

보건의료 빅데이터 활용의 법제도, 윤리, 사회경제적 쟁점

연구공동체 건강과대안
이 상 윤

보건의료 빅데이터란?

실시간으로 새롭게 생성, 기록, 집적되는 '건강' 및 '생체' 관련 다양한 데이터

보건의료 빅데이터의 예시

- 전통적인 보건의료 빅데이터
 - 국민건강 행정(보험 등)을 위한 자료
 - 의료기관의 전자의무기록
 - 의료기관의 생체 조직
 - 연구기관의 유전정보, 각종 -omics 자료
- 새로운 보건의료 빅데이터
 - 소셜미디어와 각종 건강 플랫폼을 통해 생성되는 데이터
 - 개인 건강 모니터링 기술 등을 통해 생성되는 데이터
 - 가정에 존재하는 센서를 통해 생성되는 데이터
 - 스마트폰 어플리케이션을 통해 생성되는 데이터
 - 온라인 포럼, 검색 엔진 등을 통해 생성되는 데이터

보건의료 빅데이터의 활용성

- 다양하고 방대한 양의 보건의료 데이터를 연계, 집적, 분석하여 개인과 특정 현상과의 관계나 패턴을 확인할 수 있음
- 보건의료 빅데이터 활용의 예
 - 개인 수준의 건강 결과를 예측 : 행동 예측, 환자 개인의 위험 층화, 건강 컨설팅
 - 사회 수준의 건강 결과 모니터링 : 특정 질병 유행 양상에 대한 모니터링, 만성질환 모니터링
 - 의료의 효율 향상 : 특정 개입요인 혹은 치료 방법의 효과 및 효율 평가, 의약품 부작용 모니터링, 임상적으로 유용한 예측 모델 개발, 치료 과정의 의사 결정 지원, 의료기관 자원 배분 관리
 - 각종 보건의료 연구 : 역학 연구, 감염병 연구, 유전체 및 유전자 연구 등

