------|
| Diagnosis and hospitalization prediction-related publications                     |                                                                              |             |          |                                                                                                   |
| Artificial intelligence for the diagnosis of heart failure [12]                   | The echocardiography images were used to extract the quantitative features. Multiple machine learning models were evaluated to detect the heart failure from normal patients. | 1198 patients | Korea    | The combination of expert selected features and data-driven features in echocardiography images using machine learning models can achieve high heart failure detection. |
| Medical concept representation learning from electronic health records and its application on heart failure prediction [13] | A novel embedding model was developed to represent heterogeneous medical concepts based on co-occurrence patterns in longitudinal electronic health records. Then, the widely used machine learning models were employed based on the embedding features to predict the heart failure. | 3884 heart failure, and 28,903 control | USA       | EHR data were converted into clinically meaningful numerical features, which can be employed in machine learning analyses to improve the heart failure prediction accuracy. |
| Early identification of patients with acute decompensated heart failure [14•]    | Four algorithms were developed, using the EHR data, to predict and identify hospitalization with a principal discharge diagnosis of ADHF. | 37,229 patients | USA       | Machine learning approaches with unstructured notes achieved best performance for ADHF prediction. |
| Diagnosis of heart failure with preserved ejection fraction: machine learning of spatiotemporal variations in left ventricular deformation [15] | The velocity, strain, and strain rate traces were measured from left ventricular (LV) echocardiographic myocardial velocity imaging to predict heart failure with preserved ejection fraction. | 100 patients | Belgium  | The spatiotemporal variations of LV strain rate during rest and exercise could be used to identify patients with HFP EF using machine learning methods. |
| Novel urinary peptidomic classifier predicts incident heart failure [16]          | The urinary proteomic profiles generated by mass spectrometric analysis were used to identify heart failure patients from non-heart failure patients. | 241 patients | Belgium  | Novel biomarkers derived from the urinary proteome can be a sensitive tool to improve risk stratification of heart failure patients. |
| Comparison of machine learning techniques for prediction of hospitalization in heart failure patients [17] | Compares the prediction performance of eight machine learning approaches, based on EHR data, for the hospitalization prediction of patients with heart failure. | 380 patients | Italy     | The generalized linear model net approach showed better performance than other machine learning approaches. |
| Identifying cancer patients at risk for heart failure using machine learning methods [18] | Predict heart failure risk of cancer patients using historical EHR data. | 143,199 patients | USA       | The gradient boosting–based model achieved the best prediction accuracy using EHR data. |
| Novel wearable seismocardiography and machine learning algorithms can assess clinical status of heart failure patients [19••] | Wearable devices that can remotely monitor patient ECG and seismocardiogram sensing can predict the risk of patients with heart failure and thereby potentially reduce hospitalizations | 45 patients | USA       | Wearable technologies recording cardiac function and machine learning algorithms can predict heart failure risks. |
| Readmission prediction–related publications                                      |                                                                              |             |          |                                                                                                   |
| Machine learning–based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics [20] | Demographics, admission characteristics, medical history, visits to emergency departments, history of medication use, and healthcare services out of hospital were used to predict heart failure. | 10,757 patients | Australia | Multi-layer perception was superior to other machine learning models. |
| Predictive modeling of hospital readmission rates using electronic medical record–wide machine learning: a case-study using Mount Sinai Heart Failure Cohort [3] | 4205 variables were extracted from EMR as the input of a multistep modeling strategy using the Naïve Bayes algorithm. | 1068 patients with 178 patients readmitted within a 30-day interval | USA       | The EMR-wide, naïve Bayes model achieved good readmission prediction. |
| 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 [21•] | EHR were used as input to deep unified networks to predict the patient readmission. | 11,510 patients with 27,334 admissions and 6369 30-day readmissions | USA       | The deep unified networks (DUNs) model outperformed the logistic regression, gradient boosting, and maxout networks. |
| A predictive analytics approach to reducing 30-day avoidable readmissions among patients with heart failure, acute myocardial infarction, pneumonia, or COPD [4] | A tree-based classification method using EHR data was proposed to estimate the predicted probability of readmission. | 2985 distinct adult patients from Veteran Health Administration (VHA) | USA       | The proposed model had better performance than random forest and support vector machine models. |
that the unstructured notes were important and could improve the prediction accuracy. Eight machine learning approaches, including generalized linear model nets, random forests, support vector machines, logistic regression, and neural networks, were evaluated for HF hospitalization prediction using patient demographic, medical, and clinical data [17]. The GLMN achieved the best performance. In another investigation aimed at patients with HF related to cancer therapeutics, EHR data were used to predict the risk of HF risk occurrence in cancer patients [18]. The results indicated that machine learning models can not only predict associated HF risk but also identify potential contributing clinical factors.

