The Effective Management of Machine Learning Using Electronic Health Record Data for Earlier Diagnosis of Type 2 Diabetes

Katharine Blancett[1] and Dr. Bruce Lazar, MBA, DM[2]

[1]Master of Health Informatics, Graduate Student, Southern Illinois University, Carbondale, United States of America
katharine.blancett@siu.edu
[2]School of Health Sciences Lecturer, Southern Illinois University, Carbondale, United States of America
bruce.lazar@siu.edu

Abstract: Type 2 diabetes is a chronic disease that affects many people worldwide and, if left untreated, can lead to more severe health issues. Specific factors can increase a person's risk of developing type 2 diabetes; screening tests are essential to identify these factors. Machine learning can use electronic health record data to help detect risk factors earlier. The systematic literature review explored strategies that healthcare informatics leaders use to implement and effectively manage machine learning using electronic health record data for earlier diagnosis of type 2 diabetes. Medical Literature Analysis and Retrieval System and Cumulative Index to Nursing and Allied Health Literature Plus with full text were used to search for research-related articles to help answer the research question while using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses as a guide. Twenty relevant articles enabled results to answer the research question. The five primary themes that appeared from the literature analysis of the articles were artificial intelligence and machine learning have predictive capabilities to help manage type 2 diabetes; data mining can be used to find other diseases commonly found in patients with type 2 diabetes, machine learning can look at laboratory data to early screen for type 2 diabetes, machine learning can help develop more personalized treatments plans for patients with type 2 diabetes, and machine learning can look for medications already on the market that might help treat type 2 diabetes. All five themes used electronic health record data. The findings implied that machine learning could be a valuable tool to help diagnose type 2 diabetes early, but more research needs to be done to determine which specific data points will lead to the best results.

Keywords—Electronic Health Record; Machine Learning; Type 2 Diabetes

1. INTRODUCTION

By the year 2040, it is predicted that diabetes, a chronic metabolic disorder, will have a significant impact on approximately 642 million individuals [1]. Type 1, Type 2, and Gestational are the three main types, and 90% of these patients have type II diabetes [1]. Hypertension, age, and obesity are some of the factors that can increase a person's risk of developing diabetes [2]. Identifying risk factors for diabetes early on is crucial because left untreated, it can cause damage to organs such as kidneys, nerves, and blood vessels [1].

Type 2 diabetes can typically be expected to develop in patients with these risk factors as they get older and live a specific lifestyle [2]. Screening tests determine if a patient currently has the risk factors for type II diabetes [3]. Screening tests look for chronic hyperglycemia in patients by looking at impaired glucose tolerance and impaired fasting glycemia [4]. Healthcare data is constantly changing, so healthcare informatics leaders need to develop strategies to analyze these tests and learn how they relate to the risk factors [5].

Machine learning has been used to detect diabetic risk factors at an earlier stage of diagnosis [2]. Machine learning can look at electronic health record data for type 2 diabetic patients and find similar patterns in patients with risk factors or positive screening tests for the disease [5]. The computer will learn the critical patterns, potentially leading to better diagnostic tests [6]. However, health informatics leaders need more appropriate strategies for implementing and managing machine learning using health data. This systematic literature review explores practical strategies that healthcare informatics leaders use to implement and effectively manage machine learning using electronic health record data for earlier diagnosis of Type II diabetes.

2. METHODS

Google Scholar was the first resource to search for articles that would help develop a research topic. The articles were then narrowed down further to create the research question; what strategies do healthcare informatics leaders use to implement and effectively manage machine learning using electronic health record data for earlier diagnosis of type 2 diabetes? The literature was searched further using Medical Literature Analysis and Retrieval System (MEDLINE) and Cumulative Index to Nursing and Allied Health Literature (CINAHL) Plus with full-text databases following the Boolean Searching research guides [28]. The methods phase followed five basic steps. These steps include finding studies related to the research topic, screening for inclusion and exclusion criteria, capturing specific data, using that data to help answer the research question, and then stating the findings [27].

