



임상문서 정보교류 시스템의 의미론적 상호운용성 확립을 위한 메타데이터 국제표준 적용방안

박유령¹ · 김혜현¹ · 안은양¹ · 김형회^{2,3} · 김주한⁴ · 박래웅⁵ · 박동균⁶ · 정은영⁶ · 김주한^{1,7*} | ¹서울대학교 의과대학 정보의학실 바이오정보의학연구실, ²부산대학교 의학전문대학원 진단검사의학교실, ³부산대학교병원 의학연구소, ⁴전남대학교병원 심장센터, ⁵아주대학교 의과대학 의료정보학과, ⁶가천대 길병원 유헬스케어센터, ⁷서울대학교 의과대학 시스템 바이오 정보의학 국가핵심연구센터

Establishing semantic interoperability in the course of clinical document exchange using international standard for metadata registry

Yu Rang Park, PhD¹ · Hyehyeon Kim, MS¹ · Eun Yang An, AB¹ · Hyung Hoi Kim, MD^{2,3} · Ju Han Kim, MD⁴ · Rae Woong Park, MD⁵ · Dong Kyun Park, MD⁶ · Eun-Young Jung, PhD⁶ · Ju Han Kim, MD^{1,7*}

¹Seoul National University Biomedical Informatics (SNUBI), Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul, ²Department of Laboratory Medicine, Pusan National University School of Medicine, Busan, ³Medical Research Institute, Pusan National University Hospital, Busan, ⁴Heart Center of Chonnam National University Hospital, Gwangju, ⁵Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, ⁶U-Healthcare Center, Gachon University Gil Hospital, Incheon, ⁷Systems Biomedical Informatics National Core Research Center, Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul, Korea

*Corresponding author: Ju Han Kim, E-mail: juhan@snu.ac.kr

Received July 2, 2012 · Accepted July 16, 2012

Around the world electronic health records data are being shared and exchanged between two different systems for direct patient care, as well as for research, reimbursement, quality assurance, epidemiology, public health, and policy development. It is important to communicate the semantic meaning of the clinical data when exchanging electronic health records data. In order to achieve semantic interoperability of clinical data, it is important not only to specify clinical entries and documents and the structure of data in electronic health records, but also to use clinical terminology to describe clinical data. There are three types of clinical terminology: interface terminology to support a user-friendly structured data entry; reference terminology to store, retrieve, and analyze clinical data; and classification to aggregate clinical data for secondary use. In order to use electronic health records data in an efficient way, healthcare providers first need to record clinical content using a systematic and controlled interface terminology, then clinical content needs to be stored with reference terminology in a clinical data repository or data warehouse, and finally, the clinical content can be converted into a classification for reimbursement and statistical reporting. For electronic health records data collected at the point of care to be used for secondary purposes, it is necessary to map reference terminology with interface terminology and classification. It is necessary to adopt clinical terminology in electronic health records systems to ensure a high level of semantic interoperability.

Keywords: Metadata; Semantic interoperability; Metadata registry; ISO/IEC 11179; Clinical document exchange

서 론

전자의무기록 시스템(electronic medical record system)의 활발한 보급으로 기존의 청구자료나 처방전, 검사결과뿐 아니라, 의무기록의 주요 구성요소인 임상문서들의 대부분이 생성단계에서부터 디지털 형식으로 생성되고 관리되게 되었다. 전자적으로 생성된 문서는 자료활용도가 높다. 종이문서와 달리 전자적으로 관리되는 문서는 여러 의료인이 동시에 열람하는 것도 가능하고, 복제나 전송, 자료분석 등에 효율적으로 활용될 수 있다.

임상문서는 흔히 ‘서식’이라는 이름으로 불린다. 전자의무기록 시스템의 높은 보급률에도 불구하고, 임상문서의 관리 체계는 여전히 부실하다. 대규모 의료기관의 경우 3,000~8,000개에 달하는 수 많은 서식을 관리하고 있다. 서식을 구성하는 각 항목은 경험적으로 정의된 것들이어서 의미론적 일관성을 갖기 어렵다. 예를 들면, 동일한 항목(예: 성별)이 서로 다른 표현(성, 성별, sex, gender 등)으로 표기된 경우, 항목 값의 입력형식의 정의가 명확하지 않은 경우, 자료처리를 위한 항목 값의 부호화(예: 남/녀, M/F, m/f, 0/1, 1/0, 1/2 등) 원칙에 일관성이 없는 경우가 흔하다. 의미론적으로 등가인 항목이 서식마다 다른 라벨, 항목값 및 부호화 규칙을 갖기도 한다. 뿐만 아니라 서식의 지속적인 변경과 추가로 일관성은 깨어진다. 같은 문서의 다른 버전이 일관성 없이 혼재한다.

이러한 임상문서 및 그 항목 정보의 의미론적 비일관성은 자료처리에도 상당한 혼란을 가져오지만, 추후 자료분석이나 임상연구, 임상 데이터웨어하우스 구축 등에는 더욱 심각한 장애요인이 된다. 기관에서 사용되는 모든 임상문서와 그 항목과 그 항목값의 의미론적 일관성 확보는 전사적 정보 자원의 운영체계를 갖추기 위한 필요조건이며, 기관을 넘어 국가 지식정보전략 차원의 해결과제다.

원활한 기관간 의료정보 교류는 임상진료의 질을 높이고 의료비용을 절감할 수 있다. 서로 다른 정보시스템 사이의 정보교류는 표준에 대한 합의와 실제 적용에 기반한다. HL-7 clinical document architecture (CDA) [1], American Society for Testing and Materials (ASTM) continuity

of care record (CCR) [2], Clinical Data Interchange Standards Consortium operational data model [3] 등 목적에 따라 다양한 의료정보교환 표준이 제시되고 있다. 그러나 임상정보의 방대하고 복잡한 특성으로 의료정보교류는 아직 개념 정립 단계에 머물거나, 진료비 청구와 같은 특수 분야에 제한되거나, 몇몇 기관 사이의 시범운영에 그치고 있다[4-6].

