보건의료 분야 빅데이터 활성화와 과제

황의동 실장
진료정보분석실

1. 왜 빅데이터인가

보건의료 분야는 고령화 · 질병 패턴의 변화 등으로 인한 보건의료 패러다임의 변화, 증가하는 의료비의 적정화, 의료 서비스의 질적 수준에 대한 소비자의 관심 증가, 의료 효율화와 성과의 극대화 등 당면한 문제의 해결에 대한 압박을 받고 있다. 또한 세계 각국은 보건의료산업의 경쟁력 강화와 신성장동력으로서의 가치에 주목하고 있다.

보건의료는 정보통신기술(Information and Communication Technologies, ICT)과의 융합이 다른 분야보다 친화적이어서, ICT와 소셜 네트워크 서비스(Social Network Services, SNS)를 통해 수집되는 빅데이터가 많고 그 활용이 용이하며 경제 · 사회에서 차지하는 비중이 커 빅데이터의 효과를 가장 극대화 할 수 있는 분야로 꼽힌다(Mckinsey, 2013). 보건의료의 당면과제 해결과 많은 파생서비스가 탄생되는 등의 새로운 가치창출은 빅데이터가 주목받는 이유이다. "빅데이터는 생명을 구한다"거나 "데이터에 답이 있다"는 말은 보건의료 빅데이터가 만드는 세상이 우리 삶의 질 향상과 행복을 위한 기회이자 동시에 위기일 수 있음을 시사한다.

그러나 우리는 IT강국임에도 불구하고, 보건의료 빅데이터 활용은 여전히 초기 단계이다. 전문가들은 법과 제도의 정비, 활용을 위한 마스터플랜 및 환경기반의 조성을 우선 과제로 꼽고 있다.

이하에서는 '빅데이터의 현상과 가치'를 바탕으로 국내 보건의료 빅데이터의 활용사례와

정부 전략을 검토하고, 건강보험심사평가원(이하 심평원)의 빅데이터를 중심으로 한 활용방안과 과제를 제시해 보고자 한다.

가. 빅데이터 현상의 이해

빅데이터는 관점에 따라 다양하게 정의되나, '크고 다양한 데이터'로부터 신속하게 의미있는 정보나 가치를 "추출해 내는 방법 · 기술 내지 프로세스"로 볼 수 있다.

인터넷과 SNS가 일상화됨에 따라 모든 분야에서 디지털 정보량이 기하급수적으로 증가하여, 기존의 관리 및 분석체계로는 감당할 수 없을 정도의 거대한 데이터의 집합을 빅데이터로 지칭(SERI, 2010)한다. 일반적으로 빅데이터는 '3V' 즉, 규모(Volume), 다양성(Variety)과 속도(Velocity)의 특징을 어느 정도 수준에서 만족해야 유용하다(Gartner, 2011). 여기에 가치(Value)를 더해 '4V'로 특징짓기도 한다.

빅데이터가 주목받는 이유는 ICT의 주도권이 데이터로 이동하고, 공간 · 시간 · 관계 · 세상을 담는 데이터가 생성되며 이는 미래 경쟁력과 가치창출의 원천이기 때문이다.

표 1. 데이터의 과거-현재-미래

저장	→	검색	→	관리	→	공유	→	분석	→	분석(전망)
DB		검색엔진		KMS		Web 2.0		빅데이터		상황인식 예측
축적		활용		확산 및 공유				가치창출		

* 출처 : NIA 빅데이터 전략연구소, 빅데이터 시대

빅데이터의 사회 · 경제적 의미는 새로운 가치의 원천(천연자원, 데이터 금맥)이고, 기업 경쟁력을 좌우하는 산업적 도구이자 정보의 홍수로 인한 새로운 재난(위험)으로 이해된다(Tyler Bell, 2011).

빅데이터의 활용 가치는 의료 · 공공행정 등에 적용 시 1%의 추가적인 경제적 효과와 산업부문별로 연간 약 0.5~1% 정도의 생산성 증가가 추정된다. 뿐만 아니라 각국은 국가의 안전 및 국민 행복을 위협하는 재난재해, 질병 등의 위기에 선제적으로 대응하기 위해서 데이터 분석을 활용하고 있다(Mckinsey, 2011).