The academic databases were searched by using the keywords electronic health record, machine learning, and type 2 diabetes. These keywords helped find the articles needed to answer the research question. For an article to meet the inclusion criteria, it had to have been (a) published between 2019 and 2023, (b) written in the English language, (c) full text only, (d) peer-reviewed journal articles, and (e) specifically focus on the use of machine learning in electronic health systems to help diagnose type 2 diabetes early. When all the inclusion criteria were met using the MEDLINE database, it resulted in 13 articles. The CINAHL Plus with a full-text database resulted in 7 articles. See Figure 1 for more details.

Figure 1

Flow Diagram
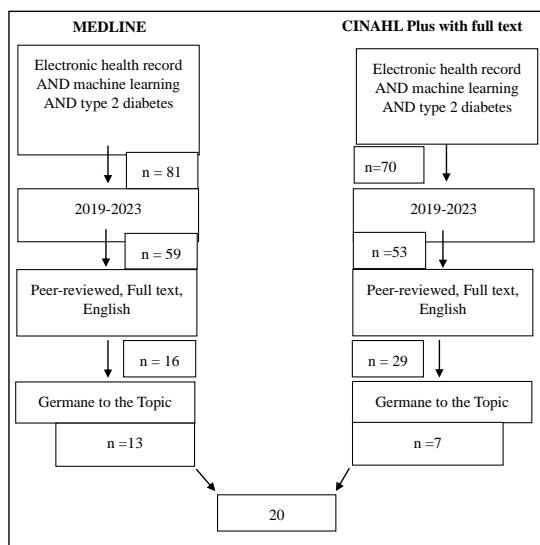
3. RESULTS

The fundamental research question was, what strategies do healthcare informatics leaders use to implement and effectively manage machine learning using electronic health record data for earlier diagnosis of type 2 diabetes? An established process for the literature search, selection of articles, and data analysis process outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [27] was followed to ensure a sound, systematic literature review occurred. By searching each database, MEDLINE and CINAHL Plus., 45 articles were collected. From this selection of articles, 20 studies were observed as relevant to address the research question (see Figure 1); by carefully assessing each study's summary findings, a conclusion was made to include or exclude each article. Table 1 shows the titles and summaries obtained from the final selected articles.

Table 1

Summarized findings of the literature.

Title	Findings
[1] Application of Artificial Intelligence in Discovery and Development of Anticancer and Antidiabetic Therapeutic Agents.	Artificial Intelligence (AI) and machine learning have predictive capabilities and are used to help manage diabetes in patients. These prediction models can help detect diabetic risk factors and reshape diabetes prevention measures. This can be done using electronic health record data.
[2] Connections between Various Disorders: Combination Pattern Mining Using Apriori Algorithm Based on Diagnosis Information from Electronic Medical Records	The Apriori algorithm was used in the data mining of an Electronic Medical Record (EMR) database to look at patients with one disorder and see if that disorder puts them at risk for one or more other disorders. Type 2 diabetes is an example of a disorder that is seen with other disorders. Using this type of algorithm is a strategy to help prevent these types of disorders. This type of algorithm was good at determining the correlation of disorders but not causality.
[3] Use of Machine Learning and Routine Laboratory Tests for Diabetes Mellitus Screening	Patients may not be aware they have diabetes when they have no symptoms. These patients might not be screened for diabetes, but they are having other routine laboratory tests completed. Machine learning models use these routine laboratory tests to determine if a patient's glycated hemoglobin values might be within a range, putting them at risk for diabetes. This type of screening tool could be used so patients know early on they need further testing to determine if they have or are at risk for diabetes.
[4] Potential value and impact of data mining and machine learning in clinical diagnostics.	Data mining and machine learning have been used on electronic health records to determine the risk for type 2 diabetes. Certain values such as age, physical activity, and family history help with these predictions. HbA1c and IL-6 can help diagnose Type 2 diabetes but is not good at diagnosing prediabetes.
[5] Automation and artificial intelligence in the clinical laboratory.	Machine learning and artificial intelligence have been used to predict the management of pre-diabetes and type 2 diabetes. These predictions helped determine new and already known risk factors.
[6] Diabetes care: is big data the future?	EHR comprises large datasets, and it can be challenging to get quality information. A type of machine learning called
[7] The future of diabetes education	Specific algorithms and machine learning can help create automatic systems and provide standardized education. These systems can be used to provide more personalized treatments for type 2 diabetic patients. These computers should not replace human interactions but could help with efficiency.
[8] Identification of repurposable drugs with beneficial effects on glucose control in type 2 diabetes using machine learning	The EHR records of patients who had just started type 2 diabetes medication were reviewed and analyzed using machine learning. This method showed that α1-adrenoceptor antagonists can help control glucose levels in type 2 diabetes, so other medications might be beneficial in treating type 2 diabetes.
[9] A novel method for causal structure discovery from EHR data and its application to type-2 diabetes mellitus	This article looked at new methods to find new causes of type 2 diabetes. This was done by looking at EHR data to help develop clinical decision support tools. This can help in finding new ways to diagnose type 2 diabetes.
[10] On using electronic health records to improve optimal treatment rules in randomized trials.	This article looked at the use of machine learning in an EHR record and compared it to non-EHR methods. The EHR method was used to look at type 2 diabetic patients and find insulin treatments specific to that patient. The individualized machine learning methods did not work as well as looking at all the type 2 diabetic patients collectively.

