

2026 Vol. 4

Privacy Report

2026 개인정보 이슈
심층 분석 보고서

의료데이터 보호 및 활용

목차

Contents

1

생성형 AI 기반 진료기록 자동화와 환자정보 보호

유승찬 연세대학교 의과대학 의생명시스템정보학교실 교수

6

의료인공지능 신뢰성의 토대로서의 학습 데이터 - 전자의무기록 데이터의 품질

윤형진 서울대학교병원 의공학과 교수/임상의학정보실장

15

다기관·국경 간 의료데이터 공유와 연합분석 거 버넌스

김주한 서울대학교 의과대학 교수/연구부총장

생성형 AI 기반 진료기록 자동화와 환자정보 보호



유승찬

연세대학교 의과대학 의생명시스템정보학교실 교수

1. 들어가며

생성형 인공지능(Generative Artificial Intelligence, 이하 생성형 AI)의 발전은 의료현장에서 진료정보를 작성하고 활용하는 방식을 변화시키고 있다. 특히 대규모언어모델(Large Language Model, LLM)은 환자의 과거 진료기록과 검사 결과를 요약하고, 퇴원기록이나 경과기록을 자동으로 작성하는 데 활용되고 있다. 이러한 기술은 의료진의 문서 작성 부담을 줄이고 환자 진료에 집중할 수 있는 환경을 제공할 것으로 기대된다.

국내에서도 전자의무기록(Electronic Medical Record, EMR)과 연계된 생성형 AI를 활용하여 응급실 퇴원기록 등을 작성하는 사례가 보고되었으며, 기록의 질을 유지하거나 개선하면서 작성 시간을 단축할 수 있음이 확인되었다.^[1] 해외에서는 의사와 환자의 대화를 분석하여 진료기록을 자동으로 작성하는 기술이 확산되고 있다.^[2]

그러나 생성형 AI의 활용 과정에서는 환자의 민감한 건강정보가 처리되므로 개인정보 보호에 대한 검토가 필요하다. 특히 기존 EMR에 저장된 환자정보를 활용하는 방식과 환자·의료진의 대화를 새롭게 수집하는 방식은 개인정보 처리의 특성이 다르므로 구분하여 접근할 필요가 있다.

2. 기존 전자의무기록을 활용하는 생성형 AI

1) 활용 현황과 기대효과

EMR 기반 생성형 AI는 의료기관이 이미 보유한 진료기록, 검사 결과, 처방 내역 등을 활용하여 의료문서의 작성을 지원한다. 새로운 환자정보를 별도로 수집하기보다는 기존 진료정보를 재구성한다는 특징이 있으며, 적절한 진료 목적의 범위에서 필요한 안전조치를 갖춘다면 적극적인 활용이 가능하다.

국내 Y-KNOT(Your-Knowledgeable Navigator of Treatment) 프로젝트는 의료기관 내부에 구축한 LLM을 EMR과 연계하여 응급실 퇴원기록과 마취 전 평가기록의 초안을 자동으로

작성하도록 구현하였다.^[3] 응급실 퇴원기록 작성 연구에서는 50개 사례와 6명의 의사를 대상으로 비교한 결과, 기록 작성시간의 중앙값이 69.5초에서 32.0초로 감소하였으며 기록의 완전성과 정확성 등에서도 긍정적인 결과가 확인되었다.^[1]

2) 의료법상 국내 보관 요건과 AI 활용

「의료법」 제23조는 전자의무기록의 안전한 관리·보존에 관한 근거를 규정하고 있다. 특히 의료기관 외부에서 전자의무기록을 관리·보존하는 경우 적용되는 「전자의무기록의 관리·보존에 필요한 시설과 장비에 관한 기준」은 “전자의무기록 시스템 및 그 백업장비의 물리적 위치는 국내로 한정한다”고 명시하고 있다.^[4]

이는 법정 전자의무기록의 보관에 관한 요건으로, 모든 의료정보의 국외 처리를 일률적으로 금지한다는 의미는 아니다. 그러나 의료기관이 외부 AI 서비스를 활용할 때 데이터의 전송·처리·보관 위치와 관련 법령을 검토해야 하는 중요한 이유가 된다.

이러한 측면에서 **원내 구축형(On-premises) 생성형 AI**는 **국내 의료환경에서 유용한 대안**이다. 의료기관 내부에서 환자정보를 처리할 수 있으며, 기존 EMR의 접근권한 관리, 사용자 인증, 접속기록 관리 등 정보보호체계를 활용할 수 있다. 적절한 요건을 충족한 국내 클라우드 기반 AI 역시 활용 가능한 방식이다.

반면 ChatGPT와 같은 외부 생성형 AI 서비스는 추가적인 검토가 필요하다. OpenAI는 국내 데이터 저장 옵션을 제공하고 있으나, 2026년 10월 현재 한국 내 모델 추론까지 보장하지 않는다.^[5] 따라서 일반 ChatGPT에 식별 가능한 진료정보를 직접 입력하는 것은 지양해야 한다. 기업용 서비스를 활용하는 경우에도 개인정보 처리위탁, 국외이전 요건, 계약상 통제 가능성을 확인해야 한다.

결국 중요한 것은 특정 AI 서비스의 사용을 일률적으로 제한하는 것이 아니라, **의료기관이 환자정보의 처리 과정을 실질적으로 통제할 수 있는 환경을 확보하는 것**이다.

3. 진료실 대화를 활용하는 생성형 AI

1) 신규 개인정보 수집과 환자 동의

최근에는 환자와 의료진의 대화를 분석하여 진료기록을 작성하는 Ambient AI Scribe 기술이 확산되고 있다. 진료 대화를 음성으로 수집하고 텍스트로 변환한 뒤 생성형 AI로 요약하는 방식으로, 의료진의 문서 작성 부담을 줄일 수 있다는 연구 결과가 보고되고 있다.^[2]

그러나 기존 EMR 기반 AI와 달리, **이 기술은 EMR에 존재하지 않았던 음성 및 대화정보를 새롭게 수집한다.** 또한 진료 대화에는 질병 및 치료정보 외에도 가족관계, 생활환경, 보호자 등 제3자의 개인정보가 포함될 수 있어 별도의 보호조치가 필요하다.

특히 진료 대화를 자동으로 기록하는 AI는 환자에게 사용 목적과 방법을 충분히 설명하고 사전에 동의를 받는 것이 바람직하다. 「개인정보 보호법」상 모든 진료 목적의 음성 수집에 일률적으로 별도 동의가 요구되는 것은 아니지만, 환자가 예상하지 못한 정보 수집이 이루어질 수 있다는 점에서 적극적인 투명성 확보가 중요하다.

실무적으로는 녹음 및 AI 처리를 시작하기 전에 환자에게 수집 목적, 정보의 종류, 외부 서비스 이용 여부, 음성정보의 보관 및 삭제 방식을 설명하고 동의를 받는 방식을 권장할 수 있다. 환자가 동의하지 않는 경우에는 기존 수기 기록 등 대체 방식을 제공하는 것이 바람직하다. 영국 NHS의 2026년 가이드라인 역시 환자에게 AI 사용 사실을 알리고 반대 의사를 표시할 기회를 제공하도록 강조하고 있다.^[6]

2) 음성정보의 보관과 생성기록의 검증

진료 대화 기반 AI에서는 음성 원본, 전사문, AI가 작성한 초안, 의료인이 최종 확정한 진료기록을 구분하여 관리해야 한다. 일시적으로 수집한 음성과 전사문은 별도의 보존 필요성과 법적 근거가 없다면 최종 기록 검증 후 지체 없이 삭제하는 것이 바람직하다.

또한 AI가 생성한 문서의 정확성과 개인정보 적정성을 검토해야 한다. 실제 연구에서는 LLM이 진료 대화에 포함된 제3자의 개인정보를 임상적으로 불필요함에도 최종 기록에 포함하는 현상이 보고되었다.^[7] 따라서 의료인은 생성된 기록을 검토·수정해야 하며, AI가 작성한 초안이 확인 없이 정식 EMR에 확정되지 않도록 해야 한다.

4. 공통 관리 원칙과 향후 정책 방향

두 유형의 생성형 AI는 개인정보의 출처와 수집 방식에서 차이가 있지만, 안전한 활용을 위한 공통 관리 원칙이 필요하다.

