



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

독립형 소프트웨어 의료기기에서  
공통 데이터 모델의 규제 활용

연세대학교 대학원

의료기기산업학과

이 나 라

# 독립형 소프트웨어 의료기기에서 공통 데이터 모델의 규제 활용

지도교수 구 성 욱 · 장 원 석

이 논문을 석사 학위논문으로 제출함

2024년 12월

연세대학교 대학원

의료기기산업학과

이 나 라

## 이 나라의 석사 학위논문으로 인준함

심사위원                구  성  욱          인

심사위원                장  원  석          인

심사위원                장  승  진          인

연세대학교 대학원

2024 년 12월

## 감사의 글

의료기기산업학과에 입학해 어느덧 2년이라는 시간이 지나 졸업을 앞두고 있습니다. 2년 동안 일과 학업을 병행하면서 포기하고 싶은 순간도 있었지만, 주변의 아낌없는 지지와 격려를 통해 무사히 석사 학위를 마칠 수 있음에 감사합니다.

먼저 대학원을 통해 배움의 기회를 주시고 아낌없는 조언과 올바른 방향을 지도해주신 구성욱 교수님, 장원석 교수님께 감사합니다. 그리고 논문 작성에 있어 꼼꼼하게 자문해주신 장승진 박사님께도 감사드립니다.

사랑하는 우리 가족, 배움의 용기를 잃지 않고 도전할 수 있도록 언제나 든든한 지원군이 되어 주신 부모님께 감사합니다. 하나뿐인 언니에게도 이 자리를 통해 고맙다는 인사를 전합니다.

그리고 학교는 다르지만, 함께 학위 생활을 하며 때로는 친구처럼 때로는 친언니처럼 포기하지 않도록 용기를 준 나리언니에게도 감사하고 서로의 도전을 응원합니다.

마지막으로 늘 옆에서 지지해주고 힘이 되어 준 사랑하는 건호에게도 감사의 인사를 전하고 싶습니다.

늘 겸손과 배움의 자세를 잃지 않으며 의료기기 산업에 도움이 되는 인재로 성장하도록 노력하겠습니다.

이나라 올림

## 차 례

그림 차례	iii
표 차례	iv
국문 요약	v
제1장 서론	1
1.1. 연구 배경 및 필요성	1
1.2. 연구 목적	3
제2장 연구 방법	4
2.1. 문헌고찰의 설계	5
2.1.1. 기초자료 수집	5
2.1.2. 연구 질문 설정	5
2.2. 문헌 검색 전략	6
2.2.1. 데이터베이스 및 검색어 선정	6
2.2.2. 문헌 선정 및 제외 기준	7
2.3. 데이터 수집 및 분석	7
제3장 이론적 배경	8
3.1. 독립형 소프트웨어 의료기기(SaMD)의 정의 및 분류	8
3.1.1. 국내 식품의약품안전처	8
3.1.2. 국제의료기기규제당국자포럼(IMDRF)	10
3.1.3. 미국 FDA	11
3.1.4. 유럽연합 MDR	13
3.2. SaMD의 기능적 특성 및 허가 현황	16
3.2.1. SaMD의 기능적 특성	16
3.2.2. SaMD 제품 인허가 현황	17

3.3. Common Data Model(CDM)의 개념 및 구조 . . . . .	18
3.3.1. OMOP-CDM . . . . .	21
3.3.2. Sentinel CDM . . . . .	24
3.4. Common Data Model(CDM) 국가별 동향 . . . . .	25
3.4.1. 한국 . . . . .	25
3.4.2. 미국 . . . . .	27
3.4.3. 유럽연합 . . . . .	27
제4장 연구 결과 . . . . .	28
4.1. 문헌 검색 및 선정 결과 . . . . .	28
4.2. 선정된 문헌의 일반적 특성 . . . . .	30
4.3. CDM의 규제적 활용 사례 . . . . .	33
4.3.1 미국의 NESTcc 프로젝트 . . . . .	33
4.3.2 실사용 증거(RWE)로의 활용 . . . . .	39
4.3.3 의료기기 개발 사례 . . . . .	39
4.3.4 의료기기 검증 사례 . . . . .	40
4.3.5 시판 후 감시(PMS) 사례 . . . . .	41
4.4. CDM 활용의 주요 패턴 및 경향 분석 . . . . .	42
4.5. CDM의 규제적 활용에 대한 잠재적 이점 및 한계 . . . . .	44
4.6. SaMD에서 CDM의 규제적 활용에 대한 국내 도입 제언 . . . . .	45
제5장 고찰 . . . . .	46
제6장 결론 . . . . .	47
참고 문헌 . . . . .	48
영문 요약 . . . . .	51

## 그림 차례

<그림 1> 스마트 헬스케어의 부상 배경 . . . . .	1
<그림 2> 독립형 소프트웨어 의료기기(SaMD) 글로벌 시장 규모 . . . . .	2
<그림 3> 인공지능·빅데이터 기반 의료기기의 특성 . . . . .	3
<그림 4> 의료기기 품목 분류 중 소프트웨어 . . . . .	9
<그림 5> 체외진단 의료기기 품목 분류 중 소프트웨어 . . . . .	9
<그림 6> Medical Device Software (MDSW) 판단 흐름도 . . . . .	15
<그림 7> CDM 데이터베이스의 구성 . . . . .	19
<그림 8> OMOP-CDM의 구조 (v5.4) . . . . .	23
<그림 9> Sentinel CDM의 구조 (SCDM v6.0) . . . . .	25
<그림 10> 심평원 청구자료 테이블과 CDM 테이블 매핑 . . . . .	26
<그림 11> 문헌 선정 흐름도 . . . . .	30
<그림 12> 선정 문헌의 년도별 출판 건수 . . . . .	31
<그림 13> NESTcc 능동 감시 프로그램 프로세스 흐름도 . . . . .	34
<그림 14> UDI 구현을 위한 IT 인프라 구축 프로세스 . . . . .	37
<그림 15> AI 기술의 임상적 검증에 필요한 7가지 핵심 단계 . . . . .	41



## 표 차례

<표 1> IMDRF SaMD Categories . . . . .	10
<표 2> IMDRF SaMD 카테고리 결정 기준 . . . . .	11
<표 3> MDR 의료기기 소프트웨어 등급분류 규칙 . . . . .	14
<표 4> 연도별(‘18~‘23)국내 소프트웨어 의료기기 인허가 건수 . . . . .	17
<표 5> 연도별(‘19~‘23) 인공지능(AI) 기반 의료기기 인허가 현황 . . . . .	18
<표 6> FDA의 SaMD 승인 현황 . . . . .	18
<표 7> 추출되는 정보의 예시 . . . . .	20
<표 8> OMOP-CDM의 설계 요소 . . . . .	22
<표 9> 문헌 검색 결과 . . . . .	28
<표 10> 문헌 1차 스크리닝 결과 제외된 문헌 건수 . . . . .	29
<표 11> 전문 검토 결과 제외된 문헌 건수 . . . . .	29
<표 12> 선정 문헌의 일반적 특성 . . . . .	31
<표 13> 선정 문헌의 요약 . . . . .	32
<표 14> 감시의 3단계에 대한 정의와 응용 사례 . . . . .	35
<표 15> NESTcc 클라우드 시스템에서 예상되는 데이터 요소 . . . . .	36
<표 16> 데이터 품질 프레임워크의 구성 . . . . .	38

## 국 문 요 약

### 독립형 소프트웨어 의료기기에서 공통 데이터 모델의 규제 활용

본 연구는 독립형 소프트웨어 의료기기(Software as a Medical Device, SaMD)에서 공통 데이터 모델(Common Data Model, CDM)의 규제적 활용에 대해 주제범위 문헌고찰 방법론을 적용하여 체계적으로 분석하는 것을 목적으로 한다. SaMD는 디지털 헬스케어의 발전과 함께 빠르게 성장하고 있으며, 소프트웨어 의료기기의 특성상 개발, 검증 및 시판 후 관리에 있어 데이터 품질 관리가 매우 중요하다. CDM은 데이터 표준화와 통합을 통해 SaMD의 개발 및 검증을 지원하며, 실사용 증거(Real-World Evidence, RWE) 생성, 시판 후 감시(Post-Market Surveillance, PMS) 측면에서 규제적 신뢰성을 높이는 데 기여한다.

연구 결과 미국의 NESTcc(National Evaluation System for health Technology Coordinating Center) 프로젝트의 능동적 감시활동과 한국의 시판 후 감시 시스템 구축 연구에서 UDI(Unique Device Identifier) 시스템 도입의 필요성에 대한 연구가 CDM의 규제 활용 주요 사례로 확인되었다. CDM의 활용은 국제 표준화, 데이터의 품질 보장 및 상호운용성을 개선하는 데 중요한 역할을 한다. 그러나 국가 간 규제 환경 차이와 데이터 통합 프로세스에서의 한계가 존재한다. 따라서, 이를 해결하기 위해 지속적인 연구와 다기관 협력이 필요하다.

본 연구는 SaMD의 규제적 효율성을 높이는 데 있어 CDM을 활용하는 것의 가치와 과제를 명확히 제시함으로써, 향후 SaMD의 규제 개선 및 데이터 기반 의사결정의 기초자료로 활용될 수 있다.

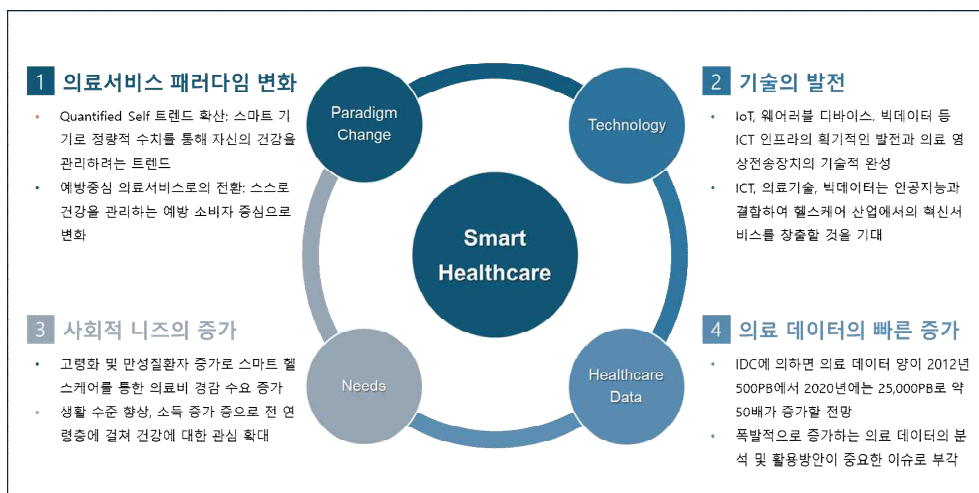
---

**핵심되는 말:** 소프트웨어 의료기기, 공통 데이터 모델, 실사용증거, 시판 후 감시

## 제1장 서론

### 1.1. 연구 배경 및 필요성

최근 의료 산업은 기존의 사후 치료 중심에서 벗어나, 의료 데이터를 이용해 사전에 예측하여 예방과 건강관리 중심의 패러다임으로 전환되고 있다 (그림 1). 특히, 소프트웨어 자체로 의료기기 정의에 해당하는 독립형 소프트웨어 의료기기(Software as a Medical Device, 이하 SaMD)의 개발이 가속화되고 있으며, “빅데이터 기반 인공지능(Artificial Intelligence, AI), 가상현실(Virtual Reality, VR) 증강현실(Augmented Reality, AR), 의료용 앱(Application, App)과 같은 형태의 소프트웨어”가 대표적인 사례이다.<sup>1</sup> 이러한 SaMD는 데이터의 수집, 분석, 그리고 이를 바탕으로 한 의료적 판단 및 처방에 중요한 역할을 한다.



\*출처: 스마트 헬스케어의 시대, 데이터 전쟁을 대비하라, 삼성KPMG 경제연구원, 2018, 재구성

그림 1. 스마트 헬스케어의 부상 배경

SaMD 시장은 의료 산업 내에서 혁신적인 부문을 대표한다. SaMD는 인공지능을 활용한 진단 도구부터 원격 환자 모니터링을 위한 웨어러블 기기에 이르기까지 광범위한 응용 분야를 포함한다. 이 시장은 개인화되고 원격 의료 서비스에 대한 수요 증가,

기술 발전, 전 세계적으로 만성 질환의 유병률 증가에 의해 주도된다. 의료의 디지털 혁신이 계속 가속화됨에 따라 SaMD 시장은 상당한 성장을 앞두고 있다. Allied Market Research에 따르면 글로벌 SaMD 시장 규모는 2023년 약 1,950만 달러로 평가되었으며, 2033년까지 9,620만 달러에 이를 것으로 연평균 성장률 17.3%로 성장할 것으로 예상하고 있다 (그림 2).<sup>3</sup> 시장 성장의 주요 요인으로는 인공지능(AI)과 머신러닝(ML)과 같은 기술 발전이 진단 능력과 치료 계획을 향상시키는 점으로 보고 있다.

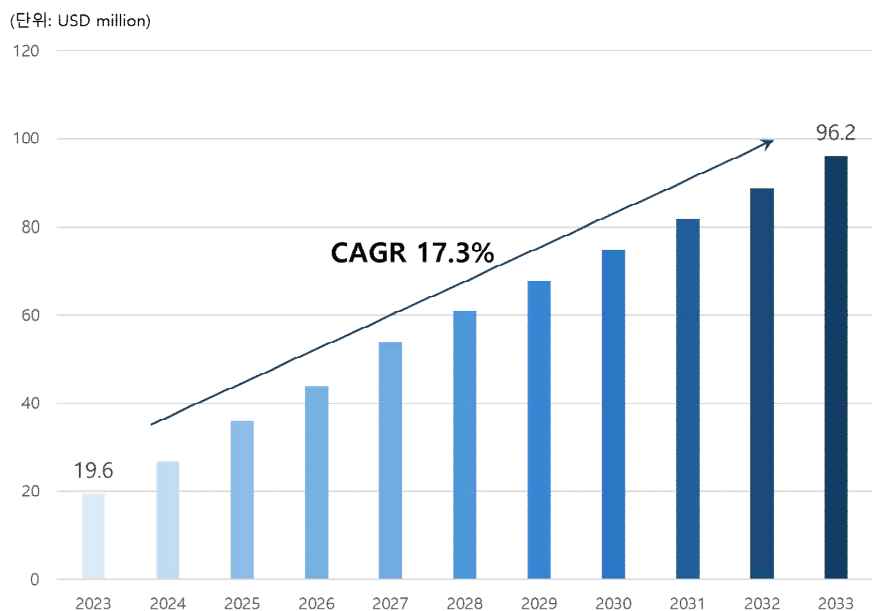
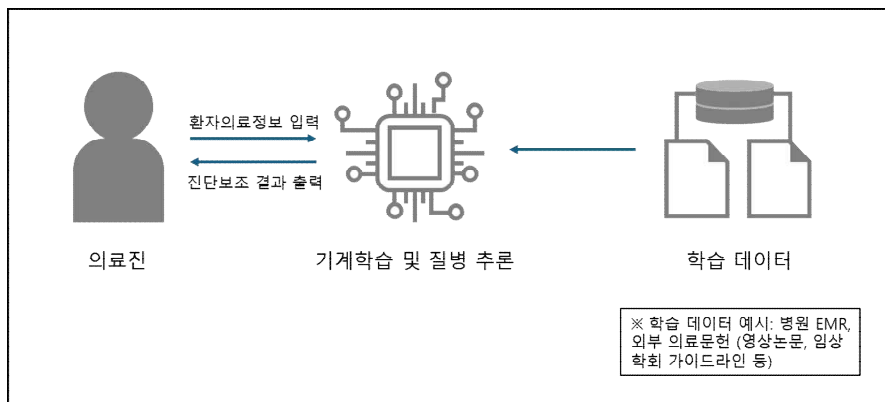


그림 2. 독립형 소프트웨어 의료기기(SaMD) 글로벌 시장 규모

SaMD의 전 주기(기획, 설계, 개발, 검증, 유지보수)에 걸쳐 효율적이고 정확한 데이터를 활용하는 것은 SaMD의 품질과 안전성을 확보하는 데 필수적이다. 특히, 공통 데이터 모델(Common Data Model, 이하 CDM)은 다양한 출처의 의료 데이터를 통합하고 표준화하여 SaMD 개발에 있어 중요한 역할을 할 수 있는 도구로 주목받고 있다. CDM은 데이터의 일관성, 재사용성, 호환성을 높여 SaMD의 전 주기에 걸쳐 데이터 활용을 극대화할 수 있는 잠재력을 지닌다.



\*출처: 신개발 의료기기 전망 분석 보고서, 식품의약품안전처, 2020, 재구성

그림 3. 인공지능·빅데이터 기반 의료기기의 특성

의료 데이터의 표준화와 상호 운용성 확보는 디지털 헬스케어 발전에 필수적이다. 특히 SaMD의 경우, 임상 데이터를 기반으로 한 알고리즘 및 소프트웨어의 신뢰성, 안전성을 보장하기 위해 데이터의 표준화 및 규제적 활용이 중요한 이슈다. CDM은 의료 데이터의 표준화된 구조를 제공하며, 이를 통해 다양한 데이터 소스 간의 통합 및 분석이 가능하다.

