Artificial Intelligent in Healthcare

Anna Meiliana[1,2,]*, Nurrani Mustika Dewi[2], Andi Wijaya[1,2]

[1]Postgraduate Program in Clinical Pharmacy, Padjadjaran University, Jl. Eijkman No.38, Bandung, Indonesia
[2]Prodia Clinical Laboratory, Jl. Cisangkuy No.2, Bandung, Indonesia

*Corresponding author. E-mail: anna.meiliana@prodia.co.id

Abstract

BACKGROUND: Giant transformations are going on currently in health care, and the greatest force behind this phenomenon is data.

CONTENT: Big data has arrived into medicine field, lead to potential enhancement in accountability, quality, efficiency, and innovation. Most updated, artificial intelligence (AI) and machine-learning (ML) techniques rapidly developed, bring forth the big data analysis into more useful applications, from resource allocation to complex disease diagnosis. To realize this, a very large set of health-care data is needed for algorithms training and evaluation, including patients' treatment data, patients respond to treatment, and personal patient information, such as genetic data, family history, health behavior, and vital signs.

SUMMARY: Precision Health involving preventive, predictive, personalized and precise. The arrival of AI and ML will enhance and facilitates the improvement of this relationship through better accuracy, productivity, and workflow, thus develop a health system that will go beyond just curing disease, but further into wellness that preventing disease before it strikes, thus the patient–doctor bond is expected to be reformed and not be eroded.

KEYWORDS: artificial intelligence, machine learning, deep learning, electronic health records, big data

Indones Biomed J. 2019; 11(2): 125-35

Introduction

Demographic, economic and technological count their own role in health care, and any changes are forcing us to reconsider its delivery, funding models, education, scientific innovation, how it affect the physician, pharmaceutical companies, payors and patients.(1) The health care circle growing to a bigger and more complex connection each other, at once send a challenge message for us whose job is to heal, innovate and educate.(1) Recently, two major facts are plastered in medicine face. The first is a failed business model, while expenditures and job demands keep increasing in healthcare, the key outcomes deteriorate. Life expectancy reduced, infant, childhood, and maternal mortality keep high. The second is data generation in massive quantities, from high-resolution medical imaging, biosensors with continuous output of physiologic metrics, genome sequencing, and electronic medical records.(2) In deeper observations, an enormous of resources wasted, a large number of serious diagnostic errors and mistaken treatments occurred, due to inefficiencies in workflow, inequities, and inadequate time between patients and clinicians.(3,4) Eager for improvement, leaders in healthcare and computer scientists have asserted that artificial intelligent (AI) might be utilized in addressing all of these problems.(2)

AI typically refers to a machine with human capabilities. Some activities such as speech recognition, learning, planning, and problem solving can be designed for AI.(5) The widespread growth of big data supporting AI in industry offers a preview of their promise and peril in medicine and biomedical research, help guide clinical decisions is a central aspect of precision medicine initiatives; machine learning (ML) may refer either to a

set of computational and statistical tools for identifying relationships in data or to the use of such tools to make predictions based on data; and deep neural networks (DNN) are a particular type of ML whose success at tasks, such as image recognition, has led to them being referred to as AI or deep learning (DL).(5) To realize such potential gains, clinicians will need to acquire a basic knowledge of the workings of ML and AI to evaluate new developments and use the tools they make possible, just as in interpreting the results of clinical trials and in conducting research.(5) AI will increase efficiency, take over many routine tasks, and provide a first and last set of eyes. Our medical teams will become increasingly multidisciplinary and consist of clinicians, engineers, and data scientists.(6) We remain optimistic that big data and AI will transform our lives as clinicians and researchers. It has the potential to reduce the time we spend on cumbersome tasks, warn us about signs of trouble, help catch potential errors, and free us to provide higher quality, lower cost care.(7)

Medical Big Data

Data collection systems has grown in advances for the last decade, as well as our ability to effectively analyze and understand it. Nowadays, information become easier to collect, analyze and understand. At some point, the data grow massively big and complex, to be dealt traditionally,

while we know that today, data is a powerful force driving many industries including health care forward.(1) Big data refer to extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions. It is often defined by '3Vs': volume (large amounts of data), velocity (high speed of access and analysis), and variety (substantial data heterogeneity across individuals and data types).(8) Big data is helping every industry become more efficient and productive in quality and cost, therefore improve outcomes and achieve new solutions (as cited from "How Big Data Keeps Transforming Healthcare," LinkedIn, May 2017; "Biomedical Data Science Initiative," Stanford news, July 2015) (Figure 1).(1,9)

Medical data grows fast, estimates 48% each year. Computing power and technical ability boost the remake of medical field with the influx of big data in advanced ways never thought before.(1) It has been long expected that big data will substantially improve health care, but why does it matter? So far, big data research application have been put into two rough groups: long-practiced analysis approaches and newer methods using machine learning and AI.(10) In such a way, big data can be used as a powerful tools to promote health care improvement (11), by tidying up remains relatively untracked and under-analyzed records, hope to mend current ineffective treatment, substantial waste, medical error, and the systemic improvement by more understanding what works and what doesn't (12).
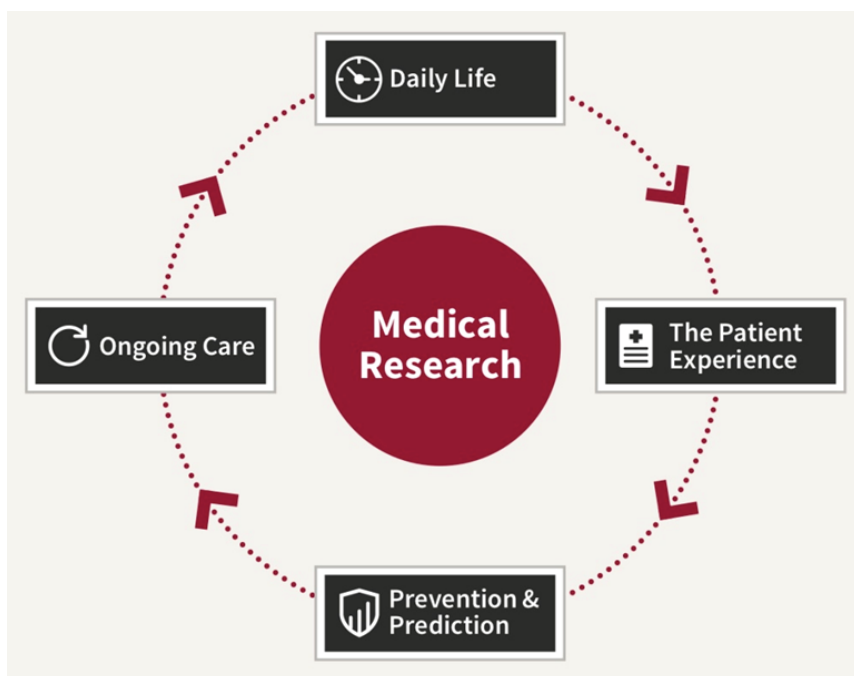

Figure 1. Data's impact throughout each step in the health care cycle: from medical research, to daily life, to the patient experience, to ongoing care, to prediction and prevention. (9) (Adapted with permission from Stanford Medicine).

