

바이오헬스 산업 내 보건 의료데이터 현황

KDB미래전략연구소 산업기술리서치센터
김은진 연구위원(ejkim@kdb.co.kr)

- I. 보건 의료데이터 개요
- II. 보건 의료데이터 국내외 정책 및 동향
- III. 보건 의료데이터 활용
- IV. 시사점

보건 의료데이터는 질병 예방, 진단, 치료, 회복 및 의료서비스 제공 과정에서 발생하는 다양한 데이터를 의미한다. 서로 다른 데이터 속성을 가진 다양한 연원에서 수집되며, 의료기술의 발전에 따라 수집되는 정보의 양과 복잡성이 증가하고 있다. 인공지능 등 디지털 기술의 발전에 따라 보건 의료데이터의 수집 및 활용에 대한 논의가 활발히 진행되고 있다.

주요국들은 데이터 활용을 촉진하기 위해 다양한 정책을 시행 중이다. 미국은 정보 공유에 대한 환자 선택권을 보장하고, 데이터 품질 및 상호운용성을 제고하기 위한 표준을 채택하였다. 유럽연합은 국가 간 의료데이터 공유를 위한 EHDS(유럽건강데이터공간)를 도입해 개인정보 주체의 권리 강화와 데이터 표준화 등 인프라 구축에 힘쓰고 있다. 핀란드는 Findata를 통해 비식별화된 데이터를 연구 등에 활용하며, 개인정보 주체의 선택에 따른 제공 동의 및 철회를 보장한다.

우리나라도 바이오헬스 산업 내 활용을 위한 보건 의료데이터의 생산, 집적, 활용 기반을 마련 중이다. 국가 바이오 빅데이터 사업과 건강정보 고속도로 운영을 통해 개인 의료데이터의 통합 관리 및 활용을 지원하고 있다. 또한 의료데이터의 표준화와 가명처리 지침 마련 등 제도적 기반도 강화하고 있다.

보건 의료데이터의 효과적인 활용을 위해서 세 가지 과제를 제안하고자 한다. 첫째, 의료기관 간 데이터 형식 차이를 줄이기 위한 표준화와 상호운용성 확보가 필요하다. 둘째, 개인정보의 2차 활용을 위해 조건부 활용 제도 도입과 함께 사회적 수용성을 높이기 위한 공론화 및 정보공개 체계가 마련되어야 한다. 셋째, 데이터 제공자와 활용자 간의 책임과 역할을 명확히 하고, 정부, 의료기관, 산업계, 국민이 참여하는 거버넌스 체계를 구축하여 데이터 활용의 기준과 방향성을 설정해야 할 것이다.

* 본고의 내용은 집필자 견해로 당행의 공식입장이 아님

I. 보건 의료데이터 개요

1. 보건 의료데이터 정의 및 유형

- 보건 의료데이터는 건강과 관련된 정보로, 질병 예방, 진단, 치료, 회복 및 의료 서비스 제공 과정에서 발생하는 다양한 데이터를 의미
 - 「보건의료기본법」에서는 보건의료와 관련한 지식 또는 부호·숫자·문자·음성·음향·영상 등으로 표현된 모든 종류의 자료로 정의
- 다양한 연원에서 수집되는 보건 의료데이터는 정보의 양과 구조적 복잡성이 지속적 증가 경향
 - 서로 다른 데이터 속성을 가진 다양한 연원에서 수집되는 보건 의료데이터는 의료기술의 발전에 따라 수집되는 양과 복잡성이 증가하는 추세
 - 이러한 데이터는 개인, 의료기관 등에서 다양한 유형의 모니터링 장치, 센서 또는 계측기를 통해 제공
 - 보건 의료데이터는 유형별 특징 보유
 - (공공기관 데이터) 전 국민 대상, 장기간의 건강 및 진료 정보, 정형화된 코드 체계에 따라 관리하여 비교적 일관성 있는 데이터
 - (진료 데이터) 진단코드, 처방정보, 검사수치 등 정형화된 구조화 데이터와 진료 소견, 증상 서술 등 자유 서술형 텍스트인 비구조화 데이터로 구성
 - (오믹스 데이터) 생물학적 데이터를 포함하며, 분자 수준에서 개인의 생물학적 특성 규명 가능
 - (건강 및 라이프로그 데이터) 실시간성과 연속성을 가지며, 일상생활 중 생성되는 생활 기반 건강정보

〈표 1〉 보건 의료데이터 유형

구분	특징	종류
공공기관 데이터	건강보험 자료 등	· 자격 및 보험료 관련 데이터, 진료내역, 건강검진결과, 사망정보 등 공공기관의 수집·보관·관리 데이터
진료 데이터	의료기관 생성 데이터	· 전자의무기록, 전자건강기록 정보, 처방정보, 입/퇴원 기록, 의료영상자료 등 데이터
오믹스(Omics) 데이터	유전체 정보 등	· 유전체(Genome), 전사체(transcriptome), 단백질체(Proteome), 대사체(Metadome), 후성유전체(Epigenome), 지질체(Lipidome) 다양한 분자 수준에서 생성된 여러 데이터
라이프로그 (Lifelog) 데이터	개인건강정보 기반 데이터	· 웨어러블, 홈 모니터링 장치, IoT, 모바일 앱 등에서 수집된 데이터로, 한 개인의 일상생활 활동에 관한 모든 데이터(체중, 심박수, 혈당, 몸무게 등)
임상연구 데이터	의료연구기관 생성 데이터	· 의약품, 임상시험 데이터, 의료기기, 유전자 연구 데이터, 인체유래물 연구 데이터, 조사관찰 연구 데이터, 개인정보를 직·간접 활용한 연구 데이터
기기 기반 데이터	의료기기 생성 데이터	· 의료기기 및 환자 모니터링 장치 기반 데이터
앱·소셜 미디어 데이터	기타 헬스 관련 미디어 데이터	· 건강 포털, SNS에서 수집된 헬스케어 관련 데이터

자료 : 유소영(2018), 헬스케어 빅데이터 딜레마와 해결 방안

2. 데이터 통합·관리

□ 보건 의료데이터 활용을 위해서는 분산되어 있는 데이터의 통합·관리가 필수적

- 보건 의료데이터는 민감정보를 포함하고 다양한 형식으로 시스템에 분산되어 있음
 - 효과적인 활용을 위해 데이터의 표준화, 통합·관리와 함께, 민감정보 보호를 위한 보안 조치 및 활용에 대한 사회적 신뢰 확보 전제가 요구
- 민감정보 보호를 위한 기술적 보안 조치 및 윤리적 과제
 - 서로 다른 데이터 체계 간 공통데이터모델¹⁾ 형식 마련 및 암호화, 익명화 등 보안 기술 적용을 통해 조직 내 데이터 책임 및 모니터링 강화
 - 윤리적 측면에서 환자의 동의, 개인 프라이버시 보호, 평등한 접근권 등 고려 필요

1) 공통데이터모델(Common Data Model, CDM)은 서로 다른 구조로 이루어진 데이터베이스의 자료 구조를 통일된 형식으로 변환하는 자료모형

○ 사회적 합의를 위한 다양한 이해관계자의 참여 요구²⁾

- 정부, 의료기관, 연구자, 환자대표 등 다양한 이해관계자의 합의와 신뢰를 구축하여 국민 인식 제고, 개인정보보호법 및 규범 정비, 환자 동의체계 정립 등 사회적 기반 마련 필요

□ 데이터를 통합·관리하고 분석 기반 마련을 위한 보건 의료데이터 플랫폼 활용

○ 보건 의료데이터 플랫폼은 공공기관, 보건의료기관 등이 수집하는 다량의 데이터를 수집, 저장, 처리, 분석, 공유하는 시스템

- 표준화된 데이터모델을 바탕으로 엄격한 개인정보 보호 규정을 준수하며 데이터 관리
- 클라우드 인프라와 분석 도구를 활용하여 분산 분석 가능

○ 암, 난치성질환 등에 대한 정밀의료³⁾의 발전, 각국의 유전체 관련 데이터 구축 프로젝트 추진에 따라 보건 의료데이터 플랫폼의 구축과 활용 확대 추세

□ 고품질의 데이터 확보를 위한 데이터 품질 거버넌스 필요

○ 데이터 수집부터 활용에 이르는 전 과정에서 표준화 및 품질 보증 체계 마련

- 국제표준화기구(ISO)는 데이터 거버넌스를 조직 내 데이터 관리 기능이 효과적으로 실현되도록 하는 평가(Evaluation), 감독(Direction), 모니터링(Monitoring) 체계로 정의
- 세계보건기구(WHO)에서는 정확성(Accuracy), 신뢰성(Reliability), 일관성(Consistency), 완전성(Completeness), 유효성(Validity), 접근성(Accessibility)을 보건 의료데이터의 품질관리 국제 기준으로 제시

