

국내 의료 시술 용어에 대한 OMOP 표준용어로의 매핑 가이드라인

아주대학교 의료정보학과

**버전 1.0
11, 2020**

TABLE OF CONTENTS

1. 서론	8
2. 의학 용어 체계	3
2-1. OMOP vocabulary.....	3
2-2. SNOMED CT	5
2-3. CPT4	6
2-4. ICD10PCS.....	7
3. 용어 매핑	8
3-1. OMOP 표준 용어.....	8
3-2. 매핑 프로세스.....	11
3-3. Usagi.....	13
4. 시술 용어 표준화 가이드라인	15
4-1. 시술 용어 표준화 결과.....	15
4-2. 3자 검토 사례 및 결과	16
4-3. 표준화된 시술 용어 적용방안	20
5. REFERENCES.....	22

1. 서론

의학용어는 의료 정보를 표현하는데 가장 기본이 되는 요소로써, 의료 행위를 식별하고, 분류하고, 소통하기 위한 목적으로 합의된 어휘 및 전문용어들을 지칭한다. 복잡한 의료행위에 대한 이해와 질병 분류를 위해서는 의학용어들이 체계적으로 정리되어야만 하고, 이렇게 체계화된 의학용어들이 모여서 의학용어집을 형성하게 된다. 의학용어집은 의료 행위(진단, 검사, 시술, 등) 및 질병분류를 명확히 이해하고 소통하기 위해서는 필수적이며, 다양한 이해관계자들(의사, 보건의료종사자, 환자 및 보험자)의 효율적인 의사소통에 기여하게 된다. 이러한 의학용어집은 각 기관 내에만 존재하는 기관코드부터 국제적으로 통용되는 용어체계(SNOMED, ICD, CPT4 등)까지 다양하게 존재하고 있으며, 동일한 의료행위일지라도 각 용어체계에서는 다양한 방식으로 표현되고 있다.

용어(Vocabulary)는 Observational Health Data Sciences and Informatics (OHDSI, "Odyssey"라고 발음함) 연구 네트워크에서 가장 기초적이고 중요한 역할을 담당하고 있다. Observational Medical Outcome Partnership(OMOP) 표준 용어는 상이한 용어로 기록된 의학적 사건에 대해 표준화된 용어로 정의함으로써, 분석 방법, 정의, 결과를 표준화하여 원격 연구를 가능하게 함과 동시에 분산 연구망(Distributed Research Network; DRN)을 가능하게 한다.

초기에 OHDSI 컨소시엄을 형성하던 대부분의 기관들은 미국을 대표하여 주로 미국에서 사용되는 의학용어 체계만을 포함하는 경향이 있었으나, 현재는 OHDSI 컨소시엄이 글로벌 네트워크를 형성하고, 확장되고 있어 다양한 국제 의학용어집을 아우르고 있다.

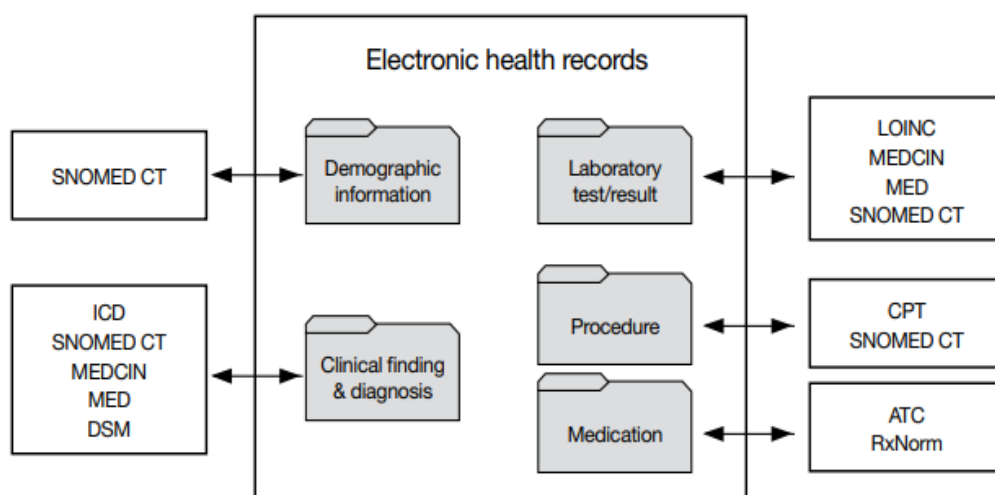


그림 1. 의료 데이터에서 널리 사용되는 용어체계

본 가이드라인에서는 시술 행위에 대해 한국의 사실상 표준(de facto standard)으로 사용되고 있는 Electronic Data Interchange(EDI)코드를 OHDSI 에서 정의하는 OMOP 표준 용어로의 매핑(Mapping) 방안을 제시함으로써, 국내 시술 의학용어체계의 매핑 방법, 어려움, 해결 방안 등을 논의하고자 한다. 더 나아가, OMOP 표준 용어집으로의 매핑을 통해 국내 시술 의학용어의 국제적 활용 가능성을 높이하고자 한다.

2. 의학 용어 체계

2-1. OMOP vocabulary

OHDSI 는 서로 다른 의학용어로 표현된 사건들을 표준화 시키기 위해 다양한 용어체계들 간의 위계를 정하고 있다. 각 용어체계들은 서로 간의 상하위 관계뿐만 아니라 수평 관계도 정해지게 되고, 표준으로 사용되는 용어와 비표준으로 사용되는 용어들이 존재하게 된다. 이렇게 규정되는 용어들을 OMOP vocabulary(OMOP 용어)라고 지칭한다. 일반적으로 OHDSI 의 Vocabulary 팀에서 용어체계들 간의 관계를 규명하는 일을 주도하고 있고 개별 연구자도 매핑 작업을 진행한 후, OHDSI 측에 공식적으로 OMOP vocabulary 편입을 검토 및 요청할 수도 있다. OMOP vocabulary 는 ATHENA (<https://athena.ohdsi.org>)에서 조회 가능하다.

전 세계적으로 사용되고 있는 대표적인 의학용어 체계는 Current procedural terminology 4th edition(CPT4), dictionary for medicines and devices(dm+d), Healthcare common procedure coding system(HCPCS), International classification of disease, 10th version(ICD10), ICD10 procedure classification system(ICD10PCS), ICD10 clinical modification(ICD10CM), Japan medical data center(JMDC), Korean standard classification of disease(KCD7), EDI, National drug code(NDC), RadLex, Structured product labeling(SPL), Systematized nomenclature of medicine clinical terms(SNOMED CT) 등이 사용되고 있다. 이들은 특정한 목적을 띄고 의약품 관리 등을 위해 정부에 의해 만들어지거나, 전자의무기록이나 의료보험 청구용과 같이 민간 주도로 만들어지기도 한다. 위에 언급된 의학용어 체계와 같이 널리 사용되고 있는 것들은 OMOP 용어 사전집에 등재되어 있으며, 각 용어 사전집에는 적게는 2 만개에서 많게는 100 만개 이상의 의학 용어들이 등록되어 있다. 특히, 의료 시술에 해당하는 의학 용어를 잘 분류하는 용어 체계는 SNOMED CT, CPT4, ICD10PS 등이 있다.

