**Annotation Guideline for Texts from KAERS (Korea Adverse Event Reporting System)**

Siun Kim, Tae Kyu Chung, Yoona Choi, YeSol Hong, and Howard Lee

Center for Convergence Approaches in Drug Development, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, Korea

(Last updated on July 20, 2021)

**Contents**

1. **Introduction**
2. **Entities**
   1. Pathological Finding
      1. Adverse Event (AdverseEvent)
      2. Disease or Indication (DiseaseOrIndication)
      3. Seriousness of Adverse Event (AESeriousness)
      4. Adverse Event at the Time of Last Observation (AEatLastObs)
   2. Drug
      1. Drug Compound (DrugCompound)
      2. Drug Product (DrugProduct)
      3. Drug Group (DrugGroup)
   3. Dosing Information
      1. Dose (DrugDose)
      2. Dosing Interval (DrugDosingInterval)
      3. Route of Administration or Drug Formulation (DrugRoAFormulation)
   4. Date
      1. DateStartorContinue, DateEnd
      2. Period (DatePeriod)
   5. Patient Information (PatientInfo)
      1. Patient Sex (PatientSex)
      2. Patient Age (PatientAge)
   6. Others
      1. Admission (EventAdmission) & Discharge (EventDischarge)
      2. Test Name (TestName), Test Result (TestResult)
      3. Non-Drug Treatment (NonDrugTreatment)
      4. Action Taken with Drug (ActionTakenwDrug)
      5. WHO-UMC Results of Assessment (CasualityAssessment)
3. **Relations**
   1. AdverseEvent ↔ AdverseEvent / DiseaseOrIndication ↔ DiseaseOrIndication
   2. AdverseEvent ↔ AESeriousness, AEatLastObs
   3. DiseaseOrIndication ↔ DrugCompound, DrugProduct, DrugGroup
   4. AdverseEvent, DiseaseOrIndication, AESeriousness, EventatLastObs, DrugCompound, DrugProduct, DrugGroup, EventAdmission, EventDischarge, TestName, NonDrugTreatment, ActionTakenwDrug ↔ DateStart, DateEnd, DateNotEnd, DatePeriod
   5. DateStart, DateEnd, DateNotEnd ↔ DatePeriod
   6. DrugProduct ↔ DrugCompound
   7. DrugCompound, DrugProduct, DrugGroup ↔ DrugDose, DrugDosingInterval, DrugRoAFormulation
   8. TestName ↔ TestResult
   9. ActionTakenwDrug ↔ DrugCompound, DrugProduct, DrugGroup
   10. CasualityAssessment ↔ AdverseEvent, DiseaseOrIndication, DrugCompound, DrugProduct
4. **Entity Lables**
   1. Pathological Finding
      1. adverse event identifier & indication identifier (MedDRA)
      2. not\_occurred
      3. seriousness
      4. event\_at\_last\_observation
      5. startcontinue\_or\_end
   2. DrugCompoud, DrugProduct & DrugGroup
      1. compound identifier & product identifier
      2. not\_occurred (not\_administered)
      3. not\_concerned
      4. startcontinue\_or\_end
   3. Others
      1. test name identifier (MedDRA)
      2. non-drug treatment (MedDRA)
      3. action taken with drug (E2BR3)
      4. WHO-UMC results of assessment (WHO-UMC)
5. **Introduction**

본 어노테이션 가이드라인은 한국의약품안전관리원에서 운영하는 KAERS 이상사례 보고시스템에서 수집한 자연어 서술자료를 일관되게 구조화할 수 있도록 어노테이션 작업을 명확히 정의하고 어노테이션 지침을 제공하는 것을 목표로 한다. 해당 어노테이션 가이드라인에 따라 만들어진 구조화된 자연어 데이터는 KAERS 내 자연어 서술자료에서 의약품 안전성 평가에 필요한 주요 임상정보를 추출하는 자연어처리 모델을 학습하는데 사용될 예정이다.

1. **Entities**

해당 절에서는 KAERS 자연어 서술자료를 구조화하는데 사용되는 개체(이하 entity)를 정의하며, 해당 entity을 일관되게 어노테이션하기 위한 규칙과 참고사항을 제공한다.

