

## 헬스케어 AI의 위험 및 부작용에 대한 실체적·절차적 규제 방법

장욱\*

연세대학교 보건대학원

## Substantive and Procedural Methods to Regulate the Risks and Side Effects of Healthcare AI

Wook Jang\*

Graduate School of Public Health, Yonsei University, Seoul, Korea

Received: Nov 6, 2024  
Revised: Nov 14, 2024  
Accepted: Dec 28, 2024

## \*Corresponding author

Wook Jang  
Graduate School of Public Health, Yonsei University, Seoul, Korea.  
Tel: +82-2-2228-2532  
E-mail: vonjang@hanmail.netCopyright © 2024 Korean Society for Medical Ethics. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Conflict of interests

No potential conflict of interest relevant to this article was reported.

## Acknowledgements

Not applicable.

## Funding information

Not applicable.

## Data availability

Upon reasonable request, the datasets of this study can be available from the corresponding author.

## ORCID

Wook Jang  
<https://orcid.org/0009-0001-7885-1732>

## Author contributions

The article is prepared by a single author.

## Abstract

As advances in artificial intelligence (AI) research and development edge closer to achieving Artificial General Intelligence (AGI), concerns about the risks and side effects of healthcare AI are growing. In response to these concerns, this study reviews the regulatory measures proposed in the lead article for this volume, which are categorized into substantive and procedural approaches to AI regulation. Special attention is given to the verification entities described in the procedural approach to critically assess their role in mitigating the risks or unintended consequences of healthcare AI.

**Keywords:** artificial general intelligence (AGI); substantive methods; procedural methods; verification entity

## I. 서론

인공지능(artificial intelligence, AI)이란 용어는 이미 수십 년 전부터 사용되어 오고 있었고, 인간의 생활을 윤곽하게 하고 보조적 역할을 하는 약한 인공지능(weak AI)를 중심으로 법적·윤리적 측면이 논의되어 왔었다. 그러다가 2022년 OpenAI사 개발한 ChatGPT가 출시되면서 인간을 대체할 수 있거나 인간의 지적 판단능력을 증가하는 범용인공지능(artificial general intelligence, AGI)의 개발이 공상과학이나 먼 미래에서나 실현가능한 일이 아닌 우리 눈 앞에 현실로 다가올 수 있음[1]을 절감하고 있다.<sup>1)</sup> 헬스케어 영역에서도 AI 기술의 접목을 통해 각종 질병의 예방, 진단, 치료, 관리 등의 다양한 의료서비스 제공에 있어서 혁신적인 변화가 이루어질 것이라고 예상되고 있다. 이러한 상황에서 AI로 인해 발생할 수 있는 위험이나 부작용에 대한 우려 또한 커지고 있으며, 이를 방지하기 위한 법적·윤리적 대응방안들이 적극적으로 논의되기

1) 영국의 컴퓨터과학자로서 AI를 통한 머신러닝 연구로 노벨물리학상을 받은 제프리 힌튼도 몇 년 전까지도 범용인공지능의 개발은 먼 미래의 일이라고 치부하였다가 ChatGPT의 개발로 이전의 견해를 철회하였다. 그리고 2023년 구글을 퇴사하고 AI에 대한 규제필요성을 주장하면서 Future of Life 재단의 인공지능 개발 일시 중단 성명서에 서명하기도 하였다.

Ethics approval  
Not applicable.

시작하고 있다.

헬스케어 AI로 인한 위험이나 부작용을 합리적으로 방지하기 위한 방안으로 김준혁은 논문에서 헬스케어 AI 윤리에서 환자·시민 참여모형을 제안하고 있다고 생각한다. 본인은 김준혁이 제안하고 있는 헬스케어 AI 윤리에서 환자·시민 참여모형에 기본적으로 공감하면서 헬스케어 AI의 위험 및 부작용과 그를 방지하기 위한 실체적 방법과 절차적 방법을 구분하여 간략히 기술하고자 한다. 그리고 그 과정에서 김준혁이 논문에서 주장하는 바와 차이가 있는 점에 대해서도 언급하고자 한다.