보건의료와 인공지능 기술

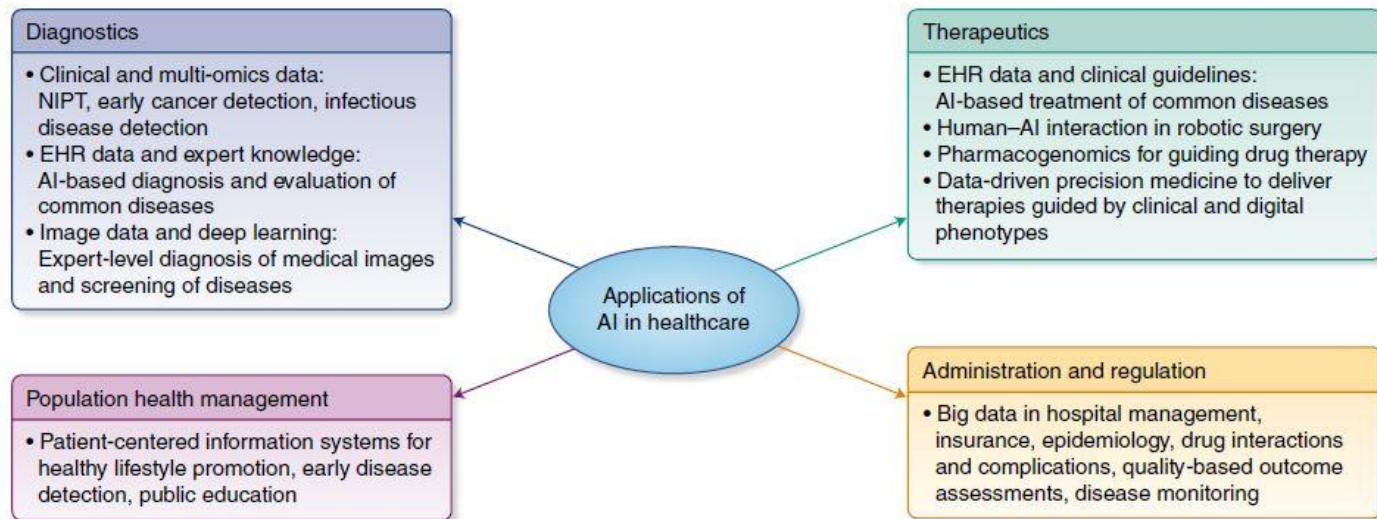


Fig. 1 | Potential roles of AI-based technologies in healthcare. In the healthcare space, AI is poised to play major roles across a spectrum of application domains, including diagnostics, therapeutics, population health management, administration, and regulation. NIPT, noninvasive prenatal test. Credit: Debbie Maizels/Springer Nature

He, Jianxing, et al. "The practical implementation of artificial intelligence technologies in medicine." *Nature medicine* 25.1 (2019): 30-36.

누가? 왜?



제약회사

신약개발,
임상실험,
판매/마케팅 등의
영역에서 필요



의료기기 회사

인공지능+딥러닝 :
왓슨 등



병원, 건강관리 회사

맞춤형 건강관리
치료과정 효율화

의료기관
performance 향상



IT업계

클라우드, 데이터,
시장 선점



공공기관

독감예측 시스템
자살예방 시스템

의료 시스템 효과
평가 및 향상



과학계, 의학계

연구의 양과 질
향상

보건의료 빅데이터 활용의 법적, 윤리적, 사회경제적 문제

- 개인정보 보호 및 프라이버시 문제
 - 고도로 민감한 개인 정보, 환자들의 취약성
- 알고리즘에 의한 결정, 자율성 침해 문제
 - 정치적 정당성, 책임성, 인권 존중
- 보건의료 빅데이터 활용 기기, 의약품 상업화로 인한 문제
 - 데이터 독점, 기술 특허, 이해관계 상충

보건의료 빅데이터의 상업적 이 용

보건의료 관련 상업적 어플리케이션과 개인정보 보호

- 모바일 어플리케이션(앱) 개발자는 정기적으로 합법적으로 사용자 데이터를 공유하고 있음.
- 대부분의 보건의료 관련 앱은 데이터 공유 관행에 대한 개인 정보 보호 차원의 확인 및 투명성 절차를 제공하지 않음.
- 보건의료 관련 앱에서 수집한 사용자 데이터는 사이버 범죄자 또는 상업용 데이터 브로커에게 특히 매력적임.
- 민감한 개인 건강 정보를 수집하는 보건의료 관련 앱은 다른 유형의 앱과 거의 동일한 방식으로 모바일 생태계 내에서 사용자 데이터를 공유함.
 - 한 연구에 따르면, 연구 대상 앱의 79%가 사용자 데이터를 다른 주체와 공유(Grundy, Quinn, et al. "Data sharing practices of medicines related apps and the mobile ecosystem: traffic, content, and network analysis." *bmj* 364 (2019))
- 소수의 회사는 사용자 데이터를 집적하고 재식별하고 있는 것으로 추정되고 있음.

보건의료 빅데이터와 개인의 자율성(Autonomy)

- 건강증진 앱 및 웨어러블 디바이스
 - 핏빗, 스마트 워치, 스마트폰 앱, 의복, 신발, 보석 등
 - behavior change techniques, 'nudge' 이론에 근거한 기술
 - 특정한 기준(norm)과 가치(value)를 전제
 - 감시, 죄책감, 자발적 강제 등에 기반 전략
 - 경쟁, 관리 등에 기반한 자아 경영
 - 책임감, 자율성, 선택 등에 기반한 신자유주의적 인간형 만들기
 - 생활세계의 식민화, 의료화
 - 대부분이 소비자주의, 상업성에 기반한 기업에 의해 제공되는 서비스라는 문제
 - 두려움, 죄책감 및 수치심의 메시지로 뒷받침되는 매혹적인 희망, 욕망, 욕망, 역능감, 성공 등을 암시하며 소비자의 감정을 조작
 - 기존의 건강 불평등을 심화하고, 불안과 낙인을 조장할 위험이 있음
 - 실제 건강 증진 효과도 미지수