첫째, 개인정보 최소화 원칙을 적용해야 한다. EMR 기반 AI에는 해당 문서 작성에 필요한 진료정보만 제공하고, 진료 대화 기반 AI에서는 불필요한 음성정보의 수집과 보관을 최소화해야 한다.

둘째, **진료기록 작성 지원과 AI 모델의 학습·성능개선을 위한 환자정보 활용은 구분해야 한다.** 진료 목적으로 환자정보를 처리할 수 있다는 사실이 모델 학습에 이를 제한 없이 활용할 수 있음을 의미하지는 않는다. 따라서 추가 학습 등에 활용할 경우 별도의 법적 근거와 필요한 보호조치를 검토해야 한다.

셋째, 접근 권한 관리와 지속적인 검증체계를 마련해야 한다. 의료진의 업무 범위에 따라 정보 접근을 제한하고, AI의 사용 및 승인 이력을 관리하며, 생성된 기록의 오류와 불필요한 개인정보 포함 여부를 정기적으로 평가해야 한다. 개인정보보호위원회의 「생성형 인공지능(AI) 개발·활용을 위한 개인정보 처리 안내서」 역시 이러한 위험 기반의 안전조치와 거버넌스를 강조한다.^[8]

향후 정책적으로는 **EMR 기반 AI와 진료 대화 기반 AI의 차이를 반영한 실무 지침**을 마련할 필요가 있다. 전자에 대해서는 국내 보관 요건과 안전한 AI 처리환경에 관한 기준을 명확히 하고, 원내 구축형 및 적절한 국내 클라우드 기반 AI의 활용을 지원하는 것이 바람직하다. 후자에 대해서는 환자 안내와 동의, 음성정보의 보관 및 삭제 등에 관한 표준적인 절차를 제공할 필요가 있다.

아울러 의료기관과 AI 서비스 제공자 간의 개인정보 처리 목적, 보관기간, 모델 학습을 위한 재이용 여부, 재위탁 및 국외이전 등에 대한 책임을 계약으로 명확히 해야 한다.

5. 맺으며

생성형 AI 기반 진료기록 자동화는 의료진의 업무 부담을 줄이고 의료서비스의 효율성과 질을 높일 수 있는 기술이다. 특히 기존 EMR을 활용하는 생성형 AI는 의료기관의 정보보호체계와 연계하여 안정적으로 도입할 수 있으며, 원내 구축형 AI는 개인정보 보호와 기술 활용을 조화시킬 수 있는 실용적인 대안이다.

반면 진료 대화를 새롭게 수집하는 AI는 충분한 사전 설명과 동의, 정보 최소화 및 적절한 삭제 절차가 중요하다. 향후 개인정보 보호 정책은 이러한 차이를 반영하여 일률적인 제한보다 활용 목적과 위험 수준에 따른 차등화된 관리기준을 마련할 필요가 있다.

궁극적으로 중요한 것은 환자의 권리와 신뢰를 보호하면서도 의료현장에서 생성형 AI의 활용 가능성을 확대하는 것이다.

참고문헌

- [1] Song JW, et al. Large Language Model Assistant for Emergency Department Discharge Documentation. JAMA Network Open. 2025;8(10). doi:10.1001/jamanetworkopen.2025.38427.
- [2] Olson KD, et al. Use of Ambient AI Scribes to Reduce Administrative Burden and Professional Burnout. JAMA Network Open. 2025. doi:10.1001/jamanetworkopen.2025.34976.
- [3] Kim H, et al. A Bilingual On-Premises AI Agent for Clinical Drafting: Implementation Report of Seamless Electronic Health Records Integration in the Y-KNOT Project. JMIR Medical Informatics. 2025;13. doi:10.2196/76848.
- [4] 보건복지부. 「전자의무기록의 관리·보존에 필요한 시설과 장비에 관한 기준». 보건복지부고시 제2023-245호. 2023.
- [5] OpenAI. Introducing Data Residency in Asia. 2025; Data Residency and Inference Residency for ChatGPT.
- [6] NHS England. Using AI-enabled Ambient Scribing Products in Health and Care Settings. 2026.
- [7] Chim J, et al. Evaluating Privacy Leakages in LLM-driven Ambient Clinical Documentation. Frontiers in Digital Health. 2026;8:1761624. doi:10.3389/fdgth.2026.1761624.
- [8] 개인정보보호위원회. 「생성형 인공지능(AI) 개발·활용을 위한 개인정보 처리 안내서». 2025.

의료인공지능 신뢰성의 토대로서의 학습 데이터 – 전자의무기록 데이터의 품질

Training Data as a Foundation for Trustworthiness of medical
artificial intelligence – Quality of Data from Electronic Medical Records



윤형진

서울대학교병원 의공학과 교수/임상외과학정보실장

최근 인공지능 에이전트들이 인간의 통제를 벗어나 예기치 않은 동작을 하는 사례들이 보고되면서 신뢰할 수 있는 인공지능에 대한 관심이 증가하고 있다. 신뢰할 수 있는 인공지능이란 신뢰할 수 있고 윤리적으로 작동하는 인공지능의 속성으로, 그 과정에서 인간의 권리, 가치, 자율성이 존중받으며, 공공의 신뢰를 손상하지 않고 증진하도록 적용하는 것을 말한다.^[1]

EU에서는 1) 인간의 행위 주체성과 감독, 2) 기술적 강건성과 안전, 3) 프라이버시와 데이터 거버넌스, 4) 투명성, 5) 다양성, 차별 금지 및 공정성, 6) 사회환경적 안녕, 7) 책임성 등을 신뢰할 수 있는 인공지능의 기준으로 제시하고 있다.^[2] 2024년 WHO에서는 1) 인간의 자율성 존중, 2) 인간의 안녕, 안전 및 공공선의 증진, 3) 투명성, 설명가능성, 이해가능성 보장, 4) 책임성과 책임성 제고, 5) 포용성과 형평성 보장, 6) 대응성과 지속가능성을 갖춘 인공지능 촉진 등을 인공지능의 의료 적용에 있어 반드시 고려하여야 할 준칙으로 제시한 바 있는데 위에서 말한 신뢰 가능한 인공지능의 기준과 상당 부분에서 중첩된다.^[3] 인간의 생명을 다루는 의료의 윤리에서 첫 번째 원칙이 “Do no harm”인 것을 고려하면 인공지능의 의료 적용에서 고려하여야 할 준칙과 신뢰할 수 있는 인공지능의 기준이 상당 부분 일치하는 것은 당연하다.^[4]

신뢰할 수 있는 인공지능의 기준 중 많은 부분이 학습 데이터와 직결되며, 인공지능의 질적 수준은 그 학습 데이터의 질적 수준이 결정한다거나, ‘garbage in, garbage out’이라는 격언은 신뢰할 수 있는 인공지능에서 학습 데이터가 그 출발점임을 쉽게 인식할 수 있게 한다. 의료인공지능 기술의 임상 적용에 필수적인 규제 당국의 승인 과정에서 안전성과 임상적 유용성을 중점적으로 검토하게 된다. 의료인공지능기술의 기능이 적정 수준에 도달하지 못하여 잘못된 조언으로 의사의 오진을 유발하거나 적절하지 않은 치료 방법을 권유하게 되면 안전성이 미흡한 것으로 볼 수 있고, 임상적 유용성의 전제 조건도 충족하지 못하게 된다. 대상 환자 집단 전체에 대한 안전성과 임상적 유용성 외에 연령, 성별, 인종, 경제적 수준 등 취약 계층에서의 안전성과 임상적 유용성도 반드시 고려하여야 한다. 이러한 부분은 인공지능의 신뢰성에 매우 결정적인 기준으로 모델 개발 이전 단계 즉 학습 데이터 단계에서부터 고려가 필요하다.

의료인공지능의 안전성과 관련하여 개인정보보호도 고려하여야 한다. 학습데이터의 개인정보에 대한 비식별 처리가 가장 기본적인 개인정보보호 방법인데, 이는 개인정보를 보호하기 위하여 인위적으로 학습 데이터의 질을 낮추는 것이지만, 한편으로는 신뢰할 수 있는 인공지능 측면에서의 전제조건을 충족시키는 것으로 목적에 부합하는 데이터 전처리라고 할 수 있다. 이처럼 매우

다양한 측면을 가지고 있는 학습데이터의 품질은 의료인공지능의 신뢰성에 대한 결정적인 요인이며, 학습 데이터의 품질에 대한 평가는 신뢰할 수 있는 의료인공지능 기술 개발의 출발점이다.

