

EMR과 디지털 뉴딜 : 데이터 댐에서 퍼 올릴 때 알아야 할 EMR의 네 가지 문맥적 특징

차원철 조교수

성균관대학교 삼성융합의과학원
디지털헬스학과



1. 들어가는 글

1) 2020년 4월 이후 '디지털 뉴딜'에 대한 관심 증가

'데이터 댐' 사업은 '데이터·네트워크·인공지능' 생태계 강화 차원에서 공공데이터 14만 개를 공개해 일종의 '댐'을 구축하는 사업으로 디지털 뉴딜 사업의 중추적 역할이 기대된다. 전자의무기록(Electronic Medical Record: EMR) 데이터에 특별한 관심이 쏠리고 있는데, 아마도 데이터 연계로 인한 잠재적 가치가 크기 때문일 것이다. EMR 데이터만 활용할 수 있다면 장밋빛 미래가 펼쳐지며 의료 및 관련 산업의 비약적 발전이 일어날 수 있을 것으로 그려지고 있다.

2) EMR 데이터 활용의 현실

현재 EMR을 이용한 많은 연구는 EMR 이전의 연구들과 크게 다르지 않은 형태로 차트리뷰로 진행하던 연구를 EMR에 맞추어 진행하는 방식이다. 자료 추출의 용이성으로 데이터의 크기와 범위가 늘어나고 있는 추세이나, 질병이나 진료과의 틀을 넘는 통합적인 SI 구축이나 외부 데이터와의 연계 등, 빅데이터 분석을 통한 혁신을 처음의 기대보다 드물게 보고되고 있다.

3) EMR 데이터 활용이 어려운 이유

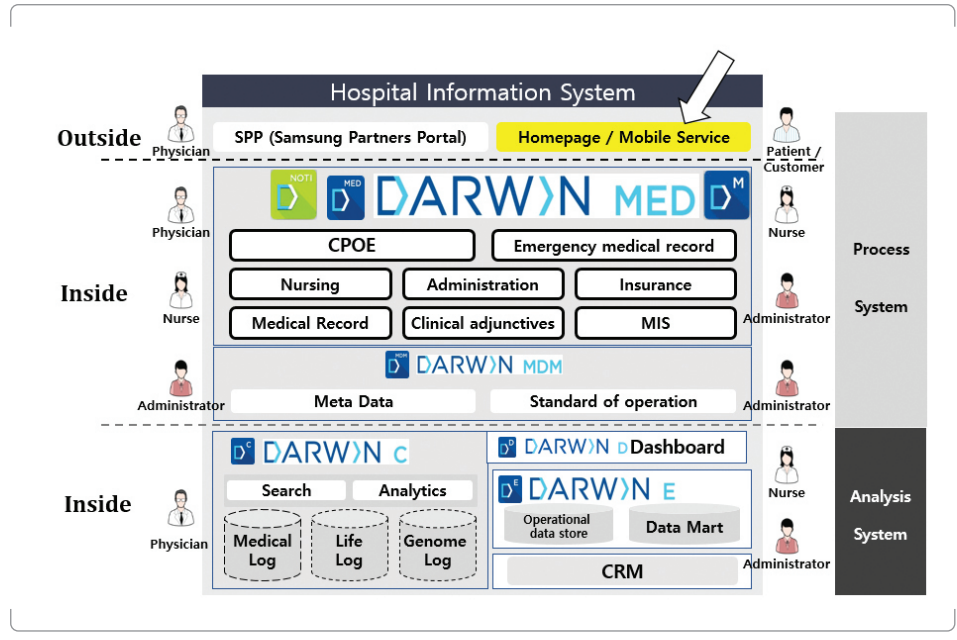
정보보안 및 개인정보 보호와 관련된 규제들은 EMR데이터의 활용을 가로막는 걸림돌로 지목되고 있으며 '데이터3법', '규제 샌드박스' 등을 통해 점차 해소될 것으로 기대된다. 규제 해소의 분위기와는 달리 실질적이며 본격적인 EMR 활용은 아직 초보단계로 보이며 산업계와의 협업도 탄력을 받고

있지는 못하다. 이는 EMR 데이터의 생성과 관련된 원초적 특징, 그 발생의 문맥과 관련이 있다 할 수 있다.