* 1. Pathological Finding
     1. Adverse Event (AdverseEvent)

이상사례 상세서술 내에서 언급된 구분가능한 **이상사례**를 지칭한다. 이때, 이상사례란 **의심약제 투약 이후 발생한 모든 증상 내지 진단된 질병**을 의미한다. AdverseEvent 개체는 **단어 수준**으로만 지정되며, 수식어의 수식을 받는 경우 수식어(e.g., chronic, acute)를 최대한 포함하여 지정한다. 환자에게 실제로 발생하거나 진단되었는지와 상관없이 이상사례 상세서술에서 언급된 이상사례명이면 모두 AdverseEvent 개체로 지정되어야 한다.

* 주의할 이상반응: 아래의 경우에도 AdverseEvent로 지정하고, 해당 MedDRA 코드로 labeling 해야 한다.

‘효과없음/Lack of efficacy’(MedDRA 10014291): 의약품을 투약한 이후 기대한 효과 및 효능이 나타나지 않았다고 기술한 경우 (

‘제품 품질 문제/Product quality issue’(MedDRA 10069327): '약품 품질 문제', 'Product physical issue', '제품/품질 관련 결함' 등을 표현한 경우

* + 1. Disease or Indication (DiseaseOrIndication)

이상사례 상세서술 내에서 언급된 구분가능한 **과거질병명**을 지칭한다. 이때, 과거질병(Disease)이란 **의심약제 투약 이전에 발생한 모든 증상 내지 진단된 질병**을 의미한다. 이중 투약한 의약품의 치료 혹은 예방 대상인 질병을 적응증(Indication)이라 지칭한다. 하지만 해당 어노테이션 가이드라인에서는 적응증을 과거질병과 구분하지 않고 DiseaseOrIndication로 어노테이션한다. 다만 Indication 개체로 인식하여 지정한 경우에는 Drug 개체와 relation을 형성해야 한다. DiseaseOrIndication 개체도 AdverseEvent와 마찬가지로 **단어 수준**으로만 지정되며, 수식어의 수식을 받는 경우 수식어(e.g., Type II diabetes)를 최대한 포함하여 지정한다. 환자에게 실제로 발생하거나 진단되었는지와 상관없이 이상사례 상세서술에서 언급된 질병명이면 모두 DiseaseOrIndication 개체로 지정되어야 한다. 그러나 실제 환자에게 발생한 질병이나 증상이 아니면 의심약제 투약 이전/이후에 관찰되었는지 판단할 수 없으므로 일관되게 AdverseEvent로 지정하도록 한다.

* + 1. Seriousness of Adverse Event (AESeriousness)

이상사례 상세서술 내에서 언급된 환자가 경험한 **이상사례의 중대성**을 의미한다. 이때, 이상사례의 중대성이란 E2B R3 서식 내 의약품 중대성 정의에 따라 **이상사례에 의한 1) 사망, 2) 생명의 위협, 3) 입원 또는 입원기간의 연장, 4) 중대한 불구나 기능저하, 5) 선천적 기형 초래, 6) 분류되지 않은 의학적으로 중대한 사건, 7) 중대하지 않은 상황** 중 하나의 상황을 지칭한다. AESeriousness 개체는 이상사례 중대성을 서술한 **문장 단위**로 지정한다. 문장으로 서술되지 않는 경우에 한해 해당 이상사례 중대성을 설명한 명사구이나 단어 단위로 지정한다. 가령, ‘Gastroenteritis 발생하여 입원하였다’와 같이 이상사례로 인해 입원이 초래된 것이 분명한 경우 ‘입원’을 AESeriouness 개체와 EventAdmission 개체로 모두 지정한다.

* + 1. Event at the Time of Last Observation (AEatLastObs)

이상사례 상세서술 내에서 언급된 **최종 관찰 시점에서의 이상사례 지속 및 회복 여부**를 의미한다. 이때, 최종 관찰 시점에서의 이상사례 지속 및 회복 여부란 E2B R3 서식 내 최종 관찰 결과 정의에 따라 **이상사례의 1) 회복됨, 2) 회복중, 3) 회복되지 않았음, 4) 회복했지만 후유증 있음** 중 하나의 상황을 지칭한다. AEatLastOsb 개체는 이상사례의 최종 관찰 결과를 의미하는 **명사구나 단어**로 지정한다. 단, 호전(회복) 후 재발한 경우, ‘호전(회복)’이 최종 결과가 아니라 하더라도, ‘호전(회복)’과 ‘재발’을 나타내는 명사구나 단어를 각각 AEatLastObs로 지정한다.