표 1. OMOP 용어집(vocabulary) 내 포함 용어체계 및 용어의 수

	OMOP 내의 표준 용어 체계	용어의 수
1	RxNorm Extension	2,031,387
2	NDC	1,016,526
3	SNOMED	886,485
4	SPL	482,100
5	dm+d	378,491
	EDI	313,431
6	RxNorm	292,671
7	LOINC	250,287
8	OSM	203,339

9	DPD	193,647
10	ICD10PCS	194,339
11	ICD10CM	96,308
12	Read	108,696
13	AMT	76,706
14	ICDO3	63,667
15	VA Product	54,768
16	CIEL	50,881
17	BDPM	44,376
18	NDFRT	37,486
19	JMDC	35,962
20	NAACCR	34,473

표 2. OMOP CDM 기록에 사용되는 용어

구분	원천 자료	OMOP CDM 표준
Conditions	KCD	SNOMED CT
Drugs	EDI 약품코드	RxNorm
Procedures	EDI 수가코드	CPT4, SNOMED CT, ICD9PCS, HCPCS
Measurements	검사 코드, 검체 코드	LOINC, SNOMED CT
Devices	보험 EDI 치료 재료 코드	SNOMED CT, HCPCS

2-2. SNOMED CT

1965 년 SNOP(Structured Nomenclature of Pathology, SNOP)의 구조적 명명법 이 시작된 이래로 내용과 기본 표현 측면에 중점을 두고 다양한 버전의 SNOMED 가 개발되었다. 초기 SNOP 및 SNOMED 버전은 사후 조정이 가능한 다축 시스템이었는데, SNOP 은 약 15,000 개의 의료대상 및 프로세스 개념으로 구성되었으며, SNOMED-2, SNOMED International 로의 발전을 거듭하여 1990 년대 중반에는 약 150,000 개의 개념을 포함하였다.

SNOMED CT 는 미국 병리학회의 SNOMED Reference Terminology (SNOMED RT)와 영국 NHS 의 CTV-3 의 내용이 통합된 것으로 2002 년 1 월에 출시되었다. SNOMED CT 의 초기 버전은 278,000 개의 유효 개념으로 구성되었으며, 2007 년부터는 IHTSDO (International Health Terminology Standards Development Organization)에서 권한을 얻어 개발 및 유지 중에 있다.

현재 약 30 만 개의 개념을 가진 가장 큰 생체 의학 Ontology 이며, 다계층 분류학적 순서로 배열되어 19 개 그룹으로 나누어져 있다. Concepts, Hierarchies, Attributes, Identifiers, Descriptions, Relationships 의 핵심요소로 구성되어 있다.

SNOMED 의 용어들은 다른 용어집에 포함된 용어들로부터 여러 개의 유의어와 동의어 관계를 가질 수 있으며 다른 용어들과 속성 관계를 가질 수도 있다.

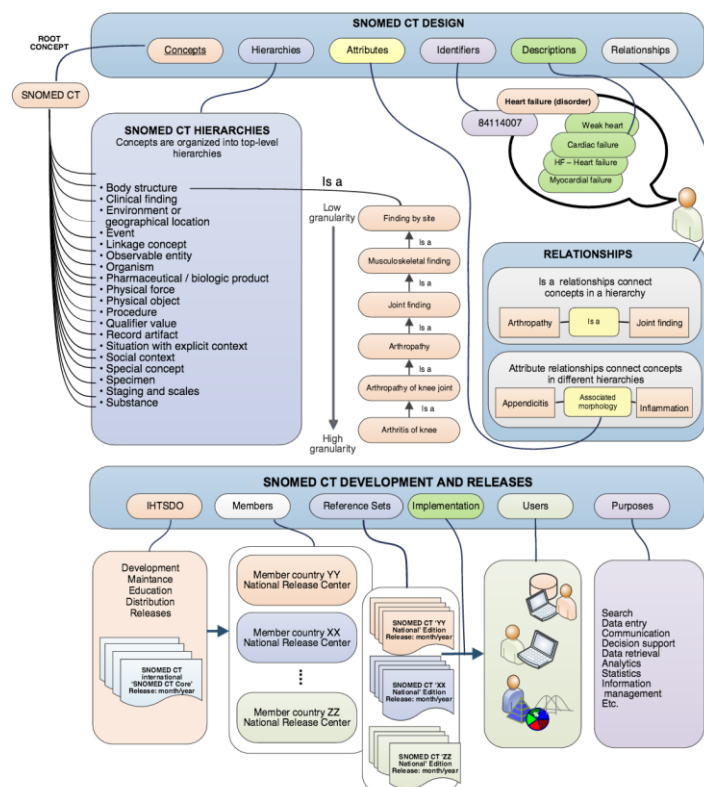


그림 2. SNOMED CT 용어체계의 구성을 나타내는 도식도

2-3. CPT4

CPT®(Current Procedural Terminology) 4th Edition (이하 CPT4)는 AMA(American Medical Association)에서 1966 년 진단 및 치료 절차를 기술하기 위해 개발한 코딩 체계로 현재 미국의 공공 및 민간의료보험에서 외래 환자의 진단, 수술, 치료와 관련된 비용청구 및 상환에서 가장 널리 활용되고 있는 용어체계 중 하나이다.

CPT4 코드의 개발 및 관리는 CPT® 편집 패널이 주도하는 엄격하고 투명한 개방적인 프로세스를 통해 이루어지며, 임상적으로 유효한 코드를 정기적으로 발급, 업데이트, 유지될 수 있도록 하여 임상 실무와 혁신을 의학에 정확하게 반영하려는 노력을 하고 있다.

CPT4 코드는 카테고리에 따라 5 자리의 숫자 또는 영문으로 구성되며, 특정시술을 하게 된 사유와 비용에 따라 코드가 다르게 부여된다. 예를 들어, 아데노이드 절제술을 시행한 경우, 1 차적 혹은 2 차적 시술인지와 12 세 이하 혹은 12 세 이상인 경우에 따라 5 번째 자리의 숫자로 구분하게 된다.

표 3. CPT4 코드의 구성 원리

CONCEPT_ID	CONCEPT_CODE	NAME	CLASS	CONCEPT	VALIDITY	DOMAIN	VOCAB
2108819	42830	Adenoidectomy, primary; age 12 or over	CPT4	Standard	Valid	Procedure	CPT4
2108820	42831	Adenoidectomy, primary; younger than age 12	CPT4	Standard	Valid	Procedure	CPT4
2108821	42835	Adenoidectomy, secondary; age 12 or over	CPT4	Standard	Valid	Procedure	CPT4
2108822	42836	Adenoidectomy, secondary; younger than age 12	CPT4	Standard	Valid	Procedure	CPT4

현재 약 16,200 여 개의 CPT4 코드가 사용되고 있으며, 연간 3 번의 회의를 통해 새롭게 추가하거나, 기존 코드의 유지 및 보수에 대해 논의를 한다. CPT4 코드는 현재 미국에서 처치 분야에 대하여 진료정보교류 전송표준 코드로 사용되고 있다.