2) OECD(2023) 연구에 따르면 다수 국가에서 “사회적 합의 부족과 신뢰의 결여”가 전자의무기록 데이터 활용의 주요 장벽으로 지적

3) 정밀의료는 각 개인의 의료정보, 개인건강정보, 유전정보, 환경 및 생활습관 정보 등을 종합적으로 분석하여 특정 질병에 대해 동일 속성을 지닌 개인들을 그룹으로 나눈 뒤 분류된 그룹에 맞춰 보다 정확한 치료, 진단 및 예방법을 적용하는 것

3. 데이터 활용

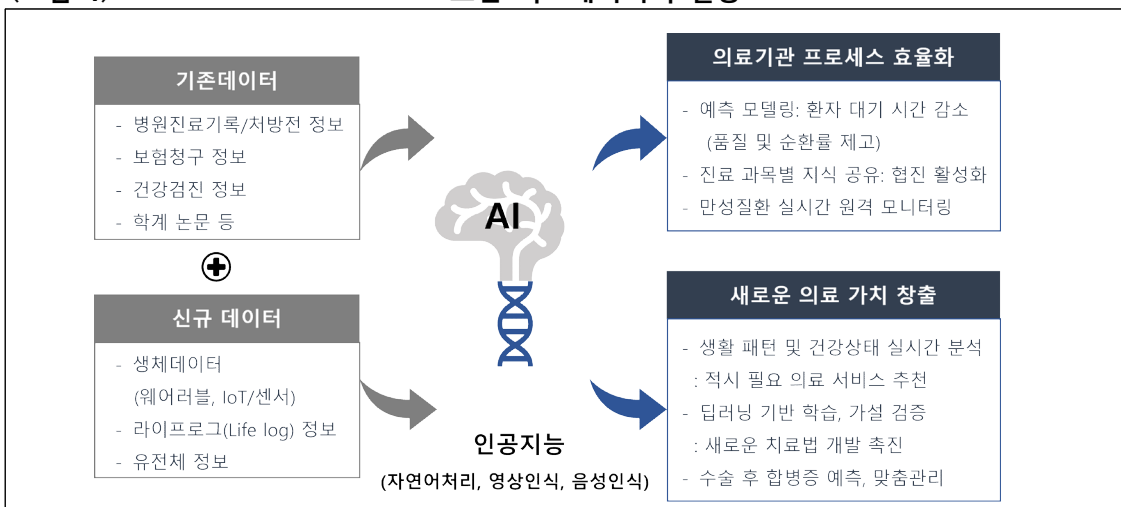
□ 바이오헬스 산업 내 보건 의료데이터의 중요성 증대

- 보건 의료데이터는 바이오헬스 산업 전반의 혁신을 견인할 수 있는 핵심 자원으로, 예방 중심의 건강관리, 정밀의료, 신약 개발 등 다양한 분야에서 폭넓은 활용 가능
- 보건 의료데이터의 연계를 통해 활용도와 가치는 급속히 증가
 - 진료기록, 유전체 정보, 생활습관 데이터 등을 통합하여 정밀한 질병 예측 및 맞춤형 치료 계획 수립
 - 공공기관 간 데이터 연계로 의료정책의 근거 확보 및 자원 배분의 효율성 향상
 - 실시간 데이터와 과거 이력의 연계를 통해 질환의 진행 상태 모니터링

□ 인공지능(AI) 기술을 활용한 고도화된 분석은 혁신적 바이오헬스 산업 가치 창출 제고

- AI 기술로 복잡하고 방대한 보건 의료데이터를 산업에 활용할 수 있는 기반 마련
 - 자연어처리, 영상인식, 거대 언어모델 등 AI 기술을 활용한 데이터 분석으로 신약개발, 진단, 임상시험 설계 등 보건의료 전반 고도화와 자동화 실현 가능

〈그림 1〉 보건 의료데이터의 활용



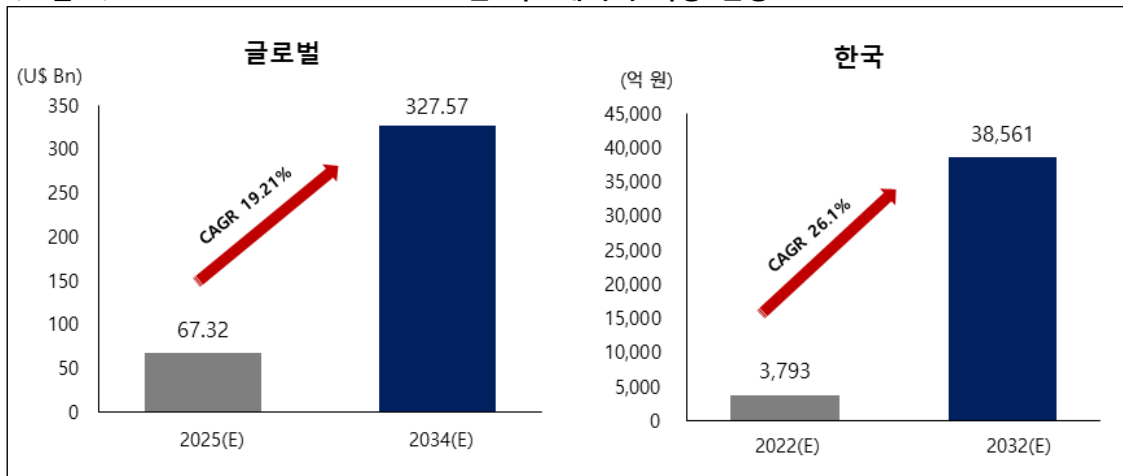
자료 : 박도휘 외 2인(2024), AI로 촉발된 헬스케어 산업의 대전환, 당행 재구성

□ 보건 의료데이터 관련 시장 규모는 개인 맞춤형 의료, 실시간 분석, 디지털 헬스케어 전환 등의 요인으로 고속 성장 전망

○ 글로벌 보건 의료데이터 시장은 2025년 673억 달러에서 2034년 3,275억 달러로 연평균 약 19.21% 성장 전망

○ 국내 보건의료 분야 데이터 분석 분야의 연평균 시장 성장률은 약 26.1%⁴⁾로, 2022년 3,793억 원에서 2032년 약 3.8조 원 규모 성장 전망

〈그림 2〉 보건 의료데이터 시장 전망



자료 : (좌)Precedence Research(2025), (우)한국보건산업진흥원(2023)

4) '2021 데이터산업 현황조사 보고서(한국데이터산업진흥원, 2022)'기준 보건의료 분야 데이터산업 시장규모 성장률

Ⅱ. 보건 의료데이터 국내·외 정책 및 동향

1. 국내 보건 의료데이터 정책 및 동향

□ 국내 보건 의료데이터 보유 현황

- (공공) 전 국민 건강보험, 인체자원, 예방접종 정보, 감염병 정보, 암발생 정보 등 방대한 보건의료 데이터 자원 보유

〈표 2〉 국내 주요기관별 공공보건 의료데이터 보유 현황

구분	보유 데이터 정보
건강보험공단	· 국민건강검진, 가입자 자격, 보험료 청구·지급내역, 진료 내역, 투약·예방접종 내역, 생활 습관, 장기요양보험, 공급자 관련 정보 (의료기관, 검진기관, 요양시설) 등 · 건강 관련 주요 빅데이터 약 3조4천억 건 (356TB) 보유
건강보험심사평가원	· 상세보험료 청구·지급 내역, 진료내역, 실시간 투약내역(DUR), 의약품 유통, 의료자원 등 · 진료현장의 상세한 데이터 약 3조건(300TB) 보유
질병관리청	· 인체자원은행, 만성질환·감염성질환·영양 등 국민건강영양조사 정보, 검역·감염병 등 공중보건에 관한 보건행정정보 다수 보유
국립암센터	· 국가 암발생 현황 통계, 암 등록정보, 암종별 레지스트리 정보, 암 유전체 정보 등 암 관련 종합적 정보 보유

자료 : 한국보건의료정보원·ABOUTGROUP(2023), 보건 의료데이터 활용 현황조사 및 분석연구

- (민간) 의료기관, 개인, 기업에서 생산하는 다양한 보건 의료데이터
 - 의료기관 내 전자의무기록시스템 구축이 보편화되어 진단명, 처방, 검사 수치, 간호기록, 검사 판독 기록, 의료 영상 등 방대한 데이터 보유
 - 임상문헌정보, 임상진료지침, 임상데이터, 의약품 제품 정보
 - 심전도, 혈당 측정기 등 의료기기에서 실시간으로 수집되는 데이터
 - 건강 관리 앱, 웨어러블 기기 등을 통해 수집되는 라이프로그