현재까지 제안된 대부분의 의료정보교류 표준은 구문론적(syntax) 표준을 지원한다. 그러나 의미론적(semantic) 표준의 적용은 진정한 의료정보교류를 위한 필수요소다. 임상문서 정보교류를 위한 대표적인 표준인 HL-7 CDA는 구문론적 표준, 즉 문서의 형식을 어떻게 표현하고 상호운용할까 하는 방안(syntactic interoperability)을 제안한다[7]. HL-7 CDA는 문서 항목의 의미론적 상호운용성(semantic interoperability)을 지원하지 못하는 한계를 극복하기 위해, 항목들을 통제용어체계인 Systematized Nomenclature of Medicine (SNOMED) [8]나 unified medical language system (UMLS) [9] 등으로 코드화할 것을 제안한다. 그러나 통제용어 체계는 개별 용어 수준에서의 개념적 일관성을 지원할 뿐이며, 여러 용어의 조합으로 표현해야 하는 상위 혹은 복잡개념의 표현을 위한 부호조합규칙(encoding rule)을 제공하지는 않는다. 결국 동일한 의미를 갖는 상위/복잡 개념들은 수많은 방식으로 서로 달리 표현되어 심각한 모호성을 유발한다. 즉 통제용어 체계는 용어 자체를 넘어서는 상위개념 표현을 위한 의미론적 일관성을 지원하지 못한다. 최근까지 수행된 HL-7 CDA를 활용한 임상문서 교환은 이러한 한계로 제한적 결과만을 도출할 수 있었다.

최근 발전한 메타데이터 기술은 의학계에서 개발된 방법만 아니지만 이러한 한계에 대해 큰 시사점을 제공한다. 메타데이터는 “데이터에 대한 데이터”로 정의된다. 전자상거래나 국가 정보관리체계 등에 널리 적용되며 발전해온 메타데이터 기술은 정보항목과 그 속성값에 대한 매우 구체적인 정보요소들을 체계적으로 기술, 관리할 수 있는 다양한 표준을 제시한다. 특히 국제표준화기구를 통해 발전한 ISO/IEC 11179 메타데이터 기술 및 등록에 관한 표준(metadata registry, MDR)은 생명의료 분야에서도 적용되기 시작했다[10]. 미국 보건성인 U.S. Department of Health and Human Services

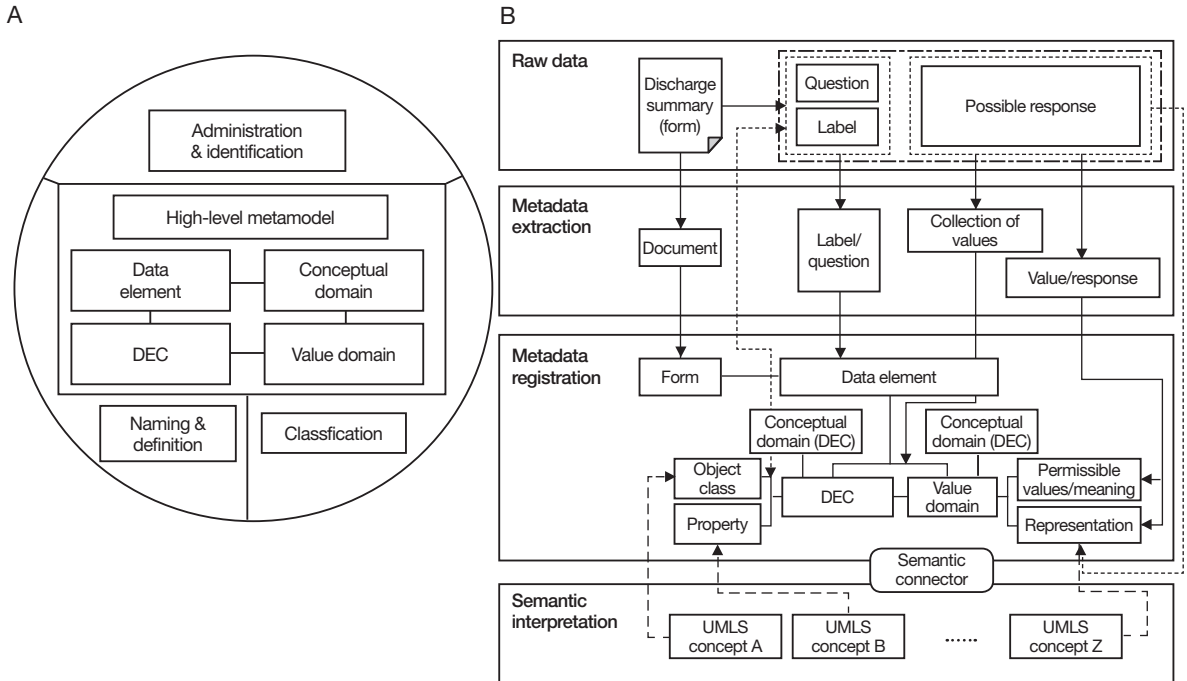


Figure 1. Creating metadata for clinical document exchange. (A) Basic architecture of ISO/IEC 11179 consists of data element, data element concept (DEC), conceptual domain and value domain. (B) Procedure diagram of metadata extraction vocabulary mapping from clinical documents. UMLS, unified medical language system.

(HHS)와 미국 국립암센터(National Cancer Institute, NCI)의 경우 임상정보의 상호운용성 확보를 위해 ISO/IEC 11179 MDR 기술 기반의 HHS U.S. Health Information Knowledgebase (USHIK) 및 NCI caDSR 시스템을 개발하여 메타데이터 생성 및 등록을 지원하고 있다[11,12].