2-4. ICD10PCS

ICD10PCS 는 입원환자 시술을 보고하기 위해 사용되는 국제 의료용어 분류 시스템으로, 미국에서 입원 환자 시술 코드를 유지 및 보수하는 The Centers for Medicare and Medicaid Services(CMS) 에서 ICD9CM 에는 포함되어 있지 않던 시술행위를 추가하기 위하여 설계 및 개발되었다.

ICD10PCS 코드는 7 개의 영어 또는 숫자로 구성되어 있으며, 각각을 "자리(Character)"라고 부른다.

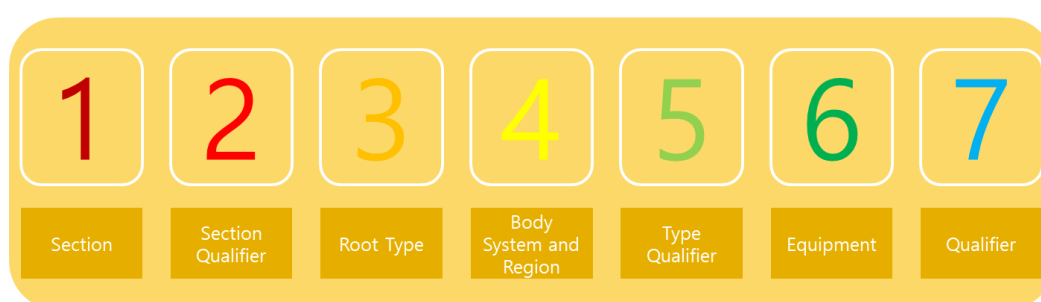


그림 3. ICD10PCS 코드의 원리

첫 번째 자리는 '섹션'으로, 환자에 대한 의료 서비스 행위를 구분한다(수술, 영상, 핵의학 등). 두 번째에서 일곱 번째까지의 자리는 각각 신체계통, 루트수술, 신체부분, 접근방법, 장치, 한정어를 의미한다. 각 자리는 개별 단위의 문자나 숫자로 표시되면 34 개의 값(Value)이 들어갈 수 있다(숫자 : 0-9, 문자 : A-H, J-N, P-Z). 숫자 0 과 1 과의 혼동을 피하기 위해 O 와 I 는 제외된다.

예를 들어, 59 세의 여성이 뇌졸중으로 입원하고 이틀 후, 실어증의 증세를 보여 보완대체의사소통을 통하여 발성치료 재활을 실시할 때 부여되는 코드인 "F06Z3MZ"는 다음과 같이 구성된다:

1. F(Physical and Rehabilitation Diagnostic Audiology)
2. 0(rehabilitation)
3. 6(Speech Treatment)
4. Z(None)
5. 3(Aphasia)
6. M(Augmentative/alternative communication)
7. Z(Zone)

ICD10PCS 의 예는 소수점이 없으며, 약 80% 이상이 첫 번째 섹션인 "의료 및 외과"에 있다.

3. MAPPING

3-1. OMOP 표준 용어

OHDSI 커뮤니티에서 OMOP vocabulary 에 포함된 OMOP 표준 용어는 분산 연구망 네트워크에서 가장 기초적인 부분이자, 공통 데이터 모델의 핵심적인 부분이다. OMOP 표준 용어는 상이한 국가와 데이터베이스마다의 의미들을 동종의 의미로 통하도록 정의함으로써 분석에 대한 정의, 방법, 결과를 표준화하여 진정한 의미의 분산 연구망을 구축 가능케 한다.

의학 용어는 새로운 기술과 약물, 재료 등이 지속적으로 개발되고 사라지고 있음에 따라 규모의 범위와 복잡성이 확대되고 있다. 또한, 각 기관의 사용 목적에 따라 의학 용어의 분류나 체계 등이 상이한데, 이러한 용어집들의 상당수는 공공기관 혹은 정부 기관에서 장기적으로 관리하고 있다. OHDSI 커뮤니티의 의학 용어집은 다음과 같은 목적을 가진다.

- (1) 커뮤니티에서 사용되는 모든 용어의 공통 저장 자료
- (2) 연구에 사용하기 위한 표준화된 매핑

이미 많은 용어집들이 관찰 데이터에서 사용되어 오고 있으며, 용어집의 복잡한 구성과 관리, 많은 이해관계자의 의견을 수렴해야 한다는 점 때문에 OHDSI 는 모든 용어를 처음부터 구축하는 것보다 기존 용어집의 내용들을 차용하고자 하였다. 이러한 이유로, OHDSI 는 몇 가지 특별한 내부 용어만 생성할 뿐, 이외의 용어들은 모두 기존에 존재하는 용어집으로부터 차용하여 쓰고 있다.

OMOP 공통데이터모델 (Common Data Model, 이하 CDM)의 모든 임상 기록은 하나의 Concept 으로 표현되고 있으며, 이는 각 의료 기록의 의미를 나타낸다. OMOP CDM 의 모든 Concept 은 OMOP vocabulary 에 정의되어 있다. 환자의 모든 의료 경험이나 의료시스템의 일부 관리 정보를 기록하기 위해, OMOP vocabulary 는 매우 포괄적인 용어를 포함하고 있다.

각 Concept 들의 ID 는 무의미한 정수로 이루어져 있으며, 원천 용어집에서 사용되던 코드를 대신하여, OMOP CDM 내에서 중복되지 않도록 구성되어있다. 이 외, 차용한 기존 용어집에서 사용되던 용어의 코드, 용어명, 용어집 명 등의 정보가 기록되어있다.

또한, 각 Concept 에는 DOMAIN_ID 필드에 도메인이 할당되는데, 정수로 구성된 Concept ID 등과 달리, 대소문자를 구분하면서 길이가 짧은 고유한 문자 ID 이다. 이러한 각 도메인의 예로는 "Condition", "Drug", "Procedure", "Visit", "Device", "Specimen" 등이 있다. 모든 Concept 들은 각 주어진 도메인에서만 기록되어야 한다.

CONCEPT_ID	313217	Primary key
CONCEPT_NAME	Atrial fibrillation	English description
DOMAIN_ID	Condition	Domain
VOCABULARY_ID	SNOMED	Vocabulary
CONCEPT_CLASS_ID	Clinical Finding	Class in vocabulary
STANDARD_CONCEPT	S	Standard, Source of Classification
CONCEPT_CODE	49436004	Code in vocabulary
VALID_START_DATE	01-Jan-1970	Valid during time interval
VALID_END_DATE	31-Dec-2099	
INVALID_REASON		

그림 4. Concept 테이블의 구성

하나의 임상 사건에 대하여 동일한 의미를 갖는 서로 다른 Concept 이 존재하기 때문에, OMOP vocabulary 에서는 표준 Concept (Standard concept)을 설정한다. 심방 세동의 예를 들면, MeSH 코드 "D001281", CIEL 코드 "148203", SNOMED 코드 "49436004", KCD7 코드 "I48", ICD9CM 코드 "427.31" 및 Read 코드 "G573000"로 각기 다르게 표현하고 있으나, "Condition" 도메인에선 SNOMED 의 코드인 "49436004"만을 표준으로 정하여 해당 질환을 의미한다. 이 외의 용어들은 모두 Non-standard Concept 로, OMOP CDM 에서는 하나의 임상 사건을 의미하는 데 사용되진 않지만, 여전히 표준 용어집의 일부를 담당하고 있으며 원천 데이터에서 발견된다. 이러한 Non-Standard Concept 들을 Standard Concept 으로 연결하는 작업을 매핑(Mapping)이라고 부른다. 각 도메인에서 표준으로 사용되고 있는 Concept 는 다음 표와 같다.

