
의료 빅데이터의 활용과 해결과제

우성희

한국교통대학교

Applications and Issues of Medical Big Data

SungHee Woo

Korea National University Of Transportation

E-mail : shwoo@ut.ac.kr

요 약

빅데이터는 디지털 환경에서 생성되는 모든 데이터를 의미하며 규모가 방대하고 생성주기가 짧고 다양한 형태를 가지는 특성이 있다. 스마트폰과 인터넷이 대중화되면서 사용자들이 남기는 데이터의 양과 종류는 점점 더 큰 규모로 생성되고 있으며 생성된 빅 데이터로 부터 사용가치가 있는 정보만을 추출하여 활용하는 시기로 전환되고 있다. 빅데이터는 또한 의료 산업이나 보건 분야에도 응용될 수 있으며 IoT, 스마트 헬스케어 등의 기술과 함께 융합되어 시너지 효과를 창출하고 있다. 그러나 방대한 데이터를 의미있고 안전하게 활용하기 위해서는 정보보호 등의 선행과제가 존재한다. 따라서 본 연구에서는 의료 빅데이터 활용사례와 기대효과, 해결과제, 마지막으로 의료 빅데이터의 미래전망을 분석한다.

ABSTRACT

Big data is all data generated in the digital environment which has a variety of large and a short life cycle. The amount and type of data are becoming more and more produced on a larger scale, as a smart phone and the internet are popular, and consequently it has been converted into time for users to take advantage and extract only the valuable and useful data from the generated big data. Big data can also be applied to the medical industry and health sectors. It has created the synergy to be fused with ICT such as IoT, smart healthcare, and so on. However, there will be challenges like data security in order securely to use a meaningful and useful vast amounts of data. In this study, we analyze the future prospects of the healthcare, applications and issues of medical big data, and the expected challenges.

키워드

Medical Big Data, Medical Information, Healthcare, ICT

1. 서 론

최근 몇 년간 IT시장의 화두는 빅데이터이다. 빅데이터 산업은 1.0시대를 지나 빅데이터로 실질적 가치를 창출하는 2.0시대로 진입하였다. 개념 또한 3v에서 4v로 축적 데이터의 가치에 초점을 맞추어지고 있다. 이것은 빅데이터를 하나의 자산으로서의 가치를 담고 독립적인 산업으로서 근거를 제시하기 위한 것으로, 빅데이터의 정의는 데이터를 수집, 저장, 처리, 분석하는 것 뿐 아니라 새로운 가치를 전 과정을 포괄하는 의미로 해석할 수

있다. 빅데이터는 디지털 환경에서 생성되는 모든 데이터를 의미하는 것으로 규모가 방대하고 생성주기가 짧으며 형태도 다양하다는 특성을 가지고 있다. 스마트폰과 인터넷이 대중화되면서 사용자들이 남기는 데이터의 양과 종류는 더 큰 규모로 생성되고 있다. 이렇게 생성된 빅데이터는 방대한 데이터에서 가치가 담긴 정보를 추출하여 활용하는 시기로 전환되었다. 그간의 개념검증과 시범도입을 통해 활용 방안을 확보함에 따라 데이터 수집, 저장을 넘어 실질적 활용에 보다 더 주목하고 있는 것이다. 빅데이터는 또한 의료산업이나 보건

분야에도 응용될 수 있으며 IOT, 스마트 헬스케어 등의 기술과 함께 융합되어 시너지 효과를 창출할 것이라 전망이다. 빅데이터 분석과 효율적인 모니터링 시스템이 더 많은 국가의 보편적 건강 보장에 기여할 것으로 본다. 또한 미래창조과학부는 올해 빅데이터 시장이 지난해 2623억원 보다 30% 증가한 3400억원에 이를 것이라고 14일 발표하였다. 미래부는 경제 혁신 3개년 계획과 K-ICT 전략 일환으로 빅데이터 시장 현황을 조사한 결과, 빅데이터 시장은 지난 2013년 이후 매년 20~30% 증가한 것으로 나타났다. 시장 성장 배경은 민관 투자 확대에 볼 수 있다. 올해는 제조·의료 산업이 빅데이터 시장을 견인할 것으로 예측하고 있다. 그러나 방대한 데이터를 의미 있고 안전하게 활용하기 위해서는 많은 정보보호 등의 문제점 및 선행과제를 해결해야만 한다[1][2]. 따라서 본 연구에서는 의료 빅데이터 활용사례와 기대효과 그리고 해결과제, 마지막으로 의료 빅데이터의 미래전망을 분석한다.