Big data can be leveraged to measure hospital quality, as in the Centers for Medicare and Medicaid Services' Hospital Inpatient Quality Reporting program (13); to develop scientific hypotheses, as in proliferating genome-wide association studies (14); to compare the effectiveness of different interventions, as in the Patient Centered Outcome Research Institute (http://www.pcori.org); and to monitor drug and device safety, as with the Food and Drug Administration (FDA) Sentinel system (15).

Applying AI methods in big data, catch the pattern to make predictions and recommendations, is rapidly developing either in clinics or in daily life.(16,17) The popularity of wearable devices, such as pedometers and heart rate monitors, is exploding. Various types of home testing, including digitized glucose tests, blood pressure and genetic testing, are allowing individuals to take own control of their health, supported with many consultation tools in many websites. Thus, long before patients see the doctors, they have already their own day-to-day health data records.

More advanced in clinical applications involves image analysis equipped with algorithms make it possible to identify cancerous skin lesions from images as accurately as trained dermatologists1.(18) IDx-DR system even received FDA approval for image-based AI diagnosis of diabetic retinopathy. Furthermore, AI can be used for prognostic purposes so immediate intervention can be done (19), or when a patient is very likely to die in within short time then might consider shifting from traditional care to palliative care (20). On other side, the new age of health care also creates a tension regarding the data sharing. Some experts agree that infrastructure and policies encouraging the essence of sharing of data to maximize its benefits, while the others care about cybersecurity attacks and patients' privacy. Thus it's remain to be solving how to get benefit from public data sharing while ensuring the safe and secure of patient information.(21)

Electronic Health Record (EHR)

A medical record is narrow contains the patient's medical history, diagnoses and treatments by a particular physician, nurse practitioner, specialist, dentist, surgeon or clinic while an EHR is more comprehensive, an individual's shareable official health document that could be accessed via mobile devices such as smartphones and tablets. EHR comprise of demographic information, diagnoses, laboratory tests and results, prescriptions, radiological images, clinical notes, and more relevant patient information.(22) More than

improving operational efficiency, some studies found more application based on EHR in hospital and ambulatory care setting, especially compared to the traditional paper way or medical and health recording, EHR give benefits over time saving, environment saving by reducing papers, costs lowering, easier access, and reducing concerns regarding illegibility.(23-26)

Since 15 years ago, foundation of AI has been laid in healthcare for innovation, started from digitalization the medical data record for large numbers of patients and asymptomatic individuals (27), and improving the hardware with powerful computing platforms based on high performance computing and cloud computing, especially in graphics processing units (28). In particular, the patient data contained in EHR systems has been used for such tasks as medical concept extraction (29,30), patient trajectory modeling (31), disease inference (32-33), clinical decision support systems (34), and more. The application of these improvement help hospitals and clinics to minimize errors, improve care coordination, and provide more data for researchers.(35) In terms of functionality, EHR systems can be categorized into: basic EHR without clinical notes, basic EHR with clinical notes, and comprehensive systems. (25). Basic EHR at least provide the information of patient's medical history, complications, and medication usage.

Machine Learning and Deep Learning

Digital medicine, defined as applying digital tools in the practice of medicine to improve one to be high definition and individualized, holds promise in revolutionizing healthcare and patients adherences. The revolution core is the development of processing and integrating vast amount of individual data at the population levels and utilize them to address the health problems and challenges faced by patients, clinicians and health systems alike.(36) Some industries come out with smarter choices, Amazon use the basis of similar customers' shopping patterns to targets product recommendations. Google updates its search results using the outcomes of previous searches as a basis. Waze optimize the directions suggestion uses information from drivers traveling similar routes.(37)

Much attention has been devoted recently to the development of ML algorithms with the goal of improving treatment policies in healthcare.(38) Different from other computer, ML transforms the algorithms input into outputs using statistical, data-driven rules that are automatically derived from a large set of examples, rather than being

explicitly specified by humans. Traditionally, ML algorithm input must be hand-crafted from raw data by a practitioner domain expertise and human engineering who have knowledge to determine explicit patterns of prior interest, to design feature extractors that transformed raw data into suitable representations from which a learning algorithm could detect patterns. The engineering process is laborious and time consuming, subjectively depend on the "black art" (39), requiring creativity, trial-and-error, and oftentimes luck for creating, analyzing, selecting, and evaluating appropriate features.

DL techniques in contrast, is a form of representation learning. The machine learn optimal features directly from the data itself, without any human guidance, allowing for the automatic discovery of latent data relationships and pattern recognition, which is composed of multiple layers of representations, that might otherwise be unknown or hidden. These layers are typically arranged sequentially and composed of a large number of primitive, nonlinear operations. At the beginning one layer is represented by the raw data input, then it is fed into the next layer and transformed into a more abstract representation. The data flows through the layers of the system, while the input space becomes iteratively warped until data points become distinguishable (Figure 2). Finally based on this manner, highly complex functions can be learned.(40)

DL performed a deeper hierarchical feature construction and in effective manner capturing long-range dependencies in data (41), this approach achieved great success applied EHR data for clinical informatics tasks (42-44) compare to traditional methods, and showed better performance while require less time-consuming preprocessing and feature engineering (Figure 3).(45)

As a subfield of ML, DL shown a dramatic resurgence in the past 6 years, especially supported by increases in computational power and datasets. The machine in advances showed its ability to understand and manipulate data, including images (46), language (47), and speech (48). Healthcare and medicine have their benefits by generated vast amount of data (150 exabytes or 1018 bytes in United States alone, growing 48% annually (1)), impact on the increasing proliferation of medical devices and digital record systems.(40)

The ability to classify images and detect objects in a picture significantly improved along with the development of convolutional neural network (CNN) layers, involving application of multiple processing layers to which image analysis filters, or convolutions.(49-50). CNN replaces the multiple steps of previous image analysis methods. It represent the abstract images by constructing each layer systematically, convolving multiple filters across the image, producing a feature map that is used as input to the following layer. This image processing take image pixels as input, and the image classification as output (Figure 4).(51)