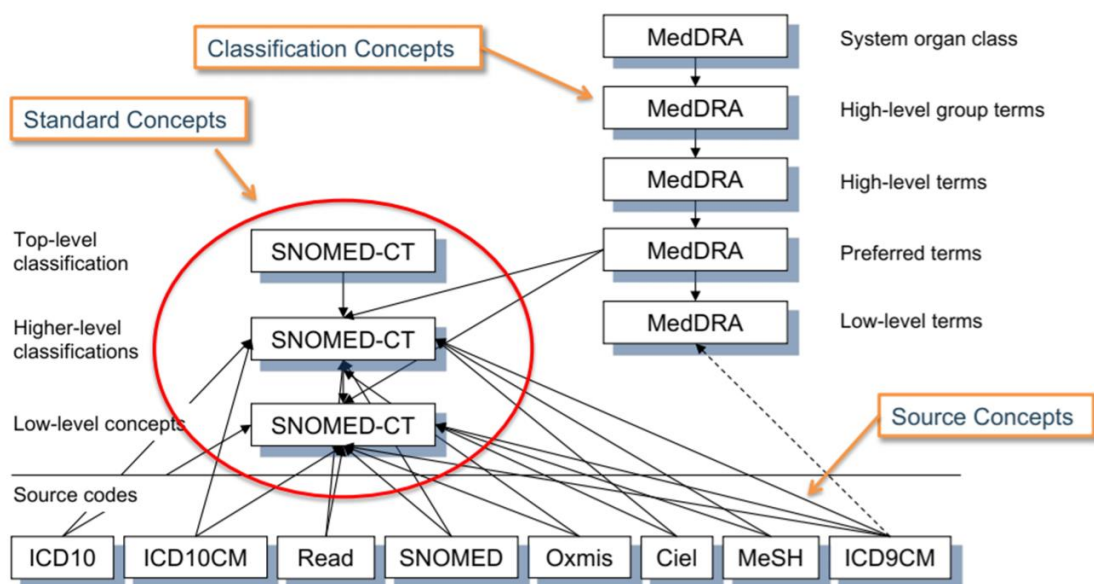


그림 5. OMOP 용어집 내 용어 체계 간 표준 설정

표 4. OMOP 에서 각 도메인 별로 사용되는 용어 체계

Domain	For Standard Concepts
Condition	SNOMED, ICDO3
Procedure	SNOMED, CPT4, HCPCS, ICD10PCS, ICD9Proc, OPCS4
Measurement	SNOMED, LOINC
Drug	RxNorm, RxNorm Extension, CVX
Devices	SNOMED
Observation	SNOMED
Visit	CMS Place of Service, ABMT, NUCC

표 5. OMOP 용어사전집의 용어체계 별 OMOP 에서 사용되는 Standard 용어의 수와 그 비율

OMOP 내 표준 용어체계	Standard 용어 수	Standard 용어의 비율(%)
SNOMED	540,077	61
CPT4	12,686	78
ICDO3	59,619	94
HCPCS	8,434	78
ICD10PCS	194,330	99
ICD9Proc	2,237	48
OPCS4	2,372	22
LOINC	105,939	42
dm+d	22,445	6
NDC	10,587	1
RxNorm	147,283	50
RxNorm Extension	1,757,038	87

이렇게 표준으로 연결된 용어들간의 관계를 매핑(Mapping) 관계라고 명시한다. 또한, 각 Concept 들은 서로 상·하위의 계층 구조로 구성되어 있고, 이러한 관계들은 CONCEPT ANCESTOR 테이블에 저장되고 있다. 예를 들어, 심방 세동의 경우 가장 상위 Concept 으로 순환기 체계의 질병이라는 Concept 이 존재하고, 아래에는 만성 심방 세동 혹은 발작성 심방 세동 등의 Concept 이 존재할 수 있다. 하지만, 이러한 관계는 Drug 과 Condition 도메인에선 높은 품질을 보장하고 있지만, Procedure, Measurement, Observation 에서는 부분적으로만 구축되어 있다.

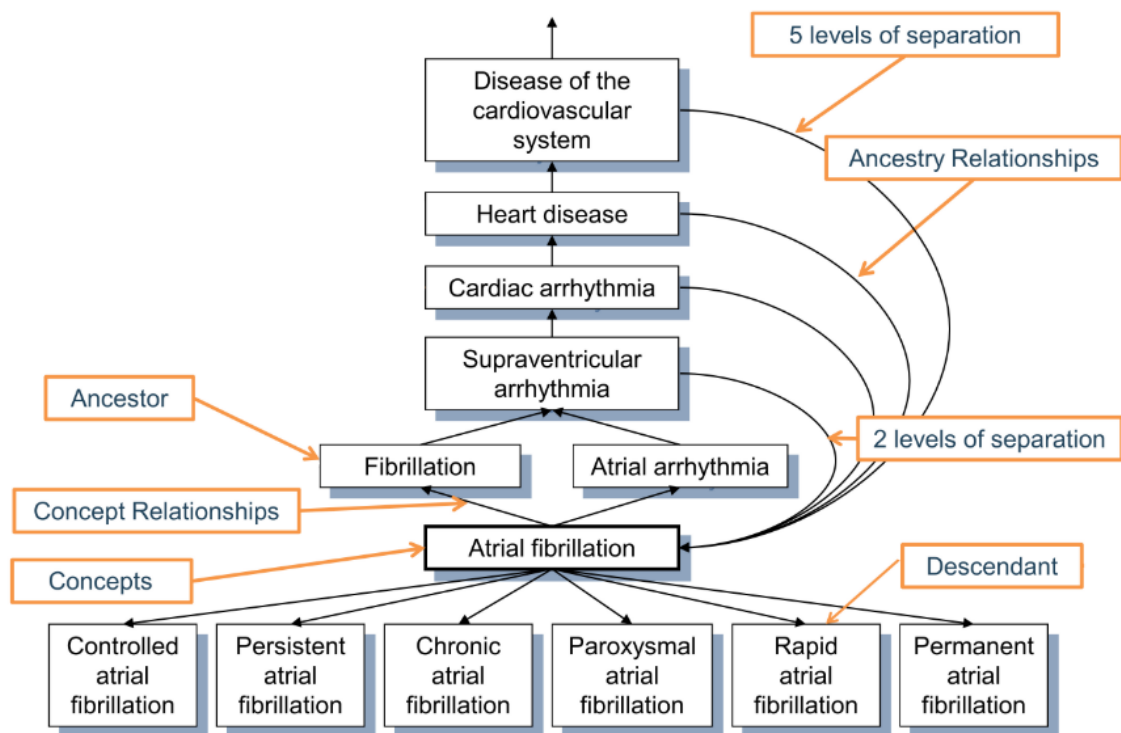


그림 6. OMOP 용어집 내 Concept 간 계층 구조

3-2. 매핑 프로세스(Mapping process)

