

보건의료 빅데이터 관리시스템 최신 동향

장 성 재

한국원자력의학원

E-mail: sjsjj@kirams.re.kr

요약문

최근 보건의료 산업이 치료 중심에서 예방 및 건강관리 중심으로 이행됨에 따라서 질병 발생 가능성을 예측하고 사전에 예방할 수 있는 관련 기술에 대한 요구가 증가하고 있으며, 이와 더불어 개인 맞춤형 정밀의료 서비스의 중요성 또한 증가하고 있는 추세이다. 또한, 건강검진 자료, 질병 자료, EMR (Electronic Medical Records), 유전체 분석 데이터 등 정밀의료 및 의료 영상을 중심으로 관련 데이터의 양이 급격히 증가하고 있다. 특히 인간유전체 염기서열 판독기술의 비약적인 발전으로 인해 생명공학의 기술 패러다임이 새로운 세대로 이동하고 있으며, 각 개인의 유전자 특성 정보를 분석할 수 있게 됨에 따라 헬스케어의 패러다임이 발병 후 치료에서 발병 전 예측 및 맞춤형 관리/치료의 개념으로 전환되고 있다. 본 동향 리포트에서는 보건의료 빅데이터 관련 법규 및 정보관리시스템에 대한 글로벌 트렌드를 소개하고 국내 현황과의 비교를 통해 관련 연구자 및 산업계에 최신 지견을 소개하고자 한다.

Key Words: 유전체, 정밀의학, 빅데이터, 개인정보, 기술 패러다임 변화

목 차

1. 서론
2. 본론
 - 2.1 보건의료 빅데이터 시장 동향
 - 2.2 보건의료 개인정보 관련 법체계
 - 2.3 보건의료 빅데이터와 인공지능의 결합
 - 2.4 개인 유전체 기반 맞춤형 의료서비스
 - 2.5 개인정보 비식별화 조치
 - 2.6 보건의료 표준화 현황

3. 결론
4. 참고문헌

1. 서론

한국생명공학연구원은 지난 2016년 12월 28일 바이오헬스 이슈를 선도하는 10대 미래유망기술을 선정하여 공개했다(표 1). 이 기술들에는 대규모 임상 유전체 정보관리 기술 및 모바일 인공지능 진단 기술 등 국내 전문가들이 선별한 향후 유망할 것으로 예측되는 기술들이 대거 포함되어 있다. 특히 단순 치료 중심에서 사전 진단·예방 및 맞춤형 치료로의 보건의료 패러다임의 변화에 부응하기 위해서는 폭발적으로 증가하고 있는 보건의료 데이터를 비롯하여 2~8번에 이르는 이슈들을 해결하기 위한 차세대 시퀀싱 기술이 필수적이며, 이 기술을 통해 생산된 방대한 데이터를 분석하기 위해서는 이들 보건의료 빅데이터를 통합하여 관리할 수 있도록 “대규모 임상 유전체 정보관리기술”가 선행되어야 할 과제 중의 하나이다[1].

“대규모 임상 유전체 정보관리기술”의 요소 기술로는 첫째, 임상-유전체 정보 데이터 웨어하우스를 통한 임상 유전체 정보의 표준화 및 관련 정보의 데이터베이스(DB)화가 필수적이며, 이를 위해 기존의 임상 정보와 유전체 정보의 통합을 위한 프로토콜의 개발 및 궁극적으로 환자의 임상 정보와 유전체 정보를 통합한 ‘임상-유전체 DB’를 구축하여야 한다. 둘째, 바이오 빅데이터 기반의 임상-유전체 정보분석 기술이 필요한데, 대용량의 암 임상-유전체 정보 기반 진단, 치료 및 예후 바이오마커 분석 등을 위한 알고리즘 및 시스템 개발이 필수적이다. 또한, 구축된 임상-유전체 DB의 기본적인 분석을 할 수 있는 인공지능 기반의 임상 의사결정지원시스템(Clinical Decision Support System, CDSS)과 같은 IT 기반의 클라우드 컴퓨팅 분석 파이프라인 및 서비스 시스템의 개발도 요구된다.

글로벌 보건의료 분야 데이터 분석시장은 2014년~2019년 연평균 25% 이상의 성장률이 예측되며 미래 경제성장을 견인하는 유망분야로 주목 받고 있다. 2013년 McKinsey 보고서에 따르면 “지난 수십 년간 의료기관의 의무기록을 디지털화하고, 제약회사가 연구개발 데이터를 전자 DB로 통합하며, 정부와 공공기관들이 저장된 데이터를 이용·검색·실행 가능하도록 투명성을 확대함에 따라서 보건의료 분야에서도 정보공개가 진행되고 있으며 이러한 데이터의 유동성 확대는 보건의료 산업을 변화시킬 것”이라고 전망했다[2, 3]. 하지만, 세계 빅데이터 시장에서 빠른 속도로 진행되고 있는 산업화와 기술 경쟁은 보건의료 빅데이터의 공익적 활용 가치와 산업적 가치 창출의 균형을 맞추기 어렵게 한다. 공적 건강보험제도에 기초한 국내 보건의료시스템에서 보건의료 빅데이터는 국가 서비스 산업 발전 전략의 일부로 추진되고 있으며, 보건의료시스템 전반에서 혁신을 유도하고 정보 생산의 주체인 국민에게 보건의료 빅데이터의 분석 편익을 돌려주는 공익적 활용의 국가전략 수립은 지연되고 있는 실정이다[4].

