보건의료 빅데이터의 보호 및 활용을 위한 법적 기반 강화 연구

2019. 5 초안

보 건 복 지 부

제 출 문

보건복지부장관 귀하

본 보고서를 “보건의료 빅데이터의 보호 및 활용을 위한 법적 기반 강화 연구”의 최종보고서로 제출합니다.

2019.

주관연구기관명 : 서울대학교 산학협력단
연 구 책 임 자 : 고 학 수
연구원 : 이 동 진
박 미 정
김 은 수
신 수 용
김 현 창
박 유 랑
차 효 성
연구보조원 : 구 본 효
김 찬 희
박 도 현
박 지 훈
손 정 구
장 승 원
정 종 구

차례 | Contents

<그림 차례>

<표 차례>

» 요약문

1. 연구의 필요성 및 목적

가. 보건의료 정보의 특성

보건의료 빅데이터는 잠재적 가치와 활용가능성이 매우 높이 평가되고 있다. 보건의료 빅데이터 분석을 활용할 수 있는 분야도 매우 광범위하다. 단일한 데이터세트에 대한 분석을 넘어 다양한 데이터세트를 결합해 분석하는 정밀의료 패러다임은 높은 잠재력을 가지고 있기도 하다.

그러나 다른 한편 보건의료 빅데이터 분석에 다량의 동종 및 이종 데이터의 결합이 요구된다는 점이 양날의 검으로 작용하기도 한다. 분석력을 극대화하여 활용가치를 높일 수 있지만 정보주체가 인지하지 못하는 활용의 빈도도 늘어나게 되어, 정보주체의 프라이버시 보호 문제가 대두된다. 기존의 개인정보 법제는 빅데이터 시대의 특성을 아직 완전하게 반영하지 못하고 있어, 어떻게 유용한 활용도를 높이는 동시에 적절하게 개인정보보호를 할 것인지에 관한 어려운 과제가 대두된다.

나. 연구 목적

이 보고서의 연구목적은 다음과 같다.

첫째, 보건의료정보 관련 국내외 입법례를 비교법적으로 고찰하여 보건의료정보의 정의, 범위, 이용목적, 절차, 사용주체 등을 검토한다.

둘째, 보건의료정보의 보호와 활용에 있어 공공성 확보를 위한 거버넌스와 효과적 운영방안 등을 파악하고, 중장기 정책수립 및 전문 관리기관 설립 필요성을 탐구한다.

셋째, 보건의료정보 보호를 위한 다양한 기술적 방안을 조사·분석하고, 정보 오남용 및 유출 등 부정사용에 대한 처벌 및 통제방안을 모색한다.

넷째, 보건의료정보 보호 및 활용을 위한 법제 개선방안 연구로서, 주로 논의되는 쟁점사항을 반영한 법률 제·개정 방안 및 법률체계를 제안한다.

다. 연구 방법

본 연구는 법제도에 대한 분석을 위주로 하되, 학제적 연구방식을 통하여 연구자의 전공 및 전문분야의 특성에 따른 연구방법을 활용하였다. 이에 문헌조사의 방식으로 국외 주요국가의 법률 및 국내 관련 법령을 분석하고, 공공기관과 대형병원의 사례를 통해 보건의료 빅데이터의 활용사례를 검토하였으며, 민간 병원과 공공기관의 실무자 등을 대상으로 자기평가식 초점집단면접조사를 시행하였다.

2. 연구의 내용

가. 보건의료 빅데이터에 관한 법적 쟁점 분석

1) 주요 개념 분석

주요 개념으로는 보건의료정보 내지 의료정보 이외에 그 주변영역에 있는 유전정보와 생체정보, 그리고 이들이 특히 개인정보 보호법제 체계의 맥락에서 문제되는 민감정보를 다룬다.

의료정보에 관하여는 「보건의료기본법」 제3조 제6호에서 정의하고 있는 '보건의료정보'가 논의의 출발점이 된다. 이는 「개인정보 보호법」의 적용 대상인 '개인정보'를 포함하나 그에 제한되지 아니한다. 매우 넓은 개념으로 경계가 분명하다고 할 수는 없다. 미국은 HIPAA가 의료정보 보호를 별도로 규정하고 있고, 일본의 차세대의료기반법은 개인식별정보가 아닌 의료정보에 대하여 별도의 규율을 가한다. 유럽은 GDPR이 '건강에 관한 정보'를 폭넓게 정의하고 있다.

유전정보는 인체 구성물 또는 이로부터 분리된 물질인 인체유래물을 분석하여 얻은 개인의 유전적 특징에 관한 정보를 말한다(생명윤리 및 안전에 관한 법률 제2조 제11호, 제14호). 유전정보는 불변성, 가족 공유성, 고유식별성, 표현형 관련성, 미지성 등의 특성에서 일반적 개인정보와 구별된다. 우리나라와 EU는 유전정보가 민감정보에 포함되는 것으로 보고, 일본은 개인정보보호법 시행령에 DNA를 구성하는 염기서열을 포함하였으며, 미국의 GINA는 유전정보의 범위를 규정하고 그 보호를 명시하고 있다. 유전정보는 차별 또는 유전자 검사의 문제, 민감정보에 포함될 경우 정보주체의 동의 필요 여부 등의 쟁점을 제기한다.

