KAIRB
ISSN 2671-4388 (Online)
Original article
Journal of KAIRB 2022;4(1):16-22

보건의료빅데이터 연구에 대한 대중의 인식도 조사 및 윤리적 고찰

[1]서울대학교병원 임상연구윤리센터, [2]가톨릭대학교 생명대학원 임상연구윤리학과

조수진[1], 최병인[2]

The Overview of the Public Opinion Survey and Emerging Ethical Challenges in the Healthcare Big Data Research

Su Jin Cho[1], Byung In Choe[2]

[1]Center for Human Research Protection Program, Seoul National University Hospital, Seoul, Korea
[2]Department of Institutional Review and Research Ethics, Nicholas Cardinal Cheong Graduate School for Life, The Catholic University of Korea, Seoul, Korea

Purpose: The traditional ethical study only suggests a blurred insight on the research using medical big data, especially in this rapid-changing and demanding environment which is called "4th Industry Revolution." Current institutional/ethical issues in big data research need to approach with the thoughtful insight of past ethical study reflecting the understanding of present conditions of this study. This study aims to examine the ethical issues that are emerging in recent health care big data research. So, this study aims to survey the public perceptions on of health care big data as part of the process of public discourse and the acceptance of the utility and provision of big data research as a subject of health care information. In addition, the emerging ethical challenges and how to comply with ethical principles in accordance with principles of the Belmont report will be discussed.
Methods: Survey was conducted from June 3th August to 6th September 2020. The online survey was conducted through voluntary participation through Internet users. A total of 319 people who completed the survey (±5.49%P [95% confidence level] were analyzed.
Results: In the area of the public's perspective, the survey showed that the medical information is useful for new medical development, but it is also necessary to obtain consents from subjects in order to use that medical information for various research purposes. In addition, many people were more concerned about the possibility of re-identifying personal information in medical big data. Therefore, they mentioned the necessity of transparency and privacy protection in the use of medical information.
Conclusion: Big data on medical care is a core resource for the development of medicine directly related to human life, and it is necessary to open up medical data in order to realize the public good. But the ethical principles should not be overlooked. The right to self-determination must be guaranteed by means of clear, diverse consent or withdrawal of subjects, and processed in a lawful, fair and transparent manner in the processing of personal information. In addition, scientific and ethical validity of medical big data research is indispensable. Such ethical healthcare data is the only key that will lead to innovation in the future.

Keywords: Bioethics, Research ethics, Big data, Consent, The right to personal information self-determination, Privacy

서론

1. 연구 배경

2016년 6월 스위스에서 열린 세계경제포럼에서는 4차 산업을 디지털 혁명(digital revolution)으로 정의하였다. 4차 산업 시대에서의 핵심 기술은 사물인터넷(internet of things), 인공지능(artificial intelligence, AI), 그리고 빅데이터(big data)이다.[1]

우리나라의 가장 전통적인 보건 의료 데이터는 각 병원마다 가지고 있는 환자에 대한 의무기록(medical records)이며, 우리나라는 2000년대 초반 의료법 개정, 그리고 이후 전자서명법의 통과로 전자문서의 형태로 의무기록을 작성, 보관, 관리

Submitted: 29 December 2021, Revised: 21 January 2022, Accepted: 25 January 2022
Corresponding author: Byung In Choe, Department of Institutional Review and Research Ethics, Nicholas Cardinal Cheong Graduate School for Life, The Catholic University of Korea, 222 Banpodae-ro, Seocho-ku, Seoul 06591, Korea
Tel: +82-2-2258-7395, Email: bichoe@catholic.ac.kr, https://orcid.org/0000-0002-7401-2798

Copyright © 2022 The Korean Association of Institutional Review Boards

할 수 있게 되었고, 전자 의무기록(electronic health/medical record)의 도입으로, 의료기관에서 발생한 다수의 의료정보 및 의료 영상 자료가 전자문서의 형태로 관리되고 있다. 이는 가장 방대한 의료 빅데이터의 기반이 된다.

또한. 우리나라는 건강보험, 국민연금 등의 사회보험제도라는 특수성을 가지고 있고, 이로 인해 건강보험심사평가원, 질병관리본부 등 공공기관이 보유/관리하는 질환별 데이터베이스, 건강보험 현황, 의약품 이용 현황, 대국민 건강 행태 등의 다양하고 방대한 데이터들이 '공공데이터의 제공 및 이용 활성화에 관한 법률'에 따라, 민간에게 제공되고 다양한 목적으로 활용되고 있다. 그 밖에도 다양한 웨어러블 스마트 기기 등을 통하여 개인별 고유의 생체 정보 등이 수집·생성되기도 하며, 일상생활에 대한 개인별 라이프로그(lifelog) 등도 다양한 보건의료 빅데이터의 소스로 활용될 수 있다.