최근 의료인공지능 기술의 개발은 대부분 전자의무기록(electronic medical records, 이하 EMR)을 이용하기 때문에, EMR에서 추출한 데이터의 질적 수준으로 본 논의를 한정하며, EMR에서 추출한 데이터를 EMR 데이터로 정의한다.

형식적 측면에서의 EMR 데이터의 질 평가 프레임워크

EMR 데이터의 질적 평가는 체계화가 미흡하여 평가 항목의 정의 등 매우 혼란스럽다. Weiskopf 등은 2013년 EMR 데이터의 질 평가와 관련한 과거 연구 결과들을 종합하여, 완전성(completeness), 정확성(correctness), 일관성(concordance), 타당성(plausibility), 시의성(currency) 등 EMR 데이터의 질 평가를 위한 5개 지표로 제안한 바 있다.^[6] 저자들은 완전성은 “EMR 데이터에 환자에 대한 사실이 존재하는가?”, 정확성은 “EMR 데이터에 존재하는 데이터 항목이 참값인가?”, 일관성은 “EMR 데이터에 존재하는 데이터 항목들은 서로 혹은 외부 관련된 데이터와 일관성을 유지하는가?”, 타당성은 “EMR 데이터의 항목은 해당 데이터 항목이 측정하고자 하는 현상에 대한 의학적 지식에 견주어 납득할 수 있는 값을 가지고 있는가?”, 시의성은 “EMR 데이터의 항목은 주어진 시점에서의 환자 상태를 잘 표현하고 있는가?” 등으로 정의하였다. Weiskopf 등의 평가지표들은 그 자체로 평가가 용이하지 않다는 한계가 있다.

2024년 Schwabe 등은 신뢰할 수 있는 인공지능에 있어 학습 데이터의 질적 수준이 가지는 중요성을 다시 확인하면서 불공정성의 주요 원인인, 데이터에 내재하고 있는 편향을 최소화하고, 강건성과 해석가능성을 제고함으로써 신뢰할 수 있는 의료 인공지능의 토대를 마련하는 것을 지원할 수 있도록 의료 데이터의 질적 평가 프레임워크를 제안한 바 있다.^[7] METRIC이라 명명된 이 프레임워크는 기존에 비교적 표면적인 질적 평가 프레임워크의 한계를 극복하고 내재된 편향 뿐만 아니라 데이터셋의 보안이나 개인정보보호도 평가항목으로 포함하는 프레임워크로서의 미가 있다. METRIC 프레임워크는 **M**Easurement Process, **T**imeliness, **R**epresentativeness, **I**nformativeness, **C**onsistency 등 5개의 클러스터와 Data management 클러스터 등 총 6개의 클러스터로 구성된다.

Measurement 클러스터는 데이터 수집 단계에서 기술 혹은 사람에 의한 영향과 관련되는데, 데이터 항목값의 정확도나 정밀도를 손상시키는 장비 오류나 사람의 실수, 결측치 비율을 평가하는 완전성, 데이터의 신뢰성을 평가하는 자료원 신뢰도 등이 포함된다. Timeliness 클러스터는 시간에 따르는 변화가 데이터셋에 적절히 반영되어 있는지 평가하는데, 데이터가 생성된 시점과 데이터가 사용되는 시점 간의 경과가 적절한지에 대한 평가와 최신성 등이 포함된다. Representativeness 클러스터는 데이터셋이 대상으로 하는 전체 인구 집단을 적절하고 포괄적으로 잘 반영하는지 여부로, 다양성은 각각의 인구통계학적 범주와 데이터 자료원을 충분히 반영하는지, 데이터의 깊이는 전체적인 자료의 양과 각 소집단에서 충분한 데이터가 확보가 되었는지, 대상 변수 균형은 대상 변수의 각

클래스에 데이터 숫자가 적절하게 확보되었는지를 평가한다. Informativeness 클러스터는 해당 데이터셋이 전달하고자 하는 정보를 얼마나 충실히 잘 전달하는지를 평가하는데, 데이터가 모호성이 없이 잘 이해가 되는지를 평가하는 이해가능성, 동일한 데이터가 한 번을 초과하여 제공되는지를 평가하는 중복성, 결측치가 추가의 정보를 함유하고 있는지 여부를 정량화하는 정보성 결측, 데이터 항목이 데이터셋 활용에 있어 추가로 유용성을 제공하는지를 평가하는 변수 속성 중요도 등이 포함된다. Consistency 클러스터는 데이터셋의 표현과 구성에 있어 일관성을 평가하는데, 내부 또는 외부의 규칙에 데이터가 얼마나 잘 부합되어 표현되는가 하는 규칙 기반 일관성, 데이터가 논리적으로 타당하고 모순이 없는지 평가하는 논리적 일관성, 서로 다른 소집단에서 통계적 속성이 유사한지를 평가하는 분포 일관성을 포함한다.

Data management 클러스터에는 데이터셋의 메타데이터와 관련된 문서화 차원, 정보보안 차원, 개인정보보호 차원, 데이터셋이 검색 가능하고, 접근 가능하며, 상호운용성이 보장되고 재사용이 가능한지(Findable, Accessible, Interoperable, Reusable; FAIR 원칙)를 평가하는 FAIR 원칙 차원 등이 포함된다. Data Management 클러스터를 제외한 5개의 클러스터에는 모두 15개의 차원이 구성되고 각각의 차원은 경우에 따라 다시 하위 차원으로 구분된다. 이러한 METRIC 프레임워크는 EMR 데이터의 질을 평가할 수 있는 포괄적이고 구체적인 방법론을 제공한다.

EMR 데이터의 편향

EMR 데이터는 대규모 조사자료에는 포함되지 않은 많은 정보를 포함하고 있고, 후향적으로 사용이 가능하며, 적절한 비식별처리를 통하여 비교적 쉽게 접근이 가능하다는 장점이 있지만, 연구목적으로 생성된 자료가 아니기 때문에 여러 가지 편향 등 한계를 내포하고 있다. 학습 데이터의 품질이 미흡하면 외적 타당성 (external validity) 및 내적 타당성 (internal validity)에 문제가 발생하게 된다. 연구에 포함된 집단이 전체 모집단과 차이가 있다면 연구 결과의 일반화에 취약점을 가지는 외적 타당성이 훼손되고, EMR 데이터의 편향 등으로 연구에 포함된 집단에서의 추정값들이 참값과 차이가 발생하면 내적 타당성이 훼손된다. 또, EMR 데이터에서는 건강의 사회적 결정요인에 대한 정보가 충분하게 수집되지 않는 경우가 많은데, 이러한 경우 교란 변수에 대한 충분한 보정이 불가능하기 때문에 추정값 등에 오류가 발생하게 된다.^{[8],[9]}

학습 데이터 편향 평가 프레임워크

Lee 등은 EMR 데이터를 이용한 연구 결과들의 편향을 분석하여 의료 서비스 제공 단계, 데이터 관리 단계, 연구 분석 단계 등 EMR 연구 과정을 따라 3단계로 분류하고, 각 단계별로 발생할 수 있는 편향의 원인을 평가하는 프레임워크를 제안하였다. 각 단계를 다시 다차원 계층으로 소분류하였는데, 의료 서비스 제공 단계는 의료서비스 접근, 의료서비스 제공, 데이터 수집 및 측정 등 3개의 계층으로, 데이터 관리 단계는 임상 데이터의 기록 및 코드화 실행과 연구 데이터 추출 등 2개의 계층으로, 연구 분석 단계는 데이터 모델과 연구 분석 등 2개의 계층으로 분류하였다.