Majority DL algorithms was built upon the framework of artificial neural network (ANN). While only the last layer of a CNN is fully connected whereas in ANN each neuron is connected to every other neurons, as ANNs are composed of a number of interconnected nodes (neurons), arranged in layers. Neurons not contained in the input or output layers are referred to as hidden units. Every hidden unit stores a set of weights which are updated as the model is trained. Using DL methodologies, predictive models are keep improving, with ambitions to equal the human interpretability, since correct clinical decision-making determining the patients' fate, though
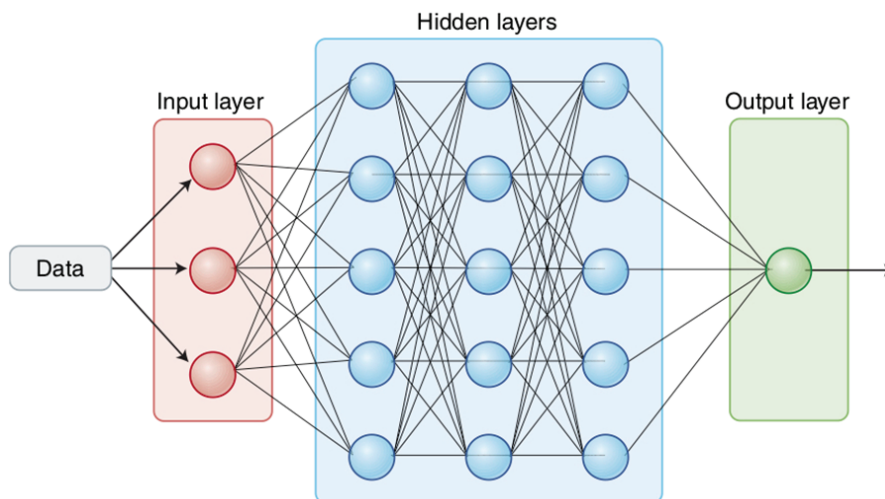

Figure 2. A deep neural network, simplified.(40) (Adapted with permission from Springer Nature).
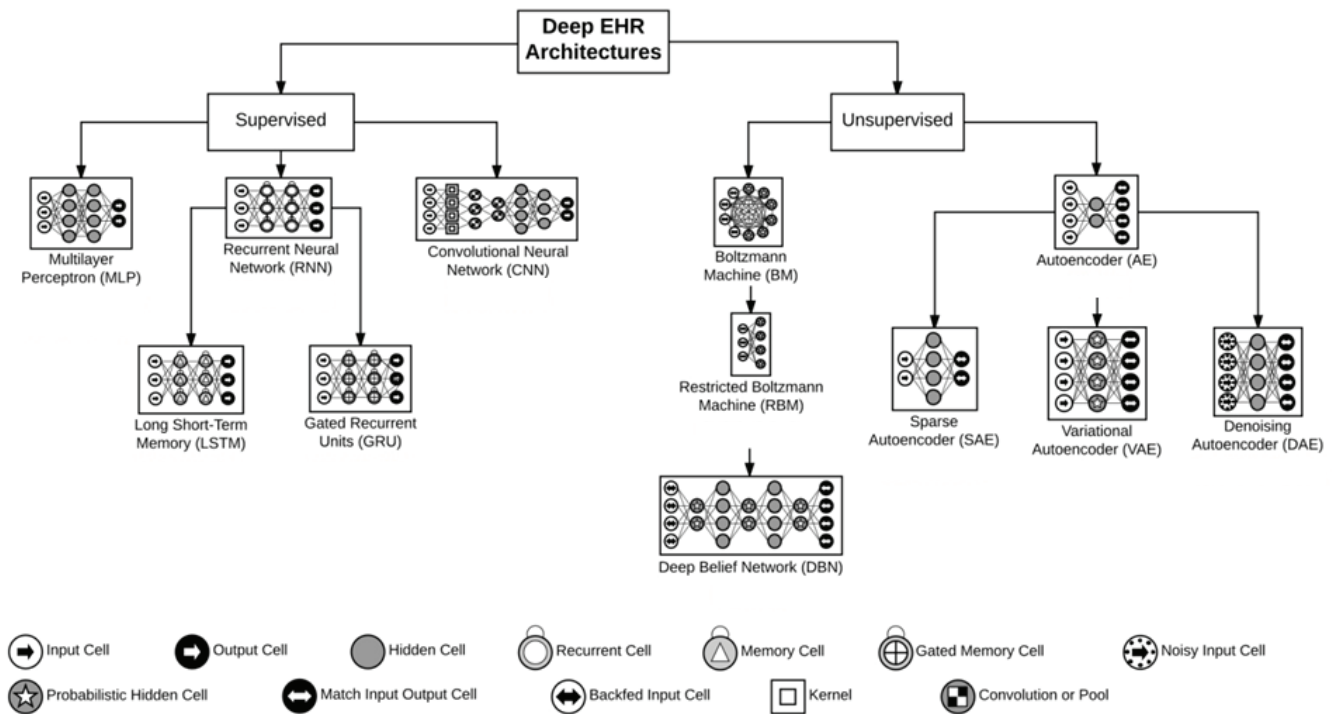
Figure 3. The most common deep learning architectures for analyzing EHR data.(45) (Adapted with permission from Association for Computing Machinery Inc).

some researchers accentuate the model performance and downplay the importance of interpretability. Currently, life-or-death decisions from systems is only marginal improved in accuracy over a human practitioner, but it warrant greater transparency than systems with near-perfect accuracy, or those with lower stakes. The transparency in deep learning will hasten the widespread adoption of such methods in clinical practice. Thus, we expect deeper EHR interpretability in the future.(45)

Physicians need to understand about the genetics of disease to recommend treatments and provide more accurate diagnoses. However, novel variants in a patient's genome which is medically relevant needs more prediction on mutations pathogenicity, and this will need features like protein structure and evolutionary conservation to train learning algorithms.(52) DL techniques with their greater power and ability to effectively integrate disparate data types, are likely to provide more accurate pathogenicity predictions than are possible today.(53) ML perform a phenotype prediction from genetic data, including diseases risk. DL with additional modalities such as medical images, clinical history, and wearable device data can enhance such models.(54) A particularly promising approach to phenotype prediction is to feed the intermediate molecular phenotypes prediction (*e.g.*, gene expression or gene splicing) into downstream disease predictors.(55) It is easier to predict

intermediate molecular states than human traits due to its larger, more proximal signals and more extensive training data, make it a good fit for DL, which has shown success at predicting splicing (56) and transcription factor binding (57). DL as an improvement iteration of ML, with huge medical data repositories and advanced learning algorithms together with physicians put together is now performing at state-of-the-art levels in previously difficult tasks, identify patterns in sparse, noisy data and requires little input-feature engineering, including image analysis, language processing, information retrieval, and forecasting. Future physicians armed with DL from models continuously trained and updated on real-world clinical data will possible to make accurate diagnoses and individually optimized treatment decisions.(37)