의료 분야에서의 빅데이터는 새로운 의료기술을 개발/개선하거나, 바이오 의약산업의 발전을 위한 새로운 신약의 개발 및 국민의 건강 안보를 지키는 신뢰성 있는 정부의 보건의료 정책을 수립하는데 활용되기도 한다.[2] 또한, 환자의 질병의 진단에 있어 인공지능(AI) 등을 통한 임상 결정 지원 및 의료 데이터와 유전정보 등을 기반으로 '환자 맞춤형' 미래 질환의 예측/관리 등 의료의 질과 환자의 건강을 향상시킬 수 있는, 국민 건강과 생명에 직결되는 핵심적인 자원이라고 할 수 있다.

2. 연구 목적

우리가 원하든 원치 않던 매일 다르게 엄청난 양의 의료 데이터가 생성되고 축적되며, 내 의도와는 상관없이 누구나 의료 데이터 정보 제공자가 되는 상황에서 의료 빅데이터를 활용한 과학과 기술의 진보의 시대적 흐름은 거스르기 어렵다. 하지만, 의료 빅데이터의 효용성과 의료산업 발전에만 맞춰져, 일반 대중이 우려하는 재식별화에 따른 프라이버시의 침해, 사회적 차별, 그리고 데이터 수집에 있어 공정성과 동의에 근간이 되는 개인의 존엄성과 자율성에 대한 윤리적 고려와 방법론은 간과되고 있다. 연구자가 아닌 일반 대중들의 입장에서 의료 빅데이터에 대한 인식과 보건의료 정보 주체자로서 빅데이터 연구의 활용/제공에 대한 수용도 등에 대하여 사회적 인식을 조사해 보고, 새로운 윤리적 관점을 고찰해 보고자 한다.

연구 방법

사회 담론의 형성 과정의 일환으로 일반 대중들이 보건의료정보에 대한 어떤 인식을 가지고 있는지, 해당 데이터의 활용 및 제공에 대하여 얼마나 포용하는지 알아보고, 보건의료 빅데이터에 대한 제공, 활용 연구에 대한 대중의 인식도를 조사하였다. 설문조사는 온라인 서베이(http://www.dooit.co.kr/)를 통해, 2020년 8월 3일부터 2020년 9월 6일까지 진행되었으며, 인터넷 사용자를 통한 자발적 참여로 이루어 졌다. 이 설문조사 연구는 가톨릭대학교 성의교정 의학연구윤리심의위원회의 승인을 받고 진행하였다(IRB 승인번호: MIRB-면20200807-001).

조사 항목은 인구학적 통계(성별, 연령, 학력, 직업)를 비롯하여, 보건의료정보의 개념, 권리, 개인정보 보호와 연구 목적 활용에 대한 인식을 조사하였고, 의료 빅데이터의 개념과 의료 빅데이터 연구에 대한 대중의 수용도, 의료 빅데이터 연구에서 뒷받침되어야 할 제도적 장치 등의 문항으로 구성하였다.

연구 결과

설문조사는 성인을 대상으로 20대부터 50대 이상까지 총 380명이 응답하였으나, 설문을 끝까지 완료하는 것을 이 연구의 자발적 참여 동의로 간주하였기 때문에, 설문을 완료한 총 319명만 설문 분석[±5.49%P (95% 신뢰수준)]을 시행하였다. 보건의료정보에 대한 인식 및 정보 주체 권리에 대한 인식, 보건의료 빅데이터에 대한 인식 및 제공/활용에 대한 수용도 등은 5점 리커트 척도 평가로 시행되었다. 또한, 보건의료 빅데이터를 통한 다양한 연구, 보건의료 정책, 의료 기술, 의학 산업 등으로 활용되기 위해서 뒷받침되어야 할 제도적/사회적 장치 등은 우선순위로 답할 수 있도록 하여 중복응답에 대해 가중치를 부여하여 분석하였다.

1. 의료정보에 대한 인식과 의료정보 주체자로서의 권리에 대한 인식

응답자 대부분은 의료정보가 민감한 개인정보(86.8%)이며, 실제 나의 질병의 치료에서 중요한 정보(81.2%)라고 응답했다. 또한 의료정보가 새로운 의학 발전을 위한 연구에 도움이 될 수 있다고(68.7%)응답하였다.

그렇지만, 본인의 의료 정보가 분실, 오용, 무단 공개 및 유출 가능성에 대한 우려가 있다고 응답한 대상자는 절반 이상의 64.3%였고, 만약 해당 의료정보가 유출된다면 나의 프라이버시가 침해당할 것이라고 응답한 수는 78.4%에 달했다. 의료기관에서 의료 정보를 적절하게 보호할 것이라는 신뢰도에 대해서는 낮은 수 42.4%만이 '대체로 그렇다' 답하였고, 신뢰할 수 없다는 응답자 또한 25.4%였다. 응답자의 대부분이 의료정보 보호와 프라이버시에 대한 우려가 크며, 의료정보 보호에 대한 기관에 대한 신뢰가 낮게 조사되었다.

