

논문 2014-51-10-17

# 환자 의료 정보 공유 및 데이터 통합을 위한 데모그래픽 데이터 활용 연구

( A Study for Sharing Patient Medical Information with Demographic  
Datasets )

임 중 우\*, 정 은 영\*\*, 정 병 희\*\*, 박 동 균\*\*, 황보 태 근\*

( Jongwoo Lim, Eun-Young Jung, Byoung-Hui Jeong, Dong Kyun Park<sup>Ⓢ</sup>, and Taeg-keun Whangbo )

## 요 약

온라인에서 공유 및 활용되고 있는 정보들이 기하급수적으로 생성되는 인터넷 정보 시대에, 개별 의료기관의 환자 정보는 의료기관 고유의 데이터베이스 구성 및 환자 사생활 정보 보호 문제 등의 이유로 인해 병원들 간의 환자 데이터 공유가 원활히 이루어지지 않고 있다. 환자 사생활 정보를 보호하면서 각 의료기관 고유의 환자 정보를 의료기관들 간에 상호 공유하는 것은 의료 정보화를 위해 아직도 해결해야 할 과제로 남아있다. 본 논문에서는 환자 사생활 정보를 보호하면서 환자의 의료정보를 공유하기 위해서, 국내외 의료정보 공유 현황 및 관련 국제 의료정보 표준안을 고찰 및 국내 의료기관의 데모그래픽 데이터를 활용하기 위해 실제 국내 의료기관의 환자 데이터 구조 및 특성을 분석하고 의료 정보 공유 시스템 구조 설계를 제안하고 자 한다.

## Abstract

Recently, although exponentially growing the quantity of information that have been used and shared on internet networks, the patient information of each medical center have not been used and shared among medical centers due to the protection of patients privacy and the different database schema. To address this problem, we have studied the data structure of the patient information, the standard of medical information for patients we propose a patient information sharing system design that each medical center is able to use and share the patient information among medical centers in spite of different patient information systems with protecting patients privacy.

**Keywords :** Demographic Data, Patient Master Index

## I. 서 론

온라인에서 공유 및 활용되고 있는 정보들이 기하급

수적으로 늘어가고 있는 인터넷 정보 시대에, IT 기술의 발전으로 각 분야에서 IT 기술을 도입을 한 결과 특히 의료정보 산업은 의료 정보시스템의 도입으로 이전보다 의료기관 서비스의 질과 업무 처리 속도가 많이 향상이 되어 공공 의료 기관 개별 단위의 정보화 수준이 높아지고 있다.

그러나 개별 의료기관 고유의 환자 정보 데이터베이스 구성으로 인한 서로 다른 환자 정보 형태 및 환자 개인정보 보호 문제로 인해 의료기관 간에 환자 데이터 같은 의료 정보 공유가 원활히 이루어지지 않고 있다.

\* 정회원, 가천대학교 IT대학 컴퓨터공학과  
(Department of IT, Gachon University)

\*\* 정회원, 가천대학교 길병원 유헬스케어 센터  
(U-Healthcare Center, Gachon University)

Ⓢ Corresponding Author (E-mail: pdk66@gilhospital.com)

※ 본 연구는 보건복지부 보건의료연구개발사업의 지원에 의하여 이루어진 것임(A112020)

접수일자: 2014년08월18일, 수정일자: 2014년09월16일  
게재확정: 2014년09월30일

이는 각 의료기관마다 고유의 데이터베이스를 구성하여 서로 다른 데이터 형태를 사용하고 있는 것이 하나의 원인이다. 그 결과 동일한 환자가 근무처나 거주지 변경 등의 이유로 다른 병원에 가게 될 경우, 환자에 대한 이전 의료 정보를 공유할 수 없을 때에는 검사를 처음부터 받음으로써 중복 검사와 같은 불필요한 절차로 인해 시간과 비용의 낭비가 발생한다.

이러한 문제점들은 표준화된 의료정보 시스템을 기반으로 각 의료기관의 환자 정보를 공유함으로써, 불필요한 개별 환자 의료 데이터의 중복 생성을 방지하여 동일 환자에 대한 불필요한 중복 진료를 피할 수 있고, 개별 환자에 대한 정확한 의료 정보를 공유함으로써 의료 기관에 종속되지 않는 연속적인 개인 맞춤형 의료 서비스를 기대할 수 있다.

따라서 개별 환자 의료 정보를 공유하기 위해서는 표준화된 의료 정보 시스템이 요구된다. 현재 국내외에서 의료 정보 표준화 제안 및 표준화 활동이 의료 문서, 환자 성명, 생년월일, 주소 등 환자 개인의 특성을 나타내는 환자 신상 정보와 환자 수술 기록, 처방 기록 같은 임상 정보 등 다양한 부분에서 이루어지고 있다. 본 논문 2장에서는 의료기관의 의료 정보 공유를 위한 국내외 동향을 설명한다<sup>[6]</sup>.