## II. 헬스케어 AI의 위험 내지 부작용

헬스케어 AI의 경우 발생가능한 위험 내지 부작용들로 여러 가지 것[2]들이 있을 수 있지만 주요하게 논의되고 있는 것들은 다음의 3가지로 요약될 수 있다.<sup>2)</sup>

첫째, 기존의 의료행위 결정과정을 보면 의료행위에 대한 표준가이드라인이 있고 동일한 질환을 가진 환자라고 하더라도 의학적 데이터 이외에 경제 사회적 다른 여러 상황을 고려하여 최선의 방법이라고 생각할 수 있는 치료법이 행해질 수 있다. 그러나 헬스케어 AI의 경우에는 축적된 의학적 데이터에 따라 동일한 질환의 경우 동일한 치료법이 제안될 가능성이 클 수 밖에 없다. 즉 헬스케어 AI가 행하는 의료행위<sup>3)</sup>가 의료윤리적 견지에서 옳다고 평가될 수 없는 측면이 있을 수 있다[3].

둘째, 헬스케어 AI에 적용되는 알고리즘은 인간이 쉽게 파악할 수 없는 딥러닝 과정을 거쳐 이루어질 것이기 때문에 개발자라고 하더라도 모든 상황을 통제할 수는 없을 것이다. 특히 최근 범용인공지능(AGI)의 개발이 가시화된 시점에서는 AI가 행하는 의료행위에 대해 인간의 통제범위를 벗어나서 이루어지지 않을까 하는 불안감은 더욱 증폭될 수 있다.

셋째, 헬스케어 AI 알고리즘의 가치편향성 및 공정성 문제이다. AI 알고리즘이 가치편향성을 가지게 되는 이유에는 개발자인 개인의 편향된 가치판단기준에 알고리즘 설계에 반영된 것일 수도 있고, 부지불식간에 사회구조적으로 편향된 데이터들이 알고리즘에 반영되어 나타날 수 있다. 이러한 알고리즘에 기반한 헬스케어 AI가 수행하는 의료행위의 경우 인종 간, 남녀 간 또는 사회소수자에 대한 차별적 성향을 가진 채 의료행위를 수행할 위험성도 제기되고 있다. 그리고 이러한 행위들이 명시적으로는 잘 드러나지 않고 부지불식간에 이루어지는 경우 사후적으로도 교정이 어렵게 되거나 고착될 우려도 있기 때문에 더욱 문제가 될 수 있다.

위에서 언급한 이러한 헬스케어 AI가 가질 수 있는 위험 및 부작용 등으로 인해 헬스케어

2) 이 부분은 나해란/김현성, 빅데이터, 인공지능시대 의료윤리, The Journal of Korean Diabetes, vol.21, 2020 내용을 참고로 재구성하여 기술함.

3) AI가 하는 의료행위에 대한 법적인 측면에서의 규제방안과 관련해서는 의료행위 영역과 비의료건강관리 영역으로 나누어서 살펴보아야 한다. 먼저 전자의 경우에는 진단, 치료, 예방 등 직접적인 의료행위와 관련되거나 보조하는 AI 기술이 관련될 수 있고, 의료법, 의료기기법, 개인정보보호법 등의 규제영역에 속한다. 또한, 이러한 의료행위 영역에 관련된 AI에 대해서는 안전성과 유효성 검증에 있어 보다 엄격한 규제방식이 요구될 것이다. 후자의 경우에는 건강증진 및 관리, 생활습관개선 등에 사용되는 AI 기술로 현행법상 의료행위로 간주되지 않은 웨어러블 기기를 통한 건강관리, 헬스케어 앱등이 포함될 수 있다. 이러한 비의료건강관리는 소비자보호법, 전자상거래법 및 개인정보보호법 등이 주요 규제영역이 될 수 있고, 의료행위 영역에서의 AI 기술보다는 상대적으로 완화된 규제를 받게 될 것이다. 헬스케어 AI의 특성에 따른 보다 구체적인 규제방안이나 그 타당성에 대한 논의는 의료윤리학회지 보다는 의료법학회지에서 논문으로 다루는 것이 타당하다고 보여줘 본 논평에서는 생략하기로 한다.