HL-7 CDA는 임상문서 정보교류의 구문론적 상호운용성을 지원하는 대표적인 국제표준이다. ISO/IEC 11179 MDR은 의미론적 상호운용성 지원을 위한 범용성 도구다. 두 표준 모두 XML을 그 기반기술로 사용한다. 그러므로 본 연구에서는 ISO/IEC 11179 표준을 지원하는 MDR을 구축하고 여러 의료기관의 임상문서를 수집하여, 각 임상문서의 항목과 항목값을 메타데이터 형식으로 추출, 등록한다. 각 임상문서의 형식은 구문표준인 CDA로 표현하고 문서내의 항목과 항목값들은 ISO/IEC 11179 표준에 따라 기술한 통합된 XML 문서를 생성하여 구문론적 상호운용성과 의미론적 상호운용성을 모두 지원하는 임상문서의 범용 정보교류체계를 구축 시험 운영한다.

재료 및 방법

메타데이터는 자료에 대한 자료를 체계화하여 자료관리 의 일관성 및 상호운용성을 높이기 위한 방법이다. 메타데이터의 표현에 대한 표준안에는 RDF, Dublin Core 등이 있지만, 기술 및 등록에 관련한 대표적인 표준은 ISO/IEC 11179 MDR이다[13]. ISO/IEC 11179의 특징은 자료의 기본단위를 데이터 엘리먼트(data element, DE)로 보아 DE와 관련한 특징들을 잘 표현할 수 있는 명세서를 제공하는 객체 모델이라는 점이다. ISO/IEC 11179 모델에서는 DE의 식별, 명명, 정의, 개념, 값, 분류 등에 중점을 두어 교환될 자료의 이용자가 의미, 표현형식 및 식별에 대한 공통된 이해를 갖기 위해 메타데이터 등록과 유지를 원활히 수행하기 위한 지침을 제공한다.

Figure 1A는 ISO/IEC 11179의 MDR 모델이 high-level metamodel, administration & identification, naming & definition, classification으로 구성됨을 보여준다. ISO/IEC

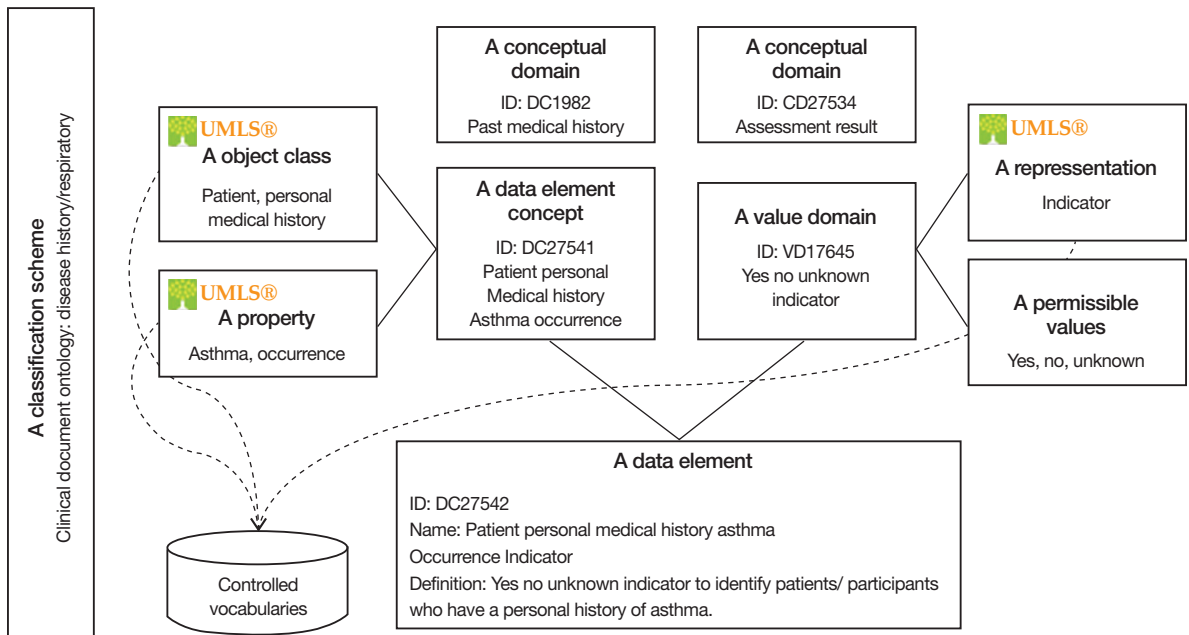


Figure 2. The example of metadata about patient personal medical history asthma occurrence indicator.

11179의 핵심 부분인 high-level metamodel은 메타데이터를 기술하기 위해 DE의 개념부인 'DE concept'와 'conceptual domain' 그리고 표현부인 'DE'와 'value domain'으로 구성된다. 나머지 administration & identification, naming & definition 및 classification은 high-level metamodel에서 기술된 메타데이터의 관리를 위한 항목들이다. Administration & identification은 등록된 메타데이터의 제출기관, 등록자, 관리항목 등 관리식별 관련 항목을 구조화한다. Naming & definition은 사용언어, 등록 대상의 범위 등 명명과 정의의 관련 항목을 구조화한다. Classification은 등록된 메타데이터의 분류체계, 분류항목 및 분류체계와 항목들간의 상관관계를 구조화한다.

Figure 1B는 임상문서에서 메타데이터를 추출하고, ISO/IEC 11179 표준에 따라 기술하고, MDR에 등록하고, 다시 표준 통제용어체계에 매핑하는 절차를 보여준다. 수집된 임상문서에서 항목의 라벨과 문항을 추출하고, 답항의 가능한 값(허용치)을 추출하여, Figure 1A에 상세히 표현된 ISO/IEC 11179 표준의 각 항목들을 완성한다. 이러한 과정은 자동화 과정과 수동 과정을 배합한 반자동 방식으로 수행된다.

모든 과정이 완료되면 DE concept의 하위 영역인 object class와 property에 할당된 속성값과 value domain의 하위 영역인 representation에 할당된 속성값을 각각 통제용어체계에 매핑하여 좀더 풍부한 의미론적 요소를 부여할 수 있다.