생체정보, 즉 바이오정보는 '지문, 홍채, 음성, 필적과 같이 개인을 식별할 수 있는

신체적 또는 행동적 특징에 관한 정보'를 의미한다(정보통신망 이용촉진 및 정보보호 등에 관한 법률 시행령 제9조의2 제1항 제1호). 생체정보는 사람의 신체와 행동 등에 관계한다는 점에서 보건의료정보와 관련되지만 기능적으로는 개인정보 보호법의 고유식별정보에 준하여 취급할 여지가 있다. 빅데이터 시대를 맞이해 IoT 기술의 발전에 따라 수집되는 수많은 (행태)정보가 생체정보에 포함되는가 하는 문제가 대두되는 중이다.

이들 정보는 상당수 민감정보에 해당한다. 국내 법제 형성에 큰 영향을 미친 유럽연합의 Directive 95/46/EC가 건강정보를 포함하여 민감정보를 정의하고 있고, GDPR은 유전정보와 생체정보도 민감정보에 해당한다고 본다. 미국은 일반적인 민감정보에 관한 규정이 없어 주로 학술적으로 논의되고 있을 뿐이지만 보훈부 법령에서는 의료기록을 민감정보로 규정하고 있다. 일본은 민감정보와 유사한 필요배려 개인정보에 의료정보가 포함된다고 해석하나, "부당한 차별, 편견 및 그 밖의 불이익"이 생길 우려가 있는 경우로 한정된다는 특징이 있다. 현행「개인정보 보호법」제23조 제1항은 민감정보를 열거하고 있고 그중 "건강에 관한 정보"가 포함되어 의료정보와 건강에 관한 정보의 관계가 논란이 될 수 있다. 헌법재판소는 민감정보의 분별기준에 대한 판시를 한 바 있으나 의료정보를 명시적으로 다루지는 않았다.

2) 법적 쟁점

보건의료정보와 관련하여 제기되는 주요 법적 쟁점은 다음과 같다.

첫째, 과학적 연구, 공익 목적의 연구, 통계 목적의 처리에 관하여 개인정보 보호법 적용의 예외가 일반적으로 인정될 수 있는지 여부

둘째, 보건의료산업에 적용되는 다양한 규제의 일원화 및 개선을 통해 산업발전을 추구하는 방안 및 그에 관한 개인정보의 보호와의 규범적 조화 방안

셋째, 정밀의료 맥락에서의 개인정보 보호법제 적용에 있어 비식별화 조치, 식별가능성, 목적 외 이용 등과 관련된 각종 쟁점

넷째, 임상시험과 임상연구에서 데이터 활용에 관한 동의의 필요성 및 비식별화 정보의 활용 가능성

3) 개인정보처리 관련 실무자 초점집단 면접조사, 이해관계자 대상 초점집단 면접조사

초점집단면접조사에서 제시된 대안과 권고를 정리하면 다음과 같다.

첫째, 개인정보처리자들이 데이터를 생성할 때부터 제공 및 이차 활용을 위한 보안 인식이 있어야 한다.

둘째, 각 이해관계자들이 보유하는 데이터가 각 기관의 고유한 보유 및 관리정책이 따르고 있는 현실에서 수집 목적 외 이차활용을 위해 제공을 위한 표준화된 절차가 필요하다.

셋째, 데이터 보안의 범위를 데이터 처리를 포함하여 내부보안/외부보안에 대하여 단계별로 정하고 대비하여야 한다.

넷째, 제3자 제공에 대한 개별 법률에 대한 정합성이 필요하다.

다섯째, 보호의 필요성이 더 강한 유전정보와 인체유래물에 대한 IRB 제도 개선이 필요하다.

여섯째, 데이터세트의 모듈화 등 활용분야와 영역에 맞는 데이터 공유정책이 수립되어야 한다.

나. 의료보건의료 빅데이터 보호 및 활용에 있어 국내외 입법례와 거버넌스 분석

1) 각국의 보건의료 정보 관련 법·제도의 현황

가) 유럽연합(EU)

유럽연합은 2018년부터 GDPR을 통해 개인정보 보호를 통일적으로 규율하고 있다. GDPR은 건강 등 특수범주 데이터의 목적 외 처리의 제한에 관한 규정을 두고 있다. 해당 규정은 데이터 처리를 위하여는 정보주체의 동의가 필요하다고 하면서도, 과학적 연구 등에 대하여 일정한 예외를 인정한다.

EU의 거버넌스 체계로는 의료정보의 처리를 위한 개인정보보호담당관 제도와 개인정보처리의 목적 및 수단을 결정하는 콘트롤러 및 프로세서의 역할 및 지위에 관한 규정을 들 수 있다. 그 이외에 유럽평의회의 보건의료 데이터에 대한 권고는 합법적인 데이터 처리의 기준을 마련하고, 유전정보도 처리를 통해 사용할 수 있도록 한다. 다만 의료서비스의 제공 및 관리 목적의 보건 관련 데이터가 공유되는 경우 개인정보주체는 사전에 통지를 받아야 한다. 또한 공공정보 재사용에 관한 지침이 제정되어 국내법에 따라 접근할 수 있는 모든 콘텐츠는 원칙적으로 상업적 및 비상업적 목적을 위하여 재사용 가능하도록 허용하고 있다. 특히 임상시험과 관련하여 데이터의 이차활용에 초점을 두고 있다.

나) 영국

영국은 Data Protection Act of 2018을 통하여 개인정보를 일반적으로 규율하고 있다. 그 기본적인 내용은 GDPR과 대동소이하다.