AI 개발에 있어서의 윤리적 원칙 및 규제 시스템을 통해 보건의료의 민주화와 공정성<sup>4)</sup>을 지키고자 하는 것이다[4]. 헬스케어 AI의 위험 내지 부작용을 방지할 수 있는 방법으로 실체적 규제방법과 절차적 규제방법으로 나누어 볼 수 있는데 이에 대해서는 단락을 바꾸어 논의하겠다.

### III. 헬스케어 AI의 위험 및 부작용 방지를 위한 규제 방법

#### 1. 실체적 규제방법

헬스케어 AI의 위험 및 부작용 방지를 위해 WHO는 2021년 ‘의료분야 인공지능윤리와 거버넌스 지침서(Ethics & Governance of Artificial Intelligence for Health)’를 통해 (1) 자율성 보호, (2) 인간의 복지와 안전 증진, (3) 투명성, 설명성, 지적 능력의 보장, (4) 책임의 함양, (5) 포괄성과 형평성 보장, (6) 대응성과 지속 가능한 AI를 촉진하는 것 등의 핵심원칙에 따라 AI가 개발되어야 함을 천명하였다[5]. 그러면서 의료진이 AI 시스템과 의사결정에 대한 완전한 통제권을 유지해야 함을 명시하고 있다. 이러한 AI 윤리는 2024.3.14 EU가 제정된 AI 규제법에도 상당부분 반영되어 있다. 그리고 우리나라에서도 2023년 8월 질병관리청과 국립보건연구원이 공동으로 ‘보건의료분야 인공지능 연구자를 위한 연구윤리 지침’(이하 ‘연구윤리 지침’이라 한다.)을 개발하여 시행하고 있다. 이 지침에서도 WHO의 지침[6]의 영향을 받아 인공지능 연구·개발에 있어 연구자와 기관의 사회적 책임을 구현하기 위한 윤리원칙으로 (1) 인간의 자율성 존중과 보호, (2) 인간의 행복, 안전, 공공의 이익 증진, (3) 투명성, 설명 가능성, 신뢰성, (4) 책무, 법적 책임, (5) 포괄성, 공정성, (6) 대응성, 지속 가능성을 제시하고 있다.<sup>5)</sup> 따라서 현재 시행되고 있는 이러한 AI 윤리원칙들이 헬스케어 AI를 개발함에 있어서 실체적 규제의 주요한 내용이 될 것이다.

#### 2. 절차적 규제방법

헬스케어 AI의 위험 내지 부작용을 방지하기 위한 절차적 방법은 위에서 언급한 실체적 방법을 어떻게 합리적 절차를 통해 구현해 나갈 것인가의 문제일 것이다. 질병관리청이 제시하고 있는 연구윤리지침에 따르면 AI 연구·개발 절차적 고려사항으로 다음의 6가지를 제시하고 있다. (1) 사전고려 단계에서 헬스케어 AI 연구·개발 착수전에 지침에 근거해서 연구의 명확한 윤리 프레임 워크를 설정했는지 여부, (2) 데이터셋 생성단계에서는 헬스케어 AI 모델개발을 위한 데이터 수집과 처리과정에서 개인정보보호 조치들이 충분히 이루어졌는지 여부, (3) 모델 개발 단계에서는 연구목적에 맞는 알고리즘을 설정하고 예비데이터를 적용하여 적절성을 판단했는지 여부, (4) 훈련, 검증, 평가 단계에서는 수집한 데이터로 알고리즘 훈련 및 검

4) 논평의 대상인 김준혁 논문에서 의료윤리 원칙에 입각한 보건의료행위를 보건의료의 민주화와 공정성을 가진 행위로 정의하고 있는 것으로 보어져 이 논평에서도 기존의 의료윤리 원칙에 입각한 보건의료의 이행이라는 측면에서 김준혁의 논문에서 쓰고 있는 용어를 그대로 원용하기로 한다.

5) 김준혁은 논문에서 헬스케어 AI 윤리는 기존의 생명윤리 원칙의 연장선상 내지 일맥 상통한 것으로 이해할 수 있다고 설명하고 있다.