용어 매핑을 통해 실제로 의미가 동일한 두 concept 은 도메인 혹은 원천 어휘집과 상관없이 특정 관계를 명시할 수 있다. 관계의 특성은 CONCEPT RELATIONSHIP 테이블의 RELATIONSHIP_ID 필드에서 대소문자를 구분하는 고유한 짧은 ID 로 표시한다. 두 concept 에 대한 관계를 나타내는 이 ID 는 CONCEPT_ID_1 과 CONCEPT_ID_2 에 대하여 상호 대칭 관계로 내용이 교환된다. 예를 들면, CONCEPT_ID_1-CONCEPT_ID_2 의 "Maps to" 관계는 CONCEPT_ID_2-CONCEPT_ID_1 의 "Mapped from"과 대칭의 관계를 갖는다. 또한, 관계의 의미 중 "동등 개념 Equivalent concept"와 같이 동일한 개념을 나타낼 때도 있지만, 더 세밀한 병명을 나타내어 계층적으로 포함된 하위 개념을 의미하는 "하위 개념 Descendant concept"을 가질 때도 있다. 일부 매핑은 둘 이상의 표준 concept 에 연결된다. 예를 들어, ICD9CM 의 070.43 "Hepatitis E with hepatic coma"는 SNOMED 의 235867002 "Acute hepatitis E"뿐 아니라, SNOMED 의 72836002 "Hepatic coma"에도 매핑이 되어있다. 이는 원천 용어집에서 명시된 의미가 조합되어 (hepatitis + coma) 선 반영된 경우가 있기 때문이다. 이 외 매핑에 관련된 자세한 규칙은 Book of OHDSI 혹은 OHDSI wiki 에서 확인할 수 있다.

이러한 매핑 관계는 기존의 OHDSI 커뮤니티 혹은 기존에 존재하던 지식을 기반으로 명시하여 공개하고 있으나, 한국에서 고유하게 사용하고 있는 KCD 나 EDI 코드의 경우는 아직 매핑이 되어있지 않다. 물론 KCD 의 경우, 대부분 ICD 와 동일한 분류 체계 및 code 를 사용하고 있기 때문에, 일부 한국에서만 사용되고 있는 특이한 용어를 제외하면 KCD-ICD-SNOMED 순서를 통해 별도의 매핑 작업이 필요 없이 관계를 지을 수 있지만, EDI 의 경우 새롭게 그 관계를 명시해야 한다.

이처럼 기존 concept relationship 테이블에 의해 정의되지 않은 매핑 관계는, source to concept map 이라고 하는 테이블에서 그 관계를 표현할 수 있다. 예를 들어, 한국의 EDI 체계에서 "HE109"는 "기본자기공명영상진단-척추-경추-일반"을 의미하는데, 이 경우 SNOMED 의 "241646009"인 "MRI of cervical spine"와 매핑 될 수 있고, 각 각 SOURCE_CONCEPT_ID 필드와 TARGET_CONCEPT_ID 필드에 명시될 수 있다.

이러한 매핑 작업은 OMOP CDM 의 품질 뿐 아니라 다기관 공동 연구의 질에 많은 영향을 줄 수 있기 때문에 정확한 코드로 매핑이 필요하다. 용어 매핑의 품질을 높이기 위해서는 2 인 1 조로 동일한 코드에 대해 독립적으로 매핑 작업을 진행하는 것이 권유된다. 이후 매핑이 완료된 결과에 대한 교차 검토를 통해, 불일치한 결과는 논의를 통해 합의하여 결정하거나, 이것이 이루어지지 않는 경우 3 자의 검토를 통해 최종 매핑 코드를 승인하는 것이 권유된다. 일반적으로 매핑 작업은 의무기록사, 간호사, 의사 등과 같은 의학적 지식을 보유한 전문가에 의해 이루어져야 한다.

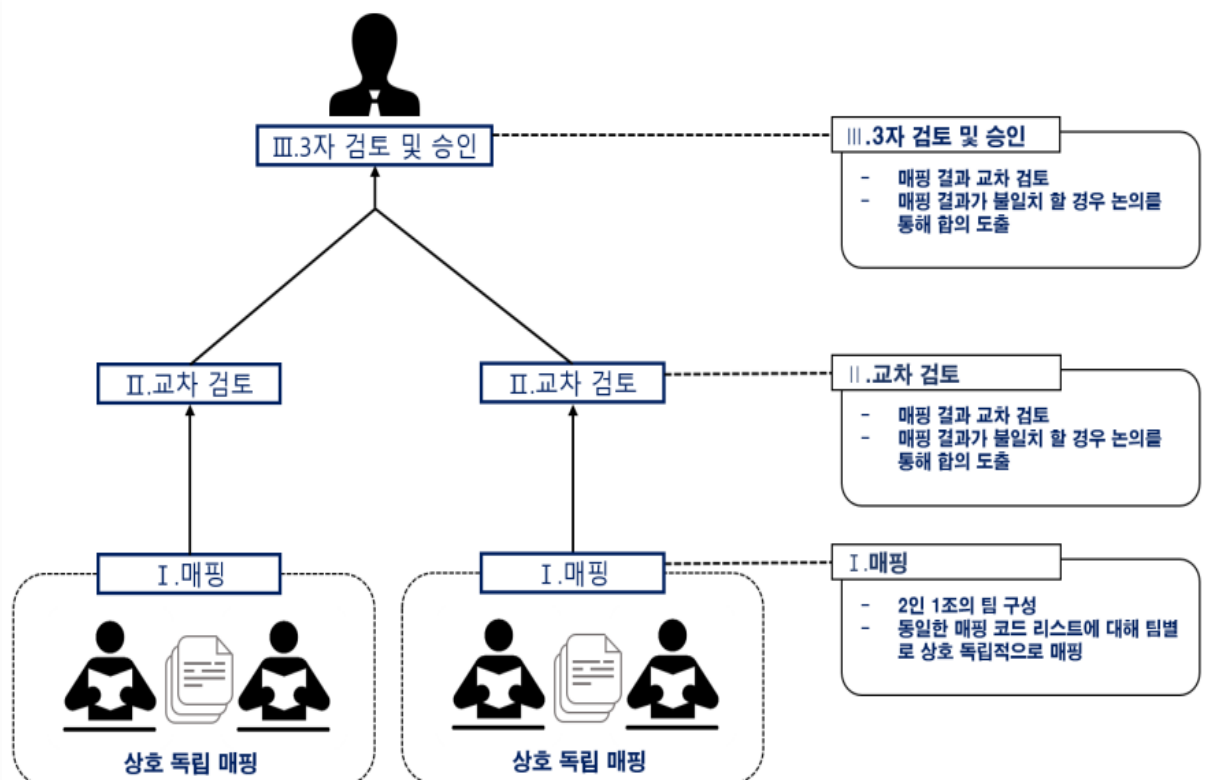


그림 7. 용어 매핑 절차

3-3. Usagi

위에서 언급한 것과 같은 매핑 작업은 엑셀이나 R 과 같은 도구를 사용하여 수기로 매핑하 방법도 있지만, OHDSI 컨소시움에서 개발한 Usagi 라는 도구를 사용할 수 있다. Usagi 는 용어 매핑 절차를 도와주는 도구로, 원천 코드 설명에 입력된 단어의 유사도에 기반하여 적절한 표준 용어를 추천해주는 도구이다. 만약 원천 코드의 설명이 영문이 아닐 경우엔, Google Translate 를 통해 해당하는 용어의 영어 번역을 확인할 수 있다. 또한, Usagi 의 용어 추천이 정확하지 않을 경우엔 사용자가 직접 적절한 표준 용어를 찾을 수 있다. Usagi 는 해당 GitHub repository (<https://github.com/OHDSI/Usagi>)를 통하여 사용할 수 있다. Usagi 소프트웨어를 사용하기 위한 순서는 다음과 같다:

