

# 의료기관 빅데이터 품질관리의 필요성과 사례 분석\*

## The Necessity and Case Analysis of Bigdata Quality Control in Medical Institution

최혜린<sup>1</sup> · 이승원<sup>2</sup> · 김영아<sup>3</sup> · 이종호<sup>3</sup> · 고 흥<sup>4</sup> · 김현창<sup>2,3\*</sup>

연세대학교 대학원 보건학과<sup>1</sup>, 연세대학교 의과대학 예방의학교실<sup>2</sup>  
연세대학교 의료원 정밀의료데이터사이언스ICT센터<sup>3</sup>, 연세대학교 의료원 소아과학교실<sup>4</sup>

### 요 약

빅데이터의 활용은 사회 전 분야에서 중요한 역할을 하고 있으며, 특히 보건의료분야에서는 국민의 생명과 건강을 다루기 때문에 빅데이터의 역할이 더욱 중요하다. 하지만 의료 빅데이터의 품질관리에 대한 관심과 인식은 현저히 떨어지는 실정이다. 저 품질 의료 빅데이터는 국가적 손실과 국민의 건강 저해를 야기시키므로 의료 빅데이터의 품질관리가 필요하다. 이에 국내의 의료 빅데이터 품질관리 사례 및 가이드라인에 대하여 문헌 조사하여 국내 의료 빅데이터 품질관리에 대한 방향성을 제시하고자 한다. 또한, 국내 한 대형 의료기관의 의료 빅데이터 품질관리 사례로 Y의료원의 ‘빅데이터 품질관리 TFT’ 활동과 데이터 관련 업무종사자 대상 설문조사 결과를 소개하고자 한다.

■ 중심어 : 의료, 빅데이터, 품질관리

### Abstract

The use of Bigdata plays an important role in all areas of society. Especially in the health care field, the role of Bigdata is very considerable because it deals with people's life and health. However, the interest and awareness of quality control of medical data is markedly low. Because the low-quality medical Bigdata leads to national loss and public health impairment, quality control of medical Bigdata is needed. The purpose of this research is to present the direction of medical Bigdata quality management by examining literature and cases of domestic and foreign medical Bigdata quality management practices. In addition, as a case of medical Bigdata quality control in the Y medical institution in Korea, activities of a Bigdata quality management TFT and results of a survey conducted for major data users in the hospital were presented.

■ Keyword : Medicine, Bigdata, Quality Control

## I. 서 론

빅데이터(Bigdata) 분석은 미래를 더 잘 예측하여 각종 분야에서 발생 가능한 손실을 줄이고 경영

효율성을 높일 수 있기 때문에 빅데이터가 다방면에서 활용되고 있다. 보건의료 분야는 데이터의 양(Volume), 다양성(Variety), 복잡성(Complexity) 등이 높아 금융, 통신과 더불어 대표적인 빅데이터

2017년 11월 23일 접수; 2017년 12월 22일 수정본 접수; 2017년 12월 26일 게재 확정

\* 이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2017-0-00163, 빅데이터 품질 평가 도구 개발).

† 교신저자 hckim14@gmail.com

생산분야에 속한다. 또한, 보건의료는 약간의 효율성 개선 만으로도 비용절감 효과가 크며, 그에 따른 비경제적 파급효과도 크기 때문에 빅데이터 분석이 매우 중요하며 필요한 분야이다. 특히 보건의료 빅데이터 품질 저하는 국민의 생명과 건강을 위협할 수 있는 심각한 문제를 유발하고, 보건의료 시스템의 효율성을 저하시킨다. 하지만 의료 빅데이터의 품질관리에 대한 인식률은 아직까지 매우 낮으며, 품질관리에 대한 모니터링 및 가이드라인 제시 등의 구체적인 활동은 이루어지지 않고 있는 실정이다[9]. 의료 빅데이터는 진료, 수술, 연구, 정책결정 등 다방면으로 사용되기 때문에 저 품질의 데이터를 사용하는 경우 그에 따른 손실이 매우 클 것으로 추정된다.

