

의료 AI의 안착 과정

자동 심전도 판독(Automated ECG

interpretation)의 발전 역사를 통해 -2부

지난 3월호 A리포트에서는 ‘자동 심전도 판독’ 기술의 발전 역사와 관련된 AI 내용을 다루었다.¹ 두 번째 시간으로 의학 연구에서 ‘내적 타당도’를 평가할 때 왜 설명력이 중요한지를 임상 의사 결정 과정의 특징을 바탕으로 살펴보도록 하겠다. 그리고 상기 심전도의 예제처럼 신의료기술이 내적 타당도를 지속적으로 높이면서 외적 타당도를 확보하려면 어떤 연구 환경이 필요한지에 대해서도 살펴보고자 한다.

글 | 정세영 syjung@snuh.org

의료 인공지능에 대한 대중들의 관심이 높아지고 있습니다. 의료 자원의 낭비를 막고 치료의 질을 높일 수 있는 방법에 대해 인공지능 기술 가운데서 찾고 있는 연구자도 급속하게 늘어나고 있습니다. 그렇다면 그동안 의료 분야에서는 인공지능 기술이 전혀 쓰이지 않았을까요? 역사적인 실례를 통해 인공지능 의료기기 같은 신의료기술이 태동하고 시장에 안착되기까지 어떤 과정을 거치는지 살펴봄으로써 의료 인공지능 연구자들이 놓치지 말아야 할 교훈에 대해 이야기 나누고 싶습니다.

다시 응급실 스냅샷으로 돌아가 보자. 다음은 쉽게 찾아볼 수 있는 신문 기사다.

<p>통증 없이 찾아올 수 있는 심근경색²</p> <p>“가슴통증 없었는데...” 방심하다 심근경색 ‘골든타임(golden time)’ 놓친다</p> <p>고령의 당뇨환자나 여성일수록 목·팔 아프거나 속 더부룩할 때 심장 의심 알고 치료시기 놓쳐...</p> <p>‘가슴통증=심장병’ 잘못된 상식 근육통·늑막염 등 증상일 수도...</p> <p>많은 사람들은 ‘돌연사의 주범’인 급성 심근경색 등 심장병이 생기면 당연히 가슴이 아플 것으로 생각한다. 그러나 치명적인 급성 심근경색임에도 불구하고 가슴 통증이 나타나지 않는 경우가 적지 않다. 이로 인해 소화불량이나 콧병 등과 같은 다른 질환으로 잘못 알아 엉뚱한 치료로 시간을 끌다 생명을 구할 수 있는 ‘골든 타임’을 놓치는 경우도 있다.</p> <p>급성 심장병 중 25%는 가슴 통증 없어...</p> <p>급성 심근경색은 심장에 혈액과 영양분을 공급하는 관상동맥이 혈전에 의해 갑자기 막혀 심장근육이 괴사하는 상태를 말한다. 이렇게 되면 심장이 멎어 병원 응급실에 후송되기 전 사망할 확률이 30%, 응급실 도착 직후 사망할 확률이 10% 정도 되는 무서운 질환이다.</p> <p>따라서 급성 심근경색이 발생하면 5분 이내 심장마사지 등 응급처치를 해야 한다. 관상동맥이 예기치 않게 막히는 것은 죽상(粥狀) 동맥경화가 주 원인이다. 죽상 동맥경화는 혈관의 가장 안쪽을 덮고 있는 내막에 콜레스테롤 침착 등의 이유로 죽처럼 묽은 ‘죽종’이 생기는 증상이다. 동맥경화는 혈관의 노화 외에도 고혈압, 고지혈증, 동물성 지방 위주의 식습관, 흡연 등에 의해 생긴다.</p> <p>그런데 이런 급성 심근경색 환자 가운데 ‘빠르하다’, ‘체한 것처럼 답답하다’, ‘고춧가루를 뿌린 것 같다’ 는 등의 전형적인 가슴 통증이 나타나지 않는 경우가 4명 중 1명꼴이 나 된다.</p>
--

