

# 의료분야에서 인공지능 현황 및 의학교육의 방향

정진섭

부산대학교 의과대학 생리학교실

## Current Status and Future Direction of Artificial Intelligence in Healthcare and Medical Education

Jin Sup Jung

Department of Physiology, Pusan National University College of Medicine, Yangsan, Korea

The rapid development of artificial intelligence (AI), including deep learning, has led to the development of technologies that may assist in the diagnosis and treatment of diseases, prediction of disease risk and prognosis, health index monitoring, drug development, and healthcare management and administration. However, in order for AI technology to improve the quality of medical care, technical problems and the efficacy of algorithms should be evaluated in real clinical environments rather than the environment in which algorithms are developed. Further consideration should be given to whether these models can improve the quality of medical care and clinical outcomes of patients. In addition, the development of regulatory systems to secure the safety of AI medical technology, the ethical and legal issues related to the proliferation of AI technology, and the impacts on the relationship with patients also need to be addressed. Systematic training of healthcare personnel is needed to enable adaption to the rapid changes in the healthcare environment. An overall review and revision of undergraduate medical curriculum is required to enable extraction of significant information from rapidly expanding medical information, data science literacy, empathy/compassion for patients, and communication among various healthcare providers. Specialized postgraduate AI education programs for each medical specialty are needed to develop proper utilization of AI models in clinical practice.

**Keywords:** Artificial intelligence, Delivery of health care, Medical education

### Corresponding author

Jin Sup Jung  
Department of Physiology, Pusan  
National University College of Medicine,  
49 Busandaehak-ro, Mulgeum-eup,  
Yangsan 50612, Korea  
Tel: +82-51-510-7081  
Fax: +82-51-510-8076  
E-mail: jsjung@pusan.ac.kr  
https://orcid.org/0000-0002-7643-0918

Received: January 31, 2020  
1st revised: March 23, 2020  
2nd revised: March 31, 2020  
3rd revised: April 7, 2020  
Accepted: April 9, 2020

## 서론

인공지능(artificial intelligence)은 옥스퍼드 영어사전에서 “시각 인식, 음성 인식, 의사결정, 언어번역 등 인간의 지능을 필요로 하는 업무를 정상적으로 수행할 수 있는 컴퓨터시스템의 이론과 개발”로 정의하고 있으며[1], 확률적 기계학습(probabilistic machine learning), 진화적 컴퓨팅(evolutionary computing), 전문가시스템(expert system), 퍼지시스템(fuzzy system) 등 다양한 방법을 포함한다. 역사적으로 인공지능 연구는 1943년 McCulloch와 Pitts [2]가 인간 뇌의 신경망을 모식화 한 전기회로모델을 제안하였고, 1947년 Alan Turing이 경험을 통해 배우는 기계를 제안하며 시발되었다[2]. 인공지능이라는 용어는 1955년 제출한 McCarthy 등[3]의 제안을 1956년 Dartmouth 회의에서 받아들임으로써 사용하게 되었다. 1958년 Rosenblatt [4]는 인공지능의 기본 개념인 퍼셉트론(perceptrons)이라는 알고리즘을 개발하였고, 이것을 패턴 인식 분야에 적용하였

다. 이후 인공지능망을 이용한 인공지능기술은 기술적 한계와 컴퓨터의 연산능력 제한 등으로 1980년대에 개발된 전문가시스템에 밀려 크게 주목을 받지 못했다. 그러나 토론토 의과대학의 Hinton 교수와 Lecun 교수 연구팀에 의해 오류 역전파(back propagation) [5], 다층신경망 구조[6], 콘볼루션 신경망(convolutional neural network) [7] 등 심층학습의 기본원리가 개발되고 그래픽스 처리장치(graphic processing unit)를 포함한 하드웨어의 빠른 발전과 빅데이터 확보 등 기반 여건이 성숙됨으로써 도약의 계기를 마련하였다. 2012년 심층신경망을 이용한 심층학습기술이 이미지 인식[8,9]과 음성인식[10]에서 인공지능 분야의 다른 기술보다 현저히 우수한 성능을 보이는 것으로 보고됨에 따라 관련 연구가 폭발적으로 확산되어 다양한 분야에 적용되었다. 한국에서 인공지능에 대한 큰 대중적 관심을 유도한 사건은 2016년 구글 딥마인드가 개발한 알파고와 프로 바둑의 세계 챔피언인 이세돌을 이기고 바둑에서 일인자가 된 것이었다[11]. 같은 연도에 구글의 심층학습모델이 안저 사진을

이용한 당뇨병성 망막증 진단에서 전문가 수준의 능력을 보임을 보고한 이후[12], 의료 관련 분야에서도 심층학습을 포함한 인공지능기술이 광범위하게 연구되고 있다.

미국 성인 외래환자의 5%가 오진을 경험한 것으로 보고되었고 [13], 병원에서 일어나는 의료사고의 6%-17%의 원인이 오진에 의한 것으로 알려져 있다[14]. 또한 의사의 번아웃이 의사의 주관적 판단을 저해하여 의료의 질을 떨어뜨린다는 보고가 있다[15]. 따라서 인공지능기술을 의료에 적용할 필요성이 제기되는 이유는 고령화 사회에서 의료인력의 부족과 폭증하는 의료비의 절감뿐만 아니라 의사의 주관적 판단에 의한 오류를 보완함으로써 환자 안전 확보를 통한 의료 질의 고도화이다.

인공지능의 확산 초기에 의사를 대체할 수 있다는 불안감이 존재하였으나 현재의 기술수준으로 미루어 볼 때 향후 상당한 기간 동안은 의사의 보조 역할에 머무를 것이다. 하지만 인공지능으로 인해 의사의 일의 성격은 바뀔 것이며[16,17], 인공지능을 활용하지 않는 의사는 인공지능을 활용하는 의사에 의해 대체될 것이라는 주장이 제기되고 있다[18,19]. 이러한 의료환경의 변화에 가장 큰 영향을 받을 미래 세대의 의료인인 의과대학생들은 의학지식 및 술기의 습득, 의사국가시험 스트레스 등으로 이 분야의 발전에 대해 정확한 정보를 얻지 못하고 어떻게 대비해야 할지 막연한 두려움 속에 있다. 또한 이러한 막연한 두려움이 졸업 후 전공의 선택에도 영향을 줄 수 있다. 2018년에 캐나다에서 의과대학생을 대상으로 한 조사에 따르면 인공지능이 영상의학과 의사를 대체할 것이라는 의견이 30% 정도였고, 대체는 하지 못하더라도 영상의학과 의사의 수요를 줄일 것이라는 답변이 68%였으며 응답자의 1/6이 인공지능에 대한 불안감 때문에 영상의학 전공 선택을 재고하였다고 답하였다[20]. 반면에 영상의학을 많이 경험하고 인공지능에 대해 이해도가 높은 학생들은 인공지능에 대한 불안감의 수준이 감소되었다[20]. 이러한 예는 인공지능을 포함한 의료기술의 발전 현황과 문제점에 대한 체계적인 교육의 필요성을 뒷받침한다.

본 종설에서는 인공지능기술의 의료분야 연구개발 현황 및 문제점을 개괄하고 이에 따른 의료의 변화에 대응하는 의학교육의 방향성을 검토해 보고자 한다.

## 인공지능의 개요 및 의료 적용분야

### 1. 인공지능 개요

Figure 1은 인공지능, 기계학습 및 심층학습의 관계를 도식화한 것이다. 기계학습은 인공지능을 구현하는 한 방법으로 데이터를 이용한 훈련에 의한 학습을 통하여 스스로 목표에 도달하는 통계학적이고 수학적 방법을 의미한다. 기계학습의 학습방법으로는 지도학습, 비지도학습, 강화학습이 있다. 심층학습을 제외한 전통적인 기계학습은 데이터의 특성을 추출하고 이를 특징하는 알고리즘을 제작하여 미지의 데이터에 대한 해법을 제공하기 위해 수동적인 코딩(coding)이 필요하다.

심층학습은 기계학습의 한 형태로 인간의 뇌가 학습하는 방식과 유사한 방법으로 다층 인공신경망을 다량의 데이터로 학습시켜 원하는 결과를 도출하는 학습방법이다. 심층학습은 기계학습의 다른 카테고리보다 달리 많은 데이터와 계산능력이 필요하고 답이 어떻게 도출되었는지 논리적으로 이해하지 못하지만 데이터의 특성(feature)을 스스로 추출하고 정확도가 높으며 다양하고 복잡한 업무에 적용이 가능한 장점이 있다.

### 2. 인공지능의 의료 적용분야

#### 1) 이미지 분석

인공지능을 이용한 이미지 분석은 영상의학(X-레이, 초음파, 컴퓨터 단층 촬영(computed tomography, CT), 자기공명영상(magnetic resonance imaging) [21,22], 병리학(원발성 및 전이성 암 조직 검사) [23,24], 소화기학(내시경영상) [25-27], 피부과학[28,29], 순

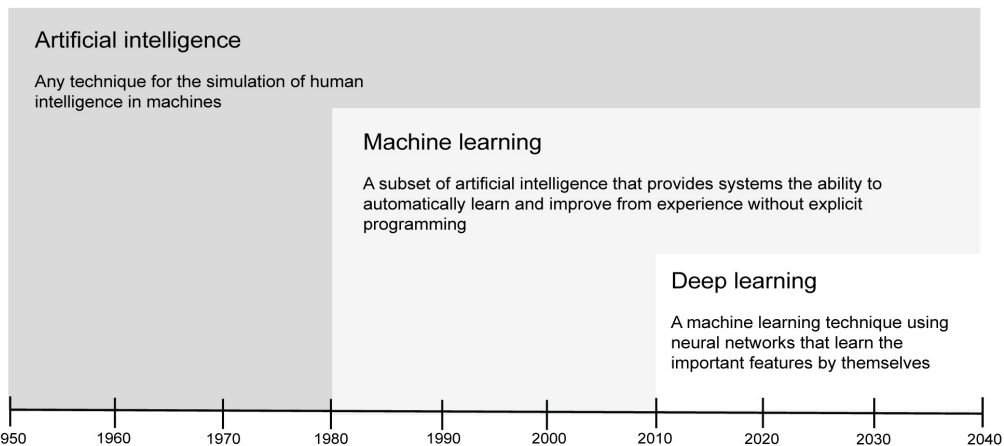


Figure 1. The relationship of among artificial intelligence, machine learning and deep learning and their development timeline.

환기학(심장영상) [30], 안과학(optical coherence tomography, OCT 및 안저 사진) [12,31,32] 등 다양한 의료분야에 적용되고 있다. 2019년 Liu 등[33]은 최근 2012년 1월부터 2019년 6월까지 게재된 심층학습을 이용한 이미지 분석 논문을 메타분석한 결과 현재 심층학습의 이미지 분석능력은 평균적으로 전문가와 동일한 수준에 도달해 있음을 확인하였다. 최근 일부 경우에서 의료전문가보다 우수한 감별능력을 가진 심층학습모델이 보고되고 있다. 구글의 저선량 폐 CT 영상 분석모델의 이전 비교 영상이 없는 경우의 폐암 진단[34]과 구글 딥마인드모델의 유방암 진단[35]은 영상의학 전문의보다 낮은 위양성률과 위음성률을 보였고, 일차진료에서 접하는 피부질환의 80%를 차지하는 26가지 피부 병변에 대하여 심층학습모델이 피부과 전문의와 유사하며 일차진료 의사보다 우수한 진단 정확도를 보임이 보고되었다[36]. 인공지능 학습모델의 개발 초기에는 흑색종, 폐암 등 특정 병적 상태 한 가지만을 분석하는 모델이 개발되어 실제 임상상황에서 적용하기에 제한점이 있었으나 최근에는 OCT를 이용한 안질환[32], 피부질환[36], 흉부 X-레이 영상[37,38] 진단에서 다양한 병적 변화를 동시에 분석하고 구별하는 모델들이 개발되어 실제 임상상황에서의 이미지 판독을 모사하는 방향으로 발전하고 있다.