의료정보에 대한 주체와 소유권에 대한 대중에 인식조사에서는 의무기록 등이 의료기관의 검사 기록이며, 의료인들의 의학적 판단에 대한 기록이기도 하지만, 보건의료정보는 나의 치료 기록에 대한 개인정보이므로 해당 정보의 소유의 권리는 환자

에게 있다고 응답한 수는 82.5%으로, 이는 많은 대중이 의료기관 내 의료정보 주체로서의 권리가 있다고 생각하는 것을 보여준다.

의료기관 내 의료정보가 새로운 의학 발전을 위한 연구에 도움이 될 수 있는지에 대하여 상당수 동의하였지만, 본인의 의료정보가 연구 목적 등으로 활용되는 것에 대하여 대부분(91.5%)이 나의 사전 동의가 필요하다고 답하였다. 또한, 나의 의료정보가 다른 연구 목적 등으로 활용되었다 하더라도, 언제든지 내 의사에 따라 사후에라도 동의를 철회할 수 있는 절차가 마련되어야 한다고 응답한 대상자 또한 86.5%에 달했다. 의료기관 등에서 동의를 얻기 위한 방법으로 연락처 등의 개인정보를 활용한 개별 연락(전화, 문자 등) 허용 여부에 대해서는 55.5%가 허용하는 입장 이였고, 개별 연락에 부정적인 입장은 15.4%였다(Table 1).

의료정보의 연구 목적 활용에 대하여 다수의 응답자(69.9%)는 새로운 의학 발전을 위하여 나의 보건의료정보가 연구 목적 등으로 제공/활용되는 것에 대하여 개별적인 사전 동의가 있다면 제한적으로 제공할 의향이 있다고 응답하였다. 만약, 의료기관이 나의 의료정보에 대한 개인정보 보호를 철저히 지켜줄 것이라는 신뢰가 있다면, 내 의료정보가 포괄적으로 활용되도록 제공할 의향이 있다고 응답한 수도 14.1%로 나타났다. 이는 의료기관이 체계적으로 환자의 개인정보 보호 방안을 마련하고, 의료정보 활용에 대한 투명성과 프라이버시 보호에 대한 신뢰도 구축의 중요성이 시사된다(Fig. 1).

2. 의료 빅데이터에 대한 인식과 제공/활용에 대한 수용도

단일 정보의 의료정보보다 무궁한 잠재력을 가진 “의료 빅데이터”에 대한 인식 및 의료 빅데이터의 구축과 공공 또는 상업적 목적의 제공/활용에 대한 대중의 수용도를 조사하였다.

의학 발전에 있어 의료 빅데이터가 핵심적 가치가 될 것이며(69.0%), 다양한 의료 빅데이터로 결합된 인공지능(AI)으로 분석된 건강 예측, 그리고 개인 유전적 특성 분석에 따른 맞춤형 정밀의료 등에 대하여 신뢰가 있다고 응답한 대상자 수는 52.6%로, 향후 의료 빅데이터 기술의 가능성과 신뢰성에 대해서는 높게 평가되었다. 하지만, 의료 빅데이터에서 나를 당장 식별할 수 있는 개인식별정보가 제거되었다 하더라도, 다른 데이터들과 결합되면 여전히 내 개인정보는 식별될 가능성이 있을 것이라고 응답한 사람은 57.1%으로 많은 대중들이 의료 빅데이터에 있어 개인정보 재식별 가능성에 대한 우려가 높은 편이였다.

의료 빅데이터에서 나의 개인정보가 비식별화 되어 있다면, 별도의 동의가 없어도 연구 목적으로 제공/활용될 수 있다고 응

Table 1. A survey on the public perceptions of health care information

Public perception on health information	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
I believe that my healthcare information is important for treating my diseases.	165 (51.7%)	94 (29.5%)	49 (15.4%)	10 (3.1%)	1 (0.3%)
I consider my healthcare information as very sensitive personal information.	193 (60.5%)	84 (26.3%)	35 (11.0%)	5 (1.6%)	2 (0.6%)
I think that my healthcare information can be helpful to research medical advancements.	86 (27.0%)	133 (41.7%)	72 (22.6%)	25 (7.8%	3 (0.9%)
I am concerned about the possibility of loss, misuse, unauthorized disclosure, and leakage of my healthcare information.	64 (20.1%)	141 (44.2%)	86 (27.0%)	27 (8.5%)	1 (0.3%)
I think my privacy can be violated if my healthcare information is leaked.	140 (43.9%)	110 (34.5%)	48 (15.0%)	19 (6.0%)	2 (0.6%)
I have confidence that medical institutions will do a good job of protecting personal information in the creation/use/management of my healthcare information.	27 (8.5%)	108 (33.9%)	103 (32.3%)	64 (20.1%	17 (5.3%)
When I visited a medical institution, I was faithfully informed about the purpose and processing of my personal information (medical and personal information) and the provision to third parties.	36 (11.3%)	89 (27.9%)	88 (27.6%)	81 (25.4%)	25 (7.8%)
I think that my prior consent is necessary for my healthcare information to be used for treatment/research/education at medical institutions.	215 (67.4%)	77 (24.1%)	21 (6.6%)	5 (1.6%)	1 (0.3%)
I think that even if my healthcare information is used for treatment/research/education in a medical institution, there should be a procedure to withdraw consent after the fact.	163 (51.1%)	113 (35.4%)	32 (10.0%)	11 (3.4%)	0 (0.0%)
I may allow individual contact (phone calls, text messages, etc.) using personal information such as contact information to obtain my consent to use healthcare information for research purposes at medical institutions.	44 (13.8%)	133 (41.7%)	93 (29.2%)	36 (11.3%)	13 (4.1%)
I think that the patient owns healthcare information because it is a personal record of mine regarding the diseases and treatments	170 (53.3%)	93 (29.2%)	42 (13.2%)	13 (4.1%)	1 (0.3%)
I think that medical institutions have ownership of my healthcare information because the information is records of medical judgments of healthcare professionals and tests performed at medical institutions, as well as the information is about my disease.	17 (5.3%)	62 (19.4%)	65 (20.4%)	110 (34.5%)	65 (20.4%)