빅데이터가 가치를 만들어내는 방식은 ① 투명성 제고로 연구개발 및 관리 효율성 향상

② 시뮬레이션을 통한 수요 포착 및 주요 변수 탐색으로 경쟁력 강화 ③ 고객 세분화로 맞춤형 서비스 제공 ④ 알고리즘을 활용한 의사결정 보완 및 대체 ⑤ 비지니스 모델과 제품, 서비스의 혁신 등이다(Mckinsey, 2011).

빅데이터는 새로운 차원의 생산성 향상과 경쟁력 제고, '발견(예측)'에 의한 문제 해결과 맞춤화의 실현, 의사결정(경영)의 과학화와 자동화, 새로운 고객 가치와 비즈니스 창출이 가능하다는 현실로 이어져 활용도가 높아지고 있다.

즉, 빅데이터는 공공부문과 보건의료 등에 있어 혁신, 경쟁력, 생산성의 핵심요소로서 미래사회의 현안과 과제를 해결하는 신 가치창출 엔진이 될 것이다.

나. 빅데이터의 기회와 위기

빅데이터에 관심이 집중되고, 차세대 이슈가 되는 것은 빅데이터가 가진 무한한 가능성과 잠재가치 때문이다.

'데이터로 열어가는 산업혁명 시대(Tyler Bell, 2011)'에서의 빅데이터 기능은 ① 차세대 산업혁명의 석탄이나 철 ② 21세기의 원유(oil) ③ 렌즈(현미경) ④ 플랫폼 역할이다(한국DB진흥원, 2014).

빅데이터는 ① 데이터의 사전처리에서 사후처리로 ② 표본조사에서 전수조사로 ③ 질보다 양으로 ④ 인과관계에서 상관관계로, 데이터 본질의 변화를 가져오므로(한국DB진흥원, 2014) 이런 현상의 이해 역시 필요하다.

국내 건강보험 빅데이터를 보건의료산업에 활용할 경우, 국내 의료비 절감 기댓값(미국의 사례에 따른 단순 계산)은 최대 8조 2,000억 원으로 전망되고 의료는 예방중심 및 환자중심으로 변화하고, 의료서비스가 개선될 것이라고 예상한다(조대곤 · 하용찬, 2014).

미국 의료서비스 부문에서 적절한 빅데이터 활용이 이루어질 경우, 매년 1,000~1,900억 달러 이상의 비용절감이 예측된다. 특히, 임상에서 비교효과성 분석, 임상 의사결정 지원, 원격진료 수준 향상, 성과의 투명성 제고 등으로 인해 750~1,500억 달러가 절감 가능하다고 추정한다(Mckinsey, 2013). 영국은 병원 · 의료서비스 산업에서 빅데이터로 인한 경제적 효과가 2012~2017년 사이 총 204억 파운드에 달하고, 같은 기간 4,000여개 일자리 창출을 전망한다(CEBR, 2012).

이러한 빅데이터 시대를 위해서는 준비가 필요하다. 우선 '빅데이터 현상'을 이해하고 정보 독점 상태를 벗어나는 즉, '상호연결과 협력'을 통해 활용도와 가치가 더욱 확장될 수 있다. 이를 위해 플랫폼(App.), 오픈소스(SW), 초고속 컴퓨팅파워(HW)의 영향력 증대가 요

구된다. 이러한 것은 '정보공유와 협업', '집단지성'을 의미하고 연결된다.

빅데이터의 핵심역량은 '창의적인 데이터 과학자(Data Scientist)'의 양성과 품질관리 등의 기반 조성이다. 데이터 안전성과 신뢰환경 기반 구축으로 빅데이터는 위기가 아닌 기회가 될 수 있다.

빅데이터 시대의 '어두운 그림자'로 사생활 침해와 정보보호, 책임원칙 훼손, 데이터 오남용 문제가 부각되고 있다. 이러한 위기는 데이터의 안전한 사용을 보장하는 정보보안과 사용가이드라인, 동의에서 책임으로, 결과 기반 책임원칙 규범과 알고리즘 접근 허용으로 해결될 수 있도록 법적 · 사회적 측면의 환경이 전제되어야 할 것이다.