표 1. 2017년 바이오 미래유망기술

(부제: 바이오헬스 이슈를 선도하는 10대 미래유망기술)

	이슈	10대 미래유망기술
1	보건의료 빅데이터 플랫폼 구축	대규모 임상 유전체 정보관리기술 (Clinico-genomic big data management)
2	개인 유전체 기반 맞춤형 치료제 개발	단일세포 유전체 분석기술 (Single cell genomics)
3	재생의료 융합화 및 기술개발 확대	생체 내 직접교차분화기술 (In vivo direct reprogramming)
4	유전체 발현 지휘 단백질 발견	후성유전학적 발생·분화조절기술 (Epigenetic regulation of development)
5	혁신적 진단기술, 액체생체검사 대두	순환 종양세포 DNA 탐지기술 (Circulating tumor DNA detector)
6	표적유전자를 편집하는 유전자 가위기술의 혁신	유전자 편집을 통한 질병치료기술 (Genome editing-based gene/cell therapy)
7	신종 바이러스와 슈퍼 박테리아의 전세계 확산	역학정보 분석기술 (Infoepidemiology)
8	모바일기기로 질병 예측·예방·진단 및 관리	모바일 인공지능 진단기술 (Mobile AI diagnosis)
9	애플·구글 등 스마트 헬스케어 전쟁	웨어러블 건강관리기술 (Wearable health device)
10	원격의료 시범사업 확산	연속식 혈당측정기술 (Continuous glucose monitoring)

[출처: 한국생명공학연구원, 바이오헬스 이슈를 선도하는 10대 미래유망기술 보도자료, 2016]

표 2. 대규모 임상 유전체 정보관리기술의 니즈

	주요 이슈
기술적 측면	2007년부터 본격적으로 도입된 NGS (Next Generation Sequencing) 기술을 이용한 유전체 정보를 얻는 비용과 시간의 획기적인 절감에 따른 대용량 유전체 빅데이터 생산기술이 급속히 발전
경제적 측면	질병의 정확한 진단과 개인에 알맞은 효과적 치료와 치료제의 선택 및 질병을 예측/예방함으로써 건강한 삶 추구하고 의료비 절감에 기여
사회적 측면	고령화 사회로 급속히 접어들고 있으며, 그에 따른 질병치료에 대한 비용 및 치료기간의 지속적 증가에 따라서 유전체 기반의 개인맞춤의료에 대한 요구가 증가

2. 본론

2.1 보건의료 빅데이터 시장 동향

보건의료 빅데이터 시장은 해외의 여러 시장 조사 보고서에서 공통적으로 매우 빠른 성장을 통해 큰 시장을 형성할 것으로 예측하고 있다. Technavio (www.technavio.com)에서 발표한 시장 조사 보고서에서는 2015년부터 2019년까지 연평균 42%의 고성장을 보일 것으로 예측하고 있다. 특히, 미국의 경우 HITECH (Health Information Technology for Economic and Clinical Health) Act와 Affordable Care Act를 위시한 정부 주도의 빅데이터 이니셔티브에 기인한 영향으로 전 세계 시장의 약 63%를 차지할 정도로 미국 중심의 시장이 형성될 것으로 예측하고 있다. 그 외 여러 시장 보고서들의 내용에서도 연평균 17.34~40.5%의 시장 증가를 예상하고 있으며, 2018년에는 100~500억 달러 규모의 시장으로 예측하고 있다. 그리고 대표적인 기업들로는 IBM, SAS, Oracle과 같은 전통적인 IT업체에서부터 Cerner, Allscripts와 같은 EMR 업체들이 언급되고 있다.

산업통상자원부의 보고서에 따르면 보건의료 빅데이터 시장은 2013년에 52억 달러 규모에 달하고 2023년에는 629억 달러에 이를 것으로 예측하고 있으며, 특히 데이터 해석-분석 분야의 성장률이 연평균 37.55%에 달할 것으로 보고 있다[6]. 하지만, 시장 규모는 데이터 저장-통합 부분이 기반 인프라 구축 차원에서 큰 투자가 필요해 당분간 가장 큰 규모를 차지할 것으로 예측하였다. 특히, 국내 시장의 경우 2013년에 857억원 규모였던 것이 2023년에는 세계시장 성장률을 웃돌아 약 5,600억원의 규모로 확대될 전망이라고 한다.

표 3. 보건의료 빅데이터 해외시장 규모

(단위: 억달러)

구분		2013	2018	2023	연평균성장률(%)
서비스	건강관리	525	1,147	1,824	9.38
	예측의료	4.4	6.7	10.5	9.47
	유전자	17	2.8	42	8.41
DB 관리	데이터 저장-통합	27.3	146.4	346.2	21.27
	데이터 해석-분석	1.6	6.9	34.4	37.55
	시각화-빅데이터	23	92.4	248	19.62

표 4. 『포스트게놈 다부처 유전체 사업 추진계획(안)』 추진 경과

일정	내용
'10.10~'10.11	유전체 연구 공동기획 추진을 위한 부처 협의
'10.11~'11.02	국과위 BT 5개 중점분야 투자전략 수립 추진
'10.11~'11.02	국과위 유전체 실무위원회 구성 및 운영
'11.02	국과위 유전체 정보·분석 연구 투자전략 확정
'11.05~'11.11	다부처 유전체사업 공동기획
'12.03.09	2012년 상반기 예비타당성조사 대상 선정
'12.03~'12.11	KISTEP 예비타당성조사 실시
'13.10	범부처 협의체 구성·운영계획 수립
'13.12	2014년 1차년도 사업예산 확정
'14.02.20	2014년도 부처 공동 사업설명회 실시
'14.03	포스트게놈 다부처 유전체사업 추진계획(안) 수립

[출처: (2014~2021) 포스트게놈 다부처 유전체사업 추진계획(안), 2014]

우리나라는 2019년에 65세 이상의 노령인구가 전체인구의 14.4%, 2026년에는 20.0%에 이르는 초고령화 사회로 진입할 것으로 예측되는 등, 세계에서 보기 드문 고령화 속도를 보이고 있어 보건의료 관련 산업의 성장 잠재력이 매우 높으며, 이를 해결하기 위해 정부에서도 국가정책으로 보건의료 빅데이터에 대한 지원을 위해 노력을 하고 있다. 특히, 2014년부터 포스트게놈 다부처 유전체사업 추진계획을 발표하고, 창조경제의 핵심분야인 유전체 사업의 연구 역량 확대 및 조기 투자 성과 창출을 위해 부처별 연계·협력을 강화하려는 노력을 아끼지 않고 있다.