Figure 2는 ISO/IEC 11179 표준에 따라 구성한 메타데이터의 예로써, '천식 발병력(patient personal medical history asthma occurrence indicator)'의 내용과 구조를 도식화한 것이다. 정의에서와 같이 이 DE는 천식병력을 가진 환자의 yes no unknown의 표기 지표다. 이 DE의 conceptual domain은 past medical history, DE concept는 patient personal medical history asthma occurrence에 해당하고, conceptual domain은 assessment result, value domain은 yes no unknown indicator에 해당한다. 해당 value domain에 허용되는 값(permissible value)은 yes, no, unknown으로 제한된 지표(indicator)로 사용된다. DE concept는 다시 patient, personal medical history라는 객체와 asthma occurrence라는 세부속성을 갖는다. Figure 2에서 UMLS 아이콘으로 표기된 세 개의 항목, DE con-



Table 1. Distribution of metadata served by BMESH server

Source type	Name	No. of documents	No. of instances
Public metadata repository	National Cancer Institute Cancer Data Standards Registry and Repository	21 Projects and 3,242 case report forms	14,331
	Agency for Healthcare Research and Quality U.S. Health Information Knowledgebase	46 Standard development organizations	39,572
Clinical documents	Clinical documents extracted from three Korean teaching hospitals	42 Departments, 8 categories, and 814 documents	33,325
	Five clinical documents from five hospitals ^{a)}	25 Documents	527
	HL-7 clinical document architecture	6 Documents	127
	CRFs from colon cancer research	1 Documents	477
	Clinical trials	66 CRFs, 5 clinical trials	6,102
	CRFs from rare disease research	92 CRFs for 19 disease	2,142
Clinical research & trials	College of American Pathologists Cancer Protocols	12 Categories 44 protocols	2,156
	American Society for Testing and Materials continuity of care record	1 Standard specification	188
Content standards	CDISC CDASH	1 Standard specification	161
Epidemiology	KFDA cohort study	68 Cohort studies	11,153
Total	14 Biomedical resources		114,308

CRF, case report form; CDISC, Clinical Data Interchange Standards Consortium; CDASH, Clinical Data Acquisition Standards Harmonization ; KFDA, Korea Food & Drug Administration. ^{a)}Five clinical document types (admission note, initial medical examination note, discharge note, emergency note, and operation note) from five major Korean hospitals (Seoul National University Hospital, Ajou University Medical Center, Pusan National University Hospital, Gachon University Gil Hospital, and Chonnam National University Hospital).

cept의 object class와 property 및 value domain의 representation을 해당 UMLS 개념에 매핑했다. Figure 2는 ISO/IEC 11179 표준에 따른 메타데이터의 상세한 기술방법의 예이지만, 이처럼 각 속성 및 세부속성 값의 표준 통제 용어체계와의 매핑을 통해 더욱 강력한 의미론적 특성을 확보할 수 있음도 함께 보여주고 있다.

이와 같은 과정을 거쳐서 추출된 메타데이터를 등록하고

운영하기 위한 MDR 서버로는 서울대학교 의과대학 정보의학실에 구축되어 있는 Biomedical Metadata Element Standard for Health (BMESH) 서버를 활용한다. Table 1은 현재까지 BMESH에 등록되어 운영되고 있는 메타데이터의 분포를 보여준다. BMESH는 2012년 현재 공공 메타데이터 저장소에서 추출한 약 45,000개와 주요 의료기관에서 추출한 33,000여 개를 포함해서 14개의 자원에서 추출한 약 114,000여 개의 메타데이터를 제공한다. 본 연구를 통해 다섯 의료기관에서 수집한 다섯 가지 임상문서에서 추출한 527개의 메타데이터가 추가적으로 등록되었다.

Figure 3은 HL-7 CDA 기반의 임상 문서 교환 시에 의미론적 상호운용성을 확보하기 위해서 국제표준 MDR을 적용하기 위한 시스템의 설계도다. 각 의료기관에서 사용되는 EMR 시스템의 서로 다른 구조를 갖는 다기관 임상 문서 교환 상황을 시나리오로 상정하며, 임상문서 교환을 위한 구분표준은 HL-7 CDA의 XML 메시징을 사용한다. 수신문서의 적합성 검증은 다음 3 단계로 수행된다. 1) XML 문서의 적합성을 표준적인 XML 적합성 검증기술로 검증한다. 2) 각 의료기관에서 생성된 CDA 문서의 적합성을 HL-7 template와 OpenEHR (Open Electronic Health Record) archetype으로 검증한다. 3) XML과 CDA 표준으로 검증할 수 없는 메타데이터 항목의 적합성을 검증 한다. CDA의 범위를 넘어서는 추가적인 정보를 표현하기 위해서 XML의 탈출태그를 적용한다. 탈출태그로 감싸진 메타데이터 DE들의 상세한 기술을 연결자가 제공하는 MDR을 통해 검색하여 이들의 ISO/IEC

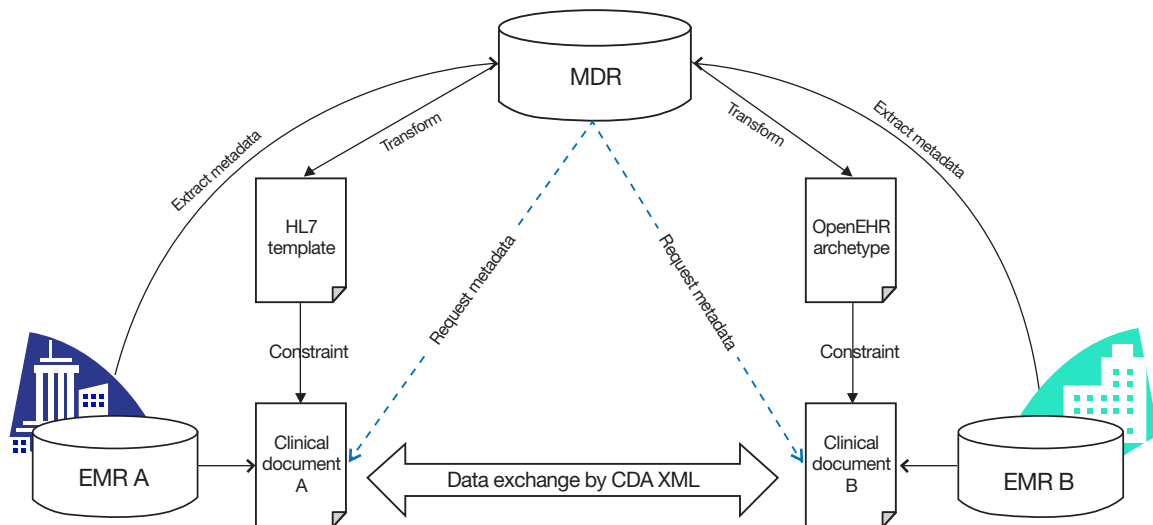


Figure 3. Metadata registry (MDR)-based semantic interoperability during clinical document architecture (CDA)-based clinical document exchange. EMR, electronic medical record; OpenEHR, open Electronic Health Record.