만약에 응급실에 방문했던 환자에게 심전도 검사를 하지 않았다면 어떤 결과가 벌어졌을까? 아마 환자는 소화제만 처방받아 귀가했을 것이고, 집에 가는 중에 쓰러졌을 수도 있다. 그렇다면 흉통이 없는 환자에게 왜 심전도 검사를 시행했을까? 위 신문 기사의 내용처럼 흉통이 없어도 심근경색 같은 관상동맥질환이 발생할 수 있기 때문이다. 그리고 응급실을 내원하는 환자는 동네 의원을 내원하는 환자들에 비해 관상동맥질환의 발병률이 높다는 사실을 경험적으로 알기 때문이다. 담당의는 응급실 표준 진료절차를 잘 지켰고, 그 덕분에 심근경색을 발견할 수 있었다. 하지만 환자가 살 수 있었던 가장 중요한 이유는 자동 심전도 판독 기술이 임상 의사 결정 과정의 일부로 자연스럽게 편입되었다는 점이다.

임상 의사 결정 과정의 특징

의학 연구나 의료 인공지능의 내적 타당도를 평가할 때 설명력이 중요한 이유는 임상 의사 결정 과정 자체가 설명력에 바탕을 둔 베이즈 추론(Bayesian inference) 과정이기 때문이다.

임상 의사 결정 과정에 잘못된 증거가 통합되면 벌어질 수 있는 일을 에이즈 진단의 예로 살펴보자.

<p>바이러스 감염 여부 검사 결과의 오류 발생 확률³</p> <p>현혈자 혈액에서 HIV(Human Immunodeficiency Virus, 에이즈 바이러스)가 있는지 선별 검사를 하고 있다. 일반 인구 집단에서는 아주 작은 확률(0.1%)로 HIV가 있을 가능성이 있다. 항체검사는 매우 정확하지만 완벽하지는 않다. 99%의 확률로 감염된 혈액을 잡아내지만 1%는 감연되지 않은 혈액을 감염된 것으로 잘못된 결론을 내릴 수 있다. 검사 결과 HIV 양성으로 나왔다면, 실제 현혈자가 HIV를 갖고 있을 확률과 검사 결과에 오류가 있을 확률(위양성)은 얼마나 되는가?</p>

계속 읽기 전에 정답이 무엇일까 생각해보자.

실제로 에이즈 바이러스 감염을 진단하기 위해서는 두 번의 선별 검사와 한 번의 확진 검사를 시행한다. 세 차례 검사를 시행하였을 때 진단의 민감도와 특이도는 각각 99%로 알려져 있다. 즉, 실제 HIV 감염자 중 양성 결과(“HIV 바이러스에 감염됨”)가 나올 확률과 실제 HIV 비감염자 중 음성 결과(“HIV 바이러스 없음”)가 나올 확률이 각각 99%라는 뜻이다. 이것이 과연 어떤 의미일까? 정답을 알아보기 전에 다른 예를 살펴보자.