따라서, 우리나라의 의료 빅데이터 품질관리의 중요성과 필요성을 강조하기 위하여, 국내외 문헌 조사를 통해 의료 빅데이터 품질관리의 현황과 가이드라인, 품질관리 실패사례와 그의 문제점 등을 조사하여 의료 빅데이터의 품질향상을 도모하였다. 문헌조사로 각 나라의 보건의료국가기관의 연구자료, 논문, 가이드라인 등을 정리하였다. 정리한 내용 바탕으로 우리나라 의료 빅데이터 품질관리에 대한 방향성을 제시하고자 하였다. 또한, 국내 한 대형 의료기관(Y의료원)의 의료 빅데이터 품질관리 개선 활동을 위하여 Y의료원 의료 빅데이터 품질관리 TFT(Task Force Team)을 구축하였다. TFT은 원내 빅데이터 품질관련 문제확인 및 개선 등 데이터 품질관리를 위한 활동을 주로 하며, Y의료원의 의료 빅데이터 품질관리 현황을 알기 위하여 설문조사를 실시하였다.

## II. 본 론

### 2.1 문헌 조사

#### 2.1.1 의료 빅데이터 품질관리의 필요성

의료기관에서 의료 빅데이터 활용의 1차 목적은 환자에게 진료서비스를 제공하기 위함이다.

이는 환자와 의료인, 환자관리, 경영지원, 병원의 재정 및 행정을 위하여 사용된다. 2차 목적은 교육, 법률, 연구, 공중보건, 정책입안, 산업 등 보건의료 관련 분야에 활용하기 위함이다. 최근 데이터를 이용하여 정보 및 지식을 창출하던 시기에 더 나아가 서비스를 개발하고 평가하는 시기로 변화하였으며, IT(Information Technology)에서 DT(Data Technology) 시대로 변화함에 따라 의료 빅데이터의 활용 영역도 확대되고 있다[4].

의료 빅데이터 활용 영역의 변화에 따라 양질의 의료 빅데이터 기반 활용의 필요성이 증대되었으며, 의료 빅데이터를 효율적으로 활용하기 위해서는 우선적으로 의료 빅데이터의 질적 수준을 향상시켜야 한다는 중요성이 강조되고 있다.

#### 2.1.2 의료 빅데이터 품질관리의 실패

2016년 ECRI(Emergency Care Research Institute) 보고서[12]에 따르면 의료 빅데이터 품질관리 실패 유형은 네 가지로 나뉜다. 첫 번째 유형은 데이터관리 시스템과 작업의 흐름이 맞지 않은 경우이다. 예를 들어, 만약 환자의 알리지 여부, 복용 약물 및 검사, 치료 등에 대한 최신 정보가 조직 내에 신속하게 전달되지 않는다면 치료가 지연되거나 실패하는 문제가 생기게 된다.

두 번째 유형으로는 환자 등록 시 잘못된 정보가 입력되거나, 다른 인물과 정보가 바뀌는 등의 환자식별 오류이다. 성(last name)과 이름(first name)을 따로 쓰는 미국의 경우 이 둘을 바꾸어 쓰는 실수가 자주 발생한다. 이 오류는 환자의 안전과 밀접한 관련이 있으며 치료나 예후에도 영향을 미칠 수 있기 때문에 patient index를 만들어 관리하는 등의 조치를 권고하고 있다[6].

세 번째로는 검사 결과의 보고를 제 때에 하지 않는 경우와 추적조사를 실패하는 경우이다. 중요한 검사결과를 보고하지 않는다면 후속조치를 하는 과정에서 환자의 안전과 건강에 해가 될 수 있다. 또한, 환자가 검사 결과가 나오기 전

에 퇴원을 하는 경우나 의료서비스 제공자 간의 의사소통이 부족한 경우, 환자의 정보를 분실한 경우 등에는 추적조사를 실패하게 되어 의료 빅데이터의 품질을 저하시킨다.