증, 연구목적에 적용가능 여부에 대한 평가가 제대로 이루어졌는지 여부, (5) 적용단계에서는 개발된 모델이 사전 설정한 윤리 프레임 워크와 법적 규제를 준수했는지 여부, (6) 사후 고려 단계에서는 헬스케어 AI 모델의 향상을 위해 사용자와 지속적인 소통과 피드백을 하고 있는지 여부이다. 이러한 사항들은 AI 연구·개발을 단계별 절차로 나누어 검토되어야 할 사항들을 적시하였다는 점에서 절차적 사항으로 볼 수 있지만, 적절한 검증기관이나 주체에 의해서 이루어졌는지에 대한 논의가 빠졌다는 점에서 제대로 된 절차적 규제방법이라고는 할 수 없다. 이러한 점에서 김준혁도 논문에서 헬스케어 AI 연구·개발에 있어 신뢰성을 보장하기 위한 윤리적 논의나 절차적 고려는 거의 없었다는 문제점을 지적하고 그 대안으로 검증 주체로서 환자·시민 참여 모형을 제시하고 있다고 생각한다. 이러한 그의 문제점 지적에 전적으로 동감하는 바이며 독립적인 외부기관에 의해 단계별로 실제적인 사항들을 고려하여 개발되고 있는지를 검토해야 한다고 생각한다. 그리고 김준혁은 전문가, 전문검증기관 및 국가의 경우에는 검증 주체로서 정당성이나 합법성을 인정할 수 없기 때문에 외부검증 주체로서 환자·시민이 주체가 되어야 함을 주장하고 있는 것으로 보인다. 그러나 그러한 그의 주장에는 일견 동의할 수 없는 측면이 있다. 환자·시민이 헬스케어 AI로 영향을 받는 주요한 이해관계자로서 검증 주체로 참여하여야 함에는 동의하지만 전문가 및 국가기관을 배제하고 주체가 되어야 함에는 동의할 수 없는 측면이 있다. 전문가, 국가기관, 그리고 환자·시민은 각각의 고유한 영역에서의 역할 및 기능이 있기 때문에 이들 모두가 참여하는 협의체를 외부검증주체로 하여 각각의 단계별로 검토가 이루어지는 것이 바람직하다고 생각한다.

#### IV. 결론

AI 연구·개발이 가속화되면서 범용인공지능(AGI) 출현이 가시화되고 있는 상황에서 헬스케어 AI로 인한 위험 또는 부작용의 우려 또한 커지고 있다. 이러한 상황에서 헬스케어 AI 규제방안은 실제적 측면과 절차적 측면이 동시에 고려되어야 한다. 실제적 규제방법은 윤리 지침 등을 통해 어느 정도 구체화되었다고 볼 수 있지만 절차적 규제방법과 관련해서는 검증주체와 관련해서 여전히 모호한 측면이 있다. 절차적 측면에서는 우선적으로 헬스케어 연구·개발 및 실용화 각 단계에서 실제적 규제 내용들이 제대로 작동하고 있는지를 검증할 수 있는 주체가 명확히 설정되어야 한다. 그러한 검증은 외부전문가 내지 전문기관, 국가기관 및 환자·소비자가 모두 참여한 협의체를 통해 이루어져야 할 것이다. ㉞

#### REFERENCES

1. Na HR, Ki HS. Medical ethics in the era of artificial intelligence based on medical big data. J Korean Diabetes 2020;21:126-129. <https://doi.org/10.4093/jkd.2020.21.3.126>
2. Yang H. Safety issues of artificial intelligence and policy responses. J Korean Inst Commun Inf Sci 2018;43(10):1724-1732. <https://doi.org/10.7840/kics.2018.43.10.1724>
3. Le WB. U.S. Antidiscrimination law in healthcare & algorithmic bias: its lessons for South Korean legislation. Public Law J 2023;25(2):175-203. <https://doi.org/10.31779/plj.25.2.202405.007>
4. Lee SB, Song J, Park AP. A trend of artificial intelligence in the healthcare. J Korea Content

Assoc 2020;20(5):448-456.

<https://doi.org/10.5392/JKCA.2020.20.05.448>

5. Korean National Institute of Health. Research ethics guidelines: for healthcare AI researchers. National Institute of Health; 2023.
6. World Health Organization (WHO). Ethics and governance of artificial intelligence for health [Internet]. WHO; 2021 [cited 2024 Oct. 30].  
<https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use>