<p>불확실한 판정에 대한 과잉 확신 사례 #1</p> <p>검사에서 양성인 나왔다는 것은 당신의 혈액 속에서 HIV 항체가 발견됐다는 것을 말한다. 이는 당신이 HIV에 감염됐다는 사실을 말해준다. 당신은 평생 감염돼 있을 것이며 다른 사람들에게 HIV를 전염시킬 수도 있다.</p> <p>- 일리노이 주 공중보건국</p>
<p>불확실한 판정에 대한 과잉 확신 사례 #2⁴</p> <p>1990년 11월 어느 날 베티의 전화기가 울렸다. 그녀는 플로리다에 사는 45세 여성으로 세 명의 10대 아들을 둔 어머니였다. 아이들의 아버지는 오래전에 죽었다. 그녀는 인근 병원에서 갑상선 질환 검사와 혈액 표본 검사를 받았는데, 전화는 그 병원에서 걸려온 것이었다. 병원을 방문한 베티는 의사로부터 에이즈에 걸렸다는 이야기를 들었다. 의사는 그녀가 얼마나 더 살 수 있는지 확실히 답하지 못했다. 의사의 말이 이어지는 동안 그녀는 병에 대해 생각하지 않으려고 텔레비전을 계속 쳐다봤다. 하지만 밤이 되자 이런 생각이 들었다. “내가 죽으면 어떻게 될까? 아이들은 누가 돌보지? 사람들이 아이들을 대해 어떻게 대할까?”</p> <p>1992년 의사는 베티에게 디다노신(didanosine)을 처방했다. 이 약은 항 HIV 약으로 구역질, 피로 같은 여러 부작용을 유발한다. 그런데 어느 날 베티가 지역 에이즈 환우회에 참석했을 때 상담가는 베티의 T-세포 수치가 지속적으로 높게 남아 있다는 것을 알아차렸다. 상담가는 베티에게 재검사를 제안했다. 1992년 12월 어느 날, 다시 베티의 전화기가 울렸다. 베티는 병원에 방문해야 하느냐고 물어보았다. 상담소 직원은 이렇게 말했다. “뭐라고 말씀드려야 할지... HIV 검사가 음성으로 나왔습니다.”</p> <p>베티는 그를 진료한 의사와 상담소 및 플로리다 주 보건 재활국(첫 검사를 수행한 기관이었다)을 상대로 소송을 걸었다. 배심원단은 2년 동안의 고통과 시련에 60만 달러의 보상금이 적절하다는 판결을 내렸다.</p>

어떤 질환에 대한 검사 결과가 양성으로 나왔다는 것이 과연 어떤 의미를 가질까? 검사 결과 그 병이 있다는 의미일까?

다시 처음 질문으로 돌아가 보자. 질문에 대한 명확한 답을 찾기 위해 아래 글과 같이 생각해보자. 글의 내용을 이해하려면 먼저 일반 인구의 HIV 감염 유병률을 알아야 한다. 우리나라는 그 숫자가 0.1%로 알려져 있다. 즉, 10만 명의 사람이 있을 때 100명 정도가 HIV 바이러스를 갖고 있다는 뜻이다. 그 확률(유병률)은 정맥주사 사용자일 경우 10%로 올라가게 된다. 정맥주사 사용자 10만 명 중 1만 명이 HIV 바이러스를 갖게 된다. 여기서 정맥주사 사용자란 안전하지 않은 방법으로 각종 주사 약물(주로 중독성 약물, 마약)을 남용하는 사람을 의미한다.

‘양성 판정’이 실제로 뜻하는 바는⁵

10만 명이 검사를 받았다고 가정하자. 그 중 100명(0.1%)이 HIV 감염자일 것이고, 99명(99%)은 검사 결과 양성일 것이다. 나머지 9만 9,900명에게는 HIV가 없겠지만 검사 결과 1%는 HIV가 있다고 잘못된 결과가 나올 것이다. 따라서 999명은 위양성결과를 보이게 된다. 정리하자면 99+999=1,098명에서 양성이 나오는데, 그 중 99/1,098=9%만 진짜 양성일 것이다. 나머지 91%는 위양성이다. 따라서 검사 결과가 양성이라도 9%만이 실제 HIV를 가진 것이 된다.