심층학습은 특정 질병과 연계된 이미지 분석 이외에도 안면형태 분석을 통한 드문 소아유전질환 진단[39], 알츠하이머병에 이르는 인지능력의 변화 평가[40], 심장의 관상동맥 혈류량 평가[41] 등 기능적인 영역으로도 활용되고 있다. 또한 인간이 볼 수 없는 이미지의 미묘한 차이를 감지하여 안저 이미지에서 나이, 성별, 흡연 여부 등 심혈관질환의 위험도[42], 알츠하이머병[43], 빈혈[44] 등을 분석할 수 있고, 조직표본으로부터 폐암의 돌연변이[45], 소화관 암의 반복서열 불안정성(microsatellite instability) [46], 흉부 CT 영상에서 폐암의 예후[47], CT 영상으로부터 난소암의 분자적 표현형과 예후를 예측[48]함으로써 환자 진료에 새로운 정보를 제공할 가능성을 제시하고 있다.

2016년 이후 급속도로 발전한 심층학습에 의한 이미지 분석은 연구개발 단계를 넘어 많은 분야에서 인허가를 획득하여 임상에 활용되는 단계에 도달했다. 미국의 경우 2014년 AliveCor가 심전도에서 심방세동을 자동분석하는 알고리즘으로 최초로 Food and Drug Administration (FDA) 허가를 받은 후 인공지능 관련 제품 허가 건수가 2016년 4개, 2017년 8개, 2018년 28개, 2019년 22개로 증가하는 추세에 있다[49]. 2019년 7월 이후 승인제품을 포함한 FDA 허가 인공지능모델들의 적용분야는 영상분석 이외에도 심전도 분석, 주의력결핍 과다행동장애, 유아 자폐증, 수면장애, 간질발작, 치매 등 정신신경학 분야, 혈당, 전체 혈구계산, 간의 철 농도 등 검사영역, 노인의 낙상, 안구운동 분석, 관상동맥혈류 계산, 입원 환자의 생리지표 모니터링, 인슐린 용량 추천, CT 영상으로부터 관상동맥 석회화 계산 등 다양한 분야를 포함하고 있다[49] (Table 1). 한국에서도 2018년 뷰노의 골 연령 판독시스템을 시작으로 2019년 현재 13개의 제품이 식품의약품안전처의 허가를 받았다 (Table 2).

인허가 제품들은 대부분 의사의 최종 판단을 필요로 하는 보조의 료기기이지만 당뇨병성 망막증을 진단하는 IDx-DR (IDx Technologies Inc., Coralville, IA, USA)은 특정촬영장비(Topcon NW400; Topcon Medical Systems Inc., Oakland, NJ, USA)로 촬영한 망막 이미지를 분석하여 환자의 상태가 전문가의 진료가 필요한지, 단순 관찰 수준인지를 의사의 개입 없이 독자적으로 판단 하도록 허용되고 있다[50]. 또한 2019년부터 FDA 승인을 받은 심층 학습모델이 영상기기에 장착되어 사용자에게 제공되고 있다. Phillips의 뇌출혈, 기흉, 척추압박골절, 관상동맥 칼슘 양, 유방암 등의 진단 영상장비가 디지털병리시스템과(전립선암 진단)에 장착 되어 제공되고 있고[51,52], GE사의 포터블 X-레이에 기흉 의심 영상을 감별하는 인공지능 제품이 포함되어 제공되고 있다[53].

**Table 1.** List of the Food and Drug Administration approved artificial intelligence models since September 2019

Approval date	Company	Description
2019. 9	DreaMed Diabetes	Insulin delivery recommendations
	GE Healthcare	Triage of pneumothorax patients in portable chest X-ray
	HeartFlow	Modeling of coronary artery disease
2019. 10	Biofourmis	Monitoring of vital signs in ambulatory patients
	HeartVista	Cardiac MRI analysis
	Omega Medical Imaging	Diminution of radiation exposure during cardiac fluorography
	Subtle Medical	Improvement of brain MRI image
2019. 11	HealthLytix	Detection of prostate cancer in MRI
	Hologic	Diminution of image volume for breast tomography
	Ultromics	Analysis of echocardiography
	Zebra Medical Vision	Pleural effusion detection in chest X-ray
2019. 12	Sight Diagnostics	Complete blood count analysis

MRI, magnetic resonance imaging.

**Table 2.** List of Ministry of Food and Drug Safety approved artificial intelligence models

Year	Company and product name	Description
2018	Infomeditec NeuroAI	Dementia prediction from brain MRI
	JLK UNISTRO	Brain infarction pattern analysis in CT and MRI
	Lunit Insight CXR1	Lung nodule detection in chest X-ray
	Vuno BoneAge	Bone age diagnosis
2019	DeepNoid Deepspine	Lumbar compression fracture in X-ray
	JLK Inspection ASTOSCAN	Brain aging prediction in brain three-dimensional MRI
	JLK Inspection Colonoscopy	Colonoscopic image analysis
	JLK Inspection Gastroscopy	Gastroscopic image analysis
	JLK Inspection Lung CT	Nodule detection in lung CT
	Lunit Insight MMG	Nodule detection in breast mammography
	Samsung Electronics ALND	Lung nodule in chest X-ray
	Vuno Chest X-ray	Image analysis of chest X-ray
Vuno Deepbrain	Dementia diagnosis	

MRI, magnetic resonance imaging; CT, computed tomography.

2) 치료 분야

환자 특성별 치료 개별화의 필요성 때문에 진단 분야에 비해 치료 분야에 기계학습의 효과적 적용이 쉽지 않아 많은 모델이 개발되지 못하고 있다[54]. 최근 임상적 효능에 대해 회의적인 의견이 제기되고 있으나 왓슨 온콜로지사가 환자의 특성에 따른 암 치료를 추천하는데 활용되고 있고[55,56], 싱가포르에서 전이성 전립선암의 환자 특성별 맞춤 치료[57]와 두경부암 방사선 치료계획 수립[58]에 기계학습이 활용되었다.

외과 수술영역에서 현재 통합하는 로봇이 개발되고 있으나[59], 직접적으로 수술에 관여하는 것보다 환자의 정보를 수술 전, 수술 동안 실시간으로 종합 분석하여 환자의 상태에 대한 정보를 제공하고[60], 수술과정의 움직임을 분석하여 술기를 분석하는 기능에 활용되고 있다[61]. 또한 만성질환의 치료약물에 대한 환자의 순응도 (adherence)를 증가시키는 데도 활용되고 있다[62,63].

3) 질병발생 위험도, 질병진단 및 예후 예측

병원에서 생성한 전자의무기록(electronic medical records, EMRs)의 원본을 분석하여 질병의 진단, 예후 및 특정 상황 발생 가능성을 예측한 다양한 예가 보고되고 있다. 2018년 구글은 동일한 심층학습모델에 캘리포니아대학교-샌프란시스코 캠퍼스와 시카고 대학병원의 EMRs을 각각 학습시켜 각 병원의 입원환자의 사망 확률, 30일 재입원 및 장기입원 확률과 최종 퇴원 시 병명 진단에 기존 모델보다 더 높은 정확한 예측결과를 보임을 보고하였다[64]. 567,498명의 중국 광저우 여성아동병원 소아 환자 EMRs을 분석하여 개발한 소아질환 진단 인공지능모델이 뇌염, 신경정신 계통 질환 외에는 대부분의 질환 군에서 전공의 3년차 및 8년 임상경험을 가진 소아과 의사보다는 정확하고 15년 이상 임상경험을 가진 소아과 의사보다는 못한 결과를 보임을 보고하였다[65].

구글 답마인드는 미국 보훈병원에 입원한 환자의 EMRs을 이용

하여 급성 신장 손상을 48시간 전에 조기 예측하는 인공지능기술을 보고하였다[66]. 이 기술은 신장투석을 요하는 심한 급성신부전의 경우 48시간 전 발생 예측률이 90%였으며, 모든 종류의 급성 신장 손상으로 확대한 경우는 예측률이 55.8%였고 세 번에 한 번의 위양성 경고를 발생시켰다[66]. 그 외에도 외과 수술을 시행한 환자의 퇴원일[67], 중환자실에서 패혈증 발생[68], 영국 바이오뱅크 자료로부터 심혈관질환 병력이 없는 사람의 심혈관질환 발생 위험도 [69], 입원환자에서 심정지 발생 확률[70], 중환자실 환자 예후[71] 등을 예측하는 모델이 있다.

4) 건강관리

다양한 웨어러블 장비나 스마트폰 등을 이용하여 획득된 다량의 연속데이터를 분석하여 환자 건강상태의 원격관리뿐 아니라 질병의 예방 및 발생 가능성의 예측에 활용될 수 있다. 그 예로, 심전도 분석을 통한 심방세동 조기진단[72,73]이나 소셜 네트워크 서비스 (social networking service) 및 스마트폰 분석을 통한 정신건강평가 변화 분석[74] 등이 있다. Zeevi 등[75]은 당뇨병이 없는 피검자의 장 세균 분석 혈당측정, 음식섭취 종류 및 양, 잠자는 시간, 스마트폰을 이용한 실시간 운동량 조사, 설문조사(식습관, 생활습관, 의학적 배경), 체중, 키, 나이, 성별, 골반 둘레, 혈액검사 결과를 모은 데이터를 이용하여 음식섭취 후 혈당 증가양상을 예측하는 기계학습모델을 제시하였다. 모델 개발 시 사용한 집단과 다른 집단에 이 모델을 적용하여 식후 혈당 변화를 예측한 결과 유사한 정확도를 보였으며, 동일한 음식에 대해서 서로 다른 식후 혈당 변화 양상을 보이는 특성을 반영한 식이를 제공한 결과 혈당반응을 완화하고 대장의 미생물 분포가 변하는 것을 관찰하였다[75].

5) 의료행정 및 관리시스템의 개선

반복되는 업무의 자동화, 챗봇(chatbot)을 이용한 다양한 안내시



시스템, 외래 및 판독 등의 의료기록 자동화 및 표준화[76], 의료기기 성능 최적화(영상기기의 영상의 질, 방사선 치료 프로토콜), 환자 분류 및 병원 자원관리, 컴퓨터 비전을 활용한 환자 안전관리(낙상, 중환자실 모니터링 등)와 수술과정의 모니터링 및 교육[77] 등 다양한 분야에 활용될 수 있다.

6) 약물 개발

약물의 고안 단계부터 독성 예측, 안전성 시험의 환자군 선택, 시판 후 부작용의 검색 등 다양한 분야에서 약물 개발에 활용될 수 있다[78-80]. 실제로 일반적으로 5년 이상 걸리는 신약 탐색기간을 인공지능을 이용하여 1년으로 단축하여 세로토닌수용체(5-HT1A)에 대한 작용제(agonist)를 개발하여 강박장애 치료를 위한 임상 1상 시험을 일본에서 시작한 예가 있고[81], 항생제에 내성이 있는 세균을 죽일 수 있는 항균 화학물질을 발굴하였다[82].