빅데이터가 기회가 될 지, 아니면 위기가 될 지는 결국 활용하는 사람과 사회의 몫이다.

2. 우리나라 보건의료 빅데이터 활용 사례와 정책

가. 활용사례

공공부문에서는 심평원과 국민건강보험공단(이하 건강보험공단)의 건강보험 및 보건의료 데이터 개방이 점진적으로 이루어지고 있다.

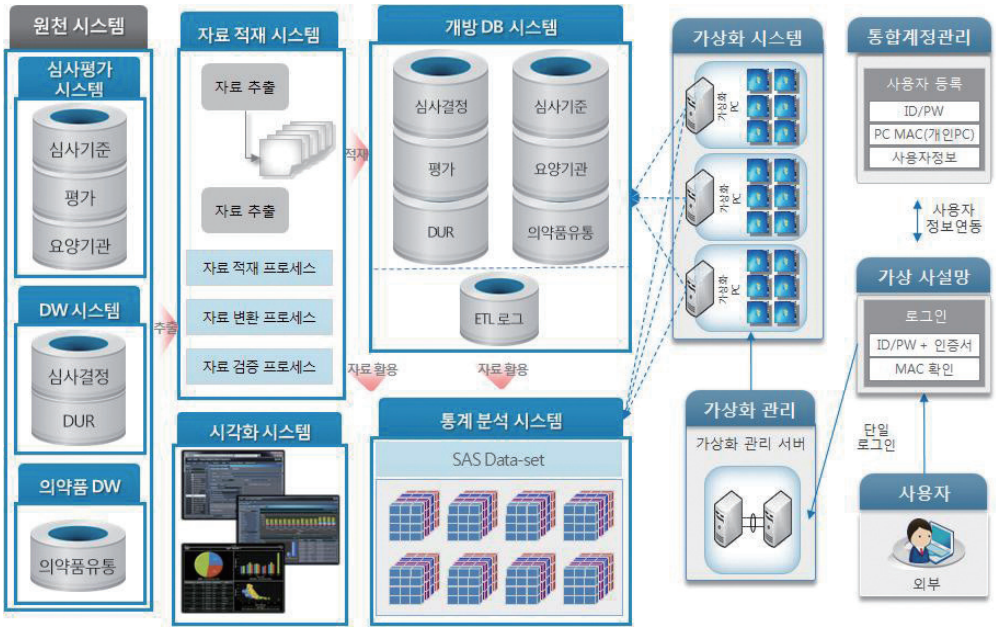
심평원은 보건의료정보의 개방 · 제공 및 활용 확대를 통해 의료의 성과를 높이고, 보건의료 생태계를 조성하고자 2013년 이후부터 의료데이터를 개방하기 시작하였다. 이를 본격화하기 위해 2014년 4월부터 '보건의료빅데이터센터'를 운영하고 있다.

심평원은 실시간 수집되는 국민들의 진료내역, 의약품 생산 · 공급 · 사용 및 처방내역, 의료인력 · 시설 · 자원 현황, 의료 질 평가정보 등 30여 종의 보건의료정보를 연간 약 14억 건(522TB)을 공개하고 있다. 이는 A4용지로 출력해 쌓을 경우 에베레스트산(8,848m)의 22.15배, 미국 의회도서관정보(20TB)의 19.3배에 달하는 규모이다. 또한 심평원의 진료이용내역 등의 빅데이터는 의료이용 · 의약품 유통 · 의료자원 현황정보가 시간의 흐름에 따라 수집 · 저장되어 그 활용가치가 매우 높다는 평가이다.

건강보험공단은 건강검진, 보험료, 진료내역 등 8,000억 건이 넘는 국민건강정보DB를 활용하여 국민건강보험 코호트DB를 마련하여 제공하고 있으며, 그 데이터의 개방에 대해서는 논의 중이다.

