

# AI 기반 건강위험 예측 도구의 윤리적 쟁점과 사회적 수용성

## 임재강<sup>1</sup>

<sup>1</sup> 전) 경운대학교

### 초록

이 연구는 인공지능(AI) 기반 건강위험 예측 도구의 윤리적 쟁점과 사회적 수용성 문제를 미국, 유럽 연합(EU), 한국, 중국을 중심으로 비교 분석한다. AI 예측 도구는 질병의 조기 발견, 맞춤 치료, 의료 자원 최적화 등 혁신적 잠재력을 지니지만, 데이터 거버넌스(프라이버시, 보안, 동의), 알고리즘 건전성(편향, 공정성), 투명성 및 설명가능성, 책임 소재 규명, 차별 가능성 등 복잡한 윤리적 과제를 동시에 제기한다. 분석 결과, 공통적으로 윤리적 우려를 공유하지만, 기본권 보호(EU), 건강 형평성(미국), 사회 안정(중국), 혁신-규제 균형(한국) 등 강조하는 가치에서 차이를 보였다. 사회적 수용성에서도 의료 AI에 대한 중국의 높은 낙관론과 미국과 EU의 우려와 한국의 신중론 등 차이를 나타냈다. 규제 방식에서도 EU의 포괄적 사전 규제, 미국의 분산적 접근, 중국의 목표 지향적 규제, 한국의 위험 기반 접근 등 서로 다른 특징을 보였다. 이러한 차이는 문화, 정치 시스템, 경제 전략, 기술 발전 속도 등 복합적 요인에 기인하며, 글로벌 AI 거버넌스, 국제 협력, 시장 통합에 중요한 함의를 지닌다.

**주요 용어:** 인공지능, 건강위험 예측, 의료 AI, 윤리적 쟁점, 사회적 수용성, 투명성

### 알기 쉬운 요약

**이 연구는 왜 했을까?** 기존 연구들이 AI 건강 예측의 기술적 측면, 특정 윤리적 쟁점, 또는 단일 국가 내 수용성 문제에 집중했던 것과 달리, 이 연구는 AI 기술의 영향이 국경을 초월하고 각국의 대응 방식이 상이하다는 점에 주목하여, 주요 국가(미국, EU, 한국, 중국)의 윤리적 논의, 사회적 수용성 동인, 규제 프레임워크를 비교 분석하여 그 차이와 유사성의 원인을 규명하고 글로벌 거버넌스에 대한 함의를 도출하고자 했다.

**새롭게 밝혀진 내용은?** 윤리적 우선순위 차이를 발견했다. EU는 기본권 보호를, 미국은 건강 형평성을, 중국은 사회 안정을, 한국은 혁신과 규제의 균형을 강조한다. 사회적 수용성 차이를 발견했다. 중국은 높은 낙관론을 보이는 반면, 미국과 EU는 우려를 표하며, 한국은 신중론을 보인다. 규제 방식에서도 EU는 포괄적 사전 규제를, 미국은 분산적 접근을, 중국은 목표 지향적 규제를, 한국은 위험 기반 접근을 채택하는 등 서로 다른 특징을 보였다.

**앞으로 무엇을 해야 하나?** 민감 건강 데이터의 책임 있는 관리와 보안 확보가 필수적이다. 데이터 편향으로 인한 건강 불평등 심화를 해결하기 위해 기술적, 조직적, 사회적 차원의 통합적 접근이 필요하다. AI의 '블랙박스' 특성으로 인한 불투명성을 해소하여 신뢰성과 책임성을 확보하는 노력이 필요하다. AI 오류로 인한 피해 발생 시 책임 소재 규명을 위한 새로운 주의 기준 정립과 법적·제도적 재구성이 요구된다. AI가 보험 및 고용 등에서 차별적으로 활용되거나, 알고리즘 편향으로 인해 건강 불평등을 심화시키지 않도록 예방해야 한다. 의료 AI는 의료진의 판단을 대체하는 것이 아니라 지원하는 도구로 신뢰와 수용성 확보가 필요하다.

■ 투 고 일: 2025. 04. 30.

■ 수 정 일: 2025. 06. 19.

■ 게재확정일: 2025. 06. 30.

# I. 서론

## 1. 연구의 필요성

인공지능(AI), 특히 기계 학습(Machine Learning, ML)과 심층 학습(Deep Learning, DL) 기술은 방대한 건강 데이터를 분석하여 질병 예측, 진단 보조, 맞춤 치료 등 의료 분야에 혁신적 변화를 가져올 잠재력을 보여주고 있다(Gerke et al., 2020). 의료 분야 인공지능(AI) 기술은 전자건강기록(EHR), 의료 영상, 유전체 정보 등 다양한 데이터를 활용한 AI 기반 건강위험 예측 도구는 만성 질환 조기 발견부터 감염병 확산 예측, 의료 자원 최적화에 이르기까지 광범위한 응용 가능성을 제시하며 정밀 의학 및 맞춤 의학 시대를 앞당길 핵심 기술로 주목받고 있다(Hendricks-Sturup et al., 2023). 이러한 기술의 눈부신 발전 이면에는 데이터 편향성으로 인한 건강 불평등 심화, 민감한 개인 건강 정보의 프라이버시 침해, 예측 모델의 '블랙박스' 문제로 인한 신뢰성 및 책임 규명 어려움, 예측 정보의 차별적 활용 가능성 등 중대한 윤리적, 사회적 과제가 산재해 있다(Gundersen & Bærøe, 2022). 이러한 과제들은 기술의 신뢰성 및 사회적 수용성(Social Acceptance) 확보와 직결되며, 성능 향상을 위한 데이터 요구와 프라이버시 및 공정성 보호라는 가치 사이에 근본적인 긴장 관계가 존재한다(Orall & Rekito, 2025). 따라서 기술 혁신과 윤리적 가치 간의 균형을 찾는 것이 필수적이며, 개발 초기 단계부터 신중한 성찰과 규제적 접근이 요구된다.

그러나 선행 연구들은 AI 건강 예측의 기술적 측면, 특정 윤리적 쟁점(예: 편향, 프라이버시), 또는 단일 국가 내 수용성 문제를 다루어왔다. 최근 AI 기술의 영향이 국경을 초월하고 각국의 대응 방식이 상이하게 나타나는 상황에서, 주요 국가들의 윤리적 논의, 사회적 수용성 동인, 규제 프레임워크를 비교 분석하여 그 차이와 유사성의 원인을 규명하고 글로벌 거버넌스에 대한 함의를 도출하는 연구는 상대적으로 부족하다.

## 2. 연구의 목적과 방법

이 연구의 목적은 AI 기반 건강위험 예측 도구의 윤리적 쟁점, 사회적 수용성, 그리고 규제 환경을 미국, 유럽 연합(EU), 한국, 중국 네 국가를 대상으로 비교하는 것이다. 미국, 유럽연합(EU), 한국, 중국을 비교 분석 대상으로 선정한 배경은 이 네 국가는 각각 AI 기술 수준, 윤리·법적 제도, 그리고 의료 AI에 대한 정책 접근에서 뚜렷한 차이를 보이기 때문이다. 미국은 민간 주도 기술 혁신과 분산적 규제를 대표하고, EU는 기본권 기반의 엄격한 규범 틀을 주도하며, 중국은 국가 주도형 산업정책과 강력한 감독 체계를 결합하고 있으며, 한국은 기술 혁신과 규제 균형을 모색하는 중도적 국가로 특징을 보이며 연구 사례로서 가치가 있다.

연구 질문은 다음과 같다.

- ① 미국, EU, 한국, 중국에서 AI 건강위험 예측과 관련하여 중요하게 논의되는 윤리적 쟁점은 무엇이며, 우선순위와 차이, 그리고 원인은 무엇인가?
- ② 각 국가에서 AI 건강위험 예측 도구에 대한 사회적 수용성 수준은 어떠하며, 수용성에 영향을 미치는 주요 요인은 무엇인가?
- ③ 각 국가의 AI 건강위험 예측 관련 규제 체계/framework)는 어떤 특징을 보이며, 이는 윤리적 문제 해결에 어떤 영향을 미치는가?
- ④ 이러한 국가별 차이는 어떤 사회문화적, 정치경제적 요인에 의해 발생하며, 글로벌 AI 거버넌스에 어떤 영향을 미치는가?

이 연구는 질적 비교 사례 연구 방법론(comparative case study)에 기반한다. 미국, EU, 한국, 중국을 분석

대상으로 선정하고, 각 국가의 AI 건강위험 예측 관련 문헌을 수집하여 분석한다. 분석의 틀은 각 지역별로 ① 윤리적 쟁점, ② 사회적 수용성 수준 및 영향 요인, ③ 규제 프레임워크 특징을 비교하는 데 초점을 맞춘다.

논문의 구성은 II장에서 AI 건강위험 예측의 기술적 개념과 메커니즘, 주요 윤리적 과제(데이터 거버넌스, 알고리즘 건전성, 투명성, 책임성, 차별 문제) 및 사회적 수용성 영향 요인에 대한 이론적 고찰을 한다. III장에서는 연구 설계를 기술한다. IV장에서는 미국, EU, 한국, 중국의 현황을 분석하고, 이를 바탕으로 윤리적 우선순위, 사회적 수용성, 규제 방식의 유사점과 차이점을 비교한다. V장에서는 주요 연구 결과를 요약한다.

## II. 이론적 고찰

### 1. AI 기반 건강 예측 도구의 개념

AI 기반 건강위험 예측 도구는 인공지능을 활용하여 다양한 건강 관련 데이터를 분석하고, 미래의 특정 건강 관련 사건(질병 발생, 악화, 예후 등)의 가능성을 예측하는 기술이다(Hendricks-Sturup et al., 2023). 활용되는 데이터는 전자건강기록(EHR), 인구통계학적 정보와 같은 정형 데이터뿐만 아니라, 의료 영상(X-ray, CT, MRI), 유전체 염기서열, 웨어러블 기기 기반 생체 신호, 생활 습관 정보, 환경 노출 및 사회경제적 요인 등 비정형 데이터까지 광범위하다. AI 기반 건강위험 예측 기술은 임상 진료부터 공중 보건 정책 수립, 의료 시스템 운영에 이르기까지 활용된다. 주요 사례는 다음과 같다.

첫째, AI는 만성질환(예: 당뇨병, 심혈관 질환) 고위험군을 증상 발현 이전에 식별하고, 예방적 개입(생활 습관 개선, 조기 검진 등)을 가능하게 한다. 의료 영상(유방촬영술, CT, MRI 등) 분석을 통해 암을 조기에 발견하거나, 유전체 데이터 분석을 통해 특정 암 발병 위험도를 예측한다(Binkley et al., 2022). 피부암 진단을 위한 이미징 시스템이나 심장마비 확률을 예측하는 스마트 센서와 같은 혁신적인 응용 사례들이 등장하고 있다(FDA, 2025). 미국 식품의약국(FDA)은 이미 1,000개 이상의 AI 의료기기를 승인했다(FDA, 2025, January 6).

둘째, AI는 고위험 환자 집단을 선별하여 개인화된 치료 계획을 제시하거나, 집단 수준의 예방 프로그램을 설계하는 데 활용된다. Owens(2023)는 AI/ML 기반 도구가 고위험 환자 식별과 진단·치료의 효율성 향상에 기여할 수 있다고 보았으며, 다만 이들 시스템의 지속적 유지보수와 품질관리가 전제되어야 함을 강조하였다.

셋째, AI는 병상 및 인력 배치, 재입원 예측 등 의료 자원의 효율적 배분을 가능케 하며, 의료 시스템의 운영 효율성을 제고하는 데 기여한다. 예를 들어, 감염병 유행 시 AI가 지역별 확산 위험을 예측하여, 의료진 배치, 이동형 선별 진료소 설치, 백신 및 치료제를 효율적으로 조정하도록 한다. 또는 질병 발병 위험이 높은 환자군을 사전에 식별하여 집중 관리가 필요한 환자에게 의료 인력, 병상, 의약품 등 자원을 우선 배정할 수 있도록 돕는다.

그러나 AI 예측 도구는 개인 정보 보호 및 데이터 보안 문제와 충돌한다(Jiang et al., 2021). 학습 데이터가 특정 집단을 과소 또는 과대 대표하거나, 데이터 수집 과정에 역사적·사회적 편견을 반영할 경우, AI 모델은 이러한 편향을 학습하고 증폭시켜 특정 집단에게 불리한 예측 결과를 나타낼 수 있다(Orall & Rekito, 2025). 따라서 AI 예측 모델 개발 초기 단계부터 신중한 접근과 관리가 필요하다.

AI 건강위험 예측 도구는 인간이 인지하기 어려운 미묘한 패턴, 특정 건강 결과와 관련된 위험 요인을 식별하기 위해 고도의 기법과 알고리즘을 사용한다. 인공지능 및 머신러닝 기반의 임상 의사 결정 지원 시스템은 환자 선택 과정에서 편향을 줄일 수 있는 가능성을 가지고 있다(Binkley et al., 2022). 이렇게 AI 기술이 발전하고 있음에도 불구하고, AI를 사용하는 사람들 중 절반 이상이 AI 챗봇의 건강 정보의 정확성에 대해 확신하지 못한다고 답했다(KFF, 2024). 이러한 결과는 정보의 신뢰성에 대한 대중의 의구심이 여전히 존재함을 시사한다.

이처럼 AI 건강위험 예측 도구는 다양한 방식으로 의료의 질을 향상시키고 시스템 효율성을 높이는 데 기여할 수 있지만, 이러한 잠재력을 실현하기 위해서는 다음에 논의될 윤리적 문제들을 신중하게 고려하고 해결하려는 노력이 반드시 병행되어야 한다.