인공지능기술의 임상 적용 시 고려해야 할 문제점

1. 심층학습의 원리 및 기술적 문제점

1) 심층학습의 원리

의료분야에 가장 많이 활용되고 있는 기계학습모델인 심층학습의 기본원리를 간략하게 설명하면 다음과 같다. 인공신경망은 신경세포의 역할을 하는 노드(node)로 네트워크를 형성하고, 노드 간의 결합 세기를 학습을 통하여 변화시켜 문제를 해결하는 모델이다. 인공신경망은 입력층(input layer), 히든층(hidden layer), 출력층(output layer)으로 구성되어 있다(Figure 2B). 입력층에는 각각의

입력변수와 1:1로 매칭되는 노드가 존재하며, 히든층의 개수가 2개 이상이 되는 경우 심층 신경망(deep neural network)이라고 한다. 히든층과 출력층에 존재하는 노드는 이전 층에서 오는 다양한 입력값(input, xi)과 각 노드 간의 전달과정(synapse)에서 입력값을 변형시키는 가중치(weight, wi)를 곱한 합(summation)을 계산하는 기능이 있으며 입력받은 노드의 가중합을 어떤 방식으로 출력하는지를 결정하는 활성화 함수(activation function) 기능을 포함한다(Figure 2A, C). 활성화 함수를 통한 출력을 조절하기 위하여 임의의 상수인 바이어스(bias)를 도입하여 보다 나은 결과를 얻도록 조절할 수 있다. 바이어스 값은 일반적으로 1.0으로 시작하지만 상황에 따라 음수값에서부터 다양하게 변화가 가능하다(Figure 2A).

인공신경망의 심층학습은 노드와 층의 수, 각 노드를 연결하는 가중치 및 바이어스의 조정(Figure 2A), 노드에 입력된 값을 어떻게 가공하는지를 결정하는 활성화 함수(activation function)의 종류(Figure 2C), 층의 배열방법과 노드 간의 연결방식(neural network architecture) 등의 조합을 통해 대량의 데이터로부터 얻은 산물의 에러를 반복되는 오류 역전파(back propagation)를 통해 최소화하여 최적화된 결과를 도출함으로써 이루어진다(Figure 2B). 인공신경망의 구조는 노드 간의 연결방식, 데이터의 전처리, 층의 구성 등에 따라 다양한 종류가 있으며[83], 대표적인 것이 물체 인식, 동영상 분석 등에 주로 활용되는 컨볼루션 신경망(convolutional neural network), 자연어 처리, 음성 인식 등에 주로 활용되는 재귀 신경망(recurrent neural network), 분석에 그치지 않고 이미지나 음성을 만드는 기능을 가진 생성적 적대 신경망(generative adversarial network) 등이다[84].

데이터 사이언스에 대한 전문지식이 없어도 이미지 분석 등 기존

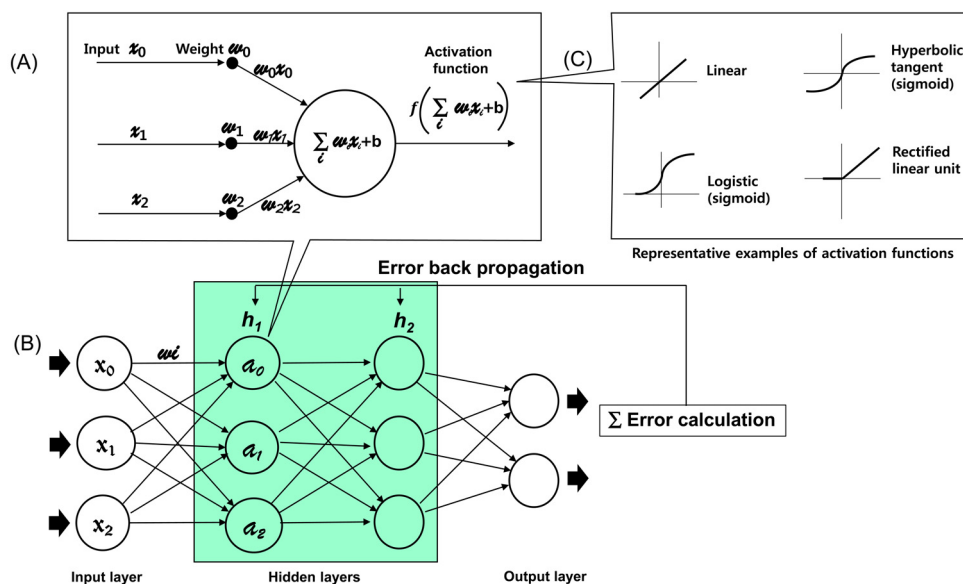


Figure 2. Schematic diagram of deep neural network. (A) The building block of deep neural networks. (B) Example of a feedforward multilayer neural network. (C) Representative examples of activation functions.

연구를 통해 심층학습모델이 정형화된 경우는 이를 활용하여 유사 업무에 적용할 수 있게 도와주는 다양한 글로벌 IT (information technology) 기업이 제공하는 자동기계학습(autoML)과 응용프로그램 인터페이스(application programming interface) 등이 제공되고 있다[85]. Faes 등[86]은 구글의 autoML을 이용하여 다양한 의학영상 이미지를 유의하게 감별할 수 있음을 보고하였다.

## 2) 심층학습에 이용되는 데이터와 알고리즘의 기술적 문제점

심층학습의 결과는 입력하는 데이터의 특성에 따라 달라진다. 심층학습모델의 학습에 이용되는 기 확보된 데이터는 데이터를 수집한 집단에 대한 대표성, 정확성, 질적 균질성 등에서 여러 가지 문제점을 가지고 있으며 대표적인 의료데이터인 EMR은 상당한 오류를 포함하고 있다[87]. 이를 극복하기 위하여 많은 수의 데이터를 사용하여 모델을 개발하나 데이터의 수를 늘리는 것만으로도 해결되지 않는 문제들이 제기되고 있다. 그 예로는 아마존의 면접프로그램이 여성을 차별한다든지[88], 구글의 이미지 모델이 흑인을 고릴라로 분류한다든지[89], 범죄 가능성을 인종에 따라 다르게 판단하는[90] 것 등이다. 의료분야에서도 흑인에서 흑색종을 잘 진단하지 못하고 [91] 미국에서 광범위하게 사용되는 건강 위험도 평가 알고리즘이 인종적 바이어스를 가진 것으로 보고되는 등[92] 입력 데이터의 대표성 오류에 의한 다양한 문제들이 노출되고 있다. 이는 특정 병원, 특정 기기로 획득한 데이터로 훈련시킨 심층학습모델을 다양한 집단, 다양한 기기, 다양한 질의 데이터를 가진 실제 의료환경에서 사용할 경우 정확도가 현저히 감소할 수 있음을 시사한다. 구글 딥마인드의 급성신부전 예측모델은 데이터를 획득한 보훈병원의 특성상 학습 데이터에 여자 환자 비율이 6.38%로 낮아 여자 환자에 대한 예측률이 유의하게 떨어진대[66].

다량의 데이터로부터 데이터의 특성을 스스로 추출하는 심층 인공지능망의 특성상 알고리즘이 어떤 방식으로 결과를 도출하는지를 파악하기 어렵다. 사용자의 입장에서 심층학습이 어떤 기준과 근거를 가지고 판단을 내리는지 설명이 제공되어야 안심하고 활용할 수 있을 것이고 심층학습모델의 오류에 대해서도 대응이 가능하게 될 것이다. 따라서 최근에는 설명 가능한 인공지능기술 개발에 많은 연구가 이루어지고 있다[93,94]. 심층학습으로 병의 진단이나 예후를 예측하는 경우 결과 판단에 중요하게 작용한 요소를 시각화하여 보여주거나[58,66], 이미지 분석의 경우 이상 부위를 이미지에 표시하는 방법으로 설명을 제공하고 있으나[34,35,37,38], 심층학습의 특성상 알고리즘의 논리구조를 완벽하게 이해하는 것이 쉽지 않다.

심층학습에서 투명성과 성능은 서로 상보적인 관계를 가지고 있어 투명성을 지나치게 강조하면 심층학습모델의 성능이 감소할 것이라는 우려가 있다. 또한 의료에 적용되고 있는 치료방법이나 약물 등도 명확한 작용 기전을 알지 못하고 경험적으로 사용하는 경우가

있으며(예, 전기충격요법이나 중추신경계 작용 약물), 아스피린이나 페니실린도 작용 기전을 알기 전에 광범위하게 사용되어왔다. 따라서 심층학습모델의 효능이 과학적이고 객관적으로 검증이 된다면 심층학습모델이 활용되는 분야에 따라 모델의 투명성에 대한 기준을 다르게 적용해야 한다는 제안을 검토해 볼 필요가 있다[95-97]. 예를 들어, 이미지의 분석, 검사결과의 해석, 자연어 처리 같은 분야는 낮은 투명성으로도 적용이 가능하도록 허용하고, 바이어스가 높을 수 있는 질병의 진단, 예후 분석, 위험도 분석 등에는 심층학습모델의 보다 높은 투명성 기준을 적용하는 것이다.

심층학습으로 데이터를 학습하는 과정에 최종 목표와 직접적인 연관성이 없는 부수적 요소를 더 중요하게 판단하여 결과를 예측하는 경우도 있다. 예를 들어, 자발성 기흉의 진단에 있어 흉관이 없는 환자에서 현저하게 낮은 진단 정확도를 보이거나[37] 포털 X-레이 기계와 폐렴을 연관시키기도 하며[98] 외과 수술용 마킹으로 흑색종의 위양성 판정을 하기도 한다[99]. 사진 속에 자(ruler) 유무, 영상 촬영이 이루어진 시기(심야, 응급) 같은 부수적 요인으로 질병을 진단하고 예후를 판단하는 경우도 보고되고 있다[100]. 또한 새로운 치료 및 진단방법의 급속한 발전으로 모델 개발 전후 집단의 의료 데이터 양상이 달라져서 이전 데이터로 개발된 모델의 정확도가 떨어질 수 있다.

이러한 심층학습모델의 문제점에도 뛰어난 활용성을 감안하면 이를 합리적으로 사용할 기준이 필요하다[97,101]. 심층학습모델의 광범위한 임상 적용 시 질과 안전성에 대한 엄격한 평가를 통해 현재 임상에서 보편적으로 적용되는 표준과 동등하거나 더 나은 수행능력을 증명해야 한다. 이를 위해 개발된 모델에 대해서 신약 개발 시의 임상시험과 유사하게 실제 훈련한 자료와 다른 다양한 의료환경에서 무작위 전향적 연구를 수행하고 이를 과학적으로 검증하는 것이 중요하다. 이를 가장 잘 보여주는 모델이 구글에서 2016년에 개발한 안저 망막 이미지 평가를 통한 당뇨병성 망막증 진단기이다[12]. 구글은 이를 개선한 모델을 2019년 태국의 의료기관에 설치하고 전향적 데이터를 분석하여 개발한 모델이 태국의 현장 평가자보다 더 우수한 평가결과를 보임을 증명함으로써 이 모델이 의료현장에서 실제 사용 가능함을 제시하였다[102]. 하지만 FDA에서 승인을 받은 인공지능 관련 의료기기를 상당수가 논문 게재를 통한 관련 기술의 객관적인 전문가 평가를 받지 않고 있다[17]. 또한 인공지능 관련 게재 논문의 상당수가 객관적 효능검정에 미흡한 것으로 보고되었다. 2012년 1월부터 2019년 6월 6일까지 출판된 이미지 분석 심층학습 관련 논문을 참여자 특성(데이터 포함/배제 기준, 성비, 인원수), 학습에 사용한 데이터의 출처 및 훈련방법의 적절성, 정확성의 확인(validation) 방법, 결과 표시방법, 수행도 평가기준, 데이터 전처리 방법, 사용한 알고리즘의 적절성 등을 기준으로 평가한 결과, 전향적 연구가 매우 적고 전체 20,530 연구 중 객관적으로 인공지능모델과 전문가의 정확성 비교평가에 사용할 수 있는

연구가 25개에 불과하며 평가 가능한 논문들도 민감도와 특이도의 역치 설정 및 정확도를 검증(validation)하는 방법 등 세부적 평가에서 여러 가지 문제점들이 관찰되었다고 보고되었다[33]. 또한 2019년 2월까지 게재된 중환자실 패혈증 예측 및 예후평가 관련 인공지능 모델 논문 25편을 분석한 결과 편향성(bias)의 위험이 낮고 임상 적용 가능성이 높은 논문이 2편에 불과하다는 보고가 있다[103]. 이러한 우려를 반영하여 기계학습 관련 논문을 평가하는 가이드라인이 제시되었고[104,105], 기계학습을 진단에 적용하는 논문의 제출 시 기본적으로 포함해야 하는 정보나 가이드라인을 제정하고 있다[106].