의료 서비스 접근 계층은 해당 의료기관의 진료권에 거주하는 인구 집단 중 해당 의료기관의 서비스를 이용하기 위하여 병원을 방문하는 환자 집단의 특성에 대한 것으로 환자들의 인구통계학적 특성, 사회경제적 상태, 생활습관, 거주 환경, 임상적 특성 등이 해당된다. 이러한 특성이 진료권 거주 인구 집단과 유의한 차이가 있다면 편향이 발생할 수 있다. 전술한 환자의 대표성 등이 여기에 해당할 수 있다. 의료 서비스 제공 계층은 의료진, 의료기관의 지역별 차이, 치료 성적이나 수준, 기관의 특성 등 서비스 제공자와 관련되어 발생할 수 있는 편향이다. 데이터 수집 및 측정 계층은 검사실 검사 프로토콜, 검사 장비 운영자, 검사 장비 정확도, 측정 오류 등으로 인하여 발생하는 편향을 의미한다. 데이터 관리 단계의 임상 데이터 기록 및 코드화 계층은 앞선 의료 서비스 제공 단계에서 발생하는 검사실 검사 결과, 병리 결과지, 영상 판독지, 진단 결과 등 다양한 데이터를 EMR 시스템에 저장하기 위하여 구조화된 스키마로 변환하는 과정에서 발생하는 편향이다. 데이터 추출 계층은 EMR 데이터를 연구 분석을 위하여 ETL하는 과정에서 발생하는 편향이다. 마지막 연구 분석 단계의 데이터 모델 계층은 용어 체계, 용어 표준화, 의미론적 변환 등의 과정에서 발생하는 편향이고, 연구 분석 계층은 코호트 정의나 분석에 포함되는 변수의 취사 선택, 분석 알고리즘 등과 관련된 편향이다. 이러한 프레임워크의 각 단계와 계층을 따라서 학습 데이터에 편향이 존재하는지 순차적으로 평가하면서 편향의 문제를 처리하면 내적 및 외적 타당성을 최대한 담보할 수 있는 인공지능 솔루션을 개발할 수 있을 것이다.^[11]

표 1 각 단계별 편향 평가 프레임워크^[11]

단계	계층	세부 사항
Healthcare System	Access to Medical Care	Demographics, Lifestyle, Geographic factors, Socioeconomic status, Clinical characteristics, Healthcare engagement factor
	Provision of Care	Provider, Geographic variations, Level of care, Institutional characteristics, Care delivery processes, Encounter type, Insurance reimbursement policies
	Data Acquisition & Management	Laboratory testing, Data recording and interpretation bias, Testing protocol, Medical device, Medical imaging, Measurement
Data Management	Clinical Documentation Coding Practice	Coding standards & processes, Clinical documentation, EHR system variations, Documentation practices
	Data Extraction	Heterogeneous data sources, Data extraction methods, Data quality & preprocessing, Data freshness & timeliness, Data format & consistency
Research	Data Model	Data standards, Data transformation, Data governance
	Data Analytics	Cohort, Data exploration & preparation, Computable phenotype, Study design, Algorithm modeling, Result/Evaluation

선택 편향

선택편향은 대상군을 선정하는 과정에서의 오류나 연구 참여에 영향을 미치는 요인에 의하여 노출과 결과변수와의 관계가 연구 참여 조건을 충족한 집단과 실제로 참여하였던 참여자 간에 차이가 나는 경우 발생할 수 있다. 실제 연구에 참여하지 않았던 집단에서의 노출과 결과변수 간의 진정한 관계는 알수가 없기 때문에 선택 편향은 추정만 가능하다.^[10]

선택편향은 결측치 문제와 많이 중복되어 비교적 관심이 낮지만 선택편향에 대하여 정확하게 평가가 이루어져야 결측치의 발생 기전을 이해할 수 있고 적절한 결측치 처리가 가능하기 때문에 선택 편향의 평가는 EMR 데이터를 이용한 연구에서 중요하다.

선택 편향에 대하여 적절하게 평가하지 않고 전체 데이터셋의 결측치 발생 기전을 하나로 단순화 하여 가정하고 결측자료를 대치하게 하면 학습 데이터에 내포되어 있는 환자나 의료진, 의료기관 등에서 선택편향과 데이터 결측을 초래하는 다양한 인자들 간의 복잡한 상호작용을 간과하여 결과적으로 선택편향을 적절하게 교정하지 못하게 될 우려가 있다.

Haneuse S와 Daniels M은 선택편향의 평가를 위한 프레임워크를 제안한 바 있다. 선택 편향을 평가하는 첫 번째 단계는 연구 목적을 달성하는 데 필요한 최적의 데이터셋을 가상으로 정의하는 것이다. 다음은 데이터 수집의 각 단계별로 '어떤 데이터가 왜 수집되었는가' 즉 데이터 계보(data provenance)를 파악하는 것이다. 각 단계별로 왜 어떤 환자는 분석에 필요한 모든 데이터 항목에 대한 데이터를 가지고 있고 어떤 환자는 그렇지 못한지 파악하는 것이다. EMR 데이터의 데이터 계보 각 단계 별로 결측의 발생 기전을 파악하고 이를 통하여 최적의 결측 대체 방법을 적용함으로써 선택 편향을 가능한 보정할 수 있다. 그러나 이러한 각 단계별 평가에 있어 명확한 기준점이 없기 때문에 연구자들은 각 단계별로 불확실성을 내포한 의사결정을 해야 한다는 것이 이 프레임워크의 가장 큰 단점이라고 할 수 있다.^[12]

결측치

EMR은 진료 과정에서 생성되는 데이터로 연구 목적으로 축적되는 데이터와는 근본적인 차이가 있는데 데이터 항목 중 결측 비율이 높은 항목이 매우 많다는 것이 그 중 하나이다. 결측치는 환자의 거부 또는 의료진의 측정 누락 등으로 발생(a lack of collection)할 수도 있지만 음성 데이터인 경우 의료진이 기록을 누락하면서 발생하는 경우(a lack of documentation)도 상당히 많다. 기록 누락으로 인한 결측치는 결측치 대체로 해결하기 어렵다. 대부분 임상 경험을 통하여 음성(negative)으로 처리하게 되는데 오류의 가능성을 배제할 수 없다.

수집 누락으로 인한 결측치는 missing-completely-at-random (MCAR), missing-at-random (MAR), non-missing-at-random (NMAR) 등으로 구분한다. 예를 들어 운석이 주차해 놓은 자동차 위로 떨어지는 바람에 환자가 병원에 오지 못해서 검사를 못한 경우는 MCAR에 해당한다.

MAR은 건강한 사람은 병원에 오는 횟수가 아무래도 만성질환을 가진 환자보다 적을 수밖에 없는 데 그 결과 혈압 측정 데이터에 결측의 확률이 높다. 따라서 이 경우에는 병원 방문 횟수와 혈압으로 단변량 분석을 하면 양의 상관관계가 나오겠지만, 현재 건강 상태에 대한 다른 변수들을 보정하는 다변량 분석을 하면 이러한 양의 상관관계가 사라진다. NMAR은 어떤 악의적인 검사실 직원이 있어 특정 성씨를 가진 사람인 경우에는 의도적으로 검사 결과를 입력하지 않는 것과 같이 전혀 측정하거나 고려하지 못하는 변수에 의하여 전적으로 발생하는 결측으로 극단적으로 예외적인 상황에서 발생한다. NMAR의 경우 편향이 없는 통계값을 추정해내는 것이 매우 어렵다. 다만 EMR에 포함되어 있는 데이터 항목은 대부분 다른 변수들과 연관되어 있는 경우가 대부분이고 결측치가 가지는 데이터 변이가 100% MCAR이나 NMAR에 기인한다는 것은 상상하기 어렵다.

결측치 처리에서 가장 간단하고 흔한 처리 방법은 결측치를 포함하는 증례를 제외하고 분석하는 complete case analysis인데, MCAR인 경우 통계적 검정력이 저하되기는 하지만 분석 결과에 편향이 발생하지는 않는다. 그러나 EMR의 결측은 대부분 나름대로의 기전을 가지고 있어 결측이 있는 증례는 결측이 없는 증례들과 구조적으로 차이가 있는, 즉 NMAR 이외의 경우가 거의 대부분이기 때문에 결측이 있는 증례를 제외하게 되면 편향이 발생하게 된다.

