CHAPTER 9

Artificial intelligence and machine learning

Ashton Pike[1], Barlas Benkli[2], Sayyed Omar Gilani[1] and Salman Hirani[1]
[1]Department of Anesthesiology and Perioperative Medicine, School of Medicine, Oregon Health & Science University, Portland, OR, United States
[2]Department of Neurology, University of Texas McGovern Medical School at Houston, Houston, TX, United States

Introduction

As medicine advances and the healthcare system experiences the unprecedented need for healthcare providers, there is a strong incentive to enhance efficiency in diagnosing and effectively treating patients. One particularly novel avenue of research is artificial intelligence's (AI) applicability to healthcare. Initially, machine learning was limited by domain expertise and human engineering needed to transform raw data into appropriate representations that a learning algorithm used to discover patterns [1]. However, deep learning, a subset of machine learning, goes further by building its representations from raw data with multiple layers of abstraction, which are then coupled with an algorithm that indicates how much change there should be compared to its previous internal parameters [1,2].

The utility of deep learning is derived from its ability to feed on large amounts of data to train itself to be better. This allows it to view trends and patterns through categorizing, understanding, and predicting data [1]. Its current impact is difficult to quantify given limited integration into clinical practice. Still, simulation data demonstrate AI's potential to increase the accuracy and efficiency of screenings, reduce workload, and even diagnose disease noninferiorly compared to strictly clinician knowledge [3–5].

Deep learning has succeeded in medical imagery, where algorithms focus on image and video comprehension when engaging with object classification, segmentation, and detection. Studies on AI applications in breast cancer screening, cardiac imaging, and melanoma screening showed promising results [3–10]. This has evolved further with the development of convolutional neural networks (CNNs), deep learning algorithms that better process data with spatial invariance (i.e., images) [1,2]. In studies that examined the diagnostic utility of CNNs in object classification, the neural networks showed near or at-human level performance, going so far as to reach physician-level accuracy on a medley of diagnostic tasks, such as identifying skin cancer, cardiovascular risk, and breast cancer in mammograms [3,11,12]. A recent study by Abd-Alrazaq on

Substance Use and Addiction Research
DOI: https://doi.org/10.1016/B978-0-323-98814-8.00007-X

AI use during the COVID-19 pandemic showed that during this period, AI was used for a wide array of applications, from vaccine/drug discovery to diagnosis [13]. Currently, most applications of AI in use actively integrate clinicians and the programs to support clinicians, though in some cases, the combo model enhances accuracy and efficiency [6].

Natural language processing

Natural language processing (NLP) systems are promising for addiction research and quality improvement projects in pain medicine, as they utilize verbal data to describe patterns and form hypotheses. These systems can swiftly analyze free text from clinical documentation without a need for direct human supervision and, out of these texts, create structured data that can be used by clinicians or researchers who need to review large amounts of information [14–16]. The structured data can also be used in downstream informatics processes to identify secondary data points and potential clinical or epidemiological studies [15]. Such applications have already been used to identify clinical trial candidates, monitor drug reactions, screen clinic-pathological documents in cancer registries, and diagnose asthma [14,15,17–19]. As part of an effort led by the FDA and CDC to create a public NLP platform (named Clinical Language Engineering Workbench) able to capture structured data from free-text clinical data, Kreimeyer et al. did a systematic review that identified 35 systems/pipelines used to process clinical notes. However, the information provided on these systems was mostly scarce [15,20]. In the same systematic review, rule-based and hybrid approaches were most frequently utilized; some of the commonly used systems were NegEx algorithm, Weka machine learning workbench, Stanford CoreNLP, NLTK, and OpenNLP [15,21–25]. Recent studies show that with the improved use of machine learning, NLP methods also continue to evolve, where the potential use of statistical NLP methods can enhance the efficiency of such platforms [19,20,26].

Applicability to addiction medicine

Automated systems for detecting opioid use have been used in numerous hospital settings and reviewed in studies. Yet, most of these systems rely on structured electronic health record (EHR) data including diagnosis, procedure notes, billing codes, prescription history, demographics, and more [27–30]. Furthermore, many studies utilize standard statistical methods to identify risk factors for opioid overdose [31–36]. However, when utilizing these data and associated outcomes as predictors of overdose risk, these systems have demonstrated suboptimal performance, potentially due to their limitations in addressing nonlinear risk prediction and characterizing interactions between predictive variables [36–38]. Therefore deep learning's utility within addiction medicine would lie in its superior capability to handle complicated interactions in large data sets and to unveil novel patterns that could better inform clinical intervention.