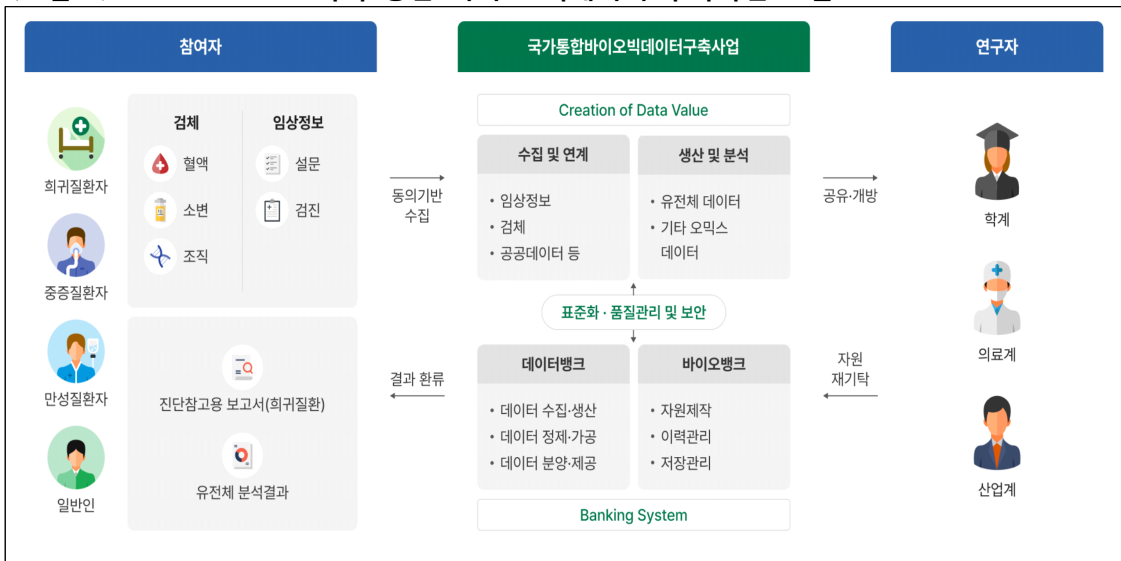
□ 2021년 보건복지부는 데이터 기반의 바이오헬스 경쟁력을 확보하고 미래의료 혁신을 추진하고자 「보건 의료데이터·인공지능 혁신전략(‘21-’25)」을 수립·시행

- 보건 의료데이터 활용을 위한 고품질 데이터 생산·개발, 플랫폼 고도화, 안전한 활용 기반 조성 등 데이터 중심의 보건의료혁신 생태계 구축

□ 관련 정책 등을 통한 보건 의료데이터 생산·집적·활용 현황

- (생산) ‘국가 통합 바이오 빅데이터 사업’으로 바이오 빅데이터 구축
 - 국가 주도 연구개발 사업으로, 참여자 검체, 임상정보, 의무기록, 공공기관 보유 데이터, 개인생성건강정보, 유전체 및 오믹스 데이터 등을 수집
 - 1단계 5개년(‘24-’28년) 간 약 77.2만 명, ‘32년까지 최종 100만 명의 임상·공공 데이터 확보⁵⁾
 - 질병의 예방, 진단 및 치료법 개발, 정밀의료, 바이오산업 혁신 연구 활용을 통해 환자·가족·사회 부담 완화 목표

<그림 3> 국가 통합 바이오 빅데이터 구축사업 흐름도



자료 : 국가통합바이오빅데이터구축사업단(2025.6.15.)

5) 100만 명 데이터에는 희귀질환자 7만, 중증질환자 22만, 일반인 71만의 임상·공공데이터와 55만 명의 전장유전체시퀀싱(Whole Genome Sequencing, WGS) 분석 자료 포함

- **(의료정보 집적·활용)** ‘건강정보 고속도로’ 사업으로 공공기관과 의료기관에 흩어진 개인의 의료정보를 ‘나의 건강기록 앱’을 통해 제공
- 여러 의료기관이 보유 중인 개인건강기록⁶⁾을 환자 본인이 열람하고, 동의 기반으로 원하는 곳에 제공할 수 있는 플랫폼
 - (데이터수집) 의료정보가 개인의 동의 하에 건강정보 고속도로 플랫폼으로 전송
 - (플랫폼) 의료기관, 연구기관, 보건의료 관련 기업 및 산업체 등에 연결되어 정보 전달
 - (활용) 건강관리에 활용, 병원 등 의료기관 및 연구기관, 보건의료 관련 기업 등 필요한 곳에 제공 활용
 - (사례) (주)메라키플레이스(나만의 닥터)는 ‘24년 과학기술정보통신부의 규제 샌드박스 지정’, ‘25년 5월 ‘건강정보 고속도로’와 연동하여 비대면진료 시 의료진과 원활한 진료 지원 기능에 활용

〈그림 4〉 건강정보 고속도로 플랫폼 및 활용 서비스



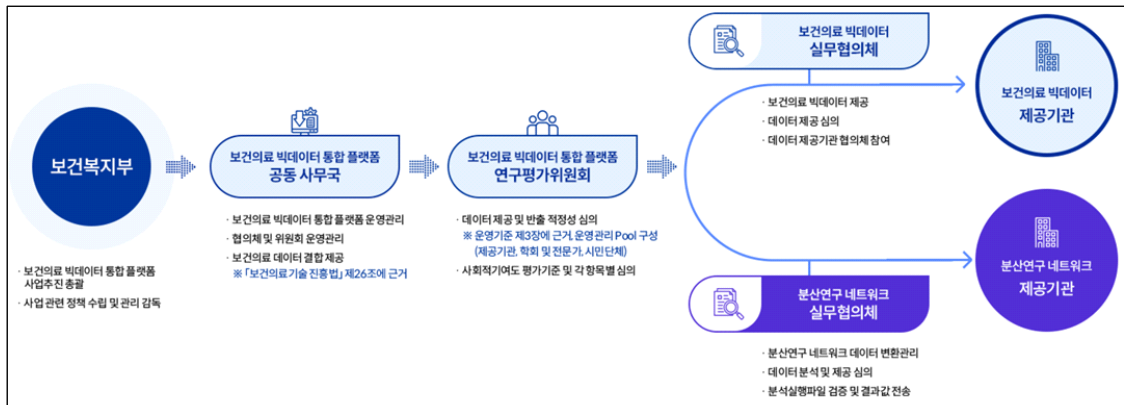
자료 : 보건복지부, 한국보건의료정보원 건강정보고속도로 포털(2025.6.15.)

6) 개인건강기록(Personal Health Record, PHR)은 개인건강과 관련한 모든 정보, 이를 바탕으로 제공되는 개인건강관리 서비스, 개인건강정보와 개인건강관리 서비스를 제공하는 플랫폼을 모두 포함하는 개념

7) 제34차 ICT규제샌드박스 심의위원회에서 건강정보 고속도로와 연계된 개인의 의료데이터를 비대면 진료 시 의사가 참고할 수 있도록 적극해석 처리

- (공공보건정보 집적·활용) 보건의료 빅데이터 통합 플랫폼을 통해 공공보건 의료데이터의 제공 및 분산연구 네트워크 활용 가능⁸⁾
 - 보건의료 빅데이터 연구평가위원회가 보건 의료데이터를 활용하고자 하는 연구과제의 연계 및 반출 적정성을 심의
 - 데이터 개방시에는 고유식별번호를 활용하며, 개별 보건 의료데이터와 제3의 신뢰기관이 발급한 랜덤키값을 결합하여 사용
 - 분산연구 네트워크에서는 원천 데이터 제공 없이 연구자의 분석실행파일 결과값만을 제공

〈그림 5〉 보건의료 빅데이터 통합 플랫폼 거버넌스 구성



자료 : 한국보건 의료정보원 보건 의료 빅데이터 통합 플랫폼(2025.6.15.)

- 의료데이터 표준화를 통한 상호운용성 제고 및 「보건 의료데이터 가이드라인」 개정으로 의료데이터 활용에 대한 제도적 기반 강화
 - 병원마다 상이한 의료데이터를 표준화하고자 데이터 교류에 필수적인 항목을 정의하여 보건 의료데이터 용어 및 전송 표준 고시('23.9.15)
 - AI 기술 성장으로 활용이 증가하고 있는 영상, 텍스트 등 비정형 의료데이터의 가명 처리 방법과 절차를 구체화하는 등 「보건 의료데이터 가이드라인」 개정('24.12)

8) 질병관리청, 통계청, 국립장기조직혈액관리원, 국립재활원, 건강보험심사평가원, 국립암센터, 국립 중앙의료원, 국민건강보험공단, 국민건강보험 일산병원 9개 기관의 데이터를 개인 단위로 결합, 공공의 목적 연구에 활용할 수 있도록 연구자에게 개방