마지막으로 데이터의 단위 혼용과 관련된 실패 사례를 들 수 있다. 예를 들어 체중의 단위인 파운드와 킬로그램이 혼용되는 경우 단위를 바꾸어 잘못 기록하는 오류가 발생할 수 있다. 30 파운드인 아이의 체중을 30킬로그램(66파운드)으로 기록되어 약물처방이 잘 못된 사례를 예로 들 수 있다. 이러한 잘못된 체중 정보로 인해 환자가 위험해질 수 있으므로 데이터 관리 실패가 환자의 안전에도 영향을 미친다는 것을 확인할 수 있다.

미국의 MRI(Medical Record Institute)에 따르면, 질적 수준이 낮은 저 품질 의료 빅데이터가 환자안전(patient safety), 공공안전(public safety), 진료의 연속성(continuity of patient care), 보건경제(health care economics), 임상연구 및 성과(clinical research and outcomes) 등 다양한 측면에서 문제를 유발한다고 한다. 저 품질 의료 빅데이터가 영향을 끼치는 분야가 다양한 만큼 그에 따른 경제적 손실도 무시할 수 없다. 미국 전체의 손실 규모는 연간 약 660조에 달하며 [11], 개별 의료기관 별 손실액은 의료기관 수입의 약 15% 수준으로 추정된다[16].

2010년 한국데이터베이스진흥원에서 공공, 금융, 통신/미디어, 유통/서비스, 제조, 의료 등 6개의 산업군에 속하는 317개의 공공기관 및 민간 기업을 대상으로 데이터 품질관리 성숙 수준을 조사하였다. 국내 기관 및 기업의 평균 데이터 품질관리 성숙 수준은 1.0 레벨이며 금융은 1.7 레벨, 통신은 1.2 레벨로 2009년도 대비 각각 0.5, 0.3 레벨이 상승해 비교적 높은 수준으로 파악되었다. 하지만 공공기관과 의료기관의 경우 전체 평균보다 낮은 0.9 레벨로 3년 연속 최하위 그룹에 속하는 것으로 보고되었다. 이는 우리나라

의료데이터의 품질관리 수준이 가장 기초적인 도입단계에도 이르지 못한 수준임을 뜻한다 [5]. 또한, 우리나라 의료기관의 병원 및 의원급의 EMR 도입율은 92.1%(전반적 이용 58.0%, 부분적 이용 34.1%)로 대부분의 의료기관이 높은 전산화 수준을 보이지만[1], 의료기관의 의료 빅데이터 품질관리를 위한 활동은 매우 미흡하다. 국내에서 의료정보 수준이 높다는 평가를 받는 병원의 EMR 데이터 유용성에 대한 연구에 따르면 EMR 데이터의 88.6%가 형태나 의미적 측면에서 적합성을 갖추고 있는 것으로 나타났다. 그에 비해 11.4%는 형태나 의미적 측면에서 부적합하였으며, 다른 병원 및 의원의 EMR 데이터의 질적 수준은 이 연구의 대상 병원에 비해 낮을 것으로 판단된다[4].