또한 동일 환자에 대한 각 병원의 데이터 형태가 서로 다르므로 환자에 대한 의료 정보를 공유하기 위해서는 의학 발전을 위한 연구를 진행할 때에도 각각 다른 병원의 데이터 포맷으로 인하여 의료 연구를 위해서 병원별로 데이터 포맷을 통일해야 하는 별도의 부담이 발생하게 되어 연구비용이 추가로 소모되게 된다.

해외의 선진국에서는 이러한 문제점을 먼저 인지하고 일부 선진국에서는 통합적인 환자정보를 관리하는 곳도 있으며, 이러한 노력으로 사용되는 기법들 중에 공개된 것은 open MPI(Master Patient Index)가 있다<sup>[15]</sup>. 그러나 이렇게 개발된 기법들은 주로 미국과 같은 언어 및 의료 환경에 적합하도록 연구 개발이 되었기 때문에 국내 의료 환경에서 바로 적용하기에는 어려운 점이 있다.

본 논문에서는 I 장 서론에 이어 II 장에서는 의료 정보 공유에 대한 국내외의 의료기관 현황을 소개하고 III 장에서는 국내 의료기관의 의료 정보 공유 시스템 개발을 위해서 의료 기관의 샘플 데이터를 기반으로 데모그래픽 데이터 특성을 분석하고 환자 의료 정보 공유를

위한 구조 설계를 제안한다. IV 장에서는 연구 결과를 토대로 결론 및 향후 연구 방향에 대해 기술한다.

## II. 국내외 의료정보 현황

IT 기술의 발전으로 기존 의료 정보 시스템 기술은 개별 의료기관의 환자정보 관리 시스템, 의료 영상 시스템, 진단 시스템 등 IT분야의 부문 기술을 활용하여 발전하였고 개별 의료기관 단위로 많은 발전을 해왔다. 이러한 개별 단위 기술의 발전과 함께 개별 의료기관의 환자 정보를 공유함으로써 보다 질 높은 의료 서비스를 할 수 있다.

이번 장에서는 국내외 의료기관의 환자 정보 공유를 위한 관련 기술 및 사업 현황을 소개함으로써 의료 정보 시스템의 국제화를 위한 환자 정보 공유의 필요성을 제시하고자 한다.

### 1. 국외 환자 의료 정보 공유 현황

영국, 호주, 캐나다, 미국 등 선진국에서는 2000년대 초 부터 국가 수준의 의료 정보화 사업을 추진하고 있다. 각 국의 환자 의료정보 공유를 위한 의료 정보화 사업을 다음과 같이 간략히 정리하여, 국내 현황과 비교하고자 한다.

#### 가. 영국

영국은 2002년부터 국가 보건 의료 정보화에 정보기술 아키텍처(enterprise architecture) 개념을 적용하였다. 영국 NHS(National Health Service)는 기본적으로 5개 지역의 분산구조로 구성되어 있으며, 진료예약, 보건 의료관리, 전자처방, 네트워킹 서비스를 중요 요소로 정의하고 환자의 건강과 진료에 관한 정보의 요소를 담고 있는 국가적 차원의 데이터베이스인 Spine을 중심으로 지역의 의료정보 시스템이 연결되는 구조를 가지고 있다.

Spine에는 영국의 유일한 NHS 고유번호와 환자성명, 주소, 생년월일과 같은 환자 개인의 특성을 나타내는 신상정보(Demographic information) 및 알레르기, 응급실 방문, 약물부작용, 수술기록, 투약기록과 같은 요약된 임상정보를 저장하고 있다. 영국 EHR(Electronic Health Record) Architecture는 논리적 아키텍처, 응용 아키텍처, 기술 아키텍처, 데이터 아키텍처로 구성되어

있으며 아키텍처 산출물에 대한 명세서를 제시함으로써 서비스 제공의 요구사항을 정의하고 있다.

#### 나. 캐나다

Canada Health Infoway는 임상자료의 평생기록관리를 위해 Blueprint v1.0에서 개념적인 데이터 및 저장소에 대한 내용을 제시하고 있으며, Blueprint v2.0에서는 관할구역에 따라 EHRi (EHR Infostructure) 저장소에 대한 논리적인 개념 모델을 간략하게 제시하고 있다. EHRs (EHR Solution) Blueprint v2.0은 캐나다의 4만여 솔루션 개발업체들이 일관성 있게 시스템을 개발, 공급할 수 있도록 지원하는 가이드라인으로써 현존하는 EHR 아키텍처의 대표적인 사례로 볼 수 있다.

EHRs Blueprint는 지역별 분산된 아키텍처를 기반으로 하지만 환자 인덱스 및 주요 진료정보는 중앙집중식으로 운영하는 형태를 가지며, 관할 지역내의 EHR 서비스를 제공하는 솔루션인 EHRs는 HIAL(Health Information Access Layer)를 통하여 환자정보의 인덱스를 활용하고, 건강정보의 저장위치에 대한 정보를 관리하는 Locator를 이용하여 다른 지역의 EHRs와 연결됨으로써 논리적으로는 통합된 형태의 EHR 서비스 제공이 가능한 구조로 구성되어 있다.