따라서 complete case analysis보다는 결측치 대체를 권장하고 있다. 결측치 대체 방법은 여러 가지가 있는데 최근에는 multiple imputation이 가장 많이 이용된다. 결측치 imputation의 세부 사항에 대해서는 다루지 않는다.^[13] 실제 결측치 대체 과정에서 가장 먼저 만나는 문제는 어느 정도의 결측률까지 결측치 대체가 가능할까 하는 것이다. 결측률 50% 정도까지는 안정적으로 대체가 가능하지만 50%~70% 사이는 주의가 필요하고 70%를 넘는 경우에는 분석 결과의 신뢰도에 상당한 영향을 미친다는 연구 결과가 보고된 바 있다.^[14] 그러나 이 연구 결과는 EMR을 이용한 연구가 아니라 100여 개 국가에서 수집되는 데이터를 기반으로 한 simulation 연구로 실제 EMR 연구에서도 그대로 적용하는 것은 한계가 있다. 결측치 대체에서 고려해야 하는 사항은 결측 데이터(missing data)와 결측 정보(missing information)를 구분하는 것이다. 결측 데이터는 어떤 데이터 항목에서 실제로 결측된 데이터의 숫자를 말한다. 결측 정보는 multiple imputation에서 생성되는 데이터셋에서 분석 결과의 결과값이 얼마나 안정적인가 혹은 불안정한 것인가를 통해 추정할 수 있다. 어떤 경우에는 결측 데이터 비율, 즉 결측률은 높지만 imputed dataset 간의 결과값의 변이가 작을 수(안정적일 수) 있는데, 이런 경우 결측 정보는 작다고 평가할 수 있다. 결측된 데이터 항목과 연관되는 다른 데이터 항목이 많은 경우에는 해당 데이터 항목의 결측률이 높더라도 결측 정보량이 작을 수 있고 반대로 연관되는 다른 데이터 항목이 별로 없는 경우에는 해당 데이터 항목의 결측률이 비교적 낮더라도 결측 정보량이 클 수 있다.^[13] 이러한 결측률이 높은 변수가 적절하게 처리되지 않는다면 그로 인한 편향이 발생할 수 있고, 만일 그러한 결측 변수가 교란변수라면 잔여 교란으로 이어질 수 있어, 분석 결과의 신뢰성을 담보할 수 없다.

정보 편향

정보편향은 데이터를 수집하는 과정에서 발생한다. 노출변수에 대한 평가나 결과변수에 대한 평가에 있어 구조적인 오류가 있는 경우 정보 편향이 발생한다. 정보 편향은 분석 대상 집단 간에 오류의 정도에 차이가 없는 경우(non-differential)와 집단 간 차이가 있는 경우(differential)가 있다. EMR 데이터 연구에서는 진단 장비나 검사의 정확도에 문제가 있을 경우 non-differential 정보 편향이 발생할 수 있고, 연구 대상 집단 간에 설문지 응답의 정확도에 차이가 있는 경우 differential 정보 편향이 발생할 수 있다. Non-differential 정보 편향은 거의 항상 통계적 차이를 상쇄시키는 방향으로 작용하고, differential 정보 편향은 통계적 차이를 증가시키기도 하고 감소시키기도 한다. 정보 편향의 흔한 유형으로 회상 편향 (과거 노출에 대한 부정확한 기억으로 인한 편향), 면담자 편향 (면담자가 연구 가설에 부합 또는 배치되는 응답에 선입관을 가지고 인터뷰 응답을 처리하면서 발생하는 편향), 관찰자 편향 (관찰자가 연구 참여자의 노출변수 상태에 대하여 알고 있는 상태에서 결과 변수를 평가하면서 발생하는 편향), 회귀희석편향 (regression dilution bias) 등이 있다.^[10] 회귀희석편향은 혈압과 같은 측정 시점에 따라 변동이 있는 연속형 변수인 노출변수인 경우 한두 번 측정한 측정값이 그 환자의 혈압 상태를 정확하게 반영하지 못하므로 결과변수와의 상관계수가 참값으로부터 차이가 나는 편향을 말한다.

EMR 데이터가 가지고 있는 이러한 편향들을 잘 이해하고 학습 데이터셋 구성이나 전처리 등 연구 수행 과정에 적절히 반영하는 것이 신뢰할 수 있는 인공지능을 개발하는 관건이 된다.

생성형 인공지능과 프라이버시 침해

학습 데이터에 개인정보가 포함되어 있을 경우 적절하게 비식별처리를 하지 않으면 data memorization이 발생하면서 프롬프트 조작 등을 통하여 개인정보가 유출될 위험이 있다. 학습 데이터의 비식별처리는 개인정보 보호를 위하여 학습 데이터의 질을 의도적으로 저하시키는 것으로 가명처리의 원칙이나 세부사항은 보건복지부가 개인정보보호위원회와 공동으로 발표하는 '보건의료데이터 활용 가이드라인'을 참고하기 바란다.

인공지능의 학습 데이터와 관련하여 비식별 처리 외에 많이 거론되는 것이 membership inference attack(MIA)이다. 기존 인공지능기술은 물론 생성형 인공지능에서도 MIA는 문제가 될 수 있다. 다수의 데이터 유출 사고로 상당히 많은 정보주체들의 정보가 다크웹 등에서 거래되고 있는 현실을 고려하면 어느 인공지능 솔루션 하나에서의 MIA보다는 이미 확보하고 있는 개인정보를 바탕으로 MIA를 시도했을 때의 위험을 고려해야 한다. 인공지능기술에 대한 MIA에서 소수 집단의 경우 공격 성공률이 나머지 집단에 비하여 더 높다는 연구 결과가 보고된 바 있어, 소수 집단의 경우에는 인공지능 솔루션의 성능에 대한 우려와 함께 MIA와 같은 프라이버시 공격 시에 더 취약하다는 점을 인공지능 신뢰성 평가에 고려하여야 한다.^[5]

인공지능 솔루션의 성능은 그 솔루션의 개발에 사용된 학습데이터만큼 좋다는 얘기가 있다. 즉 학습 데이터의 질은 인공지능 솔루션의 기능적 상한선이 된다. 소위 real world data는 우리 인간의 일상생활에 대한 거울이기 때문에 여러 불합리한 관행이나 습관 등이 모두 반영되어 있다. 인공지능 솔루션 개발에서 이러한 부분을 적절하게 처리하지 못하면 성능과 공정성이 담보되지 못하는 인공지능 솔루션을 개발하게 되어 이런 불합리를 더욱 급속하고 광범위하게 증폭시킬 수 있다. 방대한 자료로 사전학습을 수행하는 생성형 인공지능의 경우도 전처리를 통하여 학습 데이터의 품질이 더 향상되면 HW 등 자원을 덜 소모하면서 우수한 성능을 낼 수 있는 모델 개발이 가능해질 것이다. 신뢰할 수 있는 인공지능 기술의 개발에 있어 고려하여야 하는 여러 사항들을 기준에 부합하도록 하기 위해서는 앞서 살펴본 구조적 질 평가 항목 외에 편향 등과 관련한 사항들을 고려하여 학습 데이터를 확보하고 전처리하는 것이 필수적이며 특히 의료인공지능 기술개발에 있어서는 더욱 강조되어야 한다.

참고문헌

- [1] Potdar A. Introduction to trustworthy AI. In: Gudur, V., Pandey, B., Patel, A. (eds) Trustworthy AI Systems. Springer, Cham. http://doi.org/10.1007/978-3-032-15606-8_1, 2026
- [2] <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- [3] Ethics and governance of artificial intelligence for health. Guidance on large multi-model models. Geneva: World Health Organization. 2024
- [4] Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nature Med*, 25:2337–2340, 2019.
- [5] Knolle MA, Menten M, Jungmann F, Meissen F, Glocker B, et al. Disparate privacy risks from medical AI. *Nature*, 656:192–198, 2026
- [6] Weiskopf N, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform assoc*, 20:144–151, 2013
- [7] Schwabe D, Becker K, Seyferth M, Klasß A, Schaeffter T. The METRIC-framework for assessing data quality for trustworthy AI in medicine: a systematic view. *npj Dig Med*, 7:203, 2024
- [8] Goldstein ND. Electronic health records in epidemiology: appropriate questions, common biases, and potential sensitivity analyses. *Cur Epidemiol Rep*, 12:11, 2025
- [9] Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med*, 178(11):1544–1547, 2018
- [10] Tripepi G, Jager K, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. *Nephron Clin Pract*, 115:c94–c99, 2010
- [11] Lee H, Xiong C, Baughman D, Dun C, Tong J, et al. A multidimensional hierarchical framework for sources of bias in real-world healthcare evidence: a scoping review. *J Biomed Inform*, 174:104989, 2026
- [12] Haneuse S, Daniels M. A general framework for considering selection bias in EHR-based studies: what data are observed and why? *eGEMs*, 4(1):16, 2016
- [13] Wells BJ, Chagin KM, Nowacki AS, Kattan MW. Strategies for handling missing data in electronic health record derived data. *eGEMs* 1(3):7, 2013
- [14] Junaid KP, Kiran T, Gupta M, Kishore K, Siwatch S. How much missing is too much to impute for longitudinal health indicators? A preliminary guideline for the choice of the extent of missing proportion to impute with multiple imputation by chained equations. *Population Health Metrics*, 23:2, 2025