## 2. AI 기반 건강 예측 도구의 윤리적 쟁점

AI 기반 건강위험 예측 도구는 윤리적 과제를 수반한다. 이러한 윤리적 쟁점들을 요약하면 데이터 거버넌스, 알고리즘 건전성, 투명성 및 설명가능성, 책임 소재, 차별가능성 및 불평등이다. 이 윤리적 과제들은 WHO, OECD, EU 및 각국 윤리 가이드라인에서 강조되고 있는 항목들이며, 실제 의료 현장에서 AI 기술 적용 시 발생할 수 있는 주요 위험성 및 민감성 이슈로 인식되고 있다. 최근 WHO(2024)는 「보건을 위한 대규모 멀티모달 모델의 윤리 및 거버넌스」 보고서를 통해, 의료 AI의 윤리 기준으로 데이터 편향 방지, 인간 감독, 설명가능성, 안전성 검증, 사회적 영향 평가 등을 제시하였다. 이는 본 연구가 설정한 윤리 범주와 상당한 정합성을 보인다. 또한 WHO는 고위험 의료 AI의 사전적 통제 필요성과 사회적 수용성 강화를 위한 투명한 거버넌스 구조를 강조하고 있다.

### 1) 데이터 거버넌스

AI 기반 건강위험 예측 도구는 높은 예측 정확도를 달성하기 위해 방대한 양의 민감한 건강 데이터를 필요로 한다. 여기에는 보호 대상 건강 정보(Protected Health Information, PHI)와 개인 식별 정보(Personally Identifiable Information, PII)가 포함되며, 이들의 수집, 저장, 공유, 분석 전 과정은 프라이버시 침해 및 기밀 유출이라는 위험과 직결된다(Gundersen & Bærøe, 2022). AI 개발에 있어 데이터의 책임 있는 관리와 보안 확보는 기술 신뢰성뿐 아니라 사회적 수용성 확보의 핵심 요소로 간주된다.

의료 데이터는 그 민감성으로 인해 해킹, 무단 접근, 내부자 오용 등에 지속적으로 노출되어 있다. 실제로 2018년부터 2023년 사이 전 세계적으로 대규모 의료 데이터 유출 사고는 102% 증가했으며, 2023년 한 해에만 1억 6,700만 명 이상의 개인정보가 유출되었다(Fox, 2024). 이에 따라 미국 보건복지부(HHS)는 의료 기관을 대상으로 기술 자산 목록 및 네트워크 맵 작성의 의무화를 포함한 문서화 기준 강화를 검토 중이다.

비식별화(익명화 혹은 가명화)된 데이터라 하더라도, 다른 정보와의 결합 또는 고도화된 재식별 기술을 통해 개인을 특정할 수 있는 가능성은 여전히 존재한다. 실제로 Gymrek 등(2013)은 유전체 정보의 일부만으로도 개인을 재식별할 수 있음을 실증적으로 제시한 바 있다. 이는 완전한 익명성이 현실적으로 매우 어렵다는 점을 시사한다.

AI 시스템의 윤리적 잠재 위험은 설계 단계에서 모두 예측하기 어렵다. 예컨대 특정 인종이나 사회경제적 집단에 불리하게 작동하는 알고리즘의 경우, 개발자조차 그 영향을 완전히 인지하지 못한 채 시스템이 운영될 수 있다. 또한 의료 환경은 지속적으로 변화하기 때문에, AI 모델이 학습한 데이터 분포와 실제 운영 환경 간의 차이(dataset shift)가 성능 저하를 유발할 수 있다. 패혈증 예측 모델의 경우, 임상 환경 변화에 따라 성능이 현저히 저하된 사례가 보고되기도 했다(Owens, 2023). 따라서 AI 시스템은 도입 이후에도 정기적인 성능 평가와 지속적인 유지보수 인프라가 필수적이다.

한편, 환자들은 자신의 민감한 건강 정보가 AI 시스템에 의해 어떻게 수집·처리·활용되는지, 이로 인한 위험은 무엇인지 명확히 인식하기 어려울 수 있다. 이러한 사항에 대한 국가 간 규제 차이는 의료 데이터를 활용한 AI 연구 및 국제 협력에 있어 추가적인 법적·윤리적 복잡성을 초래한다(양대승, 육소영, 2022).

## 2) 알고리즘 건전성

AI 알고리즘의 편향(Bias) 문제는 예측 정확성과 공정성을 심각하게 저해할 수 있으며, 이는 알고리즘의 설계와 운영 전반에 걸쳐 발생할 수 있다. 특히 의료 분야에서는 데이터 자체에 내재된 구조적 불평등과 사회적 편견이 AI 모델에 그대로 반영되거나 증폭될 가능성이 높다(Shuford, 2024). 예컨대, 의료 기록은 과거 특정 인종, 성별, 혹은 사회경제적 지위를 가진 환자들이 차별적으로 진단·치료받은 결과를 반영하고 있으며, AI는 이러한 역사적 편향을 '사실로 학습할 수 있다.

가장 흔한 편향의 원인은 훈련 데이터의 대표성 부족이다. 특정 인구 집단(예: 소수 인종, 여성, 노인, 장애인, 농촌 거주자 등)이 데이터에 과소 또는 과대 대표되는 경우, 해당 집단에 대한 예측 정확도가 낮아지며, 이는 임상적 의사결정 오류로 이어질 수 있다(Galvin & Hogan, 2024). 또한, 알고리즘을 설계하고 훈련하는 과정에서 개발자의 무의식적 편견이 개입될 수 있으며, 이는 개발·검증·해석 등 시스템 생애주기 전반에서 다층적인 편향으로 이어진다(Verma et al., 2024).

이러한 편향은 기존의 건강 불평등(Health Disparities)을 고착화하거나 오히려 심화시킬 수 있다. 예를 들어, 피부암 진단을 위한 AI 알고리즘이 어두운 피부톤을 가진 환자에 대해 낮은 정확도를 보이는 사례는 알고리즘 편향이 실제 임상에 미치는 영향을 보여주는 대표적 사례다(The Guardian, 2021.11.9). 결과적으로 이는 특정 집단에 대한 오진, 진단 지연, 치료 기회 박탈로 이어질 수 있으며, 의료 시스템 전반에 대한 소외 계층의 불신을 심화시킨다(염현주, 2025.4.10.).

이 문제를 해결하기 위해서는 기술적, 조직적, 사회적 차원의 통합적 접근이 필요하다. 기술적 측면에서는 다양한 인구 집단을 포괄하는 데이터셋을 구축하고, 알고리즘이 다양한 하위 집단에 대해 일관된 성능을 발휘하는지를 검증하는 것이 필수적이다. 정기적인 모델 감사와 성능 모니터링을 통해 잠재적 편향을 식별하고 수정해야 하며(Fang et al., 2024), 공정성 지표(Fairness Metrics)를 평가 및 모델 개선 과정에 통합하는 기술적 방법도 함께 사용되어야 한다(Herlihy et al., 2024).

그러나 기술적 조치만으로는 편향 문제를 근본적으로 해결하기 어렵다. 알고리즘 편향의 근본 원인은 종종 사회 구조적 불평등과 제도적 차별에 뿌리를 두고 있으므로, AI 윤리와 공정성을 확보하기 위해서는 데이터 기반 기술 개발을 넘어, 사회 정의(social justice) 관점에서의 제도 개선과 교육, 다분야 협력적 거버넌스가 병행되어야 한다.

## 3) 투명성 및 설명가능성

AI, 특히 심층 학습(Deep Learning) 기반의 복잡한 인공 신경망 모델은 높은 예측 성능을 제공하는 반면, 그 내부 작동 원리를 인간이 이해하기 어려운 '블랙박스(Black Box)' 특성을 지닌다(Fan et al., 2020). 이러한 불투명성은 AI가 제시하는 건강 위험 예측이나 치료 권고의 신뢰성을 저하시킬 수 있으며, 의료 현장에서의 수용성과 책임성 확보에 본질적인 장애 요소로 작용한다.

투명성(transparency)과 설명가능성(explainability)은 의료 AI의 신뢰성과 책임성을 확보하는 핵심 개념으로, 각각 AI 시스템의 전반적인 작동 방식에 대한 이해 제공과 특정 결정이 도출된 이유를 명확히 밝히는 것을 의미한다(Priyansh & Kaur Saggi, 2025). 투명성은 AI 시스템이 어떤 원리로 작동하며 어떤 자료에 근거하여 판단을 내리는지를 설명하고, 설명가능성은 특정 예측 결과나 권고가 어떤 요소들에 의해 도출되었는지를 명확하고 이해하기 쉬운 방식으로 제시하는 것을 목표로 한다(Nalepa et al., 2019).

설명가능성과 예측 성능 간에는 본질적인 긴장이 존재한다. 일반적으로 AI 모델의 복잡성이 높을수록 예측 정확도는 향상되지만, 설명가능성은 저하된다. 특히 딥러닝 모델의 경우, 수백만 개의 파라미터 간의 비선형 상호

작용을 설명하는 것은 기술적으로 매우 어렵다. 이에 따라, 모델 성능을 유지하면서도 설명가능성을 확보하려는 다양한 기술적 시도(e.g., LIME, SHAP, attention mechanisms 등)가 연구되고 있으나, 아직까지 의료 분야에 적합한 일관된 기준이나 최적의 접근법은 정립되지 않았다(National Conference of State Legislatures, 2025).

#### 4) 책임성

AI 기반 건강위험 예측 도구가 오류를 발생시켜 환자에게 피해를 입힌 경우(예: 질병 위험 간과로 인한 진단 지연, 부정확한 예측에 따른 부적절한 치료 권고 등), 그 책임을 누구에게 귀속시켜야 하는지는 현재 의료법 체계 내에서 매우 복잡한 문제로 떠오르고 있다(정완, 2023). 이는 AI 시스템의 개발부터 적용까지 다양한 이해관계자가 관여하며, 단일 주체의 과실로 귀결되기 어려운 구조 때문이다. 알고리즘 개발자, 학습 데이터를 제공한 기관, AI를 도입·운영한 의료 기관, 예측 결과를 임상 판단에 활용한 의료진, 심지어는 AI 시스템 자체까지도 책임의 연쇄 고리에 포함될 수 있다.

기존 의료 과실 책임 이론이나 제조물 책임법(Product Liability)은 AI의 특수성—자율 학습 능력, 블랙박스 문제, 데이터 의존성—을 충분히 반영하지 못할 수 있으며, 따라서 현재의 법적 틀은 AI 기반 의료 기술의 복잡성과 불확실성을 효과적으로 다루기에 한계가 있다(김운명, 2023). 이에 따라 새로운 형태의 주의 기준(Standard of Care) 정립과 더불어, AI 관련 의료 사고 발생 시의 과실 입증 책임 및 손해배상 책임 분담 방식에 대한 법적·제도적 재구성이 요구된다.

책임 분산이 아닌 명확한 거버넌스 체계의 구축, 그리고 AI 시스템의 결정 과정 및 데이터 흐름에 대한 추적 가능성(traceability) 확보는 책임 규명을 위한 핵심 인프라로 간주된다(Mohammad et al., 2023). 이러한 체계가 갖추어지지 않으면, 책임 회피 가능성이 높아지고, 피해자는 실질적인 법적 구제를 받기 어려운 구조에 놓이게 된다. 더불어, 의료 현장에서 AI 오류로 인한 사고 발생 시 최종 법적 책임이 의료진에게 전가되는 ‘책임의 함정(Liability Sink)’ 문제가 제기되고 있으며, 이는 의료진의 AI 활용 기피로 이어질 수 있다(Relias Media, 2024).

또한, AI 시스템이 의료 전문가의 판단을 보조하거나 대체하는 과정에서 ‘탈숙련화(Deskilling)’ 현상이 발생할 가능성도 우려된다. 이는 의료진이 AI에 지나치게 의존하면서 임상적 직관이나 비판적 사고 능력이 저하되는 현상으로, 장기적으로 의료의 질과 환자 안전에 부정적 영향을 줄 수 있다(Beger, 2023). 반면, AI에 대한 이해 부족이나 불신으로 인해 기술을 충분히 활용하지 못하거나 오용하여 또 다른 오류를 유발할 위험도 존재한다. 이러한 상황 속에서 책임 구조가 불명확하다면, 기술 발전은 오히려 의료 환경의 불안정성과 리스크를 증가시킬 수 있다.

#### 5) 차별 가능성 및 불평등

AI 기반 건강위험 예측 도구는 잠재적으로 의료의 질과 효율성을 향상시킬 수 있는 강력한 도구이지만, 동시에 건강 불평등을 고착화하거나 심화시킬 수 있는 차별적 결과를 초래할 가능성도 존재한다.

가장 우려되는 사례 중 하나는 보험 및 고용 등에서의 차별적 활용이다. AI가 생성한 건강 위험도 예측 결과가 보험사에 의해 보험료 산정, 가입 거절, 보장 범위 제한 등에 사용될 경우, 특정 질병 소인 또는 유전적 위험 요소를 가진 개인이 구조적으로 불리한 조건에 놓일 수 있다(Thulasi Vel, 2023). 마찬가지로, 기업이 채용이나 인사 평가 과정에서 지원자의 미래 건강 위험을 고려할 경우, 고용 기회에서의 불이익이라는 새로운 형태의 건강 기반 차별이 발생할 수 있다. 이는 AI 기술이 예기치 않게 사회적 약자에 대한 구조적 배제를 강화하는 도구로 기능할 수 있음을 의미한다.