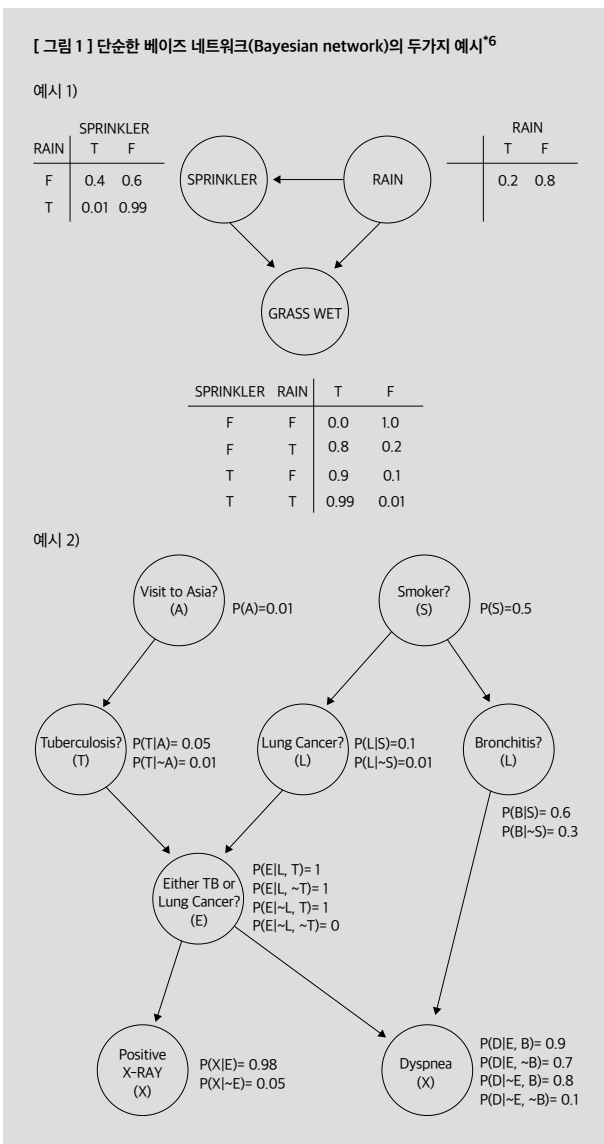
같은 예를 정맥주사 사용자에게 적용해 보자. 정맥주사를 사용할 경우 HIV 유병률이 10%다. 다시 10만 명이 검사를 받는다고 가정하자. 그 중 1만 명(10%)이 실제 HIV를 갖고 있을 것이고 9,900명(99%)은 검사 결과가 양성으로 나올 것이다. 나머지 9만 명은 HIV를 실제 갖고 있지 않지만 1%인 900명에서 HIV 검사 결과가 양성으로 나올 것이다. 총합하면 9,900+900=10,800명은 양성이 나올 것이고 9,900/10,800=92%의 진양성결과를 보일 것이다. 나머지 8%의 양성결과는 위양성이다. 검사 결과가 양성이라면 92%의 확률로 HIV를 갖고 있을 가능성이 있는 것이다.

따라서 첫 번째 질문의 정답은 헌혈자가 HIV 바이러스를 실제 갖고 있을 확률과 검사에 오류가 생겨서 양성 결과가 나올 확률은 각각 9%와 91%이다. HIV가 감염될 위험 행동을 전혀 하지 않은 환자는 확진 검사까지 받아도 그 검사의 정확성은 9%에 지나지 않으며 이는 동전 던지기 확률(50%)에도 한참 미치지 못한다.

민감도와 특이도는 진단 방법의 성능이다. 이진 분류(binary classification) 모형의 성능지표로 많이 활용되는 ROC(receiver operating characteristic, 수신기 운영 특성) 곡선은 민감도와 특이도의 교환 관계를 나타낸 그래프이며, 이 또한 그 모형의 성능을 의미한다. 양성예측도와 음성예측도는 특정 질환이 있거나 없을 사전확률에 진단 방법의 성능을 결합한 지표이다.

어떤 질병에 대한 검사가 양성으로 나왔다고 가정하자. 이 때 실제로 그 병을 가지고 있을 확률은 얼마나 될까? 즉, 그 결과의 양성예측도는 얼마나 되는가? 앞서 얘기했듯이 양성예측도를 계산하려면 질환의 유병률(사전확률)을 알아야 한다. 위의 HIV 진단 예제에서는 일반 인구와 정맥주사 사용자의 HIV 감염 사전확률(유병률)은 100배 차이가 난다. 그에 따라 양성예측률도 10배 차이가 나게 된다.