2. EMR 데이터의 문맥적 특징

2.1. EMR은 매우 다양한 주체 때문에 사용되며, 각 직역의 업무 문맥에 따라 상이한 관점에서 쓰일 수 있다

[그림 1] 삼성서울병원 EMR시스템 개념도: 여러 단위의 부서 업무가 EMR로 통합·연결



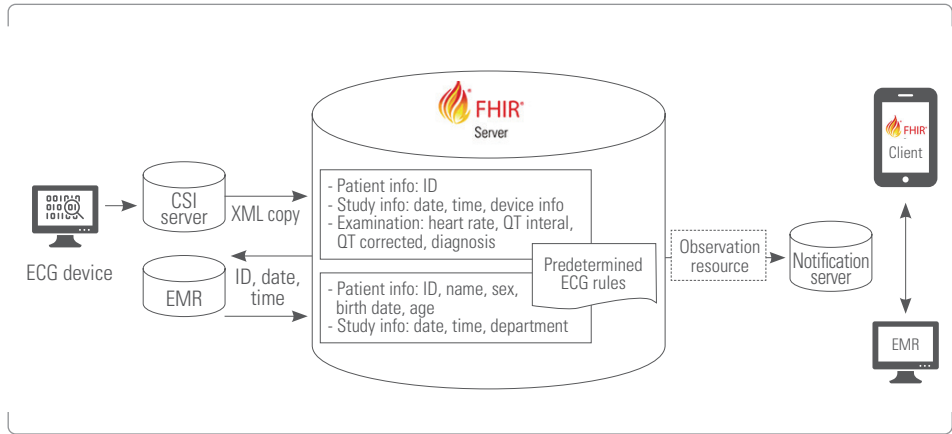
EMR 이 의료진, 특히 의사에 의하여 입력되는 데이터로 이루어져 있다고 여기는 경우가 있으나 이는 전체 데이터의 일부에 불과하다. 의사에 의해 직접 입력되는 데이터는 진료기록과 검사, 투약, 및 처치 처방 등에 국한되며 진단명 입력 역시 의사만의 권한인 경우가 많다.

간호사는 의무기록 중에 매우 중요한 부분을 담당하고 있다. 특히 입원 및 응급실 관련 기록은 간호사의 기록 없이는 환자의 상태를 재구성할 수 없을 정도로 중요한 정보이나 많은 경우 비정형데이터로 구성되어 재활용이 어렵다.

이 외에도 각종 검사실, 원무, 총무, 나아가 인사담당 부서까지 데이터가 연결되어 있으며 실제 데이터 분석에는 이러한 부서별로 나누어진 데이터 테이블에서 데이터를 조합해야 하는 경우가 많다.

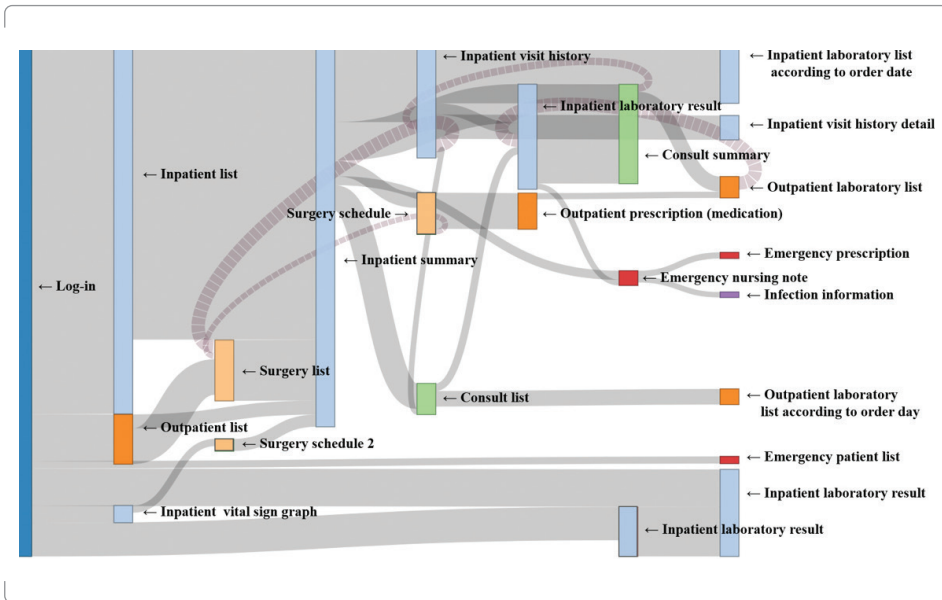
진료제공자(caregiver)들이 서로 다른 관점에서 환자를 대하기 때문에 해석에 주의해야 할 경우가 많은데 예를 들어 입원환자의 “퇴원” 시간을 정의할 때 의사의 퇴원처방, 간호의 퇴실기록, 원무의 퇴원기록, 그리고 주차장의 출차 기록이 상이하게 다른 경우가 흔히 발생한다. 진단명의 경우도 같은 환자를 돌보는 의료진의 영역에 따라 “주진단” 명이 달라질 수 있으며, 극단적으로는 같은 환자에게 같은 시각에 작성된 두 소견서에 완전히 다른 진단명 목록이 존재할 수도 있다.

