

보건의료 빅데이터의 연구목적 사용에 대한 법제 개선방안

박대웅* · 정현학** · 정명진*** · 류화신****

I. 서론

II. 보건의료 빅데이터와 연구목적 사용

1. 보건의료 빅데이터의 개념
2. 보건의료 빅데이터의 종류
3. 보건의료 빅데이터의 특성
4. 연구목적 사용의 의미

III. 보건의료 빅데이터의 연구목적 사용에 대한 국내 규제 체계

1. 보건의료정보 이용의 원칙적인 모습
2. 보건의료정보의 연구목적 사용을 위한 예외 규정
3. 국내 규제체계의 한계

IV. 보건의료 빅데이터의 연구목적 사용에 대한 해외 규제 동향

1. 해외 규제 개요
2. 연구목적 사용에 대한 근거 법제
3. 비식별화 방식에 대한 규제
4. 연구목적 사용을 위한 연계 제도
5. 옵트아웃 제도에 대한 논의 동향
6. 해외 사례의 시사점

V. 보건의료 빅데이터의 연구목적 사용을 위한 국내 규제 개선 방향

1. 연구목적 사용 규정의 필요성
2. 비식별화 방식의 구체화
3. 보건의료 빅데이터 연계 방안 마련
4. 옵트아웃으로의 전환

VI. 결론

* 논문접수: 2016. 11. 30. * 심사개시: 2016. 12. 14. * 게재확정: 2016. 12. 29.

* 변호사, 한국보건산업진흥원 책임연구원.

** 한국보건산업진흥원 연구원.

*** 한국보건산업진흥원 보건산업기획단장(교신저자).

**** 충북대학교 법학전문대학원 부교수(교신저자).

* 본 연구는 보건복지부의 재원으로 한국보건산업진흥원의 보건의료기술연구개발사업 지원에 의하여 이루어진 것임(과제고유번호: H116C1000).

I. 서론

세계적인 경영컨설팅 업체인 맥킨지 앤 컴퍼니(Mckinsey & Company)는 빅데이터(Big Data) 기술이 활용될 대표적인 영역으로 보건의료를 지목하였다. 이 회사는 보건의료 빅데이터 기술이 미국에 미칠 잠재적 경제 가치를 연간 3천억 달러로 추산하고,¹⁾ ‘Right Living’, ‘Right Care’, ‘Right provider’, ‘Right value’, ‘Right Innovation’ 등 다섯 가지 경로를 통해 보건의료 패러다임을 변화시킬 것으로 예측하였다.²⁾ 이러한 보건의료 빅데이터를 효과적으로 활용하기 위해서는 분산되어 있는 보건의료정보를 통합하여 조사·연구에 사용가능한 형태로 이용자에게 제공할 수 있어야 한다. 분산된 보건의료정보를 통합하는 영국의 디지털국민보건서비스(National Health Service, NHS Digital)나, 민간 의료보험회사들의 정보통합플랫폼을 조정하는 미국 보건부 산하의 국가조정사무소(Office of the National Coordinator, ONC) 등에서 보건의료 빅데이터의 활용 인프라를 구축하고 있다.³⁾ 우리 정부도 빅데이터 연계와 개방을 통한 융복합서비스 창출, 공공기관 보건의료 빅데이터 연계 플랫폼을 구축하여 질병 재발률 분석, 약물반응 예측 등에 분석·활용 지원 등 보건의료 빅데이터 활용 전략을 지속적으로 추진하고 있다.⁴⁾ 하지만 부작용에 대한 우려가 없는 것은 아니다. 지난 2016년 9월, 보건의료인단체연합은 정부가 건강보험 빅데이터를 연구목적 등으로 민간에게 개방하는 정책들이 「

1) McKinsey Global Institute, Big data: The next frontier for innovation, competition, and productivity, 2011, 39.

2) Right Living이란 웰빙을 촉진하는 정보에 기반한 라이프스타일의 선택과 치료에 대한 소비자 자신들의 적극적 참여, Right Care란 환자에게 안전을 보장하면서 필요한 결과를 제공하는 증거기반치료, Right Provider란 개별 환자에게 가장 적합한 임상적 효과를 전달할 의사나 간호사를 포함한 의료공급자와 시스템, Right Value란 동일하거나 더 나은 결과를 비용 절감을 통해 달성하여 보건의료 가치 향상의 지속가능성 보장, Right Innovation이란 발견, 개발 그리고 안전 분야에서 연구개발의 생산성을 높이고 의학의 경계를 확장시키는 혁신을 의미한다[Mckinsey & Company, The ‘big data’ revolution in healthcare, 2013, 7].

3) 강희정, “보건의료 빅데이터의 정책 현황과 과제”, 보건복지포럼(통권 제238호), 2016, 57면.

4) 관계부처 합동, 서비스경제 발전전략, 2016.; 관계부처 합동, 보건산업 종합발전전략, 2016. 등.

개인정보보호법》을 위반하고 있다는 성명서를 발표한 바 있다.⁵⁾ 보건의료정보는 개인의 의료·질병과 관련된 정보가 포함되어 있는 민감정보이고 재식별 위험이 높은 정보임에도 불구하고 당사자의 동의 없이 민간에 개방되어서는 안 된다는 것이다. 이처럼 보건의료 빅데이터의 활용은 개인정보의 보호라는 가치와 긴장관계에 놓여있다. 보건의료 빅데이터를 연구목적으로 효과적으로 활용하기 위해서는 법적 기반이 갖추어져야 한다.

이를 위하여, 본 논문은 먼저 보건의료 빅데이터라는 규제 대상을 명확하게 확정하고 국내의 규제 현황과 한계, 그리고 해외의 규제 동향을 파악하고자 한다. 그리고 보건의료정보에 특화된 규제체계, 개인정보의 비식별화 방법, 분산되어 있는 보건의료정보의 연계방안, 개인정보 제공에 대한 동의 방식 등을 검토하고자 한다. 이를 통해 보건의료 빅데이터의 연구목적 사용과 개인정보 보호를 조화시킬 수 있는 법제 개선방향을 도출하는 것이 이 논문의 목적이다.

II. 보건의료 빅데이터와 연구목적 사용

1. 보건의료 빅데이터의 개념

우리 법제는 보건의료 빅데이터에 대한 직접적 정의를 내리고 있지 않다. 다만 「보건의료기본법」에 따르면 보건의료정보는 “보건의료와 관련한 지식 또는 부호·숫자·문자·음성·음향·영상 등으로 표현된 모든 종류의 자료(제3조 제6호)”를 의미하고, 이때의 보건의료를 “국민의 건강을 보호·증진하기 위하여 국가·지방자치단체·보건의료기관 또는 보건의료인 등이 행하는 모든 활동(제3조제1호)”으로 정의한다.⁶⁾ 다시 말하면 국민의 건강보호 및 증진에 관

5) 라포르시안, “건강보험 빅데이터를 제약사·보험사에 돈벌이에 제공?”, 9.9, 2016, <http://www.raportian.com/news/articleView.html?idxno=28549>, (2016.11.20. 확인).

6) 한편, 대법원은 의료행위에 대해 “의학적 전문지식을 기초로 하는 경험과 기능으로 진료, 검안, 처방, 투약 또는 외과적 시술을 시행하여 하는 질병의 예방 또는 치료행위 및 그밖에 의료인이 행하지 아니하면 보건위생상 위해가 생길 우려가 있는 행위(대법원 2004.1.15.

한 모든 행위를 보건의료행위로 볼 수 있고, 이러한 행위로 발생하는 모든 정보가 보건의료정보가 된다. 또한, 보건의료정보 중 개인을 식별할 수 있는 정보는 개인정보보호법상의 개인정보 및 민감정보에 해당할 수 있다. 개인정보보호법에 의하면 개인정보란 “살아있는 개인에 관한 정보로서 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보”를 의미하며, 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것도 포함한다(제2조제1호).⁷⁾ 그리고 개인정보보호법은 “사상·신념, 노동조합·정당의 가입·탈퇴, 정치적 견해, 건강, 성생활 등에 관한 정보, 그 밖에 정보주체의 사생활을 현저히 침해할 우려가 있는 개인정보로서 대통령령으로 정하는 정보(제23조제1항)”를 민감정보로 정의하고 있다.

빅데이터의 경제적 가치에 대한 국내외의 기대에도 불구하고, 빅데이터의 개념 자체는 단일하게 확립되어 있지 않다. 포함되는 개념의 범위가 이를 정의하는 주체마다 다양한데, 미국의 시장 조사기관인 가트너(Gartner)의 정의가 가장 일반적으로 사용되고 있다. 이에 따르면 빅데이터는 “더 나은 의사결정, 시사점 발견 및 프로세스 최적화를 위해 사용되는 새로운 형태의 정보처리가 필요한 대용량(Volume), 초고속(Velocity) 및 다양성(Variety)의 특징을 가진 정보 자산”으로 정의된다.⁸⁾ 여기서 데이터 규모의 방대성(Volume), 데이

2001도298 등)”라고 판시하고 있는 바, 이러한 대법원의 개념설정은 의료가 보건과 중첩성이 있음을 보여주고 있다. 의료에는 단순히 질병의 치료행위 뿐만 아니라 질병의 예방행위, 더 나아가 ‘보건상의 위해 발생우려’라는 개념표지까지 포괄된다고 판단하고 있는 것이다(박지용, “보건의료에 대한 헌법적 기초로서 개념적 및 역사적 접근”, 헌법학연구(제19권 제4호), 한국헌법학회, 2013, 515면).