이에 따라 여러 규제기관에서 의료기기의 안전성과 유효성을 평가하기 위해 공통 데이터 모델을 도입하려는 시도가 이뤄지고 있다. 그러나 SaMD에서 CDM을 규제적 목적으로 활용하는 방법은 여전히 초기 단계에 있으며, 그 효과와 제한점에 대한 연구가 부족한 실정이다. 따라서, SaMD에 대한 CDM의 규제적 활용 현황과 성공 사례를 종합적으로 검토하고, 이를 통한 규제 효율성 증진 방안을 모색하는 것은 중요한 과제로 남아있다.

## 1.2. 연구 목적

본 연구의 목적은 독립형 소프트웨어 의료기기(SaMD)에서 공통 데이터 모델(CDM)이 규제적 목적으로 어떻게 활용되고 있는지를 체계적으로 분석하는 데 있다. 이를 위해 SaMD에서 CDM을 활용한 주요 사례를 파악하고, 각 사례가 SaMD의 규제적 접근에 기여한 방식과 그 시사점 및 한계점을 분석하고자 한다.

## 제2장 연구 방법

주제범위 문헌고찰(scoping review)은 특정 연구 분야의 주요 개념을 신속하게 파악하고, 이용 가능한 주요 자료 출처와 증거 유형을 탐색하는 것을 목적으로 한다. 특히 복잡하거나 포괄적으로 검토된 적이 없는 영역에 대해 독립적인 프로젝트로 수행될 수 있는 연구 방법이다.<sup>4</sup> 주제범위 문헌고찰을 통해 주제의 전반적인 이해를 제공하고, 향후 심층 연구나 체계적 문헌 검토의 기초를 마련하는 데 기여할 수 있다.

Arksey와 O' Malley의 주제범위 문헌고찰 가이드라인에 따른 구체적인 연구 방법 프레임워크는 아래와 같다.<sup>4</sup>

### (1) 연구 질문 설정:

체계적 문헌고찰 연구 방법과 마찬가지로 해결해야 할 연구 질문을 식별한다. 연구 질문은 명확하게 정의해야 하며, 광범위한 검토를 위해 범위는 넓어야 한다. 연구 질문을 설정하는 방법으로는 PICO (Population, Intervention, Comparison, Outcome), PCC (population, concept, context) 등이 있다.

### (2) 문헌 검색 전략 수립:

검색 전략은 체계적이어야 하며, 다양한 출처(예시: 전자 데이터베이스, 인터넷 및 연구 등록부, 참고 문헌 목록, 주요 저널의 수동 검색, 기존 네트워크, 관련 기관 및 컨퍼런스)를 통해 연구 증거를 검색하는 전략을 채택해야 한다.

### (3) 문헌 선정 기준:

체계적인 검토와 의사결정의 일관성을 보장하기 위해 프로젝트 시작 시 특정 연구 질문을 기반으로 포함 및 제외 기준을 개발해야 한다.

### (4) 데이터 추출 및 분석:

문헌 선정 다음 단계는 검토 중인 문헌에서 얻은 정보를 주요 이슈와 주제에 따라 자료를 선별, 차트화 및 분류하여 질적 데이터를 종합하고 해석하는 과정이다. 이 과정은 데이터베이스 프로그램인 Excel을 사용할 수 있다. 추출될 수 있는 정보의 예시는 아래와 같다.

- 저자(들), 출판 년도, 연구 지역
- 개입 유형 및 비교자(있는 경우), 개입 기간
- 연구 대상자(보호자 그룹, 보호 대상자 그룹)
- 연구 목적
- 방법론
- 중요한 결과

#### (5) 결과 요약 및 해석:

이 단계에서는 결과를 정리, 요약 및 보고하는 것이 포함된다. 체계적 문헌고찰과 달리 주제범위 문헌고찰은 증거를 종합하거나 다양한 연구의 결과를 집계하지 않는다. 주제범위 문헌고찰은 검토된 모든 자료의 개요를 잘 제시하는 것이 중요하다.

본 연구는 주제범위 문헌고찰 연구방법론을 기반으로 독립형 소프트웨어 의료기기(SaMD)에서 공통 데이터 모델(CDM)의 규제적 활용을 광범위하게 검토하고자 하였다. 이를 통해 SaMD의 규제적 관점에서 CDM이 수행하는 역할을 체계적으로 고찰하여, CDM의 규제적 활용에 대한 포괄적인 기초자료로 기여하고자 하였다.

## 2.1. 문헌고찰의 설계

### 2.1.1. 기초자료 수집

연구질문 설정에 앞서 본 연구의 핵심 키워드인 독립형 소프트웨어 의료기기와 공통 데이터 모델 각각의 개념 및 현황에 대한 이론적 배경에 해당하는 기초자료 수집을 진행하였다.

### 2.1.2. 연구 질문 설정

본 연구는 ‘공통 데이터 모델(CDM)이 독립형 소프트웨어 의료기기(SaMD)의 규제 목적으로 어떻게 활용되고 있으며, 이로 인해 어떤 결과가 도출되는지’에 대한 연구 질문을 분석하기 위한 주제범위 문헌고찰 연구이다.

## 2.2. 문헌 검색 전략

본 연구는 주제범위 문헌고찰 가이드라인(Arksey & O' Malley, 2005)과 PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) 체크리스트에 따라 수행되었으며 PICO (Population, Intervention, Comparison, Outcome)를 다음과 같이 설정하였다.<sup>4,5</sup>

- 연구집단(Population):  
독립형 소프트웨어 의료기기(Software as a Medical Device, SaMD)
- 중재(Intervention):  
공통 데이터 모델(Common Data Model, CDM)의 규제적 활용
- 비교중재(Comparison):  
공통 데이터 모델을 사용하지 않은 규제적 접근
- 중재결과(Outcome):  
공통 데이터 모델을 활용한 SaMD의 규제적 장단점 및 시사점

### 2.2.1. 데이터베이스 및 검색어 선정

문헌 검색은 최근 10년 (2015년 1월 1일부터 2024년 9월 30일까지) 동안 출판된 연구물을 대상으로 하여 SaMD의 규제 적용이 활발히 진행된 최신 정보를 수집하였다. 문헌 선정의 편향을 줄이기 위해 학위논문, 학술논문 및 회색 문헌 등을 모두 포함하였다. 국외 문헌은 PubMed와 Google scholar를 통해 검색하였고, 국내 문헌의 경우 국가학술정보 검색시스템(<https://cloud.nanet.go.kr/>), 한국의학논문데이터베이스(KMBASE) 및 학술연구정보서비스(RISS)로 총 5개의 데이터베이스를 통해 검색하였다.

검색은 MeSH 용어와 일반 키워드로 AND/OR 및 절단 검색을 적용하여 진행하였다. 약어가 있는 경우에는 풀어쓴 용어와 약어를 함께 사용하여 검색 범위를 넓혔다. 검색에 사용된 주요 핵심어는 (“Common Data Model” OR “CDM”) AND (“Software as a Medical Device” OR “SaMD” OR “Standalone software medical device” OR “Medical



device”) AND (“Regulatory” OR “Government Regulation” OR “post-market surveillance” OR “product surveillance” OR “safety assessment” OR “clinical evaluation”) AND (“efficacy” OR “effectiveness” OR “compliance” OR “guideline adherence”)로 조합하여 검색하였다. 한국어로는 (“공통 데이터 모델”) AND (“독립형 소프트웨어 의료기기” OR “소프트웨어 의료기기” OR “의료기기 소프트웨어”) AND (“규제” OR “인허가”)로 조합하여 검색하였다.

### 2.2.2 문헌 선정 및 제외 기준

문헌 선정 및 제외의 구체적인 기준은 아래와 같다.

#### (1) 선정 기준

- 소프트웨어 의료기기 대상 연구
- 공통 데이터 모델의 규제적 활용에 대한 연구
- 2015년 1월 1일부터 2024년 9월 30일까지 출판된 연구
- 한국어 또는 영어로 표기된 연구

#### (2) 배제 기준

- 의료기기 대상이 아닌 연구
- 기존의 연구를 활용한 문헌고찰 및 메타분석 연구
- 원문을 확보하기 어려운 연구

## 2.3 데이터 추출 및 분석

선정된 문헌의 원문을 검토하여 저자, 출판연도, 연구국가, Publication type에 대한 일반적 특성 데이터를 수집하였다. 이후 심층적인 분석으로 연구 목적, CDM의 활용 형태, CDM 활용의 시사점 및 한계점에 대한 자료를 중심으로 데이터를 수집하였다. 수집된 데이터는 표와 차트로 정리하였으며, 각종 표와 차트는 Microsoft 365 Excel과 Power point를 사용하여 작성하였다.

## 제3장 이론적 배경

### 3.1. 독립형 소프트웨어 의료기기(SaMD)의 정의 및 분류

#### 3.1.1. 국내 식품의약품안전처

식품의약품안전처의 「의료기기 소프트웨어 허가·심사 가이드라인」에서 소프트웨어 의료기기의 정의를 확인할 수 있다. 가이드라인에 따르면 의료기기 소프트웨어란 「의료기기법」 제2조에 해당하는 목적으로 사용하기 위해 개발·제조된 소프트웨어로 내장형 소프트웨어, 독립형 소프트웨어, 모바일 의료용 앱 등을 말한다. 의료기기에서 생성된 데이터를 서버 등에 전송하여 저장·분석 및 신호·영상처리 등을 수행하는 소프트웨어를 포함한다.<sup>6</sup>

그 중, 독립형 소프트웨어(소프트웨어 의료기기)란 전자·기계장치 등 하드웨어에 결합되지 않고 범용 컴퓨터 등과 동등 환경에서 운영되며 그 자체로 「의료기기법」 제2조에 해당하는 목적으로 사용되는 독립적인 형태의 소프트웨어로 정의하고 있다.

「의료기기법」 제2조 내용은 다음과 같다.

“의료기기”란 사람이나 동물에게 단독 또는 조합하여 사용되는 기구·기계·장치·재료·소프트웨어 또는 이와 유사한 제품으로서 다음 각 호의 어느 하나에 해당하는 제품을 말한다. 다만, 「약사법」에 따른 의약품과 의약외품 및 「장애인복지법」 제65조에 따른 장애인보조기구 중 의지(義肢)·보조기(補助器)는 제외한다. <개정 2018. 12. 11.>

1. 질병을 진단·치료·경감·치치 또는 예방할 목적으로 사용되는 제품
2. 상해(傷害) 또는 장애를 진단·치료·경감 또는 보정할 목적으로 사용되는 제품
3. 구조 또는 기능을 검사·대체 또는 변형할 목적으로 사용되는 제품
4. 임신을 조절할 목적으로 사용되는 제품

식품의약품안전처는 2020년 의료기기 품목 분류에 (E) 소프트웨어 대분류를 추가하여 소프트웨어 의료기기를 관리하고 있으며 (D) 체외진단 의료기기는 삭제되고 「체외진단의료기기법」으로 분류되었다. (E) 소프트웨어 대분류는 12개의 중분류로 분류되며, 각각 심혈관 진료용 소프트웨어, 치의학 진료용 소프트웨어, 이비인후과학 진료용 소프트웨어, 위장병학 및 비뇨의학 진료용 소프트웨어, 병원진료용 소프트웨어, 신경과학 진료용 소프트웨어, 산부인과학 진료용 소프트웨어, 안과학 진료용 소프트웨어, 정형외과학 진료용 소프트웨어, 재활의학 진료용 소프트웨어, 방사선종양학 및 영상의학 진료용 소프트웨어, 피부과학 진료용 소프트웨어로 분류된다(그림 4, 그림 5). 이에 따른 세부 품목은 총 109개로 분류된다.

대분류	중분류
(A) 기구 기계	01000 심혈관 진료용 소프트웨어
(B) 의료용품	02000 치의학 진료용 소프트웨어
(C) 치과 재료	03000 이비인후과학 진료용 소프트웨어
(D) <삭제>	04000 위장병학 및 비뇨의학 진료용 소프트웨어
(E) 소프트웨어	05000 병원진료용 소프트웨어
	06000 신경과학 진료용 소프트웨어
	07000 산부인과학 진료용 소프트웨어
	08000 안과학 진료용 소프트웨어
	09000 정형외과학 진료용 소프트웨어
	10000 재활의학 진료용 소프트웨어
	11000 방사선종양학 및 영상의학 진료용 소프트웨어
	12000 피부과학 진료용 소프트웨어

그림 4. 의료기기 품목 분류 중 소프트웨어

대분류	중분류
(I) 검체 전처리 기기	01000 체외진단 소프트웨어
(J) 임상화학 검사기기	02000 소인관련 체외진단 소프트웨어
(K) 면역 검사기기	03000 예후, 예측 관련 체외진단 소프트웨어
(L) 수혈의학 검사기기	04000 병리검사용 영상 처리 소프트웨어
(M) 임상미생물 검사기기	
(N) 분자진단기기	
(O) 조직병리 검사기기	
(P) 체외진단 소프트웨어	

그림 5. 체외진단 의료기기 품목 분류 중 소프트웨어

### 3.1.2. 국제 의료기기 규제당국자 포럼(IMDRF)

국제 의료기기 규제당국자 포럼(International Medical Device Regulators Forum, 이하 IMDRF)에서는 “의료기기로서의 소프트웨어 (Software as a Medical Device)”를 하드웨어 의료기기의 일부가 되지 않고, 하나 이상의 의료 목적을 수행하기 위해 사용될 의도로 개발된 소프트웨어로 정의하고 있다.<sup>7</sup> 하드웨어 의료기기의 일부가 아닌 소프트웨어란 하드웨어 의료기기가 의도한 의료 목적을 달성하는 데 필수적이지 않은 소프트웨어를 의미한다. 소프트웨어가 하드웨어 의료기기를 구동하는 것이 목적이라면, SaMD의 정의에 해당하지 않는다.

SaMD는 일반적인 컴퓨팅 플랫폼(비의료 목적)에서도 실행될 수 있으며, 하드웨어 의료기기 및 다른 SaMD 소프트웨어뿐만 아니라 일반 소프트웨어와도 인터페이스 될 수 있다. 위의 정의를 충족하는 모바일 앱도 SaMD로 간주된다.

IMDRF에서는 SaMD를 환자 또는 공중 보건에 미치는 영향 수준을 기반으로 네 가지 카테고리(I ~ IV)로 분류한다(표 1, 표 2). 카테고리 IV는 가장 높은 수준의 영향을 미치고, 카테고리 I은 가장 낮은 영향을 미친다. SaMD가 여러 의료상황 또는 상태에서 사용될 수 있다고 명시되어 있는 경우, SaMD 정의에 포함된 정보에 따라 가장 높은 카테고리로 분류된다.<sup>8</sup>

표 1. IMDRF SaMD Categories<sup>8</sup>

State of healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV	III	II
Serious	III	II	I
Non-serious	II	I	I

\*출처: Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations, IMDRF, 2014

표 2. IMDRF SaMD 카테고리 결정 기준<sup>8</sup>

Category	Criteria
I	<ul style="list-style-type: none"> <li>• 심각하지 않은 상황이나 상태에서 질병이나 질환의 임상 관리를 추진하기 위한 정보를 제공하며 <u>영향이 낮은 것</u>으로 간주됨</li> <li>• 심각한 상황이나 상태에서 질병이나 질환의 임상 관리를 알리기 위한 정보를 제공하며 <u>영향이 낮은 것</u>으로 간주됨</li> <li>• 심각하지 않은 상황이나 상태에서 질병이나 질환의 임상 관리를 알리기 위한 정보를 제공하며 <u>영향이 낮은 것</u>으로 간주됨</li> </ul>
II	<ul style="list-style-type: none"> <li>• 심각하지 않은 상황이나 상태에서 질병이나 상태를 치료하거나 진단하기 위한 정보를 제공하며 <u>중간 정도의 영향</u>을 미치는 것으로 간주됨</li> <li>• 심각한 상황이나 상태에서 질병이나 상태의 임상 관리를 추진하기 위한 정보를 제공하며 <u>중간 정도의 영향</u>을 미치는 것으로 간주됨</li> <li>• 위급한 상황이나 상태에서 질병이나 상태의 임상 관리를 알리기 위한 정보를 제공하며 <u>중간 정도의 영향</u>을 미치는 것으로 간주됨</li> </ul>
III	<ul style="list-style-type: none"> <li>• 심각한 상황이나 상태에서 질병이나 상태를 치료하거나 진단하기 위한 정보를 제공하며 <u>영향이 큰 것</u>으로 간주됨</li> <li>• 위급한 상황이나 상태에서 질병이나 상태의 임상 관리를 추진하기 위한 정보를 제공하며 <u>영향이 큰 것</u>으로 간주됨</li> </ul>
IV	<ul style="list-style-type: none"> <li>• 중대한 상황이나 상태에서 질병이나 증상을 치료하거나 진단하기 위한 정보를 제공하며 <u>영향이 매우 큰 것</u>으로 간주됨</li> </ul>

\*출처: Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations, IMDRF, 2014

### 3.1.3. 미국 FDA (U.S. FOOD & DRUG ADMINISTRATION)

FDA는 그 자체로 의료기기인 소프트웨어, 즉 소프트웨어 의료기기(SaMD)는 의료기기와 관련된 세 가지 유형의 소프트웨어 중 하나로 IMDRF의 정의를 기반으로 SaMD를 관리하고 있다.

소프트웨어 의료기기(SaMD)는 스마트폰에서 진단 목적으로 자기공명영상(MRI) 의료기기에서 얻은 이미지를 볼 수 있게 하는 소프트웨어부터 유방암 탐지를 돕기 위해 이미지 후처리를 수행하는 컴퓨터 보조 진단(CAD) 소프트웨어까지 다양하다.