Artificial Intelligence

AI implementation grows rapidly in many industries including healthcare industries. it was expected to augment the ability of humans to provide healthcare, mimic the diagnostic abilities of physicians.(47,51) AI involves the development of computer algorithms, including but not limited to ML, representation learning, DL, and natural language processing to perform tasks typically associated
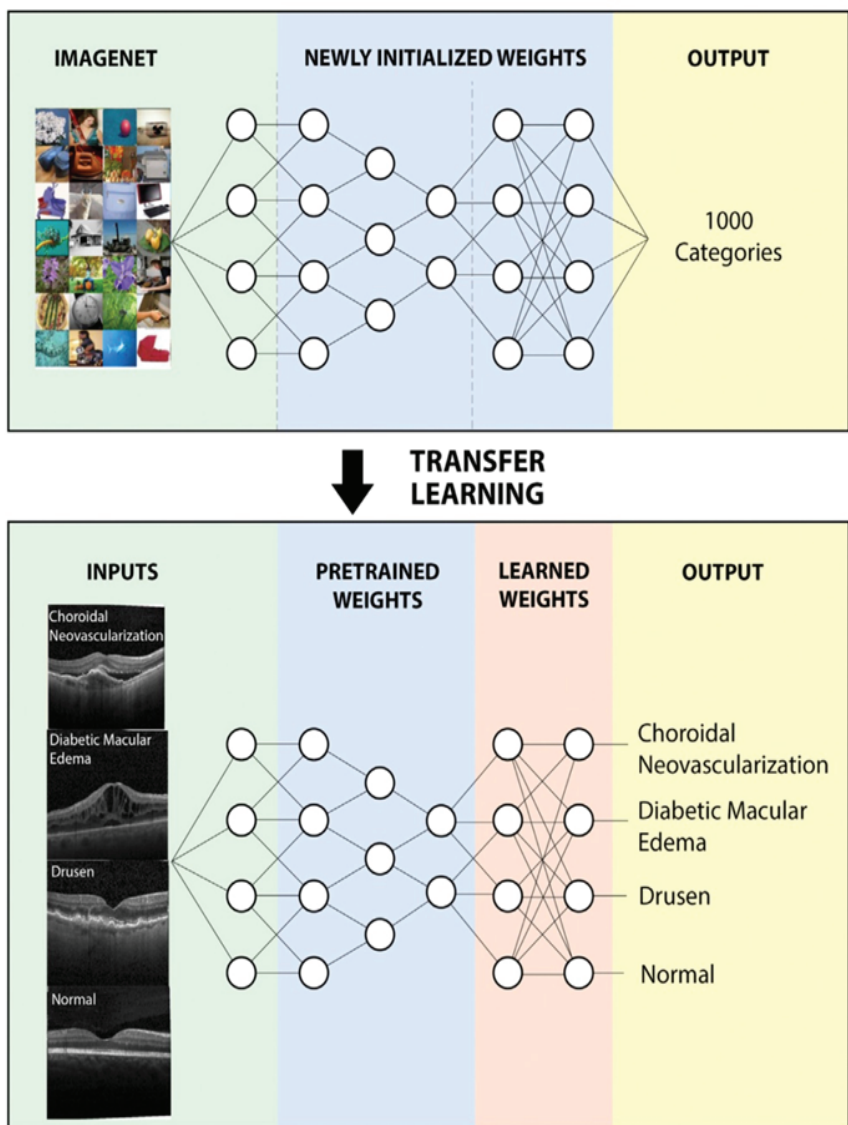
Figure 4. Schematic of a Convolutional Neural Network.(51) (Adapted with permission from Elsevier).

with human intelligence.(41) Despites of the techniques used here, involving this technology in medicine aims to harness computer algorithms to uncover relevant information from data and use it to assist clinical decision-making, from diagnosis generation and therapy selection, making risk predictions and stratifying disease, reducing medical errors, and improving productivity.(58-60) AI is more reliable in terms greater consistency, speed, and reproducibility than humans, so AI have a chance in human replacement scenario, but unlikely to replace the entirely system. AI can be applied in estimating bone age on radiographic exams (61), diagnosing treatable retinal diseases on optical coherence tomography (51,62), or quantifying vessel stenosis and other metrics on cardiac imaging (63). Tasks automation using AI can incredibly saving labor and time, and the human capital may be freed for an improved used.

Studies have demonstrated better results by synergizing clinicians and AI work together. Clinicians augment real-time clinical decision support, resulting in improved efforts toward precision medicine with help of AI.(64-66) Table 1 shows some AI that have been approved by FDA.

AI algorithms implementation has been effectively proven in many clinical settings, such as facilitating stroke, autism or electroencephalographic diagnoses for neurologists (67,68), helping anesthesiologists avoid low oxygenation during surgery (69), diagnosis of stroke or heart attack for paramedics (70), finding suitable clinical trials for oncologists (71), selecting viable embryos for in vitro fertilization (72), help making the diagnosis of a congenital condition via facial recognition (73) and pre-empting surgery for patients with breast cancer (74). Using AI, 250 images could be processed per day for the cost of

Table 1. FDA AI approvals are accelerating.(2) (Adapted with permission from Springer Nature).

| Company	FDA Approval	Indication
Apple	September 2018	Atrial fibrillation detection
Aidoc	August 2018	CT brain bleed diagnosis
iCAD	August 2018	Breast density via mammography
Zebra Medical	July 2018	Coronary calcium scoring
Bay Labs	June 2018	Echocardiogram EF determination
Neural Analytics	May 2018	Device for paramedic stroke diagnosis
IDx	April 2018	Diabetic retinopathy diagnosis
Icometrix	April 2018	MRI brain interpretation
Imagen	March 2018	X-ray wrist fracture diagnosis
Viz.ai	February 2018	CT stroke diagnosis
Arterys	February 2018	Liver and lung cancer (MRI, CT) diagnosis
MaxQ-AI	January 2018	CT brain bleed diagnosis
Alivecor	November 2017	Atrial fibrillation detection via Apple Watch
Arterys	January 2017	MRI heart interpretation

about $1,000 (75), representing time and cost saved. AI-assisted image interpretation can improve the workflow, reduce workforce for administrative jobs such as coding and billing, scheduling of operating rooms and clinic appointments, and staffing, finally increase the productivity (Figure 5).(2)