To further improve prediction accuracy using EHR data, novel embedding approaches [13] have been developed to convert EHR data into clinically meaningful numeric vectors/features. Prediction accuracy can be improved by applying machine learning models to these numeric features. Moreover, wearable equipment and sensors are being developed to acquire real-time data from HF patients to monitor potential risks remotely [19]. For example, wearable devices that can remotely monitor patient electrocardiography (ECG) and seismocardiogram sensing [19] can predict the HF risk of patients, and thereby potentially reduce patient hospitalizations and mortality. In addition to the EHR and imaging data,
novel biomarkers derived from urinary proteomics data [16] of HF patients have also been investigated. These biomarkers may have potential to accurately predict HF risk in machine learning models that combine them with EHR and imaging data.

Machine Learning and Deep Learning for Prediction of Hospital Readmissions

Hospital readmission rate is a significant challenge in the management of HF patients. It is widely accepted that HF outcomes and healthcare expense can be improved if HF patients with a high risk of readmission can be accurately identified and targeted with management algorithms. Several machine learning approaches, most involving the use of EHR data, have been employed to identify HF patients at high risk for readmission. In one study, the naive Bayes model was used to predict HF readmission using data from patient primary encounters [3]. Specifically, the top associated features in HF readmission for each of subset of the patient cohort were selected and combined as the input of the predictive naive Bayes model. A tree-based model, adopted from the random

![Fig. 1](image_url) Some initial statistics of the collected 335 papers about machine learning and deep learning in patients with heart failure. a The paper publishing trends from year 2015 to 2020. b The top 20 journals in which papers were published. c The word cloud of the paper titles from the 335 papers.
Machine Learning and Deep Learning for Mortality Prediction

Accurate mortality prediction is critical to effective therapeutic decision-making in HF patients. This prediction is challenging because of the lack of stable marker factors, the noise in the data, and the prevalence of imbalanced data sets. In one study, several machine learning approaches, including random forests, logistic regression, AdaBoost, decision trees, and support vector machines, were used for the HF mortality prediction using EHR data [29]. A similar set of machine learning models were evaluated for mortality prediction of HF patients using a comprehensive set of data, including all baseline demographic, clinical, laboratory, electrocardiographic, and symptom data [27]. The random forest model achieved the best performance among these models.

In other recent studies, deep learning models have been used to improve mortality prediction in HF patients. In one study, a novel deep learning model, Feature Rearrangement based Convolution Net (FReaConvNet), was used to predict in-hospital, 30-day, and 12-month mortality by mining the most important features. Feature importance analysis is important in improving the prediction accuracy in unbalanced data sets. Using EHR and laboratory data, machine learning analysis identified two important features, i.e., serum creatinine and ejection fraction [31]. Using only these two features obtained better prediction results than using all other evidential EHR data [31]. In another study, novel and complex computer vision models, using convolutional networks to calculate heart motion trajectories, accurately predicted patient survival using 4D imaging of heart (3D MRI images + time) [30].

Conclusions

HF is associated with high morbidity and mortality, as well as excessive associated healthcare cost. Using EHR data, including demographic characteristics, medical treatment history, medical diagnosis notes, laboratory results, image data, and genetic and proteomic profiles, machine learning and deep learning have been employed for the prediction of HF readmission and mortality. These models are essential to personalized and precision prevention, treatment, and management of HF patients. A set of machine learning and deep learning models have been evaluated for related prediction analyses with considerable success using large variable data sets derived from the EHR. Nonetheless, there are still challenges to be resolved, and novel machine learning models are still needed to integrate diverse and heterogeneous data in the quest to more accurately identify high-risk HF patients. The diverse and heterogeneous attributes of clinical EHR data include variable data format (longitudinal data versus fixed data; structure versus non-structured data; text data versus complex image data), the measurement of different aspects of the diseases, the data noise, and the predominance of imbalanced HF versus control study group samples. Novel machine learning models for systematic data representation, integration, and prediction have potential to revolutionize model prediction accuracy.

