약물 부작용 사례 코딩용어체계: MedDRA를 중심으로

서울대학교병원 임상시험센터¹, 서울대학교 의과대학/서울대학교병원 의학연구협력센터²,
서울대학교 의과대학 예방의학교실³

안희정¹, 김선하^{2,3}, 한서경², 박병주^{1,2,3}

=Abstract=

A Review on Coding Dictionary for Safety Data: with Focus on the MedDRA

Hee Jung Ahn¹, Seon Ha Kim^{2,3}, Seokyoung Hahn², Byung Joo Park^{1,2,3}

Clinical Research Center, Seoul National University Hospital¹, Medical Research Collaborating Center,
Seoul National University College of Medicine/ Seoul National University Hospital²,
Department of Preventive Medicine, Seoul National University College of Medicine³

Costart, Whoart, and MedDRA have been widely used for classification of drug safety data for the purpose of analysis and retrieval. MedDRA is, however, different from the other dictionaries in many aspects, mainly in size, granularity, and its structure. The MedDRA is characterized by its hierarchical and, multiaxial features and its regular update, which can have an impact on the safety data analysis as well as a presentation of the data. The MedDRA is globally preferred as it is continuously updated by reflecting users requests, and many pharmaceutical companies and regulatory authorities required MedDRA as a compulsory for electronic transmissions. Therefore, understanding the coding system of MedDRA is very important for clinical investigators before they use it to record and report their drug safety data. An appropriate use of coding system should enhance the applicability of drug safety database.

Key words : Coding dictionary, Drug safety data, MedDRA, Clinical trial, Adverse event

서론

약물안전성자료는 약물시판 전에는 임상시험 과정을 통해서, 시판 후에는 자발적인 부작용 보고사례나 재심사제도를 통해 수집된다. 시판 전에 안전성과 효능이 입증되어 시판 후에 널리 사용되던 약물 가운데 예상치 못한 심각한 부작용이 보고되고, 약물과의 인과성이 확인되어 시장에서 퇴출되

는 경우가 자주 발생하고 있다. 시사프라이드의 경우를 예로 들면, 임상에서 다년간 소화기 운동능력을 돕기 위하여 널리 사용되었지만, 비정형 심실성 빈맥(Torsade de pointes)으로 인한 사망사례가 보고되었고, 다른 약물과의 상호작용에 의한 약물유해반응임이 확인되면서 사용상 주의를 요하게 되었고 결국 시장에서 퇴출당했다. 이런 경우 해당 약물과 부작용에 대해 각국의 행정당국과 제약회

교신저자 : 박병주

소 속 : 서울대학교 의과대학 예방의학교실

주 소 : 서울특별시 종로구 연건동 28 (110-744)

전 화 : 02-740-8325, 팩스 : 02-747-4830 E-mail : bjpark@snu.ac.kr

사, 또는 연구자들 사이에 신속하고 정확한 정보교환과 국제적인 공조는 매우 중요하다. 각국의 행정당국이나 제약회사에서는 수집된 약물부작용사례들을 데이터베이스화하여, 부작용 발생빈도를 분석하고 자료검색과정을 통해 유용한 실마리정보(Signal)로 파악하고 있다. 자료의 활용도를 높이고 신속하고 정확한 정보교환을 위해 구조화된 데이터베이스 양식을 개발하여 자료를 입력하고 표준화된 코딩용어체계를 이용하게 된다. 약물 부작용 사례를 코딩하기 위해 널리 이용되는 용어체계로 COSTART(Coding Symbol for a Thesaurus of Adverse Reaction Terms), WHOART(World Health Organization Adverse Reaction Terminology), MedDRA(Medical Dictionary for Regulatory Activities) 등을 들 수 있다. 여러 문헌을 통해 이 세 가지 코딩용어체계를 개괄적으로 살펴보고, 그 중 국제적으로 표준화된 신약개발을 이끌고 있는 국제조화회의(International Conference on Harmonization, ICH)에 의해 개발되어 최근 표준적 코딩용어체계로 사용되고 있는 MedDRA의 특징을 중점적으로 살펴보고자 한다.