* 1. Drug

이상사례 상세서술 내에서 언급된 구분가능한 **약물명**을 총칭한다. 해당 약물이 성분명으로 서술된 경우 DrugCompound로, 의약품명(상품명)으로 서술된 경우 DrugProduct로, 구체적인 성분명이나 상품명을 언급하지 않고 치료목적을 공유하거나 기타의 이유로 묶이는 약물군으로 서술된 경우 DrugGroup으로 어노테이션한다.

* + 1. DrugCompound

이상사례 상세서술에서 언급된 구분가능한 **성분명 형태로 서술된 약물**을 지칭한다. DrugCompound 개체는 **단어 수준**으로만 지정한다. 본 어노테이션 가이드라인에서는 약물의 용량, 투약 간격, 제형, 투여 경로는 별도의 개체로 지정하기 때문에 해당 내용을 나타내는 어구나 단어는 DrugCompound로 지정하지 않는다. ‘의심 약물’이나 ‘복용 약물’ 등 구체적인 약물 정보가 전달하지 않는 경우 DrugCompound로 지정하지 않는다. 특정 의약품명이 성분코드집과 의약품코드집 모두에 기재된 경우 DrugCompound로 지정하는 것을 원칙으로 한다.

* + 1. DrugProduct

이상사례 상세서술에서 언급된 구분가능한 **의약품명(상품명)으로 서술된 약물명**을 지칭한다. DrugProduct 개체는 **단어 수준**으로만 지정한다. 본 어노테이션 가이드라인에서는 약물의 용량, 투약 간격, 제형, 투여 경로는 별도의 개체로 지정하기 때문에 해당 내용을 나타내는 어구나 단어는 DrugProduct로 지정하지 않는다. 가령, 이상사례 상세서술에서 ‘가스티딘정150mg’, ‘그라나텍점안액0.4%’이라 표현된 경우 ‘가스티딘정’, ‘그라나텍점안액’만을 DrugProduct로 지정하고 ‘150mg’, ‘0.4%’는 DrugDose 개체로 지정한다. ‘의심 약물’이나 ‘복용 약물’ 등 구체적인 약물 정보가 전달하지 않는 경우 DrugProduct로 지정하지 않는다. 만약, 용량 단위 표시 없이 의약품명(상품명)의 일부로 숫자가 표시된 경우에는 해당 숫자를 포함하여 entity로 지정하고, 용량은 따로 지정하지 않는다.

의약품명이 성분코드집과 의약품코드집에 등록되지 않은 개별 한방병원 또는 한의원 조제 한약의 경우에도 DrugProduct로 지정한다.

* + 1. DrugGroup

이상사례 상세서술에서 언급된 구분가능한 **치료목적을 공유하거나 기타의 이유로 묶이는 약물군으로 서술된 약물명**(e.g. 당뇨약, 혈압약, 항히스타민제)을 지칭한다. 본 어노테이션 가이드라인에서는 약물의 용량, 투약 간격, 제형, 투여 경로는 별도의 개체로 지정하기 때문에 해당 내용을 나타내는 어구나 단어는 DrugProduct로 지정하지 않는다. DrugGroup 개체는 **단어 수준**으로만 지정한다. ‘의심 약물’이나 ‘복용 약물’ 등 구체적인 약물 정보가 전달하지 않는 경우 DrugGroup로 지정하지 않는다.

* + 1. 기타: 항암제 병합요법

항암제의 병합요법이 통용되는 이름(주로 머리글자의 조합)으로 기술된 경우, 가능한 개별 성분명에 해당하는 머리글자를 글자단위로 따로 DrugCompound로 지정한다. (e.g. COPADM: C/O/P/AD/M - Cyclophosphamide, Oncovin (Vincristine), Prednisone, ADriamycin (Doxorubicin), Methotrexate)

* 1. Dosing Information

의약품 안전성 평가를 수행하기 위해서는 투약 약물 정보뿐 아니라 상세한 투약 정보가 필요하다. 이를 위해 KAERS 자연어 상세서술 어노테이션에서는 투약 용량(Dose), 투약 간격(Dosing Interval), 투여 경로(Route of Administration), 약물 제형(Drug Formulation)을 구조화한다.