7) 한편, 개인을 식별할 수 없어 법률상의 개인정보에는 해당하지는 않지만 ‘개인과 관련된’ 정보를 생각해볼 수 있다. 이러한 ‘개인과 관련된 정보’는 거리사진·항공사진·통계데이터와 같이 공개가 가능한 정보(비식별·공개가능)와 이동·구매·열람 등 개인의 행동이력과 같이 공개가 어려운 정보(비식별·비공개)로 분류할 수 있다. 하지만 프라이버시 보호 논의와 데이터 분석 기술의 발전으로 ‘개인과 관련된 정보’도 개인정보로 포섭되는 경향이 있다[보건복지부, 보건복지분야 빅데이터 서비스 제공방안 연구, 2013, 230면]. 즉, 개인 정보는 시대와 기술의 발전에 따라 그 범위가 확장되고 있다.

8) Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation

터 처리 및 분석의 속도(Velocity), 그리고 데이터 종류의 다양성(Variety) 즉, 3Vs를 빅데이터의 가장 일반적인 특징으로 꼽는다.⁹⁾

이상의 논의를 종합하면, 보건의료 빅데이터란 ‘국민의 건강보호 및 건강증진 행위로 인하여 축적되었으며, 대용량, 초고속 및 다양성의 특징을 바탕으로 더 나은 의사결정, 시사점 발견 및 프로세스 최적화에 활용될 수 있는 정보자산’ 정도로 정의할 수 있다. 본 논문에서는 이러한 보건의료 빅데이터 중 개인을 식별할 수 있어 개인정보보호법의 적용을 받는 빅데이터를 중심으로 논하도록 하겠다.

2. 보건의료 빅데이터의 종류

보건의료 빅데이터의 활용을 위해서는 빅데이터의 원천이 되는 보건의료정보의 종류에 대한 파악이 필요할 수 있다. 다양한 방식으로 분류할 수 있는데, 정보의 생산 및 보유 주체에 따라 공공기관과 민간기관으로, 데이터 유형에 따라 유전체 데이터, 진료정보 데이터, 그리고 라이프로그(Life Log) 데이터¹⁰⁾로 분류할 수 있다. 포함범위에 따라서는 ‘일반적인 개인정보’, ‘진료정보’, 의료기관에서 가공된 정보가 포함된 ‘협회의 의료정보’, 개인의 건강증진 및 예방에 대한 정보도 포함된 ‘건강정보’로 분류가 가능하다.¹¹⁾ 이러한 보건의료정보 중에서 우리나라는 공공기관, 진료정보 및 일반 개인정보 중심으로 빅데

[<http://www.gartner.com/it-glossary/big-data/>, (2016.11.20. 확인)].

9) 3Vs에 더하여 추가적인 특징 요소를 제시하는 경우도 있다. 가령 오라클(Oracle)은 빅데이터에 대한 정의와 특징으로 가치(Value)를 언급하고 있다. 즉, 데이터마다 경제적 가치가 다르기 때문에 빅데이터에 내재된 가치 있는 정보를 파악해내는 것이 중요하다는 것이다. 또한 가트너(Gartner)는 3Vs중에 2개 이상이 결합되어 문제가 발생하는 복잡성(Complexity)을 추가하여 빅데이터를 정의하고 있으며, 포레스터(Forrester)는 빅데이터 환경에서 그 데이터의 형태가 점차 증가하고 있음을 의미하는 변동성(Variability)를 그 특징으로 들고 있다[이명진·김우주, “빅 데이터를 위한 고급분석 기법과 지원 기술”, *Entrue Journal of Information Technology*(제11권 제1호), LG CNS, 2012, 47면].

10) 라이프로그 데이터란 웨어러블 기기, 이동통신, SNS 등으로 획득되는 취미, 건강, 여가 등 개인 생활 전반의 기록을 의미한다.

11) 김재선, “의료정보의 활용과 개인정보의 보호”, *행정법연구*(제44호), 행정법이론실무학회, 2016, 274면.

이터가 활용되는 편이다.¹²⁾ 우리나라의 대표적인 공공 보건의로 빅데이터는 국민건강보험공단의 표본연구DB,¹³⁾ 건강보험심사평가원의 환자데이터셋,¹⁴⁾ 질병관리본부의 한국인체자원¹⁵⁾ 등이 있다.

3. 보건의로 빅데이터의 특성

앞서 살펴본 대로 보건의로 빅데이터는 개인정보보호법에 따른 개인정보자 민감정보의 성격을 갖는다. 그리고 대부분의 보건의로 빅데이터를 공공기관에서 생산 및 보유하고 있다는 점에서 공개를 통한 공익적 활용이 유도되는 공공정보의 특성 또한 보유하고 있다.¹⁶⁾ 하지만 방대성, 속도 등과 같은 빅데이터의 특성으로 인해 정보 유출에 따른 피해 또한 증가할 수 있다.¹⁷⁾ 빅데이터의 유출은 즉시 그리고 대량으로 이루어지며, 따라서 피해 구제가 어렵고, 유

- 12) 민간영역의 임상데이터는 아직 국가적으로 활용될 만한 제도적·물리적 기반이 취약하고, 유전체 데이터는 외국에 비해 소규모인데다가 진료정보 등과의 연계가 제한되어 연계가 있다[강희정, 앞의 논문, 57면]. 라이프로그 데이터도 전자의무기록과의 연계가 제한되어 이를 활용한 건강관리 서비스를 제공하기 어렵다[김수영·김현경, “디지털헬스케어환경에서 개인정보의 활용과 규제의 합리적 조화방안 연구”, IT와 법연구(제12집), 2016, 236면].
- 13) 표본연구DB는 동일 대상자에 대해 사회·경제적 변수(거주지, 사망년월, 사망사유, 소득수준 등)가 포함된 자격자료, 진료내역 및 건강검진자료 등을 연결한 코호트 자료이다 [국민건강보험공유서비스 홈페이지, <https://nhiss.nhiss.or.kr/bd/ab/bdaba016iv.do>, (2016. 11.30. 확인)].
- 14) 환자데이터셋은 건강보험 청구자료를 기초로 하여 진료개시일 기준 1년간 진료 받은 환자를 대상으로 표본 추출한 연구목적의 자료이다[보건의로빅데이터개방시스템 <http://opendata.hira.or.kr/op/opc/selectPatDataAplInfoView.do>, (2016.11.30. 확인)].
- 15) 질병관리본부의 한국인체자원은행 사업으로 축적된 DB로, 동 사업은 인체유래물과 함께 해당 인체유래물 기증자의 임상, 역학정보 및 이로부터 분석된 유전정보 등을 아울러 인체자원으로 정의한다[질병관리본부 홈페이지, <http://cdc.go.kr/CDC/contents/CdcKrContentView.jsp?cid=61065&menuIds=HOME001-MNU1136-MNU1826-MNU1827>, (2016.11.30. 확인)].
- 16) 공공기관이 보유한 정보는 「공공데이터의 제공 및 이용활성화에 관한 법률」(이하 ‘공공데이터법’이라 한다)에 따라 규율된다. 이는 “공공데이터의 민간 활용을 통한 삶의 질 향상과 국민경제 발전에 이바지함을 목적(공공데이터법 제1조)”으로 하고 있다.
- 17) 「社会保障分野サブワーキンググループ」及び「医療機関等における個人情報保護のあり方に関する検討会」、医療等分野における情報の利活用と保護のための環境整備のあり方に関する報告書, 2012, 8.

출사례가 표면화되지 않을 수 있는 문제점이 그것이다.¹⁸⁾

4. 연구목적 사용의 의미

보건의료정보는 1차적으로 환자의 진료나 처방을 목적으로 사용된다. 하지만 연구, 공중보건, 건강보험, 마케팅 등 그 밖의 목적으로도 사용될 수 있다. 이를 보건의료정보의 2차적 이용¹⁹⁾이라고 하는데, 연구목적 사용이 대표적이다. 만약에 보건의료정보의 연구목적 사용이 허용된다면 우선 ‘연구’가 무엇인지에 대한 정의가 필요할 수 있다. 미국 연방규정에 따르면 연구는 “일반화할 수 있는 지식을 도출하거나 기여할 수 있도록 설계된 체계화된 조사”를 의미하며 체계화된 조사에는 연구개발, 시험 및 평가가 포함된다고 한다.²⁰⁾ 우리 법제는 연구를 직접적으로 정의하고 있지는 않다. 그런데 「생명윤리 및 안전에 관한 법률」(이하 ‘생명윤리법’이라 한다)은 ‘인간대상연구’를 정의하면서, ‘개인을 식별할 수 있는 정보를 이용하는 연구’를 포함하고 있다.²¹⁾

다만, 보건의료정보의 사용 형태에 있어, 연구목적 사용은 의료기관간 진료

18) 실제로 지난 2016년 4월 미국에서는 컴퓨터 네트워크 설정 실수로 655,000건의 헬스케어 데이터가 인터넷에 노출되는 사건이 발생한 바 있다[Healthcare IT News, Bon Secours says data breach affects 655,000 patients, 8.15, 2016, <http://www.healthcareitnews.com/news/bon-secours-says-data-breach-affects-655000-patients>, (2016.11.30. 확인)]. 또한 2015년 3월 미국의 대형 의료보험회사가 1,100만명의 개인정보가 유출되었다고 발표한 바 있는데, 조사결과에 의하면 첫 유출은 9개월 전인 2014년 5월이 발생한 것으로 추정된다[The New York Times, Premera Blue Cross Says Data Breach Exposed Medical Data, 3.17, 2015, http://www.nytimes.com/2015/03/18/business/premera-blue-cross-says-data-breach-exposed-medical-data.html?_r=0, (2016.11.30. 확인)].

19) 의료법 제21조 제2항은 「국민건강보험법」에 따른 급여비용 심사·지원, 「감염병의 예방 및 관리에 관한 법률」에 따른 역학조사 등 다양한 2차적 사용 목적을 열거하고 있다. 하지만 연구목적 사용에 대해서는 언급하고 있지 않다.

20) 45 C.F.R. §46.102 (d) Research means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge.