II. 의료 빅데이터의 활용사례와 기대효과

빅데이터의 활용이 가장 좋은 부문은 의료이다. 빅데이터는 새로운 서비스의 제공과 효율적인 의료서비스의 제공을 가능하게 하기 때문이다. 의료 부문의 주요 빅데이터는 제약회사의 연구개발 데이터, 임상데이터, 청구 및 비용 데이터, 환자행동 및 감성 데이터로 효율적인 의료서비스의 제공, 신약 및 새로운 의료기법의 연구개발, 신규 보건 의료 서비스의 제공, 보건의료 정책의 의사결정에 활용될 수 있다. 또한 의료진의 의사결정 지원과 개인정보를 바탕으로 한 질병 예측 등 신규서비스 개발을 가능하게 한다. 의료산업적인 측면에서 또한 신약과 신규 의료기술의 개발을 위한 연구에 활용되어 비용절감과 개발기간 단축에 활용될 수 있으며, 제약회사의 효율적인 재고관리, 수요 예측에도 유용하다. 보건의료정책을 결정함에 있어 빅데이터는 통합적인 질병감시 및 관리, 의료 수요 파악을 통한 의료 수급 조절, 효과적인 치료법의 선별과 보급, 부당청구 적발을 통한 효율적인 의료보험서비스 제공에 기여한다. 의료부문의 빅데이터 활용사례[3] 일부를 보면 다음 표 1과 같다.

표 1. 의료 빅데이터 활용사례와 기대효과

활용 사례	활용내용	기대효과
N I H Pillbox	소비자가 알기 어려운 약의 약물에 대한 복용량, 부작용, 주의사항 등 의학 정보 제공	연간 5,000만 달러 이상의 비용을 절감함
Explorys	- 14개의 주요 통합 의료시스템으로 구성된 네트워크를 구축 - 임상, 병원 경영 등	- 진료예약 최적화로 진료대기시간 20% 줄임 - 화학치료를 위

	에 관한 데이터 수집 - 응급환자수송, 수요 예측, 환자안내 등에 활용	한 환자 대기 시간을 1/3로 줄임 - 맞춤형 의료서비스 제공
IBM Watson	- 6십만 건의 의학적 근거와 2백만 장의 전공서적 및 2만 5천 건의 사례 등 빅데이터를 분석 - 클라우드 기반으로 서비스에 가입한 의료진은 언제, 어디서든 모바일기기를 통해 Watson을 활용가능	빅데이터 분석 능력을 극대화, 의료진의 데이터 활용도 향상
Fitness Tracker	- 생활 밀착형 건강관리 모바일 기기 - 다양한 건강관리 기기가 클럽, 암밴드 등의 형태로 착용 - 사용자의 걸음걸이, 이동거리, 칼로리 소모량, 수면의 질 등의 정보 측정 - 스마트폰과 연동	- 건강관리에 대한 관심과 수요 증가 - 고도화된 센싱 기술 발전 가속화 - 모바일 생태계와의 연계를 기반으로 건강 데이터의 종합적 분석, 수요자 맞춤형 서비스의 상용화
HBP(Human Brain Project, EU)	- 뇌공학 연구 프로젝트 - ICT 플랫폼을 중심으로 뇌지도 작성·시각화, 대규모 데이터 분석을 통한 수리적 모형 구축 및 시뮬레이션	뇌 연구로 데이터 처리에 대한 이해를 넓히고 인공지능을 발전시켜 빅데이터 분석 플랫폼의 진화를 이끌어내는 선순환 구조 형성
익스프레스 스크립츠	환자들에게 저렴하게 약제를 보충할 방법 소개, 맞춤형 메시지 전송	환자의 약품 관리 지원 정보 처리 과정에 빅 데이터 활용, 행동 양식 변화와 프로세스 개선 효과
인마틴스어	- 유타주 내 22 곳의 병원과 185 곳의 의료단체, 연계 보험사 지원 시스템 - 소속 의료 공급자들의 전자 의료 기록을 활용하는 2개의 틀을 이용, 건강 결과 분석	건강 결과 평가
유아드스어	잠재적 의료 보험 사기 및 신원 도용 가능성 감지 도구로 소셜 네트워크 연구	보험사기 감지, 질병 치료, 고객 만족 지원
파트너 헬스케어	- EHR용 지식 시스템 - 기관들의 금융 시스템, 운영 시스템, 임상 분석 시스템을 통합	분석자료 수집 및 통합
구독 감보 서비스	독감, 인플루엔자 등 감기와 관련된 검색어 쿼리의 빈도를 조사, 독감 확산 조기경보 서비스	미국 보건당국보다 앞서 시간 및 지역별 독감 유행 정보 제공