Artificial Intelligence in Medical Practice

DL, known as a type of ML models consisting of interconnected nodes called neurons arranged in a network layout, inspired by how human brains works and connected. Neurons in DL. Mimicking human brains, a neuron in DL is a mathematical approximation of a biological neuron, which transforms the data input to an output value. In Figure 6, electrocardiogram (ECG) input activated the DL neurons, processing it to classify heart rhythms and impaired left ventricular systolic function.(28,76,77) By predicting the outcome faster, theoretically, hospital care resources can be more efficient and precise managed, including the patients discharge.(2)

Some DL algorithms has been FDA approved, enable the public to take their healthcare into their own hands. In 2017, a smartwatch algorithm using photoplethysmography and accelerometer sensors was FDA-cleared to detect atrial fibrillation, and subsequently in 2018 the Apple Watch Series 4 approved.(78,79) The devices learn user heart rate at rest and with physical activity, equipped with a haptic warning to record an ECG via the watch when a significant deviation noticed. Recorded ECG will accurately detect whether there is high potassium in the blood, give benefits for patients with kidney diseases.(2) Smartphone apps AI for medical adherences also developed rapidly, including skin lesions and rashes, ear infections, migraine headaches, and retinal diseases such as diabetic retinopathy and age-related macular degeneration. AiCure (NCT02243670) demand the patient to take selfie video as they swallow their prescribed pill. Other apps use image recognition of food for calorie and nutritional content, with multimodal data AI guide an individualized diet and has a chance to be developed into a personal medical coach in thef uture.(80) The matchmaking apps has been reformed to matchmaking doctors and patients with higher levels of trust.(81)
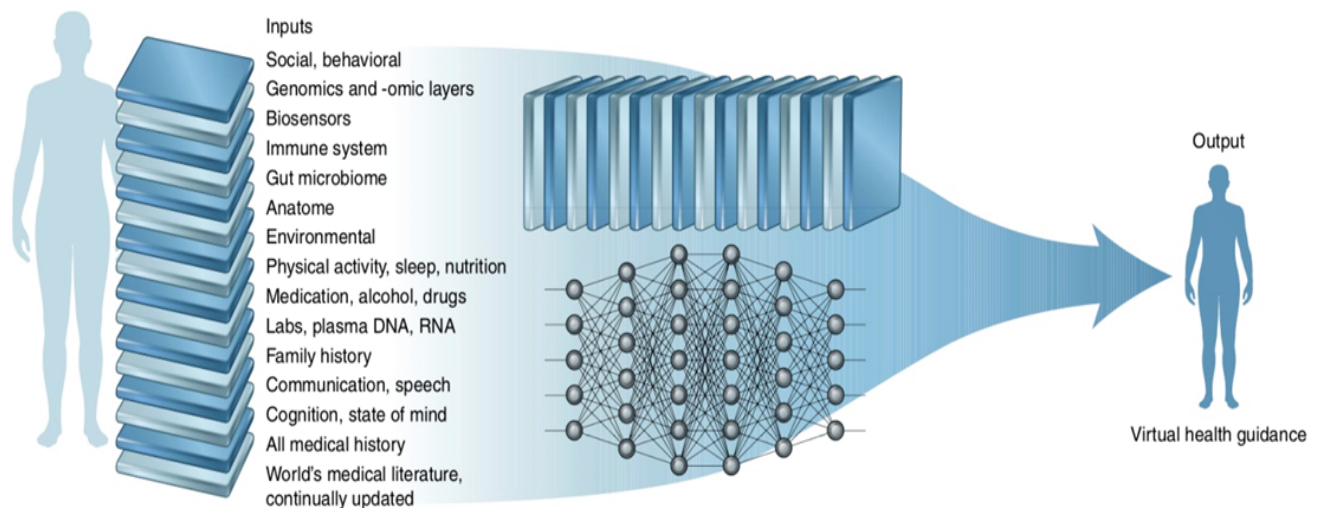
Figure 5. Examples of AI applications across human lifespan.(2) (Adapted with permission from Springer Nature).
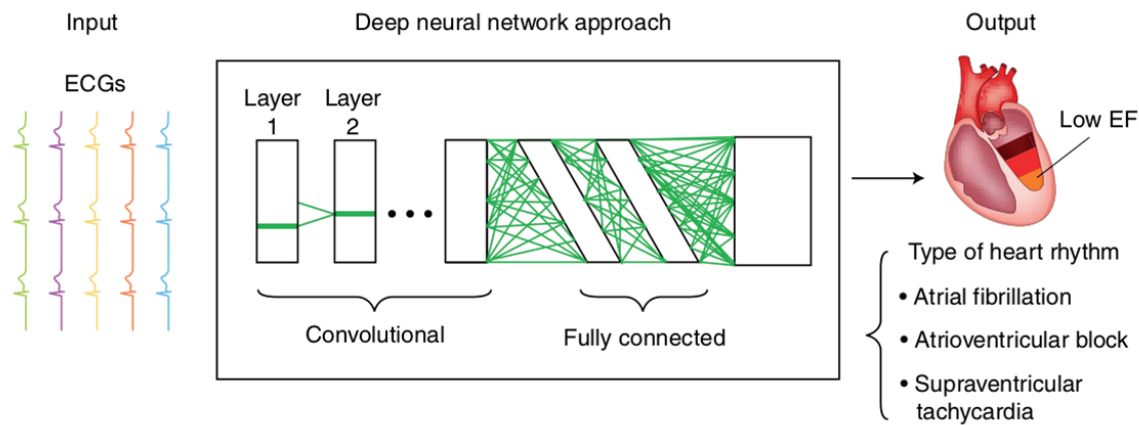
Figure 6. A deep neural network approach for analyzing electrocardiograms.(28) (Adapted with permission from Springer Nature).

One study recorded a continuous sensing of blood-glucose (for 2 weeks) along with assessment of the gut microbiome, physical activity, sleep, medications, all food and beverage intake, and a variety of lab tests, and apply the data to predict the individual glycemic response to specific foods, a physiologic pattern that is remarkably heterogeneous among people and significantly driven by the gut microbiome.(82-84). Calibrated continuous glucose sensors will precede the need for finger-stick glucose calibrations, and give an interesting fact that post-prandial glucose spikes commonly occur, even in healthy people without diabetes (85,86), suggesting an association with gastrointestinal barrier dysfunction (87,88) in experimental models.