약물 안전성 자료의 코딩용어체계 비교(Table 1)

COSTART는 미국 식약청(Food and Drug Administration, FDA)에 의해 1969년 DART(Dictionary of Adverse Reaction Terms)로부터 개발되었고, 1995년 제5차 개정판 이후로 더 이상 갱신되지 않고 있다. 이 용어체계는 12개 Body System, 200여 개의 Mid level Term, 1,200여 개의 Preferred Term, 약 5,500여 개의 Glossary Term으로 구성되고, 가장 상위 수준인 Body System은 신체기관별로 분류된다¹⁾. 숫자가 아니라, 상징적 영어(coding symbol)로 코딩되는데, 예를 들면 abdominal pain은 ABDO PAIN으로 코딩된다¹⁾.

WHOART는 COSTART와 같은 해인 1969년 세계보건기구(World Health Organization, WHO)에서 만든 코딩사전으로 2003년 말까지 새로운 용어가 추가되면서 버전이 갱신되었다. 32개 System Organ Class, 180여 개 High Level Term, 2,000여 개의 Preferred Term과 약 3,000여 개의 Included Term로 구성되고, COSTART와 같이 4단계 계층구조를 갖고 있으나, 가장 낮은 단계인 Included

Table 1. 안전성 자료 코딩 용어체계 비교

	COSTART	WHOART	MedDRA
개발기관	FDA	WHO	ICH
갱신여부	1969년 개발 1995년 제5차 개정판 이후 갱신 없음.	1969년 개발 2003년 12월 3개월마다 용어 추가 ²⁾	1994년 개발 2005년 8월 version 8.1 6개월마다 갱신
관리기구		WHO collaborating center	MSSO
용어구조	Body System(12) Mid level Term(200) Preferred Term(1,200) Glossary Term(5,585)	System Organ Class(32) High Level Term(180) Preferred Term(2,000) Included Term(3,032)	System Organ Class (26) High Level Group Term (332) High Level Term(1,683) Preferred Term(16,976) Lowest Level Term(62,960)
비용	비교적 저렴	비교적 저렴 ³⁾ 각국 정부에 무료사용 허용	매년 고가의 사용료 부담 ⁴⁾

Term이 약 3,000개로 용어 수가 상대적으로 적은 편이다. 영어, 프랑스어, 독일어, 스페인어, 포르투갈어, 이탈리아어 등 6개국 언어로 출시되어 특정 언어와 관계없이 코딩할 수 있게 각 단계별 용어에 숫자 코드가 배정되어 있다²⁾. 비교적 가격이 저렴하고³⁾ 각 나라의 정부기관에는 무료로 배포되어 전세계적으로 널리 사용되고 있으며, WHO 협력센터인 읍살라 모니터링센터(Uppsala Monitoring Center, UMC)의 350만 자발적 유해사례보고 데이터베이스에 적용되고 있다.

1969년 WHOART와 COSTART가 개발된 이후 유럽연합에 의해 요구되는 유해사례 코딩에는 주로 WHOART가 이용되었고, 미국 FDA에 의해 요구되는 코딩에는 주로 COSTART가 이용되어 왔다.

MedDRA는 1994년 미국, 유럽, 일본의 제약업체와 행정당국으로 구성된 국제조화회의에서 개발하였다. 다른 용어들과 달리 5단계 계층구조를 갖고 있고, 가장 낮은 단계인 Lowest Level Term(LLT)에 6만 여 개 이상의 용어가 포함되어 여러 사람에 의해 코딩이 되더라도 좀 더 정확하고 일관되게 코딩될 수 있다. 하지만 매년 고가의 사용료를 지불하여야 한다⁴⁾.

연구자가 보고한 유해사례가 세 가지 용어체제로 각각 어떻게 코딩되는지 예를 들어 보겠다. 연구자가 ‘Diabetic neuropathy’를 유해사례로 보고한 경우, COSTART의 PT(Preferred Term)은 Neuropathy로, WHOART체계에서도 PT Neuropathy로 코딩되는데, MedDRA PT로는 연구자가 기재한 그대로 Diabetic neuropathy가 된다(Table 2). 이것은 MedDRA 체계가 셋 중 가장 용어가 풍부하다는 것을 보여주는 사례이다. 보고된 용어는 일단 코딩되면 코딩된대로 분석이 진행되므로 코딩의 정확성이 분석의 정확성에 큰 영향을 미치게 된다. 따라서 코딩자료는 원자료의 정보를 빼거나 더하지 않고 정확한 의미를 반영해야 하는데, 이런 측면에서 보면, 다른 사전에 비해 용어가 많은 MedDRA가 안전성 자료코딩 용어체제로 가장 적합하다고 하겠다.