2. 해외 보건 의료데이터 정책 및 동향

1 미국

□ 미국의 보건 의료데이터 정책은 환자의 개인정보 보호, 데이터 상호운용성 제고, 데이터 활용 촉진을 중심으로 구성

- (개인정보 보호) HIPAA⁹⁾ 개인정보 보호 규정을 통해 개인의 의료 기록 및 개인 식별 건강정보 보호
 - 건강보험 관련 기관, 의료정보 중계기관, 의료 제공자 등 기관에 적용
 - 개인 건강정보 보호를 위한 안전장치 마련 의무, 개인 동의 없는 정보 이용 및 제공 제한, 환자의 정보 자기결정권 보장을 주요 내용으로 함
- (상호운용성 및 표준화) 국제전송기술표준 FHIR¹⁰⁾을 채택하여 다양한 형태의 건강정보 데이터 구현 및 폭넓은 상호운용성 지원
 - FHIR 표준방식은 데이터 교류 시 리소스 단위로 교류를 하여 전송이 빠르고 효과적이며, 모바일·클라우드 등 다양한 플랫폼에 대응 가능
 - 표준화된 데이터 형식, 상호운용성, 환자 동의 관리로 데이터 교환 시 개인정보 보호 강화 등 보건 의료데이터 교환과 관리의 효율성 향상
- (건강정보교환) 의료기관은 환자의 건강 정보 전자적 교환 가능, 환자에게 정보 공유에 대한 선택권 제공
 - 환자 전체 병력의 추적과 공유로 의료진의 진단 및 의료기관 관리 업무 효율성 강화, 비용 절감
 - 환자 데이터 추세 파악을 통한 전염병 확산 등 질병 발생 모니터링, 새로운 의료 기술 및 서비스의 적용·확산, 보건의료 기술 인프라 기반 제공 등 활용 가능

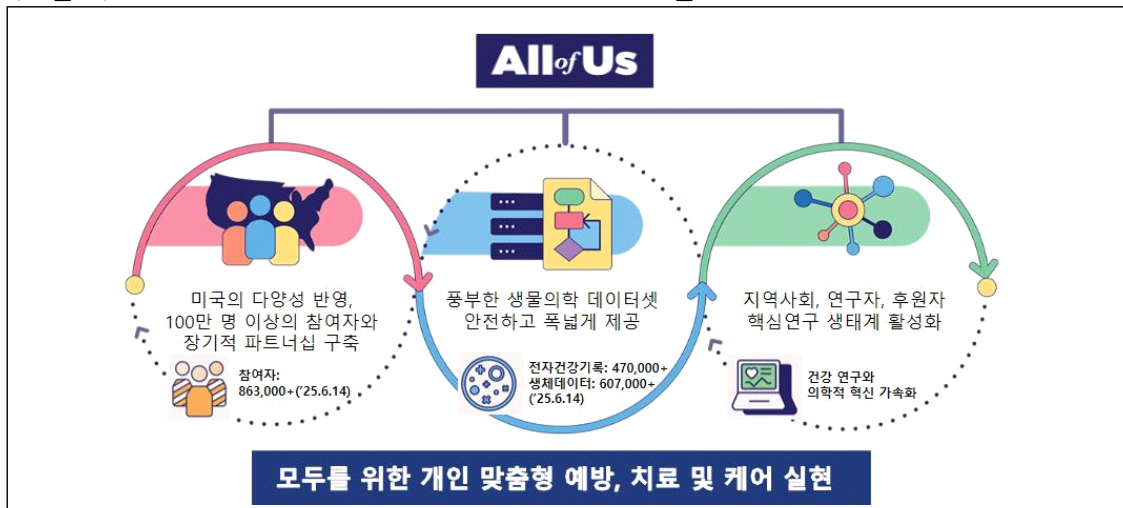
9) 건강보험 이전과 책임에 관한 법(Health Insurance Portability and Accountability Act, HIPAA)은 1996년 미국 연방 정부에서 제정된 환자 개인정보보호의 근간이 되는 법

10) 빠른 의료 상호 운용성 리소스(Fast Healthcare Interoperability Resources, FHIR)는 차세대 국제 전송표준체계로, 기존 문서단위 교환 표준 방식의 단점을 보완, 앱·클라우드 등 다양한 환경에서 적용 가능한 차세대 전송 표준

□ 미국은 정밀의료 이니셔티브, 블루버튼 이니셔티브 등을 통해 보건 의료데이터 구축·활용

- (정밀의료 이니셔티브) 2015년부터 개인 맞춤형 치료와 예방을 목표로 정밀의료 계획추진
 - (단기적 목표) 암 유전체 연구의 확장을 통해 다양한 암의 예방 및 치료법 개발, 정밀 암 치료를 위한 임상시험과 신약 개발 지원
 - (장기적 목표) 정밀의료 연구범위를 모든 질환으로 확대, 2026년까지 100만 명 코호트 구축 통해 전자건강기록, 생체데이터를 수집하는 ‘All of Us’ 프로그램 운영

<그림 6> ‘All of Us’ 프로그램

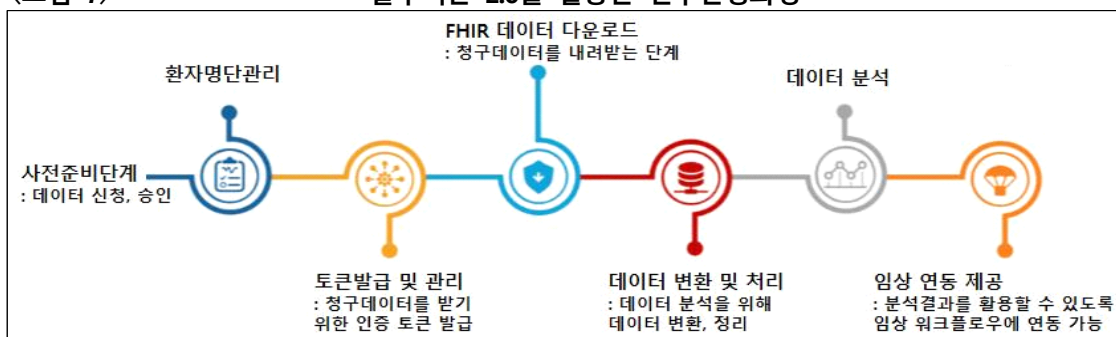


자료 : National Institute of Health, All of Us Research Program, 당행 재구성

- (블루버튼 이니셔티브) 개인의 건강의료데이터 직접 관리 및 공유 지원 캠페인
 - 2010년 다운로드 기반 도입, 2018년 블루버튼 2.0 출시를 통해 FHIR 기반 API로 전환
 - 메디케어, 메디케이드, 민간보험가입자, 의료기관 이용자 대상 복용 중인 약, 알레르기, 진료 정보, 검사 결과, 건강검진 정보, 건강보험 청구정보 등 데이터 제공
 - iBlueBotton 등 모바일 앱 출시 등 앱 연동 활성화, 표준화 및 상호운용성 강화 등 확장 노력 지속

<그림 7>

블루버튼 2.0을 활용한 연구진행과정



자료 : ARCADIA CareJourney(2020), 당행 재구성

- (기타) 국립보건정보기술조정국은 정밀의료 이니셔티브 지원을 위한 Sync for Science, Sync for Genes 운영
 - 전자의무기록 데이터, 유전체 데이터 등 표준화된 데이터 교환, 환자 중심의 데이터 활용, 정밀의료 지원 목표

<표 3>

Sync for Science와 Sync for Genes

프로젝트	기간	주요 목적	내용
Sync for Science (S4S)	2016-현재	· 개인의 전자의무기록을 연구자에게 안전하게 제공	· 개인이 자발적으로 건강데이터를 연구 목적으로 제공 · 데이터 호환성과 보안 보장을 위한 기술 표준 기반
Sync for Genes	2017-2023	· 유전체 데이터의 표준화된 공유 구현	· 정밀의료와 유전체 기반 진료·연구의 핵심 인프라 역할

자료 : Sync for Science 웹사이트, ASTP(2023), Sync for Genes 참고, 당행 재구성

2 유럽연합

□ 유럽 전역에서 유전체학 및 임상 데이터를 안전하게 공유, 연구에 활용하기 위해 '1+Million Genomes Initiatives' 운영

- 2018년 유럽 집행위원회 주도로 시작되어, 2027년까지 유럽 내 100만 명 이상 유전체 데이터 수집 및 공동 활용을 목표로 진행 중
 - (참여국) EU 회원국 25개국과 노르웨이, 영국 등을 포함하며, 각국 보건당국, 연구기관, 병원, IT 인프라 제공자 등이 협력
 - (기대 효과) 맞춤형 치료 구현 기반 강화, 정확한 진단 및 예방, 보건의료 비용 절감 및 유럽 내 의료데이터 주권 확보, 글로벌 유전체 연구 및 데이터 공유 표준 선도
- 법적, 기술적, 운영적 요소 전반을 고려하여 2단계 로드맵 구성
 - (1단계) 2018년-2022년, 국가 간 데이터 접근을 위한 공통 법·지침 수립, 데이터 표준화, 분산형 데이터 인프라 초기 설계, 유전체 데이터 품질 지침 초안 제시
 - (2단계) 2023년-2027년, 인프라의 실제 구형 및 시범 운영, 고품질 유전체 데이터 추가 생산 및 연계 강화 등