빅데이터와 AI 알고리즘 활용 의료 기기에 대한 허가, 평가 기준

- 알고리즘의 효과성, 안전성 등을 평가하는 틀, 방식의 문제 (Parikh, Ravi B., Ziad Obermeyer, and Amol S. Navathe. "Regulation of predictive analytics in medicine." *Science* 363.6429 (2019): 810-812.)
 - 의미 있는 임상적 유용성을 입증해야 함 : 생명 연장, 검사의 효과 증진, 검사의 오류 감소 등 "임상적으로 의미 있는" 결과에 기반한 평가가 이루어져야 함. 알고리즘에 의한 결정이 낳을 수 있는 bias에 대한 충분한 고려가 이루어져야 함
 - 기존 진단, 개입, 치료 방식 등과의 대조군 연구가 이루어져야 함
 - 호환가능성, 일반화 가능성의 문제 : 특정 병원, 특정 환자, 특정 EHR 등을 이용한 알고리즘의 문제
 - 적절한 개입 방식과 연관되어야 함
 - 사후평가 메커니즘의 존재

보건의료 빅데이터 AI 기기 개발, 현실 적용시 요구사항

- 어떤 데이터를 사용하여 개발하였는지 오픈할 것(개인정보 보호)
- 기기 개발시 어떤 방식으로(데이터 라벨링 등) '기계 학습'을 시켰는지 오픈할 것(투명성)
- 기기의 알고리즘과 결과에 대해 사후적으로 "해석가능함"을 증명할 것
- 기기의 오작동으로 인한 피해 발생시 책임주체를 명확히 할 것

보건의료 AI 기기의 실제 사용의 현실적 장애물

- 데이터 공유 수준 : 의료기관에 인센티브가 있는가
- 데이터 표준화 : 다양한 의료기관 데이터가 얼마나 표준화될 수 있을 것인가
- 알고리즘 표준화 : 개별 업체의 알고리즘이 하나의 워크 플로우에 통합될 수 있을 것인가
- 재정 문제 : 엄청난 규모의 재정 투입이 필요

유럽연합 GDPR과 보건의료 AI(2)

GDPR 제15조 개인정보주체의 열람권

1. 개인정보주체는 본인에 관련된 개인정보가 처리되고 있는지 여부에 관련해 개인정보처리자로부터 확답을 얻을 권리를 가지며, 이 경우, 개인정보 및 다음 각 호의 정보에 대한 열람권을 가진다.
 - (a) 처리 목적
 - (b) 관련된 개인정보의 범주
 - (c) 개인정보를 제공받았거나 제공받을 수령인 또는 수령인의 범주, 특히 제3국 또는 국제기구의 수령인
 - (d) 가능한 경우, 개인정보의 예상 보관 기간 또는, 여의치 않은 경우, 해당 기간을 결정하는 데 사용되는 기준
 - (e) 개인정보처리자에게 본인의 개인정보에 대한 정정 또는 삭제를 요구하거나 개인정보주체 본인에 관한 처리의 제한이나 반대를 요구할 권리
 - (f) 감독기관에 민원을 제기할 수 있는 권리
 - (g) 개인정보주체로부터 개인정보를 수집하지 않은 경우, 개인정보의 출처에 대한 모든 가용한 정보
 - (h) 제22조(1) 및 (4)에 규정된 프로파일링 등 자동화된 의사결정의 유무. 최소한 이 경우, 관련 논리에 관한 유의미한 정보와 그 같은 처리가 개인정보주체에 가지는 중대성 및 예상되는 결과
2. 개인정보가 제3국이나 국제기구로 이전되는 경우, 개인정보주체는 제46조에 따라 적절한 안전조치에 대해 고지 받을 권리가 있다.
3. 개인정보처리자는 처리가 진행 중인 개인정보의 사본을 제공해야 한다. 개인정보주체가 추가 사본을 요청하는 경우, 개인정보처리자는 행정적 비용에 근거하여 합리적인 비용을 청구할 수 있다. 개인정보주체가 전자적 방식으로 해당 요청을 하는 경우, 관련 정보는 통상적으로 사용되는 전자적 양식으로 제공되어야 한다.
4. 제3항에 규정된 사본을 입수할 권리는 제3자의 권리와 자유를 침해하지 않아야 한다.