MedDRA 개발과 이용 현황

MedDRA는 1994년 ICH에서 개발하였는데, 당시 영국의 자발적 부작용사례보고 데이터베이스에 적용되고 있으면서 가장 체계적이라고 평가되던 ADROIT(Adverse Drug Reactions Online Infor-

Table 2. Diabetic neuropathy에 대한 각 용어체계의 코딩 사례 비교

Verbatim : Diabetic neuropathy		
COSTART	WHOART	MedDRA
Glossary Term	Included Term	Lowest Level Term
Neur opathy	Neur opathy	Diabetic neuropathy
↓	↓	↓
Preferred Term	Preferred Term	Preferred Term
Neur opathy	Neur opathy	Neur opathy
↓	↓	↓
Body System	System Organ Class	System Organ Class
Nervous system	Central & peripheral nervous system	Nervous system disorders
		Endocrine disorders
		Metabolism & nutrition disorders

mation Tracking)용어로부터 개발되었다^{5),6)}. 이는 용어 수가 적어 다양한 부작용사례의 정확한 의미를 반영하지 못하던 기존 용어사전의 제한점을 극복한 체계라고 할 수 있다. 이후 MSSO(Maintenance and Support Services Organization)에서 용어체계를 관리하면서, MedDRA 사용자의 요구사항을 검토하고 의견을 반영하여, 용어체계를 지속적으로 갱신하고 있는데^{5),7)}, 2005년 9월 현재 버전 8.1이 출시된 상태이다.

MedDRA는 COSTART, WHOART, ICD9(International Classification of Disease Ninth Revision), ICD9-CM(Clinical Modification) 등을 포함하고 있어 각각 다른 용어체계로 맵핑(mapping)이 가능하고, 자료분석과 보고, 전자적 전송이 용이한 구조이다. 유해사례, 증상, 징후, 절차, 검사, 적응증, 의학적, 사회적 과거력 등의 코딩에 이용되나, 약물을 코딩하진 않는다^{5),7)}.

현재 임상시험을 수행하는 경우에 반드시 MedDRA를 사용하도록 요구하는 행정당국은 없지만, 유럽연합에서 수행되고 있는 모든 임상시험에서 발생한 약물부작용사례의 신속보고에 MedDRA를 사용하도록 지침이 제정된 상태이다⁸⁾. 일본에서는 2000년 4월 이후, 유럽연합에서는 2003년 1월 이후 약물유해반응(ADR) 사례보고서의 모든 전자문서 제출 시 MedDRA로 코딩하도록 행정당국에서 규정하고 있다⁹⁾. 그 외, 미국 FDA의 유해사례보고시스템데이터베이스(AERS database), 유럽의약품평가청(European Agency for the Evaluation of Medicinal Products, EMEA)의 EudraVigilance, 일본의 처방사례모니터링시스템(Prescription event monitoring system) 등에 MedDRA가 이용되고 있고, 다국적 제약업체와 각국 행정당국 사이에 의심되는 약물유해반응에 대한 정보의 전자교환을 위해 이용되고 있다^{8),10)}.

MedDRA 특징

MedDRA의 특징은 크게 세 가지로 요약할 수 있다. 첫째는 5가지 계층구조를 갖는다는 점, 두 번째는 다층적인 구조라는 점이고, 세 번째는 버전이 정기적으로 갱신된다는 점이다.

이 세 가지 MedDRA의 특징을 정리해보고, MedDRA로 약물부작용사례를 코딩할 때 고려해야 할 점에 대해 살펴보겠다.

1. 5단계 계층구조

MedDRA는 그림 1에서와 같이 5단계 계층구조를 갖는다. SOC가 가장 상위수준이고 LLT가 가장 하위수준의 용어이다. 제8.1차 개정판에는 26개의 SOC, 332개의 HLGT, 1,683개의 HLT, 16,976개의 PT 그리고 62,950개의 LLT가 포함되어 있다⁷⁾(Fig. 1).

