

보건의료를 위한 빅데이터의 활용

박 병 주^{1,2} | ¹한국의약품안전관리원, ²서울대학교 의과대학 예방의학교실

Application of big data for public health

Byung-Joo Park, MD^{1,2}

¹Korea Institute of Drug Safety and Risk Management, Seoul, ²Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea

빅데이터는 단순히 방대한 자료가 아니라 큰 크기로 인하여 처리와 분석이 용이하지 않으나, 이를 이용하여 의사결정에 도움이 되는 유용한 정보를 생산해낼 수 있는 자료, 그리고 그 자료를 활용할 수 있는 일련의 시스템을 뜻한다[1]. 빅데이터의 활용이 보건의료 분야에 가질 수 있는 가치는 매우 커서[2], 본 특집을 통하여 빅데이터의 국내외 현황을 살펴보고, 우리나라에서 보건의료관련 구체적인 사례라 할 수 있는 약물감시에서의 응용사례를 살펴보고, 보건의료 전 분야에서 적극적 활용을 위하여 필요한 조치와 함께 제도 정책적 인프라 구축 방안을 살펴보기로 한다.

최근 전 세계적으로 정부차원에서 빅데이터에 대한 투자가 활발하며, 우리 정부도 2013년부터 창조경제, 정부 3.0을 국정과제로 내세우고 공공분야 빅데이터 활용을 위하여 노력하고 있다. 미국은 2012년 'Big Data R&D Initiative'에서 적극적인 투자의지를 밝혔고, 일본은 2012년 '빅데이터 활용 기본전략'을 발표하고 정부주도의 발전전략을 실행하고 있다. 우리나라는 미래창조과학부 중심의 '빅데이터 산업 발전

전략'을 발표하고 주요 공공분야의 자료를 공개하고 인력·인프라 양성에 박차를 가하여, 현재 다수의 공공분야 빅데이터 프로젝트가 진행 중에 있다. 하지만, 아직 빅데이터 선진국과 기술격차를 따라잡고, 전문인력을 양성, 개인정보를 보호하면서도 새로운 가치를 창출할 수 있는 서비스를 기획해 내는 역량강화라는 과제가 남아있다.

보건의료 분야에서 빅데이터에 대한 공공분야 투자는 비감염성질환이 증가하면서, 이로 인한 의료비지출 증가를 억제하기 위한 최선의 과학적 근거를 가진 의료를 찾는 노력에서 촉발되었다[3]. 이런 배경으로 유럽에서는 보건의료기술평가(healthcare technology assessment), 북미에서는 비교효과연구(comparative effective research)가 활성화되었고, 구미 선진국에서는 각 국가의 상황에 맞게 국가기구를 창설하고 비용효과적인 의료서비스를 제공하기 위하여 전국민 대상의 의료서비스 관련 전산자료를 구축하여 활용하고 있다. 미국의 경우 오바마 행정부가 빅데이터 활용의 하드웨어를, 비영리기술단체가 소프트웨어를, 그리고 비영리·비정부단체가 이들의 활동을 조율하고 있다. 특히, 기존 구축 자료 간의 네트워크를 개발하고 행정기관과 의료기관의 개별환자를 연결시킨 데이터베이스를 구축하는 것이 보건의료 분야 문제해결을 위한 주된 인프라로서 강조되고 있다.

주어진 자료원들을 활용하여 효율적이고 효과적인 보건 의료서비스를 제공하기 위한 근거창출을 하려면 자료를 생산하는 기관에 분산되어 있는 빅데이터를 연계하는 것이 필

Received: April 16, 2014 Accepted: April 30, 2014

Corresponding author: Byung-Joo Park
E-mail: bjpark@snu.ac.kr

© Korean Medical Association

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

수적이다[4]. 이미 미국, 영국, 호주, 캐나다 등 의료선진국에서는 보건의료분야 연구에 연계자료를 활발히 활용하고 있으며, 미국의 SEER (Surveillance, Epidemiology, and End Results)-Medicare data 그리고 호주의 Western Australian Data Linkage System 등이 그 예이다. 국내에서 국민건강보험공단, 건강보험심사평가원에서 자료를 공개하였지만, 이 자료들은 연계가 되지 않은 단일기관자료원으로 그 활용에 한계가 있었다. 하지만, 한국보건의료연구원은 이들 자료와 병원진료자료, 중앙암등록자료, 통계청사망자료를 연계하여 전립샘암 검진법의 양성예측도와 경제성을 분석하여 국가암 검진 프로그램에 전립샘암 검진을 포함할 것인지에 대한 과학적 근거를 생산함으로써 국내 다양한 보건의료 자료원들의 연계 활용이 유용함을 보였다. 그리고 보건의료기술진흥법 개정을 통하여 국가기관 및 공공기관 민감정보 및 고유식별정보를 수집하여 연계할 수 있는 법적근거를 마련하였으나, 자료보유기관의 협조 하에 각 기관의 자료를 검토하는 방대한 작업을 수행할 자원을 확보하는 것이 필요하다.