The latest challenge encounter by ML and DL recently is the analytics of genomic and other -omics biology datasets. Open-source algorithms mostly in single -omic approach are available for classifying or analyzing whole-genome sequence pathogenic variants (89-91), somatic cancer mutations (92), gene–gene interactions (93), RNA sequencing data (94), methylation (95), prediction of protein structure and protein–protein interactions (96), the microbiome (97), and single cells (98). Facilitates by algorithmic prediction of CRISPR guide RNA activity (99) and off-target activities, multi-omic algorithms has been developed and integrate the datasets (100). AI also transform the drug discovery in many levels, including biomedical literature searching, data mining of millions of molecular structures, designing and making new molecules, predicting off-target effects and toxicity, predicting the right dose for experimental drugs, and developing cellular assays at a massive scale.(101-104) Animal testing may be reduced by the implication of ML prediction of toxicity (105), and AI cryptography make it possible to combine large proprietary pharmaceutical company datasets and discover previously unidentified drug interactions (106).

As promising as it was, AI still face some obstacles and pitfalls, especially when it comes to validation and readiness for implementation in patient care. A recent example is IBM Watson Health's cancer AI algorithm (known as Watson for Oncology). Based on n a small number of synthetic, non-real cases with very limited input (real data) of oncologists, this algorithms have been used by hundreds hospitals around the world, sending many error output for treatment recommendations, such as suggesting the use of bevacizumab in a patient with severe bleeding, which represents an explicit contraindication and black box warning for the drug.(107) We expect for software that will ingest and meaningfully process massive sets of data quickly, accurately, and inexpensively also for machines that will see and do things that are not humanly possible. This will ultimately lay the foundation for high-performance medicine, truly data driven, reducing our dependences in human resources, and the symbiosis should take us beyond the sum of the parts of human and machine intelligence, and lastly is realistic to be implemented widely.(2)

Conclusion

Will AI replace doctors and researchers? No, but a clinician who uses AI will replace clinicians who do not. Provided big data along with enhanced computing power, the availability of cloud storage, has enable the implications of AI and DL including in healthcare industry. For the beginning, clinicians benefit from rapid and accurate image interpretation, the health system improved their workflow and reduced medical errors, and patients enable

to process their own data to promote health. Overtime, accuracy, productivity, and workflow will have a remarkable improvement, and the patient–doctor relationship is expected to be reformed and not be eroded.

References

1. Stanford Medicine. Stanford Medicine 2017 Health Trends Report: Harnessing the Power of Data in Health. Stanford: Stanford University School of Medicine; 2017.

2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019; 25: 44-56.

3. Singh H, Meyer AND, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. BMJ Qual Saf. 2014; 23: 727-31.

4. Berwick DM, Hackbarth AD. Eliminating waste in US health care. JAMA. 2012; 307: 1513-6.

5. Rodriguez F, Scheinker D, Harrington RA. Promise and perils of big data and artificial intelligence in clinical medicine and biomedical research. Circ Res. 2018; 123: 1282-4.

6. Yeung S, Downing NL, Fei-Fei L, Milstein A. Bedside computer vision - moving artificial intelligence from driver assistance to patient safety. N Engl J Med. 2018; 378: 1271-3.

7. Castelvecchi D. Can we open the black box of AI? Nature. 2016; 538: 20-3.

8. Executive Office of the President. Big data: seizing opportunities, preserving values. Washington D.C: The White House of Barack Obama; 2015.

9. Linked In [Internet]. Marr B. How Big Data Keeps Transforming Healthcare. Linkedin May 2016 [updated 2016 May 27; cited 2019 Jan 10]. Available from: https://www.linkedin.com/pulse/how-big-data-keeps-transforming-healthcare-bernard-marr/.

10. Price WN, Cohen IG. Privacy in the age of medical big data. Nat Med. 2019; 25: 37-43.

11. Hoffman S. Electronic Health Records and Medical Big Data. New York: Cambridge University Press; 2016.

12. Kohn LT, Corrigan JM, Donaldson MS, eds. To Err is Human: Building a Safer Health System. Washington DC: National Academies Press; 2000.

13. Centers for Medicare and Medicaid Services [Internet]. Hospital Inpatient Quality Reporting Program [updated 2017 Sep 19; cited 2019 Jan 10]. Available from: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html.

14. Kohane IS. Using electronic health records to drive discovery in disease genomics. Nat Rev Genet. 2011; 12: 417-28.

15. Behrman RE, Benner JS, Brown JS, McClellan M, Woodcock J, Platt R. Developing the sentinel system—a national resource for evidence development. N Engl J Med. 2011; 364: 498-9.

16. Price WN. II Black-box medicine. Harv JL Tech. 2016; 28: 419-67.

17. Terry NP. Appification, AI, & healthcare's new iron triangle. J. Health Law Policy. 2018; 21: 3020784. doi: 10.2139/ssrn.3020784.

18. 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: 115-8.

19. Liu NT, Holcomb JB, Wade CE, Batchinsky AI, Cancio LC, Darrah MI, *et al.* Development and validation of a machine learning algorithm and hybrid system to predict the need for life-saving interventions in trauma patients. Med Biol Eng Comput. 2014; 52: 193-203.

20. Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. BMC Med Inform Decis Mak. 2018;18(Suppl 4):122. doi: 10.1186/s12911-018-0677-8.

21. Data Center Knowledge [Internet]. Riccio K. Big Data Experts in Big Demand. [updated 2017 May 30; cited 2019 Jan 15]. Available from: https://www.datacenterknowledge.com/archives/2017/05/30/big-data-experts-big-demand.

22. Birkhead GS, Klompas M, Shah NR. Uses of electronic health records for public health surveillance to advance public health. Annu Rev Public Health. 2015; 36: 345-59.

23. Botsis T, Hartvigsen G, Chen F, Weng C. Secondary use of EHR: data quality issues and informatics opportunities. Summit Transl Bioinform. 2010; 2010: 1-5.

24. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet. 2012; 13: 395-405.

25. The Office of the National Coordinator for Health Information Technology [Internet]. Henry J, Pylypchuk Y, Searcy T, Patel V. Adoption of Electronic Health Record Systems among U.S. Non-Federal Acute Care Hospitals: 2008-2015 [updated 2016 May; cited 2019 Jan 18]. Available from: https://dashboard.healthit.gov/evaluations/data-briefs/non-federal-acute-care-hospital-ehr-adoption-2008-2015.php.

26. Yang N, Hing E. National Electronic Health Records Survey: 2015 Specialty and Overall Physicians Electronic Health Record Adoption Summary Tables. Atlanta: Centers for Disease Control and Prevention; 2017.

27. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, *et al.* The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018; 562: 203-9.

28. Minch'olé A, Rodriguez B. Artificial intelligence for the electrocardiogram. Nat Med. 2019; 25: 22-3.

29. Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. Extracting information from textual documents in the electronic health record: a review of recent research. Yearb Med Inform. 2008; 47: 128-44.

30. Jiang M, Chen Y, Liu M, Rosenbloom ST, Mani S, Denny JC, *et al.* A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries. J Am Med Inform Assoc. 2011; 18: 601-6.