[그림 2] 병원 EHR 통신데이터 시스템¹



사람의 의해서 생성된 데이터 외에도 의료 현실을 재구성하기 위해서는 시스템 내부의 통신데이터에 유의해야 한다. 병원 EMR은 많은 sub system과 인터페이스가 있고 이들 간의 통신 결과를 데이터로 저장하는 경우가 많다(그림 2). 이러한 깊은 데이터를 통해 남다른 통찰(insight)을 얻을 수 있는데, 예를 들어 특정 혈액 검사 결과값(Value) 보다 이 처방이 발행 → 접수 → 수납 → 수행 부서에서 확인 → 환자 내원하여 접수 → 바코드 출력 → 검사 시행자 수행 사인 → 이송용 트레이 바코드 → 검사시작 → 검사 결과 → 검사결과 확인(confirm)되는 각 과정에 대한 데이터를 활용하여 전체적인 시스템 최적화나 특정 검사의 속도를 높이는 연구를 수행할 수도 있다.

1. IT 리소스간의 통신데이터를 포함하고 있음. 이런 경우, EMR내에서의 데이터에 차이가 발생할 수 있으며, 본 건의 경우는 데이터 중 일부가 데이터표준 중 하나인 FHIR를 기반으로 송수신됨을 볼 수 있음

[그림 3] EMR 을 통해 본 의료진의 관심 동선²

주: EMR상의 Workflow를 따라가 보면 의료진별 off-line 업무를 유추할 수 있으며, 각 진료 주체가 해석하고 확인하는 정보가 제한적일 수 있음을 시사함

EMR은 의료행위를 정확하고 효율적으로 보조하기 위해 운영되고 있으나, 의료행위의 주체는 각 진료과나 부서 단위로 이루어지는 경우가 많기 때문에 이와 같은 진료과와 부서에서 만들어진 데이터 간에 모순이 발생하는 경우가 있다. 진료 주체들은 각자의 워크플로우를 가지고 정보를 확인하므로, 해당 워크플로우, 혹은 담당 범위에서 벗어나는 곳에 주요 정보가 있다면 누락될 가능성이 있다. [그림 3]

예를 들어 암 환자의 경우 신환으로 등록했을 경우에는 당시 상태가 비교적 객관적으로 등록이 되나, 치료가 진행되고 상태가 바뀌면서 치료 주체가 여럿이 되는 경우 금표준(gold standard)이 되는 병기를 정하기 어려운 사례가 발생한다.