**2. 인공지능의 임상 적용 시 고려해야 할 문제점**

**1) 인공지능모델의 임상적 효용성 검토 필요성**

모델의 타당성이나 정확도 평가 이외에도 진단 정확도를 향상시키는 인공지능모델이 초기 병변을 더 많이 감지하여 과다 진단에 의한 불필요한 치료를 유도할 가능성이 있다[107]. Marcus [108]는 애플 와치를 이용하여 뇌경색의 위험인자로 알려진 심방세동의 무증상 환자를 스크리닝한 연구가 임상적으로 유용한지를 판단하기 위해서 더 많은 연구가 필요하다고 하였다[73]. 그 이유로는 검사방법의 높은 위양성률에 의한 불필요한 스트레스 유발과 무증상 심방세동 환자의 항응고제 치료에 의한 부작용 등을 제시하였다. 따라서 인공지능모델이 실제 자신의 의료환경(사용 의뢰기기 종류의 차이, 인공지능이 학습한 데이터와 자신의 환자군의 유사성 등)에서 동일한 결과를 재현할 수 있는지, 인공지능이 제공한 정보로부터 자신의 환자 진료계획이 영향을 받고 이러한 변화가 실제 환자의 예후나 병의 경과 개선, 진료비의 절감 등 임상적 효용성을 가지는지를 비판적으로 판단하여 인공지능의 사용 여부를 결정해야 한다[104]. 최근 인공지능모델에 의한 과다 진단의 문제를 줄이기 위해 양성/악성으로의 이분법적 판정보다 경계판정(판단불가판정)을 추가해야 한다는 의견도 제시되고 있다[107].

**2) 인공지능 관련 제품의 정보활용 범위와 기술의 질 관리를 위한 제도 확립**

인공지능 학습의 특징상 이를 기반으로 한 모델들은 지속적인 학습을 통하여 알고리즘이 변화할 수 있다. 일정한 알고리즘을 유지하여 서비스할 수도 있으나 질병의 정의 및 분류기준, 새로운 치료 진단방법의 개발 등이 계속 바뀌는 의학의 특성상 한 번 허가받은 제품이라도 지속적인 업그레이드 없이는 지속적인 사용이 불가능하다. 따라서 한 번 허가받은 제품에서 알고리즘의 지속적인 변경을 허용할 경우 이를 평가할 새로운 기준이 필요하다. 2019년 미국 FDA에서 이 문제에 대해 새로운 기준을 제안하며 제품 하나하나에 대한 평가보다 개발업체의 제품 개발과정에 대한 신뢰도를 평가하여

사전승인(pre-certification)을 부여하고 시판 후 성능을 지속적으로 평가하는 테스트 프로그램을 9개 회사를 대상으로 운영하고 있다 [109].

그러나 Szabo [110]는 현재 FDA에서 승인한 인공지능 관련 제품 상당수가 관련 기술의 근거를 기술보호라는 명분 아래 논문으로 발표하지 않거나 기술에 대한 전문적인 평가가 미흡한 상태로 시장에 출시되고 있으며, 인공지능제품처럼 위험도가 낮다고 평가되어 501K라는 완화된 기준으로 인허가를 받은 기존의 다른 제품들에서 심각한 위험을 초래하는 경우가 있음을 지적하였다. 따라서 기존 제품의 개발과정을 신뢰할 수 있으면 변경을 허용하고 사후 평가하는 제도에 대해 우려의 목소리가 제기되고 있어[110,111], 한국에서도 인공지능 관련 제품의 안전성 및 효능 검증기준 설정 시 검토가 필요하다.

**3) 개인정보 보호 및 보안문제**

인공지능의 학습을 위해 의무기록이 병원에서 개발회사의 서버로 공유되는 과정에서 개인정보 유출에 대한 보호관리정책과 보안에 대한 세심한 고려가 필요하다. 2018년 구글, 캘리포니아대학교-샌프란시스코 캠퍼스, 시카고 의과대학에서 주도한 EMR을 이용한 예측연구[64]와 관련하여 구글과 시카고 의과대학이 개인정보 보호 위반으로 고소되었다[112]. 비록 의무기록이 익명화 처리되었으나 병원에서 진료받은 날짜가 연구에 이용되어 구글의 위치정보 등 다른 정보와 합쳐지면 개인을 특정하는 것이 가능하다는 이유 때문이었다. 스마트네이션을 지향한 싱가포르 최대 의료기관인 Singhealth에서는 리센싱 총리 정보를 포함한 150만의 개인정보가 유출되고 이 중 상당수의 의료기록도 유출되었으며[113], 인공지능 모델의 노이즈 이미지를 첨가하거나 단순히 이미지를 회전시키는 등 조그마한 조작에 의해 양성 사마귀를 악성으로 진단하거나 다른 물체로 인식하는 등 다른 결과를 도출할 수 있으므로 악의적 공격이나 조작으로부터 EMR의 보안을 강화해야 한다[114,115].

**4) 기타 사회적 문제**

**(1) 인공지능의 판단오류에 대한 법적 책임문제**

인공지능기기들이 의료에 다양하게 적용이 되면 인공지능의 오류에 따른 다양한 법적 문제가 제기될 수 있다. 이는 의사와 환자 간의 문제만이 아니고 인공지능제품의 도입 결정에 관련한 병원, 인공지능제품 제조사, 의료사고 관련 보험사 등도 연관될 수 있는 복잡한 문제이다. IDx-DR처럼 인공지능이 최종 판단을 하는 의료기기의 경우 기술적 오류에 의한 책임은 제조사에 있다. 하지만 인공지능이 확률적으로 추천을 하고 의사가 최종 판단을 하게 하는 경우 의사가 법적으로 문제가 되지 않으려면 일반적으로 통용되는 표준기준에 따라 의료행위를 했는지가 현재는 중요한 판단기준이



될 것이다[116]. 다양한 인공지능제품이 의료에 적용될 경우 의사는 어떤 상황에서 인공지능모델을 사용할지, 인공지능의 추천에 어느 정도의 신뢰성을 부여할지에 대해 판단할 수 있어야 한다. 의료에 이용되는 인공지능모델은 인증기관에 의해 기본적인 질 보장이 되고 있으나, 소속기관에서도 각 기관의 특수성에 맞추어 인공지능모델의 수행능력을 지속적이고 체계적으로 점검할 필요가 있다.

(2) 알고리즘과 인간과의 상호작용을 통해 파생될 수 있는 문제에 대한 이해

영국 국가보건의료서비스(National Health System, NHS)의 의뢰에 의하여 영국 왕립의대학술원(Academy of Medical Royal Colleges)에서 작성한 보고서[117]와 Char 등[118]은 의료에 이용되게 되는 고도화된 알고리즘과 인간과의 상호작용에서 예측하지 못한 문제가 발생할 수 있음을 지적하였다. 인공지능에 의한 영상이나 조직 진단시간의 단축이 더 많은 환자를 처리하게 하여 과다 진단에 의한 부작용이 생길 수 있고 자동화된 시스템에 의존하여 관습적으로 결론을 내리는 자동화 편향(automation bias)이 생길 수 있다. 앱을 통해 보급되는 인공지능을 이용한 증상검사 챗봇들과 인터넷을 통해 전문지식을 획득한 환자와 의사 간의 관계에도 변화가 일어날 수 있다[119]. 병리조직검사서에서 인공지능과 의사의 협업이 일반적으로 진단의 정확성을 개선하는 데 기여하나 인공지능이 잘못된 판단을 하는 경우 오히려 의사의 진단 정확도를 감소시키는 것으로 보고되었다[120]. 진료의 표준에 대한 판단이 의사 개인의 경험에 의한 것보다 문헌과 시스템에 축적된 데이터로 결정되는 현재의 시스템에서 의사들이 인공지능모델의 진단과 치료에 대한 권고에 더 의존하게 될 수 있다. 인공지능모델의 개발이 환자 의료의 질 개선 측면보다 대중의 평가와 보험 지급 가능성, 이익 창출 등 시스템의 목표에 맞추어 이루어질 가능성도 제기되고 있다. 또한 인공지능모델이 의료 취약지역에 의료의 질을 향상하는 데 기여할 수도 있지만 인공지능모델이 일부 병원에 제한적으로 제공되어 일부 계층에만 사용됨으로써 의료의 차별성을 악화시킬 수도 있다.

**의학교육 변화의 필요성 및 방향**

향후 의료시스템에서 기계학습의 활용이 더욱 확산되어 의료시스템의 한 구성원으로 자리매김하게 되면 의료의 큰 변화가 예상된다. Wartman과 Combs [121]는 앞으로의 의료의 특징은 (1) 환자의 위치에 무관한 의료제공 환경 구축, (2) 의사, 간호사, 사회복지사, 가정방문간호사, 물리치료사, 가족 등 다양한 의료제공자에 의한 팀 케어, (3) 인공지능 프로그램을 포함한 다양한 소스로부터 제공되는 자료의 활용, (4) 의료에서 인공지능모델의 역할에 대한 효율적 관리일 것으로 요약하였다. 이러한 변화에 능동적으로 대처하기 위해 필요한 역량으로 세계경제포럼에서 제시한 4C의 중요성이 강조

되고 있다(비판적 사고: critical thinking, 소통: communication, 협업: collaboration, 창의력: creativity) [122,123]. 이를 효과적으로 개발하기 위해서는 관련 교육의 변화가 필수적이다. 부산지역 2개 의과대학 및 의예과 학생을 대상으로 인공지능 교육의 필요성에 대한 설문조사결과 응답자의 97% (149/153명)가 필요하다고 하였다[124]. 스탠포드 의과대학에서 시행한 미국의 의사, 전공의 및 의과대학 학생을 대상으로 한 설문조사결과, 학생의 73%가 미래 의료혁신에 더 잘 대비하기 위하여 추가적인 교육의 필요성을 느낀다고 하였고 교육이 필요하다고 생각하는 분야로 44%가 고급 통계학 및 데이터 사이언스를, 36%는 인구집단건강(population health)을 선택하였다[125]. 하지만 의료의 새로운 기술에 대비하는데 현재 받은 교육이 얼마나 도움이 되었는지를 묻는 질문에 학생과 전공의의 18%만이 매우 도움이 된다고 답하여 변화에 상응하는 교육프로그램 개발이 필요함을 보여주었다.

최근 인공지능기술 발전에 따른 의료 변화에 대응하기 위해 미국 의학원(National Academy of Medicine) [95], 미국의사협회 대의원회(House of Delegates) [126]와 영국 NHS [127] 등에서 제시한 교육내용을 요약하면 다음과 같다: (1) 의학의 원리를 이해하여 폭증하는 의학지식에서 유익한 정보를 추출 및 통합하여 해석하고 다양한 상황에 응용하는 역량; (2) 결과를 분석하고 수행능력을 지속적으로 개선하는 자기주도적 평생학습 역량; (3) 기계학습의 원리를 이해하여 기계학습이 제공하는 정보를 해석하고 이의 한계점을 인지하며 이를 효과적이며 비판적으로 환자에게 전달하는 통계학적 지식을 포함한 역량; (4) 인공지능모델, 실시간 모니터링 등을 통해 얻은 많은 데이터를 기반으로 보다 정밀한 치료를 제공하는 역량; (5) 다양한 의료현장에서 적절한 기계학습모델을 취사 선택하고 적용하는 역량; (6) 인공지능모델을 포함한 의료시스템의 구성원들과 소통하고 이들 사이의 복잡성을 관리하며 협업하는 역량; (7) 개별적으로 특이적인 환자의 상황을 공감(empathy/compassion)하며 양방향으로 소통하는 역량.