#### 다. 호주

호주의 HealthConnect는 비즈니스 아키텍처를 강조하고 있으며 특히 전염병 감시 및 등록 관리에 중점을 두고 있다. HealthConnect를 구성하고 있는 데이터 저장소에는 환자의 다양한 사건요약 정보를 저장하고 국가적 차원의 자료 저장소를 관리하며, 환자의 신상정보 및 식별을 위한 등록저장소, 의료제공자의 등록 및 관리를 위한 저장소, 그리고 건강정보의 효과적인 수집 및 활용을 위한 Metadata 저장소를 제공하고 있다.

#### 라. 미국

의료서비스의 상당 부분을 민간 의료기관이 담당하고 이에 따라 다양한 이해 당사자들이 존재하는 미국의 경우에는 연방 정부에서 제시한 표준 가이드라인을 준수한다는 전제 하에 주 정부가 개별적으로 EHR 시스템을 개발하고 있다.

미국 NHIN은 서로 다른 시스템 간 또는 서로 다른 지역 간에 의료정보 전달 및 교환을 위해 공통 표준 및

아키텍처 모델을 제시하고 이를 준수하도록 독려하고 있다. 즉 국가에서 표준이라는 최소한의 요건만을 제시하고 의료계와 산업계의 자율적인 EHR 시스템 개발 및 운영을 정보화 예산의 형태로 지원하는 방식을 채택한다. EHR 시스템은 기본적으로 분산형태이며, 진료문서 교환을 목적으로 개발된 HL7 (Health Level 7)의 CDA(Clinical Document Architecture) 표준과 ASTM (American Society for Testing and Materials)의 CCR (Continuity of Care Record)을 기본으로 구성하고 있다<sup>[1]</sup>.

## 2. 국내 환자 의료 정보 공유 현황

국내 의료 정보산업은 연평균 12%의 고속성장으로 의료기관 개별 단위의 정보화 수준은 높아지고 있으나 각 의료기관간의 의료 정보 공유는 아직도 원활히 진행되고 있지 않다.

국내에서는 대형 의료기관을 중심으로 협력 의료기관과 인터넷 홈페이지를 통해 표준화되지 않은 의료정보를 공유하는 정도이고 보건소의 경우 동일한 EMR (Electric Medical Record)로 구축하여 보건복지 정보개발원을 통해 통합관리 구축중이나 표준안을 반영하지 않아서 보건소이외의 의료기관과는 의료정보 공유가 어려운 상황이다.

또한 국내 의료기관의 의료 정보 시스템은 국제 표준 기반 정보교환에 대한 지원을 고려하지 않고 설계되어 있으며, 표준 기반 데이터의 생성은 DICOM(Digital Image Communication in Medicine) 등 일부분의 데이터에 한정되어 있다.

따라서 국내 의료기관 시스템은 국내 의료 정보 산업의 국제화 및 미래 맞춤형 의료서비스를 제공하기 위해서 국제 표준을 기반으로 하는 의료 정보 공유 시스템을 고려해야 하는 과제를 가지고 있다.

이를 위해 개별 의료기관 간 데이터를 공유하기 위한 인터페이스 작업이 필요하며, 인터페이스의 설계는 개별 의료기관의 서로 다른 시스템을 지원하면서, 새로 추가되는 의료기관의 적용이 용이하도록 데이터베이스 연결과 데이터 객체를 분리하여 관리해야 하고, 의료기관과 환자의 식별을 위한 고유 식별 방법을 제공할 수 있어야 한다.

의료기관 간 데이터를 공유하기 위해 고려해야 할 필수 요소로는 환자 식별, 문서 서식 표준, 표준 프로토콜,

문서 서식 용어 매핑, UI 연결 인터페이스 등이 있다.

### III. 환자 의료정보 공유를 위한 구조 설계 제안

환자 의료정보를 공유하기 위해서는 의료정보 데이터에 대한 국내외 표준안을 기반으로 하는 의료정보 시스템을 설계해야 한다. 이러한 설계를 위해서는 먼저 국내외 의료정보 시스템 표준안에 대한 분석이 필요하다<sup>[2~3, 5]</sup>. 또한 국제 표준기반의 의료정보 통합 및 공유를 위해서는, 의료정보 시스템에서 발생하는 모든 데이터를 처리할 수 있어야 하며, 네트워크를 통한 통합 및 공유를 위해서는 통합 저장소와 공유네트워크에 대한 데이터보안이 정의되어야 한다.

공유 데이터는 표준 데이터 포맷으로 저장되어야 하고, 데이터의 열람 및 전송에 대한 추적이 가능해야 한다. 마지막으로 다양한 정보 제공자로부터 수집된 데이터에 대해서, 환자에 대한 UID를 식별할 수 있어야 한다<sup>[7~8]</sup>.

본 논문에서는 환자정보를 공유하기 위한 의료정보 시스템의 환자에 대한 UID 식별부문을 중심으로 하는 의료정보 공유 시스템 구조를 제안하고자 한다.

1절에서는 국제 의료정보 표준안에 대한 현황을 조사 및 분석하고 2절에서는 국제 표준안을 기반으로 환자정보 공유를 중심으로 하는 의료정보 시스템의 설계를 제안한다.