임상 의사 결정 과정은 교과서적 지식, 의사의 개별 경험, 검사(혈액, 영상, 병리 등), 신체 진찰, 병력 청취로 얻은 정보를 결합하여 최종 결정을 내리는 베イズ 추론 과정이다. 질환의 유병률을 사전확률로 결합하여 사후확률을 도출하는 앞의 예제는 베イズ 추론의 정말 단순한 예다. 각각의 증거는 하나의 노드를 구성하게 되며 각각의 노드는 네트워크 형태로 얽혀 있다. 임상 의사는 진료 현장에서 한정된 시간동안 얻을 수 있는 모든 정보를 동원하여 머릿속에서 베イズ 추론 과정을 거쳐 최종 의사 결정을 하게 된다. 하나의 검사 결과는 하나의 증거로서 하나의 노드를 구성하게 되며, 각 노드간의 이동확률은 조건부 확률로써 규정된다. 베イズ 추론 과정을 구성하는 근거(노드)들이 명시적으로 드러날 때 그 추론 과정은 설득력을 얻게 되고 그에 따른 최종 결정에 정당성이 부여된다. 임상 의사 결정 과정 자체가 통합되는 증거들의 ‘설명력’ 없이는 이루어질 수 없는 것이다. 임상 의사 결정 과정의 중간에 모호한 증거가 결합되면 최종 결정은 그만큼 힘을 잃게 된다. 위에서 언급한 베티의 예처럼 검사의 성능(민감도, 특이도)에 대한 충분한 이해 없이 잘못된 증거를 의사 결정 과정에 통합하면 의료진은 환자에게 치명적인 결과를 안겨주는 잘못된 결정을 할 수도 있다. 의료 인공지능이 도출한 결과를 임상 의사 결정 과정에 통합할 때도 마찬가지로 과정을 거칠 수 밖에 없다. ‘인공지능’으로 도출한 결과라도 100% 정확한 모형은 없기 때문이다.



예시 1)은 강우와 스프링클러 작동 여부가 잔디를 젖게 하는 확률을 표현하였다. 예시 2)는 폐결핵, 폐암, 기관지염 발생 확률을 선행확률로 놓고, 흉부 X-ray에서 이상 소견이 발견될 확률과 환자가 호흡곤란(dyspnea)을 호소할 확률을 표현하였다.

어떤 단일 인공지능 알고리즘의 결과만 보고 최종 판단을 내리는 상황을 가정해보자. 그 알고리즘에는 고려해야 할 모든 변수가 포함되어 있는가?(예컨대, HIV 감염진단의 예처럼 HIV 감염 유병률에 대한 정보도 포함되었는가?) 그 전에 ‘고려해야 할 모든 변수’를 아는게 가능이나 할까?

어떠한 단일 검사 하나만 보고 최종 진단을 내리게 되면 오류를 범할 가능성이 높아진다. 예컨데, 2번의 선별검사와 1번의 확진검사를 통한 HIV 검사의 진단의 민감도와 특이도가 99%라고 하여 위의 예제처럼 선불리 HIV 감염판정을 내리면 안된다. 유병률이 낮은 질환일수록 더욱 조심해야 한다. 그 환자가 HIV에

감염되었을 만한 증거를 더 모아야 한다. 동성 연애자인지, 위험성 행위를 하지는 않았는지, 혹은 정맥주사 사용자인지 등을 의심해 보아야 한다. 급성 심근경색 환자의 예도 마찬가지다. 그 환자를 보았던 인턴 의사는 머릿속에서 환자의 병력과 신체 진찰 결과를 통합하면서 직관적으로 심장질환에 대한 사후확률을 떨어뜨렸을 것이다. 하지만, 자동 심전도 판독의 결과가 베イズ 네트워크에 결합되어 추론되었을 때 급성 심근경색의 확진에 이를 만한 강력한 사후확률이 도출되었고, 그 덕에 환자는 살 수 있었다.