Table 2. Number of metadata instances extracted from five document types from five hospitals

Hospitals	No. of metadata instances in specific clinical documents					Total
	Admission note	Initial medical report	Discharge note	Emergency note	Operation note	
Hospital A	81	33	62	79	32	215
Hospital C	24	28	14	23	22	70
Hospital G	58	26	34	53	11	136
Hospital P	133	40	22	32	12	176
Hospital S	12	6	10	140	9	157
Total	219	114	99	343	50	599

11179 표준에 대한 적합성을 검증한다. 이러한 세 단계를 통해 다기관에서 전송되는 이중 임상문서의 구문론적, 의미론적 상호운용성을 확보할 수 있으면서도 구문표준인 XML 과 CDA의 규칙을 위배하지 않을 수 있다. 이와 같은 MDR 기반 임상문서 교환체계 구축을 위해서는 메타데이터 추출 및 통합, 메타데이터 변환, 표준 기반 임상문서 검증, 임상문서 교환의 네 가지 과정을 거쳐야 한다.

가장 대표적인 5개 임상문서 서식(입원기록지, 초진기록지, 수술기록지, 응급기록지, 퇴원기록지)을 다섯 개의 국내 의료기관(서울대학교병원, 아주대학교의료원, 부산대학교

병원, 가천대 길병원, 전남대학교병원)에서 수집했다. ISO/IEC 11179에서 정의한 방법에 따라 각 임상서식의 각 항목(문항과 답항)들을 메타데이터로 추출한다. 추출된 메타데이터는 메타데이터 구조에 따라 DE concept, value domain 및 DE 수준에서 정보통합을 시행한다. 통합은 3인의 전문가의 검증으로 수행되었고 정제된 메타데이터는 MDR에 등록한다.

MDR에 등록된 메타데이터는 임상 문서 정보교류 시 상호운용성 확보를 위한 기존의 방법 중 HL-7 template와 open Electronic Health Record (OpenEHR) archetype 모델로 변환한다. HL-7 template 검증은 XML 문서의 규칙 기반 검증 언어의 하나인 Schematron을 사용하고, OpenEHR archetype은 자체 언어인 Archetype Description Language (ADL)를 사용한다. 메타데이터에서 검증 모델로 변환하기 위한 변환규칙을 생성하여 프로그램으로 구현한다. 끝으로 다섯 의료기관 사이의 메타데이터 기반 임상문서 정보교류 시스템을 구축하고 교류 문서의 구문론적, 의미론적 상호운용성 검증을 시행한

Table 3. Number of common metadata between hospitals

	Hospital A	Hospital C	Hospital G	Hospital P	Hospital S
Hospital A	215	25	51	38	32
Hospital C	25	70	21	22	16
Hospital G	51	21	136	35	30
Hospital P	38	22	35	176	34
Hospital S	32	16	30	34	157

Table 4. Number of metadata shared in five clinical document types from five hospitals

No. of common metadata	No. of metadata shared in specific clinical documents					Total
	Admission note	Initial medical report	Discharge summary note	Emergency note	Operation note	
Shared by 5 hospitals	1	1	0	0	1	4
Shared by 4 hospitals	0	1	1	7	6	14
Shared by 3 hospitals	14	8	8	22	4	42
Shared by 2 hospitals	29	17	20	27	12	54
Specific in one hospital	175	87	70	287	27	445
Total	219	114	99	343	50	559

다. 현 시점에서 실제 기관 대 기관의 임상문서 정보교류 시스템 운영에는 임상문서의 소유권과 같은 법률적 제한사항이 존재하므로, 본 연구에서는 환자 본인이 의료기관에 진료 기록 사본을 직접 요청하는 시나리오를 가장한 시스템을 설계하고, 다섯 의료기관의 다섯 개 서식에 대한 교환을 지원하는 스마트폰 앱을 개발하여 검증한다.

결 과

1. 임상문서 메타데이터

다섯 의료기관에서 수집한 다섯 서식에서 총 559개의 메타데이터를 추출했다. 거의 모든 문서에 공통인 인구학적 자료 항목은 생략했다. Table 2는 각 의료기관의 서식 별로 추출된 메타데이터의 분포를 보여준다. 가장 많은 경우인 215개에서 가장 적은 경우인 70개까지의 분포를 보인다. 서

식 별로는 응급기록지에서 가장 많은 343개, 수술기록지에서 가장 적은 50개가 추출되었다.

Table 3은 서로 다른 의료기관에서 공통적으로 사용되어 의미론적으로 매핑 가능한 메타데이터의 분포를 보여준다. A와 G 의료기관 사이의 의미론적 매핑률이 가장 높아 51항목 37.5%에 달했다. S와 C 기관 사이의 매핑률은 가장 낮은 경우는 16항목 10.1%였다. 평균적으로 30.4개 약 21%의 매핑률을 보여 의료기관간 임상문서의 차이가 상당한 수준임을 확인할 수 있었다.

임상문서의 종류별로 공통적으로 사용된 메타데이터의 분포는 Table 4에서 볼 수 있다. 인구학적 자료 항목을 제외하면, 모든 의료기관이 함께 공유하는 메타데이터는 전체 559개 중 4개로 희소했다. 두 개 이상의 의료기관이 공유하고 있는 메타데이터는 총 114개 (=4+14+42+54), 전혀 공유되지 않는

특정 기관 고유의 메타데이터는 445개로 약 1:4의 비율을 보였다. 가장 매핑률이 낮은 임상문서는 응급기록지로 약 1:5의 비율을 보였고 수술기록지는 가장 높은 약 1:2의 비율을 보여, 수술기록지의 자연발생적 표준화율이 다섯 문서 중 가장 높은 것으로 볼 수 있다.