31. Ebadollahi S, Sun J, Gotz D, Hu J, Sow D, Neti C. Predicting patient's trajectory of physiological data using temporal trends in similar patients : a system for near-term prognostics. AMIA Annu Symp Proc. 2010; 2010: 192-6.

32. Zhao D, Weng C. Combining PubMed knowledge and EHR data to develop a weighted bayesian network for pancreatic cancer prediction. J Biomed Inform. 2011; 44: 859-68.

33. Austin PC, Tu JV, Ho JE, Levy D, Lee DS. Using methods from the data-mining and machine-learning literature for disease classification and prediction: A case study examining classification of heart failure subtypes. J Clin Epidemiol. 2013; 66: 398-407.

34. Kuperman GJ, Bobb A, Payne TH, Avery AJ, Gandhi TK, Burns G, *et al.* Medication-related clinical decision support in computerized provider order entry systems: a review. J Am Med Inform Assoc. 2007; 14: 29-40.

35. Knake LA, Ahuja M, McDonald EL, Ryckman KK, Weathers N, Burstain T. Quality of EHR data extractions for studies of preterm birth in a tertiary care center: guidelines for obtaining reliable data. BMC Pediatr. 2016; 16: 59. doi: 10.1186/s12887-016-0592-z.

36. Editor of Nature Medicine. Medicine in the digital age. Nat Med. 2019; 25: 1. doi: 10.1038/s41591-018-0322-1.

37. Noreot B, Glicksberg BS, Butte AJ. A call for deep-learning healthcare. Nat Med. 2019; 25: 14-5.

38. Gottesman O, Johansson F, Meier J, Dent J, Lee D, Srinivasan S, et al. Evaluating reinforcement learning algorithms in observational health settings. New York: Cornell University; 2018.

39. Domingos P. A few useful things to know about machine learning. Commun ACM. 2012; 55: 78-87.

40. Esteva A, Robicquet A, Ramsudar B, Kuleshov V, DePristo M, et al. A guide to deep learning in healthcare. Nat Med. 2019; 25: 24-9.

41. Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge: The MIT Press; 2016.

42. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. Sci Rep. 2016; 6: 26094. doi: 10.1038/srep26094.

43. Jagannatha AN, Yu H. Structured prediction models for RNN based sequence labeling in clinical text. Proc Conf Empir Methods Nat Lang Process. 2016; 2016: 856-65.

44. Jagannatha AN, Yu H. Bidirectional RNN for medical event detection in electronic health records. Proc Conf. 2016; 2016: 473-82.

45. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. New York: Cornell University; 2018.

46. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. Int J Compute Vis. 2015; 115: 211-52.

47. Hirschberg J, Manning CD. Advances in natural language processing. Science. 2015; 349: 261-6.

48. Hinton G, Deng L, Yu D, Dahl G, Mohamed A, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process Mag. 2012; 29: 82-97.

49. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017; 60: 84-90.

50. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. Lect Notes Comput Sci. 2014; 8689: 818-33.

51. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell. 2018; 172: 1122-31.

52. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014; 46: 310-5.

53. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics. 2015; 31: 761-3.

54. Dudley JT, Listgarten J, Stegle O, Brenner SE, Parts L. Personalized medicine: from genotypes, molecular phenotypes and the quantified self, towards improved medicine. Biocomputing. 2014; 2015: 342-6.

55. Leung MKK, Delong A, Alipanahi B, Frey BJ. Machine learning in genomic medicine: a review of computational problems and data sets. Proc IEEE. 2016; 104: 176-97.

56. Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, Yuen RKC, et al. The human splicing code reveals new insights into the genetic determinants of disease. Science. 2015; 347: 1254806. doi: 10.1126/science.1254806.

57. Alipanahi B, Delong A, Weirauch MT, Frey BJ. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nature Biotechnol. 2015; 33: 831-8.

58. Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA. 2013; 309: 1351-2.

59. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017; 21: 230-43.

60. Johnson KW, Torres Soto J, Glicksberg BS, Shameer K, Miotto R, Ali M, et al. Artificial intelligence in cardiology. J Am Coll Cardiol. 2018; 71: 2668-79.

61. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, et al. Canadian association of radiologists white paper on artificial intelligence in radiology. Can Assoc Radiol J. 2018; 69: 120-35.

62. De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med. 2018; 24: 1342-50.

63. Slomka PJ, Dey D, Sitek A, Motwani M, Berman DS, Germano G. Cardiac imaging: working towards fully-automated machine analysis & interpretation. Expert Rev Med Devices. 2017; 14: 197-212.

64. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep Learning For Identifying Metastatic Breast Cancer. New York: Cornell University; 2016.

65. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology. 2017; 284: 574-82.

66. Sitapati A, Kim H, Berkovich B, Marmor R, Singh S, El-Kareh R, et al. Integrated precision medicine: the role of electronic health records in delivering personalized treatment. Wiley Interdiscip Rev Syst Biol Med. 2017; 9: e1378. doi: 10.1002/wsbm.1378.

67. Petrone J. FDA approves stroke-detecting AI software. Nat Biotechnol. 2018; 36: 290. doi: 10.1038/nbt0418-290.

68. The Atlantic [Internet]. Hsu J, Spectrum. AI could make detecting autism easier. In The Atlantic [updated 2018 Jul 14; cited 2019 Jan 18]. Available from: https://www.theatlantic.com/technology/archive/2018/07/ai-autism-diagnosis-screening-bottleneck/564890/.

69. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. Nat Biomed Eng. 2018; 2: 749-60.

70. Fast Company [Internet]. Peters A. Having a heart attack? This AI helps emergency dispatchers find out. In Fast Company [updated 2018 Nov 1; cited 2019 Jan 18]. Available from: https://www.fastcompany.com/40515740/having-a-heart-attack-this-ai-helps-emergency-dispatchers-find-out.

71. Patel NM, Michelini VV, Snell JM, Balu S, Hoyle AP, Parker JS, et al. Enhancing next-generation sequencing-guided cancer care through cognitive computing. Oncologist. 2018; 23: 179-85.

72. Dailymail Online [Internet]. De Graaf M. Will Al replace fertility doctors? Why computers are the only ones that can end the agony of failed IVF cycles, miscarriages, and risky multiple birth [updated 2018 Oct 10; cited 2019 Jan 18]. Available from: https://www.dailymail.co.uk/health/article-6257891/Study-finds-artificial-intelligence-better-doctor-crucial- stage-IVF.html.

73. Gurovich Y, Hanani Y, Bar O, Fleischer N, Gelbman D, Basel-Salmon L, et al. DeepGestalt—identifying rare genetic syndromes using deep learning. New York: Cornell University; 2017. https://arxiv.org/abs/1801.07637.