* + 1. Dose (DrugDose)

이상사례 상세서술에서 언급된 구분가능한 **투약 용량**을 지칭한다. 이때, 투약 용량은 **복용량 내 유효성분 용량(e.g. ‘10mg’, ‘1g’), 주사제나 점안액 등의 농도(e.g. ‘5%’, ‘50g/L’), 복용횟수나 복용 사이클수(e.g. ‘1회’, ‘2 cycles’), 복용기간(e.g. ‘1주일간’, ‘1년’)**을 포함한다. DrugDose 개체는 **단어 단위**로 지정하며, 숫자로 표현되는 용량과 용량 단위를 모두 포함하여 지정한다. 의약품 개체를 지정하는 방식과 유사하게 실제 투여한 용량이 아니더라도 용량에 관한 단어나 명사구(e.g. ‘500cc 중 480cc 버림’ 중 ‘500cc’, ‘480cc’)를 포함하여 지정한다.

단, 단위가 명시되지 않고 숫자만 기술된 경우, 해당 숫자를 DrugDose로 지정하지 않는다. (e.g. 상품명의 일부로 용량 단위 없이 숫자만 포함된 경우, 한약의 배합비율)

* + 1. Dosing Interval (DrugDosingInterval)

이상사례 상세서술에서 언급된 구분가능한 **투약 간격**을 지칭한다. 이때, 투약 간격은 **기간 내 복용횟수(e.g. ‘1일 2회’), 투약 간격(e.g. ‘6시간마다’), 투약 간격을 표현하는 어구(e.g. ‘once a daily’)**를 포함한다. DrugDosingInterval 개체는 **단어 단위**로 지정하며, 숫자로 표현되는 투여 간격과 시간 단위(ex. ‘시간’, ‘일’, ‘주’)를 모두 포함하여 지정한다.

* + 1. Route of Administration or Drug Formulation (DrugRoAFormulation)

이상사례 상세서술에서 언급된 구분가능한 **투약 경로나 약물제형**을 지칭한다. 이때, 투여 경로는 **경구 투여, 주사 투여 방법(e.g. ‘iv’, ‘intramuscular’), 기타 투여방법 (e.g. ‘흡입’, ‘topical administration’) 등**을 포함한다. 약물 제형은 **경구용 제형의 세부 종류(e.g. ‘서방형’, ‘SR’, ‘캡슐’), 도포제의 세부 종류(e.g. ‘연고’, ‘크림’, ‘로션’, ‘젤/겔’) 등**을 포함하며, 투여 경로나 약물 제형이 별도의 어절(띄어쓰기 단위)로 표현된 경우에만 DrugRoAFormulation로 지정한다. 만약, 제형이 의약품명(상품명)의 일부로 의약품을 나타내는 단어와 띄어쓰기 없이 기술된 경우에는 해당 제형을 별도의 entity로 지정하지 않고, 제형을 나타내는 단어를 포함한 의약품명(상품명) 전체를 DrugProduct로 지정한다. (e.g. ‘Meropen inj’ 또는 ‘메로펜 주사’는 ‘Meropen’과 ‘메로펜’을 DrugProduct로, ‘inj’와 ‘주사’는 DrugRoAFormulation로 지정하고, ‘메로펜주’의 경우, ‘메로펜주’ 전체를 DrugProduct로 지정한다.)

단, 추론을 통해 유추할 수 있는 단어(e.g. ‘접종’)는 DrugRoAFormulation 로 지정하지 않는다. 또, ‘터부헬러’, ‘오로스(정)’와 같이 특정 제품에만 한정된 제형은 DrugRoAFormulation로 지정하지 않는다. (이 경우에도 띄어쓰기 없이 제품명의 일부로 기술된 경우에는 해당 어절 전체를 DrugProduct로 지정한다.)

* 1. Date

이상사례 상세서술 내에서 언급된 구분가능한 **날짜 및 시간 정보**를 총칭한다. 이때, 날짜 및 시간 정보는 **년월일 단위의 날짜와 시분 단위의 시간 정보**를 의미한다. Date 개체는 **단어 또는 명사구 단위**로 지정하며, ‘6/25일’, ‘2016년 12월’ 등 날짜를 나타내는 단어나 구둣점을 모두 포함하여 설정한다. 단 ‘6월 26일 오전 10시’처럼 표현되는 경우 날짜 정보와 시간 정보를 구분하여서 ‘6월 26일’과 ‘오전 10시’ 두 개의 entity로 나누어 설정한다. 또한 ‘다음날’, ‘같은날’, ‘알 수 없는 날’ 등의 표현을 포함한다. 해당 날짜 및 시간에 발생한 의학적 사건에 따라서 DateStartorContinue, DateEnd, DatePeriod로 나눈다.