예시는 다음과 같다:<sup>10</sup>

- SaMD는 하드웨어 의료기기, 다른 소프트웨어 의료기기, 일반 소프트웨어와 인터페이스 될 수 있다. 예를 들어, 선형 가속기에 사용되는 정보를 제공하는 치료 계획 소프트웨어는 SaMD에 해당된다.
- 의료 목적을 가진 소프트웨어가 의료용 목적이 아닌 일반 컴퓨팅 플랫폼에서 작동하는 경우에도 SaMD로 간주된다. 예를 들어, 소비자용 디지털 카메라의 내장 프로세서에서 작동하는 3축 가속도를 이용해 질병을 진단하는 소프트웨어는 SaMD로 인정된다.
- 하드웨어 의료기기과 연결되어 있지만, 그 하드웨어 의료기기가 의도된 의료 목적을 달성하기 위해 필요하지 않은 소프트웨어는 SaMD이며, 하드웨어 의료기기의 부속품이 아니다.
- SaMD는 일반적인 컴퓨팅 플랫폼(비의료 목적)에서 실행될 수 있다. 이러한 플랫폼에서 실행되는 SaMD는 하드웨어 의료기기에 있을 수도 있다.

SaMD가 아닌 소프트웨어의 예시:<sup>10</sup>

- 하드웨어 의료기기의 일부인 소프트웨어  
(예: 주입 펌프에서 모터를 구동하거나 약물 펌핑을 제어하는 소프트웨어, 혹은 이식형 심박 조율기와 같은 하드웨어 의료기기에서 폐쇄 루프 제어에 사용되는 소프트웨어는) 이러한 소프트웨어는 "임베디드 소프트웨어", "펌웨어", 또는 "마이크로코드"라고도 불린다.
- 하드웨어 의료기기의 의도된 의료기기 용도를 수행하기 위해 필요한 소프트웨어. 이는 하드웨어와 별도로 판매되더라도 마찬가지다.
- 의료기기에서 데이터를 가져오지만 의료 목적이 없는 소프트웨어  
(예: 데이터를 암호화하여 전송하는 소프트웨어)
- 임상 의사소통 및 워크플로우를 지원하는 소프트웨어  
(예: 환자 등록, 방문 예약, 음성 통화, 화상 통화 등의 기능을 가진 소프트웨어)
- 장비의 성능 또는 적절한 기능을 모니터링하여 장비의 서비스 목적으로 사용되는 소프트웨어(예: X선 튜브 성능을 모니터링하여 교체 필요성을 예측하는 소프트웨어)

어), 또는 실험실 품질 관리 데이터를 분석해 오류나 보정 추세를 파악하는 소프트웨어

- SaMD에 입력되는 매개변수를 제공하지만, 의료 목적이 없는 소프트웨어  
(예: 검색 및 쿼리 기능이 있는 데이터베이스)

### 3.1.4. 유럽연합 MDR (Medical Device Regulation)

Medical Device Regulation (이하 MDR) 제2조 1항에서 의료기기를 정의하고 있으며, 이 정의에 독립형 소프트웨어 의료기기가 포함된다. MDR의 Annex VIII는 의료기기의 등급분류 기준을 설명하고 있다.

세계법제정보센터에서 제공하는 Medical Device Regulation 2017/745 국문 번역본에 따르면 내용은 다음과 같다.

- (1) ‘의료기기’란 다음의 특정 의료목적 중 하나 이상에 대해 단독 또는 조합하여 인간에게 사용되도록 의도된 모든 기구, 장치, 기기, 소프트웨어, 임플란트, 시약, 재료 또는 기타 제품을 의미한다.
  - 질병을 진단, 예방, 모니터링, 예측, 예후, 치료 또는 완화
  - 부상 또는 장애의 진단, 모니터링, 치료, 완화 또는 보완
  - 해부학 또는 생리학적 혹은 병리학적 프로세스나 상태의 연구, 대체 또는 변경
  - 장기, 혈액 및 조직기증을 포함하여, 인체에서 채취한 검체의 체외검사를 통해 정보제공
- (4) “능동형 의료기기”란 해당 목적을 위해 인체 또는 중력에 의하여 생성된 에너지를 제외한 에너지원에 의존하고 밀도를 변경하거나 해당 에너지를 변환하여 작동하는 모든 기기를 의미한다. 에너지, 물질 또는 매개변수의 중대한 변화 없이 능동형 의료기기와 환자 사이에 에너지, 물질 또는 그 밖의 요소를 전달하기 위해 사용되는 기기는 능동형 의료기기로 보지 아니한다.  
소프트웨어도 능동형 의료기기기로 간주되어야 한다.

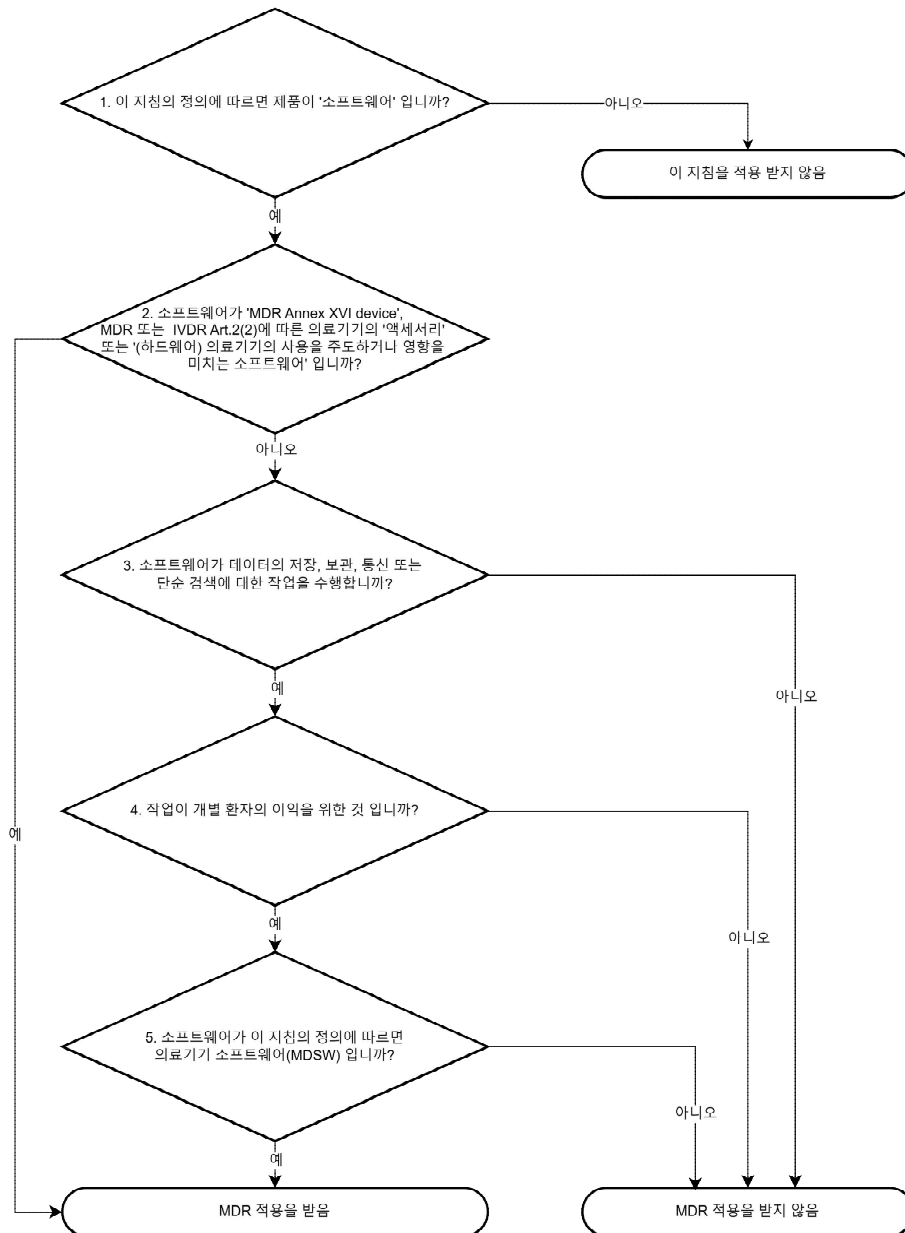
MDR Annex VIII 분류 기준에 따른 적용은 2019년 10월 개정된 \*MDCG 2019-11 지침에서 자세하게 설명하고 있다. “의료기기 소프트웨어(Medical Device Software, MDSW)”로 지칭하고 있으며, MDR 또는 IVDR (In Vitro Diagnostic Regulation) 에서 정의한 의료기기의 목적을 위해 단독 또는 조합하여 사용하도록 의도된 소프트웨어로 정의한다. 소프트웨어 등급분류 규칙은 (표 3)과 같으며 (그림 6)의 판단 흐름도를 활용한다.<sup>11</sup>

표 3. MDR 의료기기 소프트웨어 등급분류 규칙<sup>11</sup>

<b>Rule 11</b>	진단 또는 치료 목적의 결정을 위한 소프트웨어 또는 생리적 과정을 모니터링하기 위한 소프트웨어
	<ul style="list-style-type: none"> <li>- 진단 또는 치료 목적으로 의사결정을 내리는 데 사용되는 정보를 제공하도록 의도된 소프트웨어는 클래스 IIa로 분류된다.</li> <li>- 단, 사망 또는 사람의 건강 상태의 돌이킬 수 없는 악화인 경우 클래스 III;</li> <li>- 또는 사람의 건강 상태의 심각한 악화 또는 수술적 개입인 경우 클래스 IIb로 분류된다.</li> <li>- 생리적 과정을 모니터링 하도록 의도된 소프트웨어는 클래스 IIa로 분류된다.</li> <li>- 단, 중요한 생리적 매개변수를 모니터링 하도록 의도된 경우, 이러한 매개변수의 변화 특성이 환자에게 즉각적인 위험을 초래할 수 있는 경우 클래스 IIb로 분류</li> <li>- 기타 모든 소프트웨어는 클래스 I으로 분류된다.</li> </ul>
<b>Rule 12</b>	물질을 투여 및/또는 제거하기 위한 능동형 의료기기
	<ul style="list-style-type: none"> <li>- 소프트웨어 장치는 물리적으로 물질을 투여 및/또는 제거할 수 없으므로 본 규칙이 적용되는 MDSW에 대한 부록 VIII의 구현 규칙 3.3을 참조한다.</li> </ul>
<b>Rule 13</b>	다른 모든 능동형 의료기기
	<ul style="list-style-type: none"> <li>- 능동형 의료기기에 적용되는 모든 구현 및 분류 규칙을 고려하면 다른 규칙이 적용되지 않는 경우 다른 모든 능동형 의료기기는 클래스 I으로 분류된다.</li> </ul>
<b>Rule 15</b>	피임에 사용되는 의료기기
	<ul style="list-style-type: none"> <li>- 피임에 사용되는 소프트웨어는 클래스 IIb로 분류된다.</li> </ul>

\* MDCG: Medical Device Coordination Group





\*출처: Guidance on Qualification and Classification of Software in Regulation (EU)  
 2017/745 - MDR and Regulation (EU) 2017/746 - IVDR, MDCG, 재구성

그림 6. Medical Device Software (MDSW) 판단 흐름도

### 3.2. SaMD의 기능적 특성 및 허가 현황

소프트웨어 의료기기의 가장 큰 특성은 “물리적 형태가 없는 무형의 의료기기”라는 점이다.<sup>12</sup> 무형이라는 특성으로 인해 제품의 업데이트, 복제 및 배포가 용이하다는 장점이 있지만, 이로 인해 소프트웨어 의료기기 제조업체에서 제공한 업데이트를 제3자가 설치할 수 있다는 위험성도 있다. 소프트웨어의 사소한 변경 혹은 결함은 예상하지 못한 심각한 문제를 발생시킬 수 있다. 따라서 소프트웨어 개발 프로세스는 소프트웨어 변경으로 인해 발생할 수 있는 예상치 않은 결과를 사전에 발견하고 예방하기 위하여 충분히 계획하고 관리해야 하며 이를 문서화해야 한다.

소프트웨어 의료기기로 인해 발생하는 문제는 사용자의 의사결정 단계를 지연시키거나 잘못된 선택이나 의사결정을 내리는데 작용할 수 있다. 이로 인해 결과적으로 환자에게 악영향을 미칠 수도 있다. 따라서 소프트웨어 의료기기는 다양한 측면(사회 기술 환경, 시스템 환경, 정보보안)을 고려하여 안전성을 확보하는 것이 매우 중요하다.

#### 3.2.1. SaMD의 기능적 특성

##### (1) 측정, 분석 및 진단보조 기능

의료기기 소프트웨어는 일반적으로 의료진이 환자를 진단하는 데 도움을 주는 역할을 수행한다. 크게 측정, 분석, 진단보조 기능으로 구분할 수 있다. 먼저, 측정 기능은 특정 항목을 정확하게 측정하고 그 결과의 신뢰성을 확보하는 것이 중요하다. 분석 기능은 분석에 사용된 지표와 원리, 그리고 결과의 정확도 및 분석 속도가 핵심 요소로 고려된다. 마지막으로, 진단보조 기능은 임상적으로 신뢰할 수 있는 결과를 제공해야 하며, 그 유효성이 명확하게 입증되어야 한다.

##### (2) 데이터 변환 및 전송/수신 기능

의료기기 소프트웨어의 데이터 변환 및 전송, 수신 기능은 여러 의료기기 간의 연결 또는 유·무선 네트워크 환경에서 작동하는 의료기기에 필수적인 요소다. 데이터 변환 기능은 변환 과정 전·후, 그리고 복구 과정에서 데이터의 무결성을 보장하는 것이

핵심이다. 즉, 데이터가 손실 없이 유지되고 원래 상태로 복원될 수 있어야 한다. 전송 및 수신 기능에서는 데이터의 유형, 송수신 방식 등이 중요한 고려사항이며, 유·무선 LAN, 블루투스, 근거리 통신(NFC)과 같은 기술들이 주로 사용된다. 이러한 환경에서는 다른 의료기기 및 시스템과의 연결이 이루어지므로, 보안 및 개인정보 보호도 중요한 문제로 다루어져야 한다. 또한, 물리적으로 연결된 장치 간에 데이터가 변환, 전송, 수신되는 경우, 서로 다른 기종 간의 인터페이스 방식도 반드시 고려해야 한다.

### (3) 디스플레이(Display) 기능

의료기기 소프트웨어의 디스플레이 기능은 환자의 생체 신호, 영상 등과 같은 중요한 정보를 모니터와 같은 출력 장치를 통해 제공하는 역할을 한다. 이때 출력 장치에 어떤 정보가 표시되는지, 그리고 그 정보가 어떤 형태로 나타나는지가 매우 중요하다. 표시된 정보가 명확하고 직관적이어야 하며, 의료진이 필요한 결정을 빠르게 내릴 수 있도록 시각적 구성과 정보 전달 방식이 신중하게 설계되어야 한다.

#### 3.2.2. SaMD 제품 인허가 현황

식품의약품안전처 의료기기 허가보고서에 따르면 “2023년 주요 허가 동향은 ‘소프트웨어 의료기기’ 품목이 증가하는 추세이며, 새로운 치료 수단으로 ‘디지털 치료기기’가 등장한 점이다. ‘의료기기 소프트웨어’ 인허가 건수는 최근 6년간 지속적으로 증가하고 있으며, 2023년에는 156개의 제품이 식품의약품안전처의 인허가를 획득했다(표 4). 이 중 국내 제조 품목이 118건으로 75.6%를 차지하여 국내 개발 제품이 강세를 보임”을 알 수 있다.<sup>13</sup>

표 4. 연도별('18~'23) 국내 소프트웨어 의료기기 인허가 건수

	연도					
	2018	2019	2020	2021	2022	2023
건수	20	36	73	111	97	156

\*출처: 식품의약품안전처

특히, 기계학습 방식의 인공지능 기술을 기반으로 빅데이터 학습을 통한 질병을 진단 혹은 예측하거나 환자에게 맞춤 치료법을 제공하는 소프트웨어의 개발이 증가하고 있다. 2023년 인증 혹은 허가된 인공지능 기반 의료기기는 총 64건으로, 전년 대비 20건 증가하였다. 이 중 수입 9건, 제조 55건으로 국내에서 인공지능 기반 의료기기 개발이 활발하게 이루어지고 있음을 알 수 있다(표 5).<sup>13</sup>

표 5. 연도별(‘19~’23) 인공지능(AI) 기반 의료기기 인허가 현황

(단위: 건)

	연도				
	2019	2020	2021	2022	2023
제조	10	45	34	40	55
수입	-	5	3	8	9
전체	10	50	37	48	64

\*출처: 식품의약품안전처

미국 FDA에서도 소프트웨어 의료기기(SaMD) 제품에 대한 인허가가 지속적으로 증가하고 있다. 2023년까지 총 562건의 SaMD가 승인되었으며, 그중 2023년에만 133건이 승인되었다(표 6). 이는 전년 대비 47건 증가한 수치로, 소프트웨어 의료기기의 시장 확장과 기술 발전이 빠르게 이루어지고 있음을 보여준다.