의약품안전성 감시 분야는 빅데이터의 활용이 특히 유용할 수 있다. 빅데이터는 의약품의 유해성을 파악하는데 소요되는 시간을 줄여줄 수 있다. 이로써, 유해한 의약품에 우리 국민들이 노출되는 기간을 줄여주고 결과적으로 의약품에 의한 위해로부터 국민들을 보호할 수 있다. 또한, 많은 경우 부족한 유해성 정보에 근거하여 시판금지 조치가 취해지는데, 빅데이터는 유해성 여부에 대한 객관적인 근거를 만들어 벤딕틴, 프로폭시펜의 예에서 볼 수 있듯이 안전하고 효과적인 약이 시장에서 퇴출되는 것을 막아 줄 수도 있다. 또한 웹 게시물을 비롯한 비정형 빅데이터를 이용할 경우 환자보고 안전성(patient reported safety outcome)까지 안전성 모니터링의 범위를 확장할 수 있어서 의료인이 발생시키는 정형 자료의 한계를 극복할 수 있는 가능성까지 갖고 있다. 이에 미국에서는 정부주도로 미니센티넬(mini-sentinel) 사업을 거쳐 하버드대학을 협연센터로 하는 분산데이터 분석체계를 구축하였고, 민간 중심으로는 Observational Healthcare Data Sciences and Informatics를 구축하여 운영하고 있다. 유럽에서는 유럽의약품청 주도로 ENCePP (European Network of Centres for Pharmacoepidemiology and

Pharmacovigilance)를 운영하고 있으며, 제약업계 협력체인 PROTECT (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium)가 유럽의약품청과 협력 하에 전자의무기록을 이용한 유해성 조기파악방법론을 개발하고 있다[5]. 국내에서도 정형화된 빅데이터인 건강보험심사평가원 자료를 이용한 약물역학 연구가 활발하며, 비정형데이터 분석 프로젝트인 '의약품안전성 조기경보서비스'가 한국의약품안전관리원, 미래창조과학부, 그리고 한국정보화진흥원의 시범사업으로 2013년 진행되어 실용화의 가능성을 확인하였다.

하지만, 자료연계에 기반한 빅데이터 활용을 통한 공공이익의 실현도 국민들의 사생활의 비밀과 자유가 보호되는 범위 내에서 이루어져야 한다. 우리나라에서는 최근까지 대규모의 개인정보 유출사고가 있어왔고, 이러한 배경 하에서 2011년 개인정보보호법이 제정되어 공중보건과 같은 공익적 목적을 위한 개인정보의 사용에 대한 논의가 법 제정 과정에서 충분하지 못하다. 그 결과 현재의 개인정보보호법은 다른 법률에 특별한 규정이 없는 경우 자료연계를 통한 학술 연구를 허용하지 않는 것으로 해석되기 때문에, 지금까지 국가에서 정책적으로 추진하고 있는 빅데이터의 적극적 활용이 개인정보보호법에 상충하는 상황이 발생할 수 있으며, 이에 대한 근본적 논의와 국민적 합의에 기반한 제도적 보완 없이는 갈등과 혼란이 발생할 소지가 크다. 유럽에서는 최근 정보보호기본규칙안을 통하여 공중보건전문가들의 의견을 반영하여 일반적인 학술연구와 구분하여 공중보건을 위한 개인정보를 사용하는 것을 분명하고 허락하고 있다[6]. 우리나라에서도 행정목적을 위하여 개인정보의 연계를 허용하는 사회복지통합전산망과 관련된 법령이 있다. 이러한 사례에서 밝혀진 공익과 사익의 조화가 공공보건 분야에서도 명백히 가능하므로 개인정보보호 속에 공공보건을 위한 개인정보 처리를 가능하게 하는 우리나라의 현실을 고려한 제도적 장치가 마련되어야 할 것이다[7].

ORCID

Byung-Joo Park, <http://orcid.org/0000-0003-4630-4942>

REFERENCES

1. Beyer MA, Douglas L. The importance of 'big data': a definition. Stamford: Gartner; 2012.
2. Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA 2013;309:1351-1352.
3. Bae JM, Park BJ, Ahn YO. Perspectives of clinical epidemiology in Korea. J Korean Med Assoc 2013;56:718-723.
4. Bradley CJ, Penberthy L, Devers KJ, Holden DJ. Health services research and data linkages: issues, methods, and directions for the future. Health Serv Res 2010;45(5 Pt 2):1468-1488.
5. Choi NK, Lee J, Park BJ. Recent international initiatives of drug safety management. J Korean Med Assoc 2012;55:819-826.
6. Carinci F, Di Iorio CT, Ricciardi W, Klazinga N, Verschuuren M. Revision of the European Data Protection Directive: opportunity or threat for public health monitoring? Eur J Public Health 2011;21:684-685.
7. Larson EB. Building trust in the power of "big data" research to serve the public good. JAMA 2013;309:2443-2444.

자율학습 2014년 4월호 정답 (노화와 조혈기능 장애)

1. ①
2. ③
3. ④
4. ②
5. ②

6. ③
7. ①
8. ③
9. ④
10. ③