〈그림 8〉 유럽연합 '1+ Million Genomes' 로드맵



자료 : EU(2023), 1+MG Roadmap 2023-2027, 당행 재구성

- 개인 맞춤형의료 구현과 의료데이터의 초국가적 연계를 목표로 유전체 데이터 기반 프로젝트 단계적 추진

〈표 4〉 유럽연합 유전체 데이터 기반 프로젝트

프로젝트명	기간	주요 내용
1+Million Genomes Initiative (1+MG)	2020-2022	· 100만 명 이상 유전체 수집 목표 · 분산된 국가 인프라를 연계한 유럽연합 전체의 데이터 연동 기반 마련 · 개인 맞춤형의학 적용 및 유전체 데이터 활용 기반 구축
Beyond 1 Million Genomes (B1MG)	2020.6-2023.10	· 1+MG 이니셔티브 실행 지원을 위한 설계 및 테스트 프로젝트 · 유럽연합 Horizon 2020 자금 지원 · 기술 프레임워크 및 구현 전략 마련
Genomic Data Infrastructure (GDI)	2022.11-2026.10	· 24개국 참여 · B1MG 결과를 기반으로 실질적인 유럽 유전체 데이터 인프라 구축 · 유럽 데이터 공간(EHDS) 연계 · 1+MG 실행 기반 강화 및 운영 지속성 확보
Beyond 1 Million Genomes Plus (B1MG+)	2025.2-2028.1	· B1MG 후속으로 국가 인프라 성숙도 향상/상호운용성 제고/지속 가능성 촉진 · 1+MG의 실행 확대 및 지원 · 새로운 데이터 활용 구조와 정책 연계 강화 예정
Genome of Europe (GoE)	2024.10-2028.3	· 27개국 참여 · 10만 개 Whole Genome Sequences(WGS)* 수집 · 범유럽 유전체 참조 데이터베이스 구축 · 다유전자 위험점수(PRS)** 활용 가능성 포함

* : 전장유전체, 한 종의 생물체에 속하는 모든 유전적 정보

유전적 원인 규명을 통한 질병 연구 및 예방, 유전자치료 및 치료법 개발 등 여러 분야에서 활용

** : 다유전자 위험점수, 전장유전체분석에서 발견된 변이들의 위험도를 합산하여 질환의 위험도를 예측하는 방법

자료: European Commission(2020), 1+Million Genomes Roadmap 2020-2022, European Commission-Beyond One Million Genomes(2023), 1+MG Roadmap 2023-2027 참고, 당행 작성

□ 2025년 3월 의료데이터 관리 및 공유·활용을 위한 플랫폼인 European Health Data Space(EHDS) 발효

- (목적) 전자의료데이터의 안전한 공유와 활용을 통한 환자 중심의 의료 서비스 개선과 관련 연구 혁신 촉진
- (법적·기술적 기반) 일반개인정보보호규정(GDPR)¹¹⁾, 데이터거버넌스법¹²⁾, 네트워크 및 정보시스템(NIS2) 규정¹³⁾ 기반
- (주요기능) 데이터에 대한 개인 권한 강화, 데이터의 2차 활용, 표준화 및 상호운용성 확보, 거버넌스 및 인프라 구축 등

〈표 5〉

EHDS 핵심 기능

기능	주요 내용
개인 권한 강화	· 개인의 전자건강기록에 디지털 방식으로 접근 가능 · 개인은 데이터를 다른 EU 국가의 의료진과 공유, 접근 제한·허가 가능
데이터 2차 활용	· 익명화된 데이터는 연구, 공공보건, 정책 수립 목적으로 활용 가능 · 활용 승인은 각국의 국가 데이터 접근 관리 기관(National Data Access Bodies)이 담당
표준화 및 상호운용성	· 전자건강기록의 기술적 표준화 및 공통 규칙 수립 · EU 전역에서 데이터 호환성과 보안성 강화
거버넌스 체계 구축	· 각 국은 EHDS 이행을 위한 디지털 보건당국 및 데이터 접근 관리 기관 설립 · EU 차원의 규정 준수 및 협력 체계 운영
인프라 구축	· MyHealth@EU, 국가 간 의료정보 교환 플랫폼 · HealthData@EU, 공공연구 및 정책 활용을 위한 데이터 공유 네트워크

자료 : The European Health Data Space(EHDS) 웹사이트 참고, 당행 작성

- 11) 일반정보보호규정(General Data Protection Regulation, GDPR)은 2018년 5월 시행된 유럽연합의 개인정보보호법
- 12) 데이터거버넌스법(Data Governance Act, DGA)은 2023년 9월 시행된 데이터 공유에 대한 신뢰를 기반으로 한 유럽 데이터 거버넌스로, 사회·경제적 이익 목적으로 회원국·부문 간 데이터를 공유할 수 있는 안전한 환경 조성을 목표로 함
- 13) 네트워크 및 정보시스템 규정(Network and Information Security 2, NIS2)은 유럽연합 회원국 공통의 사이버 보안 수준 구축을 목표로 함

3 핀란드¹⁴⁾

□ 2019년 「Act on the Secondary Use of Health and Social Data(의료 및 사회보장 데이터 2차 활용에 관한 법률)」 제정에 따라 보건 의료데이터 활용

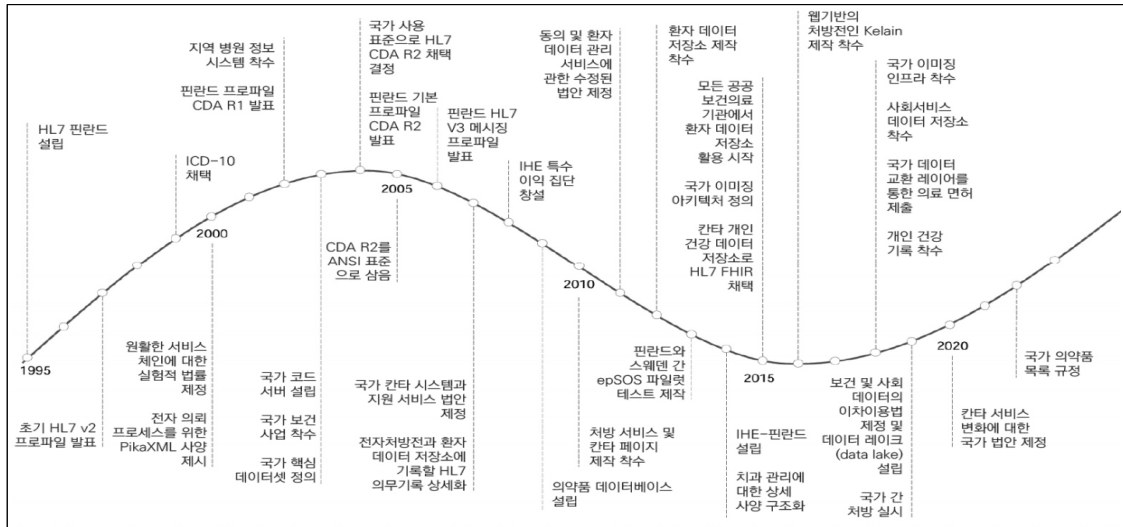
- 2007년 ‘국가e헬스 로드맵’을 통한 환자 데이터 디지털화 및 인프라 구축 추진¹⁵⁾
 - 2008년 정부 주도의 의무기록 통합시스템인 칸타(OmaKanta) 서비스 도입 후, 1950년대부터 축적된 보건의료 정보 저장·관리
- 의료 및 사회복지 서비스로 생성된 등록 데이터를 1차 수집·보관 목적 외 과학적 연구·통계작성·교육 등 2차 목적 활용 가능
- 데이터 활용 허가기관인 Findata¹⁶⁾를 통해 산재해 있는 관련 데이터 연계·제공
 - 데이터 요청 시 여러 기관의 데이터를 수집, 익명 처리 후 결합하여 제공하는 데이터 관리시스템 운영
- Findata는 개인정보 및 사생활 보호를 위해 최소한의 필요한 데이터 제공
 - 예외적 경우를 제외하고 개인정보 보호 및 사생활 보호를 위한 데이터 비식별화
 - 개인정보는 범위·유형·이용 대상 등에 대한 선택적 개인동의(opt-in) 및 철회(opt-out) 보장으로 정보주체 통제권 강화와 개인정보 활용 기회 확보

14) 핀란드는 유럽 국가 중 최초로 건강데이터의 2차적 활용에 관한 법 제정을 통해 보건 의료 데이터의 활용 체계를 선도적으로 구축함에 따라 별도 사례로 제시

15) 공공부문이 우세한 보건의료전달체계로 국가 주도의 보건의료 디지털화의 수용도가 높은 구조적 특성 보유

16) 보건복지부 산하 핀란드 국립보건복지연구원(Finnish Institute for Health and Welfare, THL)의 독립된 부서