Some studies build on this attribute by developing prediction models that predict overdose or stratify risk factors for opioid misuse through deep learning modeling on electronic health records [36,39–41]. One study noted almost a one-third increase in the prevalence of opioid misuse through NLP-assisted review compared to traditional surveillance by ICD-9 coding [42]. This could better inform clinicians on patients that warrant further workup and early intervention for misuse, with the goal being successful addiction treatment or prevention.

Compared to traditional models or statistical analyses, specific deep learning models can also better characterize trends in data [36]. Lo-Ciganic et al. compared various traditional and deep learning models in predicting overdose risk among Medicare beneficiaries in the three months after treatment with opioid prescriptions. These models included multivariate logistic regression, least absolute shrinkage and selection operator-type regression, random forest, gradient boosting machine (GBM), and deep neural network (DNN), with DNN and GBM models ultimately outperforming the other three. The DNN model went a step further by stratifying patients into low, medium, and high-risk categories, where greater than 90% of overdoses occurred in the high and medium-risk subgroups. Tseregounis et al. conducted a review of clinical prediction models assessing opioid overdose risk and found that models using machine learning techniques performed better than those utilizing standard logistic regression [36,41]. In designing potentially more accurate risk stratification in these populations, novel deep learning algorithms such as these demonstrate the potential of continually refined predictive models.

In the face of continuous development and refinement of modeling modalities within AI, there is still no consensus on which approach is superior in predicting opioid misuse or overdose. This can be attributed to the numerous parameters and variables set within each system that might be better optimized depending on the modality. For example, recurrent neural networks (RNNs) theoretically show promise in predictive modeling with EHR data processing large data sets without domain knowledge to include more clinical features into the picture, overcoming prior limitations of other deep learning modalities [43]. However, although DNNs can factor in the interaction between variables and are capable of adding a temporal component to their modeling, standard machine learning models using logistic regression, XGBoost, or random forest showed better performance in some studies, though predictive capabilities across machine learning were better than guideline-driven approaches [36,41,44]. One particularly useful feature of machine learning processes is their capability of being retrained to accommodate different populations, rendering them useful in characterizing data sets in an infinite number of ways [44]. Ultimately, this makes it difficult to conclusively recommend one particular modality of machine learning over the other. The decision of which modality to utilize will be based on the population of interest and variables under consideration.

If such modeling eventually becomes integrated with clinical decision-making, there is an opportunity to better intervene in patient care. Theoretically, predictive modeling could extract various clinical characteristics that would suggest a progression from short- to long-term opioid use with misuse/overdose and, in doing so, provide cues for clinicians to consider alternative patient management strategies. The financial implications of integrating these models into the prediction of opioid overdose or misuse are significant, as proper risk stratification can enable more targeted intervention for higher risk patients, allowing for resource allocation optimization for patients, insurance agencies, and medical providers.

Limitations and considerations

Despite the evidence demonstrating the potential for more accurate opioid overdose/misuse prediction, machine learning still has limitations. Many modalities to date fall under the umbrella of supervised learning, given the need for domain expertise and definitions in developing these models, and as such, autonomous modeling is quite rare [1,45]. Calibration is also a crucial consideration in prediction models, where inadequate calibration could over or underestimate prediction estimates, leading to inappropriate conclusions and downstream inefficient resource allocation [41]. Continuous training and calibration also remain difficult given their time-consuming nature, need for a plethora of data, and possible waste of useful information from prior modeling [46,47]. Updating trained models becomes a relatively tedious necessity.

One key limitation is that many of these models utilize patient databases retrospectively, as few are actively being utilized in clinical practice. This is not unexpected given that many predictive models for opioid misuse/overdose lack external validation in studies, which would be necessary to assess how they might perform in different populations [36,41,43]. For those that demonstrate good performance when externally validated, the limited evidence behind the benefits and harms of screening makes the role of automated algorithms unclear. Additionally, though many of the referenced studies demonstrated improved performance compared to traditional predictive models, the clinical significance of this is difficult to assess given the limited number of prospective studies evaluating actual clinical intervention informed by these predictions [48].