선진국을 포함한 대부분의 국가에서 의료정보 빅데이터 활용에서 공통적인 면은 공익적 목적으로 빅데이터의 활용을 제한함으로써 위험을 감소시키는 한편, 민감한 개인정보 보호와 데이터 보안을 강화하는 전략을 통해 국민의 신뢰와 참여를 유도하고 있다는 것이다.



그림 1. 통합 건강관리 시스템(Integrated Healthcare) 모식도
 [출처: MaRS Market Intelligence 2014]

· 미국

미국은 65세 이상 노인에 대한 국가건강보험 프로그램인 메디케어(Medicare; <http://www.medicare.gov/>)와 저소득층에 대한 주정부 관장 의료보호제도인 메디케이드(Medicaid; <http://www.medicaid.gov/>)를 제외하고는 대부분을 민간의료시장에 의존하고 있다는 특성을 가지고 있다[3]. 하지만, 미국의 초기 보건의료 빅데이터 사업은 보건부의 NIH (National Health Institute)와 AHRQ (Agency for Healthcare Research and Quality)를 통한 비교효과성 연구(Comparative Effectiveness Research, CER) 사업으로 시작되었고, 의학과 의료 제공에서 혁신을 유도하도록 근거 생산 연구를 확대를 통한 환자중심통합의료(Patient-Centered Integrated Care) 기반을 구축하는 국가적인 전략을 추진하고 있다. 미국은 민간 보험사들을 중심으로 다양한 정보 통합 플랫폼이 활성화되도록 간접적인 조정자의 역할로서 보건부 산하에 ONC (Office of the National Coordinator)를 설치하여 데이터 상호운용성(Interoperability)을 높이기 위한 정부 제공과 분류 표준화 등을 담당하게 하고 있다. 또한, 의료질향상과 비용 절감을 통한 시스템 효율화를 목적으로 공적 시스템에서 다양한 정보활용사업(환자 중심 의료연계 실적을 포함한 의료질 평가)을 추진하고 있다.

· 영국

영국은 전 국민에게 무료로 의료서비스를 제공하는 조세 기반 국가보건의료서비스(National Health Service, NHS)를 운영하고 있으며, 보건의료 빅데이터 활용에 대해서는 미국과 동일하게 환자중심통합의료 기반을 구축하는 국가적인 전략을 추진하고 있다[5]. 특히, 영국 보건부는 2013년에 Personalized Health and Care 2020을 발표하여 환자와 시민의 의료 및 복지 정보에 대한 통제권을 강화시키고 분산된 사회보장 데이터들을 수집, 저장, 연계, 분석하는 독립조직인 HSCIC (Health and Social Care Information Center)를 설립한 바 있다. 또한, NHS에서 더 나은 의료서비스를 제공하기 위해 의료정보의 활용(The Care Data program을 통한 의사 간 환자 진료정보 공유)을 목적으로 국민의 참여와 선택을 유도하고 있다.

· 일본

일본은 전 세계적인 빅데이터 활용 움직임에 일단 보조를 적극적으로 맞추고 있다. 2015년 1월 “의료 빅데이터 시대의 개막”이라는 보고서를 의학전문 출판사인 의학서원(醫學書院)에서 출간하였는데, 일본에서는 연간 약 900만건의 데이터가 축적되는 진단군 분류 자료와 진료보수 명세서 자료를 중심으로 연구자가 구축한 증례자료와 생명공학 DB 등 포괄적인 의료정보 빅데이터에 대한 정비 작업을 추진 중이다. 이의 일환으로 진료보수 명세서 및 특정 건강검진 결과에 대한 국가 DB 구축을 시작으로 각각 독립적으로 구축해 온 DB를 연계해 개인 건강 상태의 인과관계를 규명하고 개개인에 대한 의료서비스의 개선은 물론, 의료정책의 수립에 이르기까지 통합 시스템을 활용하도록 정부 차원에서 지원하고 있다.

2.2 보건의료 개인정보 관련 법체계

2.2.1 미국

· 일반사항

미국의 경우 의료정보의 수집 및 사용에 균일하게 적용되는 개인정보보호 원칙은 따로 존재하지 않으며 기본적으로 연방법 차원의 HIPAA (Health Insurance Portability and Accountability Act)의 Privacy rule에서 주로 규정하고 있고, GINA (The Genetic Information Nondisclosure Act of 2008), 42 CFR Part 2를 비롯하여 기타 각 주별로 법률체계에서 관리하고 있다[7].

· 관련 법체계

미국은 개인정보보호에 관하여 각 산업별로 규정하고 있는데, 1966 HIPAA의 제정을 통해

의료영역에서 개인정보보호를 규율하게 되었다. HIPAA 제정 당시의 주목적은 건강보험의 이동성과 책임을 규정하는 법으로 피고용자가 퇴직이나 이직을 하는 경우에도 의료보험의 보장성을 유지하기 위한 것이었다. HIPAA Privacy rule은 개인건강정보의 이용 및 공개 방법, 범위 등에 관한 법규 명령으로 2002년 8월 14일 제정하여 2003년 4월 14일부터 발효되었다. 또한, HIPAA Security rule은 전자개인의료정보에 한해 규정하고 있다. 일반적으로 HIPAA에서 개인의료정보의 이용과 공개를 위해서는 아래와 같은 경우를 제외하고는 개인의 서면 동의가 원칙이나 진료 및 의료행위에 필수적으로 수반되는 진료, 지불 또는 기관운영을 위해 필요한 경우에는 예외로 한다고 규정되어 있다.