표 6. FDA의 SaMD 승인 현황

	연도					
	~2018	2019	2020	2021	2022	2023
건수	168	30	62	83	86	133

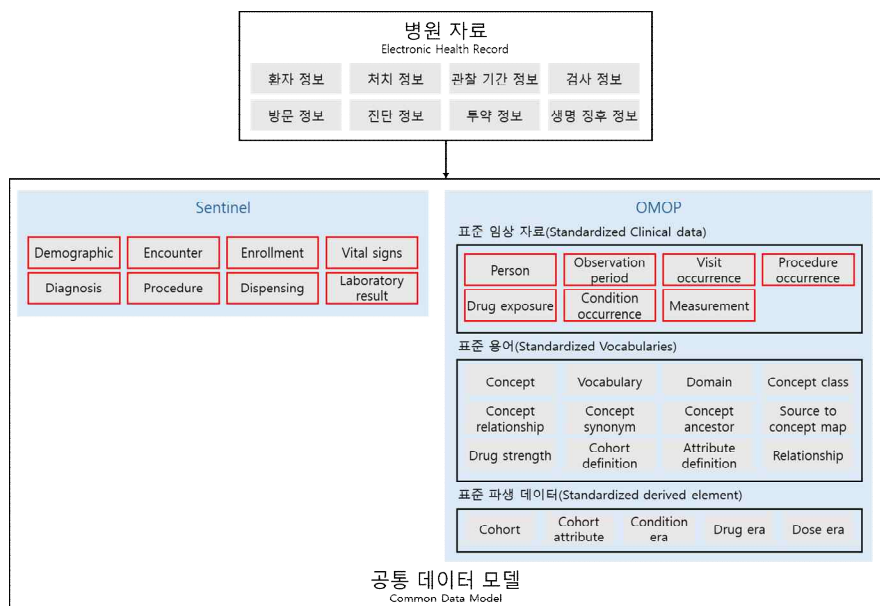
\*출처: 미국 식품의약청(FDA)

### 3.3. Common Data Model(CDM)의 개념 및 구조

공통 데이터 모델(Common Data Model, 이하 CDM)은 서로 다른 구조의 데이터베이스를 공통된 형식으로 변환하는 데이터 모델로, 광범위한 빅데이터를 표준화함으로써 용도에 따라 다양한 목적의 연구를 설계할 수 있도록 한다. 또한, 데이터베이스에 직

접 접근하지 않고도 분석이 가능하기 때문에, 데이터의 보안을 유지하면서도 효율적이고 체계적으로 데이터를 활용할 수 있다. 특히 민감한 개인의 의료정보를 다룰 때, CDM을 사용하면 정보 보안과 데이터 활용이라는 두 가지 목표를 동시에 달성할 수 있어 유용하다.

CDM은 연구 목적에 따라 데이터 구성항목이 다르며, 목적에 따라 OMOP-CDM, Sentinel CDM, PCORnet-CDM 등으로 구축된다. 각 모델은 고유한 도식(schema)을 바탕으로 테이블과 변수 구조를 가지며, 지속적으로 업데이트되어 데이터 분석의 효율성을 높인다. CDM 데이터베이스에는 다양한 테이블들이 포함되며, 이들 테이블은 자료의 구조와 관계를 나타낸다 (그림 7).



\*출처: 병원자료 분석네트워크 (<https://moa.drugsafe.or.kr/>), 재구성

그림 7. CDM 데이터베이스의 구성

병원의 전자의무기록(EMR)을 활용하여 CDM 데이터베이스를 구축하려면, 추출-변환-적재(ETL; Extraction-Transformation-Loading) 과정을 거쳐 데이터를 CDM 형식에 맞게 변환해야 한다. 이 과정에서는 데이터의 표준화와 함께 각 데이터 필드가 CDM의 구조에 맞게 매핑(Mapping)되는 작업도 필수적이다. 이를 통해 다양한 출처의 데이터

를 체계적이고 일관성 있게 분석할 수 있는 기반을 마련할 수 있다.

- **추출 (Extraction)**

병원 자료로부터 CDM 구축에 필요한 정보를 추출하는 단계는 다양한 의료 데이터를 표준화된 CDM 구조에 맞게 준비하는 과정이다. 이 단계에서는 병원의 전자 의무기록(EMR) 시스템 등에서 다양한 유형의 데이터를 수집하여 CDM을 구성하는데 필요한 요소들을 도출하게 된다. 추출되는 주요 정보들은 다음과 같다:(표 7)

표 7. 추출되는 정보의 예시

항목	예시
환자 정보	환자의 고유 식별번호, 성별, 나이, 출생일, 인종, 민족 등 환자의 인구통계학적 데이터
진단 정보	각 환자에게 내려진 진단의 세부 정보(예: 진단명, 진단 코드, 진단 날짜)
약물 정보	환자에게 처방된 약물에 대한 정보(예: 약물명, 복용량, 투여 경로, 처방 날짜, 복용 기간)
검사 정보	혈액검사, 영상검사 등 각종 검사의 시행 내역과 그 결과(예: 검사명, 검사 수치, 검사 일자)
시술 및 수술 정보	환자에게 시행된 의료 시술, 수술 관련 정보(예: 시술명, 시술 코드, 시술 날짜)
입원 및 퇴원 정보	입·퇴원 내역과 입원 기간, 퇴원 시 상태 등의 정보
의료 기록	의사의 소견, 진료 기록, 메모와 같은 임상 기록
보험 및 결제 정보	환자의 보험 청구 데이터, 의료비 지불 내역, 코드화된 청구 항목

- **변환 (Transformation)**

추출된 원시 데이터(raw data)를 CDM 형식에 맞게 변환하는 과정이다. 병원 자료에서 추출된 데이터는 다양한 형식과 구조를 가질 수 있으며, 이를 표준화된 CDM의 테이블 및 변수 구조에 맞춰 변환해야 한다.

- **적재 (Loading)**

변환된 데이터를 데이터베이스에 저장하여 실제로 사용할 수 있도록 하는 과정이

다. 이 단계에서는 변환 작업을 거친 데이터를 표준화된 CDM 형식으로 데이터베이스 관리 시스템(DBMS)에 업로드하여 CDM 데이터베이스를 구축하게 된다. 이때 사용되는 DBMS는 각 의료기관의 환경에 따라 PostgreSQL, Oracle, MySQL 등으로 선택할 수 있다.

- 매핑 (Mapping)

매핑은 의료 현장에서 사용되는 다양한 용어 체계를 통일하는 작업을 의미한다. 이는 약물, 진단, 검사 등의 정보를 일관되게 기록하고 관리하기 위해 필수적이다. CDM에서는 이러한 매핑 작업을 통해 표준 용어 체계를 적용함으로써, 데이터의 호환성을 높이고 국제적인 이슈에 대한 분석을 가능하게 한다.

CDM 데이터베이스는 각 협력 기관이 고유의 방화벽 내에서 자료를 안전하게 관리할 수 있는 분산형 자료망 형태로 설계되어 있다. 이 구조는 환자 정보를 철저히 보호하며, 보관과 취급 과정에서도 보안이 유지될 수 있도록 되어 있다. 병원 자료(전자 의무기록, Electronic Medical Record)를 CDM으로 변환하는 과정에서는 모든 개인정보가 비식별화된다. 이 비식별화 작업을 통해 환자의 개인 신상정보나 의무 기록이 외부로 유출될 위험을 최소화한다. 따라서 CDM 시스템을 통해 데이터 분석이나 연구가 이루어질 때도 환자의 개인정보가 안전하게 보호되며, 정보 유출에 대한 우려 없이 데이터 활용이 가능하다. 이러한 보안 조치는 의료기관과 연구자들이 데이터의 가치와 개인정보 보호를 동시에 고려할 수 있는 기반을 마련해준다.

### 3.3.1. OMOP-CDM

2008년 미국 정부의 지원으로 Observational Medical Outcomes Partnership (OMOP)이 결성되었으며 이로부터 Observational Health Data Sciences and Informatics (이하 OHDSI) 협의체가 파생되었다. OMOP-CDM(Observational Medical Outcomes Partnership Common Data Model)은 여러 의료 시스템 내에서 발생하는 데이터의 구조, 코드 체계 및 용어 등을 표준화하여 연구자들이 데이터를 공통의 기준으로 비교하고 분석할 수 있도록 설계되었다.<sup>14</sup> 이로 인해 여러 데이터 소스 간 협업 연구가 가능해졌으며, 대규모 인구 수준에서 의료 행위와 정책의 영향을 평가하거나 질병 및 치료 특성을 분석하는 데 기여하고 있다.

## (1) 기본 개념 및 목적

OMOP-CDM은 여러 의료 시스템 내에서 발생하는 관찰 데이터를 체계적으로 정리하여 표준화된 데이터 구조로 제공함으로써, 연구자들이 데이터를 일관되게 분석할 수 있도록 한다. 다양한 데이터 출처에서 데이터를 직접 이동시키지 않고 분석 프로세스를 공유하는 방식으로 분산 연구를 지원한다. 이는 환자 데이터를 보호하면서도 광범위한 연구 협력이 가능하게 한다.

## (2) 설계 원리 및 구조

OMOP-CDM은 사람 중심(person-centric) 관계형 데이터 모델로 설계되어 모든 임상 사건을 PERSON 테이블을 중심으로 연결한다. 이를 통해 환자별로 의료 기록에 대한 종적 관찰이 가능하다. 데이터는 특정 도메인(예: 질환, 약물 노출 등)으로 분류되어 저장되며, 도메인별로 테이블 필드가 구분된다. 모든 기록은 표준화된 개념으로 정의하고 있다. 또한, OMOP-CDM은 특정 기술에 종속되지 않고 다양한 데이터베이스 플랫폼에서 활용할 수 있다. 예를 들어 관계형 데이터베이스(Oracle, SQL server 등)와 분석 데이터 플랫폼(SAS 등)에서 모두 구현이 가능하다. OMOP-CDM은 구체적인 설계 요소는 다음과 같다(표 8).

표 8. OMOP-CDM의 설계 요소<sup>14</sup>

설계 요소	설명
목적에 대한 적합성	데이터 분석 연구에 최적화된 데이터 제공
데이터 보호	개인정보 제한
도메인 설계	사람 중심 관계형 데이터 모델
도메인의 이론적 근거	모든 데이터는 개체-속성-값 구조를 가진 Observation 테이블에 관찰 데이터로 저장
표준화된 용어	표준 건강관리 개념을 포함하는 표준 용어에 의존
기존 용어 재사용	국가 및 산업 표준, 기관이나 협회 용어를 사용하기도 함
원천 코드 유지 관리	정보 소실 방지를 위해 원천 코드 저장
기술 중립성	특정 기술에 치우치지 않음
확장성	다양한 데이터 수용
이전 버전과의 호환성	모든 변경 사항은 Github에 저장

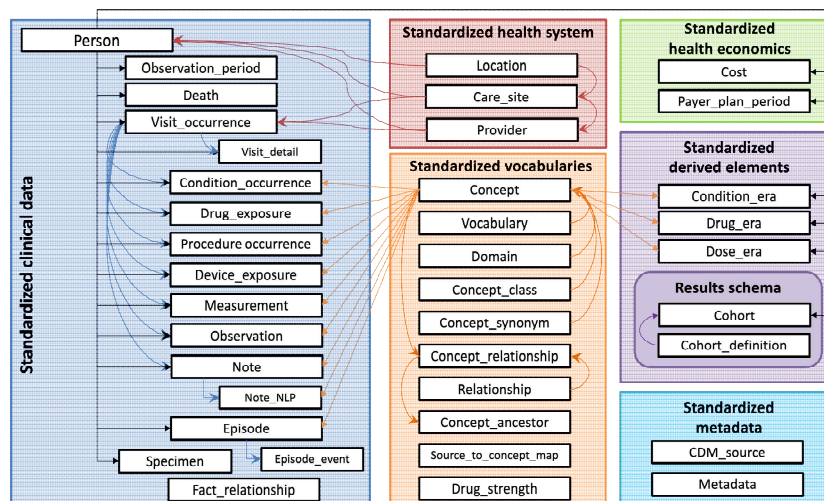


### (3) 주요 요소

OMOP-CDM은 표준화된 의료 용어를 사용하여 데이터를 정규화하고 일관성을 유지한다. 예를 들어, 모든 임상 사건은 특정 **Concept\_ID**를 통해 기록된다. 이를 통해서 데이터의 일관성 및 재현 가능성을 보장하며, 다양한 데이터베이스에서 분석 결과를 비교할 수 있다. 데이터를 변환하기 위해서는 원천 데이터를 추출-변환-적재하는 과정이 필요하다. 이 과정을 통해 다양한 형식의 데이터를 표준화된 CDM 구조로 정리할 수 있다.

OMOP-CDM은 총 39개의 테이블로 구성되어 있으며(그림 8), 주요 범주는 다음과 같다:

- Clinical Data: 환자의 임상 사건 정보
- Derived Elements: 임상 정보로부터 파생된 요소들
- Results schema: 결과 스키마
- Health System Data: 보건 시스템 데이터
- Health Economics: 보건 경제학 데이터
- Vocabulary: 표준 용어 체계
- Meta-data: 메타데이터 (CDM의 메타 정보를 포함)



\* 출처: <https://ohdsi.github.io/CommonDataModel/>

그림 8. OMOP-CDM의 테이블 구조 (v5.4)

#### (4) OMOP-CDM의 활용

OMOP-CDM은 “지금까지 개발된 CDM 중 임상 정보를 가장 광범위하게 포함할 수 있는 데이터 구조를 가지도록 설계되어 진화”하고 있다.<sup>17</sup> 임상적 특성 분석, 인구 수준 추정, 환자 수준 예측 등 다양한 의료 연구에 활용될 수 있으며, 이를 통해 치료 경로와 질병의 자연사, 약물의 안전성 및 효과 등을 체계적으로 분석할 수 있도록 한다.

ODHSI 커뮤니티는 모든 핵심 리소스를 오픈 소스로 제공함으로써, CDM을 활용하여 전 세계의 연구자들이 협력하고 데이터를 공유할 수 있도록 지원한다. 다학제적 접근을 지원하여 보건 의료 데이터 분석의 혁신을 주도하고 있다.

#### 3.3.2. Sentinel CDM

Sentinel은 Sentinel Initiative의 일환으로 주요 목표는 약물, 생물학 및 기기를 포함한 FDA 규제 의료 제품의 안전성을 모니터링하기 위한 국가적 공중 보건 감시 시스템을 구축하고 운영하는 것으로 2009년 설립되었다. Sentinel Operations Center(SOC)는 sentinel data partners 네트워크를 조정하고, data partners가 로컬 데이터에 대해 분산 프로그램을 빠르게 실행할 수 있도록 하는 표준 데이터 구조인 Sentinel Common Data Model(SCDM)의 개발을 주도하고 있다.

SCDM(Sentinel Common Data Model)은 Sentinel 활동에 필요한 데이터 요소를 체계적으로 담기 위해 개발된 데이터 모델로, SOC(Sentinel Operations Center)에서 Mini-Sentinel Common Data Model Guiding Principles에 따라 설계되었다. 현재 SCDM에는 Sentinel 활동에 필요한 데이터를 포함하는 19개의 테이블로 구성되어 있으며, 여러 식별자(고유 개인 식별자인 PatID, 고유 제공 식별자인 ProviderID, 고유 시설 식별자인 FacilityID, 고유 접촉 식별자인 EncounterID)를 사용하여 테이블 간에 레코드를 연결한다(그림 9). 모든 데이터 파트너는 필수 테이블을 공통적으로 구축해야 하지만 부가적인 테이블은 모든 파트너에게 요구되지 않으며, 주로 보험회사를 제외한 의료기관에서만 CDM 형식으로 구축된다. 이러한 구조는 기관의 필요와 역할에 따라 다르게 적용되며, 각 기관의 데이터 활용 목적에 맞춘 유연한 시스템을 제공하는 데 기여한다.

또한, 센티넬은 약물 안전성 모니터링이라는 공익적 목적으로 운영되는 미국 FDA의 지원기관으로, 센티넬 CDM, SAS(Statistical Analysis Software, Cary, NC, USA) 코드로 작성된 분석 모듈, 그리고 질 평가 도구(quality assessment tool)를 센티넬 홈페이지에 공개하고 있다. 이 시스템은 의약품 부작용 모니터링을 위해 구축된 CDM 기반 다기관 분산형 분석 시스템으로, 약제 효능 평가, 공중 보건 감시(public health surveillance), 임상 의사결정(clinical decision making), 비교 효과 연구(comparative effectiveness research) 등 다양한 분야에서 활용할 수 있는 방안을 모색하고 있다.<sup>15</sup>

Administrative Data							Mother-Infant Linkage Data	Auxiliary Data		Feature Engineering Data
Enrollment	Demographic	Dispensing	Encounter	Diagnosis	Procedure	Prescribing	Mother-Infant Linkage	Facility	Provider	Feature Engineering
Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Mother ID	Facility ID	Provider ID	Patient ID
Enrollment Start & End Dates	Birth Date	Provider ID	Encounter ID & Type	Encounter ID & Type	Encounter ID & Type	Encounter ID	Mother Birth Date	Facility Location	Provider Specialty & Specialty Code Type	Encounter ID
Medical Coverage	Sex	Dispensing Date	Service Date(s)	Provider ID	Provider ID	Provider ID	Encounter ID & Type			Feature ID
Drug Coverage	Postal Code	Rx	Facility ID	Service Date(s)	Service Date(s)	Order Date	Mother Admission & Discharge Date			Feature
Medical Record Availability	Race	Rx Code Type	Etc.	Diagnosis Code & Type	Procedure Code & Type	Rx	Child ID			FE Code Type
	Etc.	Days Supply		Principal Discharge Diagnosis	Etc.	Days Supply	Childbirth Date			Etc.
		Amount Dispensed				Rx Route of Delivery	Mother-Infant Match Method			
						Etc.	Etc.			