Though several studies demonstrate the potential promise of AI, the field is ultimately one in relative infancy. Simple deep learning algorithms show the capability of enhanced stratification and identification of opioid misuse compared to traditional models. Yet, many such studies show a dependence on domain definitions, implying potential human error requiring multiple attempts to refine the domain. Some previously described limitations in other studies noted the potential for nonprescription opioid consumption that could not be accounted for in modeling algorithms. This crucial data point could change prevalence and risk estimates [36,44]. Despite such limitations, machine learning's applicability to addiction medicine seems to currently focus on its role of clinical decision support for diagnosis and management, which might also be suggested by the limited number of studies demonstrating its role beyond this.

Furthermore, most studies mainly account for opioid overdose and misuse based on documented ICD coding and prescription disbursement. However, advances in machine learning through NLP or RNNs could sift through unstructured clinical data that might otherwise be missed. An example of this would be connecting patient comorbidities into machine learning algorithms to predict adverse outcomes post opioid administration, with one study demonstrating better prediction performance compared to using prescription documentation alone [44]. Such clinical data could, for example, identify higher risk neurocognitive or personality profiles in free text that then would be treated with targeted intervention programs [49,50].

Given the importance of domain knowledge and effective data, it is likely that data collection will be an important focus as machine learning becomes more readily implemented into analysis and clinical decision-making [51]. As the field continues to incorporate efficient and nimble approaches to data collection, machine learning algorithms could simultaneously stand to improve further, generating more accurate and reliable predictive values than what is available now. As the field advances, it is reasonable to believe that many of the current limitations of machine learning may be resolved. Ultimately, deep learning's influence on medical research and patient management will continue to expand.

Conclusion

As medicine advances and the healthcare system faces historic provider shortages, there is a tremendous motivation to improve diagnostic and treatment efficiency. The use of AI in healthcare is a new study area. Human engineering and domain expertise were initially required to turn raw data into acceptable representations for learning algorithms. A subset of machine learning, deep learning, builds its representations from raw input, with an algorithm indicating how much change should be made relative to its prior internal parameters. Deep learning's value comes from its capacity to learn from enormous volumes of data. It can see trends and patterns by classifying, analyzing, and forecasting data. Its present impact is difficult to evaluate because of poor clinical integration. Still, simulation data show AI's potential to improve screening accuracy and efficiency, reduce effort, and even identify illnesses noninferior to professional expertise. This is especially true in medical imaging, where algorithms use deep learning to classify, segment, and detect objects in images and videos. Promising outcomes were seen in AI-based breast cancer, cardiac imaging, and melanoma screening studies. CNNs are deep learning algorithms that better analyze input with spatial invariance (images). The accuracy of CNNs in diagnosing skin cancer, cardiovascular risk, and breast cancer in mammography was close to or at the physician level in trials examining their diagnostic value in object categorization. An Abd-Alrazaq research on AI use during the COVID-19 pandemic found that AI was employed for everything from vaccine/drug discovery to diagnosis. Now, most AI systems actively merge physicians and algorithms, albeit, in some circumstances, this approach improves accuracy and efficiency. The value of domain expertise and data collecting is projected to increase as machine learning is more readily applied to the analysis and clinical decision-making.

References

[1] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436−44.
[2] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med 2019;25(1):24−9.
[3] McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature 2020;577(7788):89−94.
[4] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542(7639):115−18.
[5] Brinker TJ, Hekler A, Utikal JS, Grabe N, Schadendorf D, Klode J, et al. Skin cancer classification using convolutional neural networks: systematic review. J Med Internet Res 2018;20(10):e11936.
[6] Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. JAMA Netw Open 2020;3(3):e200265.
[7] Kwon JM, Jeon KH, Kim HM, Kim MJ, Lim SM, Kim KH, et al. Comparing the performance of artificial intelligence and conventional diagnosis criteria for detecting left ventricular hypertrophy using electrocardiography. Europace 2020;22(3):412−19.
[8] Houssami N, Kirkpatrick-Jones G, Noguchi N, Lee CI. Artificial Intelligence (AI) for the early detection of breast cancer: a scoping review to assess AI's potential in breast screening practice. Expert Rev Med Devices 2019;16(5):351−62.
[9] Haenssle HA, Fink C, Toberer F, Winkler J, Stolz W, Deinlein T, et al. Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions. Ann Oncol 2020;31(1):137−43.
[10] Brinker TJ, Hekler A, Enk AH, Klode J, Hauschild A, Berking C, et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. Eur J Cancer 2019;113:47−54.
[11] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging 2018;9(4):611−29.
[12] Kooi T, Litjens G, van Ginneken B, Gubern-Merida A, Sanchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal 2017;35:303−12.
[13] Abd-Alrazaq A, Alajlani M, Alhuwail D, Schneider J, Al-Kuwari S, Shah Z, et al. Artificial intelligence in the fight against COVID-19: scoping review. J Med Internet Res 2020;22(12):e20756.
[14] Tucker TC, Durbin EB, McDowell JK, Huang B. Unlocking the potential of population-based cancer registries. Cancer 2019;125(21):3729−37.