2.2. EMR 관찰 기록은 환자의 임상적 문맥을 반영한다

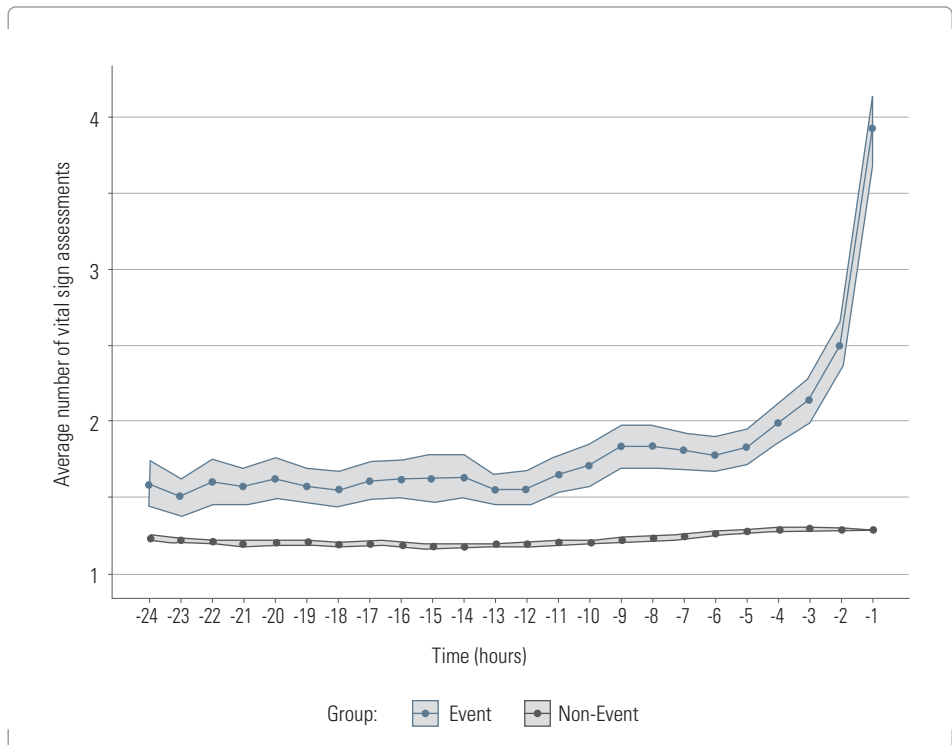
활력징후는 기계적으로 객관적일까? 거의 모든 의료데이터는 목적성을 가지며 이를 측정하는 의료진과 환자간의 상호작용을 내포하고 있어 이를 해석하고 활용하기 위해서는 데이터가 수집되는 상황적 문맥을 이해하여야 한다. 활력징후는 맥박, 호흡, 혈압, 체온으로 이루어져 있는데 의료진이 환자의

2. 같은 EMR을 사용하는 주체들 간에 정보는 제한적으로 공유, 갱신되며 공유되는 정보일수록 정확한 정보를 정의하기 어려움. [그림 3]

상태를 수치화하여 판단하는 가장 기초적인 정보이며 세계적으로 거의 같은 포맷을 사용하기에 활용하기 용이한 편이기에 많은 연구에서 활용되고 있으며 대부분의 악화예측모델에 포함된다. 활력징후는 의료진이 환자를 파악하려는 의도를 가지고 측정하고 기록하는 정보이므로 상태가 불안정하거나 주의가 요구되는 경우(예: 특정한 시술이 이루어진 전후)에도 집중적으로 이루어진다. 측정 그 자체가 의미를 가진다는 점으로 마치 금융신용조회가 개인의 신용평가에 영향을 미치는 것과 비슷한 맥락으로 이해할 수 있다. 실제 연구 결과, 활력징후 측정 횟수 자체가 상태 악화와 관련 있음을 알 수 있었다.

환자의 상태를 확인하기 위한 의료진의 문진 역시 임상적 예측을 기반으로 이루어지며 질문을 하고 자료를 기록하는 과정에서 이와 같은 소견(impression)이 영향을 미침을 인지해야 한다. 예를 들어 발열 환자가 요로감염이 의심되는 경우 비뇨기계 증상 등을 추가로 질문할 것이며 환자가 없다고 답한다면 의무 기록에 해당 증상이 없다고 기록할 것이다. 반면, 현재 진단 및 감별진단과 거리가 먼 문진 내용은 실제 문진이 이루어졌다 하더라도 기록되지 않을 가능성이 크다. 다시 말해 EMR 기록이 보는 환자 상태는 의료진의 추론을 거친 주관적 결과일 수 있음을 알아야 한다.