21) 생명윤리법의 규정과 미국 연방 규정을 고려하면, 개인식별정보를 이용하여 일반화된 지식을 도출하는 다양한 형태의 보건의료 빅데이터의 연구목적용 상정해 볼 수 있다. 인간유전체 데이터를 이용한 맞춤형 의약품의 개발, 진료정보 분석데이터를 활용한 건강보험 관리 시스템 개선, 라이프로그 데이터와 유전체 데이터를 연결한 만성질환 예측·관리 기술 등이 그것이다.

정보의 교류나 빅데이터를 활용한 건강관리업무와는 구별할 필요가 있다. 의료기관간 진료정보의 교류는 환자의 개인의료정보를 의료기관이나 약국 등이 공유하여 치료의 안전성과 효율성을 높이기 위한 것으로 일반화할 수 있는 지식의 도출과는 거리가 있다. 건강관리업무도 일반화된 지식을 얻는 것이 주목적이 아니라는 점에서 마찬가지이다. 일본 후생노동성에 의한 ‘의료 등 분야에 의 정보의 활용과 보호를 위한 환경정비에 관한 보고서’에 의하면, 일본은 의료 정보연계기반 구축에 있어서 의료기관간 진료정보 교류, 환자 본인의 건강관리업무, 연구목적 사용 등 3가지 취지에 따라 각각의 네트워크시스템을 구축하는 방안을 검토하고 있다고 한다.²²⁾ 미국의 「건강보험 정보의 이전 및 그 책임에 관한 법률(The Health Insurance Portability and Accountability Act, HIPAA)」 프라이버시규칙(Privacy Rule)도 연구활동과 건강관리업무를 구분하고 건강관리업무에 대해서는 프라이버시규칙상의 제한을 가하지 않는다고 한다.²³⁾

III. 보건의료 빅데이터의 연구목적 사용에 대한 국내 규제 체계

1. 보건의료정보 이용의 원칙적인 모습

가. 개인정보보호법 및 “개인정보 비식별조치 가이드라인”

개인정보보호제도는 헌법상의 개인정보자기결정권에 근거하고 있다. 헌법 재판소에 따르면 개인정보자기결정권은 “자신에 관한 정보가 언제 누구에게 어느 범위까지 알려지고 또 이용되도록 할 것인지를 그 정보주체가 스스로 결정할 수 있는” 권리이고, “인간의 존엄과 가치, 행복추구권을 규정한 헌법 제10

22) 김진영·홍태식, “의료분야에 있어 빅데이터의 활용 동향과 법적 제문제”, 법학논총(제40권 제3호), 단국대학교 법학연구소, 2016, 347~348면.

23) 박종현, “의료연구에 있어서 개인건강정보 보호 및 활용”, 법과사회(제37권), 법과사회이론학회, 2009, 256면.

조 제1문에서 도출되는 일반적 인격권 및 헌법 제17조의 사생활의 비밀과 자유"에 의하여 보장된다고 한다.²⁴⁾ 이러한 개인정보자기결정권을 보호하기 위하여 공공부문과 민간부문²⁵⁾을 망라하는 개인정보보호법이 2011년 9월 30일부터 시행되었다.²⁶⁾

보건의료정보는 개인정보이자 민감정보에 해당하기 때문에 원칙적으로 개인정보보호법에 따라 수집되어야 한다. 이에 따라 개인정보보호법은 개인정보처리자가 개인정보를 수집할 수 있는 경우를 열거하고 있다. 정보주체의 동의를 받은 경우, 법률에 특별한 규정이 있거나 법령상 의무를 준수하기 위하여 불가피한 경우, 공공기관이 법령 등에서 정하는 소관 업무의 수행을 위하여 불가피하게 필요한 경우, 정보주체와의 계약의 체결 및 이행을 위하여 불가피하게 필요한 경우 등이 그것이다(개인정보보호법 제15조). 원칙적으로는 보건의료정보의 수집 시 정보주체의 사전 동의가 필수적이다. 이를 옵트인(opt-in) 방식이라고 한다.²⁷⁾ 하지만 대다수의 보건의료 빅데이터를 생산 및 보유하는 공공기관의 경우 두 번째나 세 번째 요건, 즉 법률이나 소관업무에 근거하여 정보주체의 동의 없이 보건의료정보를 수집하는 경우가 많다.

이렇게 수집된 보건의료정보는 그 수집 목적의 범위 내에서 이용될 수 있고

24) 헌재결 2005.7.21, 2003헌마282.

25) 민간부문의 경우 「정보통신이용촉진 및 정보보호 등에 관한 법률」, 「신용정보의 이용 및 보호에 관한 법률」, 「금융실명거래 및 비밀보장에 관한 법률」 등이 영역별로 개인정보 활용을 규율하고 있다.

26) 또한 개인정보보호법에는 OECD가 권고한 개인정보보호 8개 원칙(OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data)이 반영되었는데, 다음과 같다. ① 수집 제한의 원칙(Collection Limitation Principle), ② 정보 정확성의 원칙(Data Quality Principle), ③ 목적 명확성의 원칙(Purpose Specification Principle), ④ 이용 제한의 원칙(Use Limitation Principle), ⑤ 안전 보호의 원칙(Security Safeguards Principle), ⑥ 개인적 참여의 원칙(Individual Participation Principle), ⑦ 공개의 원칙(Openness Principle), ⑧ 책임의 원칙(Accountability Principle)[<https://www.oecd.org/sti/ieconomy/oeecdguidelinesontheprivacyandtransborderflowsofpersonaldata.htm>, (2016.11.30. 확인)].

27) 반대로 당사자가 개인정보 처리에 동의하지 않겠다고 표시할 때부터 개인정보처리가 금지되는 방식을 옵트아웃(opt-out)이라고 한다. 옵트인·옵트아웃은 스펀메일을 규제하는 방식에서 유래하였는데, 옵트인은 수신자의 사전 동의가 있어야 메일 발송이 가능하고 옵트아웃은 수신자가 수신거부 의사를 밝혀야만 메일 발송이 금지된다.

(동법 제15조 제1항), 정보주체의 동의를 받은 경우 또는 목적 범위 내에서 개인정보를 제3자에게 제공할 수 있다(제17조 제1항). 개인정보의 목적 외 이용 및 제3자에 대한 제공은 원칙적으로 허용되지 않는다(제18조 제1항). 하지만 동조 제2항에서 예외적으로 목적 외 이용이나 제3자에 대한 제공이 가능한 경우를 열거하고 있는데(개인정보보호법 제18조), 그중 하나가 통계작성 및 학술연구 등의 목적을 위하여 필요한 경우로서 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 경우이다. 따라서 연구자가 연구목적으로 보건의료 빅데이터를 사용하기 위해서는 특정 개인을 알아볼 수 없는 형태로의 변화, 즉 비식별화가 필요하다.

이와 관련하여 개인정보를 비식별 조치하여 이용 또는 제공하려는 사업자 등이 준수하여야 할 기준을 제시한 “개인정보 비식별 조치 가이드라인”(이하 ‘비식별조치가이드라인’이라 한다)이 관계부처 합동으로 지난 6월 30일 발간되었다. 비식별조치가이드라인의 절차를 거쳐 비식별된 개인정보는 개인정보가 아닌 것으로 추정하여 당사자의 동의 없이 이용하거나 제3자에게 제공할 수 있다는 것이 가이드라인의 골자이다.²⁸⁾

나. 보건의료관련 법제

헌법 제36조 제3항은 “모든 국민은 보건에 관하여 국가의 보호를 받는다.”라고 하여 국민이 개개인의 건강을 영위하기 위해 국가에 필요한 권리를 주장할 수 있는 보건권을 명시하고 있다. 이를 구현하기 위해 보건의료기본법이 제정되었고, 보건의료의 모든 관련 법률을 포괄하는 일반법이라고 할 수 있다(보건의료기본법 제9조). 하지만, 보건의료기본법에는 “보건의료에 관한 통계와

28) 비식별조치가이드라인에 의하면 비식별조치의 절차는 개인정보에 해당하는지 여부를 검토하는 사전검토 단계, 정보집합물에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체하는 등의 방법을 활용하여 개인을 알아볼 수 없도록 하는 비식별조치 단계, 관련분야의 전문가로 구성된 ‘비식별조치 적정 평가단’이 다른 정보와 쉽게 결합하여 개인을 식별할 수 있는지 평가하는 적정성 평가 단계, 비식별 정보 활용 과정에서 제식별 방지를 위해 필요한 조치를 수행하는 사후관리 단계로 구성된다.

정보를 수집·관리하여 이를 보건의료정책에 활용할 수 있도록 필요한 시책을 수립·시행하여야 한다(제53조)”는 수집·관리에 대한 원론적인 규정만 있을 뿐 구체적인 개인정보 수집 절차나 방법 등은 규정하고 있지 않다. 따라서 개별 보건의료관련 법률에서 그 타당성을 찾아야 한다.²⁹⁾ 보건의료 연구와 관련해서는 생명윤리법과 의료법을 검토해 볼 필요가 있다.