다기관·국경 간 의료데이터 공유와 연합분석 거버넌스

- 유럽연합의 Gaia-X와 EHDS의 제도적 결합 및 IHDSI의 국제협력 과제 -



김주한

서울대학교 의과대학 교수/연구부총장

1. 유럽연합 의료데이터 연합분석의 배경

유럽 의료데이터공간(EHDS, European Health Data Space)은 유럽연합(EU) 전역의 보건 의료 데이터를 안전하게 공유하고 활용하기 위해 EU가 추진 중인 핵심 디지털 전환 프로젝트다. EU는 코로나19 팬데믹 당시 국가 간 보건의료 데이터가 파편화되어 있고 실시간 공유가 어려워 신속하고 긴급한 공중보건 대응에 한계를 겪었다. EU GDPR(일반 데이터 보호 규정)이 환자의 데이터 권리를 보장하고 있지만, 국가 간 의료시스템과 법적 해석의 격차로 실제 의료현장이나 국경을 넘어설 때 실효성 있게 작동하지 못했다. 이에 EU는 2020년 '유럽 데이터 전략(A European Strategy for Data)'을 발표하며 단일 데이터 시장을 구상했고, 의료를 가장 우선 분야로 선정했다.

EHDS의 비전은 시민과 과학을 위한 안전한 디지털 의료데이터 연합 체계다. EU 시민이 어디서나 자신의 의료데이터에 안전하게 접근하고 통제할 수 있는 권리를 누리며, 연구자와 혁신가가 방대한 고품질 데이터를 바탕으로 의학적 혁신을 선도하는 세계 최대 규모의 안전한 의료데이터 환경 조성이다. EHDS는 1차 활용과 2차 활용, 그리고 시장통합을 핵심목표로 삼았다.

- **1차 활용 (Primary Use):** 환자 중심의 의료서비스 개선
- 개인이 자신의 전자의무기록에 실시간, 무료 접근 및 제어할 수 있는 권리 강화.
- EU 역내 다른 회원국 여행, 이동 시에도 전자처방전(ePrescriptions)이나 진료요약서(Patient Summaries) 등을 원활하게 연계하여 의료서비스의 연속성 제공.
- **2차 활용 (Secondary Use):** 연구, 혁신 및 정책 수립 촉진
- 보안이 확보된 폐쇄적 환경(HealthData@EU) 내에서 과학연구, 의료AI 개발, 보건의료정책 수립 및 규제 목적을 위한 의료데이터의 안전한 재사용.
- **디지털헬스 단일 시장:** 전 EU 가맹국에 통용되는 전자의무기록 시스템의 상호운용성 확립

EHDS는 현재 점진적 도입 단계에 진입했다; 2027년 제도·규격 정비, 2029년 주요 의무 적용, 2031년 데이터 범주 확대. EHDS는 2025년 3월 공포 발효되었지만, 제105조는 일반 적용시점을 2027년 3월 26일로 정했다.^[1] 유럽연합 내 약 24,000개 병원의 의료데이터가 유기적으로

연합되고 활용되는 시간표가 째깍째깍 흐르고 있다. 제55조 제6항은 지정된 HDAB의 명칭 등 식별정보를, 제75조 제1항은 2차 이용의 국가 연락처의 명칭과 연락처의 집행위원회 통보를 의무화한다. 제70조의 신청, 허가, 요청 서식, 제73조 제5항의 안전처리환경(SPE, Secure Processing Environment) 요건, 제75조 제12항의 HealthData@EU 기술 요건, 제77·78 조의 메타데이터와 품질 표시 관련 이행법령도 핵심 대상이다. 2029년부터는 카탈로그와 허가 절차를 실제 서비스로 운영할 역량이 필요하다. 제60조의 데이터 제공은 허가 또는 승인된 요청에 따른 HDAB의 요구와 연결된다. 제50조는 자연인과 소기업 법인에 대한 제외 및 회원국 법에 따른 적용 가능성을 규정한다. 보유기관은 자료 준비 인력과 품질관리 비용을 확보하고, 접근기구는 심사자연과 반출병목을 관리해야 한다.

표 1 EHDS 단계별 적용과 연합분석 운영의 준비 과제

시점	법정 변화	운영 준비의 의미
2027.3.26	일반 적용. 디지털보건당국·HDAB·국가연락처의 지정·통보, 주요 이행법령 채택 기한	역할과 절차 확정, 카탈로그·허가 서식·안전처리환경·감사 체계 준비
2029.3.26	2차 이용 제4장 대부분 적용. 1차 이용은 환자요약·전자처방·전자조제의 우선 범주 적용	실제 신청·심사·제공·분석·반출을 운영하고 권리 행사와 감독을 연결
2031.3.26	1차 이용 확대: 의료영상·검사결과·퇴원보고서. 2차 이용 확대: 유전체·오믹스 등	다중모달 의미 정합성, 고차원 식별 위험, 연구자료의 재이용 조건 관리
2035.3.26	제3국·국제기구 HealthData@EU 참가: 제75조 제5항 적용	공식 참가 요건·상호성·보호 수준을 검토하는 후속 국제연계 단계

제3국과 국제기구의 HealthData@EU 공식 참가조항의 적용시점은 2035년이다. 그러나 EHDS 헤드쿼터인 룩셈부르크 국립보건원(LIH, Luxembourg Institute of Health)이 2025년 7월 LIH와 한국 국립암센터가 공동주도하고 네이버클라우드, 오케스트로, 미국 싸이퍼룸사가 기술 파트너로 참여한 협약으로 한국은 EU 회원국 외 최초의 EHDS 참가국이 된 것은 시사점이 크다.

2. 유럽연합 의료데이터 연합활용의 관점을 데이터셋 공유에서 공동연구 수행 체계로

정밀의료의 병목은 데이터의 절대량보다 서로 다른 기관 및 국가에 분산된 자료가 하나의 연구질문에 적법하고 재현가능하게 분석되는가에 달려있다. 원자료의 이동을 제한하는 기술만으로 국가 간 협력연구가 활성화되지 않는다. 완전 중앙집중화나 데이터 절대 이동금지도 답은 아니다. 다기관 암 연구에서 진단명만 같다고 비교가능한 모집단이 될 수는 없다. 병기평가 시점, 치료선 정의, 검체 채취 조건, 유전체 검사 플랫폼, 추적 관찰의 시작과 종료가 다르면 같은 변수명도 서로 다른 임상사건을 의미한다. 여기에 기관별 연구윤리 심의, 개인정보 처리 근거, 계약과

비용 산정이 중첩된다. 공통 코호트 정의, 실행가능한 분석코드, 검증된 통계량, 재사용 가능한 메타데이터와 로그 감사 기능도 핵심 자산이다.

EHDS는 연구승인, 데이터의 의미적 정합성, 실행제어, 결과반출 책임 배분 등을 검증가능한 하나의 통합절차로 제도화하고, 분석결과의 해석가능성과 권리보호를 동시에 보장하는 협력구조를 구축했다. 의료데이터공간은 단일 제품명이나 저장소가 아니다. 참여기관의 신원증명과 자격검증, 데이터의 소재와 의미, 접근목적과 허용범위, 분석환경과 출력조건을 공통규칙으로 제도화한 사회기술적 체계다. 데이터가 물리적으로 분산되어 있어도 연구자는 데이터 카탈로그에서 자료를 찾고, 규정된 절차로 접근신청을 하며, 승인된 분석을 수행하고 결과를 반출한다.

Gaia-X는 벨기에법상 비영리협회를 중심으로 발전한 신뢰와 상호운용성의 규격체계다. 2025년 결성된 IHDSI(International Health Data Space Initiative)는 한국·룩셈부르크 간 연구기관과 기술기업이 이 원칙을 국제 공동연구에 적용한 성과다. 데이터 연합성은 두 가지 의미를 갖는다. 하나는 참여기관이 권한과 책임을 유지하면서 공통규칙에 따라 협력하는 ‘거버넌스의 연합성’이며, 다른 하나는 계산을 각 데이터 보유 환경에서 수행하는 ‘연산의 연합성’이다. 두 속성 모두 중요하다.