그 이외에 주목되는 것으로 Health and Social Care Act 2012가 있다. 같은 법은 NHS의 보건 서비스와 사회보장 서비스를 통합하여 서비스를 효율적으로 제공하고 불필요한 비용을 절감하는 것을 목적으로 한다. 그 일환으로 위 과정에서 따라 생성된 보건의료정보를 수집할 공공기관의 설립 근거도 두었는데, 그 결과 설립된 것이 NHS Digital이다. 다른 한편, 보건의료정보 이용 맥락에서 공공기관이 보유한 데이터에 대한 일반인의 접근을 인정하는 Freedom of Information Act 2000이 존재한다.

한편, 영국은 National Data Guardian을 설립하여 시민들의 기밀정보가 안전하게 보호되고 적절하게 활용될 수 있도록 하고 있다. Fiona Caldicott의 주도하에 발간된 일련의 보고서는 data guardian의 필요성과 정보관리에 관한 Caldicott 원칙, 암묵적 동의의 한계, 연구목적 데이터의 구체적 관리요건, 데이터 보안과 동의의 기준 등에 관해 정리하였다.

끝으로 바이오뱅크를 구축하여 45~69세 영국 국민 50만 명의 인체유래물과 함께 그들의 신체 계측정보와 환경인자와 과거 및 현재 건강상태, 질환위험, 신체적 특징 등에 관한 데이터를 수집하고 있다. 이 과정에서 동의는 인체유래물 제공시의 포괄적 동의 및 철회권의 보장 방식으로 이루어지고 있다. 그 외에도 UK 100,000 Genome Projects를 통해 100만 명을 목표로 유전자 시퀀싱 데이터 구축을 진행해왔다.

다) 프랑스

프랑스의 정보처리·문서 및 자유에 관한 법률(Loi a l'informatique, aux fichiers et aux libertés)은 정보기술의 형태와 개인정보 보호를 정하고 있다. 같은 법은 개인정보 보호를 위한 독립기관을 설치하여 정보시스템을 감독하게 하는 한편, 익명화된 데이터의 이차 활용은 정보 보호법의 범위에서 제외된다고 보아 이를 허용하고 있다.

디지털 공화국법률(Loi pour une République numérique)은 9개의 주요 주제(망중립성, 데이터 이동성, 접속 유지 권리, 사적 편지 기밀 유지 등)를 다루는 30개의 조항으로 구성되어 있다. 공공데이터와 민간데이터 외에 '공익성'이라는 근거를 바탕으로 분류되는 공익데이터를 규정하여, 민간에 속하는 데이터임에도 불구하고 공공

정책을 향상시키기 위하여 공개하여야 하는 데이터의 범주를 마련한다.

건강시스템 현대화에 관한 법률(Loi de modernisation de notre système de santé)은 연구 목적 건강정보 이용의 활성화를 위해 사전절차의 완화와 이용자 책임의 강화를 위한 기준을 마련하였다. 이 법에 근거하여 국가건강보험기금은 국가건강정보시스템(système national des données de santé)을 구축하였다. 또한 건강정보의 제공에 관한 원칙이 공중보건법전(Code de la santé publique)에 포함되었다.

한편 프랑스는 국가차원의 윤리위원회를 최초로 설립하고 위원회의 윤리적 견해를 채택하여 생명윤리에 관한 법률(Loi de la bioéthique)을 제정하였는데, 같은 법률은 유전적 특징에 대한 검사(examen des caractéristiques génétiques)에 필요한 동의 요건을 규정하여 수술시 채취된 인체유래물의 사후 동의 및 이의 표명이 없는 경우 이차 활용을 인정한다.

거버넌스 차원에서 중요한 것은 국가정보처리자유위원회(Commission Nationale de I'Informatique et des Libertés; CNIL)이다. CNIL은 개인정보 보호를 위한 중요 규제권한을 행사한다. CNIL은 독립행정관청(autorité administrative indépendante)으로서 독자적 행정조치를 취할 권한이 있다. CNIL은 개인식별자로서의 사회보장번호에 관한 실무그룹을 결성하고 민감정보 처리에 관하여 각 법률에서 수권 받은 바와 같이 심의 및 승인 등을 진행하고 있다.

라) 미국

미국은 1996년에 HIPAA(Health Insurance Portability and Accountability Act)를 마련함으로써 건강정보의 이동(移動)을 현대화하고 개인 식별가능정보를 보호하고 있다. 특히 HIPAA의 Privacy Rule은 개인의 건강기록과 건강정보를 보호하는 국가적 기준을 정립하여, 식별가능한 건강정보에만 Privacy Rule을 적용하고, 비식별조치에 관하여 상세한 규정을 두고 있다. HIPAA Security Rule은 건강정보의 보호 맥락에서 국가적 기준으로서의 일반적인 보안 요건을 설정하고 있다.

한편, HITECH Act(Health Information Technology for Economic and Clinical Health Act)는 경제 활성화를 위하여 제정된 연방법으로서 전장기록 및 이에 관한 기술의 이행을 권장하기 위해 제정되었다. 법률 위반시의 법적 제재의 수준을 높이고 더 엄격한 법집행을 가능하게 하여 프라이버시와 보안의 범위를 확장한다. 그 Subtitle D에서는 건강정보가 전자기록 형태로 이동할 때 발생하는 보안 문제를 다루고 있고, 나아가 전자기록으로의 전환을 위한 금전적 인센티브를 제공하여 전자기록 시스템의 채택을 유도하고 있다.

2016년에 입법한 21세기 치료법(21st Cure Act)은 의약품 개발을 촉진하여 환자 치료의 효율성을 높이고자 신속한 허가를 도모하기 위하여 제정되었다. 이에 데이터의 공유를 활성화와 함께 데이터의 보호에 대한 내용을 포함하고 있으며, 식별가능한 민감정보를 규정하여 그에 관한 프라이버시의 보호와 관련된 미 보건부의 가이드라인을 발표하도록 하고 있다.