74. Bahl M, Barzilay R, Yedidia AB, Locascio NJ, Yu L, Lehman CD. High-risk breast lesions: a machine learning model to predict pathologic upgrade and reduce unnecessary surgical excision. Radiology. 2018; 286: 810-8.

75. Beam AL, Kohane IS. Translating artificial intelligence into clinical care. JAMA. 2016; 316: 2368-9.

76. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and

classification in ambulatory electrocardiograms using a deep neural network. Nat Med. 2019; 25: 659.

77. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, *et al.* Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat Med. 2019; 25: 70-4.

78. Health News Review [Internet]. Victory J. What did journalists overlook about the Apple Watch 'heart monitor' feature? [updated 2018 Sep; cited 2019 Jan 25]. Available from: https://www.healthnewsreview.org/2018/09/what-did-journalists-overlook-about-the-apple-watch-heart-monitor-feature/.

79. Apple Insider [Internet]. Fingas R. Apple Watch Series 4 EKG tech got FDA clearance less than 24hours before reveal [updated 2018 Sep 18; cited 2019 Jan 25]. Available from: https://appleinsider.com/articles/18/09/18/apple-watch-series-4-ekg-tech-got-fda-clearance-less-than-24-hours-before-reveal.

80. diatribe Learn [Internet]. Levine B, Brown A. Onduo delivers diabetes clinic and coaching to your smartphone. [updated 2018 March 12; cited 2019 Jan 25]. Available from: https://diatribe.org/onduo-delivers-diabetes-clinic-and-coaching-your-smartphone.

81. Han Q, Ji M, de Rituerto de Troya IM, Gaur M, Zejnilovic L. A hybrid recommender system for patient–doctor matchmaking in primary care. New York: Cornell University; 2018.

82. Zmora N, Zeevi D, Korem T, Segal E, Elinav E. Taking it personally: personalized utilization of the human microbiome in health and disease. Cell Host Microbe. 2016; 19: 12-20.

83. Korem T, Zeevi D, Zmora N, Weissbrod O, Bar N, Lotan-Pompan M, *et al.* Bread affects clinical parameters and induces gut microbiome–associated personal glycemic responses. Cell Metab. 2017; 25: 1243-53.e5.

84. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, *et al.* Personalized nutrition by prediction of glycemic responses. Cell. 2015; 163: 1079-94.

85. Hall H, Perelman D, Breschi A, Limcaoco P, Kellogg R, McLaughlin T, *et al.* Glucotypes reveal new patterns of glucose dysregulation. PLoS Biol. 2018; 16: e2005143. doi: 10.1371/journal.pbio.2005143.

86. Albers DJ, Levine M, Gluckman B, Ginsberg H, Hripcsak G, Mamykina L. Personalized glucose forecasting for type 2 diabetes using data assimilation. PLoS Comput Biol. 2017; 13: e1005232. doi: 10.1371/journal.pcbi.1005232.

87. Thaiss CA, Levy M, Grosheva I, Zheng D, Soffer E, Blacher E, *et al.* Hyperglycemia drives intestinal barrier dysfunction and risk for enteric infection. Science. 2018; 359: 1376-83.

88. Wu D, Hu D, Chen H, Shi G, Fetahu IS, Wu F, *et al.* Glucose-regulated phosphorylation of TET2 by AMPK reveals a pathway linking diabetes to cancer. Nature. 2018; 559: 637-41.

89. Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, *et al.* A universal SNP and small-indel variant caller using deep neural networks. Nat Biotechnol. 2018; 36: 983-7.

90. Sundaram L, Gao H, Padigepati SR, McRae JF, Li Y, Kosmicki JA, *et al.* Predicting the clinical impact of human mutation with deep neural networks. Nat Genet. 2018; 50: 1161-70.

91. Zhou J, Theesfeld CL, Yao K, Chen KM, Wong AK, Troyanskaya OG. Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. Nat Genet. 2018; 50: 1171-9.

92. Behravan H, Hartikainen JM, Tengström M, Pylkäs K, Winqvist R, Kosma V, *et al.* Machine learning identifies interacting genetic variants contributing to breast cancer risk: a case study in Finnish cases and controls. Sci Rep. 2018; 8: 13149. doi: 10.1038/s41598-018-31573-5.

93. Lin C, Jain S, Kim H, Bar-Joseph Z. Using neural networks for reducing the dimensions of single-cell RNA-seq data. Nucleic Acids Res. 2017; 45: e156. doi: 10.1093/nar/gkx681.

94. Angermueller C, Lee HJ, Reik W, Stegle O. DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning. Genome Biol. 2017; 18: 67. doi: 10.1186/s13059-017-1189-z.

95. AlQuraishi M. End-to-end differentiable learning of protein structure. Cell Syst. 2019; 8: 292-301.e3. doi: 10.1016/j.cels.2019.03.006.

96. Espinoza JL. Machine learning for tackling microbiota data and infection complications in immunocompromised patients with cancer. J Intern Med. 2018; 284: 189-92.

97. Van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, *et al.* Recovering gene interactions from single-cell data using data diffusion. Cell. 2018; 174: 716-29.e727.

98. Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM, *et al.* Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. New York: Cornell University; 2018.

99. Listgarten J, Weinstein M, Kleinstiver BP, Sousa AA, Joung JK, Crawford J, *et al.* Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. Nat Biomed Eng. 2018; 2: 38-47.

100. Smalley E. AI-powered drug discovery captures pharma interest. Nat Biotechnol. 2017; 35: 604-5.

101. Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018; 17: 97-113.

102. Chakradhar S. Predictable response: finding optimal drugs and doses using artificial intelligence. Nat Med. 2017; 23: 1244-7.

103. Lowe D. AI designs organic syntheses. Nature. 2018; 555: 592-3.

104. Luechtefeld T, Marsh D, Rowlands C, Hartung T. Machine learning of toxicological big data enables read-across structure activity relationships (RASAR) outperforming animal test reproducibility. Toxicol Sci. 2018; 165: 198-212.

105. Hie B, Cho H, Berger B. Realizing private and practical pharmacological collaboration. Science. 2018; 362: 347-50.

106. STAT+ [Internet]. Ross C, Swetlitz I. IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show [updated 2018 Jul 25; cited 2019 Jan 25]. Available from: https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe- incorrect-treatments/.

107. Healthcare IT News [Internet]. Miliard M. As FDA signals wider AI approval, hospitals have a role to play [updated 2018 May 31; cited 2019 Jan 25]. Available from: https://www.healthcareitnews.com/news/fda-signals-wider-ai-approval-hospitals-have-role-play.