Registry Data			Inpatient Data		Clinical Data		Patient-Reported Measures (PRM) Data	
Death	Cause of Death	State Vaccine*	Inpatient Pharmacy	Inpatient Transfusion	Lab Result	Vital Signs	PRM Survey	PRM Survey Response
Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Measure ID	Patient ID
Death Date	Cause of Death	Vaccination Date	Encounter ID	Encounter ID	Result & Specimen Collection Dates	Measurement Date & Time	Survey ID	Encounter ID
Date Imputed Flag	Source	Admission Date	Rx Administration Date & Time	Transfusion Administration ID	Test Type, Immediacy & Location	Height & Weight	Question ID	Measure ID
Source	Confidence	Vaccine Code & Type	National Drug Code (NDC)	Administration Start & End Date & Time	Logical Observation Identifiers Names and Codes (LOINC)	Diastolic & Systolic BP	Etc.	Survey ID
Confidence	Etc.	Provider	Rx ID	Transfusion Product Code	Etc.	Tobacco Use & Type		Question ID
Etc.		Etc.	Route	Blood Type				Response Text
			Dose	Etc.				Etc.
			Etc.					

\*출처: <https://www.sentinelinitiative.org/>

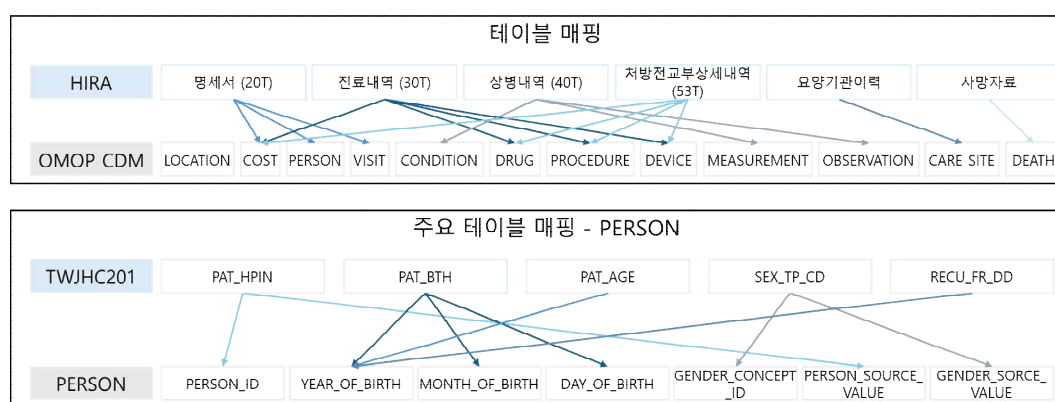
그림 9. Sentinel CDM의 구조 (SCDM v6.0)

### 3.4. Common Data Model(CDM) 국가별 동향

#### 3.4.1. 한국

국내에서는 건강보험심사평가원(이하 심평원 또는 HIRA; HEALTH INSURANCE REVIEW & ASSESSMENT SERVICE)이 ‘보건의료빅데이터개방시스템’을 통해 연구자들에게 보건의료 빅데이터를 제공하기 시작하면서 보건의료 빅데이터 활용이 본격화되었다. 심평원은 전 국민의 실사용 데이터(Real World Data, RWD)를 포함해 개별 환

자의 세부 의료이용 정보를 관리하며, 이를 표준화된 청구 서식을 기반으로 수집하고 있다. 이러한 데이터를 CDM(Common Data Model)으로 변환할 수 있는지 및 활용 가능성을 평가하기 위해, 2017년 하반기부터 심평원의 일부 빅데이터를 CDM 형식으로 전환해 연구하는 프로젝트가 추진되었다. HIRA CDM은 OMOP-CDM 형태로 구축되었으며, 건강보험 청구자료를 변환하여 “임상테이블(a), 보건의료 시스템 데이터 테이블(b), 보건의료 경제 데이터 테이블(c), 기타 의료정보 테이블(d), 의료용어 정보 테이블(e)로 구성”하고 있다(그림10).<sup>16</sup>



출처: HIRA ISSUE vol. 33, 건강보험심사평가원, 2024, 재구성

그림 10. 심평원 청구자료 테이블과 CDM 테이블 매핑<sup>16</sup>

또한, 아주대학교 박래웅 교수에 의해 2014년 6월 국내에 OHDSI가 알려지기 시작했고, 이후 국민건강보험공단, 가천길병원 등이 참여하였다. 그 결과 2016년 ODHSI에서 개별 국가를 위한 포럼을 최초로 개설하고 한국의 ODHSI 참여가 본격화되면서 현재까지 활발히 활동하고 있다.<sup>14</sup>

산업통상자원부는 2018년부터 CDM 기반 분산형 바이오헬스 빅데이터 플랫폼 구축 사업인 “FEEDER-NET” 프로젝트를 추진해왔다(그림 11).<sup>17</sup> 이 사업에서는 병원의 전자 의무기록 데이터를 CDM으로 변환해 바이오헬스 통합 빅데이터 네트워크를 구축하고, 다양한 비정형 의료 데이터(예시: 유전체, 의료영상, 생체신호 등)를 포함할 수 있는 CDM 확장 모델을 개발 중이다. FEEDER-NET이라고 명명하는 중계 플랫폼을 통해 데이터 수요자와 공급자를 연결하고 다기관 분석을 지원하며, 보건의료 연구자와 관련 기업들이 이를 활용해 혁신적인 서비스를 개발할 수 있도록 돕고 있다.

이 외에도 여러 정부 부처에서 연구과제를 통해 공공목적(의학, 정책, 의료기술 등)의 CDM 활용을 개발하기 위해 지원하고 있다.

### 3.4.2. 미국

미국은 OMOP-CDM을 개발하고 이를 바탕으로 OHDSI를 주도하는 국가로, 약 200여 개의 기관이 참여하고 있다. 현재 약 19억 명의 전자의무기록(EMR)과 보험청구자료가 CDM으로 변환되어 있으며, 이 중 20%는 EMR, 80%는 보험청구자료이다. 미국의 주요 기관들, 예를 들어 National Institutes of Health(NIH), Food and Drug Administration(FDA), National Cancer Institute(NCI), eMERGE 등이 OMOP-CDM을 공식 데이터 모델 중 하나로 채택하고 있다. 또한, All of Us 연구 프로그램에서도 OMOP-CDM을 공식 데이터 모델로 채택하였으며, 100만 명분의 임상정보와 유전체 정보를 수집하여 활용하고 있다.<sup>17</sup>

### 3.4.3. 유럽연합

혁신의학이니셔티브(Innovative Medicines Initiative, IMI)는 유럽연합 및 유럽 제약산업 협회와의 공공-민간 파트너십으로 유럽연합 내의 다양한 연구 프로젝트를 지원하고 있다. IMI2는 2018년 11월에 European Health Data & Evidence Network (EHDEN) 프로젝트를 출범하였으며, “이 프로젝트는 유럽 11개국 25개의 기관들이 보유한 의료데이터를 CDM으로 변환하는 프로젝트로써 옥스퍼드 대학, Odysseus Data Services, 유럽환자연합 등과 같은 기관에서 인증된 의료 데이터를 통합하여 1억 3천명분 이상의 환자 데이터를 익명화하여 공통 데이터 모델로 표준화”했다.<sup>17</sup>

이 프로젝트를 통해 “서로 다른 지역, 데이터 유형, 소스 어휘, 인구 규모, 그리고 서로 다른 전문성을 가진 서로 다른 팀 구성에도 불구하고 EHDEN에서 사용하는 일관된 워크플로우는 유럽 전역에서 관측 데이터의 성공적인 표준화를 지원”하는 데 적합한 것을 입증했다.<sup>19</sup> EHDEN 프로젝트는 앞으로도 수집된 실제 임상 데이터를 바탕으로 질병, 예후 및 새로운 치료법에 대한 연구를 진행하며, 이를 통해 임상적 근거를 제공할 계획이다.

## 제4장 연구 결과

### 4.1 문헌 검색 및 선정 결과

문헌 검색은 2024년 10월 21일 시행하여 10월 31일 완료하였다. 문헌 검색은 국외 데이터베이스 2개와 국내 데이터베이스 3개를 통해 진행하였으며, 국외 데이터베이스에서 503건의 문헌과 국내 데이터베이스에서 85건의 문헌이 검색되었다(표 9). 검색된 문헌 총 588건 중에서 중복된 문헌 34건을 제외하여 총 554건의 문헌이 선정되었다.

표 9. 문헌 검색 결과

데이터베이스		건수
국외	PubMed	1건
	Google Scholar	502건
국내	국가학술정보 검색시스템	32건
	한국의학논문데이터베이스(KMBASE)	22건
	학술연구정보서비스(RISS)	31건
		총 588건

중복을 제외하고 난 544건의 문헌을 일차적으로 논문의 제목과 초록을 검토하여 문헌 선정 및 배제 기준에 따라 스크리닝(screening) 작업을 실시하였다. 문헌 스크리닝 시 공통 데이터 모델과 동일한 약어 ‘CDM’을 갖는 Clinical Data Management, Clinical Decision Making, Chronic Disease Management 에 대한 연구도 본 연구의 범위에 해당하지 않는 것으로 판단되어 제외하였다. 스크리닝 결과 의료기기를 대상으로 하지 않는 문헌 354건, 공통 데이터 모델의 규제적 활용에 대한 문헌이 아닌 66건, 2015년 이전에 출판된 문헌 10건, 한국어 또는 영어 이외의 언어로 작성된 문헌 12건, 기존의 연구를 활용한 문헌고찰 및 메타분석 연구 13건, 원문이 확보되지 않은 문헌 37건으로 총 481건이 제외되었다. 스크리닝 과정 이후, 총 73건의 문헌이 1차적으로 선정되었다(표 10).

표 10. 문헌 1차 스크리닝 결과 제외된 문헌 건수

기준	건수
의료기기를 대상으로 하지 않는 문헌	344건
공통 데이터 모델의 규제적 활용에 대한 문헌이 아닌 경우	66건
2015년 이전에 출판된 문헌	10건
한국어 또는 영어 이외의 언어로 작성된 문헌	12건
기존의 연구를 활용한 문헌고찰 및 메타분석 연구	13건
원문이 확보되지 않은 문헌	37건
총 481건	

1차 스크리닝을 통해 선정된 문헌에 대해서는 전문을 확인하여 추가적인 선정 및 제외 작업을 진행하였다. 전문 검토과정에서 **SaMD** 의료기기를 대상으로 하지 않는 문헌 34건, 공통 데이터 모델의 규제적 활용에 대한 문헌이 아닌 20건, 기존의 연구를 활용한 문헌고찰 및 메타분석 연구 10건이 제외되었으며, 최종적으로 9건의 문헌이 선정되었다(표 11). 최종적으로 선정된 9건의 문헌은 모두 본 연구의 목적과 일치하는 내용을 포함하고 있으며, **SaMD**에서 **CDM**의 규제적 활용을 구체적으로 다룬 문헌들로 구성되었다.

표 11. 전문 검토 결과 제외된 문헌 건수

기준	건수
<b>SaMD</b> 의료기기를 대상으로 하지 않는 문헌	34건
공통 데이터 모델의 규제적 활용에 대한 문헌이 아닌 경우	20건
기존의 연구를 활용한 문헌고찰 및 메타분석 연구	10건
총 64건	

문헌 검색 및 선정의 전체 흐름은 (그림 11)에 상세히 나타나 있으며, 단계별 선정 및 제외 기준과 절차를 명확히 보여주고 있다.

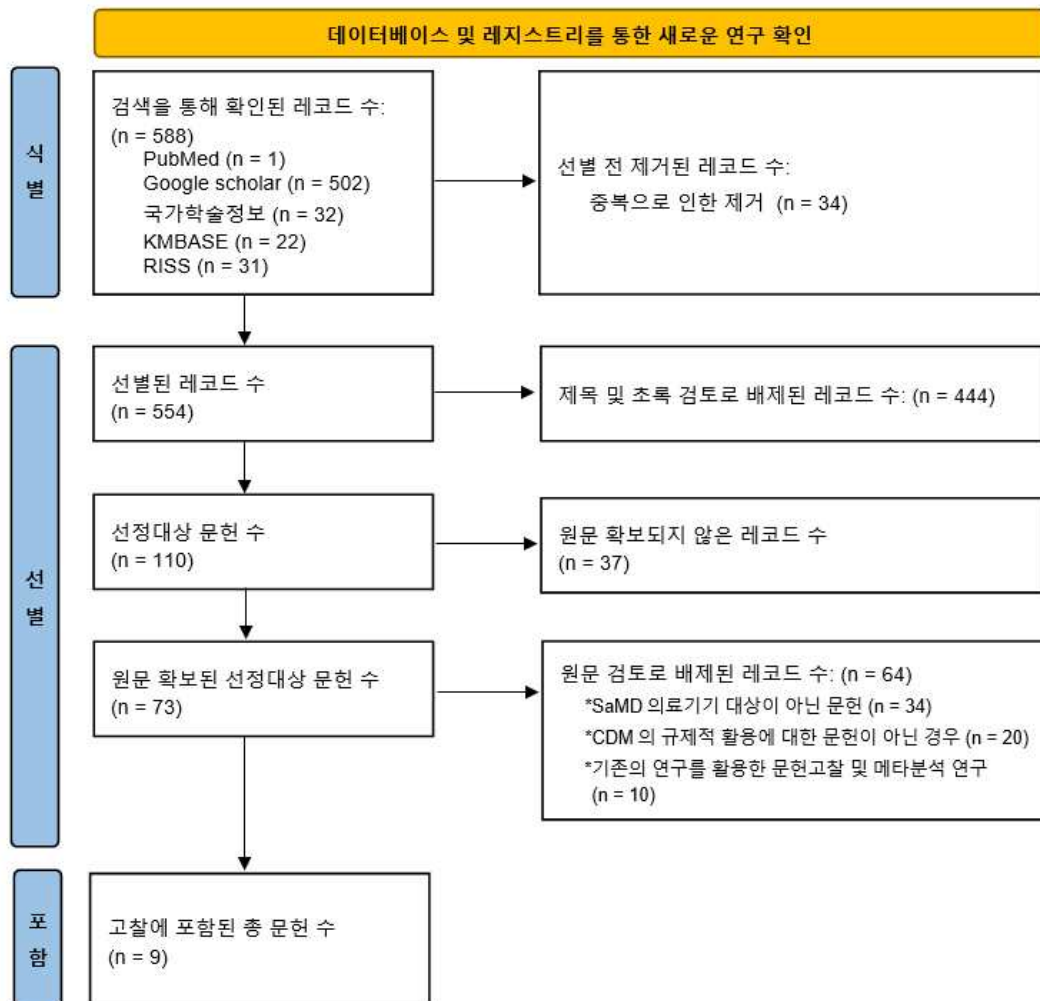


그림 11. 문헌 선정 흐름도

## 4.2. 선정된 문헌의 일반적 특성

최종 선정된 9건의 문헌에 대한 일반적인 특성은 아래 표와 같다(표 12). 9편의 문헌 중 5편은 학회지에 게재된 논문이었으며, 3편은 공식적인 보고서, 1편은 상업적 목적으로 발간된 인터넷 사설 자료였다. 총 9편 중 6편이 미국에서 실시된 문헌으로 가장 많았으며, 다음으로 유럽국가에서 실시된 문헌이 2건, 한국이 1건이었다. 공식적인



보고서 3편은 미국의 NESTcc 프로젝트로 발간된 보고서였다.<sup>18,20,22</sup> 년도별 출판 건수로 보면 2024년에 3편으로 가장 많은 건수의 문헌이 출판되었으며, 2015년에서 2024년까지 점차 증가하는 추세를 확인하였다(그림 12).

표 12. 선정 문헌의 일반적 특성

No.	Author (year)	Country	Publication Type
1	NESTcc (2021) <sup>20</sup>	USA	Report
2	Fleurence et al. (2019) <sup>21</sup>	USA	Journal
3	NESTcc (2020) <sup>22</sup>	USA	Report
4	Choi et al. (2024) <sup>23</sup>	Korea	Journal
5	NESTcc (2023) <sup>24</sup>	USA	Report
6	S. Wen et al. (2024) <sup>25</sup>	EU (Germany, Netherlands)	Journal
7	Matthew J. Page (2015) <sup>26</sup>	USA	Editorial
8	Pinsky et al. (2024) <sup>27</sup>	USA	Journal
9	Tsopra et al. (2021) <sup>28</sup>	EU	Journal

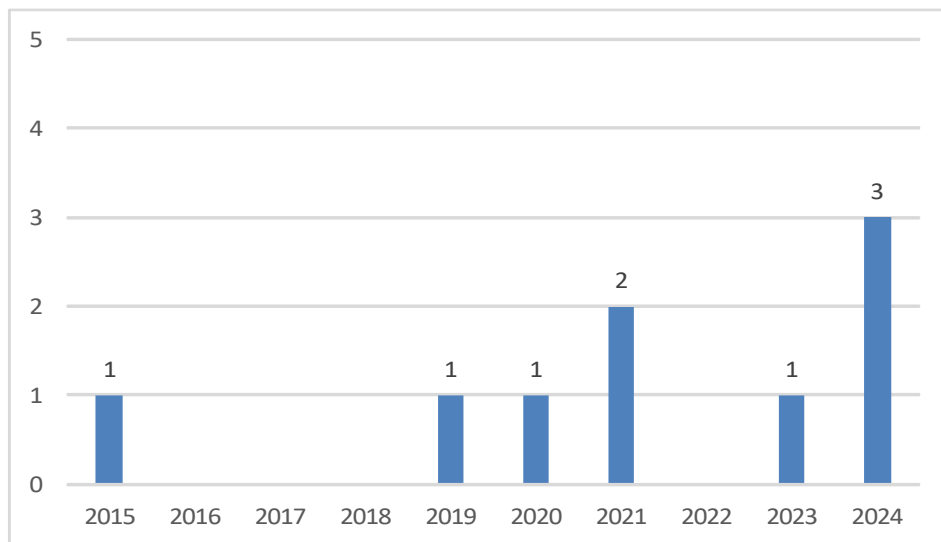


그림 12. 선정 문헌의 년도별 출판 건수

선정된 문헌의 연구 목적에 따른 CDM의 규제적 활용 방안에 대해 분류한 결과, 실사용 증거(Real-World Evidence, RWE)로 활용한 문헌이 4편으로 가장 많았으며, 의료기기의 사후관리(Post-Market Surveillance, PMS) 활동과 의료기기의 검증으로 활용한 문헌이 각각 2편, 의료기기의 개발에 활용한 문헌이 1편이었다(표 13).