#### 1. 국제 의료정보 표준안 현황

##### 가. HL7(Health Level 7)

미국에서는 보건 의료기관 및 관련단체들이 의료정보의 공유 필요성을 공감하여 자발적으로 보건 의료정보의 전자적 교환을 위한 표준 설계를 하기 위하여 HL7이라는 단체를 1987년에 설립하였다. 표준 조직으로서의 HL7은 표준의 설계, 개발을 담당하고 있는 표준개발 기구로 서로 다른 의료 데이터에 대한 획득, 추진, 처리시스템을 위한 응용계층 통신을 프로토콜을 개발하는 것으로 1994년 ANSI(American National Standards Institute)에 의해 표준개발 기구로 인증 받았다<sup>[1, 4]</sup>.

HL7의 L7은 ISO의 OSI 모델의 응용계층(Application Layer)과 상응하는 개념으로 의료 환경에

서 전자적 데이터 교환을 위한 어플리케이션 프로토콜을 의미한다. 1987년 version 1.0을 시작으로 2013년 현재 version 3.0까지 업데이트되어 있다. 구성 요소로는 메시지 기반의 데이터 교환표준, Desktop Computer에서 작동되는 소프트웨어를 연결하는 CCOW(Clinical Context Object Workgroup), 임상 의무 기록을 위한 CDA(Clinical Document Architecture), 임상 의사결정 시스템을 위한 Arden Syntax등이 있다.

또한 RIM(Reference Information Model)이라는 정보 모델을 기초로 개발되고 있으며, 이 모델은 모든 HL7 메시지가 다루는 데이터의 기초가 되는 정보를 일관적으로 표현, 공유하는 모델이다. HL7 프로토콜은 메시지 구조, 코딩 규칙, 트리거 이벤트 등으로 구성되어 있다.

##### 나. CDA(Clinical Document Architecture)

HL7의 RIM에서 HDF (HL7 Development Framework)에서 생성된 문서로써, 의료 영역에서 사용하는 모든 임상문서에 대한 구성과 의미를 규정한 국제 표준 기록 표기이다. CDA는 의료 관찰의 기록을 XML(Extensible Markup Language)로 표현하며, 텍스트, 이미지, 사운드, 멀티미디어 데이터를 포함할 수 있는 규격화된 임상 문서구조이다.

따라서 CDA는 의료 영역에서 발생하는 모든 행위와 결과물을 디지털 콘텐츠로 표현 할 수 있다. CDA는 정보의 단위를 클래스로 표현하는데 4개의 주요 클래스는 다음과 같다.

또한 CDA는 점차 보편화되고 있는 임상분류코드를 담을 수 있는 데이터 타입을 지원 지원함으로써 동일한 분류에 속해 있는 행위 및 데이터는 동일하게 처리할 수 있게 되고, 그 자료는 통계적 자료로 활용되어 보다 의미 있는 정보로 재가공 및 활용 할 수 있게 된다.

CDA는 Header와 Body로 구성되어 있으며, Header는 문서정보(문서의 이름, 생성시간, 문서버전 id등), 환

표 1. CDA 주요 클래스  
Table 1. CDA Main Class.

클래스 이름	내용
Entity	환자, 의료기관, 기기등 표현
Role	환자, 의료기관의 역할 표현
Participation Relationship	Act와 Role을 연결하는 표현
Act	진료, 관찰, 절차 등의 행위 표현

자정보(환자이름, 환자주소, 환자 전화번호 등), 의료 행위 정보, 문서에 대한 책임자 등에 대한 구체적인 내용을 포함하고 있고 Body는 문서에 나타내려는 본문부분(환자진료, 약물처방, 절차, 관찰 등)의 내용을 구체적으로 기술하고 있다.

#### 다. CCR(Continuity of Care Record)

CCR은 국제 표준화기구(ASTM)에서 표준화한 진료 정보 시스템 표준이다. CCR은 여러 의료기관이 가지고 있는 진료 정보를 환자의 지속적인 건강관리 및 치료를 목적으로 환자의 건강 상태를 요약된 형태로 표현하는 데이터 집합으로 환자 정보와 의료 서비스제공자 정보, 보험정보, 환자의 건강상태정보, 진료 계획 등의 정보를 포함한다.

미국 재료 시험학회는 2006년 7월 E2369-05 표준문서에서 개인의 지속적인 건강관리를 위한 데이터구조를 정의한 CCR 표준을 제정하였으며, 현재 Google Health를 포함해 다양한 개인건강 기록 서비스에서 가장 많이 활용되는 표준규격이다.

CCR 원활한 데이터 교환을 위해 월드 와이드 웹(W3C)의 XML을 사용하였고 문서 구조는 Header, Body, Footer로 구성되어 있다. Header는 환자의 정보, 문서를 생성하게 된 목적, 문서 생성 시간 및 문서 정보 버전등을 포함하고 Body에는 개인 의료 정보가 포함되어 있고 마지막으로 Footer에는 CCR문서에 존재하는 모든 데이터와 관련된 개인, 기관, 시스템에 등에 대한 정보(Actor), References, Comments, Signature등이 포함되어 있다.