**1. 학부교육**

1) 적응형 평생학습자(adaptive lifelong learner)로 교육

의료환경과 의학지식의 발전속도가 가속화되고 있는 상황에서 의학교육은 교과서적인 지식을 전달하고 기억하는 것이 아니라 변화하는 환경에서 자신에게 필요한 지식을 스스로 찾아 학습할 수 있는 역량을 키우는 데 초점을 맞추어야 한다[128]. 2013년 미국의학협회의 주도로 설립된 “The Accelerating Change in Medical Education Consortium”에서도 이를 향후 의학교육 개선의 주요 목표로 설정하였다[129]. 반더빌트 의과대학 교육부장인 Cutrer 교수는 (1) 학습에 대한 계획, (2) 목표 달성을 위한 적절한 방법을 이용한 학습, (3) 학습결과에 대한 자신 및 외부의 평가 및 피드백, (4) 문제점에



대한 성찰과 학습방법 조정의 4단계 모델을 현장에서 적응형 평생학습자로 교육하기 위한 모델로 제시하고 이를 효과적으로 적용하기 위해서는 멘토나 코칭, 학습자의 수행 성과에 대한 시기적절하고 정확한 평가결과 제공 등이 필요하다고 하였다[130,131]. 또한 하버드 의과대학이 2015년 입학생부터 의학교육에 플립러닝을 전면 도입한 것처럼 세계 주요 의과대학의 의학교육은 강의를 통한 지식의 전달에서 벗어나 플립러닝, 소그룹토론, 학생연구 등을 통한 학생주도학습, 임상 조기 노출, 현장 중심 교육 등을 통해 적응형 평생학습 능력 개발을 지향하고 있다[132]. 또한 이미 알려진 교과서적 의학지식 외에 복잡한 의료 관련 사회문제 및 전공을 포괄하는 의학적 문제에 대해 강의, 현장방문, 저널클럽, 세미나, 소그룹모임 등 다양한 방법으로 학생주도적 해법을 찾거나 멘토와의 만남 등을 통해 그동안의 학습성과를 성찰하는 교육프로그램들을 교육과정에 도입하고 있다(Table 3) [133-137]. 반더빌트 의과대학(Vanderbilt University School of Medicine)은 “VSTAR”이라는 평생학습 플랫폼을 사용하여 학습자와 멘토가 학습자 개인의 학습성과를 시기적절하게 확인하고 성찰하거나 피드백 함으로써 효과적인 적응형 평생학습 역량 개발을 지원하고 있다[138].

평생학습, 비판적 사고와 과학적 사고 역량 개발에 학생 연구과정이 중요한 기여를 할 수 있으며, 각 대학의 자원과 대학의 사명, 인재양성 목표에 따라 다양한 프로그램을 필수 및 선택과정으로 운영하고 하고 있다. 미국의 연구역량 상위권 의과대학은 전체 학생을 대상으로 대부분 12주 이상의 심화 학생 연구프로그램을 운영하고 있으며, 이와 연계한 다양한 교육프로그램을 전 학년에 걸쳐 운영하고 있다. 캘리포니아대학교-샌프란시스코 캠퍼스 의과대학의 경우 (1) 핵심 과정(1, 2학년 동안 1주일에 1번 해당 기간에 학습한 기본의학과과정 내용과 연계한 저널클럽, 사례탐구, 세미나) [139], (2) inquiry immersion (1학년 중간 2주, 지속적으로 진화하고 불완전한 의학지식의 다양한 분야를 토의) [140], (3) deep explore(필수 임상실습 후 12-20주 이 기간 중 4주의 임상연구설계

교육을 통해 향후 수행할 프로젝트의 계획서를 작성함)로 구성된 3단계 학생 연구 관련 교육프로그램을 1학년부터 4학년까지 편성하여 자기주도적으로 연구과제를 도출하는 역량을 배양하고 있다 [141]. 우리나라 의과대학도 대학에 따라 필수 및 선택과정으로 의학 연구를 도입하고 있으나 실질적으로 학생들의 비판적 사고와 자기주도학습 역량 개발에 기여하고 연구과제의 학생주도적 개발에 필요한 연계교육프로그램의 설계 등 운영방안의 검토가 필요하다.

2) 데이터 사이언스 교육

데이터 사이언스는 많은 데이터로부터 유용한 패턴을 찾아내기 위해 이용되는 일련의 원리, 문제의 정의, 알고리즘, 과정들을 포괄하며, 수학, 통계, 컴퓨터과학, 정보과학, 기계학습 관련 기술 등 많은 분야에서 도출한 기술과 이론을 이용한다[142]. 2018년 Kolachalama와 Garg [143]는 의과대학 학부교육과정에서 기계학습에 능숙해지는 것이 목표가 아니라 기계학습을 이해하고 활용하는 디지털 문해력(digital literacy)을 목표로 해야 한다고 제안하였다. 따라서 의과대학의 제한된 교육시간 등을 감안하면 기계학습의 알고리즘을 구체적으로 교육하는 것보다 (1) 의료에 활용되는 기계학습의 작동원리, (2) 기계학습의 기초가 되는 데이터의 특성, 검색, 유의한 정보 추출원리, (3) 기계학습에 제안하는 확률을 비판적으로 해석하고 환자에게 정확한 정보를 전달할 수 있는 지식, (4) 인구집단건강(population health)과 질병예측, 위험도 평가, 관리 등에 정보기술의 활용, (5) 기계학습의 의료에서의 활용범위를 이해하고 기술적, 윤리적, 법적 문제점 등에 초점을 맞추어 교육내용을 편성하고 실제 이미 개발된 인공지능 수단을 활용하는 실습을 통하여 향후 어떤 기준으로 인공지능기술을 의료에 사용할지, 사용할 때 유의해야 할 점이 무엇인지를 파악할 수 있도록 교육해야 한다. 2019년 Paranjape 등[144]은 이와 관련한 의학교육 단계별 교육프로그램을 제안하고 있다.

학부교육과정에서 이 분야 교육프로그램 개발이 초창기인 관계

Table 3. List of interdisciplinary programs for adaptive learning

Medical school	Name of program	Format	References
Harvard University	Advanced integrated science course	4 weeks, 2 courses during 3rd and 4th year	[133]
Johns Hopkins University	Topics in interdisciplinary medicine	3-4 days, 16 times over 1-4 years	[134]
	Translational science intersession	1 day, 4 times during core clinical clerkship, 8-week intervals	[135]
Stanford University	Reflections, research, & advances in patient care	1 day, 8-week intervals during core clinical clerkship	[136]
University of California, San Francisco	Inquiry immersion	2 weeks mid 1st year	[137]

Table 4. Examples of artificial intelligence related elective courses

Medical school	Course name	Reference
Boston University	Machine learning for biomedical applications	[145]
Harvard University	Computationally-enabled medicine	[146]
University of Pittsburgh	Artificial intelligence and machine learning in healthcare	[147]
Stanford University	Precision medicine and big data	[148]

로 아직 정규 교육과정에 필수과목으로 편성되어 교육하는 예는 없으나 선택과목 개설, 기존에 개설된 통합 교과목을 활용(세미나, 저널클럽, 사례연구), 관련 연구과정 참여 등으로 다양한 교육이 이루어지고 있다. 미국 의과대학에서 개설된 선택과목의 예는 Table 4에서 제시하였으며[145-148], 국내에는 연세대학교 및 울산대학교 등에서 선택과목으로 관련 내용 교육을 시행하고 있다[149]. 아직 인공지능 관련 교육이 체계적으로 이루어지지 못하는 이유 중의 하나는 의과대학의 교원만으로는 관련 내용을 교육할 전문가가 부족한 것이다. 미국 의사협회는 체계적인 인공지능교육을 위하여 의과대학에 데이터 사이언스나 소프트웨어 전문가를 영입할 것을 권고하고 있다[126].

### 3) 소통과 공감능력 교육의 강화

Topol [150]은 인공지능모델에 의해 의사의 행정업무 부담, 의사 결정 시간 등 기존 업무에 할당하는 시간이 줄어들면 환자와 직접 소통하는 시간이 늘어나 디지털화되며 훼손된 의료의 인간성을 회복할 것으로 예측하였다. 환자와 소통하는 시간이 늘어나면 환자와의 공감을 통해 인공지능모델이 획득할 수 없는 정보를 얻어 환자의 상황에 맞는 맞춤치료를 하는 데 도움이 될 것이다. 소통의 측면에서는 기존의 교육과정에 포함된 환자, 의사, 사회 간의 소통뿐 아니라 인공지능 학습모델의 적용과 환자 개인이 보유한 의료정보 활용 필요성 등을 반영한 소통교육 개발이 필요하다.

### 4) 교육과정과 운영의 검토

우리나라의 의과대학은 2012년 2주기 인증평가를 계기로 성과바탕교육을 통하여 2014년 대한의사협회에서 제정한 한국의 의사상에서 제시한 환자진료, 소통과 협력, 사회적 책무성, 전문직업성, 교육 및 연구역량을 포함한 다양한 역량 개발을 교육목표로 하고 있다[151,152]. 또한 2019년부터 적용된 의학교육 평가기준은 미래 의료환경 변화에 대비한 교육을 권장하고 있다[153]. 그러나 의예과에서 기초의학 과목을 강의하는 경우를 포함하면 2-2.5년의 임상실습 전 교육기간에 많은 양의 의학지식을 강의 위주로 교육하고 있고, 딱 짜인 교육과정으로 새로운 교육내용을 추가하는 것이 쉽지 않아 창의적 자기주도학습 역량, 의학지식/환자진료역량 외 다른 역량 개발, 미래환경에 대비한 교육 등이 미흡하다. 또한 취득 학점과 진급 여부가 의학지식의 성적에 따라 좌우되며 의사국가시험 대비에 대한 압박 등으로 상당수의 학생들이 학습량 과다에 따른 스트레스를 호소하고 있다. 학습스트레스와 번아웃 사이의 관련성이 보고되고 있고[154], 번아웃된 학생은 의사의 사회적 책무성에 대해서 이타적 관점을 보다 적게 가지며[155], 번아웃이 환자에 대한 공감능력을 감소시키며 오류가 증가하는 것으로 보고되었다[156]. 이러한 결과들은 성과바탕학습에서 개발하고자 하는 각 요소 간에 상호작용이 있음을 시사한다.

인공지능과 유전체학 등 의료기술의 발전으로 초래될 미래 의료 환경에 적절히 대응할 수 있는 미래 의료인력을 양성하기 위해서는 데이터 사이언스 관련 교육프로그램을 단순히 추가하고 공감교육을 강화하는 것 등의 부분적인 교육과정 수정만으로 가능하지 않다. 임상실습 전 교육기간에 집중된 의학지식 교육의 양과 내용을 학부 교육기간 내 균형적으로 재배분하고 각 역량이 균형적으로 개발되며 미래 변화에 적극적으로 대응하는 교육이 이루어지기 위해서는 학생들의 역량 및 수준에 따라 개별화된 교육내용이 적용될 수 있고 졸업 후 교육과 연계된 융통성 있고 유기적인 교육과정의 개발이 필요하다. 또한 지나친 의학지식/환자진료 위주 학습형태 변화를 유도하기 위하여 국내 일부 의과대학에서도 도입하고 있는 절대평가제 및 의학지식/진료역량 외 성과에 대한 다양하고 실질적인 평가기법을 도입하고 학생들에게 적극적으로 피드백하여 균형 있는 역량 개발이 이루어질 수 있도록 운영하는 것이 필요하다.