생명윤리법은 인간을 대상으로 하는 연구에 있어서 연구수행 전 연구대상자로부터 개인정보 보호 및 개인정보 제공에 관한 사항이 포함된 서면 동의 획득을 요구하고 있다(제16조). 그리고 개인정보 제공에 대한 연구대상자의 서면 동의를 획득할 경우 기관생명윤리위원회의 심의를 거쳐 익명화한 후 제3자에게 제공이 가능하며, 연구대상자가 개인식별정보를 포함하는 것에 동의한 경우에는 익명화하지 않을 수 있다(제18조). 인체유래물연구에 필요한 동의 절차 및 제3자에의 제공절차도 이와 유사하다(제37조 및 제38조). 다만, 주민등록번호는 개인정보보호법이 “법령에서 구체적으로 주민등록번호의 처리를 요구하거나 허용한 경우”에 허용하고 있는데, 생명윤리법에는 이러한 근거 규정이 없어 처리가 불가능하다고 한다.³⁰⁾

의료인은 의료법에 따라 정보주체의 동의 없이 개인정보의 수집이 가능하다. 의료법에 의하면 의료인은 ‘진료기록부 등’을 갖추어 두고, 환자의 주된 증상, 진단 및 치료 등의 의료행위에 관한 사항과 의견을 상세히 기록하고 서명하여야 하는데, 이러한 진료기록부에는 진료를 받는 사람의 주소·성명·연락처·주민등록번호 등 인적사항 등이 포함된다(의료법 제22조). 수집한 의료정보의 2차적 이용과 관련하여서는 제21조의 기록 열람 규정에서 정하고 있다. 정

29) 보건의료법제의 영역은 매우 광범위한데 크게 보건과 의료로 나누어 볼 수 있다. 보건분야의 법률은 질병의 예방 등 국민보건과 관련된 법률이다. 「전염병예방법」, 「국민건강증진법」, 「식품위생법」, 「정신보건법」, 「공중위생관리법」 등이 이에 해당한다. 반면, 의료분야의 법률은 의료이용, 질병 치료 및 비용 지불 등과 관련된 법률로서 그 기능에 따라 ‘의료제공과 이용’, ‘의료비용’, ‘보건의료제조’, ‘보건의료 육성발전’ 등으로 세분화할 수 있다[한국보건산업진흥원, 보건의료산업시장분석 및 규제개선방안에 관한 연구, 2014, 21면].

30) 보건사회연구원, 보건의료빅데이터활용을 위한 기본계획수립, 2015, 790~791면.

보주체가 아닌 자에 대한 의료정보의 제공은 원칙적으로 불가하고, 예외사유를 동조 제2항 제1호부터 제16호까지 열거하고 있다. 그 가운데, 개별 법률에 구체적인 조항이 있는 경우로서 대표적인 예가 국민건강보험법상 급여비용 심사·지급·대상여부 확인·사후관리 및 요양급여의 적정성 평가·가감지급 등을 위하여 국민건강보험공단 또는 건강보험심사평가원에 제공하는 경우이다. 그러나 동법에서 정하고 있는 2차적 이용의 범위에 연구목적은 규정되어 있지 않다.

2. 보건의료정보의 연구목적 사용을 위한 예외 규정

개인정보 및 민감정보 성격을 갖는 보건의료 빅데이터를 연구와 같은 2차적 목적으로 사용하기 위해서는 원칙적으로 당사자의 동의가 필요하다. 하지만 데이터 규모의 방대성(Volume)이라는 빅데이터의 성격상 모든 당사자의 동의를 받기가 쉽지 않다. 이러한 문제점을 극복하고자 한정적인 범위에서 연구 목적으로 개인정보를 처리할 수 있도록 하는 규정들이 있다. 우선 「보건의료기술진흥법」에 의하면 한국보건의료연구원의 경우 국가기관 및 공공기관에 연구에 필요한 자료의 제출을 요청할 수 있다. 요청을 받은 기관은 개인 식별이 가능한 부분을 삭제하여 제출하는 것이 원칙이나, 연구를 위하여 두 개 이상의 국가기관 및 공공기관이 보유한 자료를 통합하여 분석할 필요가 있는 경우에는 개인 식별이 가능한 부분을 포함한 자료를 제출받아 자료의 통합작업을 수행할 수 있다(제26조). 또한 국가는 보건의료기본법 시행령 제15조에 따라 만성질환 장기역학조사에 관한 사업 수행을 위하여 불가피한 경우 건강정보 및 주민등록번호의 처리가 가능하다.

3. 국내 규제체계의 한계

우리나라의 보건의료 빅데이터 규제체계는 연구목적의 사용과 관련하여 다음과 같은 한계를 가지고 있다. 우선 개인정보보호에 있어서 연구목적의 특수

성을 고려하지 못하고 있다. 대표적인 사례가 2014년 8월 7일 시행된 개인정보보호법이 이 법 시행 당시 처리하고 있는 주민등록번호를 이 법 시행일로부터 2년 이내에 파기하도록 함으로써,³¹⁾ 보건의료 연구자들이 기존에 유지하고 있던 코호트 연구자료를 더 이상 사용할 수 없게 된 사건이다.³²⁾ 목적 외 이용이나 제3자에의 제공 시 당사자의 사전 동의를 요구하는 이른바 옵트인 원칙도 보건의료 빅데이터와는 부합하지 않는다. 민감도에 관계없이 모든 개인정보에 옵트인 원칙을 요구함으로써 방대성(Volume)의 성격을 갖는 보건의료 빅데이터의 사용이 매우 어렵게 된다. 학술연구 등의 목적을 위하여 필요한 경우에는 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 이른바 비식별화 방식을 통해 가능하다는 규정이 있기는 하다. 하지만 어느 정도의 수준으로 비식별화를 해야 할지, 재식별화에 따른 문제는 어떻게 처리할 것인지 등의 문제가 명확하게 해결되지 않아 아직까지도 논란이 있다. 비식별조치가이드라인이 이러한 문제의식에서 발간되어 있기는 하지만, 법적 구속력이 약하다는 점에서 분명히 한계가 있다. 마지막으로 분산되어 있는 보건의료 빅데이터를 연계할 수 있는 법적 기반이 약하다.³³⁾ 한국보건의료연구원, 만성질환과 같이 일부 기관이나 질환에 한해 연계가 가능한데, 연계 범위를 확장시키기 위해서는 법적 근거가 필요하다.

31) 개인정보보호법 부칙(2013.8.6.) 제2조(주민등록번호 처리 제한에 관한 경과조치) ① 이 법 시행 당시 주민등록번호를 처리하고 있는 개인정보처리자는 이 법 시행일로부터 2년 이내에 보유하고 있는 주민등록번호를 파기하여야 한다. 다만, 제24조의2제1항 각 호의 개정규정의 어느 하나에 해당하는 경우는 제외한다.

32) 보건사회연구원, 앞의 보고서, 789~792면.

33) 개인정보보호법은 연구목적으로 개인정보를 제3자에게 제공할 때 비식별화를 요구하는데(제18조), 고유식별기호가 삭제되어 데이터의 연계가 어려워진다.

IV. 보건의료 빅데이터의 연구목적 사용에 대한 해외 규제 동향

1. 해외 규제 개요

빅데이터라는 개념 자체가 비교적 최근에 등장한 만큼, 각국의 개인정보보호 관련법이 보건의료 빅데이터 연구에 관해서 직접적으로 언급하는 경우가 많지는 않다. 하지만 개인정보보호 관련법이 보건의료정보의 보호 목적으로 발전한 경우에는 해당 법제가 보건의료정보의 연구목적 사용에 대해서도 비교적 상세하게 규정하고 있다. 빅데이터 시대의 본격적 등장으로 빅데이터를 사용한 연구 전반에 활용될 수 있는 제도의 도입도 다양하게 논의되어 법제에 반영되고 있다. 개인정보보호 관련법의 적용 예외를 위한 개인정보의 비식별화 방법, 빅데이터의 활용가치를 크게 높일 수 있는 데이터 간의 연계 방안, 빅데이터의 기본적 성격과 배치되어 이를 사용한 연구의 가장 큰 장애로 여겨져 왔던 옵트인 제도의 옵트아웃으로의 전환 등이 그것이다. 아래에서는 해외 선진국들의 규제 동향에 대하여 검토하기로 한다. 이는 우리 법제 개선 방향에 중요한 시사점을 줄 수 있기 때문이다.

2. 연구목적 사용에 대한 근거 법제

가. EU

2016년 4월 4일 EU 역내 개인정보를 규율하는 「개인정보보호규정(General Data Protection Regulation, GDPR)」이 공포되었다(2년 후 시행 예정). 이는 과거의 지침(directive)³⁴⁾이 아닌 규정(regulation)에 해당하는 것

34) EU는 1995년 10월에 개인정보지침(EU Data Protection Directive)을 채택한 바 있다. 개인정보지침은 개인정보의 처리와 이동에 있어 개인정보보호에 대하여 규정하고 있는데, 특히 건강 또는 성생활에 관한 데이터의 처리를 금지하고 있다[DIRECTIVE 95/46/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL Article 8 The processing of special categories of data]. 하지만 예방의학, 의학적 진단, 치료 및 처치의 제공 또는 보건의

으로서 회원국들에 대한 구속력을 가지고 있다. GDPR은 기술의 발전 및 사회상을 반영하여 보건의료정보의 처리 유형을 유전정보, 고유 식별이 가능한 생체 인식 데이터, 건강정보, 성생활이나 성적 취향으로 보다 구체화하였다.³⁵⁾ 아울러 법률 또는 보건 전문가와의 계약에 의한 경우에, 안전장치가 취해진다면, 예방·직업 의학적 목적, 직원의 작업능력 평가, 의료 진단, 보건의료·사회복지 서비스 제공 및 관리는 개인정보 보호 규제의 적용을 받지 않는다.³⁶⁾ 또한 국외의 의료적 위기사태로부터의 보호, 의료·의약품·의료기기의 품질 및 안정성에 대한 높은 기준을 보장하기 위한 사용도 예외로 하고 있으며,³⁷⁾ 특히 법률에 근거하여 공익적, 과학적, 역사적 또는 통계적 목적으로 개인정보를 활용할 수 있다.³⁸⁾ 이는 기본적으로 GDPR이 공익을 위해 공중보건영역에서 정보주체의 동의 없는 개인정보처리의 필요성을 인정하고 있기 때문이다. GDPR에 의하여 유럽 각국의 법률에도 보건의료정보의 연구목적 사용에 대한 규정들이 삽입될 것으로 예상된다.

나. 미국 연방

미국은 1966년 HIPAA를 제정하여 의료영역의 개인정보보호를 규율하고 있다.³⁹⁾ HIPAA 프라이버시규칙은 보호받는 보건의료정보(Protected Health Information, PHI)의 이용과 공개에 대해서도 다루고 있는데, 특히 정보주체의 동의 또는 거부의 권한이나 기회 없이 사용 및 공개가 가능한 경우를 규정하고 있다.⁴⁰⁾ 이를 연구목적으로 한정하면, 기관심의위원회(Institu-

료 서비스 관리에 있어 정보처리가 요구되는 경우는 예외로 하고 있다.