심평원과 건강보험공단의 빅데이터는 전 국민을 대상으로 한 정보로서 대규모 데이터 분석을 통해 학술연구, 보건의료정책 및 관련 산업 활성화 등 다양한 방면에서 활용될 수 있어 그 잠재성이 무궁무진하다는 평가이다.

표 2. 심평원 빅데이터 시스템



아직 초기단계이나 국내 많은 병원에서 다양한 빅데이터 연구 분석이 진행 중이다. 서울대병원은 맞춤형 건강관리 프로그램 '헬스온 샤인(Health-On Shine)'을, 세브란스병원은 '후헬스케어(H∞H Healthcare)' 서비스를 개발하였다. 가톨릭대학교의과대학은 '아이러브인슐린(www.iloveinsulin.co.kr)' 서비스, 빅데이터 네트워크를 만들고 홈 헬스케어 사업을 추진 중에 있다.

심평원은 연구중심 10개 대학병원과 MOU를 체결하여 빅데이터를 제공하고, 연구개발을 지원하고 있다. 특히 우리나라 의료분야의 빅데이터는 전 국민을 대상으로 한 환자진료기록, 임상데이터, 의료영상 이미지 뿐 아니라 유전자 통계, 질병정보, 생활 · 소득 · 환경 정보 등 그 범위가 방대하다. 정형화된 정보 뿐 아니라 비정형 정보를 활용할 경우에 의료진의 임상적 의사결정, 맞춤형 의료제공, 유전성 질환의 조기발견, 예방과 질환 관리, 전염성 질환 등 공공영역에서 보건의료 혁신과 의료의 질 향상에 활용할 가치가 매우 높다.

이렇듯 빅데이터는 의료서비스 분야에 많은 기여를 할 수 있다. 생산성 측면에서는 만성질환 관리와 진료프로세스 효율화에, 발견(예측)측면에서는 질병예방 · 게놈(유전)질환정보 해독이나 맞춤화된 질병관리에, 새로운 고객가치 차원에서는 환자들 스스로에 의한 의료서비스와 건강관리 및 온라인 의료관리 등에 기여할 것이다(Booz & Company, 2011).

대형병원을 위주로 차세대 전자의무기록(Electronic Medical Record, EMR) 도입과 임

상연구를 위한 임상 DW 빅데이터 플랫폼 구축을 진행하고 있어 심평원 및 건강보험공단 빅데이터와 연계 · 활용될 경우, 보다 효과적인 진료와 임상 연구개발을 촉진시킬 수 있을 것으로 기대된다.

예컨대 건강위험요인 예측모델 구축과 맞춤형 건강관리서비스 등을 통한 연구 성과는 국민의 건강증진에 도움을 줄 수 있을 것이고, 생체신호 감지기술과 웨어러블 스마트기기 등 첨단기술을 기반으로 한 의료 IT는 또 다른 빅데이터를 만들 수 있으며, 이는 내 새로운 산업과 일자리 창출을 가속화 할 수 있을 것이다.

한편, 심평원 · 건강보험공단 · 병원 · 질병관리본부 등은 각각 보건의료 · 건강정보를 보유하고 있어 이를 공유 · 연계하면 빅데이터의 유용성과 활용성을 크게 높을 수 있다. 이에 따라 기관 간 정보공유 · 연계 작업이 점차 활발해지고 있다.

나. 정부 정책동향과 발전전략

2011년 국가정보화전략위원회는 DNA/의료데이터 융합시스템 구축 등 맞춤형 의료시대를 통한 국민 삶의 질 향상을 위해 '빅데이터를 활용한 스마트 정부 구현(안)'을 발표하였다.

2012년 관계부처 합동으로 '빅데이터 마스터플랜'을 통해 범정부 추진체계 마련, 데이터 개방 · 공유 확산 및 데이터 전문 인력 양성을 추진하고 있다. 보건의료 관련하여 '유전자/의료데이터 분석을 통한 국민 건강 증진'을 위해 DNA/진료기록 · 진료정보의 융합 분석을 통한 건강진단과 질병 · 발생 가능성 예측, 암 · 당뇨 · 비만 등 100여개 질환의 위험도 예측 등 맞춤형 예방 서비스 제공을 추진하고 있다.