〈그림 9〉 핀란드 보건 의료데이터 표준화 및 상호운용성 발전 타임라인



자료 : 손슬기, 박나영(2024), 핀란드와 호주의 보건 의료 마이데이터 도입 및 활용

□ 칸타 서비스와 핀젠 연구 프로젝트(FinnGen Research Project)를 통해 보건 의료데이터 구축

- 칸타 서비스는 핀란드 정부가 추진하는 의료데이터 통합관리 및 활용 플랫폼
 - 건강·의료·사회복지 데이터의 안전한 저장·관리·공유 지원 시스템

〈표 6〉 칸타 서비스의 주요 특징과 기능

구분	내용
보유 데이터	· 처방 및 의약품 데이터, 진료기록, 영상기록, 과거 기록 등 환자 데이터 · 동의서, 사전연명의료의향서, 장기기증서약서, 예방접종 기록 · 스마트기기 연계 개인건강정보
표준화	· 국제 표준 FHIR 적용하여 데이터 호환성 및 연계성 강화
데이터 공유	· 의료 및 사회복지 서비스 제공자 간 데이터 접근 및 활용 가능 · 이용자 동의 시 활용을 위한 데이터 전송 가능
환자 중심 통제	· 데이터 접근·공유에 대해 동의/거부 선택 가능, 동의 상태 상시 변경 가능 · 개인정보 보호 및 비식별화 처리
활용	· 연구, 개발, 통계 등 공익적 목적에 한하여 Findata 승인 후 활용 가능
이용자 편의성	· 환자가 웹/앱을 통해 직접 건강정보 열람 및 관리 · 검사 결과 확인, 처방전 갱신, 장기기증동의 등 다양한 기능 제공
연계성	· 일부 유럽 국가와 전자처방전 연계 가능

자료 : Kanta 웹사이트 등 참고, 당행 작성

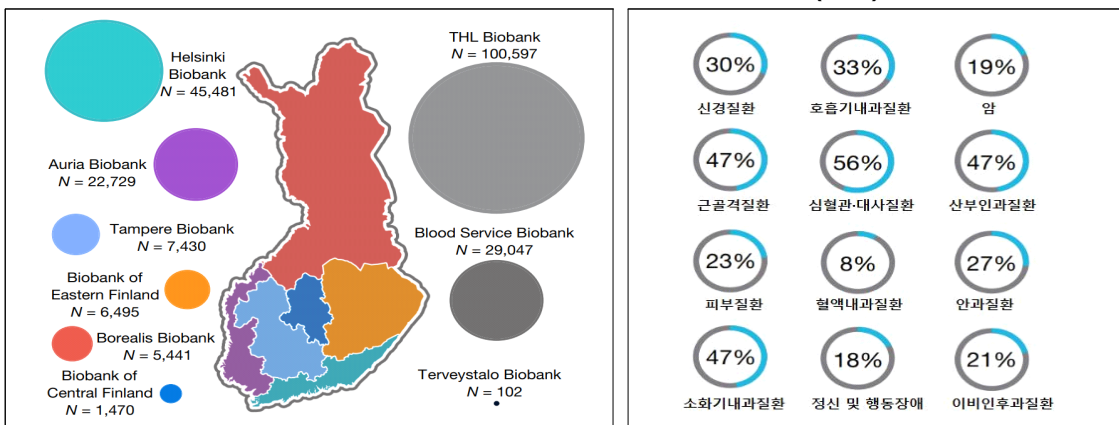
- 핀젠 연구 프로젝트는 3단계로 진행되는 민관 파트너십¹⁷⁾ 유전자 데이터 기반 프로젝트
 - (목적) 핀란드 국민 50만 명의 유전체 데이터와 건강데이터 수집·분석을 통해 유전 질환의 연관성과 생물학적 질병 기전 이해 및 개선된 치료법 개발
 - (활용) 연구, 신약개발, 질병 예측, 맞춤형 치료 등 다양한 의료·과학 연구 활용, 2024년 4월 기준 1,100여 편의 논문과 다수의 특허 산출

〈표 7〉 핀젠 연구 프로젝트 단계별 주요 내용

단계	기간	주요내용
FinnGen 1	2017.8-2020.8	· 인프라 구축 및 대규모 데이터(유전체, 건강정보) 수집 시작 · 핀란드 바이오뱅크와 보건 의료데이터에서 샘플 및 데이터 수집 · 데이터 수집·유전체 분석의 기반 마련
FinnGen 2	2020.8-2023.8	· 데이터 수집 및 인프라 확장 계속 · GWAS, PheWAS 등 유전체-질병 연관분석 본격화 · 분석 결과를 연구자 커뮤니티에 제공
FinnGen 3	2023.8-2027 (예정)	· 기존 데이터 기반 연구에서 질병 기전 및 진행과정 심층 연구로 전환 · 임상데이터 및 오믹스 데이터 통합 · 질병 진행 및 치료 반응, 장기 추적 분석 · 신약 개발, 정밀의료 실현을 위한 기초 연구 강화

자료 : FinnGen 웹사이트 등 참고, 당행 작성

〈그림 10〉 핀젠 연구 프로젝트 지역별 수집 샘플 수 〈그림 11〉 샘플(일부)의 주요질환 유병률



자료 : FINNGEN 홈페이지

자료 : Mitja IK et al.(2023), 당행 재작성

17) 핀란드 정부, 대학, 의료기관, 글로벌 제약사 등 다양한 기관이 참여, 조정기관(coordinating organization)은 헬싱키 대학교, 재정은 비즈니스 핀란드(Business Finland), 재정지원 연구파트너는 글로벌제약사(애브비, 아스트라제네카, 바이오젠, 베링거인겔하임, 세엘진/브리스톨마이어스스콧, 제넨텍, 글락소스미스클라인, 얀센, 메이즈테라퓨틱스, 머크, 노바티스, 화이자, 사노피)가 담당

Ⅲ. 보건 의료데이터 활용

1. 의료서비스 측면

□ 데이터 기반 진료, 질병 예측, 맞춤형 치료를 통한 환자 만족도 제고 및 의료비 절감

○ 데이터 기반 헬스케어를 통한 환자 중심 서비스 강화

- 전자의무기록, 유전체 정보, 개인건강데이터 등 활용을 통한 의료진-환자 간 소통 개선, 건강 상태에 대한 환자의 이해도 향상
- 환자 특성에 맞는 맞춤형 교육 자료 및 치료 정보 제공을 통한 환자 진료 경험의 질과 환자 만족도 제고

○ 보건 의료데이터 분석을 통한 의료비 절감

- 보건 의료데이터 통합·분석을 통한 질병의 조기 예방과 고위험군 환자의 선제적 관리, 의료 자원의 효율적 배분과 불필요한 의료비 지출 절감
- 데이터 기반 임상 의사결정 지원 시스템을 통한 불필요한 검사, 부적절한 약물 처방, 비효율적인 치료 경로 사전 차단으로 진료의 질 제고 및 비용 절감

○ 예측 분석 기반 질병 예방

- AI와 머신러닝 기반 예측 모델 활용으로 고위험 환자 조기 식별, 예방 중심의 맞춤형 건강관리 전략 수립
- 암, 알츠하이머병, 심혈관 질환 등 주요 질환에 대한 빠르고 정확한 AI 진단 도구를 통한 조기 치료로 치료비 절감과 생존율 향상에 기여

○ 정밀의학 기반 맞춤 치료 제공

- 유전체 정보, 단백질체 및 대사체 분석, 환경적·생활 습관 요인, 과거 치료 반응 데이터의 통합 분석으로 환자 개인의 특성을 고려한 맞춤형 치료 가능
- 정밀의학적 접근은 치료 효과 극대화 및 진료 효율화로 임상 효과 향상과 함께 의료비 절감 효과 실현 가능

〈표 8〉 의료서비스 측면에서 보건 의료데이터 활용 사례

기관	사례	주요 효과
보건복지부	의료데이터중심병원 OMOP CDM 구축	· 43개 의료기관 참여 · 암, 순환기 질환 등 질환 특화 데이터베이스 구축 · 최근 5년간 데이터 활용 연구 751건 수행
Lunit+ Massachusetts General Hospital(美)	AI 기반 폐결절 검출 시스템	· 폐결절 검출 94% 민감도, 83% 특이도 달성 · AI알고리즘을 보조 판독으로 활용 가능 증명
Geisinger Health(美)	Mycode Initiative 유전체 스크리닝	· 350,000명 참여, 5,000명 이상에게 임상적으로 유용한 유전체 결과 제공 · BRCA1/BRCA2*, Lynch 증후군** 등 35개 건강 상태에 대한 위험 평가
다기관 연구(美)***	빠른 전장유전체 시퀀싱(rWGS)	· 소아 환자 진단 시간, 입원 기간 단축 · 불필요한 검사 감소 · 환자당 \$16,730~\$28,061 절약 효과