커먼룰(Common Rule)은 과거 벨몬트 보고서에서 발표한 원칙을 대폭 반영하여 제정된 연방 규정 및 지침으로 인간을 피험자로 하는 연구에 대한 여러 가지 윤리적 원칙을 제시한다. 특히 최근 개정에서 임상시험 및 인간대상자, 인체유래물의 정의를 보다 세밀화하는 한편 포괄동의 방식의 도입, Single IRB의 의무화 등을 새로이 규정하고 있다.

거버넌스 측면에서는 퇴역군인 등을 위한 Million Veteran Program을 운영하면서 유전체가 건강에 미치는 영향력을 연구하기 위한 데이터를 수집하고 있고, 국립보건연구소의 주도로 Electronic Medical Records and Genomics Network를 창설하여 인체유래물을 수집하고 있다. 나아가 HealthData.gov를 통하여 다양한 연방기관의 데이터베이스에 접근할 수 있도록 하면서, 이때 데이터를 받기 전 데이터 이용 제한에 관한 동의를 표시하도록 하고 있다.

마) 일본

일본은 최근 개인정보보호법 개정을 통하여 개인정보 개념을 명확화하고, 익명가공정보 개념을 도입하는 한편, 개인정보의 국외이전에 대한 규정을 정비하고 개인정보보호위원회를 신설하였다.

한편 차세대의료기반법은 의료정보의 익명가공을 통해 연구개발의 활용을 도모하는 법률로, 활용체계를 크게 ① 제공·위탁과 ② 이용 및 ③ 파기의 단계로 나누어 규정하고 있다. 차세대의료기반법은 개인정보보호법의 익명가공정보 개념을 그대로 도입하고 있는데, 일종의 opt-out 방식으로 익명가공정보를 활용하는 것을 허용한다는 점에 특징이 있다.

그 이외에 건강·의료전략추진법은 유전검사와 유전진단에 대한 가이드라인을 제시하여 개인유전정보 취급원칙을 다룬다.

거버넌스의 측면에서는 2013년 설립된 건강·의료전략추진본부가 일본 건강의료전략을 수립하는 사령탑으로 기능하고 있다. 또한 차세대 의료 ICT 기반 협의회를 창설하여, 디지털 데이터의 취합과 활용 및 표준화, 의료정보 제도 정비, 차세대 의료 ICT 도입 문제를 전담한다.

바) 한국

우리나라에서는 2011년 제정된 「개인정보 보호법」이 개인정보 보호의 일반법으로 기능하고 있다. 개인정보 보호법상의 개인정보는 원칙적으로 개인을 식별할 수 있는 정보인데, 그 정보의 연계나 결합으로 인해서 개인을 식별할 식별 가능성이 생기는 경우에도 개인정보로 인정된다. 이에 따라 개인식별 가능성을 제거하는 비식별처리(de-identification)를 통해 정보를 활용하는 것에 관한 논의가 이루어져 왔다. 하지만 가이드라인에 따른 비식별처리를 한다고 반드시 재식별이 불가능한 익명정보가 되는 것은 아니라는 주장이 제기되었다.

개인정보는 원칙적으로 수집한 당시의 목적을 위하여서 처리하여야 한다. 예외적으로 그러한 목적을 벗어나 처리할 수도 있는데, 이는 법에서 정한 사유가 있는 때로 한정되어 있다. 이때 어떠한 처리가 예외적 요건을 충족하는지 여부에 대하여서는 구체적 · 개별적으로 판단하여야 한다.

민감정보의 경우 원칙적으로 개인정보 처리 목적을 충분히 설명하고 별도의 동의를 받아 처리하여야 한다. 이와 관련하여 개인정보의 이차 활용 및 미래의 미지의 연구목적을 위해 정보주체의 동의를 얻는 어려움이 문제가 되어왔다.

한편, 2004년 제정된 생명윤리 및 안전에 관한 법률은 이후 전면개정을 통해 유전정보까지 그 규율대상에 포함시키고 있다. 개정법은 유전정보 및 유전자은행 개념을 인체유래물 및 인체유래물은행으로 확대하여 연구용 자원에 대한 종합적인 관리체계를 마련했다. 인체유래물 연구는 연구계획서에 대한 기관위원회의 심의, 인체유래물 제공자의 서명동의를 받는 경우에 한하여 가능하다.

의료법에 따르면 의료인은 의료법에 따라 정보주체의 동의 없이 개인정보의 수집이 가능하나 이러한 의료정보는 공개를 목적으로 생성되는 것이 아니어서 일반적인 개인정보와는 다른 특별한 보호를 받는다. 수집한 의료정보의 이차 활용과 관련된 규정은 제21조 기록열람 규정 등에 의해, 정보주체가 아닌 자에 대한 의료정보 제공은 원칙적으로 불가하고, 예외사유를 열거하고 있으나 연구목적은 포함되어 있지 아니하다.

한편, 공공데이터의 제공 및 이용활성화에 관한 법률에 따라 국민건강보험공단 및 건강보험심사평가원의 데이터가 오픈데이터 정책의 일환으로 개방이 되고 있다. 또한 보건의료기술 진흥법에 따라 설립된 한국보건의료연구원은 민감정보 및 고유식별정보 등의 개인정보가 포함된 자료의 제출을 요청할 수 있고, 요청을 받은 국기기관 및 공공기관은 개인식별이 가능한 부분을 삭제하여 제출하여야 한다.