유럽연합 GDPR과 보건의료 AI(2)

GDPR 제22조 프로파일링 등 자동화된 개별 의사결정

1. 개인정보주체는 프로파일링 등, 본인에 관한 법적 효력을 초래하거나 이와 유사하게 본인에게 중대한 영향을 미치는 자동화된 처리에만 의존하는 결정의 적용을 받지 않을 권리를 가진다.
2. 결정이 다음 각 호에 해당하는 경우에는 제1항이 적용되지 않는다.
 - (a) 개인정보주체와 개인정보처리자 간의 계약을 체결 또는 이행하는 데 필요한 경우
 - (b) 개인정보처리자에 적용되며, 개인정보주체의 권리와 자유 및 정당한 이익을 보호하기 위한 적절한 조치를 규정하는 유럽연합 또는 회원국 법률이 허용하는 경우
 - (c) 개인정보주체의 명백한 동의에 근거하는 경우
3. 제2항 (a)호 및 (c)호의 사례의 경우, 개인정보처리자는 개인정보주체의 권리와 자유 및 정당한 이익, 최소한 개인정보처리자의 인적 개입을 확보하고 본인의 관점을 피력하며 결정에 대해 이의를 제기할 수 있는 권리를 보호하는 데 적절한 조치를 시행해야 한다.
4. 제2항의 결정은 제9조(2)의 (a)호와 (g)호가 적용되고, 개인정보주체의 권리와 자유 및 정당한 이익을 보호하는 적절한 조치가 갖추어진 경우가 아니라면 제9조(1)의 특정 범주의 개인정보를 근거로 해서는 안 된다.

유럽의 GDPR과 건강 정보 보호

- 건강 정보의 정의
 - “personal data related to the physical or mental health of a natural person, including all data about his or her past, current, or future health status collected in the course of the registration for or the provision of healthcare services to him or her”
- 건강정보는 원칙적으로 처리 금지, 다음의 경우 예외적으로 허용
 - 명시적 동의
 - 의료 행위, 공중보건 영역에서 공공의 이익(Public interest)를 위한 경우, 과학적 연구 등의 경우 개인 동의 없이 처리 가능
- 보다 엄격한 동의 규정(Informed consent), 고지 의무(notification duties)
- Data protection impact assessment
- 자동화된 데이터 처리와 알고리즘에 의한 의사 결정이 이루어 질 때 이에 대해 환자가 알고 개입할 권리 보장
- 자신의 건강정보 삭제를 요청할 권리
- Data protection by design
 - 건강정보를 다루는 기관이 스스로 엄격한 개인정보 보호 룰을 적용하도록 하고, 어겼을 경우 엄하게 처벌 : 전세계 연간 매출의 2% 또는 1천만 유로(한화 약 130억원)의 과태료, 대규모 위반시 전세계 연간 수익의 4% 또는 2천만 유로(한화 약 260억원)의 과태료

건강정보와 과학적 연구

과학적 연구 예외 조항과 관련된 쟁점

개인의 동의 없이 건강정보를 처리할 수 있는 과학적 연구의 범위는 각국에 위임

최근 개별적으로 제정되고 있는 사례는 투명하고 책임성 있는 거버넌스 기구에서 과학적 연구의 범위를 결정하거나, 과학적 연구의 범위를 한정하는 추세

- 아일랜드 : 건강 연구의 경우 원칙적으로 개인 동의 없이 수행 불가. 예외적인 경우 국가가 운영하는 위원회에서 허가를 받아야 함
- 캘리포니아 : 익명화된 정보에 한해, 상업적 목적으로는 처리 불가, 개인 정보 수집시 원래 목적에 비례한 연구에 한해 개인 동의 없이 처리 가능