* : 유전성 유방암과 난소암의 발생 위험과 관련된 대표적인 유전자

** : 유전성 비용종성 대장직장암, 대장암, 자궁내막암, 난소암, 위암 등 다양한 암의 발생 위험을 증가시키는 유전성
암 증후군

*** : Project Baby Bear

자료 : Christy M et al.(2025), Hyunsuk Y et al.(2020), Juliann M et al.(2025) 등 참고, 당행 작성

〈그림 12〉 의료서비스 측면에서 보건 의료데이터 활용 효과



자료 : 당행 작성

2. 산업적 측면

□ 최근 의약품, 의료기기 등 산업 내 보건 의료데이터를 활용한 실사용데이터(RWD)¹⁸⁾/실사용증거(RWE)¹⁹⁾ 사용 장려

- (미국) ‘21세기 치료법(2016)’ 기반 ‘RWE 프레임워크(2018)’를 통해 RWE를 의약품·의료기기 등 인허가 및 사후관리에 공식 활용
 - 미국 식품의약품(FDA)은 RWD/RWE를 활용해 의약품의 시판 후 안전성 모니터링 및 효과성 평가에 사용, 최근 데이터 접근성과 분석기술의 발전으로 활용 가능성 확대
 - 세계 최대의 다기관 분산형 의료데이터 네트워크²⁰⁾인 Sentinel Initiative²¹⁾를 통해 신약 등 의약품의 이상사례, 장기적 효과, 안전성 이슈 등을 신속하게 탐지·평가
- (유럽) ‘RWE 가이드(2024)’ 발간, DARWIN EU 네트워크 시스템 등을 규제에 활용
 - 유럽의약품청(EMA)은 전자의무기록, 보험청구자료, 병원·등록데이터 등 RWD/RWE를 규제 의사결정의 주요 근거로 적극 활용 중
 - 약물의 사용 실태 분석, 이상반응 발생률 평가, 인과관계 평가, 질병 분석 및 위험 예측 모델 개발 등 활용
 - DARWIN EU® 네트워크는 EU 전역의 연합형 RWD 분석 네트워크로, 병원, 바이오뱅크 등 다양한 데이터 파트너를 확보하여 분석에 활용
- (한국) 보건 의료데이터의 안전한 활용과 신뢰성 확보를 위한 정책 추진
 - ‘의료기기 RWE 가이드라인(2023)’, ‘디지털의료제품법(2025년 제정)’을 통해 RWE 기반 허가 심사 절차 명확화
 - 건강보험심사평가원은 약제 평가에 RWE를 활용하기 위한 가이드라인 마련 착수

18) 실사용데이터(RWD, Real World Data)는 실제 의료 환경에서 발생하는 환자의 진료 정보, 처방 기록, 전자의무기록, 보험청구데이터 등 다양한 출처에서 수집된 데이터로, 임상시험의 제한된 환경과는 달리 실제 환자의 복합적인 건강 상태를 반영

19) 실사용증거(RWE, Real World Evidence)는 RWD를 가공·분석하여 도출된 의약품·의료기기의 사용 및 잠재적 이익 또는 위험에 관한 임상적 증거

20) 보험청구데이터, 전자의무기록, 질병 레지스트리 등 다양한 실사용데이터를 각 데이터 파트너 기관이 보유·관리하며, FDA가 표준화된 방식으로 데이터 질의·분석

21) 미국 FDA가 의약품, 백신, 생물의약품, 의료기기 등 규제 대상 의료제품의 안전성을 실시간 모니터링하기 위해 구축한 국가 단위의 전자시스템

□ 신약개발 전주기에서 비용절감과 성공률 제고를 이끄는 핵심 전략으로 활용

- 인공지능, 연합학습, 가상임상 등의 기술과 융합, RWD/RWE 활용 촉진으로 보건 의료 데이터의 신약개발 활용 급속 확대
 - 표적 발굴, 임상시험 설계, 약물 재창출, 사후 안전관리, 규제 대응 등 신약개발의 주요 단계에서 보건 의료데이터 활용

〈표 9〉 신약개발 보건 의료데이터 활용 분야

활용 분야	활용 방식	의의
표적 발굴	· 유전체임상 데이터 기반 표적 구조 도출 · AI 기반 후보 물질 예측	· 후보물질 탐색 기간 단축, 정확도 향상
임상시험 설계	· 전자의무기록 기반 환자 선별 · 가상 대조군, 분산형 임상	· 빠르고 정확도 높은 환자 선별 · 개발 기간 단축 · 임상시험 비용 감소 및 환자 편익성 증대
약물 재창출	· 실사용데이터, 유전체 정보 등을 통한 새로운 적응증 발굴	· 개발 비용 절감 및 기간 단축
사후 안전관리	· 실사용데이터 기반 이상반응 조기 탐지 · 장기 추적	· 승인 이후 이상사례 효능 모니터링 강화
규제대응	· 실사용데이터 기반 약물효과 입증 허가자료로 활용	· 시판 승인 기간 단축

자료 : 한국보건의료정보원 · ABOUTGROUP(2023), FDA(2025) 등 참고, 당행 작성

- 보건 의료데이터는 신약의 적응증 확대, 임상 가이드라인 반영, 시판 후 안전성 모니터링 등에 다양하게 활용

〈표 10〉 신약개발 보건 의료데이터의 실사용증거 활용 사례

제약사	의약품	활용 방법
Pfizer	IBRANCE	· 전자의무기록 활용, 남성 유방암 환자에 대한 적응증 확대
Novartis	Entresto	· 심부전 환자 레지스트리, 관찰연구 · 실제 효과, 안전성 입증으로 임상 가이드라인 반영
AstraZeneca	TAGRISSO	· 비소세포폐암 환자 실사용데이터 활용 · 장기 생존율, 중추신경계 전이 효과 등 다양한 임상적 가치 입증

자료 : Quanticate(2024) 참고, 당행 작성

□ 보건 의료데이터는 의료기기 산업의 설계부터 사후관리까지 전 단계에서 활용되며 개발·생산 효율성과 품질 향상에 기여

- (설계) 환자건강정보·시장 피드백 등 활용 제품 요구사항 도출 및 리스크 사전예측
- (임상시험) 실사용데이터(RWD), 디지털트윈 등을 통해 시험 기간·비용 절감 및 실제데이터 반영
- (생산) 예측정비, 공정 최적화 등 IoT·AI 기반 스마트제조, 데이터 기반 제조설비 관리로 생산성 향상
- (사후관리) 원격모니터링 및 실시간 환자 데이터 분석을 통한 제품 성능 개선

〈표 11〉 의료기기 산업 보건 의료데이터 활용 사례

기업/제품	활용데이터	의미
Medtronic/ BioButton®	· 다중매개변수 웨어러블 기기 · 일 최대 1,440개의 생체신호 측정값 수집 (심박·호흡수·피부온도 등)	· 지속적·연속적 생체데이터 수집으로 연속적 건강 관리 · AI 및 예측 알고리즘으로 재입원 위험 예측 및 조기 개입을 통한 의료비 절감 · 임상 의사결정 지원을 위한 데이터 기반 플랫폼 기능
삼성메디슨/ HERA Z20	· AI 기반 산부인과 초음파 시스템 실시간 AI 자동분석 · 3D 영상 분석	· AI 기능으로 일관된 결과 도출, 진단 정확도 향상 · 실사용 데이터 기반 피드백으로 진단 알고리즘 정교화 및 기능 업데이트
Medtronic, NVIDIA/ GI Genius™	· AI 기반 대장내경 보조 시스템 · NVIDIA Clara™/IGX 엣지 플랫폼 도입 (데이터 실시간 처리 및 알고리즘 추론 가속화)	· 방대한 영상데이터 AI 훈련으로 진단 정확도 향상 · 임상평가 단계에서 RWE 활용으로 개발 리스크 감소 · FDA 가속 승인 · NVIDIA 플랫폼 통합하여 추가 AI 기능 확장 가능 · 실사용 데이터 기반 업데이트 및 사후관리 체계 강화

자료 : Weller GB et al.(2024), NVIDIA(2023), James AD et al.(2022) 등 참고, 당행 작성

IV. 시사점

□ 인공지능 등 디지털 기술의 발전에 따라 보건 의료데이터의 중요성 강조 및 데이터 수집 및 활용 논의 활성화

- 보건 의료데이터는 질병 예방, 진단, 치료, 회복 및 의료서비스 제공 과정에서 발생하는 다양한 데이터를 의미
 - 서로 다른 데이터 속성을 가진 다양한 연원에서 수집되는 보건 의료데이터는 일상적으로 수집되는 양과 복잡성이 증가하는 추세
- 보건 의료데이터 활용을 위해서는 분산되어 있는 데이터의 통합·관리가 필수적
 - 데이터를 통합하기 위해서는 민감정보의 특성 상 데이터 주체에 대한 신뢰 확보, 철저한 보안 및 이해관계자의 참여를 통한 사회적 합의 필요
 - 일관성 있고 신뢰할 수 있는 데이터 확보를 위한 데이터 품질관리 체계 필요
- 보건 의료데이터는 바이오헬스 산업 전반의 혁신을 견인할 수 있는 핵심 자원으로, 예방 중심의 건강관리, 정밀의료, 신약 개발 등 다양한 분야에서 폭넓은 활용 가능