이 외에도 영국의 사례를 보면 동식물 및 인체 전염병 확산 대응을 위해 USN, 데이터마이닝, 유전공학, 생물정보학 등을 이용하여 HIV/AIDS, 말라리아 등 다양한 전염병에 대한 전망과 대응방안 마련하였고 비만대책 수립 사례로 영국정부는 2050년까지 남자의 60%, 여자의 50%가 비만이 될 것을 예측하고 매년 499억 파운드가 소요될 것으로 전망하고 있다. 빅데이터를 통한 과학적 근거를 기반으로 정부, 기업, 시민사회의 공조하여 지

속가능한 비만대책을 수립하여 향후 40년간 영국 사회의 비만문제를 효과적으로 대응하고 있다. 국내에서도 보건의료 빅데이터 활용에 적극적인 관심을 보이고 건보공단은 전 국민의 건강보험자격 및 보험료, 건강검진 결과, 진료내역, 요양기관 현황, 암 및 희귀난치성 질환자 등록정보, 노인 장기요양보험 자료등 1조3000억 건에 달하는 방대한 데이터를 기반으로 2012년 국민건강정보 DB를 구축했다. 2013년에는 성별·연령별·소득분위별로 대표성 있는 표본(100만명)을 추출해 연구용 표본코호트DB를 마련했다. 모든 자료는 공익적 활용을 위해 외부에 공개하고 있다. 심평원 또한 지난 15일 진료 정보, 의약품, 지역별 요양기관 등에 대한 의료 분야 공공데이터 약 3258억건을 보건의료빅데이터개방시스템에 개방해 빅데이터 상용화 시대를 열었다. 이처럼 보건의료 분야에서 빅데이터의 활용가능성은 매우 크다.

III. 해결 과제와 이슈

3.1 의료정보보안

보건의료 분야의 빅데이터는 많은 양의 정보를 담고 있어 임상 및 연구에서 활용 가능성이 크지만 개인정보 유출, 정보의 상업적 이용에 따른 의료법 위반 문제가 발생할 수 있다. 국민건강보험공단과 건강보험심사평가원이 빅데이터를 주도하지만 실제 병·의원에서는 긍정적이지는 않다. 가장 큰 문제점은 개인정보 유출이다. 빅데이터의 상용화 초기단계부터 개인정보 유출에 대한 문제점이 제기되었지만 건보공단과 심평원은 경쟁적으로 빅데이터의 활용 방안을 마련하는 것에만 관심을 가질뿐 정보보안의 심각성은 고려하지 않았다. 최근 사회적 이슈가 되고 있는 주요 신용카드사의 대량의 개인정보 유출 사건은 이러한 측면에서 빅데이터 구축 및 활용을 제약하고, 정부 3.0의 구현을 어렵게 만드는 한계로 작용할 수 있다. 따라서 보건의료 빅데이터가 활성화되기 위해서는 개인정보 보호법, 기술적인 보호방법들이 법적제도[4][5]로 강화되어야 할 것이다.

3.2 데이터 표준화

빅데이터를 도입한 외국의 경우 빅데이터를 통해 의료질의 향상과 속도증가가 되었다는 결과를 얻었지만 또 다른 단점은 데이터 표준화이다. 데이터 용어 체계의 수립 즉, 통합의료 정보시스템 용어와 코드 표준을 준수해 데이터에 대한 이해력을 제고하고 상호 운용성을 확보해야 한다. 그리고 비정형 데이터의 정형화이다. 기록지나 양식에 있는 비정형 데이터를 뽑아서 분석에 사용하고자 하는 요구가 많아지고 있다. 미국 보건의료 연구 및 질 관리 기구 어니스트 모이 교수는 “미국에서의 빅데이터 사용 현황은 의료비용 및 활용 프로젝트, 의료비 패널 조사 이외에 다양한 모델을

이용한 자료 수집 및 분석 중이지만 너무 많은 측정법들이 존재해 데이터 표준화가 필요하였고 환자의 사용제한과 측정도구의 미흡한 평가 등이 단점으로 나타났다고 이야기 하였다. 국내에서도 데이터 표준화를 위한 해결방안을 모색 중에 있다. 또한 각 기관들 간에 분류개념이 다른데서 오는 데이터 표준화[3] 문제를 인지하고 있어 개선점을 찾고 있다.