[11] Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records.	A new model called LightGBM uses machine learning and the EMR to analyze patients with type 2 diabetes and normal albumin in the urine and determine their risk for diabetic kidney disease. This model appears to have the potential to help diagnose that risk.
[12] Predicting unplanned medical visits among patients with diabetes: translation from machine learning to clinical implementation.	Machine learning was used in an EHR system to determine why patients with type 2 diabetes might show up to a clinic visit unexpectedly. Certain data points, such as HDL and BP, contributed to the increased risk. This is an example of using a predictive type of model to help manage type 2 diabetes.
[13] A model-agnostic approach for understanding heart failure risk factors	Machine learning was used on an EHR database to determine the risk of type 2 diabetic patients developing heart failure. The method used appeared to be more accurate than other methods used in the past.
[14] Identifying myoglobin as a mediator of diabetic kidney disease: a machine learning-based cross-sectional study.	Machine learning was used on EHR data to discover if certain biomarkers, such as serum myoglobin can be used to identify type 2 diabetic patients at risk for diabetic kidney disease earlier on. This method could benefit patients in the future, but only if the serum myoglobin is at a higher level.
[15] Personalized treatment options for chronic diseases using precision cohort analytics.	Machine learning was used on EHR data to analyze chronic diseases such as type 2 diabetes to find other treatment options. This method might help create individual-type treatments for these patients in the future.
[16] Development and implementation of patient-level prediction models of end-stage renal disease for type 2 diabetes patients using fast healthcare interoperability resources	This article discovered that future studies need to be done to analyze how machine learning can use EMR data to predict if type 2 diabetic patients are at risk of developing end-stage renal disease within five years of diagnosis. These methods could help determine treatment early on to decrease that risk.
[17] Comorbidity clusters associated with newly treated type 2 diabetes mellitus: a Bayesian nonparametric analysis.	This article used machine learning to analyze what other chronic diseases patients with type 2 diabetes are likely to have, especially early on in diagnosis. This method appeared to be beneficial in identifying those diseases.
[18] Assessment of medical management in Coronary Type 2 Diabetic patients with previous percutaneous coronary intervention in Spain: A retrospective analysis of electronic health records using Natural Language Processing.	This review looked at a technology called *EHRead®* to identify patients with type 2 diabetes using specific variables such as lab results and medical history. This method was able to positively identify patients with this disease.
[19] Machine learning and deep learning predictive models for type 2 diabetes: a systematic review.	Machine learning was used to analyze data to help discover patients with type 2 diabetes and their risk for problems. The proper methods need to be used to determine the right data.
[20] Matched Learning for Optimizing Individualized Treatment Strategies Using Electronic Health Records	This article looked at the use of machine learning on large EHR records to discover specific treatment methods for patients with type 2 diabetes. This method could determine new ways of treating the disease.

After reviewing all 20 articles, the data was classified into five shared themes. The themes are directly associated with the research question, and that information is listed below in an affinity matrix (Table 2) to calculate how often these themes occurred in the reviewed literature. The five common themes are: (a) artificial intelligence and machine learning have predictive capabilities to help manage type 2 diabetes when using EHR data, (b) data mining used to look at other diseases seen in patients with type 2 diabetes using EHR data, (c) machine learning used to look at laboratory data on an EHR system to early screen for type 2 diabetes, (d) machine learning using EHR data could assist in developing more personalized treatment plans for patients with type 2 diabetes, (e) using machine learning to look for medications already on the market that might help treat type 2 diabetes.