## 2. 졸업 후 교육

미국 의학원 보고서에 지적한 것처럼 졸업 후에는 다양한 임상환경에서 어떤 종류의 인공지능 알고리즘을 어떻게 적용할지를 판단하고, 실제 환자와 가족에게 관련 내용을 적절히 전달 및 토의하고, 인공지능 실제 적용 시의 유용성을 평가하고 발생할 수 있는 여러 가지 문제점을 파악하며 적절히 해결할 수 있는 역량을 갖추어야 한다[95]. 이를 위해서는 졸업 후에도 이와 관련된 체계적인 교육이 필요하다. 영국의 보건복지부 산하 Health Education England에서 작성한 보고서는 교육환경의 기반 확충을 위해 정부와 관련 기관의 체계적인 교육프로그램과 인력양성 프로그램이 필요하며 각 전공학회에서도 전공에 맞는 교육프로그램 개발이 필요함을 제시하였다[127]. 인공지능을 사용하는 병원은 지속적으로 사용자에게 인공지능기술이 적절하게 적용될 수 있도록 교육하고 관련 교육의 질 향상을 위해 해당 분야 전문가를 확보해야 함을 권고하였다. 미국과 영국에서는 의료분야에 디지털 헬스케어 관련 전문가를 육성하는 프로그램을 개발하여 운영하고 있다. 2019년 영국에서는 Topol [150]의 리뷰 권고에 따라 이 분야의 임상전문가를 양성하기 위하여 2019년 디지털 펠로우쉽 과정을 개설하였다[127]. 이 과정에서는 임상 의사, 약사, 병원행정가, 물리치료사, 임상연구자, 간호사 등 의료분야 종사자 중 17명을 훈련대상자로 선발하여 평소 담당 업무와 함께 해당 기관의 디지털화 개선과 혁신을 위한 프로그램을 수행하고 관련 종사자를 교육하는 업무를 하는 데 필요한 교육을 실시하고 있다[157]. 미국에는 비영리법인이 설립되어 의사를 포함한 의료 관계 인력에 대해 의료 관련 인공지능교육과 시험을 통해 미국 의료 인공지능사(American Board of Artificial Intelligence in Medicine) 자격증을 부여하는 제도가 도입되었다[158].

## 결론

심층학습을 포함한 인공지능기술이 급속도로 발전하여 질병의 진단 및 치료뿐 아니라 질병의 예후 및 발생 위험도 예측, 건강지표 모니터링, 약물 개발, 의료업무의 효율 개선 등 다양한 의료분야에 적용되는 기술이 개발되고 있다. 이들 기술 중 일부는 규제 당국의 공식 인가를 받아 의료현장에 활용되고 있으며, 이러한 변화가 가속화되면 의료환경의 큰 변화가 예측된다. 그러나 인공지능기술이 의료의 질을 향상시키는 긍정적 방향으로 적용되기 위해서는 인공지능 모델의 개발 환경과 다른 다양한 실제 의료현장에서의 전향적 연구를 통해 훈련 데이터의 편향성, 알고리즘의 투명성을 포함한 알고리즘의 기술적 문제의 유무 판단과 객관적인 효과 검증이 필요하며 나아가 이들 모델이 의료의 질과 환자의 임상적 결과를 개선할 수 있는지도 검토되어야 한다. 또한 인공지능 의료기술의 안전성을 확보하기 위한 규제 및 관리시스템의 정비, 인공지능기술의 확산에 따라 생길 수 있는 윤리적 및 법적 문제, 의사-환자 및 사회관계의 변화 등 다양한 분야에 대한 검토가 필요하다. 이러한 의료환경의 급속한 변화에 능동적으로 적응할 수 있는 의료인력에 대한 체계적인 교육이 필요하다. 다량의 정보로부터 의미 있는 지식을 추출하는 역량, 데이터 사이언스 능력, 환자와의 공감과 다양한 의료 관련 집단과의 소통능력 등을 효과적으로 교육하기 위해서는 의과대학 학부교육과정의 전반적인 검토가 필요하며 졸업 후에도 실제 임상환경에서 인공지능모델을 적절히 활용하는 데 필요한 각 전문 분야에 맞는 특화된 교육프로그램이 필요하다.

## 저자 기여

정진섭: 자료조사 및 논문작성 등 전반적인 활동 수행

## 감사의 글

이 논문은 부산대학교 기본연구지원사업(2년)에 의하여 연구되었다.

## REFERENCES

1. Lexico. Meaning of artificial intelligence in English [Internet]. Oxford: Lexico, Oxford University Press; c2020 [cited 2020 Jan 9]. Available from: [https://www.lexico.com/definition/artificial\\_intelligence](https://www.lexico.com/definition/artificial_intelligence).
2. Piccinini G. The first computational theory of mind and brain: a close look at McCulloch and Pitts's "logical calculus of ideas immanent in nervous activity". *Synthese*. 2004;141(2):175-215.
3. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth summer research project on artificial intelligence, August

- 31, 1955. *AI Mag*. 2006;27(4):12.
4. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev*. 1958;65(6):386-408.
5. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323(6088):533-6.
6. Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput*. 2006;18(7):1527-54.
7. Le Cun Y, Jackel LD, Boser B, Denker JS, Graf HP, Guyon I, et al. Handwritten digit recognition: applications of neural net chips and automatic learning. In: Sanchez-Sinencio E, Lau C, editors, *Artificial neural networks*. Piscataway (NJ): IEEE Press; 1992. p. 463-8.
8. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1097-105.
9. Le QV, Ranzato MA, Monga R, Devin M, Chen K, Corrado GS, et al. Building high-level features using large scale unsupervised learning [Internet]. Ithaca (NY): arXiv, Cornell University; 2012 [cited 2020 Jan 9]. Available from: <https://arxiv.org/abs/1112.6209>.
10. Hinton G, Deng L, Yu D, Dahl GE, Mohamed AR, 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(6):82-97.
11. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, et al. Mastering the game of Go with deep neural networks and tree search. *Nature*. 2016;529(7587):484-9.
12. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-10.
13. Singh H, Meyer AN, 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(9):727-31.
14. Balogh EP, Miller BT, Ball JR; Committee on Diagnostic Error in Health Care; Board on Health Care Services; Institute of Medicine, et al. *Improving diagnosis in health care*. Washington (DC): National Academies Press; 2015.
15. Dewa CS, Loong D, Bonato S, Trojanowski L. The relationship between physician burnout and quality of healthcare in terms of safety and acceptability: a systematic review. *BMJ Open*. 2017;7(6):e015141.
16. Budd K; Association of American Medical Colleges. Will artificial intelligence replace doctors? [Internet]. Washington (DC): Association of American Medical Colleges; 2019 [cited 2020 Jan 9]. Available from: <https://www.aamc.org/news-insights/will-artificial-intelligence-replace-doctors>.
17. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56.
18. Coiera E. The fate of medicine in the time of AI. *Lancet*. 2018; 392(10162):2331-2.
19. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94-8.
20. Gong B, Nugent JP, Guest W, Parker W, Chang PJ, Khosa F, et al. Influence of artificial intelligence on canadian medical students' preference for radiology specialty: a national survey study. *Acad Radiol*. 2019;

- 26(4):566-77.
21. Choy G, Khalilzadeh O, Michalski M, Do S, Samir AE, Pinykh OS, et al. Current applications and future impact of machine learning in radiology. *Radiology*. 2018;288(2):318-28.
  22. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*. 2019;292(1):60-6.
  23. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology: new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol*. 2019;16(11):703-15.
  24. Chen PC, Gadepalli K, MacDonald R, Liu Y, Kadowaki S, Nagpal K, et al. An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. *Nat Med*. 2019;25(9):1453-7.
  25. Luo H, Xu G, Li C, He L, Luo L, Wang Z, et al. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *Lancet Oncol*. 2019; 20(12):1645-54.
  26. Mori Y, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, et al. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. *Ann Intern Med*. 2018;169(6):357-66.
  27. Le Berre C, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smail-Tabbone M, et al. Application of artificial intelligence to gastroenterology and hepatology. *Gastroenterology*. 2020;158(1):76-94.
  28. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. *J Dermatolog Treat*. 2019:1-15.
  29. Dick V, Sinz C, Mittlbock M, Kittler H, Tschandl P. Accuracy of computer-aided diagnosis of melanoma: a meta-analysis. *JAMA Dermatol*. 2019;155(11):1291-9.
  30. Cuocolo R, Perillo T, De Rosa E, Ugga L, Petretta M. Current applications of big data and machine learning in cardiology. *J Geriatr Cardiol*. 2019;16(8):601-7.
  31. Ting DS, Peng L, Varadarajan AV, Keane PA, Burlina PM, Chiang MF, et al. Deep learning in ophthalmology: the technical and clinical considerations. *Prog Retin Eye Res*. 2019;72:100759.
  32. 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(9):1342-50.
  33. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health*. 2019;1(6):e271-97.
  34. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. 2019;25(6):954-61.
  35. 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.
  36. Liu Y, Jain A, Eng C, Way DH, Lee K, Bui P, et al. A deep learning system for differential diagnosis of skin diseases. *Nat Med*. 2020 May 18 [Epub]. <https://doi.org/10.1038/s41591-020-0842-3>.
  37. Majkowska A, Mittal S, Steiner DF, Reicher JJ, McKinney SM, Duggan GE, et al. Chest radiograph interpretation with deep learning models: assessment with radiologist-adjudicated reference standards and population-adjusted evaluation. *Radiology*. 2020;294(2):421-31.
  38. Hwang EJ, Nam JG, Lim WH, Park SJ, Jeong YS, Kang JH, et al. Deep learning for chest radiograph diagnosis in the emergency department. *Radiology*. 2019;293(3):573-80.
  39. Hsieh TC, Mensah MA, Pantel JT, Aguilar D, Bar O, Bayat A, et al. PEDIA: prioritization of exome data by image analysis. *Genet Med*. 2019;21(12):2807-14.
  40. Bhagwat N, Viviano JD, Voineskos AN, Chakravarty MM; Alzheimer's Disease Neuroimaging Initiative. Modeling and prediction of clinical symptom trajectories in Alzheimer's disease using longitudinal data. *PLoS Comput Biol*. 2018;14(9):e1006376.
  41. Dey D, Gaur S, Ovrehus KA, Slomka PJ, Betancur J, Goeller M, et al. Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study. *Eur Radiol*. 2018;28(6):2655-64.
  42. Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. 2018;2(3):158-64.
  43. Sharafi SM, Sylvestre JP, Chevrefils C, Soucy JP, Beaulieu S, Pascoal TA, et al. Vascular retinal biomarkers improves the detection of the likely cerebral amyloid status from hyperspectral retinal images. *Alzheimers Dement (N Y)*. 2019;5:610-7.
  44. Mitani A, Huang A, Venugopalan S, Corrado GS, Peng L, Webster DR, et al. Detection of anaemia from retinal fundus images via deep learning. *Nat Biomed Eng*. 2020;4(1):18-27.
  45. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-67.
  46. Kather JN, Pearson AT, Halama N, Jager D, Krause J, Loosen SH, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med*. 2019;25(7):1054-6.
  47. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med*. 2018;15(11):e1002711.
  48. Lu H, Arshad M, Thornton A, Avesani G, Cunnea P, Curry E, et al. A mathematical-descriptor of tumor-mesoscopic-structure from computed-tomography images annotates prognostic- and molecular-phenotypes of epithelial ovarian cancer. *Nat Commun*. 2019;10(1):764.
  49. Mesko B. FDA approvals for smart algorithms in medicine in one giant infographic [Internet]. [place unknown]: The Medical Futurist; 2019 [cited 2020 Jan 10]. Available from: <https://medicalfuturist.com/fda-approvals-for-algorithms-in-medicine>.
  50. Stark A. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems [Internet]. Silver Spring (MD): The U.S. Food and Drug Administration; 2018 [cited 2020 Jan 10]. Available from: <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye>.
  51. Elfin D. Philips to offer Paige.AI's prostate cancer detection tech [Internet]. Washington (DC): MEDTECHDIVE; 2019 [cited 2020 Jan