의무기록은 병원 내에 보관하여야 하고, 전자의무기록 수단으로 보관하는 경우에도

의료기관 내부에서만 가능하며, 외부 인터넷과의 연결이 금지되어왔다. 그러나 2016년 2월 의료법 시행규칙 개정으로 의료기관이 클라우드를 활용하여 전자의무기록을 외부 장소에서 보관·관리할 수 있게 되었다.

거버넌스의 차원에서는 보건복지부가 2010년 500병상 이상의 의료기관에 대하여 '의료기관 개인정보보호 가이드라인'을 발표했고, 인간대상연구의 경우 임상시험심사위원회(Institutional Review Board, IRB) 제도 및 인간연구대상자 보호 프로그램(Human Research Protection Program, HRPP) 관련 가이드라인이 있다. 유전자 검사기관과 유전자 치료기관은 보건복지부 장관에게 신고하여야 하는데, 비의료기관 유전자 검사에 관한 여러 문제가 지적되어왔다. 이에 보건복지부 산하 'DTC 유전자 검사 제도개선 민관협의체'가 개선안을 마련하였다.

2) 개인정보보호 영향평가 제도 및 법적 제재 규정

유럽연합 GDPR은 개인정보보호 영향평가 제도를 두고, 개인정보보호 담당관을 통해 법적 의무의 준수를 담보하고 있다. 한편 법적 제재에 관하여 GDPR은 위반행위에 대한 과징금 수준을 대폭 상향조정 하는 한편, 과징금의 대상이 되지 않는 위반사항은 회원국이 처벌에 관한 별도의 입법을 하도록 규정하고 있다.

영국은 DPA 2018 개정으로 개인정보보호 영향평가 제도가 필수사항으로 새롭게 도입되었으며, 데이터의 처리가 개인의 자유와 권리에 높은 위험성을 일으킬 가능성이 있는 경우에 데이터 관리자(controller)는 데이터 처리 이전에 개인정보보호 영향평가를 실시하여야 한다. 법규위반에 대한 제재 조치로는 강력한 과징금 제도를 적용하고 있다.

프랑스의 CNIL은 GDPR이 제시하는 개인정보 영향평가를 위하여 실제적인 데이터 보호 영향평가를 실현하고자 개인정보보호 영향평가 지침서를 제공하고 있다. 또한 CNIL은 대상기관의 개인정보 침해 여부에 대해 조사·심의할 수 있고, 법률위반 행위에 대해 강력한 제재를 할 수 있다.

미국의 HIPAA와 HITECH Act는 개인정보보호 영향평가에 대한 내용을 포함하지 않고 있다. 하지만 E-Government Act of 2002라는 연방법에서 정부기관에 한하여 개인정보보호 영향평가를 명시적으로 규율하고 있다. 한편, 제제와 관하여 HIPAA는 주로 프라이버시 위반에 대해 행정규제 위주의 집행 체제를 구축하고 있다.

일본에는 일반적인 개인정보 영향평가 제도가 없다. 그러나 특정 개인을 식별하기 위한 번호의 이용 등에 관한 법률에 따라 특정 개인정보에 한하여 보호평가 제도가 있다. 법적 제재와 관련하여서는 개정 개인정보보호법이 전체적으로 위반행위에 대한

처벌조항을 강화하였다고 평가받고 있다. 차세대의료법은 형사처벌을 포함하는 벌칙 규정을 두고 있다.

우리 개인정보 보호법 제33조는 일정 기준에 해당하는 개인정보파일의 운용으로 인하여 정보주체의 개인정보 침해가 우려되는 경우 위험요인의 분석과 개선 사항 도출을 위한 평가를 하도록 의무화하고 있다. 그러나 그러한 의무적 영향평가 시행대상은 공공기관으로 제한된다. 한편, 의료법 제21조는 기록열람사유를 한정적으로 열거하고 있지만, 연구, 교육 등의 목적을 위하여 기관내부의 기관생명윤리위원회(IRB) 승인을 거쳐서 연구할 수 있다. 생명윤리 및 안전에 관한 법률은 유전자 연구의 경우 유전자은행은 수집한 모든 유전정보 등을 익명화하여 보관·관리하여야 하고 개인정보 보호를 위한 책임자를 두어야 한다고 규정한다.

3) 보건의료 빅데이터 활용 및 보호 기술 현황

유럽연합은 EU 역사상 최대 규모의 연구재정지원 프로그램인 Horizon 2020 프로젝트의 세부주제로 My Health My Data(MHMD) 시스템을 도입하여 건강정보 공유와 관련하여 정보 보호 및 보안을 확보할 수 있는 블록체인 기반 플랫폼을 마련하였다. 한편, 디지털 싱글마켓(Digital Single Market) 전략에는 건강관리와 보살핌을 위해 디지털 기술이 필수적임을 강조하면서 다양한 프로젝트를 후원하고 있다. 또한 유럽연합은 데이터 가이드라인의 차원에서 AEGLE project를 통한 빅데이터 연구의 현황에 관한 보고서를 수집·공개하고 있고, 비식별처리 가이드라인을 배포하여 특히 익명화와 가명화에 대한 여러 규정을 정리하고 있다.

영국 NHS는 2018년 12월 ‘Making NHS data work for everyone’이라는 보고서를 통해 NHS data를 기업체들과 어떻게 활용해야 하는지에 대해서 권고사항을 제시하였다. 특히 UK Biobank가 성공적인 협업모델로 운영되고 있으며, NHS Digital은 ‘Data sets’라는 웹사이트를 통하여 데이터세트를 공개하고 있다. 그밖에 데이터처리 가이드라인으로 2012년 디지털 서비스 설계 원칙(Government design principles)과 ICO 익명화가이드라인이 있다.