[15] Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. J Biomed Inform 2017;73:14−29.
[16] Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020;26(8):1183−92.
[17] Wang G, Jung K, Winnenburg R, Shah NH. A method for systematic discovery of adverse drug events from clinical notes. J Am Med Inform Assoc 2015;22(6):1196−204.
[18] Ni Y, Kennebeck S, Dexheimer JW, McAneney CM, Tang H, Lingren T, et al. Automated clinical trial eligibility prescreening: increasing the efficiency of patient identification for clinical trials in the emergency department. J Am Med Inform Assoc 2015;22(1):166−78.
[19] Dharmage SC, Perret JL, Custovic A. Epidemiology of asthma in children and adults. Front Pediatr 2019;7:246.
[20] Blumenthal W, Alimi TO, Jones SF, Jones DE, Rogers JD, Benard VB, et al. Using informatics to improve cancer surveillance. J Am Med Inform Assoc 2020;27(9):1488−95.
[21] Apache OpenNLP. Available from: <http://opennlp.apache.org/>.
[22] Witten IH, Frank E, Hall MA. Data mining: practical machine learning tools and techniques. 3rd ed. Burlington, MA: Morgan Kaufmann; 2011. xxxiii, 629 pp.
[23] Bird S, Klein E, Loper E. Natural language processing with Python. 1st ed. Beijing; Cambridge, MA: O'Reilly; 2009. xx, 479 pp.
[24] Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 2001;34(5):301−10.
[25] Manning C.D., Surdeanu M., Bauer J., Finkel J., Bethard S., McClosky D., editors. The Stanford CoreNLP Natural Language Processing Toolkit. Association for Computational Linguistics; 2014.
[26] Jagannatha A, Liu F, Liu W, Yu H. Overview of the first natural language processing challenge for extracting medication, indication, and adverse drug events from electronic health record notes (MADE 1.0). Drug Saf 2019;42(1):99−111.
[27] Kunze KN, Karhade AV, Sadauskas AJ, Schwab JH, Levine BR. Development of machine learning algorithms to predict clinically meaningful improvement for the patient-reported health state after total hip arthroplasty. J Arthroplasty 2020;35(8):2119−23.
[28] Karhade AV, Bongers MER, Groot OQ, Cha TD, Doorly TP, Fogel HA, et al. Development of machine learning and natural language processing algorithms for preoperative prediction and automated identification of intraoperative vascular injury in anterior lumbar spine surgery. Spine J 2021;21(10):1635−42.
[29] Thompson HM, Hill K, Jadhav R, Webb TA, Pollack M, Karnik N. The substance use intervention team: a preliminary analysis of a population-level strategy to address the opioid crisis at an academic health center. J Addict Med 2019;13(6):460−3.
[30] Blackley SV, MacPhaul E, Martin B, Song W, Suzuki J, Zhou L. Using natural language processing and machine learning to identify hospitalized patients with opioid use disorder. AMIA Annu Symp Proc 2020;2020:233−42.
[31] White AG, Birnbaum HG, Schiller M, Tang J, Katz NP. Analytic models to identify patients at risk for prescription opioid abuse. Am J Manag Care 2009;15(12):897−906.
[32] Bohnert AS, Valenstein M, Bair MJ, Ganoczy D, McCarthy JF, Ilgen MA, et al. Association between opioid prescribing patterns and opioid overdose-related deaths. JAMA 2011;305(13):1315−21.
[33] Becker WC, Fiellin DA, Desai RA. Non-medical use, abuse and dependence on sedatives and tranquilizers among U.S. adults: psychiatric and socio-demographic correlates. Drug Alcohol Depend 2007;90(2−3):280−7.
[34] Webster LR, Webster RM. Predicting aberrant behaviors in opioid-treated patients: preliminary validation of the Opioid Risk Tool. Pain Med 2005;6(6):432−42.
[35] Ives TJ, Chelminski PR, Hammett-Stabler CA, Malone RM, Perhac JS, Potisek NM, et al. Predictors of opioid misuse in patients with chronic pain: a prospective cohort study. BMC Health Serv Res 2006;6:46.