- 10]. Available from: <https://www.medtechdive.com/news/philips-to-offer-paigeais-prostate-cancer-detection-tech/568514>.
52. Stein A. Zebra Medical Vision's solutions now available on Philips IntelliSpace AI workflow suite, enabling more AI capabilities for radiologists [Internet]. San Francisco (CA): Businesswire; 2019 [cited 2020 Jan 10]. Available from: <https://www.businesswire.com/news/home/20191203005716/en/Zebra-Medical-Vision%E2%80%99s-Solutions-Philips-IntelliSpace-AI>.
  53. Kincaid E. FDA clears GE Healthcare's AI triage algorithm on X-ray device [Internet]. New York (NY): Medscape; 2019 [cited 2020 Jan 10]. Available from: <https://www.medscape.com/viewarticle/918340>.
  54. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-58.
  55. Somashekhar SP, Sepulveda MJ, Puglielli S, Norden AD, Shortliffe EH, Rohit Kumar C, et al. Watson for oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board. *Ann Oncol*. 2018;29(2):418-23.
  56. Liu C, Liu X, Wu F, Xie M, Feng Y, Hu C. Using artificial intelligence (Watson for oncology) for treatment recommendations amongst Chinese patients with lung cancer: feasibility study. *J Med Internet Res*. 2018;20(9):e11087.
  57. Pantuck AJ, Lee DK, Kee T, Wang P, Lakhota S, Silverman MH, et al. Modulating BET bromodomain inhibitor ZEN-3694 and enzalutamide combination dosing in a metastatic prostate cancer patient using CURATE. AI, an artificial intelligence platform. *Adv Therap*. 2018;1(6):1800104.
  58. Nikolov S, Blackwell S, Mendes R, De Fauw J, Meyer C, Hughes C, et al. Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy [Internet]. Ithaca (NY): arXiv, Cornell University; 2018 [cited 2020 Jan 10]. Available from: <https://arxiv.org/abs/1809.04430>.
  59. Sennaar K. Machine learning in surgical robotics: 4 applications that matter [Internet]. Newton (MA): EMERJ; 2019 [cited 2020 Jan 10]. Available from: <https://emerj.com/ai-sector-overviews/machine-learning-in-surgical-robotics-4-applications>.
  60. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg*. 2018;268(1):70-6.
  61. Wang Z, Fey AM. SATR-DL: improving surgical skill assessment and task recognition in robot-assisted surgery with deep neural networks. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:1793-6.
  62. Bain EE, Shafner L, Walling DP, Othman AA, Chuang-Stein C, Hinkle J, et al. Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a phase 2 clinical trial in subjects with schizophrenia. *JMIR Mhealth Uhealth*. 2017;5(2):e18.
  63. Eggerth A, Hayn D, Schreier G. Medication management needs information and communications technology-based approaches, including telehealth and artificial intelligence. *Br J Clin Pharmacol*. 2019 Jul 4 [Epub]. <https://doi.org/10.1111/bcp.14045>.
  64. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med*. 2018;1:18.
  65. Liang H, Tsui BY, Ni H, Valentim CC, Baxter SL, Liu G, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. *Nat Med*. 2019;25(3):433-8.
  66. Tomasev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*. 2019;572(7767):116-9.
  67. Safavi KC, Khaniyev T, Copenhaver M, Seelen M, Zenteno Langle AC, Zanger J, et al. Development and validation of a machine learning model to aid discharge processes for inpatient surgical care. *JAMA Netw Open*. 2019;2(12):e1917221.
  68. Van Steenkiste T, Ruysinck J, De Baets L, Decruyenaere J, De Turck F, Ongenaes F, et al. Accurate prediction of blood culture outcome in the intensive care unit using long short-term memory neural networks. *Artif Intell Med*. 2019;97:38-43.
  69. Alaa AM, Bolton T, Di Angelantonio E, Rudd JH, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK Biobank participants. *PLoS One*. 2019;14(5):e0213653.
  70. Kwon JM, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest. *J Am Heart Assoc*. 2018;7(13):e008678.
  71. Meiring C, Dixit A, Harris S, MacCallum NS, Brealey DA, Watkinson PJ, et al. Optimal intensive care outcome prediction over time using machine learning. *PLoS One*. 2018;13(11):e0206862.
  72. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. *JAMA Cardiol*. 2018;3(5):409-16.
  73. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med*. 2019;381(20):1909-17.
  74. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, et al. Artificial intelligence for mental health and mental illnesses: an overview. *Curr Psychiatry Rep*. 2019;21(11):116.
  75. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. *Cell*. 2015;163(5):1079-94.
  76. Kent J. Amazon introduces machine learning medical transcription service [Internet]. Danvers (MA): HealthITAnalytics; 2019 [cited 2020 Jan 10]. Available from: <https://healthitanalytics.com/news/amazon-introduces-machine-learning-medical-transcription-service>.
  77. Maier-Hein L, Vedula SS, Speidel S, Navab N, Kikinis R, Park A, et al. Surgical data science for next-generation interventions. *Nat Biomed Eng*. 2017;1(9):691-6.
  78. Zhu H. Big data and artificial intelligence modeling for drug discovery. *Annu Rev Pharmacol Toxicol*. 2020;60:573-89.
  79. Basile AO, Yahi A, Tatonetti NP. Artificial intelligence for drug toxicity and safety. *Trends Pharmacol Sci*. 2019;40(9):624-35.
  80. Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. *Trends Pharmacol Sci*. 2019;40(8):577-91.
  81. Sumitomo Dainippon Pharma: discovering and designing drugs with artificial intelligence [Internet]. Annecy: MarketScreener; 2020 [cited 2020 Jan 10]. Available from: <https://www.marketscreener.com/SUMITOMO-DAINIPPON-PHARMA-6492512/news/Sumitomo-Dainippon-Pharma-Discovering-and-designing-drugs-with-artificial-intelligence-30106469/>.

82. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. *Cell*. 2020;180(4):688-702.
83. Van Veen F. The neural network zoo [Internet]. Utrecht: The Asimov Institute; 2016 [cited 2020 Jan 10]. Available from: <https://www.asimovinstitute.org/neural-network-zoo/>.
84. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks [Internet]. Ithaca (NY): arXiv, Cornell University; 2014 [cited 2020 Jan 10]. Available from: <https://arxiv.org/pdf/1406.2661.pdf>.
85. Borkowski AA, Wilson CP, Borkowski SA, Thomas LB, Deland LA, Grewe SJ, et al. Comparing artificial intelligence platforms for histopathologic cancer diagnosis. *Fed Pract*. 2019;36(10):456-63.
86. Faes L, Wagner SK, Fu DJ, Liu X, Korot E, Ledam JR, et al. Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. *Lancet Digit Health*. 2019;1(5):e232-42.
87. Schulte F, Fry E. No safety switch: how lax oversight of electronic health records puts patients at risk [Internet]. [place unknown]: Kaiser Health News; 2019 [cited 2020 Jan 10]. Available from: <https://khn.org/news/no-safety-switch-how-lax-oversight-of-electronic-health-records-puts-patients-at-risk>.
88. Dastin J. Amazon scraps secret AI recruiting tool that showed bias against women [Internet]. New York (NY): Reuters; 2018 [cited 2020 Jan 10]. Available from: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.
89. Simonite T. When it comes to gorillas, Google photos remains blind [Internet]. Boone (IA): Weird; 2018 [cited 2020 Jan 10]. Available from: <https://www.wired.com/story/when-it-comes-to-gorillas-google-photo-s-remains-blind/>.
90. Polonski V. AI is convicting criminals and determining jail time, but is it fair? [Internet]. Cologny: World Economic Forum; 2018 [cited 2020 Jan 10]. Available from: <https://www.weforum.org/agenda/2018/11/algorithms-court-criminals-jail-time-fair/>.
91. Lashbrook A. AI-driven dermatology could leave dark-skinned patients behind [Internet]. [place unknown]: The Atlantic; 2018 [cited 2020 Jan 10]. Available from: <https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/>.
92. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-53.
93. Arrieta AB, Diaz-Rodriguez N, Ser JD, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI [Internet]. Ithaca (NY): arXiv, Cornell University; 2019 [cited 2020 Jan 10]. Available from: <https://arxiv.org/abs/1910.10045>.
94. Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B. Interpretable machine learning: definitions, methods, and applications [Internet]. Ithaca (NY): arXiv, Cornell University; 2019 [cited 2020 Jan 10]. Available from: <https://arxiv.org/abs/1901.04592>.
95. Matheny M, Israni ST, Auerbach A, Beam A, Bleicher P, Chapman W, et al. Artificial intelligence in health care: the hope, the hype, the promise, the peril [Internet]. Washington (DC): National Academy of Medicine; 2019 [cited 2020 Jan 10]. Available from: <https://nam.edu/artificial-intelligence-special-publication/>.
96. Bender E. Unpacking the black box in artificial intelligence for medicine [Internet]. Cambridge (MA): Undark; 2019 [cited 2020 Jan 10]. Available from: <https://undark.org/2019/12/04/black-box-artificial-intelligence/>.
97. Wang F, Kaushal R, Khullar D. Should health care demand interpretable artificial intelligence or accept “black box” medicine? *Ann Intern Med*. 2020;172(1):59-60.
98. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS Med*. 2018;15(11):e1002683.
99. Winkler JK, Fink C, Toberer F, Enk A, Deinlein T, Hofmann-Wellenhof R, et al. Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. *JAMA Dermatol*. 2019;155(10):1135-41.
100. 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):115-8.
101. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. 2019;17(1):195.
102. Raumviboonsuk P, Krause J, Chotcomwongse P, Sayres R, Raman R, Widner K, et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. *NPJ Digit Med*. 2019;2:25.
103. Schinkel M, Paranjape K, Nannan Panday RS, Skyttberg N, Nanayakkara PW. Clinical applications of artificial intelligence in sepsis: a narrative review. *Comput Biol Med*. 2019;115:103488.
104. Liu Y, Chen PC, Krause J, Peng L. How to read articles that use machine learning: users’ guides to the medical literature. *JAMA*. 2019;322(18):1806-16.
105. Doshi-Velez F, Perlis RH. Evaluating machine learning articles. *JAMA*. 2019;322(18):1777-9.
106. Collins GS, Moons KG. Reporting of artificial intelligence prediction models. *Lancet*. 2019;393(10181):1577-9.
107. Adamson AS, Welch HG. Machine learning and the cancer-diagnosis problem: no gold standard. *N Engl J Med*. 2019;381(24):2285-7.
108. Marcus GM. The Apple Watch can detect atrial fibrillation: so what now? *Nat Rev Cardiol*. 2020;17(3):135-6.
109. Digital Health Software Precertification (Pre-Cert) Program [Internet]. Silver Spring (MD): U.S. Food & Drug Administration; 2019 [cited 2020 Jan 10]. Available from: <https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-pre-cert-program>.
110. Szabo L. Artificial intelligence is rushing into patient care: and could raise risks [Internet]. New York (NY): Scientific American; 2019 [cited 2020 Jan 10]. Available from: <https://www.scientificamerican.com/article/artificial-intelligence-is-rushing-into-patient-care-and-could-raise-risks/>.
111. Lee TT, Kesselheim AS. U.S. Food and Drug Administration precertification pilot program for digital health software: weighing the