Table 2
Frequency of occurrence in the literature.

Theme	Occurrences	Instances of Attributes (n)	Percentage (%)
Artificial intelligence and machine learning have predictive capabilities to help manage type 2 diabetes when using EHR data	1, 4, 5, 6, 9, 10, 11, 12, 13, 18, & 19	n=11	55%
Data mining is used to look at other diseases seen in patients with type 2 diabetes using EHR data	2, 17	n=2	10%
Machine learning is used to look at laboratory data on an EHR system to early screen for type 2 diabetes.	3, 14	n=2	10%
Machine learning using EHR data could assist in developing more personalized treatment plans for patients	7, 15, 16, & 20	N=4	20%

The research findings showed that 55% of articles showed evidence that artificial intelligence and machine learning could use EHR data to predict and manage type 2 diabetes [1, 4, 5, 6, 9, 10, 11, 12, 13, 18, 19]. 10% of the articles provided proof that data mining can be used to identify other diseases seen in patients with type 2 diabetes when using EHR data [2, 17]. 10% of articles documented that machine learning can be used to look at laboratory data within an EHR system to help screen for type 2 diabetes earlier on [3, 14]. 20% of articles provided evidence that machine learning of EHR data helped assist in

developing a more personalized treatment plan for patients with type 2 diabetes [7, 15, 16, 20]. Finally, the data showed that 5% of the articles mentioned machine learning to look for medications already on the market that could help treat type 2 diabetes, and the data was extracted from an EHR system [8].

4. DISCUSSION

The study aimed to determine what strategies healthcare informatics leaders use to implement and effectively manage machine learning using electronic health record data for earlier diagnosis of type 2 diabetes. The systematic literature review drew on 20 peer-reviewed journal articles from 2019 to 2023 to analyze types of electronic health record data for earlier detection and diagnosis of type 2 diabetes. The research uncovered five primary themes, which included artificial intelligence and machine learning having predictive capabilities to help manage type 2 diabetes when using EHR data, data mining used to look at other diseases seen in patients with type 2 diabetes using EHR data, machine learning used to look at laboratory data on an EHR system to early screen for type 2 diabetes, machine learning using EHR data could assist in developing more personalized treatment plans for patients with type 2 diabetes, using machine learning to look for medications already on the market that might help treat type 2 diabetes. The data is taken from an EHR system, using the EHR System to find new ways of diagnosing type 2 diabetes.

Approximately 55% of the articles demonstrated that artificial intelligence and machine learning have predictive capabilities to help manage type 2 diabetes when using EHR data [7, 10, 11, 12, 15, 16, 17, 18, 19, 24, 25]. Numerous scholars identified that artificial intelligence and machine learning models were able to detect risk factors in type 2 diabetic patients and could be done through the extraction of essential data points such as age, physical activity, and family history [7,10,11,12,15,16,17,18, 19, 24, 25]. In contrast, other data sets, such as HbA1c, helped diagnose patients with type 2 diabetes but were subpar at predicting pre-diabetes [10]. As such, health informatics leaders should continue to research which data points will aid in predicting risk factors earlier on to help prevent or decrease the number of patients with type 2 diabetes.

Approximately 10% of the articles revealed that data mining using EHR data could discover other types of diseases that type 2 diabetes patients are prone to having [8, 23]. Data mining can use algorithms and follow other models to group specific disorders in diabetic patients [8, 23]. The algorithms employed were beneficial in determining correlations but had issues discovering causality [8]. Correlation is the first step in recognizing which disorders are most seen in patients with type 2 diabetes which could lead to finding a cause to prevent these comorbidities.

In 10% of the articles, machine learning was used to find laboratory data on an EHR system that matched most type 2 diabetes patients to help screen the disease [9, 20]. The laboratory data used was extracted from routine tests and have

led to discovering specific biomarkers specific to diabetic patients who end up with diabetic kidney disease [9, 20]. Some models could not locate suitable biomarkers for patients with diabetic kidney disease, but machine learning has helped determine which data points can [20]. Providers can use this information to discover which laboratory tests are critical in screening for potential type 2 diabetes [9].