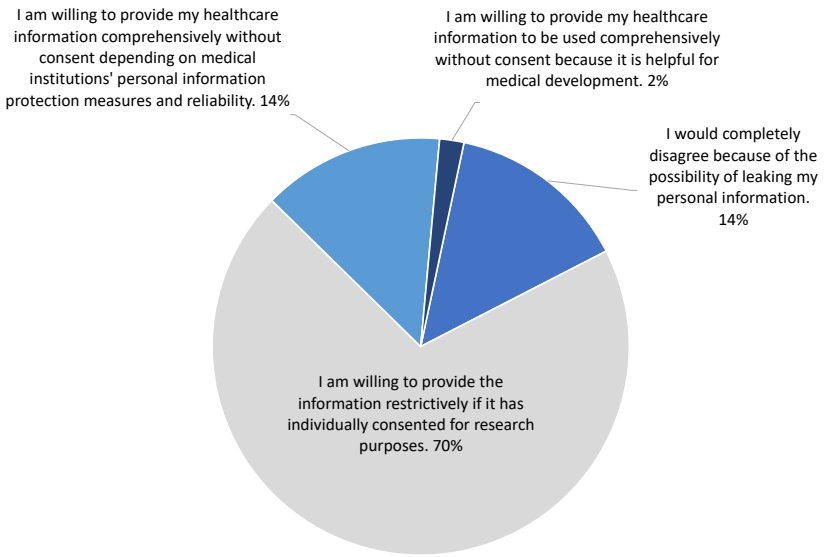

Fig. 1. Acceptance on utilization of healthcare information for research purposes.

Table 2. A survey on the public perceptions of health care big data

Public perception on healthcare-related big data	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
I understand the concept of various databases that can be used as healthcare-related big data.	27 (8.5)	110 (34.5)	122 (38.2)	49 (15.4)	11 (3.4)
I think that healthcare-related big data will become a core value for medical innovation, improvement of public health, and medical development.	59 (18.5)	161 (50.5)	75 (23.5)	19 (6.0)	5 (1.6)
To diagnose and treat my disease and manage my health, big data such as medical information of various institutions, genetic information and health information of public institutions, and personal wearable lifelog must be shared and managed in an integrated manner.	46 (14.4)	140 (43.9)	82 (25.7)	38 (11.9)	13 (4.1)
I can trust my health prediction information and personalized health management services analysed using artificial intelligence combined with various healthcare-related big data.	24 (7.5)	144 (45.1)	104 (32.6)	36 (11.3)	11 (3.4)
I think that if my personal information is de-identified through pseudonymization and anonymization, the probability of my personal information being identified is low.	21 (6.6)	107 (33.5)	116 (36.4)	57 (17.9)	18 (5.6)
I think that even if the personal data identifier is removed, it will still be possible for my personal data to be identified when combined with several different data sets.	51 (16.0)	131 (41.1)	96 (30.1)	35 (11.0)	6 (1.9)
I think that public institutions should create, acquire and manage health data, and the data should be fairly open to and provided to any third party without data imbalance.	41 (12.9)	110 (34.5)	74 (23.2)	68 (21.3)	26 (8.2)
I think that institutions need to obtain individual consent for using healthcare-related big data in research, technology, and industrial development from the person involved with healthcare information (i.e., patients, etc.).	125 (39.2)	124 (38.9)	53 (16.6)	15 (4.7)	2 (0.6)
I think that if personal information from healthcare-related big data is de-identified, it can be provided/used without individual consent.	27 (8.5)	76 (23.8)	85 (26.6)	78 (24.5)	53 (16.6)

Values are presented as number (%).