보건의료 데이터 품질관리 업무의 문제점으로는 500병상 이상 급 의료기관의 독립된 의무기록실에서 주로 데이터에 대한 품질관리 업무가 이루어진다는 것이다. 비교적 큰 규모의 병원에서 의무기록실은 하나이며, 하나의 의무기록실이 전체 병실의 데이터를 관리하기 때문에 품질관리의 효율성과 정확성이 떨어질 수밖에 없다. 그리고 의료 빅데이터에 대한 품질관리 사업을 하는 의료기관에서는 EMR의 전 데이터의 품질을 관리하는 것이 아니라 대부분 퇴원요약자료만을 대상으로 하며, 진료가 종료된 이후에 작업하는 후향적 방식으로 이루어지고 있다. 현재 운영되고 있는 품질관리 부서의 경우, 건강보험심사평가원의 청구와 관련된 질 지표 관리나 관리적 프로세스 개선에 중점을 두고 있다. 병원정보시스템의 의료 빅데이터를 총체적으로 품질관리하는 것에 목적을 두지 않기 때문에 의료 빅데이터의 품질을 관리하는 순기능을 실행하지 못하고 있다. 이 부분에 대하여 미국의 The Joint Commission에서는 진료시점에 기록하고 모니터링하는 것을 강조하고 있지만 국내에서는 국제적인 기준에 부응하지 못하는 실정이다[2].

의료 빅데이터를 관리하기 위한 품질관리 시스템이 없기 때문에 수작업으로 데이터를 검토하고 개선하는 업무 방식이 주로 사용된다. 이에 따라 의료기관에서 의료 빅데이터 품질관리 사업을 한다 해도 효율성과 정확성이 떨어진다. 실제 병원정보시스템의 의료 빅데이터를 이용한 2차 자료 연구 시 자료추출, 검증, 수정, 데이터 분석, 활동 단계를 거치는데, 이 중 데이터 검증에 가장 많은 시간이 소요되는 실정이다[2].

국내 의료 빅데이터의 품질관리 수준이 낮은 이유는 기관의 CEO를 비롯하여 전체 조직원의 데이터 품질관리에 대한 인식 부족을 가장 큰 이유로 들 수 있다. 중요성에 대한 인식이 저조하기 때문에 많은 의료기관에서 의료 빅데이터 품질관리를 위한 전문 인력이나 조직, 예산 확보가 충분히 이루어지지 않고 있다. 의료 빅데이터 품질관리 조직이 따로 없기 때문에 의료기관 내부 및 외부사용자의 데이터 사용에 대한 요구사항을 제대로 파악하지 못하여 데이터에 문제가 있어도 개선하지 못하는 경우가 많다. 품질관리 전문가의 부족으로 인해 데이터의 생성과정이 명확하게 정의되지 않고 데이터 간의 연계성 관리가 제대로 이루어지지 않는 점도 의료데이터의 품질을 저하시키는 요인이다[5].

따라서 EMR 등 병원정보시스템에 입력되는 의료 빅데이터의 품질관리에 대한 구체적인 관리방안 및 전략, 전문인력이 부족한 것을 전반적인 문제점으로 들고 있다. 이 문제점들을 해결하기 위해서는 데이터를 표준화된 형태로 수집하기 위한 데이터 수집방안이 마련되어야 하며, 의료 빅데이터의 정확성과 완전성에 대한 필요성을 인식하고 이해를 돕는 교육과 전체적인 가이드라인이 개발되어야 한다.

### 2.1.3 현재 의료 빅데이터의 품질관리에 대한 가이드라인

세계보건기구(WHO)[14]에서는 2003년에 보건 의료데이터 품질 향상을 위한 가이드라인을

발간하였다. 가이드라인은 직급별 의료종사자와 의료정보 관리자에게 일련의 지침을 제공하여 데이터의 수집 및 생성과 관련된 모든 분야에 대한 이해를 돕는다.

가이드라인에 따르면 데이터의 품질관리 방안은 데이터의 중요성(importance), 정확성(accuracy), 타당성(validity), 신뢰성(reliability), 완전성(completeness), 가독성(legibility), 적시성(currency and timeliness), 접근성(accessibility), 비밀보장성(confidentiality and security) 을 고려하여 개발된다. 데이터의 표준 문서 구현방안, 구성요소 데이터 및 정보의 정의, 임상정보의 관리 및 품질 개선방안 등의 관리 방안을 구분하여 각 대상별 관리 목적, 기준, 방법, 절차에 대해 설명하였다. 이는 관리 대상별로 산출물 샘플을 명시하여 관리 방법에 대한 이해를 도왔다.