1. 원천 시스템에서 사용하는 용어를 업로드
2. Usagi 의 단어 유사도 계산법을 이용하여 추천된 표준 용어로의 매핑
3. 매핑 결과를 용어의 SOURCE_TO_CONCEPT_MAP 으로 내보냄

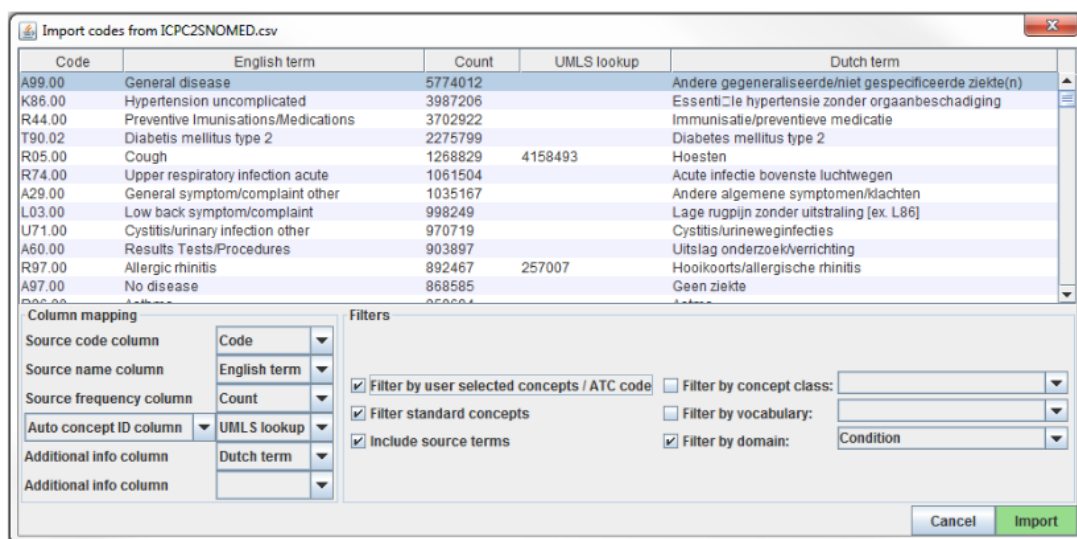


그림 8. Usagi 프로그램의 원천 용어 코드 입력 화면

4. 시술 용어 표준화 가이드라인

4-1. 시술 용어 표준화 가이드라인

그 동안 국내에서 의료용어 표준화를 위해 수 많은 노력을 했지만, 실제로 의료 현장에 보급되어 사용되고 있는 코드체계는 드문 상황이다. 또한, 각 의료기관별로 OCS, EMR, PACS 등과 같이 의료정보를 수집하고 저장하는 환경이 상이하기 때문에 의료정보를 기록하는 그 식별자 역시 대부분이 상이하다. 이러한 의료정보 환경을 고려했을 때, EDI 라는 사실상 표준으로 사용되는 코드 체계를 이용하는 것이 가장 효율적인 방법 중 하나이다.

표 6. EDI 코드 상의 의료정보 구분 및 분류

EDI 구분	분류유형	내용
의·치과	1	수가(공상 포함)
	2	준용수가
	3	보험등재 의약품
	4	원료, 조제(제제) 의약품
	7	협약재료
	8	치료재료
한방	A	수가
	B	준용수가
	C	약가
	G	협약재료
	H	치료재료
	P	공상

EDI 는 국내의 단일보험자체제의 특성상 전국민에 대한 건강보험 청구 및 상환을 위한 목적으로 개발되었으나, 결과적으로 한국 의료용어의 실질적인 표준으로 자리잡게 되었다. 또한, OMOP 용어집에 EDI 코드가 등재되어있기 때문에, EDI-시술 용어와의 매핑 관계를 공식적으로 등재하여 관리할 수 있다. 사실상 표준으로 사용되는 EDI 코드와 OMOP 용어 간의 매핑 관계를 정의하게 된다면, 국내 대부분의 의료기관으로의 적용이 가능하기 때문에 재사용성의 효율이 높다. EDI 코드(총 8 자리)는 **기본코드(5 자리)**와 **산정코드(3 자리)**로 구성 되어 있는데, 이 중 기본 코드와 OMOP 표준 용어로의 매핑을 우선시 한다.

예시) N0723A00 절제관절성형술[지관절], 만 1 세미만

N0723A01 절제관절성형술[지관절], 만 1 세미만 제 2 의 수술
 N0723A10 절제관절성형술[지관절], 만 1 세미만 야간

ID ▼	CODE ▼	NAME ▼	CLASS ▼	CONCEPT ▼	VALIDITY ▼	DOMAIN ▼	VOCAB ▼
42263933	HE538806	Abdominal MRA,Under 72 months radiologist reading	Procedure	Non-standard	Valid	Procedure	EDI
42365283	R4534	Abdominal Pregnancy	Proc Hierarchy	Non-standard	Valid	Procedure	EDI
42177139	R4534080	Abdominal Pregnancy,clinic, dental clinic(including Public Health and Medical care center) Nighttime, Saturday, holiday	Procedure	Non-standard	Valid	Procedure	EDI
42178947	R4534081	Abdominal Pregnancy,clinic, dental clinic(including Public Health and Medical care center) Nighttime, Saturday, holiday second surgery	Procedure	Non-standard	Valid	Procedure	EDI
42319213	R4534020	Abdominal Pregnancy,emergency	Procedure	Non-standard	Valid	Procedure	EDI
42303006	R4534040	Abdominal Pregnancy,Emergency holiday	Procedure	Non-standard	Valid	Procedure	EDI

그림 9. ATHENA 에서 확인한 EDI 코드의 OMOP vocabulary 등재 모습

Procedure 도메인에서는 SNOMED, CPT4, ICD10PCS 세 가지 모두 표준으로 사용하고 있지만, 미국에서 청구 관리 목적으로 사용하고 있는 CPT4 와 ICD10PCS 에 비해 SNOMED 가 좀 더 체계적으로 임상 용어를 분류하고, 구조 자체가 기술로직을 기반으로 되어있기 때문에 더 안정적이다. 따라서, 시술에 대한 EDI 코드를 OMOP 표준 용어로의 매핑 시 1 순위로 SNOMED, 2 순위로 CPT4, 3 순위로 ICD10PCS 용어집의 용어로 매핑하고 있다. 또한, 각 나라별로의 의료 보험 체계 혹은 정책이 상이하기 때문에 정확히 일치하는 시술이 없을 경우나 다른 정보에서 의료 행위를 식별할 수 있는 경우, 좀 더 상위 개념에 해당하는 용어로 매핑한다.