<서면동의 없이 이용 및 공개가 가능한 경우, HIPAA>

- 법에 의해 요구되는 경우
- 공중보건을 위한 활동
- 중독, 방임 또는 가정폭력의 신고
- 보건 감독 활동
- 사법절차, 행정절차
- 법 집행 목적의 활동(범죄수사 등)
- 사망자에 관한 정보를 검사관, 의사 등에게 제공하는 경우
- 장기, 안구, 조직 기증의 목적
- 연구목적으로 기관심의위원회(Institutional Review Board; IRB) 또는 Private Board의 승인을 받아 PHI (Protected Health Information) 이용 및 제공 가능, 연구 준비를 위해 필요한 경우 해당기관 내부에서 이용 가능, 사망자에 대한 연구가 목적인 경우 등의 일정한 유형의 연구 활동
- 공중의 건강 안전에 대한 중대한 위해(危害)를 방지하기 위한 경우
- 군사, 교정 활동 등 특별한 정부의 기능을 위한 경우
- 산업재해보상청구를 위한 경우

HIPAA Privacy rule(연방프라이버시보호법)은 행정기관이 보유하는 개인에 관한 정보를 대상으로 개인정보에 대한 정의를 하고 있기 때문에 식별가능성에 대한 사항보다는 식별된 개인에 관한 정보를 개인정보로 인정한다. 반면, 미국 아동의 온라인 프라이버시보호법과 그 시행령은 개인을 식별할 수 있는 정보라는 점을 전제하면서 개인정보의 범위에 포함되는 정보를 열거하고 있다. 이는 EU나 영국 등에서 개인정보에 대한 개념을 추상적으로 접근한 것에 비하여 상당히 구체적인 방법으로 접근하고 있다는 것을 알 수 있다. 다만, 아동에 대한 보호규정이기 때문에 다른 법률 규정에 비하여 좀 더 유연하게 이루어지고 있는 것으로도 볼 수 있다. 캐나다의 연방프라이버시보호법과 개인정보보호 및 전자문서법에서도 '식별 가능한' 또는 '신원을 확인할 수 있는' 이라는 용어를 사용하면서 개인정보의 핵심으로써 식별성을 필요로 한다는 점을 명시하고 있다. 이러한 기준은 우리의 개인정보보호법상 개인정보의 식별성에 관한 사항과 비교할 때, 우리의 경우는 좀 더 상세한 기준이 필요하다는 점을 시사하고 있다고 볼 수 있다.

또한, GINA는 일반적으로 보험 관련 의사결정을 내릴 때 수혜자의 유전정보를 제공하는 것을 금지하고 있으며, 일반적으로 고용주가 유전정보를 바탕으로 한 차별과 고용 결정에 유전정보를 사용하는 것을 금지하고 있다. 42 CFR Part 2는 연방정부가 약물남용 예방프로그램, 메디케어(Medicare)에 참여하는 제공자, 미국 마약단속 관리 번호를 가지고 있는 자에게 적용되며, 제공자는 정보공개를 위하여 환자의 사망원인, 아동학대 보고, 특정 연구를 위한 국가 위임 요구에 따른 정보공개를 예외로 하고 반드시 환자의 동의를 얻어야 한다.

또한, 최근 하원을 통과한 21st Century Cure Act (2015년 7월)는 생물의학 연구에 대한 현재의 규제에 대한 개혁과 매년 FDA 및 NIH에 추가적인 세출을 제공하는 법률안으로써 보건부장관에게 Patient Experience Data를 활용하여 Risk-Benefit Analysis를 도입하여 FDA 규제 프레임워크를 개선에 활용하고 있다. 특히, 최근 미국 내 빅데이터 연관산업의 발전에 따라서 데이터 거래산업도 성장하고 있는 추세로 미연방통상위원회(Federal Trade Commission, FTC)는 제품의 디자인 단계별 개인정보보호 체계를 마련할 것을 촉구하고 있다.

2.2.2 유럽연합

· EU의 개인정보보호법

EU의 경우 그 동안 개인정보보호에 관한 기본법적 역할을 했던 1995년 개인정보보호지침(Directive)을 폐지하고 새로운 법률인 개인정보보호규정(Regulation)을 제정하는 거의 마지막 단계에 와 있다. 개인정보의 범위와 관련한 두 법률은 기본적으로 접근방식은 동일하지만 2014년 개인정보보호규정에서 좀 더 상세한 접근을 하고 있다. 구체적으로 보면 원칙적으로 식별된 또는 식별 가능한 자연인과 관련된 여하한 정보는 모두 개인정보에 해당한다. 여기에서 식별 가능한 자연인은 직접적 또는 간접적으로 식별될 수 있는 사람이다. 그런데 특별히 본 규정에서는 식별성과 관련해서 이름, 식별번호, 위치정보, 온라인 식별자 또는 그 사람의 하나 또는 이 이상의 요소로서 특정한 신체적, 생리적, 유전적, 정신적, 경제적, 문화적, 사회적 정체성에 의하여 식별할 수 있는 것이라고 규정한다. 이와 마찬가지로 영국의 경우에는 그 내용에 있어 표현방식의 차이가 있을 뿐, 거의 유사한 내용을 담고 있으며 독일의 연방개인정보보호법도 마찬가지이다. 식별성과 관련해서 EU 개인정보보호지침(제26조) 및 개인정보보호규정(제23조)은 모두 식별성의 해석에 있어 합리적으로 판단해야 한다는 점을 강조한다.

· 영국의 일반사항

영국의 의료서비스 제공자는 환자 정보와 관련하여 보통법(Common Law) 상의 비밀유지 의무를 가지고 있기 때문에 환자 정보를 제3자와 공유하기 위해서는 해당 환자의 동의를 받아야 하는 것이 원칙이다. 보건의료 정보에 대한 접근과 이용에 관해서는 보통법 상의 비밀유지 의무 외에도 Data Protection Act 1998, Health and Social care Act 2012, Care Act 2014에 의해 규정하고 있다.

