Open Access

Editorial

Artificial Intelligence in Haematology

Nadeem Ikram

[1]Department of Pathology
Akhtar Saeed Medical College,
Rawalpindi

Address for Correspondence
Dr. Nadeem Ikram
Department of Pathology
Akhtar Saeed Medical College,
Rawalpindi
drnadeemikram@gmail.com

The increase in knowledge in the medical sciences is making decision making complex. Automation of the accumulated data and then its utilization for diagnosis, prognostication, treatment designing and many other patient's related matters leads to emergence of automation based tools. All these comes under the umbrella of Artificial Intelligence (AI). Artificial intelligence describes computational programs that simulate and mimic human intelligence, such as problem solving and learning. The AI has many subsets. Two important one are Machine Learning (ML) and Deep Learning (DL). ML involves automatic discovery of patterns within data by using statistical methods. On the other side of the spectrum lie Deep Learning (DL) models, which are multilayered networks of artificial 'neurons' designed to create accurate models from raw data. ML has two main categories, Supervised and Unsupervised. Supervised category needs input and output data, while unsupervised does not need it. In medical practice supervised ML is commonly employed. In supervised ML labeled in-put and out-put data can be applied by different algorithms. Then it can be used to accomplish the tasks of classification (e.g, support vectors, classifiers) and regression. In unsupervised learning, the data under scrutiny is investigated through the use of additional information not included in clustering model. The incorporation of the ensemble method (combining multiple algorithms to produce a more accurate model) creates "decision trees" (like random forest, XG Bost Baggina) which give robustness to these algorithms. In DL algorithms can deduce data's features/patterns through multiple layers of processes (i.e., Artificial Neural Network approach).

As the data increased it became quite a tiresome and difficult task for medical science experts to prepare a guide that will be used in the process of reading, simplifying and classifying. Getting information from the databases (data-mining) is the kind of method used for solving these problems. Statistical evaluations, like Cox's regression, Random Survival Forests, Concordance Index (c-index), and others are employed to get a meaningful and reproducible inference.

ML has an ability to process information from different diagnostic modalities and functions to diagnose, characterize, suggest therapeutic strategies and predict prognosis. Understanding of cancer, in general, has improved with the ever increasing amount of genetic and genomic data, still it is required to implement these large and complex data to it's fullest. ML approaches have shown tremendous potential in this regard. ML approaches have shown potential in the analysis of complex genetic data. Analysis genetic and genomic data form thousands of samples, from large number of studies, made it possible to develop "classifiers" that can detect haemtological malignancies, most likely leukaemias, in a near automated and low-cost fashion. Classifiers give better result by interpreting data with relevant data base. Various freely accessible online data sets (like "Leukaemia Gene Atlas", Beat-AML", "Cancer Genome Atlas",etc) are available to evaluate genetic risk profiles or identify novel genetic targets for individualized therapy. There is a vast multitude of machine learning algorithms and plate forms to implement. The resulting models can be evaluated based on their interpretability, tuning parameters, performance accuracy, computational resources, connectedness to cloud, Amazon web series, Microsoft – Azure- Login and others

Artificial Deep Neural Networks (D.N.N) are a subset of ML that imitate the neural structure of the brain.

Interconnected artificial neural networks can be applied for computer vision purposes, especially object detection, image segmentation, and classification. D.N.N can identify critical proteins associated with FLT3-ITD (FMS like tyrosine kinase 3 and internal tandem duplication (ITD). After adequate pre-processing of image data, D.N.Ns can be used in computer – aided diagnosis. Key steps in D.N.N – based assessment of bone marrow are cell – segmentation, extraction, quantification of cell-specific features and many others. Especially in leukaemias, D.N.N.s by using various segmentation techniques (filtering, enhancement, edge detection, feature extraction, etc), can analyze whole slides with automated focusing. In addition to D.N.N Support Vector Machines (SVM) and k-means clustering are other ML tools helpful in this regard. At first microscopic images are acquired from blood smears of patients with leukaemias and normal cases. After applying image processing, colour segmentation strategy is applied for segmenting white blood cells from other blood components and the discrimination features, i.e., irregularity, nuclear-cytoplasmic ration, Hansdorff dimensions, shape- colour and texture features are extracted from the entire nucleus in the whole images containing multiple nuclei. Because of the overlapping feature of blast cells, ML task is challenging in leukaemias. As compared to radiology, digitalization of slides is difficult in haematology, more so in higher magnification slides. AI algorithms, using deep convolational neural network, can automatically screen whole bone marrow smear for metastatic carcinomas cells. Images are classified to cancerous and non-cancerous images by binary support vector machine (S.V.M) classifier with 10- fold cross validation technique classifier performance is evaluated

ML can be used to develop novel prognostic indices or refine the understanding of already established prognostic mutational markers. ML can use large genomic data sets to predict targets for therapeutic agents. D.N.Ns have shown tremendous potential in identifying biomarkers and druggable targets and in the assessment of potential therapeutic molecules.