한국 OHDSI 의 공식 GitHub 인 OHDSI-Korea(<https://github.com/OHDSI-Korea>) 에서 위와 같은 가이드라인 하에 EDI-OMOP 표준 용어 간의 매핑 결과를 제시하고 있고, issue 란을 통하여 매핑 관계가 잘못되거나, 이 외 문제점들을 보고할 수 있다.

4-2. 3 자 검토 사례

2 인 1 조로 이루어진 매핑팀에서 실시한 상호독립적인 매핑 결과가 불일치하고 합의가 도출되지 않거나, OMOP 표준 용어로의 매핑되어야 할 적절한 코드를 찾지 못한 경우, 의학적 지식을 가졌으면서 OMOP 표준 용어에 대한 전문가와 3 자 검토를 실시하여 적절한 코드를 도출해낼 수 있다. 다음 일부 사례는, 매핑 연구원끼리 합의가 되지 않거나, 혼동되는 일부 사례이다.

Case 1. A1100(의료최고도 count 3-6, 의사 2·간호 5 등)

구분	내용
매핑 용어	연구원 A: 4154551(Activities of daily living management; SNOMED) VS

	연구원 B: 4126495(Activities of daily living assessment; SNOMED)
3 자 검토 결과 및 사유	수가로 청구를 받는 이유에 가급적 맞춰서 매핑 이 경우 assessment 를 진행 후 management 하는 과정에 대한 청구인데, 청구를 받는 이유는 해당 ADL level 에 대한 관리 때문이므로 4154551(Activities of daily living management; SNOMED)로 매핑

Case 2. AC301(소아 진정관리료-상급종합병원)

구분	내용
매핑 용어	연구원 A: NA VS 연구원 B: NA
3 자 검토 결과 및 사유	Pediatric sedation management 를 ATHENA 에서 찾을 수 없음 이럴 경우, 1. 상급 종합병원 등 종별에 대한 정보에 대한 매칭을 하지 않고 2. 이후에도 찾을 수 없다면 환자 나이에 해당되는 정보를 무시하기로 함 (나이에 대한 정보는 환자 나이를 통해서 추정 가능) 결과적으로 4219502(Sedation; SNOMED)으로 매핑

Case 3. AZ202(암환자 교육·상담료-수술 후_종합병원)

구분	내용
매핑 용어	연구원 A: NA VS 연구원 B: NA
3 자 검토 결과 및 사유	해당 수가는 수술 후 상담을 위한 것이며, 암 수술이란 정보는 환자의 진단명 또는 수술 기록을 통해 유추할 수 있기 때문에, cancer education 대신 4055248(Postoperative procedure education)에 매핑

Case 4. AC611(뇌졸중 집중치료실 입원료-상급종합병원)

구분	내용
----	----

매핑 용어	<p>연구원 A: 43021817(Stroke care management; SNOMED)</p> <p>VS</p> <p>연구원 B: 44807037(Acute stroke care management; SNOMED)</p>
3 자 검토 결과 및 사유	<p>Sub-ICU Patient Care-Stroke unit 임을 고려하고, acute 에 대한 정보는 cohort definition 시에 정의 할 수 있을 것으로 생각됨. 따라서, 43021817(Stroke care management; SNOMED)로 매핑</p>

Case 5. AC612(고위험임산부 집중치료실 입원료-상급종합병원)

구분	내용
매핑 용어	<p>연구원 A: 4237498(Perinatal care management; SNOMED)</p> <p>VS</p> <p>연구원 B: 44805394 (Gynaecology inpatient care management; SNOMED)</p>
3 자 검토 결과 및 사유	<p>Gynaecology inpatient care management 는 산부보다 부인과에 초점이 맞춰진 concept 으로 생각되기 때문에, 산부에 초점이 더 맞춰진 4237498(Perinatal care management; SNOMED)으로 매핑</p>

Case 6. AG113(신생아모유수유간호관리료-상급종합병원)

구분	내용
매핑 용어	<p>연구원 A: 4257051(Breast feeding support management; SNOMED)</p> <p>VS</p> <p>연구원 B: 4234170(Procedure related to breastfeeding)</p>
3 자 검토 결과 및 사유	<p>신생아모유수유간호관리료라는 개념이 모호하고, Procedure related to breastfeeding 가 Breast feeding support management 에 비해 보다 더 포괄적인 정보를 담고 있기 때문에 4234170(Procedure related to breastfeeding; SNOMED)로 매핑</p>

Case 7. AI200(심장통합진료료)

구분	내용
매핑 용어	<p>연구원 A: 4295949(Cardiac care management; SNOMED)</p> <p>VS</p> <p>연구원 B: 4042673(Procedure on heart)</p>
3 자 검토 결과 및 사유	<p>심장통합진료료는 관상동맥질환, 판막질환, 선천성 심장질환, 말기 심부전 등에 대한 청구이기 때문에 심장질환이 발생한 후 관리를 하는 개념인 Cardiac care management 보다 내/외과적인 시술에 대해 포괄적으로 사용되는 4042673(Procedure on heart)로 매핑</p>

Case 8. AJ100(상급종합병원-3 등급간호관리료적용 일반중환자실입원료)

구분	내용
매핑 용어	<p>연구원 A: 46285817(Provision of specialist intensive care registered nurse escort; SNOMED)</p> <p>VS</p> <p>연구원 B: 46285813(Provision of specialist intensive care medical escort)</p>
3 자 검토 결과 및 사유	<p>상급종합병원이라는 정보는 visit 에서 확인 할 수 있으므로, 상급종합병원에 대한 용어 매핑을 제외하고, 44804818(Critical care medicine care management; SNOMED) 로 매핑</p>

Case 9. AN200(가정간호기본방문료[방문당]-상급종합병원)

구분	내용
매핑 용어	<p>연구원 A: 4295046(Home health aide service management; SNOMED)</p> <p>VS</p> <p>연구원 B: 42537943(Respite care case management; SNOMED)</p>
3 자 검토	4295046(Home health aide service

결과 및 사유	management; SNOMED)
---------	---------------------

Case 10. AQ600(가정간호기본방문료[방문당]-상급종합병원)

구분	내용
매핑 용어	연구원 A: 4295046(Radiation therapy treatment management; SNOMED) VS 연구원 B: 42537943(Making of shielding block for radiotherapy; SNOMED)
3 자 검토 결과 및 사유	4259696(Isolation of subject emitting radiation; SNOMED)

4-3. 표준화된 시술 용어 적용 방안

표준화된 시술 용어는 원천 헬스케어 데이터베이스를 OMOP CDM 으로 변환하는 ETL(Extract Transform Load) 과정을 실시할 때 사용된다. 기본적으로 OMOP CDM 에 입력되는 모든 임상 자료들은 표준용어로 입력되어야 하는데, 원천 자료에 입력된 임상 기록들이 해당 기관에서만 고유하게 사용하는 로컬 코드이거나, 표준이 아닌 경우가 존재한다. 이 때, 로컬 코드-표준 용어 사이의 관계를 정의하기 위해 source to concept map 이라는 테이블이 사용된다. 이 테이블은 SOURCE_CODE 라는 필드에 원천 자료에서 임상적인 사건을 기록하기 위한 코드를 지정하고, OMOP 표준 용어로 변환되어야 할 코드를 TARGET_CONCEPT_ID 라고 하는 필드에 지정한다.