* + 1. DateStartorContinue, DateEnd

DateStartorContinue, DateEnd 개체는 순서대로 **약물 투약이나 이상사례 발현 등이 시작/지속되거나, 종료한 날짜 및 시간을** 지칭한다. 가령, ‘5월 25일 영양 보급을 위해 smoflipid 20%를 10시경 부터 주입 시작함’이라고 서술된 경우 ‘5월 25일’, ‘10시경’부터 약물 주입이 시작되었기 때문에 두 표현을 DateStartorContinue 개체로 지정한다. 반면, ‘해당일 저녁 18시경 clopidogrel 복용을 중단함’이라고 표현된 경우 ‘저녁 18시경’ clopidogrel 복용은 중단(종료)했기 때문에 DateEnd 개체로 지정해야 한다. 약물 투약이나 이상사례 발현이 시작, 지속, 종료되지 않은 날짜 및 시간 정보더라도 해당 정보를 담은 단어나 명사구는 단어는 DateStartorContinue 개체로 지정한다.

* + 1. DatePeriod

이상사례 상세서술 내에서 언급된 구분가능한 **이상반응, 과거 질병, 또는의약품 투약 등의 지속기간이나 특정 날짜나 시점 이후 지난 시간이나 날짜(e.g. ‘2시간 후’, ‘이틀 이후’)**를 지칭한다. 다만 ‘1월 2일부터 3일까지’와 같이 날짜 정보를 포함하는 기간의 경우에는 DatePeriod을 사용하지 않고 이외 Date 개체를 사용하여 ‘1월 2일’과 ‘3일’로 나누어서 설정한다.

* 1. Patient Information (PatientInfo)

이상사례 상세서술 내에서 언급된 구분가능한 **환자의 기본 정보**를 지칭한다. 본 가이드라인에서는 **환자의 성별, 나이**를 구조화한다.

* + 1. Patient Sex (PatientSex)

이상사례 상세서술 내에서 언급된 구분가능한 **환자의 성별**을 지칭하며, 성별을 나타내는 **명사구나 단어(e.g. ‘여성’, ‘남성’, ‘M’, ‘F’) 단위**로 지정한다. 명확하게 성별을 나타내는 경우에만 사용하며 의학적 추론을 통해서 환자의 성별을 유추할 수 있는 단어(e.g., ‘난소암’)는 포함하지 않는다.

* + 1. Patient Age (PatientAge)

이상사례 상세서술 내에서 언급된 구분가능한 **환자의 나이와 출생년도, 생년월일**을 지칭하며, 나이를 지칭하는 **명사구나 단어(e.g. ‘생후 8개월’, ‘25세’, ‘1968년생’, ‘1988년 7월 2일 출생’) 단위**로 지정한다. 범위의 형태로 나이를 나타내는 경우(e.g. ‘30대’)에도 지정하나, 나이를 추정할 수 있게 하는 단어(e.g. ‘청소년’, ‘노인’)나 질환명(e.g. ‘신생아 당뇨병’)에는 사용하지 않는다.

* 1. Others

해당 절에서는 지금까지 언급된 정보 외 이상사례 상세서술 내에서 언급된 구분가능한 주요 의약품 안전성 정보와 관련된 개체를 정의한다.

* + 1. Test Name (TestName), Test Result (TestResult)

이상사례 상세서술 내에서 언급된 구분가능한 **검사명(e.g. ‘심전도 검사’)과 검사결과(e.g. ‘정상’, ‘양성’, ‘127mmHg’, ‘병변이 확인되지 않음’)**를 지칭한다. 정확한 검사명이 언급되지 않더라도 수행된 검사의 종류를 추측할 수 있을만한 단어가 서술된 경우 이를 TestName으로 지정한다. 가령 ‘BP가 90에서 120’, ‘체온이 36.8’이라고 표현된 경우 TestName은 ‘BP’, ‘체온’으로 설정하며 TestResultsms ‘90에서 120’, ’36.8’로 설정한다. 단, ‘고열’ 또는 ‘Fever’ 등과 같이 vital sign과 관련된 병리학적 소견을 나타내는 단어와 함께 구체적인 vital sign에 해당하는 숫자가 표시된 경우(e.g. 39도의 고열), 해당 단어를 TestName으로, 숫자를 TestResult로 지정해야 하며, 동시에 해당 단어는 AdverseEvent로도 지정해야 한다. 단, 구체적인 검사 수치 자체는 TestName and/or TestResult로만 지정하고 AdverseEvent로 지정하지 않는다. 감소, 증가 등 변화추세를 명료하게 언급한 경우에는 맥락에 따라 AdverseEvent로 지정할 수 있다. (e.g. '호산구 28.2%' -> AdverseEvent로 지정하지 않음, '호산구 감소' -> TestName, TestResult로 지정하고, 맥락에 따라 AdverseEvent로도 지정함)