답한 수는 32.3%로 낮은 반면, 41.1%는 '그렇지 않다'로 응답하였다. 또한, 의료 빅데이터 등을 보관/관리하고 있는 기관 등이 향 후 연구, 기술 및 산업 개발에 활용하기 위해서는 보건의료 정보주체, 즉 의료기관인 경우 환자로부터 개별적 동의가 필요하다고 응답한 수는 78.1%로 높게 나타났다. 이는 앞선 설문과 같이, 기관 내 의료정보뿐 아니라 비식별화된 의료 빅데이터 정보 또한, 다양한 연구와 기술 개발 및 산업적 목적으로 활용되기 위해서는 동의가 필요하다는 대중의 인식을 나타낸다 (Table 2).

또한, 의료 빅데이터 연구사업 활성화 우려가 있다고 응답한 대상 수 65.8%에 달하며, 우려되는 여러 가능성에 대하여, (1) 비식별화 조치가 된 데이터라도 나의 건강정보가 유출될 가능성, (2) 나의 건강정보 등이 유출된다면 현재 나의 질병/유전자/건강 상태에 따라 고용, 보험 등의 사회적 차별 가능성, (3) 소득 수준에 따른 빅데이터 기반 개인 유전자 검사와 고가의 개인 맞춤형 건강관리 서비스 등의 의료서비스 제공의 불평등 심화 가능성을 순으로 나타났다. 반면에 빅데이터를 이용하여 개발한 인공지능의 건강 예측, 개인 맞춤형 정밀의료 등의 불확실성에

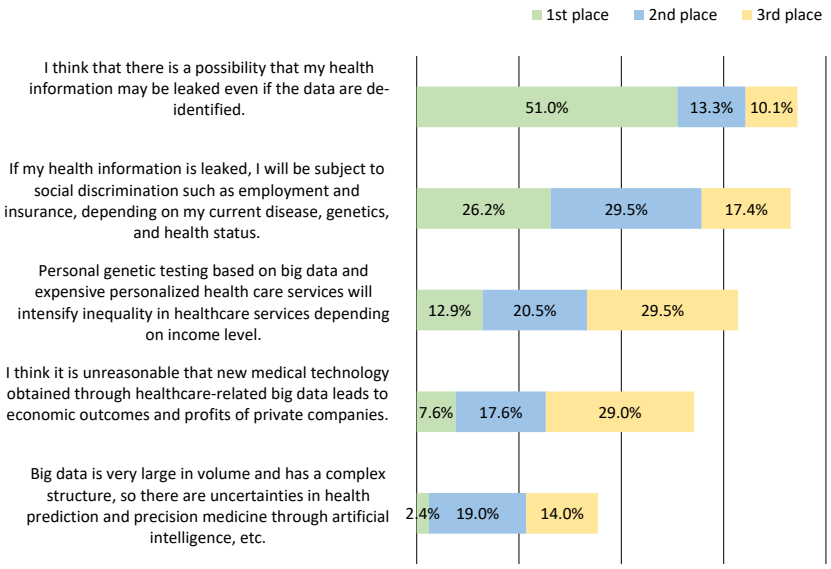

Fig. 2. Concerns on invigorating health care big data research and projects.

대한 우려는 전반적으로 높지 않게 조사되었다(Fig. 2).

의료기관이 보유/관리하고 있는 의료 빅데이터 등을 공중 보건 증진 등 공익적 목적으로 연구, 국가 안보, 의료정책 등에 제공/활용에 대하여 응답자 중 55.2%가 '대체로 동의 한다'라고 응답한 반면, 민간 기업의 상업적 목적으로 제공/활용되는 것에 대해서는 31.3%만이 '대체로 동의한다'라고 응답하였다. 이는, 의료 빅데이터의 공익적 목적의 활용보다는 민간기업의 상업적 목적에 제공 및 활용되는 것에 대한 대중의 수용도는 낮은 것으로 판단한다.

마지막으로 의료 빅데이터를 통한 다양한 연구, 보건의료 정책, 의료 기술, 의학 산업 등으로 활용되기 위해서 뒷받침되어야 할 제도적/사회적 장치 등에 대한 대중의 의견을 알아보았다.

응답자의 68%가 "개인정보 유출되지 않도록 의료기관 및 공공기관의 보안 기술 필요"를 꼽았으며, "개인정보 유출 피해 발생 시 책임의 소재 명확화 및 보상 제도"(58%), "개인정보가 유출되지 않도록 완벽한 비식별화 방안 제시"(55%), "개인정보 유출 및 남용에 대한 정기적 모니터링 제도"(44%)을 제도적/사회적 필요 장치로 꼽았다.