표 13. 선정 문헌의 요약

No.	Author (year)	연구 목적	CDM의 활용 방안
1	NESTcc <sup>20</sup> (2021)	의료기기의 활동적 감시(Active Surveillance) 체계 구축	RWE PMS
2	Fleurence et al <sup>21</sup> (2019)	의료기기의 실사용 증거(Real-World Evidence, RWE) 생성에 있어 등록 시스템(Registry)의 발전 방향을 제시	RWE
3	NESTcc <sup>22</sup> (2020)	NESTcc 데이터 품질 프레임워크는 의료기기 생애 주기 전반에 걸쳐 고품질 RWD의 생성과 사용을 지원하는 것을 목적	RWE
4	Choi et al <sup>23</sup> (2024)	한국 의료기기 사후관리 시스템에서 NHIC 청구 코드 기반 감시의 한계를 파악하고, UDI(Unique Device Identification) 시스템 도입의 필요성을 입증	PMS
5	NESTcc <sup>24</sup> (2023)	의료 시스템 내에서 의료기기 고유기기식별자(Unique Device Identifier, UDI)를 통한 의료기기 데이터 관리 최적화	RWE
6	S. Wen et al <sup>25</sup> (2024)	어린이 암 치료에서 인공지능(AI)을 효과적으로 활용하기 위한 실용적 지침을 제시 특히 소아 암 치료에서 AI를 활용하는데 필요한 데이터 수집 및 AI 모델 구현의 규제와 방법론적 가이드라인 부족을 해결하는데 목적을 둠	의료기기 개발
7	Matthew J. Page <sup>26</sup> (2015)	실제 의료 현장에서 사용되는 약물과 의료기기의 장기적 효과성과 안전성을 입증하기 위한 RWE의 역할을 탐구	RWE
8	Pinsky et al <sup>27</sup> (2024)	집중 치료(intensive care unit, ICU) 환경에서 AI 기반 임상 의사결정 지원 시스템(CDSS)의 가능성	의료기기 검증

No.	Author (year)	연구 목적	CDM의 활용 방안
		과 관련된 주요 장벽을 규명하기 위함	
9	Tsopra et al <sup>28</sup> (2021)	정밀의학에서 사용될 AI 기술의 임상 검증에 위 한 신뢰할 수 있는 평가 프레임워크 개발	의료기기 검증

\* RWE: Real-World Evidence, PMS: Post-market Surveillance

### 4.3. CDM의 규제적 활용 사례

#### 4.3.1. 미국의 NESTcc 프로젝트

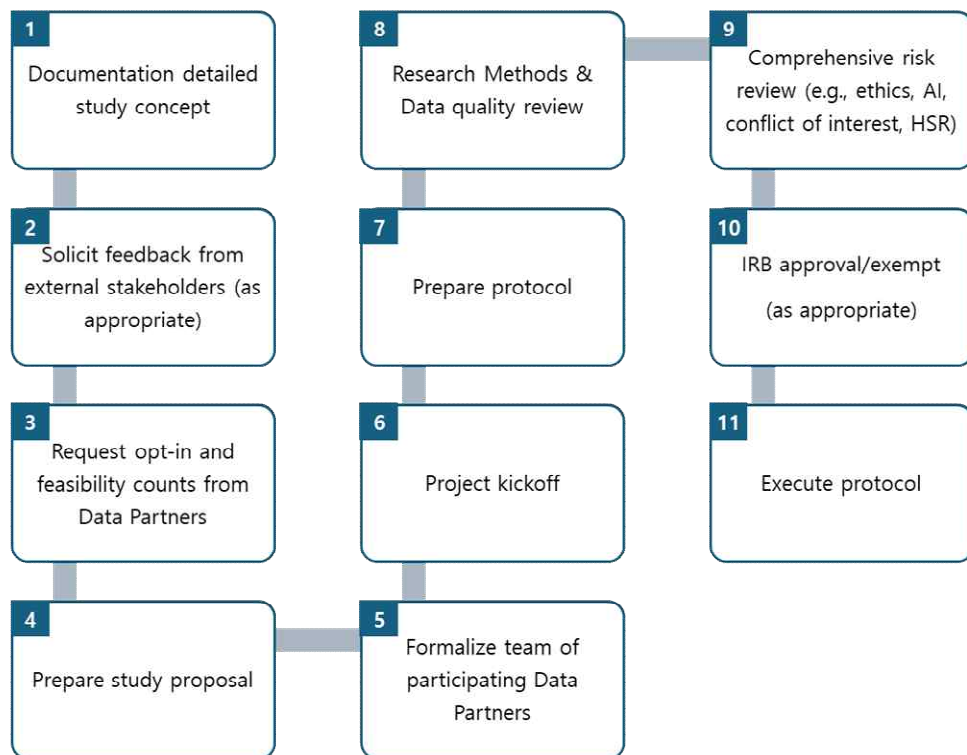
의료기기 혁신 컨소시엄(Medical Device Innovation Consortium, MDIC)으로부터 2016년에 설립된 NESTcc(National Evaluation System for health Technology Coordinating Center)는 의료기기 평가를 알리기 위해 \*실사용 데이터(Real-World Data, RWD)를 사용하여 품질과 효율성을 높이고, 전체 제품 수명 주기(TPLC) 동안 임상, 환자, 규제 및 환불 결정을 지원하는 업무를 맡고 있다. FDA는 2018년 NESTcc에 “일상적인 임상 실무에서 의료기기 사용과 관련된 기기의 성능 및 임상 결과에 대한 방대한 데이터셋을 지속적으로 생성, 액세스 및 평가”하고 “특정 기기와 부작용을 연결하는 FDA의 능력을 개선”하여 환자 안전을 위한 새로운 신호를 보다 신속하게 감지하는 시판 후 능동적 감시 기능을 구축하도록 업무를 맡겼다.

NESTcc는 사용 가능한 의료기기 유형의 다양성과 시판 전 및 시판 후 환경에서 데이터의 다양한 사용을 반영하는 21개의 테스트 케이스 프로젝트를 발표했다. NESTcc의 테스트 케이스에는 510(k) 및 시판 전 승인 규제 경로, 10개 질병 영역 및 의료기기 TPLC(Total Product Life Cycle) 전반에 걸친 프로젝트가 포함됐다. 테스트 사례는 목표에 따라 데이터 소스 및 기타 설계 특징이 다르며, 14개 테스트 사례는 라벨 확장 및 시판 전 승인 또는 허가를 포함한 시판 전 결정을 뒷받침할 수 있는 질문과 데이터를 탐구했으며, 10개는 시판 후 또는 감시 사례를 다뤘다.

\* FDA에서는 RWD를 "전자 건강 기록(EHR), 청구 및 청구 데이터, 제품 및 질병 등록부 데이터, 가정용 설정을 포함한 환자가 생성한 데이터, 모바일 기기와 같이 건강 상태를 알려줄 수 있는 다른 출처에서 수집한 데이터에서 일상적으로 수집된 환자 건강 상태 및/또는 의료 제공과 관련된 데이터"로 정의하고 있다.

### (1) NESTcc 클라우드<sup>20</sup> (Active Surveillance)

NESTcc는 의료기기에 대한 적극적 감시의 필요성을 해결하기 위해 “NESTcc 클라우드”를 개발했다. NESTcc 클라우드의 의료기기 능동 감시 프로그램 프로세스는 (그림 13)과 같다.



\*출처: Active Surveillance Roadmap, NESTcc, 재구성

그림 13. NESTcc 능동 감시 프로그램 프로세스 흐름도

NESTcc 클라우드는 다중 이해 관계자, 국가 시스템을 통해 RWE 사용을 발전시키려는 FDA의 전략을 지원하는 기반 인프라 역할을 하며, 이 시스템을 통해 능동적 감시를 활용하고, 공공 및 민간 연구에 모두 사용할 수 있도록 했다. 이는 의료기기의 비용, 시간 및 전문성을 포함해 데이터 투명성과 추적성을 보장하기 위한 인프라, 분석, 워크플로우, 데이터 품질을 보장하고 데이터 소유자 간의 데이터 공유 위험 노출을 제한하기 위한 프레임워크를 제공한다.

의료기기 시판 후 안전감시는 Sentinel Initiative에서 추진하는 것처럼 신호 감지, 신호 정제 및 신호 평가의 세 단계로 구성된다(표 14).

표 14. 감시의 3단계에 대한 정의와 응용 사례<sup>20</sup>

단계	설명 및 응용 프로그램
신호 감지 (알려진 신호 없음)	<ul style="list-style-type: none"> <li>• 상황: 사전 지정된 장치 노출에 대한 신호 없음 - 부작용 신호 쌍이 존재하지만 안전 신호에 대해 장치를 모니터링하는 것이 중요하다고 믿을 만한 가설이나 이유가 있음</li> <li>• 회고적 및/또는 전향적 데이터 수집이 사용될 수 있음</li> <li>• 응용 프로그램: 통계적 방법을 적용하여 기기와 부작용 보고 간의 연관성이 안전 신호일 가능성이 있는지 확인</li> <li>• 미래 응용 프로그램: 기계 학습 분석을 활용하여 의료기기 사용 경험과 신호(부작용 또는 이득) 존재, 잠재적 결과 간의 연관성을 다양한 범위에서 식별</li> <li>• 향후 응용 프로그램: 장치 또는 장치 클래스의 반복적이고 자동화된 모니터링 및 신호 분석을 구성하여 주요 안전 결과를 위해 자동으로 개선이 필요한 신호를 탐색</li> </ul>
신호 정제 (알려진 신호)	<ul style="list-style-type: none"> <li>• 상황: 장치와 하나 이상의 부작용 사이에 잠재적 신호가 식별되어 신호의 존재(있는 경우)를 확인하기 위해 추가 조사 및 데이터 수집이 필요한 가설 생성</li> <li>• 회고적 및/또는 전향적 데이터 수집이 사용될 수 있음</li> <li>• 가설의 검증을 목적으로 환자가 기기 또는 기기 종류에 노출된 것과 관련이 있는 것으로 생각되는 결과(들)에 관한 데이터를 수집하고 통계적으로 분석</li> <li>• 응용 프로그램: 검토 중인 “새로운 신호”의 정제, FDA 522 연구를 뒷받침하는 시판 후 감시, 장치를 사용하는 회귀 사건 또는 소규모 인구와 관련된 저주파 신호 등</li> <li>• 향후 응용 분야: 알려진 고위험 장치에 대한 범용 모니터링</li> </ul>
신호 평가	<ul style="list-style-type: none"> <li>• 상황: 장치 노출-부작용 신호 쌍이 존재하지만 인과 관계가 확립되지 않음</li> <li>• 응용 프로그램: 장치 노출과 특정 부작용 결과 간의 인과 관계의 존재 및 특성을 결정하기 위한 전체 역학 분석</li> </ul>

NESTcc는 데이터 분석의 효율적이고 일관된 실행을 위해 하이브리드 공통 데이터 모델(CDM)을 사용하여 의료기기의 특성과 같이 표현이 부족한 데이터에 대한 테이블과 데이터 요소를 포함하도록 확장했다. NESTcc CDM의 핵심은 기존 CDM 4개(OMOP-CDM, Sentinel, I2B2, PCORnet) 중에서 선택한 주요 기능을 활용하는 것이다. 데이터 모델을 평가하는 기준은 지속해서 개발 중이며, CDM이 일반적으로 사용되는 데이터 유형 및 목적, 데이터 테이블 및 요소(표 15), 예상 공변량을 수용하기 위한 CDM 확장 가능성, NESTcc와 협력하려는 CDM 관리 조직의 관심, 데이터 변환, 처리, 쿼리 및 분석을 용이하게 하는 오픈 소스 도우미 도구 및 코드의 가용성 등이 포함될 것으로 예상된다.

표 15. NESTcc 클라우드 시스템에서 예상되는 데이터 요소<sup>20</sup>

	Phase I	Phase II
<b>Data Source</b>		
EHR	O	O
Registry		O
Claims		O
Patient-generated health data (PGHD)		O
<b>Data Element</b>		
Demographics	O	O
Comorbidities	O	O
Vital signs, assessments (e.g., pain, physical therapy, mobility)	O	O
Device exposure, characteristics	O	O
Medications (orders and inpatient administration)	O	O
Labs (relevant orders and results)	O	O
Diagnostic procedures	O	O
Surgical procedures	O	O
Supplies	O	O
Outcomes: Mortality, re-hospitalization, re-operation	O	O
Outcomes: Project-specific		O
Other: As needed by use cases		O

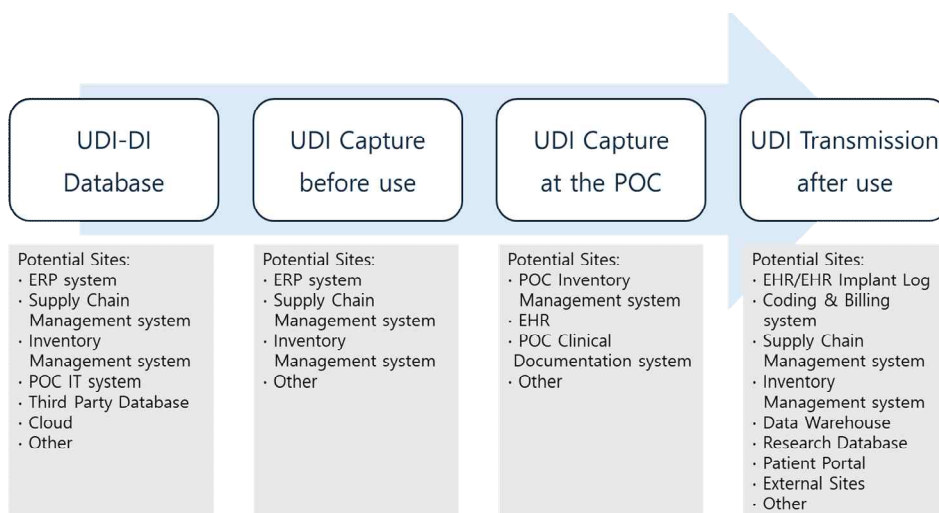
분산 데이터 네트워크가 제공하는 규모의 이점은 연구 중인 특정 장치 또는 일부 장치에 대한 노출에 초점을 맞춰야 할 필요가 있다. 이상적으로는 RWD 소스에 의약품의 국가 의약품 코드(NDC) 고유 식별자와 유사한 의료기기 고유 장치 식별자(UDI)

또는 UDI 장치 식별자(UDI-DI)가 포함되어야 한다.

## (2) UDI 시스템 구현<sup>24</sup>

UDI(Unique Device Identifier)는 의료기기의 라벨과 포장에 사람이 읽을 수 있는 형태와 기계가 읽을 수 있는 형태로 표시되는 고유 코드로, 장치 식별자(DI)와 생산 식별자(PI) 두 가지 구성 요소로 이루어져 있다. 2013년 FDA는 의료기기 제조업체가 의료기기에 UDI를 표시하도록 법적으로 의무화했다. UDI를 활용하면 임상 치료, 리콜 관리, 공급망 관리, 재무 및 수익, 부작용 보고, 연구에 UDI를 사용하면 의료기기와 관련된 정확성, 효율성, 그리고 가장 중요하게도 환자 안전이 향상될 수 있다.

UDI는 환자에게 사용된 의료기기를 정확하게 식별하고 관련 임상 결과를 추적하여 노출과 임상 결과 확인을 모두 지원할 수 있는 도구로, UDI를 의료 시스템 IT 시스템에 통합하는 것은 의료기기에 대한 RWE를 생성하는 데 중요한 요소이다. 따라서 NESTcc는 의료기기의 RWE 생성 시 RWD 소스에서 UDI 가용성을 증진하고자 건강 시스템에서 UDI 도입에 대한 플레이북을 발간했다. 해당 플레이북에서는 UDI 구현을 위한 IT 인프라 구축 프로세스를 안내하고 있다(그림 14).



\*출처: A Playbook for Health System Unique Device Identifier Implementation at the Point of Care, NESTcc, 재구성

그림 14. UDI 구현을 위한 IT 인프라 구축 프로세스

이 플레이북에서는 RWE 연구에서 UDI를 사용한 성공 사례를 소개하고 있다. NESTcc의 ThermoCool 프로젝트와 BUILD (Building UDI into Longitudinal Data for Medical Device Evaluation) 이니셔티브 사례로, 이를 통해 UDI와 CDM의 통합이 의료 데이터의 표준화와 상호운용성을 증대시켜 임상 연구 및 규제 목적을 위한 RWE 생성을 크게 지원한다고 평가하고 있다.