이에 따라 2013년 미래창조과학부는 보건의료 분야에서 질병예방 서비스, 의약품 안전성 조기경보 서비스, 심실부정맥 예측 · 인플루엔자 위험 예측 서비스, 입원병상 배정 서비스 등 빅데이터를 활용한 인터넷 신산업을 본격 활성화하기 시작하였다.

2014년 8월 기획재정부는 '보건의료 분야 투자 확대계획'을 통해, 공공기관들이 보유하고 있는 국민건강 관련 정보를 연계하는 DB구축, 의료정보 보호 · 교류 · 활용 체계 구축 계획을 밝혔다.

2014년 9월 관계부처 합동 '공공데이터 개방 발전전략' 에서, 2017년까지 데이터 개방 5대 강국 진입을 목표로 보건의료 분야에서도 고수요 데이터 우선 개방 및 공공데이터 품질 강화, 민간 · 공공 상생의 데이터 생태계 조성 사업을 밝혔다.

국가질병관리 DB, 식의약 DB, 심평원의 국가 의료정보 DB 등을 국가 중점 개방 데이터로 개방하고, 활용 기반을 조성하는 사업 등을 진행한다.

이어 「정부3.0」과제로 심평원, 질병관리본부 등 관계기관 간 DB 공유 및 협업, 데이터 연계 활용한 맞춤형 서비스 제공 등이 제시되고, 현재 그 구체화 작업을 추진 중이다.

3. 보건의료 빅데이터 활성화와 과제

보건의료 분야에서 빅데이터의 공유와 활용 가치는 무궁무진하다. 하지만 빅데이터를 제대로 활용하기 위해선 "왜, 무엇을 위해, 어떻게 할 것인가"에 대한 논의가 필요하다.

보건의료 빅데이터를 개방하고 활용하기 위해 심평원의 '보건의료빅데이터센터' 운영 경험과 사용자의 요구 수요를 감안하여, 빅데이터 활성화 방안과 과제를 '정책적 이슈, 관리적 이슈와 기술 차원의 이슈'를 중심으로 제시해 본다.

가. 전략적 접근을 통한 '개방'에서 '활용' 체제로

심평원의 보건의료 빅데이터의 유용성은 매우 크고, 고가치 · 고수요 데이터로 평가되고 있다(안전행정부, 2014). 그러나 단기간 내 데이터 개방의 확대와 민간 · 학계 등의 이용 증가세에도 불구하고, 양적 개방 확대에 치우쳐 보건의료 이용서비스나 산업 분야에서의 활용도는 아직 저조하다. 이는 심평원이 선정한 개별 데이터 중심으로 개방되어 활용이 어렵고 데이터 구조 및 양식이 민간에서 바로 쓰기 어려우며 개방 데이터 검색 · 활용 절차가 복잡하기 때문으로 평가된다. 또한 제공되는 데이터는 APP 개발에는 유용하나 웹 서비스 개발과 데이터 분석 · 활용에는 제한적이고, 민간 · 학계에 대한 실질적 활용 지원이 미흡한 것으로 평가된다.

심평원 데이터는 국민의 일상생활과 밀접하고, 비즈니스에 직접 활용될 수 있는 일정 수준의 품질을 갖춘 국가 대용량 데이터이다. 개인정보 · 민감 정보 등의 구성 특성상 개방이 불가하더라도, 마스킹 처리 등을 통해 개방 가능한 항목을 선정하여 개방할 수 있다고 본다. 또한 심평원 · 건강보험공단 · 복지부 · 질병관리본부 · 식약청 등이 각각 보유하고 있는 보건의료 · 국민건강 정보 등은 연관성 높은 데이터로 연계 · 융합하면 대용량 데이터가 생성되어, 개방 시 활용성과 가치창출에 매우 효과적일 것이다. 이 경우 기관마다 다른 기준으로 개방하는 데이터에 대해 핵심 데이터셋(dataset)의 표준화와 통합DB를 구축해야 하는 과제가 있다.

이를 위해 정부와 민간이 공동으로 공공데이터 개방표준을 조기에 마련하고, 공공데이터 품질보장 체계를 마련하는 것이 시급하다.