2. 메타데이터 변환

임상문서 정보교류 시 기본적인 XML 문서 적합성 평가 후 문서의 구조와 내용을 검증하기 위한 대표적인 방법은 HL-7 template와 OpenEHR archetype 검증모델을 적용한 평가법이다. Table 5는 추출하여 MDR에 등록된 메타데이터를 HL-7 template 및 OpenEHR archetype으로 변환하기 위한 변환규칙 정의서다. 임상문서 교환은 HL-7 CDA를 기반으로 이루어지기 때문에 HL-7 CDA에서 정의한 검증규칙(constraint rules)을 기반으로 각 모델의 구조에 맞춘

Table 5. Constraint rules for HL-7 template and OpenEHR archetypes

HL-7 CDA constraint rules	Metadata registry	HL-7 template	OpenEHR archetypes
Document	DE Set	Document level template	None
Header	Admin record of DE set	Error link	None
Unique identifier	Item identifier	Error link	Header and description section
Language	Language	Error link	Header and description section
Version	Version in admin record	Warning link	Header and description section
Section	Section in DE Set	Section level template	Template
Cardinality	Section	Error link	Definition section
Narrative block (title)	Section title	Warning link	Definition section
Vocabulary	DE set restricted VD	Error link	Definition section
Entry	DE	Entry level template	Archetype
Cardinality	DE relationship	Error link	Definition section
Narrative block	VD	Warning link	Definition section
Vocabulary	VD	Error link	Definition section

OpenEHR, open Electronic Health Record; CDA, clinical document architecture; DE, data element; VD, value domain.

변환규칙을 정의했다. HL-7 template의 경우 문서(document level), 섹션(section level) 혹은 항목 수준(entry level)을 사용하여 하나의 임상문서에 대한 다계층적 검증을 수행한다. MDR 항목과 HL-7 template의 수준을 맞추는 변환수행 결과 총 25개의 문서 수준, 75개의 섹션 수준, 559개의 항목 수준 템플릿들이 생성되었다.

하나의 OpenEHR archetype은 하나의 개념을 표현하는 단위로서, 여러 개의 archetype의 집합을 템플릿으로 정의하고 있지만, 이는 공식적으로 사용되고 있지 않다. MDR의 항목과 OpenEHR archetype의 ADL의 구조에 맞추는 변환 결과 559개의 archetype의 생성되었다. Figure 4는 메타데이터 ID가 DE:25262인 “환자 등록번호”를 HL-7 template (Figure 4A)와 Open-EHR archetype (Figure 4B)으로 변환한 결과를 보여준다.

본 연구의 제안인 HL-7 CDA와 ISO/IEC 11179 국제표준

을 적용한 구문론적 상호운용성과 의미론적 상호운용성을 지원하는 임상 문서 교류 시스템의 검증은 실제 의료 기관에서 사용되는 임상문서를 대상으로 수행했다. 임상문서의 소유권 등의 법률적 제한사항을 피하여 시연은 환자 본인이 의료기관에 자신의 진료 기록 사본을 요청하는 시나리오로 수행되었으며, 다섯 의료기관의 다섯 개 서식에 대한 교환을 지원하는 가상 서버와 안드로이드 및 iOS환경에서 동작하는 스마트폰 앱 클라이언트를 개발했다. 클라이언트는 서울대학교 의과대학 정보의학실에서 개발한 개인건강기록 관리시스템인 Health Avatar CCR+의 추가기능으로 구현되었다. Figure 5는 Health Avatar CCR+의 화면으로 가천대 길병원에서 받은 기록지 중 2011년 8월 31일 발행된 입원 초진 기록지를 수신 받는 장면이다.

결론 및 고찰

전 세계적으로 임상문서뿐만 아니라 다양한 생명과학 자료의 공유와 교류를 위한 상호운용성 확립 노력이 지속되고 있다[14-16]. 최근 빠르게 발전한 메타데이터 기술과 국제 표준화기구를 통한 표준화의 진전은 기존의 구문론적 상호운용성을 넘어선 의미론적 상호운용성 확보를 위한 좋은 기회로 판단된다. 최근 NCI와 USHIK은 의료계에서 메타데이터 국제표준에 기반한 의료정보의 의미론적 상호운용성 확립을 위한 기반기술을 제공하기 시작했다[17]. 그러나 연구 자료를 넘어서 임상문서의 교류에서 이와 같은 시도를 수행한 것은 본 연구가 처음이다. 본 연구는 ISO/IEC 11179 표준을 지원하는 MDR인 BMesh를 개발 구축하고 상당수의 메타데이터를 축적 지원해온 경험을 기반으로 메타데이터의 활용이 단순히 학술자료의 의미론적 일치의 지원을 넘어


```

A
<!-- Patient registration number in Demographic Information section: 1,2,410,113883,1,25262 -->

<!-- errors -->

<pattern id='p-1,2,410,113883,1,25262-errors' see='#p-1,2,410,113883,1,25262-errors'>
  <title>Patient registration number - errors validation phase</title>
  <rule context='[cda:templateId/@root="1,2,410,113883,1,25262"]'>
    <assert test="cda:title">The patient registration number SHALL contain title.</assert>
    <assert test="cda:text">The patient registration number SHALL contain a narrative block.</assert>
  </rule>
</pattern>

<!-- warning -->

<pattern id='p-1,2,410,113883,1,25262-warning' see='#p-1,2,410,113883,1,25262-warning'>
  <title>Patient registration number - warning validation phase</title>
  <rule context='[cda:templateId/@root="1,2,410,113883,1,25262"]'>
    <assert test="string-length(@text) &lt; 5">There is not enough letters in the patient registration number.</assert>
    <assert test="string-length(@text) &gt; 8">There is too much letters in the patient registration number.</assert>
  </rule>
</pattern>

<!-- manual -->

<pattern id='p-1,2,410,113883,1,25262-manual' see='#p-1,2,410,113883,1,25262-manual'>
  <title>Patient registration number- manual validation phase</title>
  <rule context='[cda:templateId/@root="1,2,410,113883,1,25262"]'>
    <report test="false()"></report>
  </rule>
</pattern>

B
archetype (adl_version=1.4)
  BMESH-DEMOGRAPHIC-CLUSTER_patient_registration_number.v01
specialise
  BMESH-DEMOGRAPHIC-CLUSTER_patient_registration_number.v01
concept
  [at0000,1] -- Patient registration number
language
  original_language = <[ISO_639-1::en]>
  >
description
  original_author = <
    ["name"] = <"Yu Rang Park">
    ["organisation"] = <"Seoul National University Biomedical Informatics">
    ["email"] = <"kirarang@snu.ac.kr">
    ["date"] = <"20/05/2009">
  >
  details = <
    ["en"] = <
      language = <[ISO_639-1::en]>
      purpose = <"Representation of a patient registration number">
      use = <"">
      keywords = <"Demographic data", "Patient registration number">
      misuse = <"">
      copyright = <"© BMESH">
    >
  >
  >