의료 빅데이터 웨어하우스와 영국의 보건복지정보센터(HSCIC, Health and Social Clinical Information Centre), 미국의 보건정보관리자협회(AHIMA, The American Health Information Management Association)에서도 보건의료 데이터에 대한 품질관리 방안에 대하여 제시하였다. 의료 빅데이터 웨어하우스에서는 품질 향상을 위해 데이터의 정확성(accuracy), 완전성(completeness), 일관성(consistency), 적시성(timeliness), 적절성(relevancy), 접근성(accessibility)이 보장되어야 한다고 보고하였다[7]. 영국의 HSCIC는 의료 데이터가 국가의 표준을 만족하는지에 관한 평가와 모니터링을 통해 지속적인 개선을 지원한다. 데이터 품질 교육 및 훈련 제공을 통해 품질관리의 인식을 높이고 정부가 나서서 데이터 관리의 투명성을 보장한다[3]. 미국의 AHIMA는 WHO의 의료 빅데이터 품질관리 가이드라인과 비슷한 품질관리 모델을 가지고 있다. 품질관리 도메인에는 수집(collection), 저장(warehousing), 분석(analysis), 적용(application)이 속한다. 이에 따른 데이터 품질 향상을 위한 요소로 정확성

(accuracy), 접근성(accessibility), 포괄성(comprehensiveness), 일관성(consistency), 현재성(currency), 정의(definition), 세분화(granularity), 정밀도(precision), 적절성(relevancy), 적시성(timeliness)을 제시하였다[15].

가이드라인에 근거한 비용-효과적인 의료 빅데이터의 품질관리가 필요하다. 이는 오류 없는 완전한 데이터를 지향하는 것이 아니라[10] 적정 비용으로 적정 수준의 의료 빅데이터 품질관리를 의미한다[13]. 품질관리의 목적, 범위, 내용을 명확하게 정의한 후 이를 기반으로 접근하는 것이 필요하며 데이터 생성 및 관리자들은 통일된 명확한 기준에 따라 표준화된 데이터 품질관리 작업을 해야 한다[8]. WHO의 가이드라인뿐만 아니라 여러 나라의 품질관리 방안과 마찬가지로 국가 차원의 의료 빅데이터 품질관리 가이드라인을 개발하고 체계적으로 접근할 필요가 있다.

## 2.2 의료 빅데이터 품질관리 활동 사례

앞서 국내 의료 빅데이터 품질관리에 대한 실패사례 및 문제점을 국내의 문헌조사를 통해 살펴 보았다. 국내 대형 의료기관 중 하나인 Y의료원 내에 의료 빅데이터 품질관리의 현황을 파악하고 이를 개선하기 위하여 데이터 실제 사용자 및 관리자를 대상으로 의료 빅데이터 품질관리 TFT를 구축하였으며, 주요 사용자 및 관리자 대상 설문조사를 실시하였다.

### 2.2.1 의료 빅데이터 품질관리 TFT 구축

의료 빅데이터(OCS(Order Communication System), EMR, 영상데이터, 유전체 등)를 활용한 연구 및 사업화에 대한 수요가 지속적으로 증가하고 있다. 이에 따라 데이터 품질관리의 중요성과 필요성도 함께 증가하고 있다. Y의료원에서는 의료 빅데이터의 품질관리 및 질 향상을 도모하기 위해 의료 빅데이터 품질관리 TFT를 구축하였다. 이 TFT는 의료 빅데이터와 밀접한 관련이