데이터셋은 그리드(grid)방식으로 개방하고, 오픈 API(Application Program Interface)로 제공하여 민간에서 쉽게 활용하게 할 필요가 있다. 이에 심평원은 오픈 API 구축 및 시스템 보완을 거쳐 내년 상반기 중 오픈 API 및 메타데이터 형태로 개방할 계획이다.

데이터 품질 진단을 거쳐 데이터 품질개선, 표준화, 컨설팅, 데이터 개방 및 활용과 관련한 각종 전문기술을 육성하여야 하고, DB구축-운영-활용 단계별 매뉴얼을 마련하여 데이터 품질을 확보하고 활용성을 높이는 과정도 요구된다. 또한 공공데이터의 자유로운 활용을 지원하는 오픈형 데이터 포털로 전환하여 개방하며, 이와 함께 공공데이터의 적시 제공과 개방 절차의 간소화 및 합리적인 이용요금의 기준을 마련할 필요가 있다.

뿐만 아니라 빅데이터의 목표와 방향성을 정립하고, 마스터플랜을 수립하는 것 역시 중요하다. 또한 빅데이터는 보건의료 패러다임의 전환과 현안과제를 해결하고, 정부 3.0 등을 지원하는 유용한 수단이나 정부의 정책이 더해져 서로 뒷받침되어야 문제를 해결할 수 있다.

또한, 건강보험 데이터에는 비급여 진료내역이 포함되지 않고, 진료 과정이나 결과 등 임상진료 내역이 반영되지 못하는 한계가 있다. 이 역시 정부에서 비급여 진료내역 수집, EMR 확대를 통한 진료정보 교류 절차와 환자진료정보의 온라인 플랫폼화 등에 필요한 정책과 법령을 마련해야 건강보험 데이터가 고품질의 완전한 빅데이터로서 기능할 수 있다.

나. 빅데이터 활용과 활성화를 위해 풀어야할 과제

보건의료데이터는 가장 민감한 개인정보이자 기밀 영역이면서도 대량 수집되어야 유용하다. '개인정보 보호'와 '데이터 활용의 자유'가 충돌되는 대표 분야로서, 이는 빅데이터의 잠재가치 이면에 있는 문제이자 활용 위기의 단서이다.

개인정보 보호(프라이버시)문제는 심각하다. 특히 빅데이터로 보편화될 맞춤형 서비스는 개인정보 보호와 이용 측면에서 문제가 크다. 또한 데이터가 많아질수록 이를 이용해 고객 · 환자의 실제 현상을 조작하는 '데이터 남용' 문제가 발생할 수 있다. 그리고 데이터 수집 · 공유 및 활용 각 단계에서 정보보호의 문제는 물론 데이터의 소유권, 사용권과 공개된 데이터의 지적 재산권에 대한 논란 등 데이터 소유자 · 활용자 및 이용자가 모두 얽히는 복잡다단한 문제를 제기한다.

현재 보건의료 빅데이터 관련 제도와 법은 의료법, 개인정보보호법, 공공데이터의 제공 및 이용 활성화에 관한 법률, 보건의료기본법, 건강보험법 등등으로 분산되어 있다. 법령 간 충돌 · 배치되거나 해석을 곤란하게 하여 빅데이터의 수집 · 활용 · 공유 등에 있어 발목을 잡는 실정이다.

의료분야에서 빅데이터 활용을 위해서는 이러한 문제에 대한 엄격하고 분명한 가이드라인 설정과 법제도 정비가 우선되어야 한다. 개인정보와 기밀정보에 대한 보안, 공익과 개인정보보호의 조화로운 운영, 데이터 집중과 오남용에 따른 우려 불식 또한 필요하다.

보건의료의 특수성과 개인정보 보호의 필요성 등을 종합하여 「건강정보 보호 및 활용 법률」등을 제정하여 정보 수집 · 보안 · 교류 및 활용범위, 환자의 동의 절차 등을 시급히 명시하는 것이 문제해결의 과제이자 빅데이터의 활용과 가치를 높임과 동시에 빅데이터 활용에 대한 인식개선의 출발점이 될 것이다.