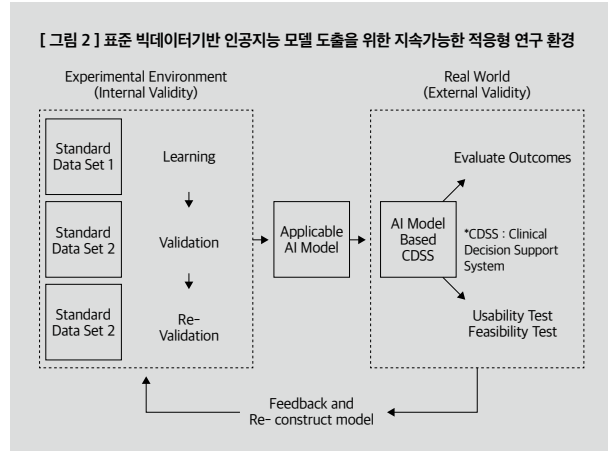
인공지능 알고리즘도 마찬가지다. 단일 알고리즘의 결과만 보고 의사 결정을 내리는 것은 사실 행위나 마찬가지다. 따라서 인공지능 알고리즘의 성능만 보고 ‘AI의 임상 적용이 카운트다운에 들어갔다’고 단정하는 것 또한 시기상조다. 의료 인공지능에 대해 지금까지 발표된 숭한 보고서들은 고작해야 ‘어떤 약물이 배양접시에서 병원균을 죽이는 것으로 확인되었다’ 정도의 의학 논문과 유사한 수준으로 봐야 한다. 그와 같은 연구 결과는 매우 흥미롭지만, 통상적인 과학 절차에 따르면 (1)연구에 사용된 방법과 재료를 상세히 기술하고, (2)후속연구에서 재현된 다음, (3)일련의 연구를 거쳐 최종적으로 대규모 임상시험을 통해 효과를 검증하는 단계를 밟아야 한다.⁷

임상 의사 결정의 바탕이 되는 베イズ 추론 과정의 본질은 무엇인가? 주어진 사전확률에 특정 정보를 결합하여 사후확률을 증대(augmentation)시키는 것이다. 그리고 의료 인공지능의 역할은 베イズ 네트워크에 결합되어 최종 확률의 정확성을 높이는 강력한 증거를 창출해내는 것이다. 이것이 바로 인공지능과 의사가 협업하여 의료의 효율성과 정확성을 증대시키는 과정이며 의료 인공지능이 자리매김 할 수 있는 형태다. 또 의료 인공지능에 대한 임상시험도 그 증대과정의 효과를 객관적으로 입증하는 과정이어야 한다.

임상 의사 결정에서 설명력이 중요한 또 다른 이유는 의료 행위가 단순히 ‘예측’을 하는 것이 아니라 ‘실행’이 필요한 영역이기 때문이다. 예를 들어, 어떤 환자가 당뇨병이 걸릴 확률이 인공지능 알고리즘에 의해 70%가 나왔다고 하자. 그럼 그 다음 그 환자는 무엇을 해야 하는가? 환자와 의사가 알고 싶은 것은 그 ‘무엇’에 대한 것이다. 의사는 환자가 개선할 수 있는 요인에 대해 권면해야 할 것이다. 비만과 좋지 않은 식습관이 당뇨병 발병확률을 높인 이유라면 환자에게 식이요법과 운동을 권할 수 있을 것이다. 환자 또한 마찬가지다. 어떤 병에 걸릴 가능성이 높은 요인에 대한 설명을 들어야 약을 먹거나 생활습관 개선의 노력을 할 수 있다.

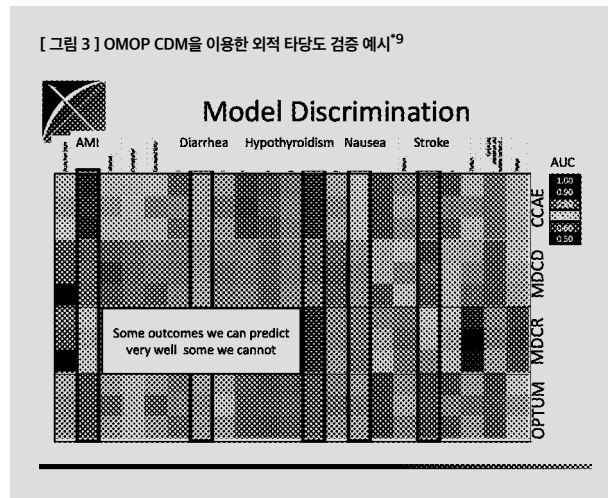
당뇨병 예측 알고리즘의 성능이 아무리 좋아도, 당뇨병 발병 확률을 계산하는데 기여했던 변수의 종류와 기여 정도를 정량적으로 알 수 없으면 실제 임상에서는 해당 모형을 활용하기 어렵게 된다.⁸