3.3 양질의 데이터 구축

우리나라 의료보장체계와 데이터 구축[3] 수준은 비교적 빠른 시간 내에 선진국 수준이 되었고 인터넷 망이 세계적인 수준으로, 현재까지 축적된 광대한 정보 데이터는 최고수준이라고 평가받았다. 의료보장 분야에서 비용 문제에 대한 논의가 있었지만 지난 2010년 이후로는 질과 비용에 대한 정책적 관심을 쏟으며 바로 양질의 데이터가 중요한 역할을 하게 되었다. 정부는 보편적 의료보장이 어느 정도 단계에 이르렀다고 판단하고 질 향상에 초점을 두고 있다.

3.4 보건의료 빅데이터 R&D 및 성장동력 창출

바이오 메디컬 분야의 새로운 발견과 환자중심의 치료법, 질병의 사전예방을 가능하게 하는 보건, 질병, 유전자 관련 고부가가치 지식베이스 및 플랫폼이 개발해야 한다. 보건의료 분야는 예방건강관리 중심으로 패러다임 전환에 따른 적응과 더불어 국민의 치료비 증가, 병원 수익구조 악화 등 어려움을 극복하기 위한 산업계의 전반적인 효율성 제고 노력이 이루어져야 하는 상황이다. 이 같은 문제점을 극복하고 산업 경쟁력 확보를 위해서는 빅데이터의 활용은 필수라고 볼 수 있다. 또한 보건의료 빅데이터 산업생태계 조성 및 전문인력을 양성[4]하고 보건의료 빅데이터의 산업기반인 클라우드 컴퓨팅, 서버, 플랫폼, 네트워킹, 시각화 기술 분야의 전문 중소기업 육성해야 한다.

3.5 빅데이터 기반의 병원정보계시스템 인프라 구성

최근 병원의 사회적 환경을 보면, 노인인구의 증가, 진단장비의 발달, 치료기술의 발달, 해외환자 증가, 정보통신기술의 발달, 연구중심병원의 지원, 의료기관 해외진출, 의료직접근용이, 의료기관 1인 1개 소, 전문병원의 증가, 수도권 집중현상, 포괄 수가제도의 확대 등을 들 수 있다. 따라서 병원 정보화에서 주목받는 기술은 세 가지로 예측과 통계 기술, 데이터 표준화 기술, 그리고 빅데이터 기술 등으로 볼 수 있다. 이를 바탕으로 병원들은 예방 관리와 조기진단 중심의 의료 체계 구축, 환자별 개인 맞춤형 서비스, 웰빙과 헬스케어 산업의 결합, 그리고 나아가 건강 수명시대를 추진할 수 있다. 하지만 병원시스템의 사용측면, 연구측면, 운영

측면에서의 문제점[4]들이 도출되고 있다.

IV. 의료 빅데이터의 미래

정보통신기술이 모든 산업과 융복합되는 '4차 산업혁명'은 의료산업에도 큰 변화를 가져올 것으로 전망하고 있다. 이것은 의료의 패러다임이 '치료'에서 '예방과 관리'로 바뀌면서 ICT 기술을 융복합한 '디지털 헬스케어' 기술이 새로운 산업을 태동시키고 있기 때문이다. 미래의료[6]는 유전자 정보를 통해 질병을 예측하고, 사물인터넷(IoT)과 센서를 통해 실시간으로 건강 상태를 점검하고 관리하는 형태가 될 것이다. 디지털 헬스케어는 이미 초기 기술개발 단계를 넘어서 데이터와 콘텐츠가 주도하는 단계로 접어들고 있으며, 다양한 기술과 서비스를 통합하는 '플랫폼' 경쟁이 본격화되고 있다. 디지털 헬스케어 서비스 방향[6]을 보면 다음 표 2와 같다.