논의

우리나라는 개인정보 보호법에 따라, 가명 처리된 데이터는 정보 주체의 동의 없이 통계 작성, 과학적 연구, 공익적 기록 보존 등의 목적으로 활용하는 것이 허용된다. 또한, 가명정보 처리 가이드라인(개인정보보호위원회, '20.9) 등을 통해 다양한 기술의 가명화, 암호화 절차와 방법 등을 제안하고 있다. 이런 정부의 방향은 개인정보가 어떤 방식이던 암호화 상태에서 개인정보를 보호하면 의료 빅데이터 연구를 할 수 있는 길을 마련하겠다는 것이다. 이는 의료 빅데이터의 무궁한 가치를 고려했을 때, 의료정보의 '활용'과 '개인정보 보호'라는 두 가지 측면에서 올바른 방향이긴 하지만, 이는 기술적인 방법이고 여전히 윤리적 논쟁의 가능성이 있다.

모든 인간 대상 연구는 벨몬트 보고서의 주요 원칙, 즉 인간 존중, 선행(프라이버시 보호), 정의(이득과 부담의 배분)의 윤리적 틀 안에서 시행되어야 한다. 의료 빅데이터 연구에서의 윤리적 논점과 윤리 원칙 준수를 위한 방법 등을 제언해 보고자 한다.

1. 의료데이터의 수집과 활용에 대한 자기결정권

의료데이터 소유 주체에 관한 문제뿐 아니라, 공공기관 등이 보유한 의료데이터는 잠재적으로 공개적인 용도로 사용될 수 있는 일종의 공공재로 간주되어야 하는가?[3] 등의 이슈가 있지만, 의료정보 주체의 권리 및 근본적인 존엄을 위해 보다 적극적인 자율성이 보장될 필요가 있다. 공공의 과학 발전과 진보가 정보 주체의 동의 생략이라는 정보인권 제약을 윤리적으로 정당화하긴 어렵다.[4] 경제협력개발기구(Organization for Economic Cooperation and Development, OECD)에서 2016년 발표한 보건의료 데이터 거버넌스에 대한 권고(Recommendation of OECD Council on Health Data Governance)에서도 '보건의료정보는 본질적으로 민감하고, 데이터의 공유와 사용 확대는 데이터 유출 및 오남용 위험을 야기하여 개인에게 개인적, 사회적, 재정적인 위해를 끼칠 수 있다'며, 개인정보 제공자에게 충분한 설명 후 동의를 구하는 적절한 절차가 필요하다고 권고했다. 또한, 어떻게 동의를 철회할 수 있는가도 명확해하여야 하며, 동의를 구하는 것이 불가능하거나 비현실적인 경우 동의를 대신할 수 있는 적법한 대안 및 예외가 명확해야 하고, 이러한

과정에 부합하는 보호 조치가 뒤따라야 한다고 제언하고 있다.[5]

개인정보 자기결정권의 핵심은 정보주체가 개인정보 처리에 대하여 실질적인 통제권을 갖는 것이며, 개인정보를 통제하는 가장 중요한 수단이 일차적으로 정보 주체의 동의라 할 수 있다.[6]

동의의 윤리적 근거는 벨몬트 보고서에 따른 인간 존중이고, 전통적 윤리적 기본 원칙은 사전 동의이며, 동의 본질은 충분한 설명에 입각한 연구에 대한 잠재적 위험에 대한 명확한 정보 제공이라 할 수 있다. 하지만, 보건의료 빅데이터 연구는 전통적인 연구와 같이 '충분한 정보'에 입각하여, 숙고한 결정을 내리기 위해 제공될 수 있는 정보에는 한계와 불확실성이 있다. 또한, 이미 축적된 의료 정보에 대하여 차후에 동의를 받은 것은 현실적으로 어려우며, 현시점 이후에 전향적으로 축적될 의료정보 또한, 지금 현재로서는 어떻게 활용될지 예측할 수 없어, 동의 범위를 정하기에도 한계점이 있다. 따라서, 빅데이터 시대에는 정보주체의 개인정보 자기결정권을 보장하는데 사전 동의권(opt-in)보다는 사후 거절권(opt-out) 방식이 적합하다는 의견이 제시되고 있다.[7]

2013년 2월 영국 National Healthcare Services는 건강데이터 '보호'와 '활용', 2개 목표를 동시에 달성하기 위하여 '보건복지 데이터의 공표를 위한 익명화 표준(annonymization standard for publishing health and social care data)'를 제정하고 Health and Social Care Information Center (HSCIC) 출범과 함께 케어닷데이터 프로그램(care.data) 실시를 발표하였다. 2014년부터는 1차 의료기관(GP)가 보관하고 있는 환자 진료정보를 사회복지 관련정보 등 여러 정보시스템과 통합하여 하나의 데이터베이스를 만드는 것을 핵심으로 시행되었으나, 영국 시민들은 본인이 알지 못하는 사이 HSCIC가 2차 목적을 위해 건강데이터가 외부에 제공해 왔다는 사실이 밝혀지고, 100만 명 이상이 해당 프로그램에서 탈퇴(opt-out) 하는 등 2016. 7월 프로그램이 잠정 중단된 바 있다.[8]