· 영국의 관련 법체계

보통법 상의 비밀유지의무에 의하면 환자 관련 정보는 원칙적으로 제3자에게 공개될 수 없으나, 환자가 자신의 정보가 어떻게 사용될 지에 대해 알고 자발적으로 정보공개에 동의를 한 경우(환자 의사에 따라서 동의의 내용 및 범위 변경 가능), 정보공개를 통해 환자 개인, 타인, 혹은 공공의 이익을 위해 필요할 경우, 정보를 공개해야 할 법률상의 의무가 있을 경우에는 개인 정보의 공개가 가능하다고 명시하고 있다. Data Protection Act 1998은 1995년에 제정된 EU의 개인정보 보호지침을 국내법으로 수용한 법률로 개인정보의 처리에 관한 일반적인 규율을 제공하고, 개인정보 처리자가 준수하여야 할 8가지 원칙을 규정하고 있다[8]. 또한, 생존하는 개인을 식별할 수 있는 정보의 처리에만 적용되며, 비식별 처리된 정보는 대상에서 제외한다. 특이한 사항으로 연구 목적에 대해서는 보유기간에 제한이 없이 예외를 인정하며, 환자의 동의를 얻은 바 없어도 기존 의무기록 등 환자의 식별 가능 정보를 이용하여 역학적 연구를 하는 것이 가능하도록 정하고 있다. Health and Social Care Act 2012는 영국의 국립보건서비스(NHS)를 개혁하기 위해 제정되었으며, 보건 의료 데이터의 연계 및 활용을 위한 독립된 권한을 가진 별도 기구로서 HSCIC(현재 NHS Digital)의 설립 근거 제공하고 있다. HSCIC는 영국 내 일반의(General Practitioner)로부터 환자의 진료기록 등의 정보를 제공받아 저장, 분석, 가공하며 이를 제3자 및 일반에 제공 및 공개하는 역할을 담당한다. 다만, 이들 정보의 제공 및 공개는 반드시 국민의 건강수준 향상, 질병예방, 보건 의료체계의 혁신을 위한 범위 내여야만 한다. Care Act 2014는 Health and Social Care Act 2012 제정 이후, 개인정보 보호가 부족하다는 비판에 따라서 HSCIC가 제3자에게 보유 정보를 제공하는 목적을 '의료·복지 서비스의 증진, 혹은 국민 건강 증진'으로 제한하였으며, 연구, 신약개발 또는 치료방법 개발 이외의 상업적인 보험회사들이 환자의 개인정보를 열람하는 것이 불가능하게 하였다.

또한, 기타 관계 법령으로 Human Right Act 1998, Health Service (Control of Patient Information) Regulations 2002, National Health Service Act 2006, Health Protection (Notification) Regulations 2010 등이 있다.

2.2.3 일본

· 일반사항

일본에서는 개인정보보호에 관한 법률로서 「행정기관이 보유하는 개인정보의 보호에 관한 법률」, 「독립행정법인 등이 보유하는 개인정보의 보호에 관한 법률」, 「정보공개·개인정보보호심사회설치법」 및 「행정기관이 보유하는 개인정보의 보호에 관한 법률 등의 시행에 따른 관계 법률의 정비 등에 관한 법률」이 2003년 5월 30일에 공포되어, 2005년 4월 1일부터 시행되고 있으며, 동시에 실시된 「개인정보의 보호에 관한 법률(이하 '개인정보 보호법)」과 마찬가지로 IT 사회에서 개인정보의 이용확대에 대응하기 위하여 제정되었다. 다만, 공적 부문 중 지방공공단체가 보유하는 개인정보에 관하여는 각 지방공공단체가 제정하는 조례에 의해 규율된다. 일본의 경우 현재 각 부처에서 분야별로 개인정보보호 가이드라인을 규정하였지만, 법률 개정으로

인해 총리실 직속으로 해당 업무가 이관될 예정이다. 특히, 의료분야는 별도로 관리번호 CRP를 도입하여 게놈 등 연구 활성화를 위해 특별법 도입을 논의 중에 있다.

· 관련 법체계

개인정보보호법이 2005년 제정된 이후 개인의 권익보호 및 공익 및 공공성의 관점에서 정보공개 양립 의견이 개선됨에 따라 2015년 9월 3일 개정안이 중의원 본회의에서 가결되었으며, 2016년 1월 1일부터 시행되었다[9]. 또한, 그 동안 분야별로 각 부처에서 관할하던 업무를 총리실 산하에 개인정보보호회를 두어 개인정보보호에 대한 업무를 총괄하도록 하였다. 개정안에는 “행정절차에 있어 특정 개인을 식별하는 번호의 이용 등에 관한 법률(2013년)”에 따라서 한국의 주민등록번호와 같은 체계가 도입됨과 빅데이터 분석의 발전에 따라서 개인정보에 ‘개인식별부호’를 추가하였고, 의료정보(병력)는 ‘필수 배려 개인정보’로 관리하도록 규정하며 구체적인 내용은 정령으로 정하도록 하였다.

후생노동성은 의료서비스의 품질 향상 및 의학 연구 발전을 추진하기 위한 정보 연계를 위해 기반이 필요할 것으로 판단하여 개인의 권리 이익의 추가 도모를 의료분야의 개별법을 추진하였다[10]. 최근 법 개정이 진행되었으나 게놈 연구 등의 확대를 위해서 현행 개인정보보호법 체계에서의 개인식별부호, 필수 배려 개인정보 등에 대한 문제점은 계속해서 제기되고 있는 실정이다.

<의료분야 주요 검토 내용>

- 환자 등의 권리와 의료 등의 서비스 제공측의 의무
- 환자 등의 자기 정보 통제 권한의 확보
- 의료 등의 서비스 제공측의 면책
- 정보 검색 및 활용
- 법의 평가 적용범위와 이행의 확보

2.3 개인정보 비식별화 조치

개인정보 비식별화(de-identification)는 데이터 값 삭제, 가명처리, 총계처리, 범주화, 데이터 마스킹 등을 통해 개인정보의 일부 또는 전부를 삭제하거나 대체함으로써 다른 정보와 결합하여도 특정 개인을 식별할 수 없도록 하는 조치이다[11]. 데이터의 활용 과정에서 개인정보 노출 및 개인 프라이버시의 침해 우려를 최소화하기 위한 방책으로 활용되며, 최근 빅데이터로 대표되는 ICT 분야의 환경 변화와 그에 따른 개인정보보호의 논란 속에서 주목 받고 있다. 국제표준기구에서도 개인정보 비식별화 표준 제정에 적극 나서는 등 공공과 산업영역에서 데이터 활용을 통해 막대한 부가가치 창출을 지원하는 동시에 개인정보보호의 법률 원칙과 시민의 프라이버시 권리를 지킬 수 있는 제도적 장치의 마련에 대한 필요성이 커지고 있다.