3. 연합학습을 포함하는 상위 개념으로서의 연합분석 방법론

연합학습(federated learning)은 분산된 데이터에서 모델의 매개변수나 업데이트를 계산하고 이를 결합하여 예측모형을 학습하는 방법이다. 대표적으로 모델평균화 접근 등이 원자료를 중앙화하지 않고도 연합학습이 가능함을 제시했다.^[2] 연합분석(federated analytics)은 공통 연구 계획 아래 분산자료의 탐색, 품질평가, 통계추정, 인과효과 분석, 모형개발과 외부검증을 수행하고 허용된 산출물만 결합하는 상위개념이다. 임상적으로 중요한 답은 종종 하나의 전역 모델을 만드는 것은 아닌 모형학습 이전의 자료특성화와 연구설계 단계에서 결정된다.

연합분석의 첫 단계는 연구가능성 평가다. 각 기관에서 동일한 선정·제외 기준으로 환자 수, 추적기간, 변수결측률과 사건 수를 산출하면 실제 데이터 반출 전에 연구의 실행가능성을 판단할 수 있다. 다음 단계인 기술통계와 역학분석에서는 공통코드로 발생률·분포·기관별 효과 추정치를 계산한다. 분산회귀에서는 필요한 요약량을 반복 교환하고, 메타분석에서는 기관별 추정치와 불확실성을 결합한다. DataSHIELD는 분석을 데이터가 있는 곳으로 보내고 비공개성을 고려한 요약 통계를 교환하는 접근을 실제 통계 분석체계로 구현한 선형 사례다.^[3]

수학적 동등성은 구체적인 조건 아래에서만 성립된다. 예컨대 동일한 변수 정의와 분석대상, 동일한 선형모형을 가정할 때 각 기관 k 가 계산한 $X_k^T X_k$ 와 $X_k^T y_k$ 의 합으로 통합 최소제곱 추정량을 구할 수 있다. 그러나 회귀 변수를 포함한 요약량은 정보 노출 위험이 발생하며, 차등프라이버시 잡음 추가나 기관별로 다른 전처리를 적용하면 중앙화 분석과 다른 결과를 얻게 된다. 로지스틱 회귀나 생존분석에서도 단순히 최종계수 평균으로 공동우도 최적화가 재현되지는 않는다. 알고리즘의 통계적 성질과 통산·보호 제약을 함께 명시해야 하는 이유다.

기관 간 이질성은 평균화로 제거할 오류가 아니라 연구대상이다. 국가별 환자구성과 치료 접근성이 다르면 기관 규모에 따른 가중평균은 대형병원의 분포를 과도하게 반영한다. 분석계획에는 목표 모집단과 추정대상(estimand), 기관별 효과, 이질성 평가, 결측자료 처리, 민감도 분석을 포함해야 한다. 모형의 성능도 전체 AUC 하나로 표현하기보다 국가·기관·성별·연령집단별 판별력과 보정도를 보고하고, 학습에 참여하지 않은 기관과 후속 시점에서 검증하는 것이 타당하다. 모델을 공유했다는 사실만으로 임상적 일반화 가능성이 입증되는 것은 아니다.

인과추론은 특히 별도의 설계가 필요하다. 치료효과 비교에서는 대상환자, 비교치료, 추적시점, 결과변수와 분석전략을 먼저 정해야 한다. 성향점수나 가중치를 각 기관에서 계산하더라도 교란 변수의 정의가 다르거나 치료선택에 관한 양성성 조건이 충족되지 않으면 통합효과 해석은 어렵다. 국가별 처방관행과 검사빈도가 만드는 선택편향, 불멸시간 편향, 경쟁위험을 검토해야 한다. 분산계산이 이러한 편향을 자동교정하지는 않는다. 연합분석의 재현성은 같은 코드의 반복실행뿐 아니라 같은 임상질문이 각 기관에서 동일하게 구현되었음을 입증함에 있다.

환자 집합이 서로 다른 수평적 연합과 동일환자의 서로 다른 정보가 흩어진 수직적 연합도 구별해야 한다. 전자에서는 표준화된 기관별 분석과 결과 결합이 중심이지만, 후자에서는 개인 단위 연계의 적법성, 일치 오류와 중복환자 처리 해결도 필요하다.

결국 분석 방식은 필요한 정보의 세밀함과 위험에 따라 선택된다. 탐색에는 승인된 집계 질의, 효과추정에는 분산통계 또는 기관별 메타분석, 예측에는 연합학습이나 독립적 외부검증이 적합할 수 있다. 하나의 HDS는 다양한 방식을 지원하면서 각 작업에 다른 접근권한과 출력 심사규칙을 부여한다. 연합학습 엔진의 도입 여부보다 연구 질문에 적합한 분석 경로를 선택하고 그 선택의 근거를 남기는 역량이 더 중요하다.

(1) **준비 단계**는 원천 데이터를 분석 가능한 연구자산으로 바꾸는 과정이다. 데이터 보유기관은 개인정보와 민감정보의 범위를 분류하고, 변수 사전과 출처, 수집 시기, 데이터 변환 이력을 정리한다. 가명처리와 직접 식별자 제거, 임상 의미의 정합성과 품질 평가도 병행한다. 원자료에서 연구용 테이블로 변환된 경로를 남겨야 오류를 발견했을 때 해당 연구와 산출물을 추적할 수 있다.

(2) **탐색 단계**는 카탈로그를 통한 발견과 실제 자료 접근을 구분한다. 연구자는 어떤 데이터가 존재하는지, 어느 기간과 환자군을 포함하는지, 어떤 제한이 적용되는지 먼저 확인한다. 공개 메타데이터를 열람하는 행위와 개인별 기록을 분석하는 행위에는 동일한 권한이 필요하지 않다. 다만 작은 코호트의 세밀한 특성을 반복 조회하면 집계 결과만으로도 특정인의 포함 여부가 드러날 수 있으므로, 탐색질의 역시 최소 집계단위와 반복질의 제한의 대상이다. ‘데이터를 보지 않는 탐색’도 무조건 무위험한 활동은 아니다.

(3) **분석 단계**에서 이동하는 대상은 승인된 작업명세와 코드이며, 실행환경은 데이터 보유기관 또는 허가된 안전처리환경이다. 실무적으로는 연구 식별자, 허가번호, 데이터 버전, 코호트 정의, 코드와 실행 이미지의 해시, 허용된 네트워크 연결, 자원 한도와 출력형식을 하나의 작업명세로

뮬을 수 있다. 각 노드는 이를 검증한 뒤 실행하고 실행이력을 남긴다. 중앙조정 기능은 작업상태와 필요 산출물을 조율하되, 원자료를 임의로 열람하거나 승인범위를 확장하는 것은 제한한다.

(4) **통제 단계**의 핵심은 결과물 반출심사다. 집계표의 작은 셀, 극단값, 자유서술문, 환자별 예측값과 학습된 모형은 서로 다른 공개위험을 가지므로 ‘검증된 결과만 이동한다’는 원칙에서 검증은 통계적 타당성과 개인정보 공개위험을 모두 포함한다. 자동필터는 반복가능한 검사를 수행하고, 경계 사례는 독립 검토자가 판단한다. 승인된 출력도 누가 어떤 목적과 조건으로 사용할 수 있는지, 재배포와 다른 데이터와의 결합이 허용되는지 함께 기록한다. 출력 게이트는 단순한 다운로드 버튼이 아니라 연구허가를 최종 산출물로 연계하는 통제지점이다.

4. Gaia-X의 신뢰체계와 분석 인프라의 구현 조건

Gaia-X의 기여는 데이터를 한곳에 모으는 데 있지 않고, 분산된 참여자와 서비스의 속성을 검증 가능한 방식으로 표현함에 있다.^[4] 의료 분야의 실제 사례인 룩셈부르크 Dataspace4Health는 2026년 공개 보고에서 HRS 병원과 LIH 사이의 Eclipse Dataspace Connector 기반 카탈로그, 정책 정의, 계약협상, 데이터전송 흐름을 실증했음을 밝혔다. HealthDCAT-AP 메타데이터와 서명된 기관의 검증가능 자격증명도 활용한다.^[5] 안전처리환경 구축은 IHDSI를 통해 구현 중이다.