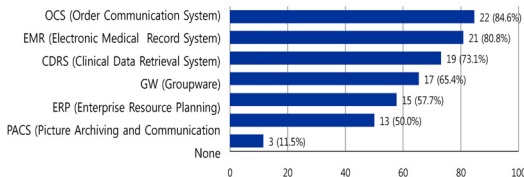
있는 의료정보실, 적정진료관리팀, 의무기록팀, 보험심사팀과 간호사, 의료진 등 총 14명으로 구성되었다. 이들의 역할은 빅데이터 품질관리 및 품질 향상과 관련 문제에 대해 확인하고 의논하는 것을 전반적으로 담당한다. 세부적으로는 의료 빅데이터 품질관리의 대상을 선정하고, 상시 모니터링, 결과분석 및 개선활동 논의, 데이터 기반 실패사례 분석, 개선활동 계획 수립, 실제적인 수행 전략 수립 및 이행, 관련 가이드라인 개발 등을 담당하고 있다. 또한, 환자의 안전 및 의료의 질 향상을 위하여 Y의료원 정밀의료 빅데이터를 활용하기 위한 방안을 모색한다.

그 첫 번째 활동으로 빅데이터 활용을 위한 접근 측면에서 익명화 및 개인정보 비식별화 프로세스를 개발하고 있다. 환자의 개인정보와 민감정보, 고유식별정보를 구분하여 익명화와 비식별화의 범위를 정하는 것이 의료 빅데이터 활용 이전에 시행되어야 하는 중요한 단계이다. 이는 연구와 같이 여러 활동을 진행함에 있어서 제약조건이 될 수 있으나 기관과 사용자를 위해 반드시 준수해야 하는 전제 사항이다. 따라서, 의료데이터를 활용한 학술연구, 질 향상 및 업무 개선 활동, 외부기관과의 협업 등을 위하여 데이터 활용 및 공유를 위한 일종의 가이드라인 마련을 위한 활동을 하고 있다.

### 2.2.2 의료 빅데이터 주요 사용자 및 관리자 대상 설문조사

Y의료원 내에 의료 빅데이터 주요 사용자와 관리자를 대상으로 2017년 9월 26일부터 10월 27일까지 설문조사를 진행하였다. 설문조사의 주요 내용은 현재 이용하고 있는 의료 빅데이터의 오류로 인해 불편했던 경험 및 개선할 사항에 대해 다루었으며 총 13개의 문항으로 구성되었다. 설문 응답자는 총 26명이며 의사 13명(50%), 연구원 4명(15.4%), 의무기록사 4명(15.4%), 간호사 2명(7.7%)이 답하였다.

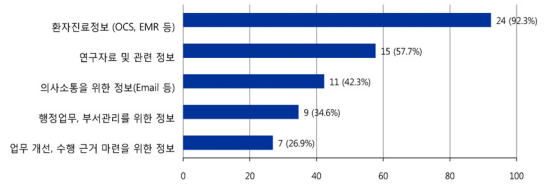
설문 응답자 중 23명(94.5%)이 병원업무를 수행하기 위해 정보시스템을 활용 중이었으며, 가장 많은 사용하는 정보시스템은 OCS(84.6%)였으며, 그 다음으로 EMR(80.8%), GW(Groupware, 65.4%), CDRS(Clinical Data Retrieval System, 73.1%), PACS(Picture Archiving and Communication System, 57.7%), ERP(Enterprise Resource Planning 또는 행정시스템, 50%) 순이었다(<그림 1> 참조).



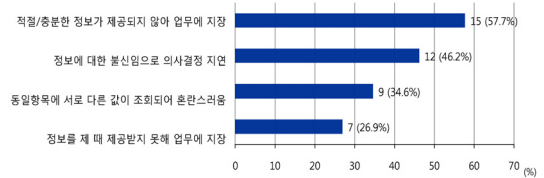
<그림 1> 정보시스템 이용률

응답자가 병원업무를 수행하기 위해 가장 중요하다고 생각한 정보는 환자진료와 관련된 정보(OCS, EMR 등, 92.3%)였고, 그 다음으로 연구자료 및 관련정보(57.7%), 이메일 등 의사소통을 위해 활용되는 정보(42.3%) 순이었다(<그림 2> 참조). 현재 업무에 이용되는 정보 중 신뢰도가 낮은 정보가 있다고 응답한 사람은 12명(46.2%)으로, 그 중 다수가 시간정보와 대상자의 자가 응답에 의한 설문정보가 신뢰도가 낮다고 답하였다.