더불어 앞서 논의된 알고리즘 편향 문제는 차별 가능성을 더욱 심화시킨다. 예측 모델이 특정 인종, 성별, 소득 계층, 지역 거주자 등 취약 집단을 과소 대표하는 데이터로 훈련되었을 경우, 이들에게 제공되는 예측 결과는

낮은 정확도와 신뢰도를 보일 수 있다. 이는 해당 집단이 조기 진단이나 예방적 개입으로부터 배제되거나, 잘못된 예측에 따른 과잉 치료 또는 과소 치료를 받게 되는 결과를 초래할 수 있으며, 기존의 의료 접근성 및 건강 결과의 격차를 더욱 확대시킨다(McCradden et al., 2020).

이러한 윤리적 문제들은 상호 독립적이라기보다는 복합적이고 상호 연동된 방식으로 작용한다. 예를 들어, AI 시스템의 설명가능성 부족은 알고리즘 편향을 발견하거나 수정하기 어렵게 만들며, 이는 차별적인 결과로 이어지고 결국 AI에 대한 사회적 신뢰를 저해한다. 또한, 프라이버시 침해 문제는 개인의 자율성과 자기결정권을 약화시키고, 데이터 제공에 대한 대중의 불안을 증폭시켜 데이터 품질 자체를 저하시킬 수 있다. 따라서 AI 기반 건강 예측 기술의 개발 및 적용 과정에서는 이러한 쟁점들을 단절된 개별 이슈가 아닌 통합된 거버넌스 과제로 간주하고, 전체론적(Holistic) 접근을 취해야 한다(Gundersen & Bærøe, 2022).

또한, AI 기술의 복잡성과 예측 불가능한 파급 효과를 고려할 때, 단순히 문제가 발생한 후 이를 규제하는 사후적 규제 방식(ex post regulation)만으로는 충분하지 않다는 비판이 제기된다. 대신, 사전 예방적 접근(Proactive Governance)의 필요성이 강조된다. 즉, 윤리적 설계 원칙(Ethics by Design)의 체계적 반영, 기술 도입 이전의 잠재적 영향 평가(Impact Assessment) 실시, 실제 사용 환경에서의 지속적 모니터링 및 독립 감사 체계 마련 등이다(Gerke & Goldberg, 2023).

이러한 능동적인 거버넌스 체계는 AI 기술의 사회적 수용성을 제고하고 기술이 의도하지 않게 사회적 불평등을 심화시키는 도구로 전락하는 것을 방지하는 데 핵심적인 역할을 수행할 수 있다.

### 3. AI 기반 건강 예측 도구와 사회적 수용성

AI 기반 건강위험 예측 도구가 의료 현장에 성공적으로 도입되고 널리 활용되기 위해서는 기술적 성능과 안전성 확보뿐만 아니라, 사회 구성원들, 특히 환자와 의료 전문가들의 신뢰와 수용이 필수적이다. 사회적 수용성은 단일 요인이 아닌, 기술적, 개인적, 사회문화적, 제도적 요인들의 복잡한 상호작용에 의해 결정된다.

#### 1) 신뢰의 중심적 역할

사회적 수용성의 가장 핵심적인 요소는 신뢰이다. 유럽연합의 AI 고위전문가그룹(HLEG)이 제시한 ‘신뢰할 수 있는 AI(Trustworthy AI)’ 가이드라인은, 신뢰 형성을 위한 핵심 조건으로 다음의 7가지 요소를 제시하고 있다. 인간 주체성과 감독, 기술적 견고성과 안전성, 프라이버시와 데이터 거버넌스, 투명성, 다양성/비차별/공정성, 환경/사회적 웰빙, 책임성 등으로, 윤리적 가치 기반 위에 설계된 기술만이 사회적 신뢰를 획득할 수 있다는 점을 강조한다(Gerke et al., 2020).

신뢰는 사용자가 AI 기반 시스템과 그 결과물이 자신에게 해를 끼치지 않고 유익할 것이라는 확신에서 비롯된다. 구체적으로, 사용자는 해당 기술이 충분한 능력(competence)을 지니고 있으며, 정직성(integrity)과 선의(benevolence)를 바탕으로 설계되었다고 인식할 때 신뢰를 형성하게 된다.

AI 건강 도구에 대한 신뢰 형성에 영향을 미치는 요인은 다음과 같은 것이다(Chen et al., 2021). AI 도구의 인지된 성능과 유용성(실제로 건강 문제 해결에 도움이 될 것이라는 믿음), 사용의 용이성, 시스템 작동 방식과 결정 근거에 대한 투명성 및 설명가능성, 안전 문제, 편향 가능성 등 인지된 위험 수준, 오류 발생 시 책임 소재의 명확성, 그리고 제도적 신뢰 등이 전체적인 수용 태도에 영향을 미친다.

반대로, 신뢰를 저해하는 요인에는 AI 시스템의 잠재적 오류 가능성(Taylor et al., 2025), 학습 데이터나 알고리즘 편향성 문제, 인간적인 상호작용 및 공감 부재, 민감한 건강 데이터의 보안 및 프라이버시 침해 가능성, AI로 인한 의료 분야 일자리 감소 또는 역할 변화에 대한 불안감, 시스템 작동 방식의 불투명성(블랙박스 문제) 등이

지적되고 있다(Von Eschenbach, 2021).

## 2) 문화적 가치와 인식의 영향

AI 기술 전반에 대한 인식과 수용 태도는 국가 또는 문화권에 따라 상당한 차이를 보인다. 예를 들어, 일부 연구에 따르면 동아시아 국가(예: 중국, 한국)의 국민들은 서구 국가(예: 미국, 유럽) 국민들에 비해 AI 기술에 대해 전반적으로 더 긍정적인 태도를 보이고, 기술 수용 의향도 더 높은 경향이 있는 것으로 나타났다(Yam et al., 2023). 이러한 차이는 단순한 경제 발전 수준이나 기술 인프라 차이에 기인한다기보다는, 각 사회의 문화적 특성, 사회 제도, 가치관, 역사적 경험 등 복합적 요인의 상호작용에 의해 형성되는 것으로 해석할 수 있다.

특히, 기술 수용에 영향을 미치는 한 요인으로 기술에 대한 의인화 경향성(anthropomorphism of technology)이 지목된다. Folk et al.(2025)은 사회적 챗봇에 대한 태도에서 동아시아 문화권 사람들은 서구 문화권 사람들보다 기술을 사람처럼 인식하고 상호작용하는 경향이 높으며, 이러한 의인화 경향성이 AI 기술에 대한 친화적 태도를 형성하는 데 기여한다고 보고하였다.

또한, 정부 및 공공기관에 대한 신뢰 수준은 AI 기술 수용성에 중요한 영향을 미친다. 일반적으로, 공공기관에 대한 신뢰가 높은 국가일수록 기술 규제나 활용에 대한 시민의 불안이 낮고, 이는 기술 수용성을 높이는 방향으로 작용한다. 반대로, 공공 제도나 기업에 대한 신뢰가 낮은 국가에서는 기술이 권력을 가진 집단의 도구로 남용될 수 있다는 우려가 높아질 수 있다.

문화심리학적 관점에서 보면, 개인주의(individualism) 성향이 강한 문화에서는 개인의 자율성과 프라이버시 보호가 중요한 가치로 작용하는 반면, 집단주의(collectivism) 성향이 강한 사회에서는 사회적 조화, 공공의 이익이 상대적으로 우선되는 경향이 있다. 이 같은 가치관의 차이는 AI 기술의 위험성과 혜택을 어떻게 평가하는지, 데이터 활용에 대한 수용성은 어느 정도인지에 중요한 영향을 준다(Barnes et al., 2024).

여기에 더해, AI 기술에 대한 사회적 담론—즉 언론, 미디어, 정치 담화 등에서 AI가 어떻게 묘사되는지—도 수용성에 큰 영향을 미친다. 어떤 사회에서는 AI가 주로 경제 성장의 동력이나 혁신의 상징으로 묘사되는 반면, 다른 사회에서는 일자리 상실, 감시 강화, 프라이버시 침해와 같은 위험 중심의 서사가 강조되기도 한다.

요약하면, AI 기술 수용성은 단지 기술 자체의 문제나 성능의 문제가 아니라, 해당 기술이 도입되는 사회의 문화적 맥락, 가치관, 제도적 조건, 역사적 경험의 총체적 산물이다. 따라서 기술의 개발과 정책 수립, 그리고 실제 적용에 있어서는 서구 중심의 보편적 접근을 넘어, 각국의 고유한 사회문화적 특성과 윤리적 기대를 반영한 맞춤형 전략이 필요하다.

## 3) 규제 환경 및 거버넌스의 역할

명확하고 신뢰할 수 있는 규제 프레임워크, 윤리 가이드라인, 그리고 이를 실효성 있게 집행할 수 있는 강력한 거버넌스 체계는 AI 기술의 잠재적 위험을 통제하고 책임 있는 활용을 보장하는 데 필수적이다. 이와 같은 제도적 기반은 대중과 전문가의 신뢰를 형성하는 강력한 촉진 요인으로 작용하며, 결과적으로 AI 기술의 사회적 수용성을 높이는 데 중요한 역할을 한다. 실제로, 현재 시행 중인 규제와 보호 장치가 AI의 오·남용으로부터 개인을 충분히 보호할 수 있다는 믿음은 신뢰의 중요한 기반이 된다(Larsson, 2020).

AI 거버넌스의 궁극적인 목표는 기술 혁신을 위축시키지 않으면서도, 안전성, 보안성, 프라이버시, 공정성, 인간 존엄성과 같은 핵심 사회 가치를 보호하고 증진하는 것이다(Center for Connected Health Policy, 2025.4.23). 이를 위해 각국은 자국의 사회적 우선순위와 정치·경제적 환경에 맞는 다양한 규제 모델을 개발·운영하고 있다. 이처럼 규제 체계와 거버넌스 모델은 각국의 정치 체제, 법률 전통, 시민사회의 요구 수준, 경제 발전 단계 등에 따라 상이하게 나타나며, AI 기술의 사회적 수용성과 향후 발전 방향에도 중대한 영향을 미친다.

#### 4) 대중 및 의료진의 인식 요인

AI에 대한 수용 태도는 연령, 성별, 교육 수준, IT 활용 역량, 소득 수준 등 다양한 인구통계학적 특성에 따라 달라질 수 있다. 일반적으로는 연령이 낮고, 교육 수준과 소득이 높으며, 디지털 기술 활용 능력이 뛰어난 학생 집단이 AI에 대해 더 긍정적인 태도를 보이는 경향이 보고되어 왔다(Bewersdorff et al., 2024). 그러나 이러한 경향은 조사 국가의 문화적 맥락, 질문의 구체적 내용, 적용된 AI 기술의 유형 및 목적에 따라 다소 변동성이 있으며, 보편적인 규칙으로 일반화하기는 어렵다.

또한, 개인의 성격 특성이나 기술에 대한 태도 역시 수용성에 영향을 미친다. 기술에 대한 개방성(Openness)이 높고, 기술 불안감(Technophobia)이 낮은 개인일수록 AI에 대한 긍정적 수용 태도를 보일 가능성이 높다(Syed et al., 2025).

의료 전문가들은 AI 기반 도구에 대해 기대 성능(Performance Expectancy), 즉 해당 도구가 실제 임상 업무에 실질적 도움이 될 것이라는 실용성 인식, 그리고 이를 원활히 활용할 수 있는 조직적·기술적 기반(Facilitating Conditions)의 존재 여부에 따라 수용 의도가 좌우된다(Yousif et al., 2024).

AI 기술에 대한 수용도는 어떤 구체적 상황에서 어떤 목적으로 사용되는가에 따라 현저하게 달라진다. 예를 들어, 행정 업무 자동화나 비침습적 건강 모니터링 시스템 등 환자에게 직접적인 위해 가능성이 낮은 응용 분야는 수용도가 비교적 높다. 반면, 중환자 예후 예측이나 수술 로봇, 약물 처방 결정 등 생명과 직결되고 고위험(high-stakes) 결정을 수반하는 영역에서는 수용성 확보가 더욱 어렵다(Center for Connected Health Policy, 2025).

의료 AI는 의료진의 판단을 대체하는 기술이 아니라 지원하는 도구로 인식되어야 한다. AI 시스템에 대한 신뢰는 해당 기술의 통제 가능성, 투명성, 복잡성, 위험도, 신뢰성 등 기술적 속성과, 사용자(의료진)의 교육 수준, 경험, 선입견 등의 인간적 요인에 의해 함께 형성된다(Asan et al., 2020). 최종 결정 권한은 의료 전문가에게 유지되어야 하며, AI는 임상 판단의 보조적 도구로 사용되어야 한다는 인식 정립이 필요하다.

### III. 연구 설계

#### 1. 분석의 틀

이 연구는 AI 건강위험 예측의 윤리적 쟁점과 사회적 수용성에 영향을 미치는 요인들을 종합적으로 고려하여 국가별 비교 분석을 위한 틀을 설정한다. 분석 틀은 크게 세 가지 차원으로 구성된다.