#### 라. CCD(Continuity of Care Document)

CCD는 미국 재료 시험학회가 2005년 CCR을 정식 표준으로 채택한 이후 CCR의 성공적인 확산으로 HL7은 미국 재료시험학회의 협의를 거쳐 CCD규격을 제정하였다. CCD는 HL7의 CDA Release 2와 미국 재료 시험학회의 E2369-05 CCR을 접목시켜 CCD를 기반으로 CCR표준 규격을 수용한 형태이다.

#### 마. IHE(Integration the Healthcare Enterprise)

HIMSS(Health Information Management Systems Society)와 RSNA(Radiological Society of North America)에서 1999년부터 IHE 프로젝트를 통해

DICOM, HL7등의 기존 표준을 이용하여 의료기관의 전반적인 업무를 표준화하여 고질적이고 반복적인 연동상의 문제를 최소화하고 좀 더 효율적이고 안전한 의료기관의 보편적인 업무 흐름을 모델링하고 표준화하고 있다. IHE는 기술적인 프레임워크, 통합 프로파일, 트랜잭션등을 개발하여 문서로 제공하고, 그 문서에 따라 잘 수행되는 지 인증하고 발표하는 작업을 한다.

#### 바. OpenEHR(Electric Health Record)

EHR에 대한 오픈 연구, 개발 및 구현을 지원하고, 새로운 패러다임을 연구하는 비영리 단체의 이름이다. OpenEHR의 명세서는 인적 정보, 임상 워크플로우, 아키타입(Archetype) 정보와 서비스 모델을 포함하며 의학-법률적으로 분산되었고 버전관리가 되는 EHR기반 구조의 기초가 되도록 설계되었다.

또한 재사용 가능한 도메인 개념의 정규 모델인 아키타입을 만들기 위해 ADL(Archetype Definition Language)을 사용한다.

OpenEHR은 서비스기반 소프트웨어 구조를 가진 다중-레벨(Multi Level Modeling)을 기반으로 하고 있다. 이 구조는 도메인 전문가들이 직접 관련 정보들을 정의할 수 있고 전문 용어체계를 쉽게 이용할 수 있도록 한다. 아키타입(Archetype) 명세서는 현재 ISO 표준(ISO 13606-2)을 따르고 있으며, 국가 e-health 정보 표준을 정하는데 있어서 여러 나라 정부에서 사용되고 있다.

## 2. 환자 의료정보 공유 시스템 구조 설계

환자 의료정보 공유 시스템은 국제 표준 기반의 의료 정보 통합 관리 표준 시스템의 일부분으로, 먼저 전체 의료 정보 통합 관리 표준 시스템 구조를 이해해야 한다. 이를 위해 먼저 국내 의료기관 데이터 구조를 분석하고, 의료정보 통합을 위한 정보공유 시스템 모델링, 그리고 의료정보 데이터 교환 모델을 설계한다.

### 가. 국내 의료 기관 데이터 구조 분석

국내 의료기관의 정보 시스템 중 환자 의료정보의 경우, 대부분이 진료에 관한 정보로써 환자 진료 목적의 데이터만으로는 환자 개별 식별이 정확하지 않다. 환자 의료 정보 데이터베이스 시스템에서 환자 식별이 가능한 테이블들은 환자 기본정보, 내원정보, 응급실환자정보, 진단정보 등의 테이블에 있는 속성값을 가지고 분

표 2. 환자 식별을 위해 참고 해야 할 환자 정보의 속성 항목

Table 2. The requirement reference properties of patients' information for identifying patients.

구분1	구분2	항목	결측률(%)	신뢰도	설명 및 비교
1차분리	환자 기본정보	성명	0	상	고유정보
		생년월일	0	상	불변정보
		성별	0	상	불변정보
2차분리	환자 기본정보	우편번호 및 주소	4	하	우편번호는 잦은 개편과 공란이 많음
		휴대전화번호	4	중	국내의 경우, 번호이동율이 높은 편
	내원정보	보험자격코드	0	중	보험사 종류에 따라 코드체계가 다름 직종이나 보험사 군집으로 분류 가능
추가요소	응급실 정보	주 진료과	0	상	첫 진료 접수시부터 기록 보관
		혈액형	N/A	상	불변정보
		화학검사	N/A	상	불변정보
	진단정보	유전자 검사	N/A	상	불변정보
		불.난치병 관련 진단정보	N/A	상	첫 진단 및 완치 일자의 경우 건강보험 심사평가원 에서 관리 및 병원간 공유가 되고 있음

석한다<sup>[9~12]</sup>.

환자 기본정보 테이블에는 현재 환자 기본정보 중 주로 사용하는 식별키 값은 주민등록번호를 사용하고 있으며, 이는 초기 환자를 식별하고 이후에는 개별 병원 시스템에서 고유한 번호로 테이블을 연결하는 고유 환자번호(이름, 생년월일, 주소, 성별, 인종 등의 데이터 값을 혼합)를 생성한다.