개인 동의 없이 이루어지는 과학적 연구의 경우 보다 엄격한 Safeguard 및 정보주체의 권리 보장을 요구

- 연구의 목적(공익성, 필요성, 비례성)
- 연구 진행 과정 및 결과의 안전성(safe haven, trusted third party)
- 투명하고 책임성 있는 거버넌스
- 연구의 전 과정에 정보 주체의 참여 및 알 권리 보장
- 매 시기마다 opt-out 권리 보장

보건의료 빅데이터와 신뢰, 정의

생의학 연구에 있어 '동의'의 중요성

사전 동의가 불가능한 상황에서 고려되는 집단의 이해관계(Group interest)와 포괄적 동의

사전에 충분히 고지된 상태의 동의는 비현실적이고, 데이터 익명화는 더 이상 효과적인 안전장치가 되지 않는 상황에서 더 강조되는 신뢰(Trust)의 중요성

- 데이터 상업화, 기타 산업적 이해관계가 개입되며 신뢰의 위기 발생

신뢰는 감정적, 규범적, 이성적 기반 위에 형성

- 오랜 시간, 경험에 의해 형성되는 것
- 경제적 이익, 집단 이익, 과학의 진보라는 이름으로 개인의 인권과 이해관계를 제약하는 방식은 정의롭지 못함
- 정책 당국은 어떻게 대중에게 개인의 인권과 이해관계를 존중하면서 보건의료 빅데이터 연구가 진행되고 있다는 확신을 줄 것인가?

잠정적 방안

- 보건의료 빅데이터 연구를 총괄하는 신뢰할 만한 데이터 중개소를 설립
- 보건의료 빅데이터 연구의 경우, 기관 데이터접근위원회(DAC)의 역할은 중앙으로 모아 한 곳에서 수행하고 기존의 기관생명윤리위원회(IRB), 연구윤리위원회(REC)는 보완적인 역할을 담당

보건의료 빅데이터 연구의 전제조건(1)

- 공공의 이익(Public benefit) :과학적 진실성, 사회적 가치
 - 연구의 이익이 정보 주체, 집단, 지역사회에 주는 해보다 큰가?
 - 질 낮은 데이터, 알고리즘 편견, 체계적 데이터 배제, 분석 방법의 부적절성, 해석의 부적절성 등이 개입되면 특정 집단에 해가 되는 연구 결과가 도출될 수도 있음

보건의료 빅데이터 연구의 전제조건(2)

- 비례성, 필요성, 최소 침해의 원칙
 - 비례성 : 해보다 이익이 큰가
 - 필요성 : 다른 연구 방법, 다른 데이터로 목적을 달성할 수는 없는가
 - 최소 침해 : 데이터 사용을 최소화, 연구자의 개인식별자 접근을 최소화

보건의료 빅데이터 연구의 전제조건(3)

- 형평성 : 사회적 연대, 호혜성
 - 특정 집단에게만 이익이 되는 연구나, 특정 집단에게 특별히 해가 되는 연구는 불가
 - 특정 집단에 대한 차별 없는 참여와 개입
 - 연구의 과정과 결과가 모든 연구자에게 호혜적이어야 함

보건의료 빅데이터 연구의 전제조건(4)

- 신뢰 : 참여

- 데이터 거버넌스는 특정 집단이 보건의료 빅데이터 접근 및 사용을 제어하고 제안된 연구의 위험과 잠재적 이점을 평가할 수 있는 메커니즘을 제공
- 특정 집단이 구조적, 체계적으로 배제되거나 거부되는 상황을 예방

보건의료 빅데이터 연구의 전제조건(5)

- 책임성 : 투명성, 정당성
 - 연구자가 데이터의 질, 안전성, 사용 등에 대해 대답할 수 있어야 함
 - 공공의 이익을 위해 투명하고 개방적으로 사용되고 있다는 인식을 주어야 함