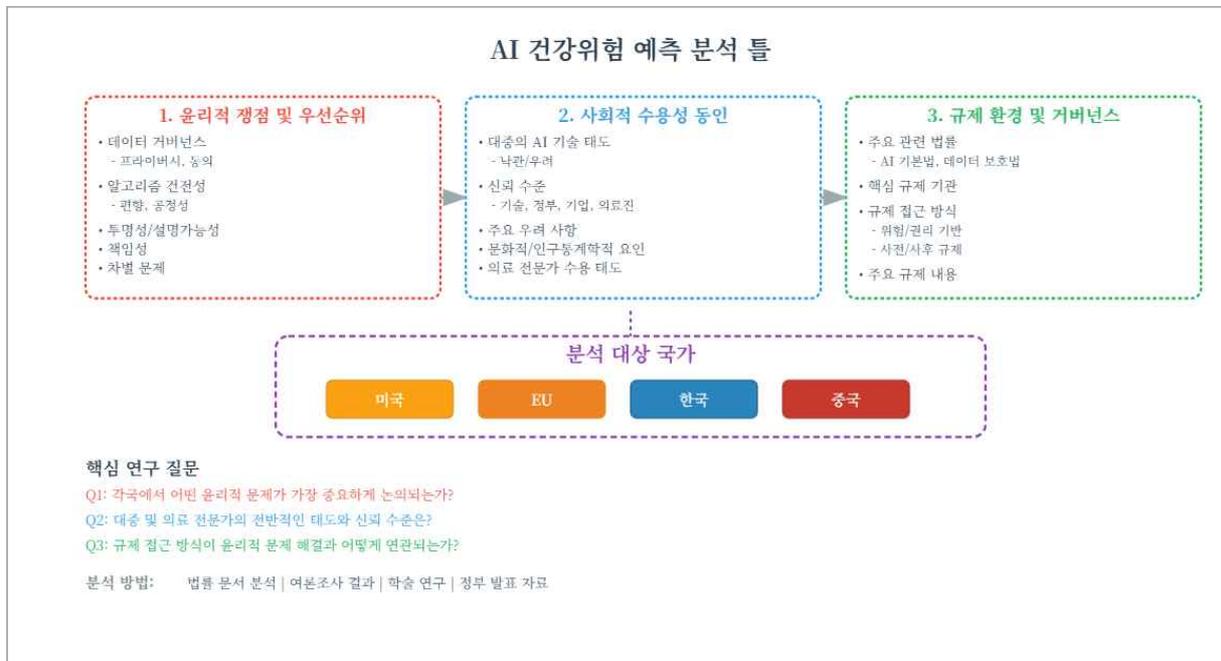
첫째, 윤리적 쟁점 및 우선순위로 각국에서 어떤 윤리적 문제(프라이버시, 편향, 투명성, 책임성, 공정성 등)가 가장 중요하게 부각되고 논의되는가? 강조점의 차이는 무엇인가?라는 연구질문을 설정하고, 비교 변수로 데이터 거버넌스(프라이버시, 동의), 알고리즘 건전성(편향, 공정성), 투명성/설명가능성, 책임성, 차별 문제를 대상으로 하여, 분석 기준으로 각 쟁점에 대한 사회적 논의의 활발성, 정책적 중요도, 관련 법규 및 지침에서의 강조점 등을 조사한다.

둘째, 사회적 수용성 동인으로 해당 기술에 대한 대중 및 의료 전문가의 전반적인 태도와 신뢰 수준은 어떠한가? 수용성에 영향을 미치는 주요 문화적, 사회적, 개인적 요인은 무엇인가?를 연구질문으로 설정하고, 비교 변수로 대중의 AI 기술 전반 및 의료 AI에 대한 태도(낙관/우려), 신뢰 수준(기술, 정부, 기업, 의료진), 주요 우려 사항, 수용성에 영향을 미치는 문화적/인구통계학적 요인, 의료 전문가의 수용 태도 및 영향 요인을 대상으로, 분석

기준으로 공개된 여론조사 결과(예: Pew, Eurobarometer, KFF 등), 관련 학술 연구 등을 조사한다.

셋째, 규제 환경 및 거버넌스로 관련 법률, 정책, 가이드라인은 어떻게 구축되어 있는가? 규제 접근 방식(예: 포괄적 vs. 분야별, 사전 vs. 사후)의 특징은 무엇이며, 이는 윤리적 문제 해결 및 수용성 증진과 어떻게 연관되는가?를 연구질문으로 설정하고, 비교 변수로 주요 관련 법률(AI 기본법, 데이터 보호법, 의료기기법 등), 핵심 규제 기관, 규제 접근 방식(위험 기반/권리 기반/분야별/수평적, 사전/사후 규제), 주요 규제 내용(데이터 사용, 알고리즘 관리, 투명성 요구, 책임 규정 등)을 대상으로, 분석 기준을 법률 및 정책 문서 내용 분석, 정부 발표 자료, 관련 연구 보고서 등을 통해 조사한다.

그림 1. 분석의 틀



분석 대상 국가는 미국, 유럽 연합(EU), 한국, 중국으로 선정하였으며, 이 지역들은 AI 기술 개발 및 활용 수준이 높고, 각기 다른 정치·경제·사회·문화적 배경과 규제 접근 방식을 가지고 있어 비교 분석에 적합하였다.

## 2. 자료 수집 방법

이 연구는 2025년 4월까지 공개된 이차 자료를 중심으로 데이터를 수집하여 활용하였다. 학술 데이터베이스 PubMed, Google Scholar, Scopus 등에서 "AI", "health risk prediction", "ethics", "social acceptance", "regulation", "governance" 및 각 국가명을 조합한 키워드로 관련 학술 논문을 검색 및 수집하였다. 정부 및 공공 기관 웹사이트로 미국(FDA, HHS, NIST 등), EU(European Commission, European Parliament 등), 한국(국가인공지능위원회, 과기정통부, 보건복지부, 식약처, 개인정보보호위원회 등), 중국(CAC, NHC, MOST 등)의 공식 웹사이트에서 관련 법률, 정책 보고서, 보도자료, 가이드라인 등을 수집하였다. 공신력 있는 연구기관 및 언론 기사로 Pew Research Center, KFF, Gallup, Eurobarometer 등 여론조사 기관의 공개 데이터, Hastings Center 등 윤리 연구기관의 보고서, 주요 언론사의 관련 기사 등을 참고하였다. 수집된 자료는 분석 틀에 따라 분류 및 정리되며, 내용 분석 방법을 활용하여 각 지역별 특징과 차이점을 도출하고 비교 분석을 수행하였다.

## IV. 연구 결과

### 1. 분석 논문의 일반적 특성

AI 기반 건강위험 예측 기술을 둘러싼 윤리적 논의, 사회적 수용성, 그리고 규제 프레임워크는 국가별로 상이한 양상을 보인다. 본 장에서는 미국, 유럽 연합, 한국, 중국을 대상으로 분석한다.

#### 1) 미국

미국의 AI 건강 예측 관련 윤리적 논의는 특히 알고리즘 편향성 문제와 이것이 건강 형평성(health equity)에 미치는 영향에 상당한 초점을 맞추고 있다. 인종 간 건강 격차를 심화시키는 구체적인 사례들이 사회적 주목을 받으며 논의를 촉발했다. 예를 들어, 의료 비용 지출 패턴에 기반한 위험 점수 알고리즘이 흑인 환자의 실제 건강 위험을 과소평가하여 필요한 의료 관리 프로그램에서 배제시킨 사례가 있다. 논문에서는 이 편향을 교정할 경우 추가 지원을 받는 흑인 환자의 비율이 17.7%에서 46.5%로 증가할 것으로 추정했다(Obermeyer et al., 2019). 이러한 문제 제기는 AI가 기존의 사회 구조적 불평등을 어떻게 반영하고 증폭시킬 수 있는지에 대한 경각심을 높였다.

데이터 프라이버시 및 보안 문제는 주로 1996년에 제정된 건강보험 이전 및 책임에 관한 법률의 틀 내에서 논의된다. 그러나 HIPAA가 AI의 복잡한 데이터 처리 방식, 대규모 데이터셋 활용, 재식별 위험 등 새로운 기술 환경이 제기하는 모든 문제를 포괄적으로 다루기에는 한계가 있다는 지적이 꾸준히 제기되고 있으며, 이에 대한 보완 및 개정 노력이 진행 중이다(Agarwal & Peta, 2024).

사회적 수용성을 보면, 미국 대중이 AI에 대해 단일한 의견을 가지지 않으며, 위험과 이점에 대한 인식에 따라 여러 세분화된 집단으로 나뉜다. 부정적(Negative) 계층(33.3%), 양면적(Ambivalent) 계층(28.5%), 미온적(Tepid) 계층(24.0%)이다(Bao et al., 2022). 공신력 있는 기관에서 실시한 최근 여론조사 결과들은 미국 대중이 AI 기술에 대해 기대감보다는 우려와 회의감을 더 크게 느끼고 있음을 보여준다. AI의 발전과 확산에 대해 흥분된다는 응답보다 우려된다는 응답의 비율이 더 높다(Kendall, 2024). 2025년에 JAMA Network Open 논문의 환자 2039명 대상 설문 조사에서는 응답자의 65.8%가 의료 AI가 책임감 있게 활용될 것이라는 점을 불신한다고 답했으며, 57.7%는 AI 시스템이 환자에게 해를 끼칠 수 있다고 우려하는 것으로 나타났다(Nong & Platt, 2025). AI가 건강 위험을 기반으로 사람들에게 대한 차별을 악화시킬 수 있다는 우려에 대해 인종별로 인식 격차가 존재한다. 흑인 응답자의 40%가 '우려' 또는 '매우 우려'한다고 답해, '약간' 또는 '전혀' 우려하지 않는다는 응답(33%)보다 높았다. 이와 대조적으로 백인 응답자의 30%가 '우려' 또는 '매우 우려'한다고 답했고, '약간' 또는 '전혀' 우려하지 않는다는 응답(44%)이 더 높았다(Beets et al., 2023). 흥미로운 점은 AI 분야 전문가들은 일반 대중보다 AI의 잠재력에 대해 훨씬 긍정적인 시각을 가지고 있어, 전문가 집단과 일반 대중 간의 인식 격차가 매우 크다는 것이다. 성별에 따른 인식 차이도 나타나, 여성이 남성보다 AI의 영향에 대해 더 신중하고 우려하는 태도를 보이는 경향이 있다. 로봇 간호사가 진단하고 투약을 결정하는 것에 대해 남성의 38%가 좋은 아이디어라고 생각한 반면, 여성은 24%만이 동의했다. 여성이 남성보다 특정 AI 적용(로봇 간호)에 대해 더 신중하고 우려하는 경향을 보여준다(Beets et al., 2023). AI와 관련된 주요 우려 사항으로는 부정확한 정보의 확산(Benzie & Montasari, 2022), 편향된 의사결정으로 인한 차별, 자동화로 인한 일자리 감소, 인간적인 관계 및 소통의 악화(Zimmerman et al., 2024) 등이 꼽힌다.

규제 프레임워크를 보면, 미국은 기존의 법률과 행정체계를 바탕으로 분야별로 분산된 규제 접근 방식을 채택하고 있으며, 이는 시장 친화적인 유연성과 자율성을 강조하는 경향을 보인다. 미국은 유럽 연합과 달리 AI 전반을

포괄하는 단일 연방 법률을 제정하기보다는, 기존의 법률 체계를 활용하고 특정 산업 분야별 규제 기관이 각자의 영역에서 AI 관련 지침을 마련하는 분야별(sectoral) 규제 접근 방식을 취하고 있다(Kuzior et al., 2024). 의료 분야에서는 식품의약국(FDA)이 AI/ML 기술을 활용하는 의료기기(Software as a Medical Device, SaMD 포함)의 허가 및 관리에 있어 핵심적인 역할을 수행한다. FDA는 AI/ML 기반 의료기기의 안전성(safety)과 유효성(effectiveness)을 보장하기 위해 위험 관리, 알고리즘 편향성 완화, 투명성 확보, 그리고 제품 전주기 관리(Total Product Life Cycle, TPLC)의 중요성을 강조하는 다수의 지침 문서를 발표하고 있다. 특히, AI 모델 개발 및 검증 과정에서 다양한 인구 집단을 대표하는 고품질 데이터를 사용할 것과, 인종, 성별, 연령 등 주요 하위 집단별로 성능을 평가하여 보고할 필요성을 강조하며 편향성 문제에 적극적으로 대응하고 있다(Mittermaier et al., 2023).

미국의 HIPAA(건강보험 이전 및 책임에 관한 법률)은 특정 공중보건 목적이나 연구 목적 등 예외적인 경우에 한해 환자의 명시적 동의 없이 PHI 사용을 허용하고 있어, 동의 기준이 GDPR보다 상대적으로 완화되어 있다. HIPAA는 건강 정보의 프라이버시와 보안을 규율하는 핵심 법률이지만, 앞서 언급했듯이 AI의 복잡한 데이터 활용 방식을 완전히 포괄하지 못할 수 있다는 한계가 지적된다(Agarwal & Peta, 2024). 질병통제예방센터(CDC)는 공중 보건 분야에서 AI를 활용할 때 고려해야 할 윤리적 지침과 건강 형평성 고려 사항을 제공한다(Shachar et al., 2020). 연방 차원 외에도 일부 주(State) 정부 차원에서 독자적인 AI 규제 법안을 제정하려는 움직임이 활발하며 (예: 콜로라도, 버지니아), 연방 의회에서도 알고리즘의 책임성을 강화하기 위한 법안(예: Algorithmic Accountability Act) 제정 논의가 지속되고 있다(MacCarthy, 2020). 전반적으로 미국은 여러 정부 부처 간 협력을 강화하고, 기술 표준 개발을 지원하며, 기존 규제 체계를 AI 기술 발전에 맞게 조정하고 확장하려는 노력을 통해 대응하고 있다.