35) EU GDPR Article 9 Processing of special categories of personal data.

36) EU GDPR Article 9 2.(h).

37) EU GDPR Article 9 2.(i).

38) EU GDPR Article 9 2.(j).

39) HIPAA의 경우, 의료보험 가입 환자들의 정보 제공 및 활용을 목적으로 제정이 추진되었으나 제정 과정에서 의료정보의 보호가 강조되어 상당부분 반영되었다고 한다[김재선, 앞의 논문, 278~279면].

40) 45 C.F.R. §164.512.

tional Review Board, IRB) 또는 개인정보보호위원회(Privacy Board)의 승인을 받은 경우, 연구 계획을 준비하기 위해 보건의료정보를 검토하는 경우, 그리고 사망자에 대해 연구하는 경우에는 동의절차가 필요하지 않다.⁴¹⁾ 그 밖에 법률에 의한 경우, 공중 보건 활동을 위한 경우 등에는 동의 없이 사용 및 공개가 가능하다.⁴²⁾ 다만, HIPAA는 연방법으로, 개별 주에서 이와는 다른 수준의 의료정보보호법을 제정할 수 있다. 실제로 2008년 뉴햄프셔주에서는 의료정보의 사용 용도를 연구나 마케팅이 아닌 치료만으로 제한하는 법안이 발의되었으나 하원에서 부결된 바 있다.⁴³⁾

다. 캐나다 온타리오주

캐나다는 주정부마다 별도의 의료정보 보호 법제를 가지고 있는데, 그중에서 온타리오주의 개인정보보호법이 가장 체계화되어 있다.⁴⁴⁾ 온타리오주의 개인정보보호법은 연구목적 공개와 연구목적 사용에 대해 규정하고 있다. 연구목적 공개의 경우,⁴⁵⁾ 연구계획서 및 연구계획을 승인하는 연구윤리기구(Research Ethics Board)의 증서를 의료정보 관리자에게 제출하는 방식으로 이루어진다. 그리고 연구윤리기구가 고려해야 할 사항,⁴⁶⁾ 관리자 및 연

41) 45 C.F.R. §164.512(i).

42) 동의 없이 이용 및 공개가 가능한 경우는 다음과 같다. (a) 법률에 의한 사용 및 공개 (b) 공중 보건 활동을 위한 사용 및 공개 (c) 학대, 방치 또는 가정폭력의 피해자와 관련된 공개 (d) 보건 감독 활동의 사용 및 공개 (e) 사법 및 행정 절차를 위한 공개 (f) 법 집행을 위한 공개 (g) 사망자에 대한 검시관, 의료심사관, 장의사를 위한 공개 (h) 사체 장기, 안구 또는 조직 기증 목적을 위한 용도로 사용 (j) 보건 또는 안전에 심각한 위협을 피하기 위한 용도의 사용 및 공개 (k) 국인 및 퇴역 군인의 활동, 국가 안보와 정보활동 등 특별한 정부의 기능을 위한 사용 및 공개 (l) 산업재해 등의 근로자 보상을 위한 공개[45 C.F.R. §164.512].

43) Nicolas P. Terry, LL.M., Legal Issues related to data access, pooling, and use, Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good: Workshop Summary, 2010, 157.

44) 방윤희·이현실·이일현, “국가별 개인정보보호법 및 의료정보보호법의 비교연구”, 한국콘텐츠학회논문지(제14권 제11호), 2014, 169면.

45) Ontario Personal Health Information Protection Act 44.

46) 연구윤리기구가 고려해야 할 사항은 해당 개인정보보호법을 사용하지 않고 연구의 목적을

구자가 지켜야할 사항⁴⁷⁾ 등을 규정하고 있다. 개인의료정보 관리자가 의료정보를 자신의 연구목적으로 사용하는 경우⁴⁸⁾에도 위와 같은 규정을 준수하여야 한다.

3. 비식별화 방식에 대한 규제

가. 미국 연방

미국은 PHI의 공개와 사용을 위해 HIPAA의 프라이버시규칙에서 보건의료정보로 개인을 식별할 수 없다고 인정하는 두 가지의 비식별화 방식을 규정하고 있다.⁴⁹⁾ 첫 번째 방식은 전문가결정 방식(Expert Determination Method)이고, 두 번째 방식은 세이프하버 방식(Safe Harbor Method)이다. 전문가결정 방식은 전문가가 데이터를 검토하고, 비식별화의 적절한 수단을 결정하며, 또한 이러한 결정을 정당화하는 분석의 방법과 결과를 기록하게 된다.⁵⁰⁾ 전문가에게는 통계 및 과학적 원리에 대한 지식이 요구되는데, 이와 관련된 다수의

달성할 수 있을지 여부, 연구 진행 시 개인의 사생활을 보호하고 정보의 기밀성을 유지하기 위한 적절한 보호장치가 마련되었는지 여부, 연구 수행 및 개인의료정보가 공개될 개인들의 프라이버시에 대한 대중의 관심, 개인의 동의를 받는 것이 불가능한지 여부 등이다.

47) 관리자는 개인정보를 공개하기 전에 연구자와 정보의 사용, 보안, 공개, 반환 또는 처분 등에 관한 조건 및 제한을 준수하는 내용의 합의를 하여야 한다. 개인정보를 제공받은 연구원은 연구 계획과 관련하여 연구 윤리위원회가 명시한 조건을 준수하여야 하고, 연구계획서에 명시된 목적으로만 정보를 사용하여야 하며, 개인을 식별할 수 있는 형태로 정보를 게시하지 않아야 한다. 그리고 법이나 규정된 예외 및 추가 요구조건을 제외하고는 개인 의료정보를 공개할 수 없고, 개인의 동의 없이 해당 개인과 직·간접 접촉이 금지되며, 관리인과의 합의나 연구원의 의무에 대한 위반이 발견되었을 경우 즉시 관리자에게 서면으로 알려야 하는 등의 제한이 있다.

48) Ontario Personal Health Information Protection Act 37(1)(j).

49) 45 C.F.R. §164.514(b).

50) 다만, 프라이버시규칙이나 미국보건성시민권국(Department of Health and Human Services Office of Civil Rights)의 이행가이드에서 어떻게 재식별화 위험이 계산되거나 계량화되어야 하는지를 제시하지 않고, 매우 적게 되어야 한다는 것 이외에 재식별화의 최소한의 허용된 위험을 제시하지는 않는다는 주장이 있다[Office of Civil Rights, Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, US Department of Health and Human Services, 2013, 5-6].

선례를 통해 전문가결정 방식이 효과적인 비식별화 수단으로 이용될 수 있다.⁵¹⁾ 세이프하버 방식(Safe Harbor Method)은 해당 기관이 ‘개인이나 친척, 고용주, 또는 개인의 가족’에 대한 데이터의 18개 특정 항목⁵²⁾을 제거하여 데이터를 비식별화하는 것이다.⁵³⁾ 세이프하버 방식은 규칙을 직접적으로 적용할 수 있고, 과정을 반복할 수 있는 장점이 있다.

나. 폴란드

폴란드는 데이터 보호에 관한 EU지침(Directive 95/46/EC)에 따라 「개인 데이터보호법(Personal Data Protection Act, PDPA)」를 제정하였다. 개인 데이터보호법 제6조에 의하면, 개인데이터란 식별되거나 식별가능한 개인과 관련된 모든 정보를 의미하고, 식별가능한 개인이란 개인식별번호 또는 신체적, 생리적, 정신적, 경제적, 문화적 또는 사회적 정체성과 관련된 하나 이상의 요소를 참조하여 직접 또는 간접적으로 식별가능한 개인을 의미한다. 특이한 점

51) Federal Committee on Statistical Methodology, Statistical Policy Working Paper 22: Report on Statistical Disclosure Limitation Methodology (Control), 2015, 8-56.

52) 여기서 말하는 18개 특정 유형은 다음과 같다. (a) 성명 (b) 주(state) 보다 작은 모든 지리적 행정구역으로서, 도로 주소(street address), 시(city), 카운티(county), 지구(precinct), 우편번호(ZIP code) 및 그와 동등한 지오코드(geocode)를 포함하나, 다음과 같은 경우는 제외 됨 (i) 처음 세 자리가 동일한 모든 우편번호를 조합하여 형성된 지리적 단위 내에 20,000명 이상 거주하거나 (ii) 20,000명 이하가 거주하는 이러한 지리적 단위에 대한 우편번호의 처음 세 자리가 000으로 변경된 경우 (c) 생일, 입원일, 퇴원일, 사망일을 포함하여 개인에 직접적으로 관련되는 일자에 대한 (연도는 제외) 모든 요소 및 89세 이상의 모든 연령과 이러한 나이를 나타내는 일자 (연도를 포함)의 모든 요소, 단, 90세 이상의 단일 범주로 종합할 수 있는 연령 및 요소는 제외 (d) 전화 번호 (e) 팩스 번호 (f) 이메일주소 (g) 사회보장번호 (h) 의료기록번호 (Medical Record Numbers) (i) 의료보험(Health Plan) 수혜자 번호 (j) 계좌번호 (k) 자격증/면허 번호 (l) 차량 번호판 번호를 포함한 차량 식별자 및 일련번호 (m) 기기 식별자 및 일련번호 (n) 웹 URL(Universal Resource Locator) (o) 인터넷프로토콜(IP) 주소 (p) 지문이나 성문(voice print)을 포함한 생체인증 식별자 (q) 얼굴 전체의 사진 및 그와 유사한 이미지 (r) 기타 고유한(unique) 식별 번호, 특징 또는 코드.