의료 빅데이터의 품질관리 미비로 인해 응답자가 겪은 어려움에 대한 항목에서는 응답자 중 적절·충분한 정보가 제공되지 않아 업무에 지장을 경험한 사람은 15명으로 57.7%를 차지하였고, 원내 시스템에서 동일 항목, 조건에 대해서도 다른 값(Value)이 추출 또는 조회되어 혼란을 경험한 사람은 12명으로 46.2%, 정보에 대한 불신임으로 의사결정이 지연된 경험을 가진 사람은 9명으로 34.6%, 요청한 정보/필요한 정보를 제 때에 제공받지 못해 업무에 지장을 초래한 경험이 있는 사람은 7명으로 26.9%를 차지하였다(<그림 3> 참조).



<그림 2> 중요하다고 생각되는 정보



<그림 3> 데이터 품질 미비로 인한 어려움

의료 빅데이터 품질의 저하로 업무수행에 어려움을 겪은 응답자가 답변한 구체적인 데이터 품질관리 실패 사례를 재분류해보면 일곱 가지로 나눌 수 있다. 첫 번째, 같은 정보를 포함하고 있는 데이터의 형식이 일관되지 않는 경우이다. 예를 들어 부서마다 나이, 키, 몸무게 등 입력 방식이 다르며, 숫자와 문자를 혼용해서 쓰는 경우가 속한다. 두 번째로는, 동일해야 하는 데이터 값이 일치하지 않는 경우이다. 부서 간 또는 기록자 간의 재원일에 대한 기준이 달라 입력된 값이 서로 다른 경우가 대표적인 사례이다. 세 번째, 다르게 입력되어야 하는 데이터가 중복 입력되는 경우이다. 이는 자료입력자의 실수로 다른 항목에 같은 값이 두 번 입력되는 경우가 속한다. 네 번째 사례로는 환자의 이름과 등록번호가 바뀌는 경우 등의 환자 식별정보 오류가 속하며, 다섯 번째는 입력된 정보가 이상치인 경우, 여섯 번째는 필요한 정보가 누락되는 경우이다. 타 시스템과의 연동이 불가능하거나 오류가 생기는 경우가 일곱 번째 사례에 속한다. 마지막으로 기타 답변에는 정보시스템 사용법의 미숙, 부서별 업무 프로토콜의 차이, 정보시스템 매뉴얼의 미비를 데이터 품질관리의 실패 사례로 답하였다.

대부분의 응답자(73%)가 위에서 언급한 어려움을 경험한 사례들이 해결된다면 업무의 정확도가 향상될 것이며 효율성이 증대될 것이라고 답하였고, 이를 통해 의료인 및 연구자 등은 물론 병원 운영 및 환자들에게도 도움이 될 것이라 응답하며 의료 빅데이터의 품질관리 사업의 중요성과 필요성을 강조하였다.

### III. 결 론

의료 빅데이터의 품질관리의 실패 사례와 국내의 품질관리 현황 등을 문헌고찰을 통해 살펴 보았다. 그 결과 대부분의 의료기관 및 공공기관에서 의료 빅데이터의 품질관리를 적절히 이행하고 있지 않았으며 저 품질로 인한 피해 사례가 많음을 알 수 있었다. 우리나라의 의료 빅데이터 품질관리 수준이 낮은 요인으로는 품질 관리에 대한 인식저조로 인해 품질관리 전문인력 및 예산이 부족한 것이 가장 큰 요인이었다. 데이터 품질관리 조직이 없기 때문에 데이터 관련 요구사항 및 개선사항이 제대로 반영되지 못하고, 데이터간의 연계성 관리가 제대로 이루어지지 않는 것도 품질저해 요인으로 보고되었다. 또한, Y의료원 내의 의료 빅데이터 품질관리 TFT의 활동과 설문조사를 통해 데이터 실제 사용자 및 관리자들의 저 품질로 인한 업무의 어려움을 겪고 있으며, 이를 개선할 시스템 및 프로세스가 필요하다는 점을 알 수 있었다. 따라서 이 연구결과를 통해 의료 빅데이터의 품질관리의 필요성과 중요성에 대한 인식과 품질관리를 위한 노력이 향상되기를 기대한다.