의료 인공지능의 외적 타당도를 높이기 위한 연구 환경



외적 타당도 검증에는 두 가지 단계가 있다. 다수의 표준 데이터로 모형을 검증하여 과적합을 방지하는 과정과 현실 세계의 사용례를 지속적으로 발굴하고 검토하여 모형의 유용성을 검증하는 과정이다.

첫 번째 과정을 위해서는 연구자가 활용할 수 있는 다수의 표준 데이터 셋을 마련해야 한다. 아울러 도출한 모델을 검증, 재검증 할 수 있는 데이터 셋이 있어야 하며, 각각의 데이터 셋은 가급적 특정 인구 집단의 편향된 정보를 담고 있지 않아야 한다.



각 열(column)은 우울증 치료를 받은 사람에게서 발생할 수 있는 22개의 질환을 의미하며, OPTUM, MDCR, MDCC, CCAE는 OMOP CDM(common data model)¹⁰⁾로 변환된 데이터 셋이다. 세 가지 방법(gradient boosting, random forest, regularized regression)을 사용하여 우울증 치료 환자의 특정 질환 발병 확률 예측 모형의 성능(area under the curve, AUC)을 비교하였다.

그림에서 보는 것처럼 각 데이터 셋마다 알고리즘의 성능이 다르게 나오는 점을 알 수 있다. 이처럼 하나의 데이터 셋에 편향된 모형을 배제하기 위해서는 개발된 인공지능 모형의 외적 타당성(external validation) 검증은 필수인 것이다.

두 번째 과정을 위해서는 인공지능 모형을 해당 분야 전문가의 철저한 모니터링 가운데 검증할 수 있는 체계가 마련되어야 한다. 앞에서 기술했듯이 임상의학의 의사 결정 특성을 고려할 때 어떤 인공지능 모델도 의료 행위로 바로 이어지는 최종 결과를 도출하는 역할을 수행할 수는 없다. 임상적 의사 결정은 정량화 할 수 있는 각종 검사 및 연구 결과와 정량화 할 수 없는 의사의 직관이나 경험이 통합되는 베이스 과정(Bayesian process)을 거치게 된다. 이 때 인공지능 모델은 선행 확률(pre-probability)로써 의사가 내린 최종 판단의 정확성을 높여주는 역할을 하게 된다. 다른 말로 표현하면 임상 의사 결정 지원 도구(clinical decision support system, CDSS)로써 기능을 하게 된다. 이런 인공지능 모델을 활용한 의료기기나 기술은 의료 현장에 적용했을 때의 효과와 발생할 수 있는 부작용을 전문가 기구에 의해 지속적으로 평가받아야 한다(outcome analysis). 아울러 기존의 해당 인공지능 기술이 임상 의사 결정 과정에 통합되었을 때의 사용성(usability)과 적용가능성(feasibility)에 대해 평가를 받아야 한다. 그리고 이런 피드백을 바탕으로 인공지능 모델을 고도화하는 과정을 다시 거쳐야 하는 것이다.

마치며

현재 기술의 진화 속도는 자동 심전도 판독이 기술적으로 발전한 속도에 비교할 수 없을 만큼 빠르다. 알파고(AlphaGo)가 등장한 후 18개월 만에 알파고 제로(AlphaGo Zero)가 나왔듯이, 현재 우리가 갖고 있는 컴퓨터 계산 능력과 인프라 수준을 고려하면 전광석화의 속도로 진화하는 기술을 현장에 적용하고 유용성을 증명할 수 있는 범·제도적 기반¹¹⁾이 절실히 필요한 상황이다.