53) 다만, 어떠한 방식을 사용하더라도, 비식별화를 수행하는 포괄기관(covered entity)은 “단일 정보만으로 또는 다른 정보와 결합하여 정보의 대상인 개인을 식별할 수 있는 실제적 지식”을 가지고 있지 않아야 한다[45 C.F.R. §164.514(c)]. 그러나 미국 정부는 단순히 제식별 기법의 존재를 아는 것은 실제적 지식을 가지고 있는 것에 해당하지 않는다고 밝힌 바 있다[Office for Civil Rights, Op. cit., 28].

은 “식별을 위하여 불합리한 정도의 시간, 비용 및 인력을 필요로 하는 경우”에
 는 식별 가능한 것으로 여기지 않는다는 점이다.⁵⁴⁾ 이와 같은 경우의 예로 제
 시되는 것이 자동차 번호판, 도로명, 보수액수 등이다.⁵⁵⁾

4. 연구목적 사용을 위한 연계 제도

가. 스코틀랜드

스코틀랜드는 스코틀랜드정보연계협력체(Scottish Informatics and Linkage
 Collaboration, SILC)를 통해 다양한 데이터를 연구목적으로 연계하고 있다.⁵⁶⁾
 SILC는 데이터 연계를 위한 보호조치로 위험 비례 관리(proportionate risk
 management), 통계적 노출 관리(statistical disclosure control), 기능분리
 (separation of functions)를 적용하고 있는데, 이 중 기능분리를 위해 ‘신뢰할
 수 있는 제3자 색인(Trusted Third Party Indexing)’이라는 자료 연계 모델을
 사용한다.⁵⁷⁾ 이는 정보 관리자(Controller), 정보 색인작성자(Indexer), 정보
 연계자(Linker)의 기능 분리를 통해 어느 누구도 전체 정보에 접근할 수 없도록
 하여 개인정보보호를 최대한 추구하면서 자료연계를 허용하는 방식이다.⁵⁸⁾ 따
 라서 스코틀랜드는 정보 관리자가 직접 데이터 연계를 하지 않고, 각 기능을 분리

54) ACT of August 29, 1997 on the Protection of Personal Data Article 6 3. A piece of in-
 formation shall not be regarded as identifying where the identification requires an un-
 reasonable amount of time, cost and manpower.

55) Practical Law a Thomson Reuters legal solution, “Data protection in Poland: overview”,
<http://uk.practicallaw.com/6-520-7945?service=fs>, (2016.11.17. 확인).

56) SILC 홈페이지, <http://www.data-linkagescotland.co.uk/how-it-works>, (2016.11.17. 확인).

57) SILC 홈페이지, <http://www.data-linkagescotland.co.uk/data-linkage-safeguards>, (2016.11.
 17. 확인).

58) 구체적인 연계 방식은 다음과 같다. ① 각 정보 관리자는 보유하고 있는 데이터베이스에
 서 색인숫자를 생성하여 개인식별자와 함께 제3의 색인자에게 보냄 ② 제3의 색인자는
 개인식별자를 이용해 양쪽 기관에서 전송받은 자료를 연결해 연결숫자를 만들어 보내고,
 별도의 연결자에게 연결숫자의 대조표를 보냄 ③ 각 정보 관리자는 색인숫자를 이용하여
 내부정보와 연결숫자를 연결하고, 이를 별도의 연결자에게 보냄 ④ 연결자는 연결숫자의
 대조표를 이용해 각 정보를 연계 후 새로운 프로젝트숫자를 부여하여 연구자들이 해당
 정보에 접근할 수 있도록 함[보건사회연구원, 앞의 보고서, 796-797면].

하여 필요시 연계하는 방식을 이용한다고 할 수 있다.

나. 미국 뉴욕주⁵⁹⁾

주보험연구조합(Statewide Planning and Research Cooperative, SPARCS)은 보건의료산업계와 주정부 간의 제휴 관계로 1979년에 설립된 모든 지불자(자기 부담과 개인보험포함)의 병원비 지급시스템이다. 응급실, 진단 및 치료센터 방문 등을 포함한 입·퇴원, 외래 환자의 청구정보를 수집한다. 수집된 정보는 재정비율 설정, 정책 개발과 평가, 역학, 보건계획, 자원분배, 치료질 평가, 보건 의료 서비스 연구, 모니터링, 이용평가에 이용되며 다른 데이터셋과 등록정보(암, 출생)의 연계에도 이용된다. 아울러 SPARCS는 주의 공보험과 민간보험 정보 및 뉴욕주의 다양한 보건의료정보에 대한 중앙 저장소인 모든지불자데이터베이스(All Payer Database, APD)로 확대 개편되고 있는 중이다. 뉴욕주건강정보네트워크(Statewide Health Information Network for New York, SHIN-NY)는 뉴욕주 전체의 보건의료정보네트워크로서 의료질 향상, 비용 절감, 공중보건 향상을 위해 의료기관, 보험회사 및 보건 당국 사이의 안전한 보건 의료정보 교환을 지원하는 네트워크이다. 이러한 네트워크는 환자 정보 교류, 협력 진료, 주 차원 정책⁶⁰⁾ 개발에 사용되어 왔다. 뉴욕주 보건국은 향후 APD, SHIN-NY, 기타 정보 등을 아우르는 개인단위 데이터셋의 통합 방안도 가지고 있다. 즉, 뉴욕주는 지불·청구정보와 의무기록 등이 모두 통합되는 완전통합형 모델을 추구한다고 볼 수 있을 것이다.

59) New York State Department of Health, Health Information Technology Workgroup Final Report, 2015, 10-12.

60) 주요 정책으로는 Delivery System Reform Incentive Payment (DSRIP), State Health Innovation Plan(SHIP)이 있다.

5. 옵트아웃 제도에 대한 논의 동향

가. 영국

영국의 NHS Digital은 보건의료정보의 연계·활용을 위해 옵트아웃 제도를 도입한 바 있다.⁶¹⁾ 하지만 Care.data에 개인정보 제공 배제를 신청한 인원이 100만 명을 넘어서는 등 문제점이 제기되었다.⁶²⁾ 이에 국가데이터보호관(National Data Guardian)은 2016년 6월 정보 보호 체계 및 옵트아웃 제도에 전반에 대하여 검토보고서를 발표하였고,⁶³⁾ 같은 해 7월에 NHS는 Care.data를 폐기하였다. Care.data는 2년의 기간 동안 스타틴의 부작용, 항암제의 효과, 항우울제와 자살의 연관성을 밝혀내는 등 보건의료 연구에 많은 기여를 하였다.⁶⁴⁾ 하지만 제도에 대한 대중의 신뢰 획득 실패로, NHS는 이 업무를 의료공동체와의 긴밀한 협력 하에 국가정보위원회(National Information Board)에서 추진할 것이라고 발표하였다.⁶⁵⁾

61) 환자들은 주치의를 통해 자신의 보건의료정보가 제공되는 곳을 선택적으로 배제할 수 있다. 요약진료기록(응급정보, 앨리지정보 공유), 진료기록공유 데이터베이스(지역 간호사 관리 등을 위한 정보 공유), 케어닷데이터(Care.data, NHS Digital의 환자 건강기록 관리 프로그램으로 연구·정책 목적의 식별가능·비식별 정보 공개 등의 역할도 수행)에 대해서 선택적으로 개인정보 제공 배제를 요청할 수 있다[보건사회연구원, 앞의 보고서, 856면].

62) The Register, One million patients have opted out of Care.data., 4.20, 2016, http://www.theregister.co.uk/2016/04/20/one_million_patients_have_opted_out_of_caredata/, (2016.11.20. 확인).

63) 보고서에 따르면 대부분의 사람들이 시스템의 신뢰성을 중요하게 생각하고, 정보 제공 목적에 민감함이 확인되었다. 특히 보험, 마케팅 목적의 활용에 대한 우려가 심각함이 파악되었다. 이에 국가데이터보호관은 NHS의 운영, 보건의료기술 개발 등 공익적 또는 연구목적에 중심으로 옵트아웃 제도를 개선할 것을 권고하였다[National Data Guardian for Health and Care, Review of Data Security, Consent and Opt-Outs, 2016, 참고].

64) The Telegraph, How the NHS got it so wrong with care. data, 7.7, 2016, <http://www.telegraph.co.uk/science/2016/07/07/how-the-nhs-got-it-so-wrong-with-care-data/>, (2016.11.30. 확인).

65) 2013년에는 민감한 환자 기록에 대한 영향을 제대로 설명하지 않았다는 비난으로 인하여 Care.data의 추진이 중지된 바 있고, 데이터 공유에 따른 위험을 포함하지 않은 Care.data 교육 캠페인을 진행하여 문제가 되기도 하였다[ComputerWeekly.com, NHS England scraps controversial Care.data programme, 7.6, 2016, <http://www.computerweekly.com/news/450299728/Caldicott-review-recommends-eight-point-consent-model-for-patient-data-sharing>, (2016.11.30. 확인)].

나. 일본

일본은 민간분야와 공공분야를 구분해서 개인정보를 규제하고 있는데, 민간분야에는 「개인정보보호법」이, 공공분야에는 「행정기관이 보유하는 개인정보보호법」 및 「독립행정법인 등이 보유하는 개인정보보호법」이 있다. 일본 개인정보보호법은 옵트아웃 제도를 두고 있으며, 따라서 민간 개인정보취급 사업자는 일정한 조건 하에 당사자의 동의 없이 개인정보를 제3자에게 제공할 수 있다.⁶⁶⁾ 그런데 민간사업자, 행정기관, 독립행정법인 등 의료기관의 설치 주체에 따라 적용 법률이 다를 경우, 민감한 개인정보가 체계적으로 관리되지 못하거나 기관 간 정보 활용이 어려워지는 등의 문제가 있어 보건의료정보에 특화된 법률의 입법이 추진되어 왔다.⁶⁷⁾ 이러한 입법을 위한 논의에서도 옵트아웃 제도의 도입이 거론되고 있다.⁶⁸⁾ 다만, 옵트아웃 제도를 도입하더라도 정보의 이용을 정보주체에 통지하여야 한다는 의견, 사용 목적에 따라 차별화된 지침을 만들기 위한 근거를 법률로 정해야 한다는 의견 등이 제시되었다.