- benefits and risks. *Ann Intern Med.* 2018;168(10):730-2.
112. Wakabayashi D. Google and the University of Chicago are sued over data sharing [Internet]. New York (NY): The New York Times; 2019 [cited 2020 Jan 10]. Available from: <https://www.nytimes.com/2019/06/26/technology/google-university-chicago-data-sharing-laws-uit.html>.
  113. Singapore healthcare cyberattack: officials will not name hackers who targeted Prime Minister Lee Hsien Loong [Internet]. Hong Kong: South China Morning Post; 2019 [cited 2020 Jan 10]. Available from: <https://www.scmp.com/news/asia/southeast-asia/article/2182222/singapore-healthcare-cyberattack-officials-will-not-name>.
  114. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. *Science.* 2019; 363(6433):1287-9.
  115. Heaven D. Why deep-learning AIs are so easy to fool. *Nature.* 2019;574(7777):163-6.
  116. Price WN 2nd, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. *JAMA.* 2019 Oct 4 [Epub]. <https://doi.org/10.1001/jama.2019.15064>.
  117. Academy of Medical Royal Colleges. Artificial intelligence in healthcare [Internet]. London: Academy of Medical Royal Colleges; 2019 [cited 2020 Jan 10]. Available from: <https://www.aomrc.org.uk/reports-guidance/artificial-intelligence-in-healthcare/>.
  118. Char DS, Shah NH, Magnus D. Implementing machine learning in health care: addressing ethical challenges. *N Engl J Med.* 2018; 378(11):981-3.
  119. The top 12 health chatbots [Internet]. [place unknown]: The Medical Futurist; 2020 [cited 2020 Jan 10]. Available from: <https://medicalfuturist.com/top-12-health-chatbots/>.
  120. Kiani A, Uyumazturk B, Rajpurkar P, Wang A, Gao R, Jones E, et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. *NPJ Digit Med.* 2020;3:23.
  121. Wartman SA, Combs CD. Medical education must move from the information age to the age of artificial intelligence. *Acad Med.* 2018;93(8):1107-9.
  122. World Economic Forum. 2015 New vision for education unlocking the potential of technology [Internet]. Cologne: World Economic Forum; 2015 [cited 2020 Jan 10]. Available from: [http://www.weforum.org/docs/WEFUSA\\_NewVisionforEducation\\_Report2015.pdf](http://www.weforum.org/docs/WEFUSA_NewVisionforEducation_Report2015.pdf).
  123. Shah NR, Lee TH. What AI means for doctors and doctoring [Internet]. Waltham (MA): NEJM Catalyst; 2019 [cited 2020 Jan 10]. Available from: <https://catalyst.nejm.org/doi/full/10.1056/CAT.19.0622>.
  124. Kwon YJ, Kim CH, Nam IH, Yoo SW, Jung JS. Current status of artificial intelligence in healthcare and awareness survey of medical students on artificial intelligence: results report for biomedical research course in College of Medicine, Pusan National University. Busan: Pusan National University; 2018.
  125. Minor LB. The rise of the data-driven physician [Internet]. Stanford (CA): Stanford Medicine; 2020 [cited 2020 Jan 10]. Available from: <http://med.stanford.edu/content/dam/sm/school/documents/Health-Trends-Report/Stanford%20Medicine%20Health%20Trends%20Report%202020.pdf>.
  126. Murphy B. AMA: take extra care when applying AI in medical education [Internet]. Chicago (IL): American Medical Association; 2019 [cited 2020 Jan 10]. Available from: <https://www.ama-assn.org/practice-management/digital/ama-take-extra-care-when-applying-ai-medical-education>.
  127. NHS England. The Topol Review [Internet]. Leeds: NHS England [cited 2020 Jan 10]. Available from: <https://topol.hee.nhs.uk/the-topol-review/>.
  128. Minor L. Tomorrow's doctors seek training in data science, but will that be enough? [Internet]. Chicago (IL): Modern Healthcare; 2020 [cited 2020 Jan 10]. Available from: <https://www.modernhealthcare.com/opinion-editorial/tomorrows-doctors-seek-training-data-science-will-be-enough>.
  129. American Medical Association. Innovations & outcomes of the consortium [Internet]. Chicago (IL): American Medical Association; [cited 2020 Jan 10]. Available from: <https://www.ama-assn.org/education/accelerating-change-medical-education/innovations-outcomes-consortium>.
  130. Murphy B. 4 Phases to making goal of lifelong physician learner a reality [Internet]. Chicago (IL): American Medical Association; 2017 [cited 2020 Jan 10]. Available from: <https://www.ama-assn.org/education/accelerating-change-medical-education/4-phases-making-goal-lifelong-physician-learner>.
  131. Cutrer WB, Miller B, Pusic MV, Mejicano G, Mangrulkar RS, Gruppen LD, et al. Fostering the development of master adaptive learners: a conceptual model to guide skill acquisition in medical education. *Acad Med.* 2017;92(1):70-5.
  132. Harvard Medical School. The future is now: medical education for the 21st century [Internet]. San Bruno (CA): Youtube; 2015 [cited 2020 Jan 10]. Available from: <https://www.youtube.com/watch?v=MRC9i85R2sY>.
  133. Pathways course info [Internet]. Boston (MA): Harvard Medical School; 2019 [cited 2020 Jan 10]. Available from: <https://meded.hms.harvard.edu/pathways-course-descriptions>.
  134. Johns Hopkins Medicine. Genes to society: a curriculum for the Johns Hopkins University School of Medicine: design and structure of the genes to society curriculum [Internet]. Baltimore (MD): Johns Hopkins Medicine [cited 2020 Jan 10]. Available from: [https://www.hopkinsmedicine.org/som/curriculum/genes\\_to\\_society/curriculum-overview.html](https://www.hopkinsmedicine.org/som/curriculum/genes_to_society/curriculum-overview.html).
  135. Johns Hopkins Medicine. Genes to society: a curriculum for the Johns Hopkins University School of Medicine: translational science intersessions [Internet]. Baltimore (MD): Johns Hopkins Medicine [cited 2020 Jan 10]. Available from: [https://www.hopkinsmedicine.org/som/curriculum/genes\\_to\\_society/year-three/translational-science-intersessions.html](https://www.hopkinsmedicine.org/som/curriculum/genes_to_society/year-three/translational-science-intersessions.html).
  136. Hosamani P, Osterberg L, Shafer A. INDE 297: reflections, research & advances in patient care [Internet]. Stanford (CA): Stanford Medicine; 2020 [cited 2020 Jan 10]. Available from: <http://med.stanford.edu/md/discovery-curriculum/clerkships/reflections.html>.
  137. Medical Education, The University of California San Francisco. Inquiry immersion [Internet]. San Francisco (CA): The University of California San Francisco [cited 2020 Jan 10]. Available from: <https://meded.ucsf>.

- edu/md-program/current-students/curriculum/foundations-1/inquiry-immersion#MiniCourse-Elements.
138. American Medical Association. Learning analytics for training in the workplace [Internet]. San Bruno (CA): Youtube; 2015 [cited 2020 Jan 10]. Available from: <https://www.youtube.com/watch?v=ykQ8qvlvtzM&feature=youtu.be>.
  139. Medical Education, The University of California San Francisco. Core inquiry curriculum [Internet]. San Francisco (CA): The University of California San Francisco; c2019 [cited 2020 Jan 10]. Available from: <https://meded.ucsf.edu/md-program/current-students/curriculum/foundations-1/core-inquiry-curriculum>.
  140. Medical Education, The University of California San Francisco. Inquiry immersion [Internet]. San Francisco (CA): The University of California San Francisco; c2019 [cited 2020 Jan 10]. Available from: <https://meded.ucsf.edu/md-program/current-students/curriculum/foundations-1/inquiry-immersion#MiniCourse-Elements>.
  141. Medical Education, The University of California San Francisco. Deep explore [Internet]. San Francisco (CA): The University of California San Francisco; c2019 [cited 2020 Jan 10]. Available from: <https://meded.ucsf.edu/md-program/current-students/curriculum/career-launch/deep-explore>.
  142. Kelleher JD, Tierney B. What is data science? [Internet]. Cambridge (MA): MIT Press; 2018 [cited 2020 Jan 10]. Available from: <https://ieeexplore.ieee.org/document/8544248>.
  143. Kolachalama VB, Garg PS. Machine learning and medical education. NPJ Digit Med. 2018;1:54.
  144. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. JMIR Med Educ. 2019;5(2):e16048.
  145. Kolachalama VB. Teaching: MS650: machine learning for biomedical applications [Internet]. Boston (MA): Boston University [cited 2020 Jan 10]. Available from: <http://sites.bu.edu/vkola/teaching/>.
  146. Harvard Medical School. AISC 610: computationally-enabled medicine [Internet]. Boston (MA): Harvard Medical School [cited 2020 Jan 10]. Available from: <https://dbmi.hms.harvard.edu/education/courses/aisc-610>.
  147. Kilic A. Artificial intelligence and machine learning in healthcare mini-elective: Spring 2020 [Internet]. Pittsburgh (PA): University of Pittsburgh School of Medicine; 2020 [cited 2020 Jan 10]. Available from: [https://www.omed.pitt.edu/sites/default/files/artificial\\_intelligence\\_and\\_machine\\_learning\\_in\\_healthcare.pdf](https://www.omed.pitt.edu/sites/default/files/artificial_intelligence_and_machine_learning_in_healthcare.pdf).
  148. Stanford University. Precision practice with big data BIOMEDIN 205: precision medicine and big data [Internet]. Stanford (CA): Stanford University; 2019 [cited 2020 Jan 10]. Available from: <https://canvas.stanford.edu/courses/106497>.
  149. Park SH, Do KH, Kim S, Park JH, Lim YS. What should medical students know about artificial intelligence in medicine? J Educ Eval Health Prof. 2019;16:18.
  150. Topol E. Deep medicine: how artificial intelligence can make healthcare human again. New York (NY): Basic Books; 2019.
  151. Korean Institute of Medical Education and Evaluation. Post 2nd period medical education accreditation standards and regulations. Seoul: Korean Institute of Medical Education and Evaluation; 2010.
  152. Korean Medical Association. Korean doctor's role. Seoul: Korean Medical Association; 2014.
  153. Korean Institute of Medical Education and Evaluation. Accreditation Standards of KIMEE 2019. Seoul: Korean Institute of Medical Education and Evaluation; 2019.
  154. Seo JH, Kim HJ, Kim BJ, Lee SJ, Bae HO. Educational and relational stressors associated with burnout in Korean medical students. Psychiatry Investig. 2015;12(4):451-8.
  155. Dyrbye LN, Massie FS Jr, Eacker A, Harper W, Power D, Durning SJ, et al. Relationship between burnout and professional conduct and attitudes among US medical students. JAMA. 2010;304(11):1173-80.
  156. Samra R. Empathy and burnout in medicine-acknowledging risks and opportunities. J Gen Intern Med. 2018;33(7):991-3.
  157. NHS England. Digital fellowships [Internet]. Leeds: NHS England; 2019 [cited 2020 Jan 10]. Available from: <https://topol.hee.nhs.uk/digital-fellowships/>.
  158. Boyko O, Chang A. American Board of Artificial Intelligence in Medicine (ABAIM) aims to educate and certify healthcare professionals in AI, and related technologies [Internet]. Beltsville (MD): CISION PRWeb; 2020 [cited 2020 Mar 10]. Available from: [https://www.prweb.com/releases/american\\_board\\_of\\_artificial\\_intelligence\\_in\\_medicine\\_abaim\\_aims\\_to\\_educate\\_and\\_certify\\_healthcare\\_professionals\\_in\\_ai\\_and\\_related\\_technologies/prweb16963954.htm](https://www.prweb.com/releases/american_board_of_artificial_intelligence_in_medicine_abaim_aims_to_educate_and_certify_healthcare_professionals_in_ai_and_related_technologies/prweb16963954.htm).