TestName 개체는 **단어 단위**로 지정하며, TestResult 개체는 기본적으로 **검사 수치와 검사 단위를 포함하여 명사구나 단어 단위**로 지정하되 검사결과가 기술적(descriptive)으로 표현되는 경우 (e.g. ‘(내시경 검사상) active bleeding은 현재 없는 상태’) **문장이나 절 단위**로 지정한다.

검사 결과의 참고치(정상 범위)는 TestResult로 지정하되, TestName과 relation은 지정하지 않는다.

* + 1. Non-Drug Treatment (NonDrugTreatment)

이상사례 상세서술 내에서 언급된 구분가능한 **비약제처치(e.g., ‘oral hydration’)**를 지칭한다. NonDrugTreatment 개체는 **단어 단위**로 지정한다. 단, 질환의 진단 수단으로서 병리학적 상태의 확인을 목적으로 시행된 처치는 NonDrugTreatment가 아닌 TestName으로 지정한다. (e.g. Computed Tomography or CT)

* + 1. Action Taken with Drug (ActionTakenwDrug)

이상사례 상세서술 내에서 언급된 구분가능한 **의약품에 관한 조치 내용**을 지칭한다. 이때, 의약품에 관한 조치 내용이란 2B R3 서식 내 정의에 따라 **1) 의약품 투여중지, 2) 의약품 투여량 감소 (용량 감소 혹은 투약 간격 늘림), 3) 의약품 투여량 증가 (용량 증가 혹은 투약 간격 줄임), 4) 투여량 변경하지 않음, 5) 모름** 중 하나의 상황을 지칭한다. ActionTakenwDrug 개체는 의약품에 관한 조치 내용을 서술한 **명사구나 단어 단위**로 지정한다.

* + 1. WHO-UMC Results of Assessment (CasualityAssessment)

이상사례 상세서술 내에서 언급된 구분가능한 **WHO-UMC 토대의 의약품과 이상사례 간 인과성 평가결과**를 지칭한다. 이때, 의약품과 이상사례 간 인과성 평가란 WHO-UMC 내 정의에 따라 **1) 확실함(certain), 2) 상당히 확실함(probable), 3) 가능함 (possible), 4) 가능성 적음 (unlikely), 5) 평가곤란(conditional/unclassified), 6) 평가불가(unaccessible/unclassified)** 중 하나의 상황을 지칭한다(각 상황에 대한 상세한 정의와 설명은 CasualityAssessment의 entity label 항목을 참고).

기본적으로 CausalityAssessment는 의약품과 이상사례 인과성을 평가한 **문장 단위**로 설정하되, 한 문장 내에서 두 가지 다른 인과성 평가가 내려진 경우에는 **적절한 절이나 명사구의 형태**로 나누어서 설정한다.

1. **Relations**

해당 절에서는 앞서 정의한 개체 간 의미를 명료하게 설정하기 위한 개체 간 관계(relation)를 정의하며, 해당 relation을 일관되게 어노테이션하기 위한 규칙과 참고사항을 제공한다.

* 1. AdverseEvent ↔ AdverseEvent / DiseaseOrIndication ↔ DiseaseOrIndication

**동일한 이상반응을 두 개 이상의 수식어가 수식하며, 수식 형태에 따라 별도로 개체를 지정해야 하는 경우** 사용하는 개체 간 관계이다. 가령, ‘손과 발의 통증’이라는 서술을 ‘손 통증’과 ‘발 통증’으로 나누어서 AdverseEvent로 지정하고자 할 때 사용한다. 이 경우 ‘손’, ‘발’, ‘통증’을 각각 AdverseEvent 개체로 지정하며, 다시 ‘손’↔’통증’, ‘발’↔’통증’을 relation으로 연결한다.