6. 해외 사례의 시사점

해외 법제 동향을 통해 많은 국가들이 보건의료 빅데이터의 활용에 있어 개인정보보호를 침해 불가한 절대적 가치로 여기지는 않다는 점을 파악할 수 있

66) 일본 개인정보보호법 제23조 제2항에 의하면 일정한 조건은 다음과 같다. 제3자 제공을 이용목적으로 한다는 점, 제3자에게 제공되는 개인데이터 항목, 제3자 제공의 수단 또는 방법, 본인의 요청에 따라 제3자 제공을 정지한다는 점을 통지하거나 용이하게 알 수 있는 상태에 둘 경우이다. 다만, 개정 개인정보보호법(2017 시행 예정)은 병력과 같은 개인의료정보가 포함된 요배려개인정보를 새로 정의하고(제2조제3호), 요배려개인정보에 대해서는 옵트아웃 방식을 금지하고 있어(제23조제2항), 보건의료 빅데이터 활용의 어려움이 가중될 것으로 보인다.

67) 「社会保障分野サブワーキンググループ」及び「医療機関等における個人情報保護のあり方に関する検討会」, Op. cit., 19.

68) 의료분야는 개인의 이익과 공익이 밀접하게 연관되어 있고, 동의 원칙의 경직된 적용은 연구 등을 통해 의료의 질을 향상시키는데 장애가 될 수 있기 때문이다[「社会保障分野サブワーキンググループ」及び「医療機関等における個人情報保護のあり方に関する検討会」(2012). Op. cit.: 13].

다. 개인정보보호와 보건의료 빅데이터 활용간의 균형을 이루려는 다양한 시도들이 확인된다. EU의 개인정보보호규정은 보건의료 연구 등의 목적으로 개인정보처리가 가능함을 규정하고 있으며, 미국·캐나다(온타리오주) 등도 보건의료에 특화된 정보보호법제를 통해 연구목적 사용의 절차를 상세하게 규율하고 있다. 비식별화를 위해서도 미국은 전문가 결정방식과 세이프하버와 같은 선택적 절차를 운영하여 보건의료정보의 활용을 유도하는 실용적 접근을 택하고 있다. 비식별화의 정도와 재식별위험성에 대한 논의 역시 활발함을 미국과 폴란드의 예에서 확인할 수 있었다. 보건의료정보의 가치를 높이기 위해 빅데이터를 연계하는 방법도 다양하게 시도되고 있다. 미국 뉴욕주 사례에서는 빅데이터 활용을 위해서 데이터들이 통합되는 경향을 확인할 수 있다. 특히, 정보 관리자, 정보 색인작성자, 정보 연계자의 기능을 분리하여 개인정보보호를 도모하는 스코틀랜드 사례를 주목할 필요가 있다. 마지막으로 옵트아웃 제도의 도입에 관한 선진국들의 논의도 활발하게 진행되고 있다. 특히 시스템의 신뢰성과 제공 목적이 주요 쟁점이 된 영국의 옵트아웃 논란을 주목해야 할 것이다.

V. 보건의료 빅데이터의 연구목적 사용을 위한 국내 규제 개선 방향

1. 연구목적 사용 규정의 필요성

앞서 살펴본 대로, 보건의료정보는 보호되어야 할 개인정보 및 민감정보의 특성과 공익적 목적으로 활용되어야 할 공공정보의 특성을 동시에 갖는다. 이러한 이질적인 성격을 조화시켜 보건의료 빅데이터를 활용하기 위해서는 보건의료정보에 특화된 정보보호법제가 필요할 것이다. 이는 보건의료정보의 2차적 이용, 특히 연구목적 사용에 대해서 절차와 책임을 상세하게 규정할 필요

가 있다. 우선 EU의 GDPR처럼 연구의 목적을 상세하게 정해야 한다. 미국이나 캐나다와 같이 연구계획을 심의하는 기구의 권한과 책임을 명시해야 하며, 특히 개인정보 관리자 및 연구자의 책임을 규정해야 한다. 앞서 설명한 비식별 조치, 데이터 연계, 업트아웃 등의 근거 규정도 필요할 것이다.

2. 비식별화 방식의 구체화

우선, 감당할 수 있는 재식별위험의 정도를 법적으로 명확하게 규정하여야 한다. 앞서 살펴본 대로, 현행 개인정보보호법으로는 빅데이터 관리자가 연구자들에게 데이터를 제공할 때 어느 수준까지 비식별화를 해야 하는지 판단하기 어렵기 때문이다. 비식별화 위험 수준에 대한 미국에서의 사례와 다양한 논의에 대해서 주목할 필요가 있다. 대표적인 비식별화 방식인 전문가결정 방식과 세이프하버 방식 모두 장점과 위험성에 대해 상세한 분석이 이루어지고 있다. 폴란드와 같이 재식별에 불합리한 시간, 비용 및 인력이 소요되는 경우에는 비식별된 것으로 추정하는 등으로 면책 범위를 법률에 명시하는 방법도 생각할 수 있다.

다음으로, 비식별화와 익명화를 명확하게 구분할 필요가 있다. 보건의료 빅데이터를 둘러싼 상당수 논란이, 개인정보를 비식별화 한 경우에 이를 개인정보로 취급하지 않고 개인정보보호법을 적용하지 않는다는 점에 있다. 하지만 비식별화를 해도 재식별 가능성이 조금이라도 남아있으면 개인정보로 보아야 한다는 주장이다. 비식별화와 개인정보를 완전히 삭제하는 익명화는 분명히 차이가 있다.⁶⁹⁾ 그러나 익명화된 정보는 보건의료 빅데이터로서의 활용가치가 거의 없는 경우가 많다. 개인정보보호 논란을 줄이고 공익적 활용도를 높이

69) 이와 관련하여 비식별화(de-identification)와 익명화(anonymization)는 명확하게 구분할 수 없다는 의견도 있다. 미국은 비식별화이라는 용어를 주로 사용하나 익명화를 혼용하기도 하며, EU나 일본은 익명화라는 용어를 사용하나 개념상 비식별화와 명확히 구분되지 않는다는 것이다. 다만, 개념 구분의 실익에 대해서는 향후 논의가 필요할 수 있다고 한다[개인정보보호위원회, 개인정보의 비식별화 처리가 개인정보 보호에 미치는 영향에 관한 연구, 2015, 8~9면].

기 위해서는 관련 법령에서 비식별화와 익명화를 구분하여 명시하고, 앞서 논의한 대로 면책되는 비식별화의 정도를 명확히 제시하여야 한다.

3. 보건의료 빅데이터 연계 방안 마련

보건의료 빅데이터의 가치 향상을 위해서는 기관별로 분산되어 있는 데이터를 연계할 수 있는 근거규정이 필요하다. 연계의 방식으로는 기관별로 플랫폼을 구축하고 자율적으로 상호 연계하는 기관분산 방식, 국가에서 플랫폼을 구축하고 이를 통해 각 기관의 데이터를 연계하여 제공한 뒤 삭제하는 연계서비스 방식, 건강관리·건강보험·연구 등 주제별로 플랫폼을 구축하여 서로 연계하는 주제통합 방식, 공공기관이 관리하는 보건의료 데이터를 모두 통합하여 보관하는 완전통합 방식을 생각해 볼 수 있다.⁷⁰⁾ 빅브라더(big brother)에 대한 사회적 우려 때문에 완전통합 방식은 어려워 보인다. 주제통합 방식은 플랫폼 간 연계에 한계가 있어, 미국 뉴욕주도 이를 통합형으로 전환하고 있다. 기관분산 방식은 표준화가 어렵고 연계 과정에서 정보 유출의 우려도 크다. 연계서비스 방식은 스코틀랜드의 예에서 보듯 기능의 분리를 통해 개인정보 유출의 위험을 최소화할 수 있다. 또, 보건의료가 아닌 타 영역과의 연계도 수월하게 제공할 수 있다. 이러한 연계서비스 방식이 가장 바람직하다고 생각된다.

4. 오픈아웃으로의 전환

앞서 설명한 바와 같이, 오픈인 방식은 빅데이터 시대의 본격적인 등장을 반영하지 못한 측면이 있다. 이러한 문제 인식 아래서 영국에서는 오픈아웃 제도를 도입하여 운영한 바 있다. 사회적 신뢰 확보의 실패로 Cara.data를 폐기하였으나, 제도의 필요성은 여전히 인정되고 있다. 일본 후생성의 ‘의료 등 분야에서의 정보의 활용과 보호를 위한 환경정비에 관한 보고서’도 오픈아웃 제도의

70) 보건사회연구원, 앞의 보고서, 872면.

도입을 논의하고 있다. 우리도 옵트아웃 제도의 도입에 대한 진지한 논의가 필요하다. 물론 개인정보는 종류에 따라 그 민감도와 성격이 천차만별이므로 옵트아웃 제도의 운영에 대한 상세한 가이드라인이 필요할 것이다. 열람청구권, 정정·삭제청구권, 처리정지 청구권 등 개인정보 보호를 위한 절차법적 보완도 필수적이다.⁷¹⁾ 무엇보다 영국의 사례에서 보듯 제도에 대한 대중의 신뢰가 필수적이다. 옵트아웃으로의 전환은 공론의 장에서 충분한 토론을 거친 뒤에 사회적 합의를 기반으로 이행되어야 할 것이다.