2.3.1 해외의 비식별화 법제 및 가이드라인 현황

· 미국

앞서 언급했듯이 미국에서는 개인정보보호 법제가 개별 분야나 사안에 따라서 적용되기 때문에 비식별화 관련 규제나 지침 역시 분야별로 다뤄진다. 미국에서 개인정보 비식별화 원칙은 온라인 개인정보 취급과 관련한 정보주체의 권리를 확립하기 위해 2012년 오바마 행정부가 제안한 '소비자 프라이버시 권리장전(Consumer Privacy Bill of Rights, CPBR)'의 개정을 통해 명확하게 정의될 전망이다[12].

비식별화 관련 지침으로는 미연방거래위원회(FTC)가 2012년 3월 발표한 '개인정보의 비식별화 가이드라인'이 있으며, 이는 특정한 소비자, 컴퓨터 및 기타 개인을 식별할 수 있는 장치들과 연관될 수 있는 것은 어떤 정보라도 보호의 대상이 되어야 한다고 규정하고 있다. FTC가 제시하는 비식별화 관련 3가지 요구사항은 다음과 같다. 첫째, 개인, 컴퓨터 및 디바이스에 대한 정보를 추론(식별)할 수 있는 데이터의 삭제, 수정, 노이즈 추가, 통계적으로 샘플링, 총계처리 등의 적절한 방법을 자율적으로 판단하여 반드시 비식별 조치를 취한다. 둘째, 데이터를 공개할 경우 해당 데이터를 비식별화하여 이용한다는 것과 향후 비식별화 상태를 유지하여야 한다는 것, 그리고 재식별하지 않을 것을 개인정보주체에게 공개적으로 약속한다. 셋째, 데이터를 제3자에게 제공할 때에는 해당 데이터를 제공받은 어떠한 이용자도 재식별하지 않도록 계약조건으로 반드시 요구해야 한다. FTC에서는 업계가 위의 3가지 조건을 이행한 경우에만 특정 데이터가 식별 가능하지 않도록 처리된 것으로 판단한다.

· 유럽연합

EU는 2015년 12월에 합의된 일반정보보호규정(General Data Protection Regulation, GDPR)을 통해 '가명화'를 통한 개인정보 비식별화 조치를 조건으로 데이터를 수집 및 이용할 수 있도록 했다[10]. 이 때 '가명화'는 '특정 개인에게 연결되지 않을 것'을 최우선의 조건으로 하며, 재식별화가 가능한 만큼 개인정보로 취급되지만 기술적, 관리적 방법을 통해 특정 목적을 위해 활용할 수 있도록 허용한다. 특히, 암호화를 통해 데이터를 인식 불가능하게 한 경우에는 데이터 유출 시 정보주체에게 통지하는 의무를 면제받도록 하고 있다. 과학, 역사, 통계 목표의 이용을 위해서는 2083조를 통해 가명화한 후 정보를 이용하도록 허용하고 있다. 또한, GDPR 제20조에서는 프로파일링에 대한 거부권을 두고 있어 개인적 특성에 대한 평가와 개인에 관한 분석 또는 예측을 위한 기준의 적용을 받지 않도록 권리를 보장하고 있는데 이는 온라인 상에서 개인정보보호를 더욱 강화하는 근거가 된다.

구체적인 비식별화 지침은 회원국별 가이드라인을 통해 제시하고 있는데, 영국에서는 정보보호기관인 ICO (Information Commissioner's Office)가 2012년 개인정보 식별화 행동강령(Anonymization: managing data protection risk code of practice, 2012)을 발표했다. ICO 강령은 법적 강제력이 없고 비식별화 해야 하는 개인정보의 범위 역시 개인 식별 가능성이

최소화되는 수준으로 관련 업계가 자율적으로 판단하여 처리하도록 하고 있다는 특징을 보인다.

· 일본

일본은 2015년 비식별화 데이터를 '익명가공정보'로 정의하고, 이를 처리 및 이용할 수 있도록 한 개인정보보호법을 개정했다[10, 11]. 이는 빅데이터를 기반으로 한 다양한 신규 비즈니스 기회를 창출하고 경제재생계획과 연결시켜 이를 실현하기 위해 기존의 빅데이터 관련 규칙을 강화한 것이 특징이다. 일본의 개인정보보호법의 경우 익명화의 개념을 적극적으로 도입해 재식별화가 불가능하도록 개인을 특정할 수 있는 요인들을 희석한 데이터에 대해서는 정보주체의 동의를 받거나, 제3자에게 제공 시 제약을 두지 않고 있다.

2.4 보건의료 빅데이터와 인공지능과 결합

2.4.1 미국

미국은 인공지능 R&D 정책을 범정부 차원에서 추진하고 있으며, 인공지능을 활용한 정밀의료 추진을 통해 의료의 질적 수준 제고에 집중하고 있다[13]. 특히, 2013년 "Brain Initiative"를 통해 백악관은 범정부 차원에서 인간의 뇌를 중심으로 체계적인 인공지능 기술개발을 통해 원천기술 확보 및 실제 인공지능을 활용하기 위한 상용화 연구의 적극적 추진을 발표하여 향후 10년간 30억 달러 규모의 투자가 이뤄질 예정이며, 2015년에는 연구교서에서 2016년 2.2억 달러(약 2,600억원) 투자를 골자로 하는 '정밀의료추진계획(Precision Medicine Initiative, PMI)'을 발표한 바 있다. 빅데이터 활용 바이오헬스 기술의 발전단계별 국가전략을 수립하여 'American Recovery and Reinvestment Act of 2009'의 HITECH Act에 IT와 헬스케어의 융합을 지원하는 것을 포함시켰으며, 국립보건원은 Pillbox (National Library of Medicine)를 통해 주요 질병의 분포, 연도별 현황 등 세부 분석을 시행하는 등 유전체 구조 및 생물학 연구, 질병 생리학 연구 등 바이오헬스 연구에 빅데이터를 활용함으로써 향후 인공지능 헬스케어 연구와의 시너지 창출을 기대하고 있다.