이상반응이나 질병명 뒤에 괄호로 동의어나 유사어를 나열한 경우, 괄호 앞의 단어와 괄호 안의 단어를 각각 별도의 entity로 지정하고, relation으로 연결한다. (e.g. '비스테로이드성 소염진통제(NSAID)' -> '비스테로이드성 소염진통제' + 'NSAID', '출혈반(Ecchymosis)' -> '출혈반' + 'Ecchymosis')

첨가제는 DrugCompound로 지정하되, 첨가제가 들어있는 의약품명(주성분)과 relation을 지정하지 않는다.

* 1. AdverseEvent ↔ Seriousness of Adverse Event (AESeriousness), Event at the Time of Last Observation (EventatLastObs)

**이상반응과 해당 이상반응의 중대성, 혹은 이상반응과 해당 이상반응의 최종 관찰 시점에서의 이상사례 지속 및 회복 여부**를 표현하는 개체 간 관계이다.

* 1. DiseaseOrIndication ↔ DrugCompound, DrugProduct, DrugGroup

**투약 약물과 해당 투약 약물의 적응증 관계**를 표현하는 개체 간 관계이다.

* 1. AdverseEvent, DiseaseOrIndication, AESeriousness, AEatLastObs, DrugCompound, DrugProduct, DrugGroup, EventAdmission, EventDischarge, TestName, NonDrugTreatment, ActionTakenwDrug ↔ DateStartorContinue, DateEnd & DatePeriod

**이상사례 및 질병의 발현 시작, 종료, 지속 상태 / 의약품의 투약 시작, 종료, 지속상태 / 이상사례의 중대성을 나타내는 사건의 발생일 / 이상사례 지속 및 회복 여부를 확인한 최종 관찰 시점 / 입원일 및 퇴원일 / 검사 실시일**을 표현하는 개체 간 관계이다. AEatLastObs, AESeriousness ActionTakenwDrug, AEatLastObs으로 지정한 문장이나 절에 날짜가 포함되어 있을 경우, 날짜를 별도의 entity로 지정하고, 관련 사건을 나타낸 entity와 relation을 지정해야 한다. (e.g. '상태 호전되어 8월30일에 퇴원조치함.' -> '상태 호전되어'(AEatLastObs), '8월30일'(DateStartorContinue), '퇴원조치함'(EventDischarge), '8월30일'~'퇴원조치함'(relation)

단, 단순 보고 일자는 relation으로 연결하지 않는다.

* 1. DrugProduct ↔ DrugCompound

**의약품명(상품명)으로 서술된 약물의 성분명 관계**을 표현하는데 사용되는 개체 간 관계이다. 가령, ‘12월 27일부터 록세핀정(록소프로펜나트륨) 1일 3회 경구투여함’이라고 기술된 경우 DrugProduct 개체인 ‘록세핀정’의 성분이 DrugCompound 개체인 ‘록소프로펜나트륨’이기 때문에 두 개체 간 관계를 설정한다. 이외 ‘1일 3회’(DrugDosingInterval), ‘경구투여’(DrugRoAFormulation) 등의 투약 정보 개체는 일괄적으로 DrugProduct인 ‘록세핀정’과 관계를 설정한다.

* 1. DrugCompound, DrugProduct, DrugGroup ↔ DrugDose, DrugDosingInterval, DrugRoAFormulation

**의약품명(상품명)으로 서술된 약물과 해당 약물의 투약 정보, 성분명으로 서술된 약물과 해당 약물의 투약 정보, 약물군으로 서술된 약물과 해당 약물의 투약 정보**를 표현하는데 사용되는 개체 간 관계이다.

* 1. TestName ↔ TestResult

**검사명와 해당 검사의 검사 결과**를 표현하는 개체 간 관계이다.

* 1. ActionTakenwDrug ↔ DrugCompound, DrugProduct, DrugGroup

**의약품에 관한 조치 내용과 해당 조치가 취해진 의약품**을 표현하는 개체 간 관계이다.

* 1. CasualityAssessment ↔ AdverseEvent, DiseaseOrIndication, DrugCompound, DrugProduct

**WHO-UMC 토대의 인과성 평가결과와 해당 인과성 평가결과와 연관된 의약품과 이상사례**를 표현하는 개체 간 관계이다. 연관성이 있는 개체가 확실하지 않은 경우, 관계를 연결하지 않는다.

1. **Entity Labels**

해당 절에서는 앞서 정의한 개체의 세부정보 및 의학용어 내 필드를 지시하는 개체 레이블(entity label)을 정의하며, 해당 entity label을 일관되게 어노테이션하기 위한 규칙과 참고사항을 제공한다.