표 2. 디지털 헬스케어 서비스의 발전방향

	Tele-헬스 (원내치료)	e-헬스 (치료 및 정보제공)	U-헬스 (치료 예방 관리)	smart-헬스 (치매 예방 복지인전)
시기	1990년대	2000년대	2006년대	2010년 이후
서비스내용	원내치료	치료 및 정보 제공	치료 예방 관리	치료 예방 복지인전
주 Player	병원	병원	병원, ICT기업	병원, ICT기업, 보험사 서비스 기업 등
주 이용자	의료인	의료인, 환자	의료인, 환자, 일반인	의료인, 환자, 일반인
주요시스템	병원운영 (HIS, PACS)	의무기록(EMR) 웹사이트	건강기록(D-R) 모니터링	개인건강기록 기반 맞춤형 서비스

1) ICT 융합이 가져오는 헬스케어 혁신

미국의 개인 유전정보 분석 업체 패스웨이지노믹스는 웰니스 애플리케이션 'OME'를 소개하였다. 이 앱은 개인 유전정보와 애플 '헬스킷'을 통해 수집한 건강상태, 운동기록, GPS 정보 등을 IBM의 인공지능 컴퓨터 '왓슨'으로 분석해 맞춤형 건강 조언을 제공한다. 유전자 정보를 활용한 개인 맞춤의료와 빅데이터, 클라우드, 인공지능 등 첨단 ICT 기능을 결합한 이 앱이 계획대로 개발된다면 디지털 헬스케어의 이상적인 형태에 가까울 것으로 보인다. 전에 완전히 다른 생태계를 갖고 있던 IT와 의료 산업이 디지털 헬스케어 시대를 맞아 한 플랫폼 안에서 본격적인 결합을 하고 있는 것이다.

2) 전 세계 '플랫폼' 확보 경쟁 치열

디지털 헬스케어 시장을 잡기 위해선 '플랫폼'을 잡는 것이 핵심이다. 이미 구글과 애플, IBM, 쿼크 등 거대 IT 기업들이 헬스케어 플랫폼 구축 경쟁에 뛰어들었고, 최근에는 중국 기업들도 이 시장에 진출해 속도를 내고 있다.

3) 의료 및 헬스케어에 클라우드 컴퓨팅 도입
의료 및 헬스케어 분야에 클라우드 컴퓨팅이 도입되면 원격진료는 물론, 환자 데이터 관리 등 다양한 이점이 생길 것이다. 보건의료연구개발사업의 일환으로 분당서울대병원과 경북대학교, SKT는 클라우드 기반의 글로벌 PHR시스템 구축을 하고 있다. 이것은 환경에 기반한 사용자 맞춤형 서비스 제공이 가능하고, 개인건강기록을 통해 국민건강관리에 활용될 것으로 기대된다. 하지만 국내 의료 산업 분야 클라우드 도입은 많지 않다. 이것은 현재의 의료법으로는 환자의 의료정보가 클라우드에 저장될 수 없고, 개인정보 유출 등을 방지하기 위한 완벽한 장치가 마련되어 있지 않기 때문이다. 따라서 현재의 의료법과 개정안은 클라우드 발전과 충돌하는 부분이 많아 의료분야 클라우드 규제 완화가 이뤄져야 할 것이다.

V. 결 론

보건의료 분야의 최근 흐름은 빅데이터와 지식재산에 집중되어 있다. 특히 ICT 분야의 대표적 기조인 빅데이터는 의료 영역과 융합하면서 의료계에 새로운 패러다임을 몰고 있다. 이 패러다임은 엄청난 데이터의 발생과 복잡성의 증가로 인해, 맞춤형 의료 서비스와 근거리 중심의 의사결정이 가능하게 할 것이며 빅데이터의 특징인 다양한 자료들을 연계하고 분석하여, 비용절감 뿐 아니라 단순 모니터링 단계에서 평가 및 예측 분야에서 성과를 거두고 있다. 따라서 의료 분야의 빅데이터는 개인의 임상 정보 및 유전체, 그리고 행동 관련 데이터를 연계하여 개인별 맞춤형 예측 시스템으로 발전할 것이라 전망하며 더욱 활성화되기 위해서는 빅데이터 공개 범위의 확대 및 전문 인력 양성, 그리고 제도 마련 등이 뒷받침되어야 할 것이다.

참고문헌

[1] 백승민, "빅데이터를 활용한 보건산업 신산업 전망 및 정책방향", 한국보건산업진흥원, vol 84, Jul, 2013.
 [2] 송태민, "우리나라 보건복지 빅데이터 동향 및 활용방안", 과학기술정책, 제23권 제 3호, Nov. 2012.
 [3] 이창현, "의료산업에서의 빅데이터활용 사례", 정보통신산업진흥원, 주간기술, Feb. 2013.
 [4] "빅데이터 기반의 병원정보계 시스템 동향과 해결방안", IDG IT world.
 [5] 정현학, 장보은, "보건의료 빅데이터 관련 각국의 법체계", 한국보건산업진흥원, vol.208, Dec. 2015.
 [6] 디지털 타임즈, 27면 Mar. 2016.