VI. 결론

규제와 현실의 괴리는 발전속도가 빠른 과학기술의 영역에서 적지 않게 발생하는 일이다. 빅데이터도 예외가 아니어서, 개인정보보호법이 경제재로서의 빅데이터의 측면을 고려하지 못하여 많은 활용상의 애로사항이 발생한다고 한다.⁷²⁾ 물론, 국민 모두에게 고유식별정보에 해당하는 주민등록번호가 부여되고, 건강보험에 강제 가입되어 전국민의 민감한 건강정보가 중앙에 집중되는 우리나라의 상황에서 개인정보보호의 필요성은 충분히 강조될 필요가 있다. 의료인을 제외하고도 의료기관의 접수계원, 검사실의 기사, 처방전을 접하는 약사뿐만 아니라 건강보험공단 및 건강보험심사평가원의 직원에 이르기까지 의료시스템의 분업화로 인한 진료기록의 합병적 열람 직종만 70여 가지에 달한다고 한다.⁷³⁾ 하지만 의료정보의 전산화, 클라우드 시스템의 도입, 빅데이터 분석 기술의 발전 등 새로운 기술의 등장은 보건의료 빅데이터의 공익적

71) 윤석진, “개인정보 보호와 빅데이터 활용의 충돌, 그 문제와 입법정책 과제”, 중앙법학(제17집 제1호), 2015, 39면.

72) 김승환, “빅 데이터 시대의 도래에 따른 개인정보 보호법의 한계와 개선방향”, 연세 의료·과학기술과 법(제4권 제1호), 2013, 127면.

73) 김주한, “진료정보 2차사용과 새로운 환자-의사 간 관계”, 의료정책포럼(제9권 제1호 통권32호), 2011, 55면.

활용 가치를 높이고 있다. 이러한 규제환경의 변화는 환자의 치료라는 1차적 목적을 중심으로 한 보건의료 빅데이터 규제체계를 기술의 발전에 부합하며 공익적 활용이 용이한 형태로 개선시켜야 할 필요성을 보여주고 있다.⁷⁴⁾

지금까지 살펴본 대로 개인정보의 보호와 연구목적 사용의 균형을 위해서 보건의료 빅데이터 법제는 정교하게 구성되어야 한다. 우선 보건의료정보에 특화되고 보호와 활용을 아우르는 법제가 필요하다. 개인정보의 비식별화도 개인정보보호법의 적용이 배제되는 수준을 명확하게 규정해야 한다. 보건의료 빅데이터를 연계할 수 있는 법적 기반의 구축을 통해 관련 연구에 시너지 효과를 창출해야 한다. 외국에서의 논의에 대한 검토와 사회적 합의를 바탕으로 옵트아웃 제도의 도입도 고려해야 한다. 하지만 무엇보다 중요한 것은 이러한 시스템에 대한 국민의 신뢰일 것이다. 자신이 제공한 개인정보가 최소한의 위험아래서 연구에 사용되고 그 연구의 과실이 다시 자신에게 돌아올 것이라는 국민의 신뢰가 확보되어야 한다. 이를 위해서 정부, 보건의료계, 산업계 등 관련 전문가들도 보건의료 빅데이터의 연구목적 사용에 대한 사회적 공감대 형성에 노력해야 할 것이다.

74) 같은 의견으로 김수영·김현경, 앞의 논문, 241~243면; 윤석진, 앞의 논문 36면 등.

[참 고 문 헌]

- 강희정, “보건의료 빅데이터의 정책 현황과 과제”, 『보건복지포럼』, 2016.
- 개인정보보호위원회, 『개인정보의 비식별화 처리가 개인정보 보호에 미치는 영향에 관한 연구』, 2015.
- 관계부처합동, 『개인정보 비식별 조치 가이드라인-비식별 조치 기준 및 관리·체계 안내』, 2016.
- _____, 『보건산업 종합발전전략』, 2016.
- _____, 『서비스경제 발전전략』, 2016.
- 김수영·김현경, “디지털헬스케어환경에서 개인정보의 활용과 규제의 합리적 조화방안 연구”, 『IT와 법연구』, 2016.
- 김승환, “빅 데이터 시대의 도래에 따른 개인정보 보호법의 한계와 개선방향”, 『연세 의료·과학기술과 법』, 2013.
- 김재선, “의료정보의 활용과 개인정보의 보호”, 『행정법연구』, 2016.
- 김주환, “진료정보 2차사용과 새로운 환자-의사 간 관계”, 『의료정책포럼』, 2011.
- 김진영·홍태석, “의료분야에 있어 빅데이터의 활용 동향과 법적 제문제”, 『법학논총』, 2016.
- 라포르시안, “건강보험 빅데이터를 제약사·보험사에 돈벌이에 제공?”, 2016.9.9.
- 박종현, “의료연구에 있어서 개인건강정보 보호 및 활용”, 『법과사회』, 2009.
- 박지용, “보건의료에 대한 헌법적 기초로서 개념적 및 역사적 접근”, 『헌법학연구』, 2013.
- 방윤희·이현실·이일현, “국가별 개인정보보호법 및 의료정보보호법의 비교연구”, 『한국콘텐츠학회논문지』, 2014.
- 보건복지부, 『보건복지분야 빅데이터 서비스 제공방안 연구』, 2013.
- 보건사회연구원, 『보건의료빅데이터활용을 위한 기본계획수립』, 2015.
- 한국보건산업진흥원, 『보건의료산업시장분석 및 규제개선방안에 관한 연구』, 2014.
- 윤석진, “개인정보 보호와 빅데이터 활용의 충돌, 그 문제와 입법정책 과제-보건 의료빅데이터를 중심으로”, 『중앙법학』, 2015.
- 이명진·김우주, “빅 데이터를 위한 고급분석 기법과 지원 기술”, 『Entrue Journal of Information Technology』, 2012.

- Federal Committee on Statistical Methodology (2015), Statistical Policy Working Paper 22: Report on Statistical Disclosure Limitation Methodology (Control).
- McKinsey & Company (2013), The 'big data' revolution in healthcare.
- McKinsey Global Institute (2011), Big data: The next frontier for innovation, competition, and productivity.
- National Data Guardian for Health and Care (2016), Review of Data Security, Consent and Opt-Outs.
- New York State Department of Health (2015), Health Information Technology Workgroup Final Report.
- Nicolas P. Terry, LL.M. (2010), Legal Issues related to data access, pooling, and use, Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good: Workshop Summary.
- Office for Civil Rights (2015), Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule.
- 「社会保障分野サブワーキンググループ」及び「医療機関等における個人情報保護のあり方に関する検討会」(2012), 医療等分野における情報の利活用と保護のための環境整備のあり方に関する報告書.

[국문초록]

보건의료 빅데이터의 연구목적 사용에 대한 법제 개선방안

박대웅(변호사, 한국보건산업진흥원 책임연구원),

정현학(한국보건산업진흥원 연구원),

정명진(한국보건산업진흥원 보건산업기획단장),

류화신(충북대학교 법학전문대학원 부교수)

빅데이터 처리기술의 발전과 함께 보건의료 빅데이터의 잠재적 가치도 크게 주목받고 있다. 그리고 이러한 잠재적 가치를 현실화하기 위해서는 보건의료 빅데이터를 이용한 다양한 연구가 필수적이다. 하지만 개인정보보호법을 중심으로 한 빅데이터 규제체계는 경제체로서의 빅데이터의 측면을 고려하지 못하고 있어, 이를 연구목적으로 활용하는데 많은 장애를 발생시키고 있다. 환자의 치료라는 1차적 목적을 중심으로 한 보건의료정보의 규제체계를 기술의 발전에 부합하며 공익적 활용이 용이한 형태로 개선시켜야 한다. 이를 위해서 개인정보의 보호와 활용의 균형에 대한 고민이 반영된 해외의 법제 동향을 살펴볼 필요가 있다. 해외 법제의 시사점을 바탕으로 우리 법제에서 다음과 같은 방향의 개선점을 도출할 수 있다. 우선 보건의료정보에 특화되고 보호와 활용을 아우르는 법제가 필요하다. 개인정보의 비식별화도 개인정보보호법의 적용이 배제되는 수준을 명확하게 규정해야 한다. 보건의료 빅데이터를 연계할 수 있는 법적 기반의 구축을 통해 관련 연구에 시너지 효과를 창출해야 한다. 외국에서의 논의에 대한 검토와 사회적 합의를 바탕으로 옵트아웃 제도의 도입도 고려해야 한다. 하지만 무엇보다 중요한 것은 이러한 시스템에 대한 국민의 신뢰일 것이다.

주제어 : 보건의료, 빅데이터, 의료정보, 비식별화, 옵트아웃, 개인정보보호법

Improving Legislation on the use of Healthcare Data for Research Purposes

Park Dae Woong

Korea Health Industry Development Institute

Jeong Hyun Hak

Korea Health Industry Development Institute

Jeong Myung Jin

Korea Health Industry Development Institute

Ryoo Hwa Shin

Chungbuk National University

=ABSTRACT=

With the development of big data processing technology, the potential value of healthcare big data has attracted much attention. In order to realize these potential values, various research using the healthcare big data are essential. However, the big data regulatory system centered on the Personal Information Protection Act does not take into account the aspect of big data as an economic material and causes many obstacles to utilize it as a research purpose. The regulatory system of healthcare information, centered on the primary purpose of patient treatment, should be improved in a way that is compatible with the development of technology and easy to use for public interest. To this end, it is necessary to examine the trends of overseas legal system reflecting the concerns about the balance of protection and utilization of personal information. Based on the implications of the overseas legal system, we can derive improvement points in the following directions from our legal system. First, a legal system that specializes in healthcare information and encompasses protection and utilization is needed. De-identification, which is an exception to the Privacy Act, should also clearly define its level. It is necessary to establish a legal basis for linking healthcare big data to

create synergy effects in research. It is also necessary to examine the introduction of the opt-out system on the basis of the discussion on the foreign debate and social consensus. But most importantly, it is the people's trust in these systems.

Keyword: healthcare, bigdata, medical record, de-identification, opt-out, Personal Information Protection