임상시험에서 보고된 유해사례는 LLT로 매칭되어 코딩되지만, 보고서는 주로 PT와 SOC로 분류되어 분석된다. PT는 하나의 의학적 개념만을 표현하는 용어이다⁷⁾. LLT는 PT의 유사어들인데, 철자가 다르거나 구어적 측면에서의 유사어들이고, 자료를 입력할 때 적용된다. 각각의 LLT들은 하나

Fig. 1. MedDRA의 5단계 계층구조⁵⁾.

Fig. 2. Tension headache에 대한 MedDRA PT와 LLT 관계의 예.

의 PT에 연결되고, 모든 PT는 LLT와 겹친다. Fig. 2에서, 유사어들인 LLT Band-like headache, LLT Head tightness, LLT Headache tension, LLT Tension headaches, LLT Tight band round head 등이 PT Tension headache로 연결된 관계를 보여주고 있다(Fig. 2).

HLT는 PT 수준에서 해부학적, 병리학적, 병인론적, 또는 기능적으로 유사하게 분류된 용어이고,

하나의 HLT는 여러 HLGT에 연결된다. HLGT는 HLT의 상위 수준으로, 하나의 HLGT는 여러 SOC에 연결된다. 5단계 중에 최상위 수준인 SOC는 대략적으로는 ‘신체기관별’ 분류를 따르나, 그렇지 않은 Investigation, Injury, poisoning and procedural complications, Surgical & Medical Procedures, Social circumstance 등도 SOC 목록에 포함된다⁷⁾(Table 3).

2. 다축성 구조(Multiaxiality)

MedDRA의 구조는 다축적이어서 하나의 유해사례는 여러 SOC로 동시에 코딩될 수 있다⁷⁾. 용어가 갖고 있는 개념에 따라 여러 축으로 연결되는 특성은 LLT수준을 제외한 모든 수준(PT, HLT, HLGT)에서 일어나는데 PT와 HLT 사이에서 가장 많이 일어난다. 각 LLT가 하나의 PT에 연결되는 반면, 하나의 PT는 여러 HLT에 연결될 수 있고, 하나의 HLT도 여러 HLGT에 연결될 수 있으며, 하나의 HLGT도 또한 여러 SOC에 연결될 수 있다(Fig. 3).

Table 3. MedDRA의 System Organ Class(SOC) 목록⁵⁾

Infections and infestations	Gastrointestinal disorders
Neoplasms benign, malignant and unspecified	Hepatobiliary disorders
Blood and lymphatic system disorders	Skin and subcutaneous tissue disorders
Immune system disorders	Musculoskeletal and connective tissue disorders
Endocrine disorders	Renal and urinary disorders
Metabolism and nutrition disorders	Pregnancy, puerperium and perinatal conditions
Psychiatric disorders	Reproductive system and breast disorders
Nervous system disorders	Congenital, familial and genetic disorders
Eye disorders	General disorders and administrations site conditions
Ear and labyrinth disorders	Investigations
Cardiac disorders	Injury, poisoning and procedural complications
Vascular disorders	Surgical and medical procedures
Respiratory, thoracic and mediastinal disorders	Social circumstances

Fig. 3. MedDRA 구조의 다축적 관계를 나타내는 그림.

결국 하나의 PT가 여러 SOC에 연결될 수 있다. 예를 들어 PT Pneumonia는 SOC Infections and infestations와 SOC Respiratory, thoracic and mediastinal disorders 양쪽에 속하게 된다. 이는 MedDRA가 의학적 정보를 고려한 코딩용어사전임을 보여주는 것이다(Fig. 4).

PT가 여러 SOC로 동시에 연결될 수 있으므로 SOC 수준에서 정보를 검색하는 경우 PT가 중복 계산되는 것을 막기 위해 각각의 PT에는 일차적 SOC가 미리 정해져 있다. PT Pneumonia의 일차적 SOC는 Infections and infestations이다. Med-

DRA에서 모든 SOC가 다축성을 갖는 것은 아니다. Investigation SOC, Surgical & Medical Procedures SOC, Social circumstance SOC는 단일축인 SOC이다.