프랑스는 국가건강정보시스템(système national des données de santé)을 구축하여 정보처리 및 자유법 제8조 제2항에 규정된 예외사항에 속하는 건강정보는 신고만으로 연구 목적으로 이용할 수 있게 하였다. 또 Epidemilogie-France 포탈은 프랑스가 보유하고 있는 260개의 보건의료 공공 DB와 500여개의 medical and economic DB, 코호트, 레지스트리와 현재 진행 중인 연구 정보들을 제공한다.

프랑스에서는 2018년 1월 23일 GDPR이 요구한 다양한 사항을 체계적으로 제시

한 새로운 개인정보보호 가이드라인(Un nouveau guide de la securite des donnees personnelles, 2018)이 발표되었다. CNIL은 2015년 WP 29에 익명화 기술에 대한 의견을 제출한 바도 있다(Clinical Trial Data Sharing: Methods and Experiences with De-Identification). 이 의견은 익명정보(Anonymized Data)를 개인정보와 구분하고, 콘트롤러가 익명화한 익명정보는 개인정보 보호법의 적용대상이 아님을 분명히 하였다. 나아가 2018년 CNIL은 기업과 그 파트너들(partenaires commerciaux)이 제3자와 데이터를 공유할 때 적법한 공유를 위해 준수해야 할 조건을 발표했다.

미국은 오픈데이터 정책을 발표하여 데이터의 활용을 활성화하고 있다. 그 중에서 Health와 관련된 데이터는 HealthData.gov를 통해서 별도로 제공되고 있다. HealthData.gov는 모두 2,981개의 dataset이 제공되고 있으며, 보건의료 데이터를 다루는 대부분의 기관과 주정부가 포함되어 있다. 그 외에 현재 공유에 따른 문제점을 지적하는 미국 보건부의 보고서 및 의료정보학회의 권고문이 발표되어 보다 용이한 정보 공유 및 활용과 데이터 결정권에 대한 보다 상세한 원칙을 마련할 것을 촉구하고 있다.

미국의 데이터 처리 가이드라인에 관하여 HIPAA Privacy Rule은 두 가지 종류의 비식별조치의 방법을 제시하고 있고, 비식별 조치에 대해 더 상세하게 설명해주는 가이드라인이 발표되었다. 이 가이드라인은 특히 전문가 판단과 세이프 하버 방법의 세부 요건을 설명하고 있다. 한편 ISO 25237:2017은 개인 건강정보의 보호 측면에서 가명화에 기초한 프라이버시 보호의 원칙과 필수 요건에 대한 내용을 포함하고 있다.

일본은 2018년 10월부터 빅데이터를 거래할 수 있는 플랫폼을 허용하였으나 보건의료 데이터의 실제 거래가 활발하게 이루어지는지 여부는 불분명하다. 한편 2012년부터 '의료혁신 5개년 전략'에서 맞춤의료를 주요 과제로 선정해 추진하고 있다. 이에 따라 2015년부터 '질병 극복을 위한 게놈 의료실현화 프로젝트'를 추진하고 있다.

일본에는 개인정보보호위원회에서 마련한 익명가공정보에 대한 가이드라인이 있다. 이 가이드라인은 개인정보를 취급하는 주체가 익명가공정보를 취급할 때 행하는 익명가공 조치가 적절하고 유효하게 시행되는 것을 목적으로 도입되었다. 이에 따라 관련 개념을 정의하고, 익명가공정보 취급사업자의 의무사항을 세부적으로 제시하고 있다. 그 이외에 실제 익명가공정보를 생성하려는 기업을 위한 Q&A 형태의 안내서를 발표하였다.

우리나라는 보건의료 빅데이터의 활용에 있어 먼저 국민건강보험공단이 관리하는 건강보험 빅데이터가 있다. 이는 공단 자료 중 정책 및 학술연구에 필요한 일부 자료를 추출하여 연구용 국민건강정보 데이터베이스로 구축한 것이다. 이들 데이터는 비

식별조치를 거쳐 일정한 조건하에 연구 목적 등으로 개방되고 있다. 또한 국민건강보험공단은 한국보건사회연구원과 공동으로 한국의료패널을 제공하고 있고, 이에 대한 패널자료가 구축되어 있다.

건강보원심사평가원의 경우 보건의료 빅데이터 개방시스템을 통해 보유하고 있는 데이터를 개방하고 있다. HIRA 빅데이터는 의료기관 등으로부터 수집된 원천 데이터를 비식별화 조치 후 별도의 망으로 분리된 개방 DB을 통해 학계·의약계·산업계 및 일반국민을 대상으로 제공된다. 빅데이터 활용 주요 서비스로서 공공데이터 목록을 제공하고 있는데, 공공데이터포털, 보건의료빅데이터개방포털을 통해 공표가능한 통계 데이터를 유형별로 제공하고 있다.

국립암센터는 암관리법에 의해 국가암등록통계로 매년 의료기관의 진료기록을 바탕으로 암환자 자료를 수집·분석, 전전년도의 암발생률, 생존율, 유병률 등을 산출하고 있으며, 국가 암관리정책 수립 및 국제비교의 근거자료로 활용된다. 또한 암빅데이터 플랫폼을 통해 국가암데이터센터 구축과 암 정밀의료 빅데이터 플랫폼 구축의 연구 및 사업을 수행하고 있다. 질병관리본부는 '인체유래물은행' 및 '국가병원체자원은행' 규정에 따른 한국인체자원은행사업을 하고 있으며 대규모 코호트 구축 운영을 통한 한국인 유전체 역학조사사업을 진행하였다.