```

Figure 4. (Continued)

```

lifecycle_state = <"AuthorDraft">
other_contributors = <>
other_details = <          ["references"] = <">      >
definition
  CLUSTER[at0001] occurrences matches {1..1} matches { -- Patient
    items cardinality matches {1..*}; unordered matches {
      ELEMENT[at0002] occurrences matches {0..1} matches { -- Patient registration number
        value matches {
          DV_TEXT matches {*}
        }
      }
    }
  }
ontology
  term_definitions = <
    ["en"] = <
      items = <
        ["at0001"] = <
          text = <"Patient">
          description = <"Patient">>
        ["at0002"] = <
          text = <"Patient registration number">
          description = <"Representation of a patient registration number">>
      >
    >
  >

```

Figure 4. Example for (A) HL-7 template and (B) open Electronic Health Record archetype on patient registration number.



Figure 5. Screenshot from clinical document exchange using Health Avatar CCR+.

실제 사용되고 있는 임상문서의 의미론적 일치와 문서교류에서의 의미론적 상호운용성을 지원할 수 있음을 제시했다는 점에서 큰 의의가 있다.

임상문서교류를 위한 가장 대표적인 국제표준은 HL-7

CDA다. 최근에는 ASTM CCR 및 HL-7 continuity of care document의 활발한 보급으로 표준 기반의 요약임상정보 교류가 빠르게 활성화되고 있다. 그러나 HL-7 CDA는 구문론적 표준만을 제공하고 의미론적 표준은 제공하지 못한다는 한계가 있다. 이를 보완하기 위해 SNOMED나 UMLS와 같은 통제용어체계를 활용할 수 있으나, 동일한 개념을 일관된 방법으로 부호화하는 규칙이 표준화되지 않아 일관성을 잃고 모호성이 증대되는 경향이 있어 실질적인 의미론적 상호운용성을 지원하지 어렵다. 국제표준에 기반한 MDR의 활용

은, 메타데이터 등록소를 구축해야 하는 과제는 남으나, 임상문서의 각 항목의 의학적 개념을 실제 사용 가능한 실용적인 수준에서 부호화하는 방안을 제공한다는 큰 장점이 있다. 서울대학교 의과대학 정보의학실에서 8년간 구축해온 BMESH

서버는 ISO/IEC 11179 표준을 충분히 지원할 뿐 아니라, 다양한 자료원을 통해 10만개가 넘는 방대한 생명과학 분야 메타데이터를 구축하여 실용적 적용이 가능한 수준에 도달했다. 본 연구의 성공적 시연은 MDR을 이용한 의미론적 상호운용성 확보가 실현 가능함을 보여준다. BMESH는 항목 값을 통제용어체계로 자동 매핑하는 것을 지원하여 기존의 시스템들과의 호환성도 폭넓게 지원한다.

HL-7 CDA를 통한 임상문서의 교환방식은 교환기관 사이에서 미리 사전 정의한 구조와 내용 안에서의 의미론적 정보교류만 가능하다[1,7]. 미리 정의되지 않은 문서나 항목이 존재하거나, 기존의 정의에서 수정된 항목을 추가하는 경우, 해당 임상문서 의미론적 정보교류는 성공하기 힘들다. 사전 정의된 임상문서의 교환만을 지원하는 시스템은 현실적으로 운영하기 불가능하므로 널리 사용되기 어렵다. 임상 의사 혹은 의료기관의 필요에 따라 다양하고 새로운 임상문서가 생성되고 교환되어야 한다. 국제표준을 따르는 MDR의 사용은 이러한 한계를 극복할 수 있다. 새로운 항목을 생성한 의료기관이 공공 MDR에 해당 항목을 ISO/IEC 11179 표준에 따라 메타데이터로 등록하면 스스로의 정보자원 관리에도 일관성을 기할 수 있다. 등록 메타데이터를 외부기관에서도 접근할 수 있도록 설정하고, 본 연구에서 제안한 방식으로 임상문서를 전송하면, 수신기관에서는 신규 등록되어 아직 인지할 수 없는 항목 정보도 MDR 조회를 통해 표준적인 방식으로 그 의미론적 특성을 인식할 수 있게 된다. 그러므로 본 연구의 방법론적 제안은 해당기관 내부에서만 사용하여 기관 내 정보자원의 관리체계의 일부로 사용하는 것도 가능하고, 공공 MDR을 사용하여 기관간 정보교류시의 의미론적 상호운용성을 확보하는 것에 활용할 수도 있다. 또한, 공공 MDR의 공개된 메타데이터의 공유는 모든 기관에서 자유롭게 참조할 수 있어 국가적인 의료정보자원 관리체계의 확립에 중요한 하부구조 역할을 할 수 있다.