[그림 4] 환자의 중증도에 따른 의료인(간호사)의 상태평가 변화



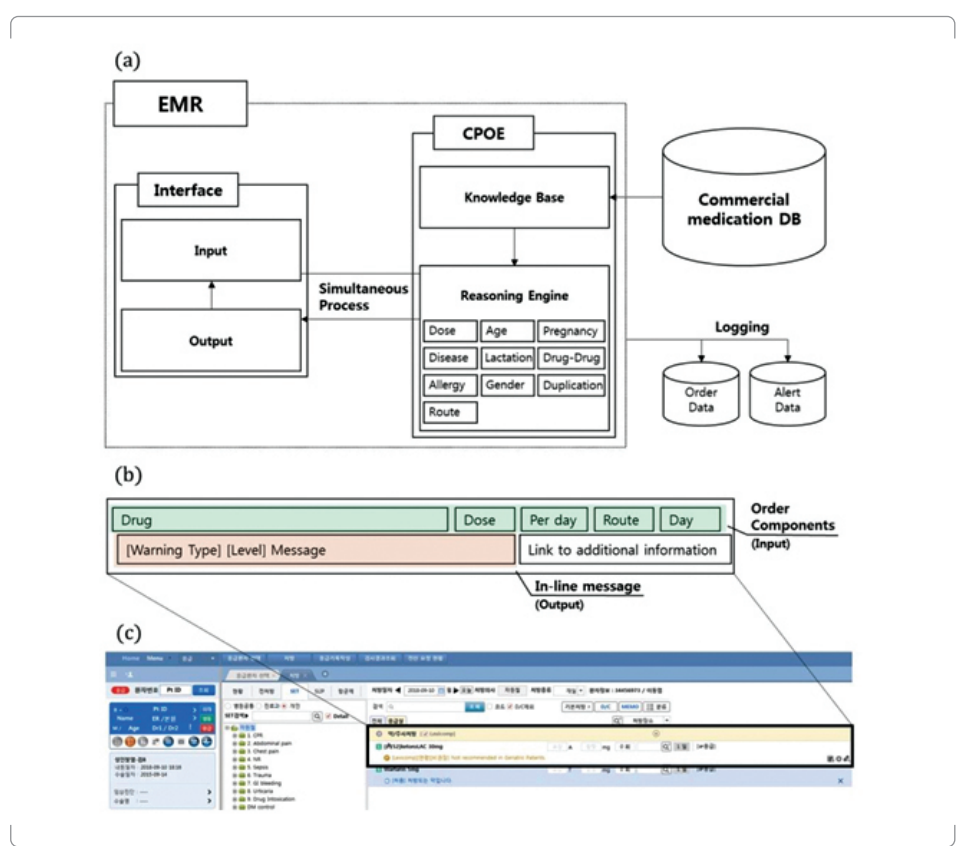
주: 환자가 상태가 급성으로 악화하기 전 간호사에 의한 상태 평가 횟수가 급격하게 늘어남. 환자 상태 측정이 목적성을 가지고 이루어짐을 시사함

2.3. 검사, 처치, 투약 처방은 질향상(Quality Improvement)의 대상이기에 기관의 정책적 문맥을 반영한다.

처치 투약 처방은 진료의 결과이기도 하며 환자에게 직접 영향을 미치는 내용이다. 또한, 행위별수가제의 특징상 수입과 직접 연결되는 내용이기에 여러 단계에서 관리의 대상이 된다. 따라서 처방 데이터는 환자안전, 보험과 관련된 국가적 혹은 기관 내 정책 변화를 고려해서 분석해야 올바른 결과를 얻을 수 있다.

관리 방법은 의료진에 대한 교육, 서면 공지 등 EMR 시스템 외의 방법뿐만 아니라 Clinical Decision Support System (CDSS)를 통한 제어 방법 등이 있을 수 있다. [그림 5] 이러한 CDSS와 처방 시 약물상호작용(interaction)을 기록하고 그 결과를 분석할 수 있다면 실제 의료진의 초기 의도와 system에 의한 중재(intervention) 결과를 구분할 수 있으나, 대부분의 경우 이러한 interaction결과는 별도로 기록되지 않고 있다.

[그림 5] 처방 제어를 위한 Clinical Decision Support System의 예시



주: 이 시스템의 경우 minimally-interruptive 방식을 사용하고 있으며, 사용자의 반응을 기록하도록 디자인되어 있음

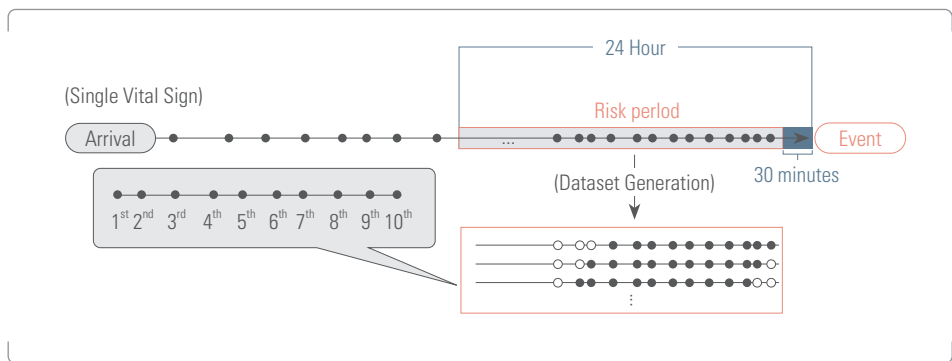
처방 데이터를 활용함에 있어서 유의할 또 한 가지는 처방세트(set)이다. 처방세트는 유사한 처방이 자주 발생할 경우, 개인 의료진이나 진료과 혹은 기관에서 여러 처방을 하나로 묶어서 처방이 일관성 있고 빠르게 이루어질 수 있도록 한 일종의 처방묶음이다. 따라서 특정 처방이 다른 처방과 반복적으로 같이 발생할 경우 처방 간의 새로운 관계를 발견한 것보다는 임의적 연결을 수리적으로 확인한 과정일 뿐일 수 있다.