환자 내원정보 테이블에는 환자가 의료기관에 방문하여 접수하는 과정에서 발생하는 정보들이 저장된다. 환자 입원 일자, 내원 환자구분, 보험자격 코드 등의 데이터 속성 값이 있다. 응급실 환자 정보의 경우 환자의 현재 건강에 관련된 값들로 구성되어 있어 환자상태 체크 및 치료에 관련된 정보로써 환자상태 식별이 주목적이다. 환자 진단정보 테이블의 경우 환자가 입원하지 않은 경우 내원일자와 거의 일치한다.

상기에 소개한 환자 식별을 위한 의료정보 데이터 시스템의 테이블들을 기본으로 하여 환자 의료정보 공유를 위해 2곳 의료기관의 의료정보 데이터베이스 시스템의 환자정보 관련 테이블을 다음과 같이 분석하였다.

두 의료기관의 환자 정보 중 환자 이름과 주민등록번호는 개인정보 보호를 위해 삭제된 상태에서 정보를 받았다.

A의료기관의 경우, 총 27,835명의 환자정보가 각 환자별로 12개의 속성으로 구성되었다. 12개의 속성 중 동일한 생년월일인 경우가 발생하고, 이메일 등의 속성 값은 공란 발생 비율(92.83%)이 높았다.

B의료기관의 경우, 총 128,319명의 환자정보가 환자

별로 6개의 속성으로 구성되어있다. A의료기관과 같은 경우가 발생하였고, 더 많은 환자정보 일수록 동일한 생년월일을 가지고 있는 환자의 비율은 증가하였다. 또한 주소의 경우 97.26%의 기록을 나타내었다.

두 의료기관의 환자 정보를 공유하기 위해서는 두 의료기관 공통의 환자정보 속성값, 예를 들면 생년월일, 성별 등으로 개별 환자를 정확히 신뢰할 수 있는 식별을 해야 한다.

위에서 알 수 있듯이, 환자 기본 데모그래픽 정보만으로는 현재 신뢰할 수 없는 식별율을 나타낸다. 따라서 환자 기본 데모그래픽 정보이외의 내원 정보, 진단 정보 등의 속성값을 추가하면 신뢰할 수 있는 식별율이 예측된다.

환자 정보 공유를 위한 선처리 작업 중 신뢰할 수 있는 환자 식별 처리가 필요하다. 환자 식별을 위해 고려해야할 환자 정보의 속성값을 표 2에서 정리하였다.

위의 표 2에서 1차 분리는 단순 매칭으로 환자를 식별 분리할 수 있는 항목들이다. 1차 분리를 통해서 완벽한 식별이 이루어지지 않는 경우 2차 분리에 속한 항목들을 추가하여 식별, 분리한다. 2차 분리 항목과 추가요소의 항목들은 개별 의료기관에서 생성되는 환자정보이다.

2차 분리 및 추가요소 항목으로 환자를 식별 분리할 수 없는 경우, 개별 의료기관의 정보가 시간에 따라 변동하기 때문에 매칭하고자 하는 항목들에 가중치를 주고 우선순위를 정하고 매칭 및 수정할 수 있는 환자 식별 및 분리 시스템을 설계해야 한다. 이러한 환자식

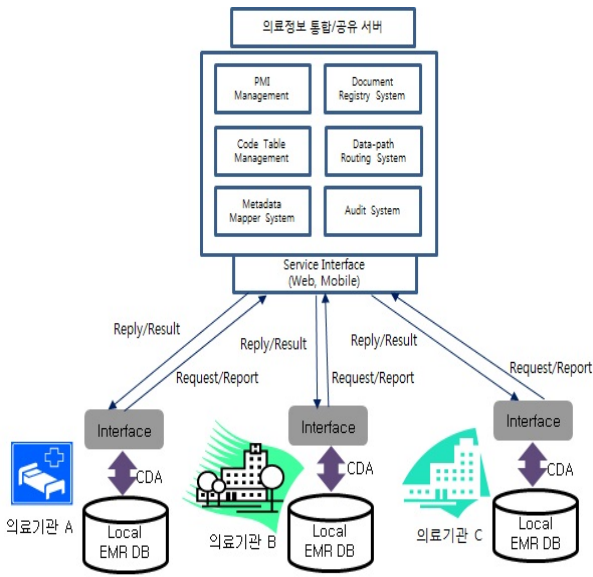


그림 1. 의료정보 공유시스템 전체 개요  
 Fig. 1. Overview of Medical information Sharing System.

별 및 분리 시스템은 그림 1의 의료정보 공유시스템 6 요소 중 PMI(Patient Master Index) Management를 구현하기 위해서 고려해야 할 필수 요소이다.

본 논문에서는 의료 정보 공유를 위한 전체 시스템 중개별 환자 정보를 식별하여 공유하기 위한 부분으로 한정하여 의료 정보 공유 시스템을 설계 및 제안한다.

그림 1에서 DICOM, CDA, HL7 등은 이전 장에서 소개한 국제 표준 인터페이스이다. 본 논문에서 제안한 의료정보 공유 구성 요소는 다음과 같다.