본 연구에서는 기존에 제안된 상호운용성 기술과의 호환성 지원을 위해 메타데이터와 HL-7 template 및 OpenEHR archetype과의 변환규칙을 제공한다. 임상문서 각 항목에 대한 메타데이터를 국제표준을 지원하는 MDR로 관리하고 다양한 상호운용성 기술을 지원하여 서로 다른 방법론을 사용

하는 의료기관 사이의 임상문서 교환 시에도 끊임 없는 의미론적 상호운용성 지원이 가능하다. 본 연구의 범위를 벗어나지만 ISO 20944: MDR interoperability & binding과 같은 표준은 ISO/IEC 11179 기반으로 개발된 MDR 간의 자료교환 방법으로 coding binding, API binding, API protocol binding 등에 대한 상세내역을 제공하고 있어서 좀더 정교한 시스템의 구축이 가능하다.

Acknowledgement

This research was supported by the basic science research program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2010-0028631). YRP's education grant was supported by the Ministry of Health and Welfare, Republic of Korea (A112020).

핵심용어: 메타데이터; 의미론적 상호운용성; 메타데이터 저장소; ISO/IEC 11179; 임상문서 교환

REFERENCES

1. Dolin RH, Alschuler L, Boyer S, Beebe C, Behlen FM, Biron PV, Shabo Shvo A. HL7 clinical document architecture, release 2. J Am Med Inform Assoc 2006;13:30-39.
2. American Society for Testing and Materials. Standard specification for continuity of care record (CCR) [Internet]. West Conshohocken: ASTM International; 2005 [cited 2012 Jun 28]. Available from: <http://www.astm.org/Standards/E2369.htm>.
3. Clinical Data Interchange Standards Consortium. Operational data model [Internet]. Round Rock: Clinical Data Interchange Standards Consortium; 2012 [cited 2012 Jun 28]. Available from: <http://www.cdisc.org/odm>.
4. Hurrell MJ, Monk TG, Nicol A, Norton AN, Reich DL, Walsh JL. Implementation of a standards-based anaesthesia record compliant with the health level 7 (HL7) clinical document architecture (CDA). J Clin Monit Comput 2012;26:295-304.
5. Otter-Nickerson B. HIE: the interoperable way to deliver quality healthcare: finally, the concept of health information exchange is being more concretely defined and accepted by the healthcare community. Health Manag Technol 2011;32:19.
6. Richesson RL, Krischer J. Data standards in clinical research:

- gaps, overlaps, challenges and future directions. *J Am Med Inform Assoc* 2007;14:687-696.
7. Smith B, Ceusters W. HL7 RIM: an incoherent standard. *Stud Health Technol Inform* 2006;124:133-138.
 8. Humphreys BL, Lindberg DA, Schoolman HM, Barnett GO. The unified medical language system: an informatics research collaboration. *J Am Med Inform Assoc* 1998;5:1-11.
 9. Cote RA, Rothwell DJ, Palotay JL, Beckett RS, Brochu L; College of American Pathologists; American Veterinary Medical Association. The systematized nomenclature of human and veterinary medicine: SNOMED international. 3rd ed. Northfield: College of American Pathologists; 1993.
 10. Solbrig HR. Metadata and the reintegration of clinical information: ISO 11179. *MD Comput* 2000;17:25-28.
 11. Agency for Healthcare Research and Quality. United States Health Information Knowledgebase [Internet]. Rockville: Agency for Healthcare Research and Quality; 2012 [cited 2012 Jun 18]. Available from: <http://ushik.ahrq.gov>.
 12. Warzel DB, Andonaydis C, McCurry B, Chilukuri R, Ishmukhamedov S, Covitz P. Common data element (CDE) management and deployment in clinical trials. *AMIA Annu Symp Proc* 2003:1048.
 13. International Organization for Standardization; International Electrotechnical Commission. Information technology: specification and standardization of data elements. Part 3. Basic attributes of data elements. ISO/IEC 11179-3. Geneva: ISO/IEC; 1994.
 14. Fridsma DB, Evans J, Hastak S, Mead CN. The BRIDG project: a technical report. *J Am Med Inform Assoc* 2008;15:130-137.
 15. Buetow KH, Niederhuber J. Infrastructure for a learning health care system: CaBIG. *Health Aff (Millwood)* 2009;28:923-924.
 16. Costa CM, Menarguez-Tortosa M, Fernandez-Breis JT. Clinical data interoperability based on archetype transformation. *J Biomed Inform* 2011;44:869-880.
 17. Clinical Data Interchange Standards Consortium. Clinical data acquisition standards harmonization [Internet]. Round Rock: Clinical Data Interchange Standards Consortium; 2012 [cited 2012 Jun 28]. Available from: <http://www.cdisc.org/cdash>.



Peer Reviewers' Commentary

대형병원은 물론 개원가까지 진료에서 전자의무기록 시스템 (EMRS)의 보급이 보편화되어 있는 현시점에 반드시 필요한 작업 중 하나가 데이터 교류를 위한 용어의 표준화 작업이다. 본 논문에서는 이러한 표준화 작업의 일환으로 현재까지 제안된 교류 표준인 HL-7 CDA 등의 체계가 구문론적 표준을 지원할 뿐 의미론적 표준이 되기 어렵다는 점을 극복하기 위해 최근 개발된 메타데이터 기술을 이용하여 의미론적 상호운용성을 지원할 수 있는 체계를 구축하였으며, 이를 5개 대형병원 임상문서의 교류를 위해 시도하였다. 본 연구는 학술자료 수준이 아닌 실제 임상자료, 즉 국내 주요 병원에서 사용하는 임상문서의 교류에서 의미론적 상호운용성의 가능성을 보여준 것으로, 이것은 아마도 세계적으로도 최초의 시도일 것이다. 본 연구에서 제안된 정보교류시스템은 각 병원의 독립성과 보안성도 고려하고 있어 실용화에도 큰 문제가 없어 보인다. 또한, 구축된 MDR을 공공 시스템으로 운영하면 국가적인 의료정보자원 관리체계 확립의 중요한 하부구조로 기여할 것이다.

[정리: 편집위원회]