처방세트에 의한 비뚤림(bias)을 최소화하기 위해서는 처방원 데이터를 확인하여 처방이 실제 어떠한 경로를 통해 이루어졌는지 확인할 수 있어야 한다. 예를 들어 a라는 처방정보 내에, 언제, 어디서, 누구에 의해, 누구에게, 어떤 진단명 하에 처방된다는 기본 정보 외에 해당 처방이 오더세트(set)에 의한 것인지, 어느 오더세트 에서 기원한 것인지 등에 대한 기원(origin)을 찾을 수 있어야 한다. 다만, 일반적으로 EMR 에서는 이러한 기원을 기록하지 않고 있는 것이 현실이다.

2.4. 해석은 개인의 단기적, 혹은 장기적 경과의 문맥 하에 분석되어야 한다

환자의 모든 데이터는 장기, 단기적 경과의 틀 내에서 분석되어야 한다. 특히 활력징후나 혈액검사와 같은 수치형 데이터에는 이러한 과정이 필수적이다. A환자의 정상 수치가 B환자에게는 비정상적일 수 있음은 물론 상황에 따라 A환자의 정상 범위가 변화하기 때문이다. 특히 시간의 축 위에서 변화하는 속도에 주의를 기울여야 한다. [그림 6]

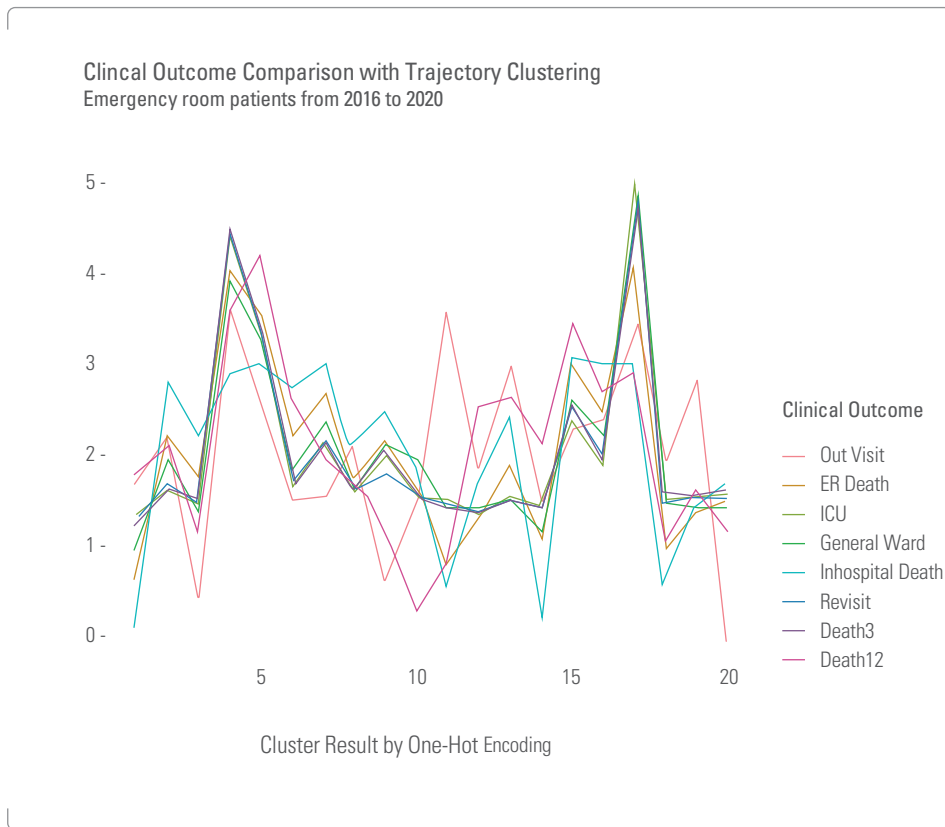
[그림 6] 활력징후를 시계열화하여 급성 악화를 예측하는 모델



주) 시계열화는 문맥적 의미를 판단하기 위한 필수적인 과정임

장기적 추이 경과에는 단기적 경과보다 해석이 어렵다. 특히 장기적으로 연속적인 수치형 데이터 대신 목록형 데이터를 활용해야 하는 경우가 많기 때문이며, 분석하고자 하는 결과지표가 광범위해질 수 있기 때문이다. 이런 경우 환자의 추이(trajecory)를 특정 결과(outcome)에 따라 군집(cluster)화하여 활용하는 방법도 있다.[그림 7] 다만, 이러한 데이터에 대한 접근은 아직도 초보적인 단계로 향후 많은 연구가 필요하다.

[그림 7] 환자의 장기적 추이 (trajectory)를 데이터로 활용한 예시



주: 이 그림에서 x축은 각 cluster, y축은 특정 outcome (응급실 방문이나 사망 등)의 확률로, 단순한 장기적 추이로 특정 outcome이 높게 나타나는 그룹을 특정할 수 있음을 보여줌(Unpublished)

3. 결론

모든 데이터, 특히 임상적 데이터는 문맥에 따라 만들어지고 있다. EMR 활용 주체의 업무에 따라 쓰이고 의료진의 임상적 지식의 문맥에 따라 취득되며 예상되는 환자 상태를 반영하고 있다. 또한 환자 안전이나 보험정책 등 처방의 외부적 요인이 반영되어 있어 이러한 상황을 고려한 접근이 필요하다. 마지막으로 해석에는 장기적, 단기적 시계열적 변화를 반영할 수 있는 모델이 필요하여 앞으로도 지속적인 연구가 이루어질 필요가 있다.

디지털 뉴딜 정책에서 EMR 빅데이터가 더 가치있게 쓰이기 위해서는 위와 같은 데이터의 특징을 반영한 분석이 선행되어야 한다. 올바른 분석은 새로운 정보와 지식을 창출하고, 나아가 혁신적인 서비스와 제품으로 이어져, 디지털 뉴딜 정책이 본래 목적을 이룰 수 있는 밑거름이 될 것이다.