2.4.2 유럽연합

EU는 2013년부터 2023년까지 10년 동안 1.8조원을 투자하여 HBP (Human Brain Project)라는 인간 두뇌의 인진 형태를 기반으로 한 지식 처리용 인공지능 기술을 개발 예정으로 6개 ICT (Information and Communications Technology) 중에서 "의료정보 플랫폼" 기술을 포함하여 유전체 분석을 기반으로 정밀의료 제공을 추진 중에 있다[13]. 또한, 영국은 암, 감염질환 치료를 목적으로 2014년부터 2017년까지 10만명 유전체 분석 프로젝트에 약 5천억원(3억 파운드)을 투자하는 정밀의료 연구개발 및 실용화를 위한 연구개발에 투자를 강화 중이다.

2.4.3 일본

일본의료연구개발기구(Japan Agency for Medical Research and Development, AMED)는 2015년 유전체 정보를 활용해 환자의 특성에 따른 최적의 맞춤형 치료 서비스를 제공하는 게놈의료의 실용화를 목표로 정부에 약 93억엔(한화 950억원)을 요구하였으며, 일본 총무성은 2045년에는 인공지능이 인간의 능력을 뛰어넘을 것이라는 유엔 보고서를 기반으로 2015년 2월 정보통신, 뇌과학 분야 전문가 12명이 참여하는 인공지능의 연구 개발 강화 방안을 모색하기 위한 “가칭 2045 연구회(인공지능화가 가속화되는 ICT 미래상에 대한 연구회)”를 발족한 바 있다. 또한, 일본 재부흥전략의 일환으로 로봇혁명의 실현을 위해 3대 핵심전략(① ‘로봇창출력’의 근본적 강화를 통한 세계 로봇혁신 거점화, ② 세계 최고의 로봇 활용사회를 목표로 ‘로봇 활용 및 보급’, ③ 빅데이터, IT 융합, 네트워크, 인공지능을 구사하는 로봇으로 세계를 주도하는 ‘로봇혁명 전개·발전’)을 선정하여 2020년까지 5년간 1,000억엔을 투자하기로 하였다[13].

2.5 개인 유전체 기반 맞춤형 의료서비스

개인맞춤의료(Personalized medicine)이란 유전정보인 DNA 분석기술의 발전, 유전정보와 질병의 발병 사이의 관계에 대한 생물학 및 의학 지식이 축적됨에 따라서 개인의 고유한 유전적 특징을 바탕으로 질병의 예방, 발견, 진단, 치료, 건강관리 등에 활용하고자 하는 의학적 노력 속에서 탄생한 개념이다[14]. 기존의 평균적인 환자에게 맞춰진 방법(One-Size-Fits-All)인 반면 맞춤의료는 의료적 처치를 각 환자의 개인별 특성에 맞추는 것을 의미하며 예방과 치료에 집중하는 것 이외에도 부작용과 비용의 감소에 크게 기여할 것으로 기대된다. 특히, 개인 유전체 분석 기반 맞춤의료 서비스를 단순화하면 개인 DNA 시퀀싱부터 유전체 정보를 의료에 적용하는 모든 과정을 포함한다.

<맞춤형 의료서비스의 적용 분야>

- 위험평가(Risk Assessment): 유전적 소인 등을 토대로 질병 발병의 위험도를 예측
- 예방(Prevention): 질병 발병 위험도 등을 토대로 질병 예방을 위한 생활 방식 등 제안
- 발견(Detection): 유전체 분석 등을 토대로 분자 수준에서 질병의 조기 발견
- 진단(Diagnosis): 유전자 이상/변이 등의 분석 결과를 토대로 질병에 대한 정확한 진단
- 치료(Treatment): 유전적 특징에 따른 진단 및 약물감수성 파악으로 효과적 치료
- 관리(Management): 질병 치료에 대한 반응 및 질병의 예후를 모니터링

2.6 의료정보 표준화 동향

국제표준화기구(Organization of International Standardization, ISO)는 전자건강기록(Electronic Health Records, EHR)을 ‘환자의 건강과 관련된 정보를 컴퓨터가 이해할 수 있는 형식으로 저장한

것'이라고 정의하고, EHR에 대하여 의료기관 수준의 의무기록, 공유 가능한 의무기록, 소비자 의무기록, 연구와 질병 등록 등과 관련된 다양한 종류로 공공의무기록의 관점에서 가이드라인을 제시하고 있다[15, 16]. 의료정보 분야의 국제 표준을 제정하는 ISO/TC 215 (Health Information)에서는 현재까지 148개의 표준이 제정되었고, 120여개의 신규 표준이 진행 중에 있다. 스마트헬스 데이터는 WG1(아키텍처, 프레임워크 및 모델), WG2(시스템 및 기기 상호운용성), WG3(시멘틱 콘텐츠), WG4(보안 및 정보보호) 등 대부분의 워킹그룹과 관련이 있다.

ISO 표준안 이외에 IHTSDO (International Health Terminology Standards Development Organization)의 SNOMED-CT (Systematized Nomenclature of Medicine–Clinical terminology)와 Regenstrief Institute의 LOINC (Laboratory Observation Identifiers Names and Codes)도 의료정보 관련 용어 표준화를 위한 중요한 국제 표준들이며, 의료 데이터의 전달 및 통합을 위한 HL7 (Health Level 7)의 CDA (Clinical Document Architecture) 표준을 비롯하여 임상 의사결정지원시스템에 대한 다양한 표준들이 개발되었다. 특히, 임상시험 데이터 표준과 관련해서는 CDISC (Clinical Data Interchange Standards Consortium)이 별도로 존재하여 다양한 국제 표준안을 개발하면서 미국 식약청에서 실제로 활발하게 사용하고 있는 중이다. CDISC에서는 open data 표준으로 CDASH (clinical data Acquisition Standards Harmonization), LAB (Laboratory Data Model), SDTM (Study Data Tabulation Model), SEND (Standard for Exchange of Nonclinical Data), Controlled Terminology 등 다양한 표준안을 제정하였다[17].