In 20% of the articles, it was found that personalized treatment plans could be developed for type 2 diabetic patients by using machine learning with EHR data [13, 21, 22, 26]. Machine learning can examine the data from groups of patients and find similarities to the current type 2 diabetic patient being treated [21]. However, machine learning should not substitute direct interaction between patients and their providers [13]. Patients with type 2 diabetes may share similar symptoms but react differently to treatment; ultimately, the patient should be treated individually.

Approximately 5% of articles found that machine learning could use EHR data to discover approved medications using different indications to treat type 2 diabetes [14]. Such as a patient with type 2 diabetes might already receive treatment with non-diabetic medicines that help control glucose levels [14]. In contrast, these benefits are specific to patients on the medication to treat a different disease. So, it is difficult to know if the same benefits would occur for someone with only type 2 diabetes, not having the secondary condition [14]. This method of researching other medication use methods could help find new ways of treating type 2 diabetes [14].

Healthcare informatics leaders successfully found strategies to implement and effectively manage machine learning using electronic health record data for earlier diagnosis of type 2 diabetes. One example of a strategy used was an Apriori algorithm that successfully identified other disorders patients with type 2 diabetes can develop [8]. Another strategy is using machine learning models to look at routine laboratory data points such as hemoglobin values, and this is a successful screening tool to diagnose type 2 diabetes earlier on [9]. Thirdly, a type of machine learning called deep learning is a strategy that can remember past predictions, which could help find information from large datasets that can be important in diagnosing type 2 diabetes earlier [12]. The results from these strategies are helpful tools in developing future screening methods for early diagnosis of type 2 diabetes.

There were a few limitations when it came to writing this research article. The articles chosen were done over twelve weeks to complete the research. The practical time frame meant that articles had to be sifted on time, and the articles reviewed could have been more specific to a geographic location. Health informatics leaders might manage machine learning differently depending on what part of the country they are in, and which diseases are more prevalent. The keywords used in the search were electronic health record, machine learning, and type 2 diabetes. Applying similar words might have brought back varying results. These limitations were

minimized by choosing terms that would bring back quality articles to answer the specific research question. This method allowed for an appropriate number of studies to research that could be managed within twelve weeks.

In the future, researchers could use this information to help use machine learning in preventing other chronic diseases. If these methods work specifically for managing type 2 diabetes, then those processes can be mirrored to help in other areas of medicine. These tools are stepping stones to improving healthcare overall.

5. CONCLUSION

This systematic literature review explored strategies that healthcare informatics leaders use to implement and effectively manage machine learning using electronic health record data for earlier diagnosis of type 2 diabetes. Machine learning has successfully detected health information from electronic health record data to help manage type 2 diabetes. However, there is a lot more that needs to be learned. The findings implied that machine learning could be a valuable tool to help diagnose type 2 diabetes early, but more research needs to be done to determine which specific data points will lead to the best results.

6. REFERENCES

[1] Afsaneh, E., Sharifdini, A., Ghazzaghi, H., & Ghobadi, M. Z. (2022). Recent applications of machine learning and deep learning models in the prediction, diagnosis, and management of diabetes: a comprehensive review. Diabetology & Metabolic Syndrome, 14(1), 1–39. https://doi-org.proxy.lib.siu.edu/10.1186/s13098-022-00969-9

[2] Ullah, Z., Saleem, F., Jamjoom, M., Fakieh, B., Kateb, F., Ali, A. M., & Shah, B. (2022). Detecting High-Risk Factors and Early Diagnosis of Diabetes Using Machine Learning Methods. Computational Intelligence & Neuroscience, 2022, 1–10. https://doi-org.proxy.lib.siu.edu/10.1155/2022/2557795

[3] Kopitar, L., Kocbek, P., Cilar, L., Sheikh, A., & Stiglic, G. (2020). Early detection of type 2 diabetes mellitus using machine learning-based prediction models. Scientific Reports, 10(1), 1–12. https://doi-org.proxy.lib.siu.edu/10.1038/s41598-020-68771-z