다른 한편, 국내의 일부 대형병원들은 자체적으로 빅데이터 센터를 두고 있고, 국내 공공기관이 보유한 빅데이터를 연계해 개방하는 플랫폼을 구축하며, 이와 병행하여 연구중심병원 등 의료기관 중심의 특화 질환별 연구 플랫폼도 구축하고 있다. 또한 산업통상자원부와 보건복지부는 대학병원 및 관련 업체 40여개로 구성된 '바이오 빅데이터 플랫폼 사업 추진단'을 2018년 5월 발족시켰다.

비식별화 관련 데이터 처리 가이드라인으로 대표적인 것은 2016년 관계부처 합동으로 만든 '개인정보 비식별 조치 가이드라인'이다.

다. 보건의료 빅데이터 활용을 위한 각국의 사례에서 도출한 시사점

첫째, 법률 체계 및 제도운영의 측면에 있어서 다양한 시사점이 발견된다.

- 먼저 개인정보보호와 보건의료정보의 관계에 주목할 필요가 있다. 현재 각국은 보건의료정보의 특수성을 고려하여 개인정보 일반에 대한 규율 이외에 보건의료정보에 관하여 별도의 상세한 규율을 하거나 규정을 두는 경우가 적지 않다. 이는 단순히 건강 관련 정보를 '민감정보'의 하나로 열거하고, 주로 인체유래물 '연구'의 관점에서 생명윤리 및 안전에 관한 법률상 유전정보를 부분적으로 다룰 뿐인 우리의 사정과 다르다. 나아가 GDPR 및 유럽 각국의 정보 보호법은 전형적으로 과학연구, 통계 등 일

정한 목적을 위해서는 민감정보라 하더라도 유연하게 활용할 수 있게 할뿐 아니라, 이때 과학연구가 광범위하게 해석되어야 함을 명문규정으로 확인하는 반면, 우리나라 법체계에선 민감정보에 해당하면 별도의 명시적 동의 또는 법률상 근거가 없는 한 이차 활용이 거의 불가능하다. 다른 한편, 보건의료정보의 보호범위가 일반적인 개인정보의 그것보다 더 넓은 경우가 있으며 보다 두터운 보호를 제공하는 경우도 있다.

- 보건의료정보의 활용을 위해서라도 보건의료정보 보호의 제고가 필요하다. 보건의료정보의 활용에는 광범위한 정보 공유와 결합이 필요하나, 이는 한편으로는 민감정보 등 높은 수준의 보호가 요구되는 종류의 보건의료정보에 대한 프라이버시 리스크를 높이기도 한다. 따라서 데이터 공유를 인정하면서도 안전장치를 확보하여야 한다. 가령 위반사항에 대한 제재를 강화하고 접근통제나 절차 등에 관한 제도를 마련하고 정비하여 정보의 오남용 가능성에 대응하여야 한다.

- 보건의료정보의 활용성을 제고하기 위한 법제도의 도입 역시 필요하다. 각국은 정보의 표준화와 상호운용성의 확보를 규율대상으로 다루고 있으며, 오늘날 EMR에 기록되는 전산화된 정보의 특성상 그 활용이 용이한 데에 반해 현재 우리나라의 경우 표준화와 상호운용성이 충분히 확보되지 않아 대책이 시급하다. 이와 관련하여 정보의 집적, 공유를 관리하는 법적 틀과 전담기구의 확보, 가명처리에 관한 규정 및 절차 마련 등에 주목해야 한다.

- 제도 운영의 면에서 IRB 심의제도의 개선과 이와 관련된 대안 연구가 추가되어야 하며, 동의 제도에 대한 개선 또한 검토가 필요하다.

둘째, 거버넌스 구축 및 감독기관의 측면에 있어서 다음과 같은 시사점이 있다.

- 보건의료 빅데이터의 활용에 관하여 대중의 참여, 참여자 보호, 기관 간 조정 등이 중요하다. 정보주체인 환자들과 치료의 주체인 병·의원 등 이해관계자가 모두 참여하여 직·간접적인 가치창출 및 가치교환이 이루어지게 하여야 하고 투명성이 확보될 필요가 있다. 이 과정에서 충분한 소통을 거쳐 동의와 자발적 참여를 유도하는 것이 필수적이다. 이들 과정을 관장하고 장, 단기의 국가전략을 수립, 집행할 기관도 필요하다.

- 빅데이터의 거버넌스 구축 과정에서 구체적인 보건의료정보 공유를 위한 조직을 어떻게 구성할 것인지의 문제가 제기된다. 현재 우리나라의 경우 국민건강보험공단과 건강보험심사평가원에 데이터가 분리되어 축적되고 있다. 관련 정보를 공유 목적으로 적절히 가공하여 일원적으로 관리하는 방안을 검토할 수 있다. 한편 정부의 재정지원으로 수행한 연구데이터 중 공유가 필요한 것을 검증, 공개, 공유하는 것 또한 고려할 수 있다. 장기적으로는 My Health Data와 같이 정보주체인 환자 자신에게 통제

권을 최대한 보장하고 이를 통하여 정보 공유와 가치 교환을 실현하는 모델을 추진할 필요가 있다.

- 감독기관의 측면에서는, 특히 연구 등 목적의 공유는 생명윤리 정책과 보건의료 정보의 활용이 모두 관계된다는 점에 주목하여야 한다.