## 2) 유럽 연합

유럽 연합(EU)의 AI 윤리 및 규제 논의는 기본권(fundamental rights) 보호, 개인 데이터 보호(특히 GDPR), 그리고 기술이 인간의 존엄성과 복지를 증진해야 한다는 인간 중심 AI(human-centric AI) 원칙을 핵심 가치이자 출발점으로 삼는다는 점에서 다른 지역과 차별화된다. 유럽연합의 일반 개인정보보호법(GDPR)은 정보주체에게 명확하고 쉽게 이해 가능한 설명을 제공하고, 언제든지 동의를 철회할 수 있는 권리를 보장한다(제7조 제3항, 제12조). 유럽연합은 위험 기반 접근 방식(risk-based approach)에 기반한 포괄적인 AI 법안을 추진하고 있으며, 이는 AI 시스템을 위험 수준에 따라 분류하고 각 위험군에 맞는 의무를 부과함으로써 예방적이고 체계적인 규제 틀을 마련하고자 한다.

EU 집행위원회가 제시한 '신뢰할 수 있는 AI를 위한 윤리 가이드라인'은 인간의 주체성 및 감독(human agency and oversight), 기술적 견고성 및 안전성(technical robustness and safety), 프라이버시 및 데이터 거버넌스, 투명성, 다양성·비차별·공정성(diversity, non-discrimination and fairness), 사회·환경적 웰빙(societal and environmental well-being), 책임성이라는 7가지 핵심 요건을 제시하며 구체적인 실천 방안을 권고한다. 특히 알고리즘의 투명성 부족, 책임성 확보의 어려움, 편향성으로 인한 차별 가능성, AI 시스템의 안전성 및 보안 문제 등이 주요 우려 사항으로 심도 깊게 논의된다(Gerke et al., 2020). GDPR의 핵심 원칙인 데이터 최소화(data minimisation) 원칙이나 목적 제한(purpose limitation) 원칙, 그리고 엄격한 동의 요건이 AI 모델 학습을 위해 방대한 데이터를 필요로 하는 기술적 요구와 상충될 수 있다는 긴장 관계 역시 중요한 문제로 인식되고 해결 방안이 모색되고 있다. EU는 궁극적으로 모든 AI 시스템이 합법적이고, 윤리적이며, 기술적으로나 사회적으로 견고한 '신뢰할 수 있는 AI'가 되어야 한다는 목표를 추구한다(Sharkov et al., 2021).

사회적 수용성을 보면, EU 공식 여론조사인 유로바로미터(Eurobarometer) 결과에 따르면, 유럽 시민들은 과학

및 기술 발전이 가져올 긍정적 영향에 대해 전반적으로 높은 신뢰(약 70% 이상)를 보이지만, 동시에 AI와 같은 신기술은 잠재적 위험을 내포하고 있으므로 신중하게 관리되고 규제되어야 한다는 인식이 매우 강하게 나타난다(약 90%의 응답자가 AI 규제 필요성에 동의). AI 기반 연구 결과의 신뢰도에 대한 질문에는 약 38%만이 신뢰한다고 응답하여, 구체적인 AI 응용 결과에 대해서는 다소 유보적인 태도를 보인다. 또한, AI 기술이 프라이버시 침해나 차별 등 인권을 위협할 수 있다는 우려도 상당한 수준(58%)으로 존재한다. AI에 대한 신뢰도나 수용성은 회원국별로, 그리고 구체적인 응용 분야별로 차이를 보이며, 특히 공공의 이익을 위해 공적 자금으로 수행된 연구 결과는 시민들이 자유롭게 접근할 수 있도록 보장해야 한다는 요구가 높게 나타난다(European Commission, 2021).

규제 프레임워크를 보면, EU는 세계 최초로 AI 기술 전반을 포괄하는 수평적(horizontal) 규제 법안인 AI 법을 제정하였다(Edwards, 2021). 이 법은 AI 시스템이 초래할 수 있는 위험의 정도에 따라 4단계로 분류한다. (1) 수용 불가능한 위험(unacceptable risk): 기본권을 침해할 명백한 위험이 되는 AI 시스템(예: 정부에 의한 사회적 점수 매기기(social scoring), 실시간 원격 생체 인식 시스템의 무분별한 사용 등)은 원칙적으로 사용이 금지된다. (2) 고위험(high-risk): 건강, 안전, 기본권에 중대한 영향을 미칠 수 있는 AI 시스템으로, 의료기기, 중요 인프라 관리, 교육, 고용, 법 집행 등에 사용되는 AI가 해당될 수 있다. 건강위험 예측 도구와 같은 의료용 AI는 대부분이 범주에 속할 가능성이 높다. 고위험 AI 시스템은 시장에 출시되기 전에 엄격한 요구사항을 준수해야 하며, 여기에는 위험 관리 시스템 구축, 고품질 데이터셋 사용 및 관리, 활동 기록 로깅, 투명성 및 정보 제공 의무, 적절한 인간 감독(human oversight) 보장, 높은 수준의 정확성·견고성·사이버보안 확보 등이 포함된다. '고위험(High Risk)' 의료 AI는 설명책임, 인간 개입 의무, 데이터 품질 등의 조건을 요구하고 위반 시 최대 3,500만 유로 또는 전세계 매출 7%의 과징금 부과를 규정하고 있다(법무부 국제법무지원과, 2024). (3) 제한된 위험(limited risk): 챗봇이나 딥페이크와 같이 사용자가 AI와 상호작용하고 있다는 사실을 인지해야 할 필요가 있는 경우, 투명성 의무(disclosure obligation)가 부과된다. (4) 최소 위험(minimal risk): 위 범주에 속하지 않는 대부분의 AI 시스템(예: 스팸 필터, 비디오 게임 AI)은 특별한 규제 의무 없이 자발적인 행동 강령 준수가 권장된다. EU의 규제 접근 방식은 잠재적 위험을 사전에 평가하고 통제하는 사전 규제(ex-ante regulation)에 중점을 둔다는 특징을 갖는다(Jiang et al., 2021).

### 3) 한국

한국의 AI 의료 분야 윤리 논의는 AI 기술 혁신을 촉진하고 관련 산업을 육성하기 위해 데이터 활용을 활성화해야 한다는 목표와, 다른 한편으로는 개인정보 보호법에 따른 민감한 건강 정보의 프라이버시를 보호해야 한다는 요구 사이에서 갈등이 존재한다(이일학, 양지현, 2023). AI 예측 모델이 오류를 발생시켜 환자에게 피해를 입혔을 경우 책임 소재에 대한 우려가 높으며, 알고리즘 '블랙박스' 문제와 이에 따른 설명가능성 확보의 필요성, 학습 데이터의 편향과 불평등 문제, 공정성 확보 방안 등도 논의 주제이다. 또한, 의료 현장에서 AI에 대한 과도한 의존이 의료진의 임상적 판단 능력 저하(탈숙련화)로 이어질 수 있다는 가능성 역시 우려 사항이다. 정부 차원에서는 '사람 중심의 AI' 실현을 국정 과제로 제시하며 기술 개발과 함께 윤리적 기반 마련과 사회적 신뢰 구축의 중요성을 강조하고 있다. AI 기술이 빠르게 발전함에 따라 이에 걸맞은 구체적이고 실효성 있는 윤리적 프레임워크와 규제 마련의 필요성이 학계, 산업계, 시민사회 등 각계에서 지속적으로 강조되고 있다(이일학, 양지현, 2023).

한국 내 의료 AI에 대한 사회적 수용성 조사 결과를 보면, 20대 한국인 대학생 40명을 대상으로 설문조사결과, 응답자들은 높은 비율로 AI 의료상담이 필요하고 직접 활용하겠다는 응답이 81%로 나타났다. 이들은 인간 의사보다 인공지능이 더 정확하고 객관적일 것으로 보았다(조민하, 2023). 이 조사는 의료 AI에 대한 특정 세대(청년층)의 초기 인식과 태도를 탐색하기 위한 예비적 연구(preliminary study)로 대표성 있는 전국 단위 조사 설계 시

참고 자료로 활용될 수 있다는 점에서 의의를 찾을 수 있다.

의사들은 생성형 인공지능(AI) ‘챗GPT(ChatGPT)’를 의료 현장에서 활용하는데 긍정적이었지만 활용 분야를 진단·처방 보조 수단으로 제한해야 한다고 봤다. 인터넷디컴퍼니가 실시한 설문조사에 따르면, 의사 1,008명중 56.8%의 의사들이 챗GPT를 활용하는 데 긍정적이었으나, 활용 분야는 제한적이어야 한다고 응답했다. 챗GPT를 의료 분야에서 활용한다면 진단·처방 보조 수단으로만 활용해야 한다는 의견이 43.8%로 가장 많았다(송수연, 2023). 직접적으로 측정한 대규모 조사는 제한적이지만, 이용 가능한 일부 설문 조사 결과들은 의료 AI의 책임 있는 활용과 안전성에 대해 상당한 수준의 불신이 존재함을 시사한다. 환자와 의료진 모두 의료 AI의 책임 소재가 명확하지 않다는 점을 가장 큰 불신 요인으로 꼽고 있다. 실제로 AI가 오진이나 오류를 낼 경우, 환자들은 누가 책임을 질 것인지에 대한 명확한 답변을 요구하고 있다.

규제 프레임워크를 보면, 한국은 AI 의료 분야의 혁신을 지원하면서도 안전성과 윤리성을 확보하기 위해 다층적인 규제 및 지원 체계를 구축하고 있다. 한국은 AI 기본법을 통해 기술 혁신을 장려하는 한편, 사회적 책임 확보를 병행하려는 중도적 접근을 시도하고 있다. 과학기술정보통신부에서는 2020년 12월 AI의 윤리에 관해 ‘사람이 중심이 되는 인공지능(AI) 윤리기준’을 마련했고, ‘인공지능 윤리 영향 평가의 도입’, ‘인공지능 윤리기준 실천을 위한 자율점검표를 개발했다(백단비, 2023). 식품의약품안전처(MFDS)는 AI 기술이 적용된 의료기기의 허가 및 심사를 담당하며, 국제 기준과의 조화를 고려하여 안전성, 성능, 임상적 유효성 검증 요건 등을 담은 가이드라인을 지속적으로 개발 및 업데이트하고 있다. 보건복지부(MOH)는 산하 국립보건연구원(KNIH)을 통해 보건의료 분야 AI 연구 수행 시 준수해야 할 윤리 지침을 발간하여 연구자들에게 방향을 제시하고, 건강보험심사평가원(HIRA)을 통해 AI 기반 혁신의료기술이 건강보험 시스템에 편입될 수 있도록 요양급여 결정 및 수가 산정 기준에 관한 가이드라인을 마련하여 기술의 현장 적용을 지원한다. 개인정보보호위원회(PIPC: Personal Information Protection Commission)는 개인정보 보호법(PIPA: Personal Information Protection Act)의 준수를 감독하는 독립 기관으로서, 특히 보건의료 데이터의 가명처리 및 이를 AI 학습 데이터로 활용하는 데 관한 구체적인 기준과 절차를 담은 가이드라인을 제공하여 데이터의 안전한 활용과 보호 사이의 균형을 모색하고 있다. 2025년 공포된 ‘인공지능 발전과 신뢰 기반 조성 등에 관한 기본법(AI 기본법)’은 AI 기술 전반에 대한 기본 원칙과 지원책을 담고 있으며, EU AI 법과 유사하게 위험 기반 접근 방식(risk-based approach)을 도입하였다. 이 법은 ‘고위험 영역 인공지능(의료 분야)에 대해 개발자 및 사업자에게 신뢰성 확보를 위한 조치(예: 위험 관리 계획 수립, 데이터 관리, 설명가능성 확보 노력, 이용자 보호 조치, 인간에 의한 감독 등)를 취하도록 의무를 부과한다. 전반적으로 한국의 규제 환경은 AI 기술 혁신을 적극적으로 촉진하면서도 발생 가능한 위험을 관리하고 윤리적 사용을 보장하려는 정책적 노력이 이루어지고 있는 단계로 평가할 수 있다.

#### 4) 중국

중국의 AI 윤리 논의는 이른바 ‘데이터 3법’이라 불리는 네트워크안전법(2017년 6월 시행), 데이터안전법(2021년 9월 시행), 개인정보보호법(PIPL: Personal Information Protection Law, 2021년 11월 시행)과 밀접한 관련이 있다(이상우, 2023). 중국의 AI 윤리 관련 논의 역시 데이터 프라이버시 및 보안(주로 개인정보보호법의 틀 안에서 논의됨), 알고리즘 편향성 문제, AI 관련 사고 발생 시 책임 소재 규명, 알고리즘의 투명성 및 해석가능성 증진 필요성, 정보주체의 동의 및 권리 보호 등 다른 국가들과 유사한 보편적인 문제들을 다룬다. 그러나 중국에서의 이러한 논의는 국가 안보나 사회 질서 유지가 더 강조된다.

사회적 수용성을 보면, 중국은 다른 주요 국가나 지역의 대중들에 비해 AI 기술에 대해 매우 높은 수준의 낙관론, 신뢰도, 그리고 수용성을 보인다. 중국은 2017년 ‘차세대 AI 발전계획’을 발표하면서 AI 분야의 중심지로 부상하고 있으며 그 배경에는 국가 권력에 의한 개인정보 침해에 대해 관대한 환경이(이찬우, 2022) 작용하고

있다. 의료 분야의 경우, 중국 종양학자 228명을 대상으로 한 설문 조사에서는 AI 도입에 대해 신중한 낙관론을 보였다. 대다수(74.13%)는 AI가 유익하며 의료를 혁신할 수 있다는 데 동의했다. AI 기술에 대한 신뢰도는 52.19%로 나타났다(Li et al., 2024).