[36] Lo-Ciganic WH, Huang JL, Zhang HH, Weiss JC, Wu Y, Kwoh CK, et al. Evaluation of machine-learning algorithms for predicting opioid overdose risk among Medicare beneficiaries with opioid prescriptions. JAMA Netw Open 2019;2(3):e190968.
[37] Hylan TR, Von Korff M, Saunders K, Masters E, Palmer RE, Carrell D, et al. Automated prediction of risk for problem opioid use in a primary care setting. J Pain 2015;16(4):380−7.
[38] Glanz JM, Narwaney KJ, Mueller SR, Gardner EM, Calcaterra SL, Xu S, et al. Prediction model for two-year risk of opioid overdose among patients prescribed chronic opioid therapy. J Gen Intern Med 2018;33(10):1646−53.
[39] Dong X, Rashidian S, Wang Y, Hajagos J, Zhao X, Rosenthal RN, et al. Machine learning based opioid overdose prediction using electronic health records. AMIA Annu Symp Proc 2019;2019:389−98.
[40] Che Z, St Sauver J, Liu H, Liu Y. Deep learning solutions for classifying patients on opioid use. AMIA Annu Symp Proc 2017;2017:525−34.
[41] Tseregounis IE, Henry SG. Assessing opioid overdose risk: a review of clinical prediction models utilizing patient-level data. Transl Res 2021;234:74−87.
[42] Carrell DS, Cronkite D, Palmer RE, Saunders K, Gross DE, Masters ET, et al. Using natural language processing to identify problem usage of prescription opioids. Int J Med Inform 2015;84(12):1057−64.
[43] Dong X, Deng J, Hou W, Rashidian S, Rosenthal RN, Saltz M, et al. Predicting opioid overdose risk of patients with opioid prescriptions using electronic health records based on temporal deep learning. J Biomed Inform 2021;116:103725.
[44] Sharma V, Kulkarni V, Eurich DT, Kumar L, Samanani S. Safe opioid prescribing: a prognostic machine learning approach to predicting 30-day risk after an opioid dispensation in Alberta, Canada. BMJ Open 2021;11(5):e043964.
[45] Mak KK, Lee K, Park C. Applications of machine learning in addiction studies: a systematic review. Psychiatry Res 2019;275:53−60.
[46] Siregar S, Nieboer D, Versteegh MIM, Steyerberg EW, Takkenberg JJM. Methods for updating a risk prediction model for cardiac surgery: a statistical primer. Interact Cardiovasc Thorac Surg 2019;28(3):333−8.
[47] Afshar M, Sharma B, Bhalla S, Thompson HM, Dligach D, Boley RA, et al. External validation of an opioid misuse machine learning classifier in hospitalized adult patients. Addict Sci Clin Pract 2021;16(1):19.
[48] Ward R, Weeda E, Taber DJ, Axon RN, Gebregziabher M. Advanced models for improved prediction of opioid-related overdose and suicide events among Veterans using administrative healthcare data. Health Serv Outcomes Res Methodol 2022;22:275−95.
[49] Ahn WY, Busemeyer JR. Challenges and promises for translating computational tools into clinical practice. Curr Opin Behav Sci 2016;11:1−7.
[50] Ahn WY, Ramesh D, Moeller FG, Vassileva J. Utility of machine-learning approaches to identify behavioral markers for substance use disorders: impulsivity dimensions as predictors of current cocaine dependence. Front Psychiatry 2016;7:34.
[51] Blanco C, Wall MM, Olfson M. Data needs and models for the opioid epidemic. Mol Psychiatry 2022;27:787−92.