연구 활용 전/후의 제도적인 동의 절차 제언에 앞서, 영국의 보건의료 빅데이터 사업인 케어닷데이터프로그램(care.data)의 실패 사례에서 보이듯이 의료 빅데이터가 개인의 민감한 건강 데이터에 기반하는 만큼, 적극적으로 대중과의 소통과 담론을 통해 의료 빅데이터의 가치와 미래 활용 잠재력을 알리고, 의료 빅데이터 활용의 공정성과 투명성(개인정보를 누가, 어떤 목적으로, 어떤 정보를, 어떤 방식으로 처리하는지를 명확하게 알리는 것)에 대한 집단적 시각과 신뢰를 갖는 것이 우선되어야 한다. 인간 존중이라는 윤리적 프레임에서 개인의 존엄성과 자율성 보장을 위해 사전 동의도 중요하지만, 본인 보건의료정보에 대한 접근, 관리 및 통제권 등 '정보 소유주체의 권리 강화'와 의료정보의 2차 목적 활용에 대한 투명한 공개를 위한 공공/의료기관의 노력과 대중의 신뢰가 우선되어야 사후 거절권(opt-out)으로 이어질 수 있을 것으로 판단된다.

2. 보건 의료 빅데이터와 프라이버시 보호

의료기관 내 보건의료정보가 의도되지 않은 사고 등으로 분실, 무단 공개되거나 고의적 또는 의도되지 않게 오용 또는 유출되었을 때 프라이버시 위반이 발생하게 된다. 프라이버시 위반이 윤리적 문제가 되는 것은 첫 번째는 결과론적인 우려 즉, 프라이버시를 침해 당한 사람에게 영향을 미치는 부정적이고, 실질적 피해의 결과이다. 보건의료정보가 오용 및 유출되었을 때 대상자가 받을 수 있는 피해는 사회적 피해, 정신적 피해, 금전적 피해(의료사기, 보험료 상승 등), 신체적 피해(상해, 치료적 불이익 등) 등 다양하다. 두 번째는 개인정보인 보건의료정보 등이 선의적인 의도였던 아니던 어떤 목적으로 사용된 경우, 정보 주체 대상자에게 직접적인 피해가 없었거나, 피해가 있었지만 피해를 인지하지 못하는 경우에도 의무론적인 관점에서 윤리적 문제가 있다고 판단될 수 있다. 이는 개인의 데이터 통제권 상실과 개인정보에 대한 제3자의 적절하지 않은 접근 자체가 개인정보 보호라는 프라이버시, 즉 기본 인격권의 침해로 볼 수 있다는 의견이다.[3] 윤리적 연구 수행에 있어 선행(beneficence)이란 대상자를 위험으로부터 보호하는 것과 그들의 복지의 보장인데, 결과론적 우려 및 의무론적 우려 모두 선행의 원칙에서 어긋나는 것이라 볼 수 있다.

앞선 설문조사에서 '의료 빅데이터'를 통한 다양한 연구, 보건의료 정책, 의료 기술, 의학 산업 등으로 활용되기 위해서 뒷받침되어야 할 제도적/사회적 장치 등에 대해서도 많은 대중이 "보건의료 관련 개인정보가 유출되지 않도록 의료기관 및 공공기관의 보안 기술 필요"를 꼽았으며, "보건의료 개인정보 유출 피해 발생 시 책임의 소재 명확화 및 보상 제도", "보건의료 관련 개인정보가 유출되지 않도록 완벽한 비식별화 방안 제시" "보건의료 개인정보 유출 및 남용에 대한 정기적 모니터링 제도"도 필요하다고 응답하였다.

공공/의료기관의 대중의 신뢰도를 구축하는 방안으로 유럽의 개인정보 보호규정(General Data Protection Regulation)와 같이, 연구에서의 구체적인 개인정보 비식별화 방법과 적절성에 대하여 평가하기 위한, 정보보호 부서(가칭) (Data Protection Officer)라는 독립적 거버넌스 구조가 제안된다. 데이터 거버넌스에는 데이터 보안에 대한 기술적 안전장치 마련뿐 아니라, 개인정보가 유출되지 않도록 비식별화 방안 제시, 보건의료 개인정보 유출 피해 발생 시 책임의 소재를 명확히 하고 이에 대한 보상 제도 및 이에 대한 정기적이고 체계적인 독립적인 모니터링 메커니즘이 필요하다. 또한, 개인정보 보호에 대한 방법은 누가 어떤 목적과 방식으로 어떤 데이터에 접근하는가에 따라 달라진다. 따라서, 보건의료정보 활용 범위와 목적에 대한 객관적 평가에 있어, 기관과 이해관계가 없는 의료 정보주체를 대변하는 환자 또는 보건의료 데이터 전문가를 포함하여 구성될 수 있을 것이다.[3]

3. 보건 의료 빅데이터의 이득과 부담

앞선 설문조사에서 의료 빅데이터 연구 · 사업 활성화에 우려로 건강정보가 유출될 가능성, 사회적 차별 가능성 및 빅데이터 기반 개인 유전자 검사와 고가의 개인 맞춤형 건강관리 서비스 등, 소득 수준에 따라 의료 서비스 제공의 불평등 심화 가능성을 꼽았다. 실제로 정보를 제공하는 대중은 이런 위험만 있을 뿐 의료 빅데이터 연구에 따른 효용과 실제적인 혜택이 분명하지 않다고 느낄 것이다.