중국은 AI를 사회 통제 및 산업 발전의 수단으로 활용하며, 특정 기술(예: 딥페이크) 및 콘텐츠(예: 사회주의 가치 저해 정보)에 대한 정밀 규제를 중심으로 한 국가 주도형 규제 체계를 운영하고 있다. 중국의 AI 규제는 국가인터넷정보판공실(Cyberspace Administration of China, CAC)이 핵심적인 역할을 수행하며, '인터넷 정보 서비스 알고리즘 추천 관리 규정', '인터넷 정보 서비스 심층 합성 관리 규정', '생성형 인공지능 서비스 관리 잠정 방법' 등을 발표했다. 이들 규정은 주로 알고리즘의 등록 및 신고 의무, 출시 전 보안 평가 요구, 플랫폼 사업자의 콘텐츠 관리 책임 강화, 학습 데이터의 합법성 및 품질 확보 의무, 사용자 권리 보호(예: 알고리즘 추천 거부권 또는 옵트아웃 권리), AI 생성 콘텐츠에 대한 명확한 라벨링 의무 등을 주요 내용으로 포함한다. 2021년에 제정된 개인정보보호법은 GDPR과 유사하게 개인 정보 처리에 대한 기본 원칙, 정보주체의 권리, 국경 간 데이터 이전 등에 대한 포괄적인 규정을 제공한다. 의료 분야에서는 국가위생건강위원회(National Health Commission, NHC)가 AI 기술의 의료 응용에 관여하며, 관련 기술 표준, 응용 시나리오 지침, 의료기기 관리 방법 등을 발표하고 있다. 또한, 과학기술 활동 전반(AI 연구 포함)에 대한 윤리 심사 절차 및 요건을 규정하는 규칙도 마련되어 있다. 종합적으로 볼 때, 중국의 AI 규제 접근 방식은 알고리즘 규제(이중희, 2023), 윤리 심사 강화, 그리고 국가 차원의 통제력 강화에 중점을 두는 것으로 요약될 수 있다.

## 2. 비교 분석

미국, 유럽 연합, 한국, 중국의 AI 기반 건강위험 예측 도구에 대한 윤리적 고려, 사회적 수용성 수준, 그리고 규제 전략에서 차이점을 고찰한다.

### 1) 윤리적 문제 비교

미국, 유럽 연합, 한국, 중국의 인공지능 기술 활용과 관련하여 개인정보(프라이버시) 보호의 중요성, 알고리즘 편향에 기인한 불공정 및 차별 문제의 해소 필요성, 알고리즘 작동 방식의 투명성과 설명가능성 확보, 그리고 AI 기술로 인한 오류 발생 시 책임 소재 명확화의 필요성 등에 대한 공통된 윤리적 우려를 공유하고 있다. 특히 환자의 안전을 보장하고, AI 기술이 초래할 수 있는 잠재적 위험을 예방 및 최소화하는 것이 모든 국가에서 핵심적인 윤리적 과제로 강조된다.

그러나 국가별로 윤리적 논의의 강조점은 다소 차이를 보인다. 유럽연합은 기본권 및 데이터 보호, 특히 일반 개인정보 보호법을 중심으로 한 윤리 및 규제 체계를 강조하며, AI 기술의 위험을 사전에 예방하는 방향으로 강력한 규제 정책을 수립하고 있다. AI 법은 사회적 점수화와 같이 민주주의 가치 및 인권을 심각하게 침해할 가능성이 있는 AI 응용 사례를 명시적으로 금지하는 등, 권리 기반의 접근 방식을 견지하고 있다.

미국은 실제 의료 현장에서 관찰되는 인종 간 건강 격차 문제에 대한 사회적 인식이 AI 윤리 논의에 반영되어, 알고리즘 편향성 완화와 건강 형평성 증진에 대한 논의가 활발히 전개되고 있다. 프라이버시 이슈는 주로 건강보험 이전 및 책임에 관한 법률 등 기존 법률 체계 내에서 다루어지며, 위험 관리 중심의 사후적 접근이 주를 이룬다. 연방 차원의 포괄적 AI 규제 법안이 부재하다는 점에서, 책임성은 미국 내 주요 윤리적 쟁점으로 부각되고 있다.

중국은 AI 윤리 논의가 사회 안정과 국가 안보 유지라는 상위 정책 목표와 밀접히 연계되어 있으며, 개인정보보호법을 통해 프라이버시 보호와 알고리즘 편향 문제를 법제화하고 있음에도 불구하고, 실제 규제 집행에서는 국가의 감독과 콘텐츠 통제가 핵심 고려 요소로 작용하는 경향이 있다. 또한 연구 개발 및 활용 과정에서 기관별

윤리 심의 체계를 강조하는 점이 특징적이다.

한국은 AI 산업 육성과 윤리적·법적 안전장치 마련 사이의 균형을 추구하는 정책 기조를 보인다. 의료 데이터를 포함한 민감 정보의 활용을 장려하면서도, 개인정보 보호법에 기반한 가명처리 가이드라인을 통해 데이터 활용과 보호 간의 균형을 도모하고자 한다. AI의 오류로 인해 발생할 수 있는 의료 사고에 대한 사회적 우려가 비교적 높은 편이며, 이에 따라 책임 소재에 관한 법제 정비 필요성이 지속적으로 제기되고 있다.

## 2) 사회적 수용성 수준 비교

의료 분야에서 AI 기술 적용에 대한 사회적 신뢰 수준을 비교하면 중국이 명확히 높은 낙관론과 높은 신뢰도를 보이는 반면, 미국과 EU, 한국은 전반적으로 의료 AI에 대한 우려와 신중론이 나타난다. 미국은 '기대보다 우려'가 높고, EU는 '유보적 태도와 규제 필요성'이 강조되며, 한국은 의료 AI의 책임 소재 불분명 문제가 불신의 주요 원인으로 지목된다(송수연, 2023). 다만 최근 젊은 세대의 수용 태도가 빠르게 변화하고 있다(조민하, 2023).

사이버 보안 위협과 개인정보 유출에 대한 우려는 모든 지역에서 공통적으로 높은 수준으로 관찰되며, 자동화로 인한 일자리 감소에 대한 불안은 특히 중국과 미국에서 상대적으로 두드러진다. 미국의 경우, AI 기술이 생성할 수 있는 부정확한 정보, 기술의 오용 가능성, 인간적인 상호작용의 감소 등이 대중과 전문가 모두에게 중요한 윤리적 문제로 인식된다.

AI 규제에 대한 인식 또한 지역 간 현저한 차이를 보인다. 중국에서는 규제의 필요성에 대한 인식이 상대적으로 낮은 반면, 기존 또는 향후 마련될 규제 체계에 대한 신뢰는 매우 높은 수준을 유지하고 있다. 반대로 미국과 EU에서는 현재의 규제가 충분하지 않다고 인식하며, 보다 강력한 통제와 윤리적 감독의 필요성을 요구한다. 이는 각국의 사회문화적 특성과 정치 제도의 차이가 AI 기술에 대한 수용성과 규제에 대한 기대를 형성하는 중요한 배경 요인임을 시사한다.

## 3) 규제 방식의 차이와 잠재적 영향

각 국가의 규제 철학과 접근 방식은 해당 국가의 고유한 사회정치적 맥락, 경제 구조, 문화적 가치를 반영한다. 미국의 분산적이고 분야별 접근 방식은 시장 중심 경제 철학과 연방 정부와 주 정부 간의 권한 배분이라는 정치적 현실을 보여준다. EU의 포괄적이고 권리 기반의 AI 법 제정은 기본권 보호와 회원국 간 규제 조화를 중시하는 유럽 통합의 가치를 반영한다. 중국의 콘텐츠, 국가 안보, 특정 첨단 기술에 초점을 맞춘 규제는 사회 안정 유지와 국가 주도의 기술 발전을 강조하는 목표와 일치한다. 한국의 접근 방식은 빠르게 변화하는 글로벌 기술 환경 속에서 국제 표준을 탐색하면서도, 자국의 산업 경쟁력 강화와 국민의 안전 및 윤리적 보호 장치 마련 사이에서 균형을 맞추려는 시도를 나타낸다.

각 국가가 채택한 인공지능 규제 방식의 차이는 해당 기술의 발전 속도와 사회적 수용성 확보에 상이한 영향을 미칠 수 있다. 유럽연합은 포괄적 사전 규제 접근(comprehensive ex-ante regulation)을 채택하고 있으며, 이는 인공지능법을 통해 특히 고위험 AI 시스템에 대해 엄격한 안전성, 투명성, 설명가능성 및 인권 보호 요건을 부과함으로써 기술에 대한 신뢰 형성과 사회적 수용성 제고에 기여하고자 한다. 이러한 접근은 장기적으로 시민의 권리 보호와 공공의 신뢰 확보에 긍정적인 영향을 줄 수 있으나, 동시에 기술 혁신의 속도를 저해하거나, 중소기업(SME: Small and Medium-sized Enterprises)에게 과도한 규제 부담을 부과할 수 있다는 우려도 제기되고 있다.

미국은 분산적 및 사후적 규제 접근(decentralized and ex-post regulation)을 택하고 있으며, 의료, 교통, 고용 등 분야별로 각기 다른 규제 기관이 역할을 수행하는 방식이다. 이는 시장의 자율성과 민간 중심의 기술 개발을 촉진하여 빠른 혁신과 산업 경쟁력 제고에 유리하게 작용할 수 있다. 그러나 통합된 연방 차원의 규제 기준 부재는 규제의 파편화(fragmentation) 및 불확실성을 야기할 수 있으며, 이는 기업 활동의 예측 가능성을 떨어뜨리고

대중의 신뢰 확보에도 제약 요인으로 작용할 수 있다.

중국은 국가 주도의 목표 지향적 규제(teleological regulation under state leadership)를 통해 전략 산업으로서의 AI를 집중 육성하고 있으며, 데이터 접근성, 대규모 실증 사업 추진, 정책 집행의 신속성 측면에서 강점을 보이고 있다. 이러한 규제 체계는 기술 상용화 속도를 가속화하고 대중의 낙관적 인식 형성에 기여할 수 있다. 다만, 개인정보 보호나 알고리즘 투명성 등의 측면에서는 상대적으로 낮은 수준의 규범적 통제가 이루어지고 있어, 국내외적으로 감시사회화, 인권 침해 등의 우려가 상존한다.

한국은 균형적 위험 기반 접근(balanced risk-based approach)을 채택하고 있으며, AI 산업 혁신을 장려하면서도 사회적 신뢰 구축과 윤리적 안전장치 마련 간의 균형을 추구하고 있다. 의료 데이터의 가명처리 기준 마련이나 ‘AI 기본법’ 제정과 같은 정책적 노력은 규제와 촉진 사이의 절충을 위한 시도로 평가된다. 그러나 이러한 접근 방식이 실질적 효과를 발휘하기 위해서는 명확한 정책 방향성 설정과 규제 간 정합성 확보, 행정 집행의 일관성 강화 등이 지속적인 과제로 남아 있다.

#### 4) 국가별 접근 방식 비교 요약

아래 표는 앞서 분석된 미국, 유럽 연합, 한국, 중국의 AI 기반 건강위험 예측 도구 관련 윤리적 우선순위, 사회적 수용성, 규제 프레임워크의 주요 특징을 요약하여 비교 제시한다.

표 1. AI 건강위험 예측 거버넌스 비교 요약

특징	미국	유럽 연합	한국	중국
주요 윤리적 우선순위	편향/형평성, HIPAA 기반 프라이버시, 책임성	기본권, GDPR 기반 데이터 보호, 금지 행위 명시, 투명성	혁신-규제 균형, 가명처리, PIPA, 책임성	사회 안정, 콘텐츠 통제, PIPL 기반 프라이버시, 기관 윤리 심의
사회적 수용성 (일반)	낮음 (우려/회의론 우세)	중간/낮음 (신중론 우세)	낮음/변화 (의료 AI 불신, 신세대 변화)	매우 높음 (낙관론/신뢰 우세)
의료진 수용성 (일반)	혼합적 (경험/연령별 차이)	혼합적 (성능/환경 의존)	낮음 (신뢰도 낮음, 경력 따라 감소)	중간 (신중한 낙관론, 연령/경험별 차이)
주요 규제 프레임워크	분야별 규제 (FDA, HIPAA), 기관 지침 (CDC), 주(State) 법률	AI 법 (위험 기반), GDPR, MDR, 윤리 가이드라인	AI 기본법 (위험 기반), PIPA, 식약처/복지부 가이드라인	특정 기술 규정 (CAC), PIPL, 위생건강위 지침, 윤리 심사 규정
접근 방식의 강점	유연성, 특정 분야 빠른 혁신 가능성	포괄성, 법적 확실성, 기본권 보호 강조	혁신 친화적 환경 조성 노력, 구체적인 데이터 활용 지침	신속한 정책 결정 및 집행, 대규모 데이터 활용 용이
접근 방식의 과제	규제 파편화, 일관성 부족, 신뢰 확보 어려움	혁신 저해 가능성, 규제 부담, 회원국 간 이행 일관성 문제	규제 간 충돌 가능성, 균형점 찾기 어려움	투명성 부족, 국가 통제 강화, 국제적 우려

## V. 결론

이 연구는 AI 기반 건강위험 예측 도구를 중심으로, 미국, 유럽연합, 한국, 중국 등 주요 4개 국가 및 지역의 윤리적 쟁점과 사회적 수용성, 그리고 규제 접근 방식의 차이를 비교 분석하였다. 분석 결과, 모든 지역에서 공통적으로 프라이버시 침해 가능성, 데이터 및 알고리즘의 편향성, 기술적 투명성 부족, 그리고 AI 관련 사고 발생 시 책임소재의 불분명성이라는 보편적인 윤리적 도전에 직면하고 있음이 확인되었다.