□ 주요국에서는 보건 의료데이터를 구축, 활용하기 위해 다양한 정책 시행 중

〈표 12〉 미국, 핀란드, 유럽연합의 보건 의료데이터 구축·활용 정책

국가	주요 내용
미국	· 개인정보보호, 데이터 상호운용성, 데이터 활용 촉진 중심 정책 구성 · 환자의 정보 공유 선택권 보장 · FHIR 표준 채택으로 폭넓은 상호운용성 지원 · 정밀의료·블루버튼 이니셔티브로 데이터 구축·활용
유럽연합	· 유전체학·임상데이터 등 초국가적 의료데이터 공유·연구 목적 프로젝트 운영 · 개인 맞춤형 의학 실현 및 데이터 연계를 위한 단계적 프로젝트 추진 · 2025년 3월부터 EHDS(유럽건강데이터공간) 발효 → 개인 권한 강화, 2차 활용, 표준화, 인프라·거버넌스 구축 포함
핀란드	· Findata를 통해 연구·통계·교육 등 2차 목적의 데이터 연계·제공 · 데이터 비식별화 후 최소한의 정보만 제공 · 활용 시점에 따라 opt-in/opt-out 방식으로 개인정보 활용 동의 관리 · 칸타 서비스 및 핀젠 연구 프로젝트로 데이터 구축·활용

□ 우리나라는 데이터 기반의 바이오헬스 분야 경쟁력 확보를 목표로 보건 의료데이터 활용 가치를 높이기 위해 데이터 생산, 집적, 활용 제도적 기반 마련

○ 보건 의료데이터 수집, 활용을 위해 국가 통합 바이오 빅데이터 구축 사업, 의료 마이데이터 사업 등 다양한 정책을 시행 중

○ 의료데이터 표준화를 통한 상호운용성 제고를 위한 고시 마련, 가명정보 활용 촉진과 데이터 기반 연구 활성화 지원을 위한 가이드라인 등 제도적 기반 마련

□ 보건 의료데이터의 활용을 위해서는 ① 데이터 상호운용성 확보, ② 개인정보의 활용 기반 합리화 및 ③ 이해관계자 간 거버넌스 체계 구축 필요

○ 데이터 상호운용성 확보를 위한 표준화 체계 정비

- 의료기관마다 형식이 다른 비정형 데이터 항목에 대한 표준화 기준 및 기관 간 데이터 연계 시 발생 가능한 매핑 오류 최소화를 위한 가이드라인 마련
- 데이터 공유 시 표준화 수준 점검·개선할 수 있는 상호운용성 평가체계 구축

○ 개인정보 활용 기반의 제도 개선과 사회적 수용성 확보

- 데이터 2차 활용 시 정보주체의 사전 동의(opt-in) 만으로는 활용에 제약이 있어, 정보 보호를 전제로 한 조건부 활용 제도 도입 논의 필요
- 데이터 활용에 대한 신뢰 제고를 위한 사회적 논의와 공론화, 활용 과정 및 결과에 대한 투명한 정보공개 및 활용 결과 피드백 체계 마련

○ 이해관계자 간 협력 기반의 거버넌스 체계 구축

- 데이터 제공자(개인, 의료기관 등)와 활용자(산업계, 연구기관) 간의 역할과 책임 구분 명확화, 데이터 제공의 적정성 기준 마련
- 정부, 의료기관, 산업계, 국민이 참여하는 보건 의료데이터 활용 협의체 구성을 통해 데이터 활용 방향에 대한 합의 기반 마련 및 세부 가이드라인 정립

참고문헌

[국문자료]

- 과학기술정보통신부 보도자료(2025), “드넓은 가능성의 신대륙, 첨단바이오 시대 개막을 위해 민관 원팀으로 바이오 역량 총결집”
- 구태언, 설기석, 방석호(2025), “의료산업과 AI, 마이 데이터 사업”, 법률신문
- 김주현, 신주연(2022), “국내 의료 분야 마이데이터 도입을 위한 해외 사례 조사”, 한국보건산업진흥원 KHIDI Brief
- 박나영 외 6인(2023), “건강정보 고속도로 플랫폼 활용 중점서비스 발굴 및 확산전략 제언”, 보건복지부, 한국보건사회연구원
- 박도휘 외 2인(2024), “AI로 촉발된 헬스케어 산업의 대전환”, 삼정 KPMG
- 보건복지부(2021), “보건의료 데이터·인공지능 혁신전략, Health Data F.L.O.W.2025”
- 보건복지부(2024), “의료 인공지능 연구개발(R&D) 로드맵(안) ('24-'28)”
- 손슬기, 박나영(2024), “핀란드와 호주의 보건의료 마이데이터 도입 및 활용”, 국제사회보장리뷰
- 식품의약품안전처 식품의약품안전평가원(2023), “의료기기의 실사용증거 적용에 대한 가이드라인”
- 심희진(2025), “나만의 닥터, 국내 최초 의료마이데이터 서비스 출시”, 매일경제
- 유소영(2018), “헬스케어 빅데이터 딜레마와 해결방안”, KAKAO AI REPORT
- 윤희정(2025), “AI를 활용한 혁신 신약개발의 동향 및 정책 시사점”, KISTEP 브리프
- 이상헌(2023), “바이오 빅데이터와 헬스케어 데이터 플랫폼의 효과적인 연계를 위한 관련 동향분석 및 시사점”, 한국보건산업진흥원 KHIDI Brief
- 이상헌, 이후석(2022), “의료 마이데이터 도입이 보건의료 데이터산업에 미치는 영향, 경제적 파급효과 분석을 중심으로”, 한국보건산업진흥원 KHIDI Brief
- 정상태(2025), “의료데이터 활용에 관한 법·제도 현황 및 이슈”, 한국제약바이오협회 Issue Report
- 최남경(2025), “RWD/RWE 외부대조군 활용 동향 및 시사점”, 한국제약바이오협회 KPBMA FOCUS
- 최윤희, 최은희(2024), “스마트헬스케어에 대한 한국의 사회적 수용성 현황과 과제”, 산업연구원
- 한국보건의료정보원, ABOUTGROUP(2023), “보건 의료데이터 활용 현황조사 및 분석연구”, 한국보건의료정보원

[영문자료]

- Assistant Secretary for Technology Policy(2023), "Sync for Genes"
- Christy M et al.(2025), "Rapid whole-genome sequencing as a first-line test is likely to significantly reduce the cost of acute care in a private payer system", The Journal of Applied Laboratory Medicine
- Davide C and Alfonso Valencia(2019), "Big data analytics for personalized medicine", Current Opinion in Biotechnology
- EMA(2024), "Real-world evidence provided by EMA"
- FDA(2025), "Real-world evidence"
- FINNGEN, "About FinnGen"
- Hyunsuk Y et al.(2020), "Validation of a deep learning algorithm for the detection of malignant pulmonary nodules in chest radiographs", JAMA Network Open
- James AD and Joseph CK(2022), "Computer copilots for endoscopic diagnosis", NPJ Digital Medicine
- Joe M(2020), "Blue Button 2.0: Everything you need to know and more!"
- Juliamn M et al.(2025), "Genomic screening at a single health system", JAMA Network Open
- Kornelia Batko and Andrzej Ślęzak(2022), "The use of big data analytics in healthcare", Journal of Big Data
- Luke S et al.(2023), "Progress on implementing and using electronic health record systems: Developments in OECD countries as of 2021", OECD
- Mitja IK et al.(2023), "FinnGen provides genetic insights from a well-phenotyped isolated population", Nature
- National Institutes of Health(2021), "All of Us"
- Quanticate(2024), "A guide to real-world evidence in clinical trials"
- U.S.Department of Health and Human Services(2024), "The HIPPA Privacy Rule"
- Weller GB et al.(2024), "A retrospective observational study of continuous wireless vital sign monitoring via a medical grade wearable device on hospitalized floor patients", Journal of Clinical Medicine

[웹사이트]

건강정보고속도로 포털(myhealthway.go.kr)

국가통합바이오빅데이터구축사업단(biobigdata.kr)

법제처 국가법령정보센터(law.go.kr)

보건의료 빅데이터 통합 플랫폼(hcdl.mohw.go.kr)

ASTP(Assistant Secretary for Technology Policy)(healthit.gov)

B1MG(b1mg-project.eu)

European Genomic Data Infrastructure(gdi.onemilliongenomes.eu)

European '1+Million Genomes' Initiative

(digital-strategy.ec.europa.eu/en/policies/1-million-genomes)

FINNGEN(finngen.fi)

Sync for Science(syncfor.science)

The European Health Data Space(EHDS)(european-health-data-space.com)

UK Biobank(ukbiobank.ac.uk)

1+MG(framework.onemilliongenomes.eu)