의료 빅데이터의 또 다른 문제는 잘못된 정보가 의학적 판단에 악영향을 미칠 수 있다는 것이다. 빅데이터는 그 자체로서 정확성과 객관성을 보장해 줄 수 없고, 특히 질병과 관련된 많은 정보들이 의학적 판단 효과가 가려지지 않는 상황에서 활용되는 경우 이런 연구 결과는 큰 위험을 유발할 가능성이 있다.[9] 의료 빅데이터 연구 수행의 전제 조건은 프라이버시 보호와 더불어 연구의 과학적 타당성과 신뢰도이다. 연구는 과학적, 사회적 가치와 공공 또는 연구 대상자의 이익과 관련하여, 위험이 최소화되고 합리적일 때 허용될 수 있다. 따라서, 의료기관 내 연구윤리위원회(Institutional Review Board)는 빅데이터 연구의 목적과 활용 계획, 체계적 데이터 수집 방법, 명확한 알고리즘의 개발과 분석 계획, 비식별화 방안, 빅데이터 보관 위치 및 접근 권한이 명확한지를 검토해야 하며, 연구 결과에 따른 경제적 이해관계가 없는지와, 이로 인하여 연구의 결과의 부적절한 영향을 미칠 가능성이 있다면 이런 영향력을 배제하는 조치가 필요하다.

결론

의료 빅데이터를 활용한 의학 기술의 진보의 시대적 흐름은 거스르기 어렵다고 본다. 의료 빅데이터 활용에 있어 부정적 시각의 과잉적 우려도 문제이지만, 공공의 이득을 위하여 무조건적인 정보제공자의 프라이버시와 권리의 침해의 정당성을 요구하는 것도 적절하지 않을 것이다. 이러한 윤리적 딜레마는 의료 빅데이터의 양면성에 대한 현대의 시대상을 보여주고 있다.

앞서 살펴본 의료 빅데이터 연구의 새로운 윤리적 논점, 그리고 의료 빅데이터 제공과 연구 활용에 대한 대중의 인식도 조사에서 보이듯이 의료정보의 '활용'과 '윤리적 함의'의 균형을 통해 의료데이터의 수집과 활용에 대한 개인정보 자기결정권, 의료 빅데이터의 프라이버시 보호, 의료데이터 연구의 과학적, 사회적 타당성이 보장되어야 일반 대중의 의료 빅데이터 활용에 대한 우려를 잠식시킬 수 있으며, 이런 윤리적 타당성에 기반한 의료 빅데이터만이 미래를 혁신으로 이끌 정보라 할 수 있을 것이다.

이해관계(CONFLICT OF INTEREST)

저자들은 이 논문과 관련하여 이해관계의 충돌이 없음을 명시합니다.

감사의 글 및 알림(ACKNOWLEDGMENTS)

이 논문은 가톨릭대학교 생명대학원 임상연구윤리학 석사학위인정보고서로 수행된 연구를 기반으로 작성되었으며(접수일자: 2020.07.06 /심사일자: 2020.11.20.), 가톨릭대학교 생명대학원 소식지 「생명마루」 제 20호 (2022년 봄)에 요약문 게시 예정입니다.

ORCID

Byung In Choe 0000-0002-7401-2798
Su Jin Cho 0000-0002-9721-8131

REFERENCES

1. Kim KH, Lee SH, Yang SY, Oh YS, Song TM. A study on the utilization of big data analysis in ICT policy. Gibonyeongu 2017:4:17-8.
2. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. Nat Rev Cardiol 2016;13:350-9.
3. Price W, Cohen I. Privacy in the age of medical big data. Nat Med 2019;25:37-43.
4. Lee SY. Healthcare big data and personal data protection, autonomy. Bio Ethics Policy 2019;3:47-58.
5. Byeon HJ. Several issues in the direction of health care big data. Health Soc 2017;8:40-61.
6. Park MJ. A study on legislative and policy measures for big health care data. J Korean Assoc Med Law 2018;26:163-92.
7. Kim GR, Lee DH. Review on healthcare big data analysis - focusing on privacy protection. Hannam J Law Technol 2018;24:57-90.
8. Park DW, Lee SM, Jeong SH. Investigation and analysis of legal trends in the latest health care big data. Cheongju (Korea): Korea Health Industry Development Institute; 2017. Report No.: 2017-45.
9. Kim KW. Usefulness and riskiness of big data in public sector. Korean J Policy Anal Eval 2013;23:1-27. Report No.: NRF-2011-330-B00034.