참고문헌

- [1] Hong S, Lee S, Lee J, Cha WC, Kim K. Prediction of Cardiac Arrest in the Emergency Department Based on Machine Learning and Sequential Characteristics: Model Development and Retrospective Clinical Validation Study. *JMIR Med Inform.* 2020;8(8):e15932. Published 2020 Aug 4. doi:10.2196/15932
- [2] Hur S, Lee J, Kim T, et al. An Automated Fast Healthcare Interoperability Resources-Based 12-Lead Electrocardiogram Mobile Alert System for Suspected Acute Coronary Syndrome. *Yonsei Med J.* 2020;61(5):416-422. doi:10.3349/ymj.2020.61.5.416
- [3] Lee S, Hong S, Cha WC, Kim K. Predicting Adverse Outcomes for Febrile Patients in the Emergency Department Using Sparse Laboratory Data: Development of a Time Adaptive Model. *JMIR Med Inform.* 2020;8(3):e16117. Published 2020 Mar 26. doi:10.2196/16117
- [4] Soh JY, Jung SH, Cha WC, et al. Variability in Doctors' Usage Paths of Mobile Electronic Health Records Across Specialties: Comprehensive Analysis of Log Data. *JMIR Mhealth Uhealth.* 2019;7(1):e12041. Published 2019 Jan 17. doi:10.2196/12041
- [5] Jung KY, Kim T, Jung J, et al. The Effectiveness of Near-Field Communication Integrated with a Mobile Electronic Medical Record System: Emergency Department Simulation Study. *JMIR Mhealth Uhealth.* 2018;6(9):e11187. Published 2018 Sep 21. doi:10.2196/11187
- [6] 유준상. 처방 임상 의사결정지원시스템: 의료인공지능 수요조사로부터 개발 및 평가에 이르기까지 단계별 사례연구 [박사학위]. [서울]: 성균관대학교 삼성융합의과학원; 2020년 8월.