셋째, 기술적 방안 및 법적 제재 측면에 있어 다음과 같은 시사점이 도출된다.

- 의료정보는 시간이 지날수록 단일 정보주체에 대한 데이터의 양이 늘게 마련이고 건강정보에는 매우 다양한 스펙트럼이 있다. 단기적으로는 보건의료 데이터에 특화된 비식별 조치 가이드라인과 절차를 마련할 필요가 있다. 다만 이러한 데이터 특성을 고려하여 너무 상세한 기술적 사항을 제시하는 것을 지양해야 한다. 정보주체의 권리 보장도, 법적인 점뿐 아니라 보건의료정보의 활용에 대한 신뢰 확보를 위해서도, 매우 중요하다. 이를 위해서는 유럽의 블록체인 기반 My Health My Data 등이 참고대상이 될 수 있다. 또한 일본이나 유럽의 데이터 공유 내지 거래를 보장하는 빅데이터 마켓 내지 플랫폼의 운영 사례 역시 참고할 만하다.

- 법적 제재의 차원에서, 기존 국내 법령은 형사처벌을 상대적으로 많이 활용하는 데 비해 전형적 위험통제이자 동태적 과정으로서의 개인정보 보호의 성격에 부합하는 개인정보보호 영향평가 제도는 일정 공공기관에 한정되어 있다. 영향평가제도를 민간 영역에도 확대하는 한편, 형사처벌 조항을 줄이고 행정적 제재권한을 강화할 필요가 있다. 나아가 보건의료정보의 특성을 고려하여 공유 주체에 제한을 두고 비밀유지의 무를 확장하여 엄격히 규제하는 것도 고려할 필요가 있다.

3. 연구의 결론 및 제언

분석결과 및 시사점을 근거로 보건의료정보 보호 및 활용에 관한 법률 개선 전략과 세부적인 개선 방안을 다음과 같이 제시하였다.

첫째, 법률체계 및 제도운영 측면

현재 많은 나라에서 보건의료 정보보호에 관한 감독의 대부분은 개인정보규제기관(Data Protection Authority; DPA)가 맡고 있으며, 별도의 감독기관을 두고 있는 입법례는 그리 흔하지 않다. 그러나 인간대상연구를 위하여 IRB의 감독이 필요하고, 이러한 차원에서도 정보보호에 관한 감독이 이루어진다. 기술적 비식별화 방법에 관한 경우 IRB의 심의를 받도록 하는 것이 바람직할 수도 있겠지만, 이를 IRB 위원들이 판

단하기 어려워 별도의 위원회를 두는 과정에서 중복규제를 받는 부작용이 나타나고 있어 그에 관한 대안이 필요하다.

둘째, 거버넌스 구축 및 감독기관 측면

공익목적의 보건의료 빅데이터의 복잡한 특성을 고려하여 데이터 분석과 관리를 위한 다각적 접근이 필요하다. 이에 보건의료 빅데이터를 수집하는 기관들은 공통 모듈로서 데이터를 체계화시킬 필요가 있다. 또 연구대상자의 이해와 인식이 중요하므로 데이터 관리에 있어 연구대상자 및 이해당사자들에게 데이터의 활용, 관리, 의사결정에 대한 이해를 도모하는 방식으로 데이터를 관리해야 한다.

셋째, 과학적 연구를 위한 빅데이터 활용

심의제도를 개선하여 IRB 체계를 단일화할 필요가 있고, 데이터 및 인체유래물의 이차 사용 동의서 제도의 개선도 필요하다. 또 IRB에서 개인정보 영향평가를 정확히 측정할 수 있도록 개선하기 위하여 기술가이드라인과 맥락을 같이 해야 한다. 빅데이터는 정보수집 당시에 그 용도를 구체적으로 특정하기 어려워 구체적이고 명시적인 동의를 받기 곤란하므로 포괄적 동의 또는 동의 없이도 당초의 수집 목적과 양립할 수 있는 범위의 활용을 가능하게 해야 한다. 다른 한편, 그러한 과정에서 프라이버시 리스크가 높아지지 않도록 제도화하여, 정보주체의 동의 및 철회 등 권리를 더욱 적극적으로 보장해야 한다. 또한 이후의 활용에 대한 지속적인 고지가 이루어져야 한다. 나아가 '신뢰할 수 있는 제3자(Trusted Third Party, TTP)'의 개입 등을 통하여 프라이버시 침해 위험을 줄여야 한다.

넷째, 이차 활용에 필요한 프라이버시 보호

이차활용 데이터 공유를 위한 협약을 마련하여야 한다. 당사자 별 권리의 종류와 범위를 정하고 정보주체와 이해관계자들의 법적 관계를 설명한 상세한 규정을 포함한 표준계약서 등의 실질적 형식의 마련이 필요하다.

다섯째, 기술적 방안 및 법적 제재 방안

연구대상자의 측면에서 특히 이차 활용과 관하여 생명윤리 및 안전에 관한 법률이 적용되는 경우 이차활용이나 제3자 제공이 엄격하게 통제되고 별도의 동의를 요하나, 이와 같은 방식에 대한 재고가 필요하다. 규제 수준을 통일하되, 프라이버시 보호 조치로서 이차활용 가능성에 대한 사전 설명, 처리정지 요구 등의 장치를 마련하는 것을 고려해 볼 수 있다. 특히 유전정보 등 인체유래물에서 도출된 정보가 복수의 정보

주체에 관련되는 점, 유출의 비가역성 등을 고려하여 법적 장치를 마련해야 한다.