### (3) 데이터 품질 프레임워크<sup>22</sup>

NESTcc는 고품질 RWD의 생성과 사용을 지원하기 위해 의료기기 산업 전반의 모든 이해 관계자가 사용할 수 있는 데이터 품질 프레임워크를 제안했다. 데이터 보증 및 품질 관리는 RWE 발전을 위한 RWD의 신뢰성에 필수적이며, 고품질 데이터는 의료기기의 시판 후 평가를 지원하고 규제 의사결정을 알리는 데 필수적이다.

NESTcc 데이터 품질 프레임워크는 데이터 거버넌스, 데이터 특성, 데이터 캡처 및 변환 접근 방식, 데이터 큐레이션 모범 사례를 포함하여 고품질 데이터 달성하도록 지원한다(표 16). 이 모델은 주로 EHR 및 기타 임상 문서화 시스템을 통해 RWD 캡처 및 관리에 사용되는 의료 기관의 거버넌스, 프로세스 및 기술을 대상으로 한다.

표 16. 데이터 품질 프레임워크의 구성

Section	Description
1. Governance	이해 관계자의 대표성을 보장하고, 편견이나 비윤리적인 행동의 가능성을 제한하며, 신뢰할 수 있는 결과와 결론을 도출
2. Characteristics of Data	필요한 데이터를 이해하고 지정하고, 해당 데이터의 잠재적 출처의 맥락과 한계를 파악
3. Data Capture and Transformation	데이터를 공통 데이터 모델(CDM)로 처리하고 분석 가능하도록 변환
4. Data Curation	큐레이션은 원시 데이터를 정리, 평가, 분석을 위한 준비 과정을 거쳐 정보로 전환
5. NESTcc Data Quality Maturity Model	성숙도 모델은 조직에서 비즈니스 역량을 평가하고, 기회를 파악하고, 용량 계획을 수행하는 데 사용



### 4.3.2 실사용 증거(Real-World Evidence, RWE)로의 활용

(Fleurence et al. 2019) 연구팀은 의료기기의 RWE 레지스트리로 분산형 네트워크 모델을 활용할 것에 대한 가능성을 제시하고 있다. 이들은 PCORnet 및 NEST와 같은 분산형 네트워크 모델이 아직 초기 단계에 있지만, 의료 현장에서 생성된 방대한 전자 데이터를 효과적으로 활용할 수 있는 모델로 부상할 것이라고 전망하고 있다. 분산형 네트워크 모델이 성공하기 위해서는 건강 시스템의 지속적인 참여가 필수적이며, 데이터 품질의 관리와 방법론적 표준의 수립 및 활용이 필요함을 설명하고 있다.

궁극적으로 이러한 노력이 의료기기에서 활용할 수 있는 실사용 증거의 품질과 양을 향상시키고, 이를 통해 환자와 환자의 건강에 보다 효과적으로 기여할 수 있을 것으로 보고 있다.<sup>21</sup>

(Matthew J. Page. 2015) Page는 무작위 배정 임상시험(Randomized controlled clinical trial, RCT) 데이터와의 비교를 통해 RWE의 가치를 설명하고 있다. RWE 연구는 무작위성이 없기 때문에 증거 계층에서 RCT보다 낮은 평가를 받지만, 주로 약물, 장치 및 기타 제품의 실제 유용성을 반영하기 때문에 RCT에 대한 실행 가능한 대안 및 보완책으로 점차 여겨지고 있다.

또한, 여러 출처(국가 내 및 국가 간)의 데이터는 여러 유형의 RWE 연구를 수행하는 데 적합하게 만들 수 있으며, 여러 데이터 소스를 결합하여 일관성 있는 RWE를 생성하기 위해서는 공통 데이터 모델(CDM)을 활용하여 개별 데이터베이스의 관련 기능을 표준화하는 것이 필요함을 설명하고 있다.<sup>26</sup>

### 4.3.3 의료기기의 개발 사례

(S. Wen et al. 2024)의 연구에서는 OMOP-CDM을 통해 소아암 환자 데이터를 표준화하고 통합하여 AI 모델 훈련에 활용할 것을 제안했다. 소아암에 AI를 적용하는 데는 크게 두 가지 단계로 구분하고 있다: 첫째는 데이터를 준비하는 단계, 두 번째는 준비된 데이터를 기반으로 AI를 구현하는 단계이다 (그림 16).<sup>25</sup>

AI 애플리케이션은 혁신적인 모델을 훈련하고 개발하는 데에 있어 데이터에 크게 의존한다. 소아암 환자의 여정 전반에 걸쳐 다양한 유형의 데이터가 생성되며, 이러한 데이터 유형의 예로는 EHRs (electronic health records) 데이터, 공중 보건 데이터, 환자 보고 결과(Patient-reported Outcomes, PRO) 등이 포함된다. 이 과정에서 CDM을 통한

접근은 표준화된 프레임워크를 제공하여 다양한 데이터 소스에서 원활한 통합과 분석을 용이하게 함으로써, 다양한 유형의 데이터를 AI에 활용할 수 있도록 한다.

유럽 건강 데이터 및 증거 네트워크(EHDEN)는 OMOP-CDM을 사용하여 소아암 환자의 데이터를 포함한 5억 개가 넘는 익명화된 건강 기록을 통합하고 조화시키고 있다. 이를 통해 다양한 출처의 데이터가 균일한 구조와 형식으로 표준화되어, AI 기반 소아암 연구에 필요한 의미 있는 데이터를 추출할 수 있도록 지원한다.<sup>25</sup>

#### 4.3.4 의료기기의 검증 사례

(Pinsky et al, 2024)의 연구는 AI와 기계 학습 모델이 실시간으로 환자의 상태를 분석하고 예측하는 데 어떻게 활용될 수 있는지, 그리고 이를 구현하는 데 필요한 데이터의 품질과 접근성 문제에 대해 다루고 있다.

Pinsky 연구팀은 \*MIMIC-IV 데이터셋을 OMOP CDM에 매핑하는 최근의 노력을 긍정적으로 평가하며, CDM은 여러 의료 기관에서 서로 다른 EHR 데이터 구조를 표준화해, 다양한 환자 집단에 대해 AI 모델을 검증하는 데 중요한 역할을 한다고 언급하였다. 또한, 이런 공개 데이터베이스를 표준 어휘를 사용하여 공통 데이터 모델에서 사용할 수 있다면 외부 검증에 대한 장벽을 해소할 수 있을 것으로 전망하고 있다.<sup>27</sup>

(Tsopra et al. 2021)의 연구팀은 ITFoC(Information Technology for the Future Of Cancer) 컨소시엄을 통해 종양학에서 치료반응 예측을 위한 AI 기술의 임상적 검증을 위한 프레임워크를 제안하였다. 이 프레임워크는 AI 알고리즘을 실제 임상 데이터와 오믹스(omics) 데이터로 검증하여 AI의 성능과 안전성을 체계적으로 평가하기 위한 목적이며, 총 7단계로 구성된다(그림 15).

Tsopra 연구팀은 AI 평가를 위한 데이터의 품질, 공유, 이식성 및 재사용성을 보장하기 위해 표준화를 강력히 권장하고 있다. 이 프레임워크에서는 임상 데이터 세트는 데이터 구조를 위한 OMOP 공통 데이터 모델과 데이터 내용을 위한 OSIRIS 모델로 표준화된다. 두 표준 모두 개념 및 소스 값 목록을 포함하며, 이는 종양학에서 임상

---

\* MIMIC-IV(Medical Information Mart for Intensive Care IV) 데이터셋은 2008년에서 2019년 사이에 BIDMC(Beth Israel Deaconess Medical Center) ICU(중환자실) 또는 응급실에 입원한 환자의 EHR 공개 데이터셋이다.

및 생물학적 데이터를 공유하는 데 필요한 최소 데이터 세트로 간주된다. 또한, 데이터 값은 ICD 10, LOINC, SNOMED CT와 같은 국제 의학 용어에 따라 구조화되어 AI 시스템 간의 상호 운용성을 보장한다.

Tsopra 연구팀은 AI의 궁극적인 성능은 평가에 사용되는 데이터의 품질에 크게 좌우된다고 보고하고 있으며, 이 프레임워크는 AI 기술의 정밀의학 적용을 가속화하는데 있어 중요한 발판이 될 것으로 기대하고 있다.<sup>28</sup>

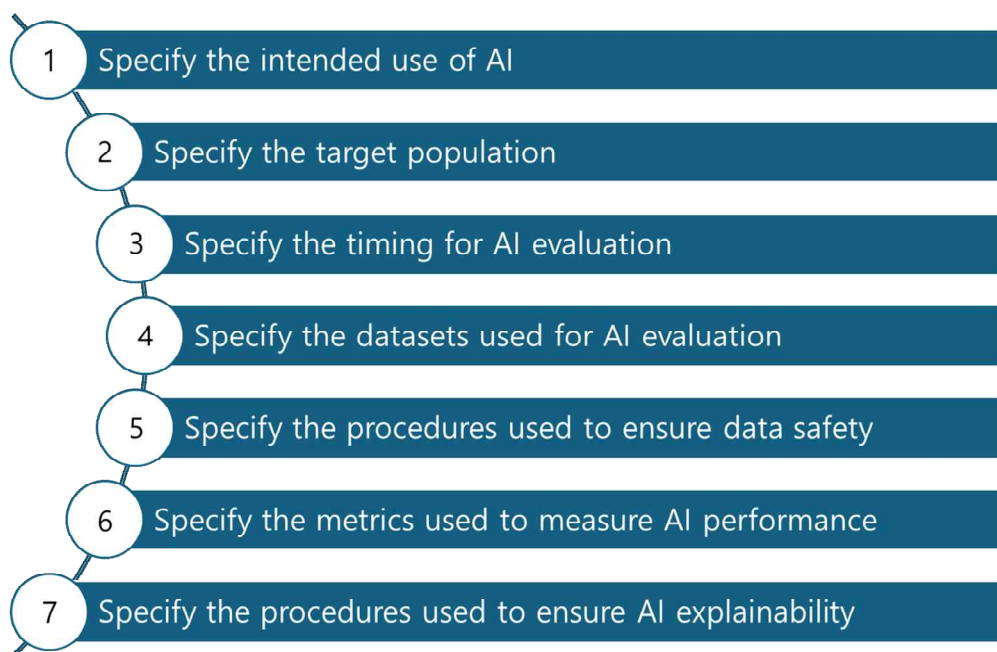


그림 15. AI 기술의 임상적 검증에 필요한 7가지 핵심 단계

#### 4.3.5. 시판 후 감시(PMS) 사례

(Choi et al. 2024) 연구에서는 한국의 CDM 기반 분산형 바이오헬스 빅데이터 플랫폼(FEEDER-NET)의 의료기기 사후관리 시스템에서 NHIC(national health insurance corporation) 청구 코드 기반 감시의 한계를 파악하고, UDI(Unique Device Identification) 시스템 도입의 필요성을 입증하고자 UDI와 NHIC 청구 코드 각각의 매칭 상황을 분

석해 의료기기 안전 감시에서의 효과성을 비교하였다.

Choi 연구팀은 NHIC 시스템 내 UDI의 통합 및 적용 범위를 평가하는 것을 목표로 구체적으로 세 가지 주요 질문을 다루었다.

- NHIC 적용 항목으로 분류된 UDI 수
- 각 NHIC 청구 코드에 포함된 UDI 수
- 각 UDI가 등록된 NHIC 청구 코드 수

연구 결과, NHIC 청구 코드는 전체 의료기기의 약 36.02%에 해당하는 UDI만을 포괄하며, 이는 (A) 의료기기, (B) 의료용품, (C) 치과 재료 범주로 제한되었으며 대부분의 소프트웨어와 체외 진단(IVD)은 제외되었다. 또한, 특정 NHIC 청구 코드에 다수의 UDI가 매핑되는 경우가 존재했으며, 이러한 다중 매핑 문제는 신호 감지 민감도를 저하시킬 수 있고 특히 결함 있는 모델을 판별하기 어렵게 하는 문제점을 발견했다. UDI 시스템은 의료기기 식별 및 추적의 정확성을 향상시켜 다양한 소스의 데이터를 통합하여 실제 증거(RWE)를 생성하는 데 중요한 역할을 할 수 있다.

결론적으로, Choi 연구팀은 NHIC 청구 코드 기반 PMS는 세분성과 신호 감지 감도에 한계가 있어 이러한 문제를 해결하기 위해서는 UDI 기반 PMS를 도입해야 함을 확인하였다.<sup>23</sup>

#### 4.4. CDM 활용의 주요 패턴 및 경향 분석

독립형 소프트웨어 의료기기(SaMD)에서 공통 데이터 모델(CDM)의 규제적 활용은 규제 기관과 의료기기 제조업체들이 임상 데이터의 통합적이고 표준화된 관리 및 평가를 위해 채택하는 중요한 수단이 되고 있다. 9건의 문헌을 통해 CDM의 규제적 활용 패턴을 분석하였으며, 주요 패턴은 (1) 실사용 증거(RWE)로의 활용, (2) 의료기기 개발, (3) 의료기기 검증, (4) 시판 후 감시 활동 4가지로 분류되었다.

특히, 실사용 증거로 활용하기 위한 연구가 가장 활발히 이루어지고 있다. CDM을 통해 고품질의 실사용 데이터(RWD)를 구축하고, 이를 기반으로 RWE를 생성하는 노력이 이루어지고 있다. 이 과정을 통해 구축된 RWE는 의료기기의 시판 전·후 규제

의사결정을 지원하는 실질적인 근거로 활용될 수 있으며, 따라서 CDM 기반 RWE의 효과성을 입증하는 것이 중요한 공통 과제로 자리잡고 있다.

#### (1) 실사용 증거(RWE)로의 활용

CDM은 실사용 증거(Real-World Evidence, RWE) 기반의 규제적 평가를 가능하게 하여, SaMD의 실질적 안전성과 유효성을 평가하는 데 중요한 도구로 작용할 수 있다. NESTcc 프로젝트는 RWE를 활용한 규제 평가와 시판 후 감시체계의 효과성을 입증하기 위한 다양한 연구를 수행하고 있으며, 이를 통해 의료기기의 품질과 안정성을 지속적으로 향상시키는 데 기여하고 있다.

#### (2) 시판 후 감시 활동

CDM은 SaMD의 시판 후 감시(PMS) 활동에서 실질적인 역할을 수행하며, 제품의 안전성과 유효성을 지속적으로 모니터링하기 위한 데이터 기반 접근을 가능하게 한다. NESTcc 프로젝트<sup>20,22,24</sup>와 한국에서의 사례(Choi et al. 2024)<sup>23</sup>를 통해 CDM 기반 PMS 구축 시 UDI(Unique Device Identification) 시스템의 도입 필요성을 확인하였다.

#### (3) 의료기기 개발 및 검증에서의 활용

CDM을 활용한 데이터 표준화는 SaMD의 개발 및 검증에서 매우 중요한 역할을 한다. 특히 인공지능(AI) 기반 SaMD의 개발에서는 신뢰성 있는 데이터 수집 및 분석이 필수적이며, CDM 기반 데이터의 통합은 이러한 요구를 효율적으로 충족할 수 있도록 한다. 선행 연구에서는 AI 기반 SaMD의 개발 및 검증 과정에서 데이터 품질을 보장하기 위한 프레임워크를 구축하는 노력이 이루어지고 있다.

미국과 유럽을 포함한 주요 국가들은 SaMD의 규제적 활용에 있어 CDM 기반 데이터 표준화 및 통합을 위한 노력을 지속하고 있으며, 이해관계자들의 협력 체계가 형성되고 있다. OMOP-CDM과 같은 모델은 전 세계적으로 채택되고 있으며, 의료기기의 규제적 활용을 강화하기 위한 기반을 제공하고 있다. 다학제간 협력을 통해 SaMD의 효율적 평가와 관리의 근간이 되고 있으며, 향후 규제적 활용이 더욱 발전할 가능성을 시사한다.

## 4.5. CDM의 규제적 활용에 대한 잠재적 이점 및 한계

선행 연구에서 공통적으로 CDM은 다양한 데이터 소스의 통합과 표준화가 가능하기에 상호운용성이 향상된다고 보고하고 있다. 이를 통해 CDM은 다양한 연구 및 보건 시스템 간의 협력과 일관성을 가능하게 한다(S. Wen et al. 2024, NESTcc. 2020). 또한, 다양한 의료 기관의 데이터 통합하여 이를 AI 기반 SaMD의 개발에 활용한다면 의료기기의 성능도 더욱 향상시킬 수 있는 이점이 있다(Tsopra et al. 2021).

각국의 규제당국에서는 의료기기의 실사용 증거에 대한 가치를 인정하고 이를 규제적으로 적용하기 위해 다양한 가이드라인을 발간하고 있다. 국내 MFDS는 2023년 7월 「의료기기의 실사용증거(RWE) 적용에 대한 가이드라인 (민원인 안내서)」<sup>29</sup>를 제정하여 실사용 증거를 규제 결정에 사용할 수 있음을 제시하였으며, 실사용 증거 적용 시 신뢰성 입증에 대해 고려해야 할 사항에 대해 제시하고 있다. 미국 FDA 역시 “Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices”<sup>31</sup> 지침 초안을 2023년 12월 발간하여 RWE의 규제적 활용을 적극 수용하고 있다. 이러한 규제와 조화를 이루어 CDM을 활용해 고품질의 실사용 증거(RWE)를 마련한다면, SaMD를 포함한 의료기기와 관련한 규제적 접근에 있어 신뢰성과 효율성을 향상시킬 수 있을 것이다. 이를 통해 규제기관의 요구사항을 충족하면서도, 혁신적인 기술과 의료기기의 안전성과 효과성을 보다 효율적으로 검증하고 활용할 수 있는 기반이 마련될 것으로 사료된다.