### 참 고 문 헌

- [1] 건강보험심사평가원, 국내의료기관 의료정보화 현황조사, 2014.
- [2] 보건복지부·한국보건산업진흥원, 보건의료정보화를 위한 진료정보교류 기반 구축 및 활성화, 2016.
- [3] 이연희, 영국 보건복지정보센터의 역할과 전략, 2015.
- [4] 조인숙, 전자의무기록 데이터의 이차활용을 위한 구조화된 데이터 질에 대한 탐색, 2009.
- [5] 한국데이터베이스진흥원, 2010년 데이터 품질관리 성숙수준 조사 보고서. 2010.
- [6] AHC Media, Incorrect registration data is a significant patient safety worry, 2015.
- [7] Bhansali Neera, Challenges of Data Quality in Medical Informatics Data warehouses, 2015.
- [8] British columbia, Quality Assurance and the Data Quality Management Plan, 2015.
- [9] Connecting for Health Common Framework. Background Issues on Data Quality. 2006
- [10] Dr. Boris Otto, Data Quality Management: Framework and Approach for Data Governance, 2007.
- [11] Eckerson, W.W., Data Quality and the Bottom Line: Achieving Business Success Through a Commitment to High Quality Data, 2002.
- [12] ECRI Institute, Top 10 list of patient safety concerns, 2016.
- [13] United States Department of the Interior, Data Quality Management Guide, 2008.
- [14] WHO, Improving Data Quality: A Guide for Developing Countries, 2003.
- [15] <http://www.ahima.org/topics/ida?tabid=certification>.
- [16] <http://www.informationweek.com/healthcare/clinical-information-systems/poor-data-management-rkqts-healthcare-providers/d/d-id/1105481?>

저 자 소 개



**최혜린(Hye Rin Choi)**

- 2017년 : 연세대학교 대학원 보건학과 (석사)
- 2017년~현재 : 연세대학교 대학원 보건학과 (박사과정)
- 관심분야 : 빅데이터, 만성질환 역학 등



**이승원(Seung Won Lee)**

- 2017년 : 연세대학교 대학원 보건학과 (박사)
- 2017년~현재 : 연세대학교 대학원 보건학과 박사후 연구원
- 관심분야 : 만성질환 역학, 심혈관계 질환 등

혈관계 질환 등



**김영아(Young Ah Kim)**

- 2003년 : 연세대학교 간호대학 (박사)
- 1997년~현재 : 연세의료원 의료정보실
- 관심분야 : 데이터 분석, 품질 관리, 표준 플랫폼 등



**이종호(Jong Ho Lee)**

- 2010년 : 강릉원주대학교 컴퓨터공학과 (학사)
- 2017년~현재 : 연세의료원 산학협력단 연구원
- 관심분야 : 데이터 엔지니어링, 대용량 데이터 처리 등



**고홍(Hong Koh)**

- 2017년 : 연세대학교 의과대학 의학과 (박사)
- 2017년~현재 : 연세대학교 의과대학 소아과학교실 부교수
- 관심분야 : 소아 소화기질환, 소아 간담체질환 등



**김현창(Hyeon Chang Kim)**

- 2003년 : 연세대학교 대학원 보건학과 (박사)
- 2015년~현재 : 연세대학교 의과대학 예방의학교실 정교수
- 관심분야 : 만성질환 역학, 임상역학 등