- (1) PMI Management: 환자의 등록과 UID 생성, 검색 관리
- (2) Document Registry System: 의료 정보 관련 서식 관리
- (3) Code Table Management: 코드 테이블 관리
- (4) Data-path Routing System: 데이터 routing 관리
- (5) Metadata Mapper System: 메타데이터 관리
- (6) Audit System: 감시/회계

본 논문은 의료 정보 공유를 위한 전체 시스템의 구조 개요를 소개하였고, 환자 데모그래픽 데이터를 활용하는 부분은 전체 시스템 구성 6가지 요소 중 PMI Management 요소이므로 이에 한정하여 환자 정보 공유 및 통합을 위한 환자 정보 식별, 분리하기 위한 모듈을 그림 2와 같이 제안한다.

PID(환자의 UID)를 이용하여 다른 의료기관 간의 상호 검색은 HL7 Application Protocol을 통해 Message

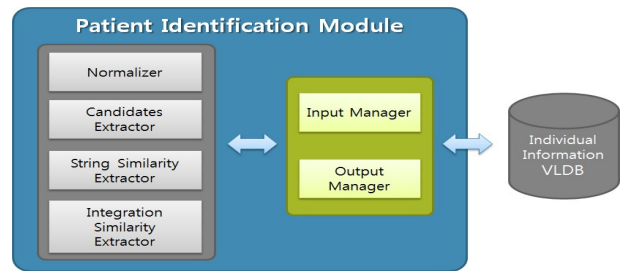


그림 2. 환자식별 모듈  
 Fig. 2. Identifying Patients Module.

를 주고받을 경우, 공유 및 통합을 위한 PID를 제공하는 Routing Map이 필요하다. 현재 EMPI(Enterprise Master Patient Index)가 구현된 다양한 국외 시스템(openEMPI 등)이 존재하며 그 특성은 다음과 같다.

- (1) 환자의 UID 중복 생성 방지
- (2) Demographic 정보 등에 대한 중복저장을 막고, 데이터 입력 오류 최소화
- (3) 환자의 기록 위치를 식별하는 서비스 가능
- (4) 응급 상황에 보다 빠른 데이터 접근 가능

그러나 EMPI는 국내 의료 기관 정보 시스템 환경 및 데이터 특성 등을 고려하지 않은 시스템이어서 적합하지 않다. 따라서 국외 시스템을 참고하여 국내 의료 기관 정보 시스템 환경에 맞는 PMI Management 시스템 구현이 필요하다. PMI Management의 핵심 역할은 주어진 의료정보를 기반으로 매칭되는 환자를 일관성있게 판단하는 것이다.

그림 2는 환자 Demographic 데이터를 이용하여 환자를 식별하는 모듈 개요도이다.

Normalizer는 입력된 데이터의 타입을 분석 가능한 타입으로 변형시키는 부분으로 예를 들면, 특수문자 제거, 분절처리, 자/모음 탈락현상 처리 등이 있다.

Candidates Extractor에서는 Soundex 알고리즘을 기반으로 2byte 체계의 언어인 한글의 특성을 반영하여 성명(스트링) 유사도 산출 대상 후보군을 추출하기 위한 Phonetic코드를 추출한다.

String Similarity Extractor에서는 Edit distance 방식을 기반으로 한글 체계에 맞게 개선하여 추출된 성명 유사도 산출 대상 후보군 내에서 입력 스트링과의 최종 유사도를 산출한다. 성명 이외의 식별 요소들에 대한 최종 유사도를 각각 산출한다.

Integration Similarity Extractor에서는 각 식별 항목 별로 산출된 유사도를 중요도 가중치를 적용하여 최종 통합 유사도를 산출한다.

최종 Output Manager에서는 입력 데이터의 유사도 높은 순위로 매칭 리스트를 출력하여 Database에 저장한다.

#### IV. 결 론

본 논문에서는 국내외 의료 정보 시스템 관련 표준안을 참고 및 기반으로, 의료 정보 공유를 위한 시스템을 구현하기 위하여 국내외의 의료정보 공유 시스템 현황을 고찰하고, 국제 표준안을 검토 및 분석하였으며, 의료 정보 공유를 위해 제안한 전체 시스템 요소 중에서 PMI Management를 위한 환자 정보를 신뢰성 있게 식별, 분리하는 방법론을 제안하였다.

또한 제안한 PMI Management의 핵심 기능인 환자 식별 기능은 현재 1byte 글자체계인 Alphabet을 사용하는 국가에서는 기존 Edit Distance Algorithm<sup>[13]</sup>과 Soundex Algorithm<sup>[14]</sup>을 2byte 글자 체계인 한글의 특성을 고려하여 스트링간의 유사도 산출의 정확도 및 변별력을 개선하였다.

개별 의료기관간의 환자 정보 공유를 위해서는 기존 개별 의료기관에서 사용하고 있는 환자 정보를 재사용하면서 전체 의료정보 시스템에서 통용되는 개별 환자의 고유한 ID를 생성해야 하고 그 고유한 ID는 환자의 사생활 정보와 연관하지 않아야 하는 과제를 해결해야 한다.