각 국가는 고유의 사회문화적 배경, 정치제도, 경제 전략, 그리고 중시하는 가치의 우선순위에 따라 이러한

윤리적 쟁점에 대한 접근 방식과 강조점에서 명확한 차이를 보였다. EU는 기본권 보호와 데이터 보호(GDPR)를 최우선 가치로 설정하며 포괄적이고 원칙 중심적인 규제 접근을 채택하고 있고, 미국은 실제 의료 현장에서의 건강 불평등 해소와 기존 법률(HIPAA)에 기반한 위험 관리 중심의 접근에 주력하고 있다. 중국은 사회 안정 및 국가 통제라는 거시적 틀 속에서 AI 윤리 논의를 전개하며, 한국은 산업 혁신 촉진과 윤리적 안전 확보 사이에서 정책적 균형을 시도하고 있다.

사회적 수용성 역시 지역별로 상이한 양상을 보인다. 중국은 높은 기술 낙관주의와 국가에 대한 신뢰를 바탕으로 가장 높은 수용성을 나타낸 반면, 미국과 EU는 상대적으로 우려 중심의 신중한 태도가 지배적이다. 한국은 동아시아 국가들의 일반적인 기술 수용성 경향과는 다르게, 의료 AI에 대해서는 대중과 의료진 모두에서 낮은 신뢰 수준이 관찰되었다. 이는 AI에 대한 신뢰가 단순히 기술적 성능이 아닌, 기술을 사용하는 주체에 대한 신뢰, 규제 환경의 안정성, 그리고 문화적 요인과 개인적 경험 등 다양한 요소들이 복합적으로 작용하는 맥락 의존적 개념임을 시사한다.

규제 방식에 있어서도 뚜렷한 차별성이 드러난다. EU는 사전 규제(ex-ante)와 위험 기반 접근(risk-based approach)을 중심으로 한 AI 법률 제정을 통해 강력한 윤리적 프레임워크를 구축하고 있으며, 미국은 기존 산업별 법률을 활용한 분산형 규제 및 시장 자율성에 무게를 두고 있다. 반면, 중국은 특정 기술 및 콘텐츠 중심의 목표 지향적 규제를 신속하게 도입함으로써 정책 일관성과 실행력을 강화하고 있으며, 한국은 AI 기본법 제정을 통해 혁신과 규제 사이의 균형을 추구하는 다층적 규제 체계를 구축하고 있다.

물론 이러한 차이는 정치 체제, 문화 규범, 법제, 산업 전략 등 복합적 요인들이 상호작용한 결과로 보아야 한다. 그럼에도 불구하고, 의료 AI는 환자의 생명과 안전에 직결되는 고위험·고신뢰(high-risk, high-trust) 영역이라는 점에서 공통점이 발견된다.

첫째, 위험 기반 접근(Risk-Based Approach)의 보편화이다. EU는 'AI법'에서 의료 AI를 '고위험군'으로 분류하고 엄격한 사전의무(위험관리, 데이터 품질, 인간 감독)를 규정하고 있으며, 한국의 AI기본법도 유사하다. 미국은 연방 단위의 포괄법은 없지만, FDA가 SaMD(Software as a Medical Device)로서 의료 AI의 안전성과 유효성을 검증하고 있어, 사실상 위험기반 규제를 적용하고 있다. 중국 역시 고위험 응용 분야에 대한 심의와 감독 절차를 강화하고 있다.

둘째, 편향성 완화와 건강 형평성 증진에 대한 우려이다. 알고리즘이 특정 집단에 불리한 결과를 초래할 수 있다는 문제는 네 국가 모두에서 윤리적 핵심 사안으로 다루고 있다. 미국은 특히 건강 형평성이라는 사회적 가치와 연결하여 편향 완화를 제도화하려는 시도가 활발하며, EU는 '비차별·공정성'을 AI 신뢰성 요건으로 규정하고 있다. 한국과 중국 역시 다양한 집단을 고려한 알고리즘 검증과 데이터 구축을 강조하고 있다.

셋째, 설명가능성과 책임성 확보에 대한 노력이다. 심층학습 기반 AI의 '블랙박스' 문제는 의료진의 임상적 판단과 환자의 수용성을 저해하는 주요 요인으로, 모든 국가에서 해결이 시급한 과제로 인식되고 있다. 특히 오류 발생 시 법적 책임 주체가 불명확하다는 점에서, 각국은 법률적으로 대응 중이다. 다만, 신뢰할 수 있는 AI(trustworthy AI)라는 개념이 보편적 공감대를 형성했다고 보기에는 무리가 있다. 해당 개념은 EU의 규범 틀에서 출발한 것이며, 다른 국가들은 이를 일부 참조하되 정책적 번역 방식이 다르다. 따라서 'explainable AI', 'auditable AI', 'responsible AI'와 같은 하위 개념에 대한 수렴이 보다 현실적인 협력 기반이 될 수 있다.

이러한 부분적 수렴은 "국제 표준으로의 일방적 통합"보다는, 보건의료라는 특수한 영역에서 각국이 최소한 공유 가능한 가치(예: 안전성, 비차별, 책임성)를 중심으로 규범적 공감대를 형성해 나가는 과정으로 이해해야 한다. 특히 일반 AI 분야와는 달리, 의료 AI에서는 인명과 신뢰가 걸린 문제이기 때문에 국가적 맥락을 반영한 유연한 규제 조화(flexible governance harmonization)가 필요하다.

따라서 향후 국제 협력은 의료 AI의 임상적 안전성·효과성 검증에 관한 기술 기준, 알고리즘 편향 감사 방법론

등 기술적 상호운용성 영역에서는 표준화를 추진하되, 데이터 동의 방식, 책임 분배, 거버넌스 모델 등 법적·문화적 요소는 국가별 자율성을 보장하는 이중적 접근이 바람직하다.

결론적으로, 본 논문은 네 국가의 비교 분석을 통해 의료 AI에 특화된 규범적 수립 가능성과 제약 요인을 동시에 도출하였으며, 단일한 표준 강제보다 다층적 거버넌스를 통한 책임 있는 AI 협력이 실질적인 대안이 될 수 있음을 제시한다.

이 연구의 가치는 각국의 윤리 우선순위와 규제 모델의 비교를 통해, 정책 입안자에게 사회문화적 맥락에 부합하는 맞춤형 AI 규제 설계의 실질적 근거를 제공한 것이다. 또한 '신뢰할 수 있는 AI'라는 보편적 개념이 문화적·정치적·경제적 조건에 따라 상이하게 구현됨을 보여주며, AI 윤리 연구에 있어 맥락 기반 접근의 이론적 중요성을 제시한다. 나아가 윤리적 가치, 규제 철학, 사회적 수용성에 대한 비교 분석 틀을 정립함으로써, 다문화·다체제적 AI 거버넌스 연구의 기초 자료와 방법론을 제공한다.

이 연구의 한계로는 질적 비교에 그쳐 이차 자료에 대한 의존하고, 분석 대상의 제한 등이 있으며, 향후 연구는 심층 인터뷰 및 설문은 물론 정량적 통계 자료 확보, 규제 효과성 평가, 문화권 확장, 종단적 분석 및 참여적 거버넌스 모델 개발로 보완될 필요가 있다.

본 연구는 미국, EU, 한국, 중국을 중심으로 의료 AI의 윤리 쟁점과 사회적 수용성, 규제 구조를 비교 분석하였으나, 향후 중남미, 중동, 아프리카 국가 등 비서구권의 의료 AI 적용에 대한 연구로 확장될 필요가 있다. 특히 브라질, 인도, 남아공 등도 의료 AI를 적극 도입하고 있어, 글로벌 남반구 국가들의 거버넌스 모델에 대한 추가 비교가 필요하다.

임재강은 경희대학교 행정학과에서 행정학 박사학위를 받았으며, 경운대학교 경찰행정학과에 재직하였다. 주요 관심 분야는 경찰, 행정, 국제개발이었으며 현재는 인공지능 정책과 행정 (AI+X)에 관해 연구하고 있다.

(E-mail: goodman2030@hanmail.net)

## 참고문헌

- 김윤명. (2023). 제조물책임 범위의 확장: SW와 AI의 적용가능성. *정보화정책*, 30(1), 67-88.
- 백단비. (2023). 보건의료분야 인공지능 가이드라인 현황 및 윤리적 쟁점. *보건산업정책연구 PERSPECTIVE*, 3(1), 60-72.
- 법무부 국제법무지원과. (2024). 해외규제 모니터링 제5호: 글로벌 AI 규범 동향.
- 송수연. (2023.05.09). 의사들, 챗GPT 활용 긍정적이지만 "보조수단이어야". *청년외사*. <https://www.docdocdoc.co.kr/news/articleView.html?idxno=3005647>
- 양대승, 육소영. (2022). 의료정보의 공유가능성에 대한 법적 검토-미국, 프랑스, 유럽연합을 중심으로-. *IT와 법연구*, 25집, 153-182.
- 염현주. (2025. 4. 10.). 의료 혁신 이끈 AI, 인종 성별 소득 따른 '편향성' 문제 지적도 *바이오타임즈*. <https://www.biotimes.co.kr/news/articleView.html?idxno=20780>
- 이상우. (2023). 중국의 생성형 AI 규제 동향. *주간 글로벌 ICT동향 & 이슈리포트*, 2023년 6월 1호. 정보통신산업진흥원.
- 이알하, 양지현. (2023). 의료 인공지능 개발, 활용의 사회적 영향과 윤리. *보건산업정책연구 PERSPECTIVE*, 3(1), 29-41.
- 이중희. (2023). 중국의 알고리즘 규제와 국가-규제의 목표, 실시 및 결과를 중심으로. *중국지역연구*, 10(2), 81-99.
- 이찬우. (2022). 중국 인공지능산업의 발전과 사회적 규제 논의-개인정보 보호와 디지털권위주의를 중심으로-. *중국연구*, 93, 269-293.
- 정완. (2023). 인공지능(AI)의 법적 쟁점에 관한 고찰. *경희법학*, 58(4), 117-143. <https://www.google.com/search?q=https://doi.org/10.15539/KHLJ.58.4.5>
- 조만하. (2023). 인공지능을 활용한 의료상담의 인식과 과제 - 20대 대학생 대상 설문조사를 통하여. *인문학연구*, 57, 35-64.
- Agarwal, S., & Peta, S. B. (2024). *Balancing technology and privacy: Securing patient data in healthcare under HIPAA regulations* [Preprint]. TechRxiv. <https://www.google.com/search?q=https://doi.org/10.36227/techrxiv.172710272.27746544/v1>
- Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: Focus on clinicians. *Journal of Medical Internet Research*, 22(6), e15154. <https://doi.org/10.2196/15154>
- Bao, L., Krause, N. M., Calice, M. N., Scheufele, D. A., Wirz, C. D., Brossard, D., Newman, T. P., & Xenos, M. A. (2022). Whose AI? How different publics think about AI and its social impacts. *Computers in Human Behavior*, 130, 107182. <https://doi.org/10.1016/j.chb.2022.107182>
- Barnes, A. J., Zhang, Y., & Valenzuela, A. (2024). AI and culture: Culturally dependent responses to AI systems. *Current Opinion in Psychology*, 58, Article 101935. <https://doi.org/10.1016/j.copsyc.2024.101935>
- Beets, B., Newman, T. P., Howell, E. L., Bao, L., & Yang, S. (2023). Surveying public perceptions of artificial intelligence in health care in the United States: Systematic review. *Journal of Medical Internet Research*, 25, e40337. <https://doi.org/10.2196/40337>
- Beger, J. (2023, July 18). *The paradox of progress: Machine learning and the risk of deskilling in healthcare*. <https://www.linkedin.com/pulse/paradox-progress-machine-learning-risk-deskilling-healthcare-beger/>
- Benzie, A. & Montasari, R. (2022). Artificial intelligence and the spread of mis-and disinformation. In *Artificial intelligence and national security* (pp. 1-18). [https://link.springer.com/chapter/10.1007/978-3-031-06709-9\\_1](https://link.springer.com/chapter/10.1007/978-3-031-06709-9_1)
- Bewersdorff, A., Hornberger, M., Nerdel, C., & Schiff, D. S. (2024). AI advocates and cautious critics: How AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. *Computers and Education: Artificial Intelligence*, 8.
- Binkley, C. E., Kemp, D. S., & Scully, B. B. (2022). Should We Rely on AI to Help Avoid Bias in Patient Selection for Major Surgery? *American Medical Association Journal of Ethics*, 24(8), E773-780. <https://doi.org/10.1001/amajethics.2022.773>