그러나 CDM의 규제적 활용에는 일부 한계도 존재한다. 첫째, 다양한 데이터 출처와의 상호운용성 문제와 국가별 규제 환경 간 차이가 있으며, 데이터 품질과 신뢰성을 보장하기 위한 추가적인 규제적 표준화와 데이터 관리 방안이 필요하다(Fleurence et al. 2019). 둘째, 데이터 통합 및 표준화 과정에서 데이터의 특수성과 다양성을 충분히 반영하지 못할 경우, 의료기기 평가와 관련한 실질적 유효성을 저해할 수 있다(Tsopra et al. 2024). 마지막으로, 이런 시스템을 구축하기 위해서는 상당한 비용과 기술적 요구가 필요하며(NESTcc, 2021), 높은 성숙도를 달성하기 위해 상당한 인프라 투자, 거버넌스, 인력 교육 또한 필요하다(NESTcc. 2020). 이런 한계점을 해결하기 위해, CDM의 규제적 활용이 SaMD 평가의 신뢰성과 효율성을 제고할 수 있도록 지속적인 연구와 표준화 노력이 필요하다. 다기관 협력 연구 및 규제기관과의 협업을 통해, CDM 기반의 데이터 활용을 더욱 정교하게 발전시키는 것이 SaMD의 안전성과 효율성을 강화하는 데 기여할 것이다.

#### 4.6. SaMD에서 CDM의 규제적 활용에 대한 국내 도입 제언

SaMD는 국내 의료기기 시장에서 빠르게 성장하고 있는 분야로, 효율적이고 체계적인 규제적 평가가 필수적이다. 이를 위해 CDM의 도입과 활용을 강화하는 것이 중요한 과제로 작용하고 있다. CDM은 다양한 의료 데이터 소스를 표준화하고 통합하여 데이터의 일관성을 확보하며 다양한 분석 가능성을 제공하는 강력한 도구로, SaMD의 규제적 활용에서도 핵심적인 역할을 할 수 있다. 최근, 식품의약품안전처의 산하기관인 한국의료기기안전정보원에서는 의료기기 사후관리 측면에서 CDM 활성화를 위한 기술세미나를 개최하는 등, 국내 규제당국에서도 의료기기에서 CDM 활용에 대해 높은 관심을 보이고 있다.

SaMD의 규제적 관리에서 CDM 활용의 구체적인 방안으로, 시판 후 감시체계를 강화하기 위한 CDM 기반의 UDI 시스템 도입을 구축하는 것이다. UDI 시스템은 SaMD 제품의 식별 및 추적을 가능하게 하여, 제품의 사용 중 발생하는 문제를 조기에 발견하고 대응할 수 있는 기반을 제공한다. CDM은 이를 보완하여 UDI 데이터를 실시간으로 수집, 표준화, 분석할 수 있는 플랫폼을 제공하며, SaMD의 안전성과 효과성을 평가하는 데 중요한 역할을 할 수 있다.

이를 위해서는 규제기관, 산업체, 학계 간의 긴밀한 협력이 필요하며, 글로벌 표준에 부합하는 규제 가이드라인을 개발하여 CDM 기반의 데이터 분석과 활용을 체계적으로 지원해야 한다. 또한, CDM을 통해 실시간 데이터 수집과 통합을 가능하게 하여 SaMD의 시판 후 감시 활동을 강화하고, 안전성과 유효성에 대한 신속한 규제적 의사결정을 지원하는 체계적인 기반을 구축할 필요가 있다.

CDM 도입은 SaMD의 시판 후 감시 및 규제적 평가의 효율성을 크게 향상시킬 뿐만 아니라, 데이터 기반 의사결정의 정밀도를 높이는 데 기여할 것이다. 이는 SaMD 시장의 신뢰성과 안전성을 강화하고, 국내 의료기기 산업의 글로벌 경쟁력을 확보하는 데 중요한 역할을 할 것이다. 또한, CDM 기반 규제 체계는 SaMD 제조업체들이 보다 신속하고 정확하게 규제 요구사항을 충족할 수 있도록 지원하며, 혁신적인 제품 개발과 시장 진출을 촉진하는 데 기여할 수 있다. 결론적으로, 국내 SaMD 산업의 지속 가능한 발전과 혁신을 위해 CDM의 도입과 활용은 필수적이다.



## 제5장 고찰

본 연구는 SaMD에서 공통 데이터 모델(Common Data Model, CDM)의 규제적 활용을 다각도로 분석하였다. CDM은 의료 데이터의 표준화, 상호 운용성 개선, 데이터 품질 보장을 통해 SaMD의 개발 및 검증 과정에서 필수적인 역할을 수행하며, 특히 실사용 증거(Real-World Evidence, RWE)를 기반으로 한 SaMD의 규제적 평가와 시판 후 감시(Post-Market Surveillance, PMS)에 있어 그 중요성이 점점 더 부각되고 있다.

본 연구는 미국, 유럽, 한국을 중심으로 CDM의 규제적 활용 사례를 분석하였다. CDM의 규제적 활용은 각 국가의 보건의료 시스템의 특성에 따라 다양하게 적용되고 있음을 알 수 있다. 미국의 NESTcc 프로젝트는 CDM을 활용한 시판 후 능동 감시(active surveillance)와 데이터 기반 의료기기 검증 사례를 보여주며, 실사용 데이터의 통합적 활용이 SaMD 규제 개선에 기여할 수 있음을 시사한다. 유럽연합의 EHDEN 프로젝트는 유럽 전역의 다양한 데이터 소스를 CDM으로 통합하여 임상적 근거를 창출하고 있다. 이 프로젝트는 다국적 협력을 통해 다양한 데이터 소스를 표준화함으로써 데이터 활용성을 극대화하고 있다. 이는 글로벌 CDM 활용의 선도적 사례로, SaMD 규제 효율성을 높이는 데 중요한 시사점을 제공한다. 한국에서는 심평원의 데이터를 기반으로 한 OMOP-CDM 구축 및 FEEDER-NET 프로젝트가 주요 사례로 언급되며, 이러한 사례는 한국의 의료 산업 특성을 반영한 CDM 활용 방안을 제시하며, 국내 SaMD 산업의 지속 가능한 발전 가능성을 보여준다.

CDM은 SaMD의 개발 및 평가를 체계화하고, 규제적 신뢰성을 확보하며, 다국가 간 데이터 비교 및 협력을 가능하게 하는 강력한 도구로 작용할 수 있다. CDM의 도입은 의료 데이터의 일관성과 표준화를 촉진하여 SaMD의 품질과 안전성을 보장할 수 있는 기반을 제공한다. 특히, CDM 기반의 데이터 활용은 SaMD의 시판 후 감시(PMS)와 실사용 증거(RWE) 생성에서 중요한 역할을 하며, 이를 통해 SaMD의 전 주기에 걸친 안전성과 유효성을 보다 체계적으로 관리할 수 있다. 그러나 본 연구는 CDM의 규제적 활용 과정에서 몇 가지 주요 과제를 확인하였다. 첫째, 국가별 규제 환경 및 데이터 보호 규정의 차이는 CDM 활용의 일관성을 저해할 수 있다. 둘째, 데이터 통합 및 변환 과정에서 발생할 수 있는 기술적 어려움과 리소스 부족은 규제적 활용의 걸림돌로 작용한다. 셋째, CDM 기반 연구의 신뢰성을 높이기 위해서는 데이터 품질 관리와 비식별화 절차가 엄격히 유지되어야 한다.



## 제6장 결론

본 연구는 주제범위 문헌고찰 방법론을 기반으로 독립형 소프트웨어 의료기기 (SaMD)에서 공통 데이터 모델(CDM)의 규제적 활용 사례와 주요 경향을 분석하였다. CDM은 데이터의 표준화와 통합을 통해 SaMD의 실사용 증거(RWE) 생성과 시판 후 감시(PMS)에서 중요한 역할을 하며, 이를 통해 규제적 신뢰성을 높이는 데 기여할 수 있음을 도출하였다. 특히 미국의 NESTcc와 같은 국제적 사례를 통해 CDM을 활용한 SaMD의 안전성 및 유효성 보장 사례를 강조하였으며, 한국에서의 UDI 기반 시판 후 감시 시스템 도입 필요성을 확인하였다.

결과적으로, SaMD에서 CDM 기반 규제 접근은 데이터의 국제 표준화 및 상호운용성 향상을 통해 SaMD 시장의 규제 효율성을 강화하는 중요한 도구로 자리 잡고 있음을 명확히 하였다. 이는 SaMD의 안전성과 성능을 보장하는 데 있어 데이터 기반의 규제적 의사결정 지원에 중추적 역할을 할 수 있음을 시사하며, SaMD 개발과 관리 과정에서 CDM의 가치를 재확인하고 향후 글로벌 규제 협력 및 다기관 연구의 필요성을 제시한다.

본 연구의 한계는 CDM의 규제적 활용이 일부 국가와 특정 프로젝트에 국한된 사례를 중심으로 분석되었다는 점이다. CDM의 다양한 활용 가능성에도 불구하고 국가별로 상이한 규제 프레임워크와 CDM의 채택 수준 차이는 분석 범위 내에서 충분히 다루지 못한 측면이 있다. 또한, SaMD의 규제 환경이 급격히 변화하고 있는 현재 상황에서 규제적 활용에 관한 장기적인 효과 분석이 다소 제한적이라는 점이 존재한다.

향후 연구에서는 SaMD에서 CDM의 활용을 더욱 다양한 국가 및 규제 환경에서 조사하고, 국제 표준화와 상호운용성 강화 방안을 도출할 필요가 있다. 또한, AI 기반 SaMD의 규제적 활용을 위한 CDM 데이터 품질 관리 방안과 새로운 규제적 지침의 장기적인 효과 분석에 대한 연구가 필요하다. SaMD의 안전성과 유효성을 지속적으로 향상시키기 위해 다양한 이해관계자들이 협력하여 규제적 데이터 활용을 더욱 정교하게 발전시켜야 한다. 이러한 노력이 SaMD 시장과 환자 중심의 의료 서비스 발전에 기여할 수 있을 것이다.

## 참고 문헌

1. 식품의약품안전처. 2020. 신개발 의료기기 전망 분석 보고서.
2. 삼성경제연구소. 2018. 스마트 헬스케어의 시대, 데이터 전쟁을 대비하라.
3. Allied Market Research. 2024. Software as a Medical Device (SaMD) Market Size, Share, Competitive Landscape and Trend Analysis Report, by Deployment Type, by Application, by End User : Global Opportunity Analysis and Industry Forecast, 2024-2033.
4. Hilary Arksey, Lisa O' Mally. 2005. Scoping Studies: Towards a Methodological Framework.. *International Journal of Social Research Methodology*. 8:1, 19-32.
5. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. 2018. PRISMA Extension for Scoping Reviews (PRISMA ScR): Checklist and Explanation. *Ann Intern Med*. 169:467-473.
6. 식품의약품안전평가원. 2023. 의료기기 소프트웨어 허가·심사 가이드라인.
7. IMDRF. 2013. Software as a Medical Device (SaMD): Key Definitions.
8. IMDRF. 2014. Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations
9. 양희영. 2019. 디지털헬스 의료기기(SaMD) 규제 기준 정립 및 진료행위 코드 마련. 보건산업진흥원. 보건산업브리프 Vol.292.
10. <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>
11. MDCG. 2019. Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 - MDR and Regulation (EU) 2017/746 - IVDR.
12. 식품의약품안전처. 2021. 소프트웨어 의료기기 제조소의 GMP 운영을 위한 민원인 안내서.
13. 식품의약품안전처. 2024. 2023년 의료기기 허가보고서.
14. 한국 오딧세이. 2019. 오딧세이 바이블 The Book of OHDSI.
15. 강민규, 강동윤, 강혜련, 이수현, 김주한. 2017. 약물유해반응 관리의 새 흐름: 수동적 유해사례 수집에서 분산형 약물감시망을 이용한 대규모 능동적 검색으로의 전환. 약물역학위해관리학회지. 9:1-10.

16. 박소정. 2024. 건강보험심사평가원 공통데이터모델(Common Data Model) 구축과 개방. 건강보험심사평가원. HIRA ISSUE vol.33.
17. 박래웅. 2019. 공통 데이터 모델과 분산연구망: 오딧세이 컨소시엄(Observational Health Data Sciences and Informatics, ODHSI) 연구사업. 대한내과학회지. 제94권 제4호, 309-314.
18. 박래웅, 유승찬. 2018. 보건의료 분야 CDM 적용 사례 및 미래 활용 방안. 건강보험심사평가원. HIRA 빅데이터 브리프. 제2권 4호.
19. Erica A. Voss, Clair Blacketer, Sebastiaan van Sandijk, Maxim Moinat, Michael Kallfelz, Michel van Speybroeck, Daniel Prieto-Alhambra, Martijn J. Schuemie, Peter R. Rijnbeek. 2024. European Health Data & Evidence Network-learnings from building out a standardized international health data network. *Journal of the American Medical Informatics Association*. 31(1), 209 - 219.
20. NESTcc. 2021. Active Surveillance Roadmap.
21. Rachael L. Fleurence, Kathleen Blacke, Jeffrey Shure. 2019. The Future of Registries in the Era of Real-world Evidence for Medical Devices. *JAMA Cardiology*. E1-E2.
22. NESTcc. 2020. National Evaluation System for health Technology Coordinating Center (NESTcc) Data Quality Framework.
23. Sooin Choi, Jin Kuk Kim, Jinhyoung Lee, Soo Jeong Choi, You Kyoung Lee. 2024. Limitations of NHIC claim code based surveillance and the necessity of UDI implementation in Korea. *nature portfolio*. 14:21014.
24. NESTcc. 2023. A Playbook for Health System Unique Device Identifier Implementation at the Point of Care.
25. Shuping Wen, Stefan Theobald, Pilar Gangas, Karina C. Borja Jim'enez, Johannes H.M. Merks, Reineke A. Schoot, Marcel Meyerheim, Norbert Graf, on behalf of UNICA4EU. 2024. A practical guide to apply AI in childhood cancer: Data collection and AI model implementation. *EJC Paediatric Oncology*. 4:100197.
26. Matthew J. Page. 2015. GET REAL: Demonstrating Effectiveness with Real World

- Evidence. *MEDPACE* (<https://www.medpace.com/>).
27. Michael R. Pinsky, Armando Bedoya, Azra Bihorac, Leo Celi, Matthew Churpek, Nicoleta J. Economou-Zavlanos, Paul Elbers, Suchi Saria, Vincent Liu, Patrick G. Lyons, Benjamin Shickel, Patrick Toral, David Tscholl and Gilles Clermont. 2024. Use of artificial intelligence in critical care: opportunities and obstacles. *Critical Care*. 28:113.
  28. Rosy Tsopra, Xose Fernandez, Claudio Luchinat, Lilia Alberghina, Hans Lehrach, Marco Vanoni, Felix Dreher, O.Ugur Sezerman, Marc Cuggia, Marie de Tayrac, Edvins Miklasevics, Lucian Mihai Itu, Marius Geanta, Lesley Ogilvie, Florence Godey, Cristian Nicolae Boldisor, Boris Campillo Gimenez, Cosmina Cioroboiu, Costin Florian Ciusdel, Simona Coman, Oliver Hijano Cubelos, Alina Itul, Bodo Lange, Matthieu Le Gallo, Alexandra Lespagnol, Giancarlo Mauri, H.Okan Soykam, Bastien Rance, Paola Turano, Leonardo Tenori, Alessia Vignoli, Christoph Wierling, Nora Benhabiles and Anita Burgun. 2021. A framework for validating AI in precision medicine: considerations from the European ITFoC consortium. *BMC Med Inform Decis Mak*. 21:274.
  29. 식품의약품안전평가원. 2023. 의료기기의 실사용증거(RWE) 적용에 대한 가이드라인.
  30. 김지현. 2021. BIO ECONOMY REPORT: 실사용증거(Real-World Evidence) 활용 동향. 한국바이오경제연구센터. 30, 1-13.
  31. FDA. 2023. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices (DRAFT).

## **ABSTRACT**

### **Regulatory Applications of Common Data Model in Standalone Software as a Medical Device**

This study aims to systematically analyze the regulatory application of the Common Data Model (CDM) in standalone software as a medical device (SaMD) by applying the scoping review methodology. SaMD is growing rapidly along with the development of digital healthcare, and data quality management is very important in development, verification, and post-market management due to the characteristics of software. CDM supports the development and verification of SaMD through data standardization and integration, and contributes to enhancing regulatory reliability in terms of real-world evidence (RWE) generation and post-market surveillance (PMS).

As a result of the study, the active surveillance system of the National Evaluation System for Health Technology Coordinating Center (NESTcc) in the United States and the necessity of the Unique Device Identifier (UDI) system in the post-market surveillance system in Korea were confirmed as major cases of regulatory utilization of SaMD based on CDM. The use of CDM plays an important role in improving international standardization, data quality assurance, and interoperability. However, there are differences in regulatory environments and limitations in the data integration process among countries. Therefore, continuous research and multi-agency cooperation are needed to address these issues.

This study clearly presents the value and challenges of using CDM in improving the regulatory efficiency of SaMD, and can be used as basic data for future SaMD regulatory improvement and data-based decision-making.

---

**Key words** : SaMD, Common data model, Real-World evidence, Post-market Surveillance