이를 위해서 국제 의료정보 표준안을 참고한 의료정보 공유 시스템의 한 요소로써, 고유 환자 정보 색인 (Patient Master Index)을 생성 관리하는 모듈이 필요하다. 기존 환자 Demographic Data를 활용하기 위해서는 개별 의료기관간의 환자 정보에 대한 정확한 매칭 모듈이 필요하다.

본 논문에서 제안한 환자식별 모듈은 향후에 환자정보 매칭 정확도를 개선한 알고리즘 및 시스템 구현을 통한 실험 결과를 제시할 계획이다.

또한 제안한 환자 의료 정보 공유를 위한 의료 정보 시스템의 설계를 바탕으로 의료 정보 표준화를 반영한 의료 정보 공유 시스템 개발 및 구현은 개별 환자에 대해 보다 정확한 진단 기능 향상 및 관련 의료 정보기술

의 발달이 예상 된다.

또한 연관된 U-health 사업이 활성화되고, 환자의 사생활 정보 공개 없이 개인 병력을 기반으로 하는 의료 정보만을 공유할 수 있게 되어 환자의 사생활을 보호받으면서 의료기관의 위치에 상관없는 양질의 의료 서비스가 예상된다.

#### REFERENCES

- [1] Donald T. mon, "HL7 PHR System Functional model" HIMSS, 2013.
- [2] Jeongeun Kim, Hongju Jung and David W. bates, History and Trends of "Personal Health record" research in Pubmed. Healthc Inform res. 2011; 17(1):3-17.
- [3] Archer n, Fevrier-Thomas U, Lokker C, mcKibbon KA and Straus Se, Personal health records: a scoping review. J Am med Inform Assoc 2011; 18: 515-522.
- [4] IEEE Draft Standard for Health informatics - Personal health device communication - Device specialization - Insulin pump, July 2013.
- [5] Paul C. Tang et al., Personal Health records: Definitions, benefits, and Strategies for overcoming barriers to Adoption, JAMIA. 2006;13:120-126
- [6] Korean Standards Association, "Prioritization of Standardization for home network healthcare service", 2006.12.
- [7] Thomas A Workman, "Engaging Patients in Information Sharing and Data Collection: The Role of Patient-Powered Registries and Research Networks," AHRQ Publication No. AHRQ 13-EHC124-EF, 2013. 9.
- [8] HIMSS(Health Information Management Systems Society), "Patient Identity Integrity", the HIMSS Patient Identity Integrity Work Group, 2009.
- [9] Magret Floy, Yuejen Zhao, John Condon, "Demographich Data Assessment for Northern Territory Public Hospitals," 2011.
- [10] Sung Sug Yoon, Mary G. George, et al., "Analysis of Data-Collection Methods for an Acute Stroke Care Registry," American Journal of Preventive Medicine, 2006.
- [11] William G. Axinn, Cynthia F. Link, Robert M. Groves, "Responsive Survey Design, Demographic Data Collection, and Models of Demographic Behavior," NSFG Survey



- Methodology Working Papers, 2009. 10.
- [12] Health Information and Quality Authority, “National Standard Demographic Dataset and Guidance for use in health and social care settings in Ireland,” 2013.09.
- [13] Navarro, Gonzal. “A guided tour to approximate string matching”, ACM Computing Surveys 33 (1): 31 - 88. 2001.
- [14] Knuth, Donald E. “The Art of Computer Programming: Volume 3, Sorting and Searching,” Addison-Wesley. pp. 391 - 92. 1973.
- [15] <http://www.medfloss.org/node/290>

---

 저 자 소 개
 

---

## 임종우(정회원)

1997년 경원대학교 전자계산학과 학사 졸업.  
 1999년 경원대학교 전자계산학과 석사 졸업.  
 2013년 University of Southern California,  
 Computer Science 박사 졸업.  
 <주관심분야 : 유헬스케어, 의료정보학, 데이터베이스, 온토로지>



## 정은영(정회원)

2002년 가천 의과대학교 보건정보학과 석사 졸업.  
 2012년 아주대학교 의과대학 의료정보학과 박사 졸업.

<주관심분야 : 의료정보학, 유헬스케어>



## 정병희(정회원)

2001년 전주대학교 생물학과 학사 졸업.  
 2010년 인천대학교 정보통신공학과 석사 졸업.  
 2014년 인천대학교 컴퓨터공학 박사 졸업.

<주관심분야 : 유헬스케어, IT융합, AI>



## 박동균(정회원)

1992년 충북대학교 의과대학 의학사 졸업.  
 2003년 인하대학교 의과대학 석사 졸업.  
 2010년 인하대학교 의과대학 박사 졸업.

<주관심분야 : 의료정보학, 유헬스케어>

## 황보택근(정회원)

1983년 고려대학교 금속공학과 학사 졸업.  
 1988년 City University of New York 컴퓨터공학 석사 졸업.  
 1995년 Stevens Institute of Technology, 컴퓨터공학 박사 졸업.

<주관심분야 : 유헬스케어, 의료정보학>