3. 정기적인 버전의 갱신

MSSO에서는 최신 의학지식이 반영되도록 MedDRA 용어를 관리하며, 6개월마다 MedDRA를 갱신하고 있다. MedDRA 사용자들의 요구사항을 검토하여, 새로운 용어를 추가하거나, 현재 상황에 맞는 용어를 선택하거나, 일차적 SOC를 변경하거나, 계층간 연결을 변화시킬 필요가 있을 때 그러한 내용을 반영하여 버전을 갱신한다. 버전이 갱신되면 같은 유해사례에 대한 코딩이 변할 수 있다. 따라서, 장기간 지속될 연구에서 MedDRA로 코딩을 할 때는 어느 버전을 사용할 지 통계분석계획에서 밝혀야 하고⁹⁾, 연구를 시작할 때부터 종료할 때까지 같은 버전으로 코딩하여야 한다. 가장 최신 버전의 MedDRA를 이용해서 부작용 용어를 코딩해야 하지만, 이전의 분석결과를 그대로 재현시켜 다른 연구결과와 비교하기 위해서는 데이터베이스에는 적어도 연구자에 의해 보고된 유해사례와 LLT가 저장되어야 하고 그 때 사용한 버전이 무엇이었는지 밝혀두어야 한다⁹⁾.

Fig. 4. Pneumonia에 대한 MedDRA PT와 HLT 사이의 다축적 연결 사례.

Fig. 5. 유해사례가 실험실적 결과인 경우 MedDRA LLT선택에 따라 달라지는 SOC의 예.

MedDRA 사용 연구에서 고려할 점

약물 안전성 자료를 MedDRA로 코딩하는 연구에서 고려해야 할 점은 일관성있는 코딩이 되어야 한다는 것이다. 유해사례는 연구자가 보고한 용어와 가장 가까운 LLT 용어로 코딩되지만 분석과 보고는 PT나 SOC에 따라 분류하게 된다¹¹⁾. 연구자가 보고한 처음 용어대로 코딩하므로 보고서에 기재된 유해사례는 정확해야 한다. 여러 코딩자에 의해 코딩되더라도 일관된 코딩을 위해 ICH에서 용어선택지침을 제시하고 있는데¹²⁾, 그 중 하나의 코딩사례를 보겠다. 연구자가 Hypoglycemia로 기재하였다면 LLT Hypoglycemia로 코딩되고, Glucose decreased로 기재한 경우는 LLT Glucose decreased가 선택되어야 한다¹²⁾. 이 경우 LLT Hypoglycemia는 PT Hypoglycemia로, 일차적 SOC는 Metabolism and nutrition disorders로 분류되지만, LLT Glucose decreased는 PT Blood glucose decreased로, SOC는 단축인 Investigations로 분류되어 해당되는 질병 SOC로는 발견되지 않게 된다 (Fig. 5).

흔히 항암제 연구에서 유해사례의 중증도 평가

를 위한 NCI-CTC(National Cancer Institute Common Toxicity Criteria)의 내용을 그대로 유해사례에 기재하는 경우가 있다. 예를 들면 Leukocyte, Platelet 등인데, 이와 같이 검사명으로 기재되는 경우 일차적 SOC는 Investigations으로 코딩된다. Leucopenia나 Thrombocytopenia로 기재되었다면 일차적 SOC는 Blood & lymphatic system disorders로 된다. 다기관 연구에서 연구자마다 Leukocyte와 Leukopenia 중 하나를 각기 다르게 기재하였다면, 연구종료 후 유해사례 빈도분석 결과에서 같은 유해사례가 다른 SOC로 나뉘게 된다. 그러므로 연구를 시작할 때 같은 유해사례가 다르게 코딩되어 빈도분석 결과가 실제보다 낮게 추정되지 않도록 표준화된 용어선택지침을 정하는 것이 필요하다.

일관된 코딩을 위해 고려해야 할 또 다른 점은, MedDRA의 경우 PT가 다른 용어체계에 비해 수도 많고, 상당히 구체적이어서 유사한 유해사례에 대해 일관성있는 코딩이 잘 안되기도 한다는 점이다⁸⁾. 즉 PT가 너무 세분화되어 단어만 약간 달라져도 다른 SOC에 속하게 되는데, 예를 들면 *Clostridium difficile*에 의한 Diarrhea가 보고된 경우

Fig. 6. PT가 너무 세분화되어 MedDRA LLT선택에 따라 달라지는 SOC의 예.