연합분석에서는 기술적 상호운용성과 의미론적 상호운용성을 함께 설계해야 한다. 전자는 메시지, 인증, 작업 실행의 연결성을, 후자는 같은 변수와 사건이 같은 임상적 의미를 갖는지를 다룬다. OMOP 공통데이터모델은 분석에 필요한 구조와 표준 어휘를 공유하는 한 방법이다.^[6]

그러나 표준형식으로 변환했다고 원 자료의 임상적 차이가 마술처럼 해소되는 것은 아니다. 진단 코드 매핑상의 손실, 검사 값의 단위 변환, 처방과 실제 투약의 차이, 병기 분류체계의 버전 등을 검증해야 한다. 영상, 병리, 오믹스까지 하나의 표에 무리하게 평탄화하기보다 각 자료의 특성을 보존하고 공통 환자군, 시간축, 출처 정보를 연결하는 방식이 필요하다. 데이터 품질은 일반적인 완전성 점수만으로 충분하지 않다. 품질은 특정 연구질문에 대한 적합성으로 평가해야 한다. 카탈로그에는 단순 보유 건수뿐 아니라 갱신주기, 표준화 수준, 검증결과와 알려진 제한을 포함해야 한다. 보안과 실행의 재현성도 중대한 해결과제다.

5. IHDSI의 연구 구현과 한국의 제도화 과제

IHDSI는 국가 간 의료데이터 연합 활용체계에 대한 질문을 구체적인 국제협력 연구로 전환한 성공사례다. 공식 발표에 따르면 2025년 7월 NIH와 한국 국립암센터가 공동으로 주도하고 네이버클라우드, 오케스트로, 미국 싸이퍼룸 사가 기술 파트너로 참여했다. 초기 과제는 전향적 방광암 코호트의 임상·다중오믹스 분석과 파킨슨병 코호트를 활용한 연합 커넥티브의 타당성 검토였다. 공동운영위원회와 기술·임상 실무그룹을 통한 추진 구조도 제시됐다.^[7]

NIH와 국립암센터는 임상질문, 코호트와 검체, 연구설계 및 임상적 해석을 맡는다. 국립암센터는 네이버클라우드의 의료SI 클라우드 컴퓨팅, 오케스트라의 보안중심 클라우드 풀스택 운영환경, 싸이퍼룸의 정밀의료 분석 플랫폼으로 구성된다. 2026년 3월 31일 고양시 창조혁신캠퍼스에서 열린 NIH 한국 대표사무소 현판식과 협약은 협력의 상시 조직기반이 마련되었음을 보여준다.^[8] 고양시는 방광암 코호트의 임상자료 수집과 공동연구를 시연했다. 방광암 연구에 적용한 연합분석 절차는 먼저 근침윤성 여부, 병기와 등급, 치료선, 재발과 진행의 정의, 검체 채취 시점을 공동 프로토콜에 명시하고, 각 기관에서 코호트 추출 가능성과 결측·사건 분포를 점검하고, 자료 수집 계획을 조율했다. 이후 사전 명시된 임상질문에 대해 기관별 통계 또는 분산추정을 수행하고, 바이오마커 모형은 별도 기관 별도 시점의 자료로 검증했다. 다중오믹스 통합분석에서는 분석성능과 표본 처리과정의 확인도 필요했다. 국가나 병원별 검체 보관법, 장비, 시퀀싱 기술 등의 차이로 예측모형이 질병특성보다 기관차이를 학습할 수 있기 때문이다. 바이오마커의 효과 추정, 예측모형의 검증, 실제 진료에서의 유용성 평가는 서로 다른 검증단계다. 탐색 결과를 임상적용으로 연결하려면 사전에 정의된 외부검증과 후속 임상평가가 필요하다.

IHDSI를 통한 유럽 의료데이터공간과 한국 국립암센터 사이의 성공적 연합분석 시연은 매우 고무적이지만 한국은 이러한 연구 과정을 반복가능한 공공역량으로 만드는 제도적 보완이 필요하다. 첫째, 기관마다 다른 신청·심사·계약을 공통항목으로 정리하되 개인정보 감독과 연구윤리 심의의 고유권한을 유지해야 한다. 둘째, 데이터 관리 담당자의 지속적 자원과 표준화·품질관리 비용을 인정해야 한다. 셋째, 승인된 분석 코드와 결과의 심사기준, 사고 대응과 재심절차를 공동 운영규칙에 포함해야 한다. 이는 EU 제도의 단순 이식보다 국내 법체계에 맞춘 기능적 정합성을 확보해야 함을 의미한다.

도입순서에 연구위험과 운영성숙도 고려도 필요하다. 24,000개 유럽연합 병원의 의료데이터가 EHDS로 연합되는 2027년을 향한 구성단위별 역할, 데이터 및 메타데이터 목록, 공동 연합분석 프로토콜과 시험환경의 마련, 2029년 의무화를 향한 반복수행 가능한 허가, 분석, 반출 절차를 실제 다기관 연구 체계구현, 2031년 확장을 향한 국가바이오빅데이터 사업의 오믹스와 연구코호트의 복합조건까지 데이터 처리역량을 확보해야 한다. 다기관, 국경 간 의료데이터 공유의 성패는 필요한 연구를 어떤 권한과 조건으로 수행했으며, 그 결과를 얼마나 신뢰할 수 있는가에 달렸다. 연합분석은 탐색과 통계, 인과추론, 모형 학습과 검증을 연결하고, 데이터 공간은 이를 반복가능한 협력으로 만드는 권리, 책임, 운영 기반을 제공한다. EHDS는 단계별 법정 일정과 공적 접근절차를, Gaia-X는 분산참여자 간 신뢰를 표현, 검증하는 기술적 틀을, IHDSI는 임상질문을 중심으로 국제협력을 구현할 기회를 제공한다.^[9] 2027년의 제도 준비가 2029년의 운영 능력으로 이어지고, 2031년의 복합 데이터 확대가 더 신뢰할 수 있는 근거 생산으로 연결되어야 한다. 이를 위해서는 독립적 승인 감독, 지속 가능한 데이터 관리, 연구자와 환자에게 설명 가능한 분석 과정을 함께 구축해야 한다. 의료데이터 공간의 최종 성과는 플랫폼의 규모가 아니라, 보호받는 권리와 국제환경에서 재현가능한 의학지식의 거대한 축적이다.

참고문헌

- [1] European Parliament and Council of the European Union. (2025.3.5.). Regulation (EU) 2025/327 of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847. Official Journal of the European Union, L 2025/327. ; 세계법제정보센터. 유럽건강데이터공간(EHDS) 규정 국문 번역본.
- [2] McMahan, H. B., et al. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), PMLR 54, 1273-1282.
- [3] Gaye, A., et al. (2014). DataSHIELD: taking the analysis to the data, not the data to the analysis. International Journal of Epidemiology, 43(6), 1929-1944.
- [4] Gaia-X Association. (2023.10.). Gaia-X Architecture Document 23.10 Release.
- [5] Alborzi, S. Z. (2026.6.). Dataspace4Health: a Gaia-X Lighthouse project moving from reference architecture to working components. Gaia-X Magazine, Edition 8.
- [6] Hripcsak, G., et al. (2015). Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. Studies in Health Technology and Informatics, 216, 574-578.
- [7] Luxembourg Institute of Health. (2025.7.22.). Global partnership launches International Health Data Space Initiative (IHDSI) to accelerate AI-driven precision medicine and cancer research. ; 전자신문. (2025.7.22.). 글로벌 의료 AI 데이터 연합 출범...“임상 데이터 공유·공동 연구로 정밀의료 가속화”.
- [8] 헤럴드K. (2026.3.31.). 고양특례시, 룩셈부르크 국립보건원(LIH) 한국 대표사무소 유치. ; Luxembourg Trade & Invest. (2026.4.3.). 고양시에 룩셈부르크 국립보건원 한국 대표사무소 개소.
- [9] Kari, A., Schurig, T., & Gersch, M. (2023.11.). European Health Data Space (EHDS), Gaia-X and Health-X dataLOFT. Freie Universität Berlin.

Privacy Report

2026 개인정보 이슈 심층 분석 보고서

『2026 개인정보 이슈 심층 분석 보고서』는
개인정보보호위원회 출연금으로 수행한
사업의 결과물입니다.

한국인터넷진흥원의 승인 없이 본 보고서의 무단전재나
복제를 금하며, 인용 출처 『2026 개인정보 이슈 심층 분석
보고서』를 밝혀주시기 바랍니다.

본 보고서의 내용은
한국인터넷진흥원의 공식 견해가 아님을 알려드립니다.

발행

발행일	2026년 10월
발행처	한국인터넷진흥원 개인정보예방팀 전남광주통합특별시 나주시 진흥길 9

Privacy Report

2026 개인정보 이슈
심층 분석 보고서

2026 Vol. 4