ML techniques validated key variables to predict post-transplant survival. The choice of conditioning regiemens and post-grafting immunosuppression can be tailored individually through large data-bases of specific immunogenetic environments of patients who underwent stem-cell transplantation

There are also data on the utility of flow cytometery-based M.L algorithms yielding diagnostic and prognostic information. Through Support Vector Machines (S.V.M) multiple algorithms are established to detect minimal residual disease in leukaemias and Myelodysplstic Syndrome, using data from multi-parameter flow cytometer.

ML and DL approaches are also been utilized in the classification of lymphomas using gene expression profile and DNA microarrays. These models are able to classify data into germinal centre B-cell and non-germinal centre B-cell with an accuracy of more than 90%.DL can distinguish diffuse large B-cell and Burkit's lymphoma based on H&E stained slides, fed into convolutional neural networks. Gene expression profiling enables diagnosis of AML without expertise intervention.

Microarray and bio-informatics approach for molecular diagnosis is a step ahead in A.I tools. A systematic approach in AI can be adopted for the monitoring of simultaneous expression of thousands of genes (Gene-profiling), using DNA microarray, independently of pervious biological knowledge

Digital imaging in pathology has undergone an exponential growth, catalyzed by changes in imaging hard ware and gains in computational processing. Digitalization of entire glass slide is now possible, at near the optical resolution limits. Whole slides can be imaged in fluorescence or by use of multispectral imaging system. Image acquisition uses web-based digital libraries or local image repositories

Artificial Neural Networks (A.N.Ns) are capable of detecting carrier status in haemoglobinopathies. They can also differentiate beta thalassaemia minor and iron deficiency. Machine learning algorithms showed predictable ferritin test results, giving credence to results of other tests, well-reflecting underlying iron-status.

Individual test results, if viewed in isolation, can lead to a conundrum in many instances. In this regard computation approaches to integrate lab tests results with clinical data and medical knowledge, judgment, and experience, offer tremendous potential to enhance

diagnostic value. In many instances accumulated test results of an individual, spanning over years, can lead to an overlook to key results. Electronic clinical decision support represents an important helping tool in this regard. AI is likely to unearth data which is not being explored or investigated in an individual patient, but is likely to help in overall management. AI narrow downs the waste of amount and energy potentially dwarfed by informationally redundant lab testing. In-built decision support system could potentially provide clinicians a list of tests predicted to be normal or abnormal at some specified confidence level. On a hind side predictors' setting can lead to a problem of "over-fitting". The issue of over-fitting can be checked by a large data base and by built- in regularization procedures

Stand – alone ML tools based on retrospective data are insufficient for widespread clinical use. In this regard an element of bias, at the level of system cannot be ruled out. Prospective validation in a proof-of-concept fashion is required. This can be ensured by an iterative workflow leading to a step-wise progression. Analysis of disagreements revealed errors in recording data and varied diagnostic criteria's are confounding. It can be subverted by AI generated "self –organized" hierarchical taxonomic classification patterns. It recognizes the place of input data in already discovered class or classes, and then can proceed to reinforce or to delete the information according to the designed criteria's.

All these applications are "data-hungry", needing a large number of samples. For a model to be useful, it should have clinical merit and generalization on populations not included in the development process. There is paucity of guidelines for AI, This is leading to a guarded approach in adopting a technology that still lacks best-practice bench marks. Inadequate data-base leads to issue of "over- fitting". Robust statistical settings are required to have a data not in any case considered as "garbage in-garbage out". It is required to identify potential biases as well as testing the model on variety of population. There is a degree of randomness in a patient's journey and events are not predictable. Dynamic system, which continues to update, based on new events, may increase the predictive confidence.