성공적인 빅데이터 활용을 위해 ① 활용할 기관 내 · 외부에서 빅데이터 자원의 지속적인 발견과 확장 ② 진화하는 빅데이터 플랫폼, 분석기술 등 프로세스와 신기술 이해 ③ 조직 차원에서 데이터 과학자 등 인재 확보를 위한 내부 역량강화 및 외부 협력 전략 마련 역시 필요하다.

4. 나가며

보건의료 분야는 급속한 고령화와 의료 패러다임의 전환에 따른 적응과 더불어 건강보험과 보건의료의 지속가능성, 병의원과 의료기기 및 의약품 등 의료산업전반과 보건의료 생태계 조성을 위한 변화 · 혁신의 도전이 거세다.

선진국들은 보건의료 빅데이터를 바탕으로 공공 · 민간부문의 데이터 공개와 연계를 통해 보건의료 분야의 생산성과 효율성을 통한 국민건강 향상과 비용절감, 기반인프라 확충을 바탕으로 한 새로운 가치창출을 선도하면서 이를 신성장동력으로 삼아 각축을 벌이고 있다.

우리나라는 세계 최고 수준의 ICT와 우수한 보건의료인재 등 빅데이터 활용 측면에서 경쟁우위에 있고, 심평원이나 정부 등 공공부문 역시 보건의료 혁신과 새로운 생태계 조성에 발 빠르게 대응하고 있다. 그러나 아직은 빅데이터 활용 초보 단계로서, 부가가치 창출을 이끌어 내기 위해서는 보건의료 빅데이터 마스터 플랜과 전략을 마련하고, 거버넌스 체계를 구축하면서 전문 인력 육성이 필요하다고 본다. 아울러, 보다 혁신을 위해서는 빅데이터 중심의 보건의료 생태계 조성을 위해 민간 · 공공의 협업과 참여를 통한 '집단지성'의 구현이 요구된다. 빅데이터 활용에 있어 미비한 법 · 제도의 정비, 개인정보 보안 기술의 개선, 보건의료정보 활용에 대한 인식개선 등이 이루어질 수 있도록 사회적 공감대 마련과 국가 의제화를 제안한다.

"구슬이 서 말이라도 꿰어야 보배"라는 의미는 보건의료 분야의 빅데이터 활용 현실과 그 미래를 위한 우리의 선택이라 본다.

참고문헌

건보공단, 건강보험빅데이터 활용방안 모색을 위한 세미나 자료, 2014.7.2.
공공데이터전략위원회, 공공 데이터 개방 발전전략, 서울: 안전행정부; 2014.
백승민, 빅데이터를 활용한 신산업 전망 및 정책 방향, 보건산업 브리프 Vol. 84
서울대 빅데이터센터 · 매일경제기획팀, 빅데이터 세상, 서울: 매일경제신문사; 2014.
심평원, 공공데이터 개방 관련 보건의료 빅데이터 활용(내부자료), 제30회 심평 포럼, 2014.9.
심평원, 보건의료빅데이터센터 안내(유인물), 서울: 진료정보분석실; 2014.
심평원, 보건의료성과향상을 위한 데이터 활용 국제심포지엄, 2014.6.12.
심평원, HIRA 빅데이터 민간활용 활성화 방안, 제30회 심평 포럼, 2014.1.21.
양금덕, 국민건강정보빅데이터산업화 길 머지 않아, 청년의사, 2014.9.2.자
한국데이터베이스진흥원, 데이터분석 전문가 가이드, 서울 · 한국데이터베이스진흥원; 2014.
한국정보화진흥원, 보건의료정보를 활용한 맞춤형 서비스 개발 컨설팅 결과보고서, 2014.1.
한국정보화진흥원, 새로운 미래를 여는 빅데이터 시대, 서울: 한국정보화진흥원; 2013.
한국IT서비스산업협회, '빅데이터 활용한 보건의료산업 활성화를 위한 사업 및 제도개선 방안 연구보고서', 2014.7.
함위근 · 채승병, 빅데이터 경영을 바꾸다, 서울: 삼성경제연구소; 2012.