3. 결론

최근 대부분의 미래학자들은 빅데이터와 사물인터넷(IoT)을 응용한 신규산업을 누가 주도하느냐에 따라 세계 경제 주도권이 달라질 것이라고 한다. '빅데이터'라는 거대한 광산에서 '유용한 정보'라는 귀금속을 캐내는 것은 매우 매력적이지만 그곳에는 무수히 많은 '개인정보'들이 포함되어 있다. 우리 현행법의 경우 과거 개인정보보호의 가치를 인격적 가치로만 이해하던 시대의 법이론을 그대로 답습하고 있어 '동의만능주의'에 빠져있다. '과학기술'이 아닌 '제도' 때문에 우리나라가 주도권 경쟁에서 뒤쳐진다는 것은 매우 불행한 일이다. 지금은 보다 선제적으로 제도개선을 위한 노력이 필요한 시점이다.

민감한 개인정보 처리에 관해서는 우리나라를 포함하여 국외의 규정이 모두 처리금지 원칙을 세우고 있다. 반면 예외에 있어서는 EU와 영국, 독일의 개인정보보호법은 매우 상세한 열거식의 규정을 운영하고 있다. 특히, 개인정보 중에서 민감 정보인지의 여부가 불분명한 경우가 존재하기 때문에 그에 대한 좀 더 상세한 기준을 제시한다는 측면에서 우리의 현행 규정에 시사하는 점이 있다고 보여진다. 이 외에 EU의 개인정보보호규정과 영국 및 독일의 개인정보보호법에서는 자동화된 개인의 의사결정에 관한 제한을 규정하고 있는데 이는 빅데이터에서의 개인정보처리와 관련하여 참고할 수 있는 규정으로 보여진다.

의료정보의 관리 효율화 및 활용을 높이기 위해서는 의료정보의 표준화가 필수적이다. 하지만 국내의 경우 의료정보 관련 표준화 작업은 외국에 비해서 활발하지 않으며, 국내 의료기관들은 자체적으로 의료정보 시스템을 구축을 진행해 왔기 때문에 국제 표준에 대한 인식 및

요구도가 매우 낮다는 문제점도 앞으로 개선해야 할 당면 과제이다. 보건의로 빅데이터를 보다 효율적으로 활용하기 위해서는 LOD (Linked Open Data) 기반 글로벌 보건의로 빅데이터 연계 플랫폼이 구축되어야 한다는 지적이 있다[3, 18]. 각기 다른 기관 간 정보 연계 및 공유가 가장 중요하며, 개인정보보호법 등 각종 규제와 관행으로 인해 공개 활용이 어렵기 때문에 웹 상의 정보를 공유/연계/재이용해 가치를 창출할 수 있는 LOD 기반이 중요하다는 뜻이다. 또한, 앞으로 보건의로 데이터 관련 기술을 개발할 때에는 기존의 개발된 표준안들을 반드시 참고로 하여 기술 개발이 이루어져야 하며, 필요에 따라서는 기존 표준안을 개선하는 것이 중요하다.

보건의로 빅데이터 관리는 의료 및 건강 정보를 대상으로 최신의 ICT 기술을 활용하여 보건의로의 질을 향상시키고 새로운 산업을 활성화하고자 하는 분야이다. 즉, 이를 위해서 전통적인 의료분야의 연구성과는 물론 유전체 분야, 그리고 딥러닝으로 대표되는 기계학습의 연구성과를 융합하기 위해서 다양한 분야의 연구자들이 협력하도록 유도하는 것이 필요하며, 미국, 영국과 같이 정부의 주도하에 정부 민간기업, 연구소의 지속적이고 체계적인 협업이 반드시 필요하다.

4. 참고문헌

- [1] 한국생명공학연구원 보도자료 “바이오헬스 이슈를 선도하는 10대 미래유망기술 발표”, 2016.12.
- [2] McKinsey&Company, “The ‘big data’ revolution in health care”, 2013.01.
- [3] 강희정, “보건의로 빅데이터의 정책현황 및 과제”, 보건복지포럼, 2016.08.
- [4] 이인재, “국내외 보건의로 빅데이터 현황 및 과제”, 정보통신기술진흥센터, 2014.07.
- [5] 영국 보건의료서비스 홈페이지 <http://content.digital.nhs.uk/>
- [6] 산업통상자원부, 개인맞춤형 건강관리 시스템 기획보고서, 2014.02
- [7] Foley and Lardner “Tapping Into the Big Value of Health Care Big Data Top Legal and Regulatory Considerations on the Path to Monetization”, 2015.
- [8] 최계영, “의료 분야에서의 개인정보 보호” 개인정보와 빅데이터, 2015.11.
- [9] Kubota, ‘사적인 정보의 이용 및 활용과 개인정보보호법의 개정’, 내각위원회조사실, 2015.01.
- [10] 후생노동성, “제1차 게놈정보를 이용한 의료 등의 실용화 추진 태스크포스” 회의록, 2015.11.
- [11] 한국인터넷진흥원, 개인정보보호 핫이슈 심층분석 보고서, 2016.05.
- [12] Gartner, Forecast Analysis: Internet of Things, Endpoints and Associated service, 2014.
- [13] 보건산업통계시스템, 보건산업브리프, vol. 208, 2015.12.
- [14] 과학기술정책연구원, “개인 유전체 기반 맞춤의료 현황과 발전과제”, STEPI INSIGHT, 제179호, 2015.12.
- [15] 국가표준코디네이터, 스마트헬스데이터 동향 및 시사점, 2016.04.
- [16] 미래창조과학부, 보도자료, “빅데이터, 산업과 생활을 바꾼다”, 2016.03.
- [17] 국제임상데이터표준화컨소시엄 홈페이지 <http://cdisc.org/>
- [18] 한국정보화진흥원, “알기쉬운 Linked Open Data”, 2015.04.

The views and opinions expressed by its writers do not necessarily reflect those of the Biological Research Information Center.

장성재(2017). 보건의료 빅데이터 관리시스템 최신 동향. BRIC View 2017-T07
Available from <http://www.ibric.org/myboard/read.php?Board=report&id=2698> (Mar 21, 2017)

Email: member@ibric.org

※ 본 콘텐츠는 **invitrogen** **applied biosystems** 의 후원으로 작성되었습니다.