The "dataism" created by AI can lead to problems of dependency of humans on algorithms and data over judgment. Algorithms are agnostic to latent covariates, which may have an impact on patient's outcome. Application of ML technologies still have limitation of prospectiveness, real-life applications, ethical concerns, logistics and many other. These algorithms and techniques come with various pitfalls and need a strict regulatory framework to ensure safe use .Issues like transparency, reproducibility, over-fitting and comparability need to be addressed. However it is reasonable to assume that the field is promising, giving due credence to considerable amount of data generated and the ability of ML and DL to analyze big data bases in a very limited time.

AI has opened up a variety of integrative approaches in the field of haematology. It appears as if M.L is surpassing human comprehension. Inexpensiveness, easiness to interpret and reliability are few of the advantages of AI. The ever growing body of clinical data poses a challenge for researchers and clinicians alike to organize and interpret all this, which ultimately will improve patients' care. Once established, these algorithms are likely to provide haematologic expertise to regions without immediate access to large medical centers and help general practitioners to adequately screen for patients in need of haematologic assessment. As promising as these results of ML may be, however, there is still a long road ahead. Clinicians should be aware of common pitfalls in ML. A thorough oversight is crucial for the proper application of computer aided diagnostic and management algorithms

There is a fierce competition between human intelligence and machines. Even the smartest one's utilize their brain not more than thirty percent to its fullest. Dealing human brain as an app, through AI is there to mimic or to an extent human intelligence may be of some help. Rest assure even the working of the AI, ML, DL and other related tools are invented by humans and are, at the end, to be used by the humans. So, it is a win/win game, most importantly for the sufferers

Bibliography

- Dogan S and Turkogln I. Iron deficiency anaemia detection from haematology parameters by using decision trees. International Journal of Science and Technology. 2008; 3(1) : 85-92.
- Eckardt JN, Bornhauser M, Wendt K. Application of machine learning in the management of acute myeloid leukaemia. Current practice and future prospects. Blood Advances. 2020;4(23):6077-85.

- Zini G. Artificial intelligence in hematology. Hematology. 2005; 10(5):394-400
- Luo Y, Szolovits P, Dighe AS. Using machine learning to predict laboratory tests results. Am J Clin Pathol. 2016; 145: 778-88
- Ghaznan F, Evans A, Madabnushi A, Feldman M. Digital imaging in pathology: Whole –slide imaging and beyond. Ann Pew Pathol,.d images. Hindan Scientific Programming.2021; doi.org/10.1155/9933481
- Mushen IM, Shyr D, Sung AD, Hashmi SK. Machine learning applications in the diagnosis of benign and malignant hematological diseases. Clinical Hematology International.2021;3(1): 13-20
- Chen P, XuRO, Chen, Zhang L. Detection of metastatic tumour cells in the bone marrow aspirate smears by artificial intelligence (A.I) –based morphogo system. Frontiers in Oncology ,2021;11: doi.10.3389
- Shouval R, Fein JA, Savani B, Monty M. Machine learning and artificial intelligence in haematology. British Journal of Haematology, 2021; 192:329-50
- Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies .The Lancet HAematology, 2020;7(7): 541-55
- Joshna F, Shovl RA. American Society of Hematology. A Reader's Guide to Machine Learning in Hematology" https://www.hematology.org/education/trainees/fellows/trainee-news/2021/a-guide-to-machine-learning-in-hematology
- Kazanci EG and Guven D. Artificial intelligence applications in hematology. Artificial Intelligence Theory and Application.2021;1:1-7
- Siddiqui NS, KLEIN a, Godera A, Varga C. Supervised machine learning algorithms using patient related factors to predict in - hospital mortality following acute myeloid leukaemia therapy. Blood.2019; 139(1): 3435-40
- Kazeni F, Najafabadi TB, Arabi BN. Automatic recognition of acute myelogenous leukemia in blood microscopic images using k-means clustering and support vector machine. Med. Signals Sens.2016; 6(3): 183-93
- Salah HT, Muhsen IN, Salama ME, Owaidah T, Hashmi SK. Machine learning applications in the diagnosis of leukemia : Current trends and future directions. Int J Lab Hematol, 2019;41: 717-25
- Shafique S, Tehsin S. Computer-aided diagnosis of acute lymphoblastic leukaemia. Computational and mathematical methods in medicine. 2018 Feb 28;2018.

  Bigorra L, Merino A, Alferez S, Rodellar J. Feature analysis and automatic identification of leukemic lineage blast cells and reactive lymphoid cells from peripheral blood cell images. J. Clin. Lab. Anal.2017;31(2):e22024.