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1. WHO-CVD. https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1.
2. Virani SS, Alvaro A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics—2020 Update: A Report From the American Heart Association. Circulation. 2020;141(9):e139–596. https://doi.org/10.1161/CIR. 0000000000000757.

3. Khader S, Johnson K, Yahi A, Miozzo R, Li L, Ricks D, et al. Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: a case-study using Mount Sinai Heart Failure Cohort. Vol. 22. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing. 2017; 276–287.

4. Shams I, Ajorlou S, Yang K. A predictive analytics approach to reducing 30-day avoidable readmissions among patients with heart failure, acute myocardial infarction, pneumonia, or COPD. Health Care Manag Sci. 2014, 2015;1:1–16. https://doi.org/10.1007/s10729- 014-9278-y.

5. Ho TK. Random decision forests. In: Proceedings of the International Conference on Document Analysis and Recognition, ICDAR. 1995.

6. Quinlan JR. Induction of decision trees. Mach Learn. 1986;1:81–106.

7. McCulloch CE, Generalized Linear Models, J Am Stat Assoc. 2001;95(452):1320–1324. https://doi.org/10.1080/01621459. 2000.10474340.

8. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995. https://doi.org/10.1007/BF00994018.

9. Hinton G. Deep belief networks. Scholarpedia. 2009. http:// scholarpedia.org/article/Deep_belief_networks.

10. Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern. 1980;36:193–202.

11. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9:1735–80.

12. Choi D-J, Park JJ, Ali T, Lee S. Artificial intelligence for the diagnosis of heart failure. npj Digit Med. 2020;3(1):54. https://doi.org/10.1038/s41746-020-0261-3.

13. Choi E, Schuetz A, Stewart W, Sun J. Medical concept representation learning from electronic health records and its application on heart failure prediction. arXiv:160203686. 2016.

14. • Blecker S, Sontag D, Horwitz L, Kuperman G, Park H, Reyetovich A, et al. Early identification of patients with acute decompensated heart failure. J Card Fail. 2018;24(6):357–62 A large cohort of adult patients ($n = 37,229$) was conducted in the USA. Results indicated that Machine learning approaches with unstructured notes achieved best performance for ADHF prediction. These findings may suggest that machine learning algorithms can help providers improve efficiency to deliver improved quality interventions.

15. Tabassian M, Sunderji I, Erdei T, Sanchez-Martinez S, Degiovanni A, Marino P, et al. Diagnosis of heart failure with preserved ejection fraction: machine learning of spatiotemporal variations in left ventricular deformation. J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr. 2018;31(12):1272–1284.e9.

16. Zhen-Yu Z, Susana R, Esther N, Wen-Yi Y, K SM, Thomas K, et al. Novel urinary peptidomic classifier predicts incident heart failure. J Am Heart Assoc. 2020;9(8):e005432. https://doi.org/10. 1161/JAHA.116.005432.

17. Lorenzi G, Sabato SS, Lanera C, Bottigliengo D, Minto C, Ocagni L, et al. Comparison of machine learning techniques for prediction of hospitalization in heart failure patients. J Clin Med. 2019;8(9):1298 Available from: https://pubmed.ncbi.nlm.nih.gov/31450546.

18. Yang X, Gong Y, Waheed N, March K, Bian J, Hogan WR, et al. Identifying cancer patients at risk for heart failure using machine learning methods. AMIA. Annu Symp proceedings AMIA Symp. 2020;2019:933–41 Available from: https://pubmed.ncbi.nlm.nih.gov/32308890.

19. • Inan OT, Baran Pouyan M, Javid A, Dowling S, Etemadi M, Dorier A, et al. Novel wearable seismocardiography and machine learning algorithms can assess clinical status of heart failure patients. Circ Heart Fail. 2018;11(1):e004313 The study was conducted on 45 patients who were fitted with wearable devices that can remotely monitor patients in the USA. Results indicated that wearable technologies recording cardiac function and machine learning algorithms can predict heart failure risks. Findings suggested that the clinical status and response to pharmacological interventions can be tracked by these techniques in the future.