LLT *C. difficile* diarrhea가 선택되면 PT *Clostridium colitis*, 일차적 SOC로 Infections and infestations로 분류되나, LLT diarrhea로 코딩되면 PT Diarrhea, 일차적 SOC는 Gastrointestinal disorders로 분류된다¹¹⁾. 그러므로 유해사례를 기재할 때는 그 임상적 의미를 정확히 담고 있는 용어로 기재하는 것이 필요하다(Fig. 6).

맺 음 말

우리나라의 약물부작용 보고건수는 1988년 자발적 부작용모니터링제도가 시행된 이후 연간 수 백건 이하로 매우 저조하다가, 2004년에는 907건으로 증가하였으나, 1만 5천~2만 건인 일본이나, 2만~3만 건인 유럽연합, 43만 건인 미국의 보고건수에 비해 현저히 적고, 약물부작용 모니터링을 위한 전담조직이나 인력이 확보되지 않아 자료가 실효성 있게 관리되지 못하고 있다¹³⁾. 그러나 국내에서도 국가 정책적으로 내설있는 데이터베이스 관리와 약물부작용정보에 대한 실질적 이용의 필요성이 제시되고¹³⁾, 전담조직을 구축할 계획을 갖고 있어, 머지않아 국내의 약물부작용사례관리를 위한 체계

적인 조직이 갖추어지고 이어서 효과적인 데이터베이스 관리체계도 정비되리라 생각된다. 약물부작용정보가 적절히 활용될 수 있도록 데이터베이스를 구축하기 위해서는 표준화되고 구조화된 자료 구축체계를 신속히 갖추어야 한다. MedDRA는 사용자 그룹의 의견을 반영하여 부단히 변화, 발전해왔고, 폭넓게 이용되며 국제적 표준용어체계로 자리잡아 가고 있다. 이런 MedDRA 용어체계에 대한 이해를 바탕으로 유해사례나 약물유해반응을 보고하고 관리할 때 좀 더 유용한 데이터베이스를 구축하고 효과적으로 활용할 수 있게 될 것이다.

참 고 문 헌

1. Andras K, Lisa S. *Automatic Translation to Controlled Medical Vocabularies 7-9*. Available From: URL:http://www.metacarta.com/docs/Kornai_stone.pdf
2. <http://www.unc-products.com/graphics/3149.pdf> [Accessed 2005 Nov 01]
3. <http://www.unc-products.com/graphics/2472.pdf> [Accessed 2005 Nov 01]
4. <http://www.meddransso.com/NewWeb2003/s>

- subscriptions/rates.htm[Accessed 2005 Nov 01]
5. Medical Dictionary for Regulatory Activities (MedDRA) Introductory Guide MedDRA Version 8.1. 2005;1-79
6. Brown EG. *Effects of Coding Dictionary on Signal Generation: A Consideration of Use of MedDRA Compared with WHO-ART*. *Drug Safety* 2002;25:445-452
7. Brown EG, Wood L, Wood S. *The Medical Dictionary for Regulatory Activities (MedDRA)*. *Drug Safety* 1999;20:109-117
8. Brown EG. *Using MedDRA: Implications for Risk Management*, *Drug Safety* 2004;27:591-602
9. Jurgen K, Richardus V, Stefan B, Winfried G, Martin H, Detlef N, Jurgen K, Karlheinz T, Peter V. *Adverse Event Analysis and MedDRA: Business as Usual or Challenge?* *Pharmacovigilance & Safety Drug Information J* 2005;39:63-72
10. Bousquet C, Lagier G, Lillo-Le LA, Le Beller C, Venot A, Jaulent MC. *Appraisal of the MedDRA Conceptual Structure for Describing and Grouping Adverse Drug Reactions*. *Drug Safety* 2005;28:19-34
11. Fescharek R, Kubler J, Elsasser U, Frank M, Guthlein P. *Medical Dictionary for Regulatory Activities (MedDRA): Data Retrieval and Presentation*. *Int J Pharm Med* 2004; 18:259-269
12. *MedDRA Term Selection: Point to consider. Release 3.4 based on MedDRA version 7.1 Available From: URL: http://www.ich.org/MediaServer.jsr?@_ID=1714&@_MODE=GLB*
13. 정영철. *의약품 안전정보 관리방안*. 보건복지포럼 2005;5:113-135 Available From: URL: <http://www.kihasa.re.kr/book/book01-9-1.jsp>