의료에서 AI 효과(AI effect) 또는 odd paradox¹²⁾라고 불리는 말의 가장 극명한 예가 바로 자동 심전도 판독이다. 의료인들이 인공지능 기술이라고 인식하지 못하는 가운데 임상 현장에서 활발하게 사용되며 그 유용성과 한계가 명확하게 알려져 있다. 그 발전 과정을 보면 제품의 성능을 올리는 노력과 함께 임상 현장에 실제 적용되었을 때 유용성을 치열하게 평가하는 과정이 병행되었다. 그리고 이제는 ‘신의료기기’가 아닌 효용성을 충분히 검증 받은 ‘의료기기’로 역할을 하고 있다. 심전도는 70년의 기나긴 발전 과정을 거쳤지만, 지금 우리는 1~2년 사이에도 전대미문의 신의료기기가 출시되는 세상에 살고 있다. 이런 기술의 발전 속도를 따라가면서 보건 의료 분야의 4차 혁명을 주도하기 위해서 정부는 연구자나 개발자가 제품의 외적 타당도를 평가할 수 있는 지속 가능한 연구 환경을 마련해야 할 것이다. 또 연구자나 개발자는 연구 결과나 출시된 제품의 내적 타당도를 평가하는 합리적인 기준을

마련해야 하며, 제품의 성능을 높이기 위해 지속적으로 치열한 노력을 해야 한다. 뿐만 아니라 의료 인공지능의 기능을 과장하지 말고, 의료 인공지능이 임상 의사 결정 과정에 자연스럽게 통합되는 사용 사례를 지속적으로 발굴해야 할 것이다.

¹⁾ 참고 | 시리포트 3월호 <의료 AI의 안착 과정> <https://brunch.co.kr/@kakao-it/220> ²⁾ 참고 | 한국일보. <http://www.hankookilbo.com/v/81b6a41a14ee404682fb24cb2923b8f4> ³⁾ 참고 | 정세영 역, 직판으로 이해하는 의학통계학, 대한의학서적 p.7 ⁴⁾ 참고 | 케르트 기가렌저, 황승식 역, 숫자에 속아 위험한 선택을 하는 사람들, 살림, P.159 ⁵⁾ 참고 | 정세영 역, 직판으로 이해하는 의학통계학, 대한의학서적 p. 8 ⁶⁾ 참고 | Wikipedia "Bayesian network" 웹사이트 <http://sujiptal.blogspot.kr/2013/07/bayesian-network-inference-with-r-and.html> ⁷⁾ 참고 | <https://www.nature.com/articles/d41586-018-03067-x> ⁸⁾ 참고 | <https://www.wired.com/story/can-machine-learning-find-medical-meaning-in-a-mess-of-genes/> ⁹⁾ 참고 | Peter R. Rijnbeek, Journey toward Patient-Level Prediction. OHDSI website. <https://www.ohdsi.org/wp-content/uploads/2016/09/OHDSI-Symposium-2016-Rijnbeek-prediction-23sept2016.pdf> ¹⁰⁾ 설명 | OPTUM, MDCR(Medicare Supplemental), MDCC(Medicaid Supplemental), CCAE(Commercial Claims and Encounters), OMOP(Observational Medical Outcomes Partnership) ¹¹⁾ 참고 | 헬스케어 빅데이터 딜레마와 해결방안. <https://brunch.co.kr/@kakao-it/221> ¹²⁾ 설명 | AI-effect (Odd paradox): 인공지능 기술이 일반 대중에게 익숙해지면 그 기술은 더이상 인공지능으로 간주되지 않고 더 새로운 기술을 '인공지능'이란 이름으로 찾게 되는 현상을 말한다. 인공지능 기술이라 하면 보통 거창한 것을 상상하지만 실상은 우리 생활 곳곳에 인공지능 기술이 이미 사용되고 있다.