- Center for Connected Health Policy. (2025, 4. 23.). *Recent AI policy developments – Can lessons be learned from telehealth policy?* Center for Connected Health Policy. <https://mailchi.mp/cchpca/recent-ai-policy-developments-can-lessons-be-learned-from-telehealth-policy>
- Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K., & Ghassemi, M. (2021). Ethical machine learning in healthcare. *Annual Review of Biomedical Data Science*, 4, 123–144. <https://doi.org/10.1146/annurev-biodatasci-092820-114757>
- Edwards, L. (2021). The EU AI Act: a summary of its significance and scope. In *Artificial Intelligence (the EU AI Act)*. <https://www.adalovelaceinstitute.org/wp-content/uploads/2022/04/Expert-explainer-The-EU-AI-Act-11-April-2022.pdf>
- European Commission. (2021). *Special Eurobarometer 516: European citizens' knowledge and attitudes towards science and technology*.
- Fan, F., Xiong, J., Li, M., & Wang, G. (2020). On Interpretability of Artificial Neural Networks: A Survey. In *IEEE Transactions on Radiation and Plasma Medical Sciences*. <https://ieeexplore.ieee.org/document/9380482/>
- Fang, S., Chen, Z., & Ansell, J. (2024). Peer-induced Fairness: A Causal Approach for Algorithmic Fairness Auditing. ArXiv, abs/2408.02558.
- FDA. (2025). *Artificial Intelligence and Machine Learning in Software as a Medical Device*. U.S. Food and Drug Administration. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>
- FDA. (2025, January 6). *FDA Issues Comprehensive Draft Guidance for Developers of Artificial Intelligence-Enabled Medical Devices*. U.S. Food and Drug Administration. <https://www.fda.gov/news-events/press-announcements/fda-issues-comprehensive-draft-guidance-developers-artificial-intelligence-enabled-medical-devices>
- Folk, D. P., Wu, C., & Heine, S. J. (2025). Cultural Variation in Attitudes Toward Social Chatbots. *Journal of Cross-Cultural Psychology*, 56(3), 219-239. <https://doi.org/10.1177/00220221251317950>
- Fox, A. (2024, Dec. 30). HHS releases notice of HIPAA Security Rule update. *Healthcare IT News*. <https://www.healthcareitnews.com/news/hhs-releases-notice-hipaa-security-rule-update>
- Galvin, A., & Hogan, A. (2024). *Bias in AI and Machine Learning: The Impact of COVID-19 in African Healthcare Communities*. Proceedings of the International Workshop on Social Impact of AI for Africa 2022.
- Gerke, S., & Goldberg, D. S. (2023, May 9). *For ethical use of AI in medicine, don't overlook maintenance and repair*. The Hastings Center. <https://www.thehastingscenter.org/for-ethical-use-of-ai-in-medicine-dont-overlook-maintenance-and-repair/>
- Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In A, Bohr & K, Memarzadeh (Eds.), *Artificial Intelligence in Healthcare* (pp. 295-336). Elsevier. <https://doi.org/10.1016/B978-0-12-818438-7.00012-5>
- Gundersen, T., & Bærøe, K. (2022). The future ethics of artificial intelligence in medicine: Making sense of collaborative models. *Science and Engineering Ethics*, 28(2), 1-16. <https://doi.org/10.1007/s11948-022-00369-2>
- Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying personal genomes by surname inference. *Science*, 339(6117), 321-324.
- Hendricks-Sturup, R., Simmons, M., Anders, S., Aneni, K., Clayton, E. W., Coco, J., Collins, B., Heitman, E., Hussain, S., Joshi, K., Lemieux, J., Novak, L. L., Rubin, D. J., Shanker, A., Washington, T., Waters, G., Harris, J. W., Yin, R., Wagner, T., Yin, Z., & Malin, B. (2023). Developing Ethics and Equity Principles, Terms, and Engagement Tools to Advance Health Equity and Researcher Diversity in AI and Machine Learning: Modified Delphi Approach. *JMIR AI*. <https://doi.org/10.2196/52888>

- Herlihy, C., Truong, K., Chouldechova, A., & Dudik, M. (2024). A structured regression approach for evaluating model performance across intersectional subgroups. *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*.
- Jiang, L., Wu, Z., Xu, X., Zhan, Y., Jin, X., Wang, L., & Qiu, Y. (2021). Opportunities and challenges of artificial intelligence in the medical field: Current application, emerging problems, and problem-solving strategies. *Journal of International Medical Research*, 49(3), 03000605211000157. <https://doi.org/10.1177/03000605211000157>
- Kendall, R. A. (2024). Public perception of AI: A review. In H. Degen & S. Ntoa (Eds.), *HCI International 2024-Late Breaking Papers. Lecture Notes in Computer Science*, 15382, 72-87. Springer. [https://doi.org/10.1007/978-3-031-76827-9\\_5](https://doi.org/10.1007/978-3-031-76827-9_5)
- KFF. (2024, August 15). *Health Misinformation Tracking Poll: Artificial Intelligence and Health Information*. <https://www.kff.org/health-information-and-trust/poll-finding/kff-health-misinformation-tracking-poll-artificial-intelligence-and-health-information/>
- Kuzior, A., Sira, M., Zozuláková, V., & Hetenyi, M. (2024). Navigating AI regulation: A comparative analysis of EU and US legal frameworks. *Materials Research Proceedings*, 45, 258-266. <https://doi.org/10.21741/9781644903315-30>
- Larsson, S. (2020). *On the governance of artificial intelligence through ethics guidelines*. *Asian Journal of Law and Society*. <https://www.cambridge.org/core/journals/asian-journal-of-law-and-society/article/on-the-governance-of-artificial-intelligence-through-ethics-guidelines/992BD33CA7CBBE83E2FBBF6B0179896C>
- Li, M., Xiong, X., & Xu, B. (2024). Attitudes and perceptions of Chinese oncologists towards artificial intelligence in healthcare: a cross-sectional survey. *Frontiers in Digital Health*, 6, 1371302. <https://doi.org/10.3389/fdgth.2024.1371302>
- MacCarthy, M. (2020). *An examination of the Algorithmic Accountability Act of 2019*. [https://www.annenbergpublicpolicycenter.org/wp-content/uploads/Algorithmic\\_Accountability\\_TWG\\_MacCarthy\\_Oct\\_2019.pdf](https://www.annenbergpublicpolicycenter.org/wp-content/uploads/Algorithmic_Accountability_TWG_MacCarthy_Oct_2019.pdf)
- McCradden, M., Mazwi, M., Joshi, S., & Anderson, J. A. (2020). *When Your Only Tool Is A Hammer: Ethical Limitations of Algorithmic Fairness Solutions in Healthcare Machine Learning*. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. <https://www.semanticscholar.org/paper/1bf03cf7bb60fa84cb4e238478d0606e70c6d9d1>
- Mittermaier, M., Raza, M. M., & Kvedar, J. C. (2023). Bias in AI-based models for medical applications: Challenges and mitigation strategies. *NPJ Digital Medicine*, 6(1), Article No. 113. <https://doi.org/10.1038/s41746-023-00858-z>
- Bashayreh, M. H., Tabbara, A., & Sibai, F. N. (2023). *The Need for a Legal Standard of Care in the AI Environment*. In *Sriwijaya Law Review*. <https://www.semanticscholar.org/paper/2bbf9765fc6848910104ab387998ad29c6854b6b>
- Nalepa, G. J., Araszkiwicz, M., Nowaczyk, S., & Bobek, S. (2019). *Building trust to AI systems through explainability: Technical and legal perspectives*. Paper presented at XAILA@JURIX: Workshop on Explainable AI in Law, Saarbrücken, Germany. <https://ceur-ws.org/Vol-2681/xaila2019-paper2.pdf>
- National Conference of State Legislatures. (2025). *Summary Artificial Intelligence 2024 Legislation*. National Conference of State Legislatures. <https://www.ncsl.org/technology-and-communication/artificial-intelligence-2024-legislation>
- Nong, P., & Platt, J. (2025). Patients' Trust in Health Systems to Use Artificial Intelligence. *JAMA Network Open*, 8(2), e2460628. <https://doi.org/10.1001/jamanetworkopen.2024.60628>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453. <https://doi.org/10.1126/science.aax2342>
- Orall, A., & Rekito, B. (2025). *US Poll Results: Public Utilization and Confidence in AI for Healthcare Information. The ObG Project*. Retrieved from <https://www.obgproject.com/2025/04/09/us-poll-results-public-utilization-and-confidence-in-ai-for-healthcare-information/>

- Owens, K. (2023, May 9). *For ethical use of AI in medicine, don't overlook maintenance and repair*. The Hastings Center. <https://www.thehastingscenter.org/for-ethical-use-of-ai-in-medicine-dont-overlook-maintenance-and-repair/>
- Priyansh, & Kaur Saggu, A. (2025). Building Trust in AI Systems: A Study on User Perception and Transparent Interactions. *International Journal on Science and Technology*, 16(1), January-March 2025.
- Relias Media. (2024, March 1). *AI Creates Liability Risks for Healthcare Organizations*. Relias Media. <https://www.reliasmedia.com/articles/ai-creates-liability-risks-for-healthcare-organizations>
- Shachar, C., Gerke, S., & Adashi, E. Y. (2020). AI surveillance during pandemics: Ethical implementation imperatives. *Hastings Center Report*, 50(3), 18-21. <https://doi.org/10.1002/hast.1125>
- Sharkov, G., Todorova, C., & Varbanov, P. (2021). Strategies, policies, and standards in the EU towards a roadmap for robust and trustworthy AI certification. *ISIJ*, 50(1), 11-22. <https://doi.org/10.11610/isij.5030>
- Shuford, J. (2024). Exploring Ethical Dimensions in AI: Navigating Bias and Fairness in the Field. *Journal of Artificial Intelligence General Science (JAIGS)*, 3(1), 124. doi:10.60087/jaigs.vol03.issue01.p124
- Syed, A. Z., Memon, Z. H., Khan, K., Hameed, I., & Nadeem, M. (2025). *Examining the behavioral determinants of AI adoption in higher education: a focus on perceptual factors and demographic differences*. On the Horizon. Advance online publication. <https://doi.org/10.1108/OTH-02-2025-0019>
- Taylor, R. A., Sangal, R. B., & Smith, M. E. (2025). *Leveraging artificial intelligence to reduce diagnostic errors in emergency medicine: Challenges, opportunities, and future directions*. Academic Emergency Medicine. <https://onlinelibrary.wiley.com/doi/abs/10.1111/acem.15066>
- The Guardian. (2021, November 9). AI skin cancer diagnoses risk being less accurate for dark skin, study finds. *The Guardian*. <https://www.theguardian.com/society/2021/nov/09/ai-skin-cancer-diagnoses-risk-being-less-accurate-for-dark-skin-study>
- Thulasi Vel, D. V. (2023). *Explainable AI (XAI) for Health Insurance Underwriting*. International Journal on Recent and Innovation Trends in Computing and Communication.
- Tin, T. T., Chor, K. Y., Hui, W. J., Cheng, W. Y., Kit, C. J., Husin, W. N. A. A., Aitizaz, A., Tiung, L. K., Afolalu, S. A., & Khattak, U. F. (2024). Demographic Factors Shaping Artificial Intelligence (AI) Perspectives: Exploring Their Impact on University Students' academic performance. *Pakistan Journal of Life and Social Sciences*, 22(2), 12248-12264. <https://doi.org/10.57239/PJLSS-2024-22.2.000876>
- Verma, S., Paliwal, N., Yadav, K., & Vashist, P.C. (2024). Ethical Considerations of Bias and Fairness in AI Models. *2024 2nd International Conference on Disruptive Technologies (ICDT)*, 818-823.
- Von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. In *Philosophy & Technology*. <https://link.springer.com/article/10.1007/s13347-021-00477-0>
- World Health Organization. (2024). *Ethics and governance of artificial intelligence for health: Large multi-modal models*. WHO guidance. World Health Organization. <https://iris.who.int/bitstream/handle/10665/375579/9789240084759-eng.pdf>
- Yam, K. C., Tan, T., Jackson, J. C., & Shariff, A. (2023). Cultural differences in people's reactions and applications of robots, algorithms, and artificial intelligence. *Management and Organization Review*, 19(5), 859-875. <https://doi.org/10.1017/mor.2023.21>
- Yousif, M., Asghar, S., Akbar, J., Masood, I., Arshad, M. R., Naeem, J., Azam, A., & Iqbal, Z. (2024). Exploring the perspectives of healthcare professionals regarding artificial intelligence; acceptance and challenges. *BMC Health Services Research*, 24, 1200. <https://doi.org/10.1186/s12913-024-11667-9>
- Zimmerman, A., Janhonen, J., & Beer, E. (2024). Human/AI relationships: Challenges, downsides, and impacts on human/human relationships. *AI and Ethics*, 4, 1555-1567. <https://doi.org/10.1007/s43681-023-00348-8>

# Ethical Issues and Social Acceptance of AI-Based Health Risk Prediction Tools

Lim, Jae Kang<sup>1</sup>

<sup>1</sup> Former Professor of  
Kyungwoon University

## | Abstract |

This study conducts a comparative analysis of the ethical issues and social acceptance of artificial intelligence (AI)-based health risk prediction tools in the United States, the European Union (EU), South Korea, and China. While such tools offer transformative potential for early disease detection, personalized treatment, and healthcare resource optimization, they also pose significant ethical challenges, including concerns over data governance (privacy, security, consent), algorithmic integrity (bias, fairness), transparency and explainability, accountability, and the risk of discrimination. The findings indicate that while these regions share common concerns, they differ in their normative priorities: the EU emphasizes fundamental rights protection, the US focuses on health equity, China prioritizes social stability, and South Korea seeks a balance between innovation and regulation. Social acceptance also varies widely: China shows high optimism for medical AI, while the West expresses concerns, and South Korea remains cautious. Regulatory approaches reflect these differences, with the EU adopting comprehensive ex-ante regulation, the US favoring a sector-specific approach, China pursuing goal-oriented control, and South Korea implementing a risk-based framework. These variations are shaped by cultural norms, political structures, economic priorities, and technological readiness, offering key insights for global AI governance, international cooperation, and cross-border policy alignment.

**Keywords:** Artificial Intelligence, Health Risk Prediction, Medical AI, Ethical Issues, Social Acceptance, Transparency