[4] Kaur, G., Lakshmi, P. V. M., Rastogi, A., Bhansali, A., Jain, S., Teerawattananon, Y., Bano, H., & Prinja, S. (2020). Diagnostic accuracy of tests for type 2 diabetes and prediabetes: A systematic review and meta-analysis. PLoS ONE, 15(11), 1–19. https://doi-org.proxy.lib.siu.edu/10.1371/journal.pone.0242415

[5] ElSeddawy, A. I., Karim, F. K., Hussein, A. M., & Khafaga, D. S. (2022). Predictive Analysis of Diabetes-Risk with Class Imbalance. Computational Intelligence & Neuroscience, 1–16. https://doi-org.proxy.lib.siu.edu/10.1155/2022/3078025

[6] Cedeno-Moreno, D., & Vargas-Lombardo, M. (2020). Mobile Applications for Diabetes Self-Care and Approach to Machine Learning. International Journal of Online & Biomedical Engineering, 16(8), 25–38. https://doi-org.proxy.lib.siu.edu/10.3991/ijoe.v16i08.13591

[7] Alqahtani, A. (2022). Application of Artificial Intelligence in Discovery and Development of Anticancer and Antidiabetic Therapeutic Agents. Evidence-Based Complementary & Alternative Medicine (ECAM), 1–16. https://doi-org.proxy.lib.siu.edu/10.1155/2022/6201067

[8] Ma, H., Ding, J., Liu, M., & Liu, Y. (2022). Connections between Various Disorders: Combination Pattern Mining Using Apriori Algorithm Based on Diagnosis Information from Electronic Medical Records. BioMed Research International, 1–16. https://doi-org.proxy.lib.siu.edu/10.1155/2022/2199317

[9] Cardozo, G., Pintarelli, G. B., Andreis, G. R., Lopes, A. C. W., & Marques, J. L. B. (2022). Use of Machine Learning and Routine Laboratory Tests for Diabetes Mellitus Screening. BioMed Research International, 1–14. https://doi-org.proxy.lib.siu.edu/10.1155/2022/8114049

[10] Saberi-Karimian, M., Khorasanchi, Z., Ghazizadeh, H., Tayefi, M., Saffar, S., Ferns, G. A., & Ghayour-Mobarhan, M. (2021). Potential value and impact of data mining and machine learning in clinical diagnostics. Critical Reviews in Clinical Laboratory Sciences, 58(4), 275–296. https://doi-org.proxy.lib.siu.edu/10.1080/10408363.2020.1857681

[11] Naugler, C., & Church, D. L. (2019). Automation and artificial intelligence in the clinical laboratory. Critical Reviews in Clinical Laboratory Sciences, 56(2), 98–110. https://doi-org.proxy.lib.siu.edu/10.1080/10408363.2018.1561640

[12] Begg, A. (2022) Diabetes care: is big data the future? Practical Diabetes, 39(3), 7–9. https://doi-org.proxy.lib.siu.edu/10.1002/pdi.2391

[13] Wake, D. J. (2019). The future of diabetes education. Practical Diabetes, 36(6), 196–200. https://doi-org.proxy.lib.siu.edu/10.1002/pdi.2245

[14] Koren, G., Nordon, G., Radinsky, K., & Shalev, V. (2019). Identification of repurposable drugs with beneficial effects on glucose control in type 2 diabetes using machine learning. Pharmacology Research & Perspectives, 7(6), 1–5. https://doi-org.proxy.lib.siu.edu/10.1002/prp2.529

[15] Shen, X., Ma, S., Vemuri, P., Castro, M. R., Caraballo, P. J., & Simon, G. J. (2021). A novel method for causal structure discovery from EHR data and its application to type-2 diabetes mellitus. Scientific Reports, 11(1), 1–9. https://doi-org.proxy.lib.siu.edu/10.1038/s41598-021-99990-7

[16] Wu, P., Zeng, D., Fu, H., & Wang, Y. (2020). On using electronic health records to improve optimal treatment rules in randomized trials. Biometrics, 76(4), 1075–1086.