20. Anse W, Bennamoun M, Soled H, Sanfilippo FM, Dwiwedi G. Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. ESC Heart Fail. 2019;6:428–35.

21. • Golas SB, Shihabara T, Agboola S, Ouki H, Satoh J, Nakae T, 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. 2018:18(1):44 Available from: https://pubmed.ncbi.nlm.nih.gov/ 29929496 A large cohort of 11,510 patients with 27,334 admissions was studied to predict 30-day readmission of patients by medical records as input to deep unified networks. Results indicated that the deep unified networks (DUNs) model outperformed the logistic regression, gradient boosting, and maxout networks. Findings may enable healthcare teams to improve overall clinical outcomes by targeting interventions for high-risk patients identified by the deep learning models.

22. Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li S-X, et al. Analysis of machine learning techniques for heart failure readmissions. Circ Cardiovasc Qual Outcomes. 2016;9(6):629–40 Available from: https://pubmed.ncbi.nlm.nih.gov/28263938.

23. • Lin Y-W, Zhou Y, Faghri F, Shaw MJ, Campbell RH. Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. PLoS One. 2019;14(7):e0218942. https://doi.org/10.1371/journal.pone. 0218942 The analysis was performed on 40,000 ICU patients available from MIMIC-III Critical Care Database. Results indicated that long short-term memory (LSTM) accurately predicted longitudinal data and outperformed other models, and thus had the ability to improve ICU decision-making accuracy. Findings suggested that machine learning and deep learning models would improve allocation of healthcare resources and patient consultation.

24. Kwon J, Kim K-H, Ki-Hyun J, Lee SE, Lee H-Y, Cho. Artificial intelligence algorithm for predicting mortality of patients with acute heart failure. PLoS One. 2019;14(7):e0219302. https://doi.org/10.1371/journal.pone.0219302.

25. Wang Z, Zhu Y, Li D, Yin Y, Zhang J. Feature rearrangement based deep learning system for predicting heart failure mortality. Comput Methods Programs Biomed. 2020. https://doi.org/10.1016/j. cmpb.2020.105383.

26. Kalscheur MM, Kipp RT, Tattersall MC, Mei C, Buhr KA, Demets DL, et al. Machine learning algorithm predicts cardiac resynchronization therapy outcomes: lessons from the COMPANION Trial. Circ Arrhythmia Electrophysiol. 2018. https://doi.org/10.1161/CIRCEP.117.005499.

27. • Angraal S, Mortazavi BJ, Gupta A, Khera R, Ahmad T, Desai NR, et al. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. JACC Heart Fail. 2020. https://doi.org/10.1016/j.jchf.2019.06.013. A cohort of 1767 patients with heart failure with preserved ejection fraction (HFpEF) from four different countries was utilized to predict
mortality and hospitalization of patients by machine learning models. Results indicated that random forest models achieved the best performance compared to 4 other machine learning models.

28. Adler ED, Voors AA, Klein L, Macheret F, Braun OO, Urey MA, et al. Improving risk prediction in heart failure using machine learning. Eur J Heart Fail. 2020. https://doi.org/10.1002/ejhf.1628.

29. Panahiazar M, Taslimitehrani V, Pereira N, Pathak J. Using EHRs and machine learning for heart failure survival analysis. In: Studies in health technology and informatics. 2015;216:40–44.

30. Bello GA, Dawes TJW, Duan J, Biffi C, de Marvao A, Howard LSGE, et al. Deep learning cardiac motion analysis for human survival prediction. Nat Mach Intell. 2019;1:95–104 Available from: https://pubmed.ncbi.nlm.nih.gov/30801055 A cohort of 302 patients from the UK was utilized for human survival prediction by using complex 4D imaging of heart (3D MRI images + time) data of patients. Results indicated that computer vision analysis using high-dimensional medical image data can efficiently predict human survival. The fast and scalable method could improve clinical decision-making accuracy and better understand mechanisms of disease.

31. Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Med Inform Decis Mak. 2020;20(1):16. https://doi.org/10.1186/s12911-020-1023-5.

32. Bird S, Loper E, Klein E. Natural language toolkit (NLTK) book: O’Reilly Media Inc; 2009. http://www.nltk.org/book_1ed/.

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