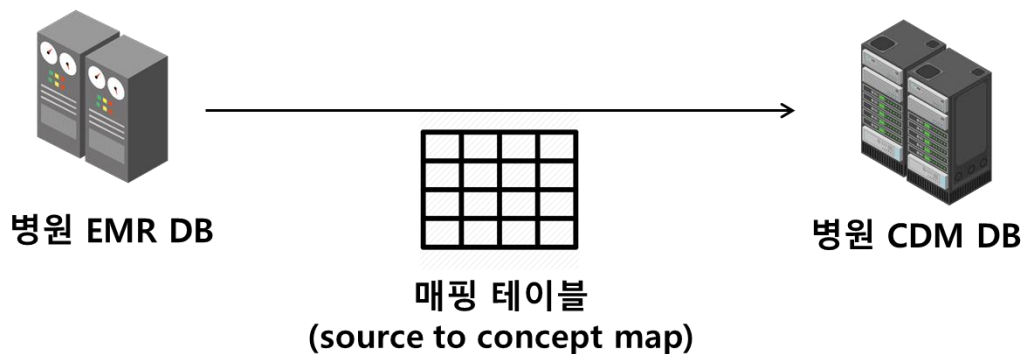


그림 10. 병원 EMR 데이터의 CDM 데이터로의 변환 시 매핑 테이블의 활용 도식도

예를 들어, A 라는 환자가 B 병원에 방문하여 "입체조형치료 (로컬코드: A; EDI 코드: HD061)", "정맥내일시주사 (로컬코드: B; EDI 코드: KK020)", "전립선내주사 (로컬코드: C; EDI 코드: HD072)" 라는 치료를 받았다고 가정하자. 이러한 경우, 우리는 직접적으로 의료기관 내에서 사용하는 코드와 표준 용어와의 매핑 관계를 지을 수 있다: "A-

762950(Three dimensional conformal radiotherapy)", "B-4059831(Intravenous injection)", "C-4059831(Injection of prostate)". 하지만, 앞서 언급한 것처럼 본 가이드라인에서는 보험 청구를 위해 사실상 표준으로 사용되고 있는 EDI 를 활용하여 매핑하는 것을 권장한다. 만약에 A 라는 환자가 C 병원에 방문하여 똑같은 치료를 받았지만, 각 각 X, Y, Z 라는 로컬 코드를 사용하는 경우 "X-762950(Three dimensional conformal radiotherapy)", "Y-4059831(Intravenous injection)", "Z-4059831(Injection of prostate)"라는 관계를 새롭게 정의해야 하기 때문이다. 따라서, EDI 코드와 OMOP 표준 용어를 직접 매핑하거나, 로컬코드-EDI 코드-OMOP 표준 용어를 거쳐 매핑하는 것을 권장한다.

표 7. SOURCE TO CONCEPT MAP 테이블 예시

SOURCE_CODE	SOURCE_CONCEPT_ID	SOURCE_CODE_DESCRIPTION	TARGET_CONCEPT_ID	TARGET_CONCEPT_NAME	INVALID_REASON
HD061	42359603	3-Dimensional Conformal Therapy; 입체조형치료 [1 회당]	762950	Three dimensional conformal radiotherapy	NA
KK020	42358203	Intravenous Injection; 정맥내일시 주사[1 일당]	4059831	Intravenous injection	NA
HD072	42357699	Intraprostatic Injection; 전립선 내주사	4059831	Injection of prostate	NA

물론, EDI 를 활용한 매핑은 모든 경우를 커버해주지 않는다. 이를테면, 청구를 위해 보고되지 않는 비급여 약물이나 치료 같은 경우는 별도의 EDI 가 존재하지 않기 때문에 반드시 로컬 코드와 OMOP 표준 용어를 매핑해야 한다. 하지만, 흔히 쓰이는 대부분의 코드들은 보험 급여 청구를 실시하기 때문에 거시적인 관점에서의 접근으로는 EDI 코드를 활용하는 것이 적합하다.

5. REFERENCES

- [1] LOINC database (Regenstrief Institute, Indiana University School of Medicine, Indianapolis, USA)
- [2] Study Plan ver.2.0. ICD-9-CM (International Classification Disease, 9th revision, Clinical Modification)
- [3] Guidelines for identifying non-personal information (2018.07, Korea).
- [4] 건강보험심사평가원 수가반영내역(19.07.01 기준).
- [5] 한국의료 QA 학회지 Volume 12, Number 1, 104-112.
- [6] 한국과학기술평가원(18.08)
- [7] LOINC user's guide. <http://www.regenstrief.org/loinc/>
- [8] NPU, LOINC, and SNOMED CT: a comparison of terminologies for laboratory results reveals individual advantages and a lack of possibilities to encode interpretive comments. *J Lab Med* 2018; 42(6):267-275. Andreas Bietenbeck et al.
- [9] Recent Developments in Clinical Terminologies - SNOMED CT, LOINC, and RxNorm. *Yearb Med Inform.* 2018 Aug;27(1):129-139. Bodenreider O, Cornet R, Vreeman DJ.
- [10] Issues in mapping LOINC laboratory tests to SNOMED CT. *AMIA Annu Symp Proc.* 2008 Nov 6:51-5. Bodenreider O.
- [11] Use of SNOMED CT® and LOINC® to standardize terminology for primary care asthma electronic health records. *J Asthma.* 2018 Jun;55(6):629-639. doi: 10.1080/02770903.2017.1362424. Epub 2017 Oct 9. Lougheed MD, Thomas NJ, Wasilewski NV, Morra AH, Minard JP.
- [12] Standardizing Germany's Electronic Disease Management Program for Bronchial Asthma. *Stud Health Technol Inform.* 2019 Sep 3;267:81-85. doi: 10.3233/SHTI190809. Sass J, Essenwanger A, Luijten S, Vom Felde Genannt Imbusch P, Thun S.
- [13] An update on the use of health information technology in newborn screening. *Semin Perinatol.* 2015 Apr;39(3):188-93. doi: 10.1053/j.semperi.2015.03.003. Epub 2015 Apr 29. Abhyankar S, Goodwin RM, Sontag M, Yusuf C, Ojodu J, McDonald CJ.
- [14] The LOINC RSNA radiology playbook - a unified terminology for radiology procedures. *J Am Med Inform Assoc.* 2018 Jul 1;25(7):885-893. doi: 10.1093/jamia/ocy053. Vreeman DJ, Abhyankar S, Wang KC, Carr C, Collins B, Rubin DL, Langlotz CP.
- [15] Single-Center Experience Implementing the LOINC-RSNA Radiology Playbook for Adult Abdomen/Pelvis CT and MR Procedures Using a Semi-Automated Method. *J*

- Digit Imaging. 2018 Feb;31(1):124-132. doi: 10.1007/s10278-017-0016-0. Sandhu RS1, Shin J, Wang KC, Shih G. Sandhu RS, Shin J, Wang KC, Shih G.
- [16] Observational Health Data Sciences and Informatics. The Book of OHDSI. <https://ohdsi.github.io/TheBookOfOhdsi/>
- [17] Observational Health Data Sciences and Informatics. Standardized Vocabulary V5.0. <https://www.ohdsi.org/web/wiki/doku.php?id=documentation:vocabulary>