[17] Dong, Z., Wang, Q., Ke, Y., Zhang, W., Hong, Q., Liu, C., Liu, X., Yang, J., Xi, Y., Shi, J., Zhang, L., Zheng, Y., Lv, Q., Wang, Y., Wu, J., Sun, X., Cai, G., Qiao, S., Yin, C., & Su, S. (2022). Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. Journal of Translational Medicine, 20(1), 1–10. https://doi-org.proxy.lib.siu.edu/10.1186/s12967-022-03339-1

[18] Selya, A., Anshutz, D., Griese, E., Weber, T. L., Hsu, B., & Ward, C. (2021). Predicting unplanned medical visits among patients with diabetes: translation from machine learning to clinical implementation. BMC Medical Informatics & Decision Making, 21(1), 1–11. https://doi-org.proxy.lib.siu.edu/10.1186/s12911-021-01474-1

[19] Miran, S. M., Nelson, S. J., & Zeng-Treitler, Q. (2021). A model-agnostic approach for understanding heart failure risk factors. BMC Research Notes, 14(1), 1–5. https://doi-org.proxy.lib.siu.edu/10.1186/s13104-021-05596-7

[20] Wu, R., Shu, Z., Zou, F., Zhao, S., Chan, S., Hu, Y., Xiang, H., Chen, S., Fu, L., Cao, D., & Lu, H. (2022). Identifying myoglobin as a mediator of diabetic kidney disease: a machine learning-based cross-sectional study. Scientific Reports, 12(1), 1–14. https://doi-org.proxy.lib.siu.edu/10.1038/s41598-022-25299-8

[21] Ng, K., Kartoun, U., Stavropoulos, H., Zambrano, J. A., & Tang, P. C. (2021). Personalized treatment options for chronic diseases using precision cohort analytics. Scientific Reports, 11(1), 1–13. https://doi-org.proxy.lib.siu.edu/10.1038/s41598-021-80967-5

[22] Wang, S., Han, J., Jung, S. Y., Oh, T. J., Yao, S., Lim, S., Hwang, H., Lee, H.-Y., & Lee, H. (2022). Development and implementation of patient-level prediction models of end-stage renal disease for type 2 diabetes patients using fast healthcare interoperability resources. Scientific Reports, 12(1), 1–9. https://doi-org.proxy.lib.siu.edu/10.1038/s41598-022-15036-6

[23] Martinez-De la Torre, A., Perez-Cruz, F., Weiler, S., & Burden, A. M. (2022). Comorbidity clusters associated with newly treated type 2 diabetes mellitus: a Bayesian nonparametric analysis. Scientific Reports, 12(1), 1–13. https://doi-org.proxy.lib.siu.edu/10.1038/s41598-022-24217-2

[24] González-Juanatey, C., Anguita-Sánchez, M., Barrios, V., Núñez-Gil, I., Gómez-Doblas, J. J., García-Moll, X., Lafuente-Gormaz, C., Rollán-Gómez, M. J., Peral-Disdie, V., Martínez-Dolz, L., Rodríguez-Santamarta, M., Viñolas-Prat, X., Soriano-Colomé, T., Muñoz-Aguilera, R., Plaza, I., Curcio-Ruigómez, A., Orts-Soler, E., Segovia, J., Maté, C., & Cequier, Á. (2022). Assessment of medical management in Coronary Type 2 Diabetic patients with previous percutaneous coronary intervention in Spain: A retrospective analysis of electronic health records using Natural Language Processing. PLoS ONE, 17(2), 1–17. https://doi-org.proxy.lib.siu.edu/10.1371/journal.pone.0263277

[25] Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. Diabetology & Metabolic Syndrome, 13(1), 1–22. https://doi-org.proxy.lib.siu.edu/10.1186/s13098-021-00767-9

[26] Wu P, Zeng D, Wang Y. Matched Learning for Optimizing Individualized Treatment Strategies Using Electronic Health Records. Journal of the American Statistical Association. 2020;115(529):380-392. doi:10.1080/01621459.2018.1549050

[27] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T., Mulrow, C. D., Shamseer, L., Tetzlaff, J., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, n71. https://doi.org/10.1136/bmj.n71

[28] Research guides: Writing as inquiry: Boolean searching. Boolean Searching - Writing as Inquiry - Research Guides at New York University. (n.d.). https://guides.nyu.edu/WAI/Boolean