* 1. Pathological Finding
     1. pathological finding identifier (MedDRA)

이상사례 상세서술 내에서 언급된 **이상사례**(**AdverseEvent), 과거질병(DiseaseOrIndication) 개체가 MedDRA 용어 체계 내 어떤 필드에 해당하는지** 표현하는 개체 레이블이다. 해당 개체 레이블은 **AdverseEvent, DiseaseOrIndication 개체를 지정하는 경우 필수적으로 설정해야 한다.** 어노테이터의 의학적 지식 및 판단에 따라 적절한 8자리 MedDRA 식별자(ex. ‘10028813’ – Nausea)를 입력한다. MedDRA 코드는 24.0 version English & Korean를 사용하며, 코드 검색은 MedDRA Desktop Browser 4.1 version을 사용하여 수행한다. 기본적으로 **MedDRA 코드 내 LLT(Lowest Level Term) 단위의 concept에 사용하여** 식별자를 설정한다. 그러나 해당 AdverseEvent나 DiseaseOrIndication 개체가 단일 PT(Preferred Term) 내 2개 이상의 LLT에 해당한다고 판단되거나, 세부 LLT를 지정하기에 정보가 충분하지 않은 경우 상위 PT를 사용한다. LLT, PT 외 상위 용어(i.e., SOC, HLGT, HLT)에는 매핑하지 않는다.

원칙적으로 pathological finding identifier를 설정할 때 어노테이터의 의학적 판단에 따라서 진단을 내리거나 해당 표현을 해석하여 레이블을 설정하는 것은 허용하지 않는다. 가령, ‘가끔은 월경 주기에 월경을 하고, 아닐때도 있습니다.’와 같은 표현을 두고 ‘불규칙 월경’으로 진단하여서 label을 지정하지 않는다.

* 주의할 질병 및 적응증: ‘구토 예방’ – MedDRA 10068079 (이처럼 특정 의약품을 예방 목적에서 투약한 경우 ‘구토’가 아닌 ‘구토 예방’으로 레이블을 지정한다.)

* + 1. not\_occurred

이상사례 상세서술 내에서 언급된 **이상사례**(**AdverseEvent), 과거질병(DiseaseOrIndication) 개체의 실제 발병 여부**를 표현하는 개체 레이블이다. 별도로 지정하지 않는 경우에 대한 상세 설명은 ‘D. 기타 고려사항: not\_occured/not\_concerned를 지정하지 않는 경우’를 참고한다.

* + 1. seriousness

**이미 지정한 이상사례의 중대성(****AESeriousness) 개체의 이상사례 중증도**를 설정하는 개체 레이블이다. 아래 표의 항목 중 하나의 값을 갖는다. 단, 중대한 이상사례에 해당되나 중대성 분류가 구체적으로 기술되지 않은 경우(e.g. ‘중대한 이상사례: 전신 발진‘, ‘본 SAE 최초보고는’)에는 ‘6\_ 분류되지 않은 의학적으로 중대한 사건(Medically Important\_unclassified’으로 지정한다.

* + - * 주의 사항
        1. 이상반응의 ‘중증도’를 표현한 단어나 명사구는(e.g. CTCAE Grade 1-4, mild/moderate/severe, 경증/중등도/중증)를 ‘중대성’과 혼동하지 않도록 주의해야 한다. 따라서 ‘중증도’를 표현한 단어나 명사구는 AESeriousness로 지정하지 않는다. (‘CTCAE\* Grade 5’는 아래 2) 항목 참고)
        2. ‘사망’한 경우, 해당 표현(‘CTCAE\* Grade 5’ 포함)을 반드시 AESeriouness로 지정하고, seriousness 1\_사망(Death)으로 labeling 해야 한다. 단, ‘원인을 알 수 없는 사망’의 경우, AESeriouness로 지정하고, 동시에 AdverseEvent로도 지정해야 한다.

\* CTCAE: Common Terminology Criteria for Adverse Events

|  |
| --- |
| E2B R3 서식에 따른 이상사례 중대성 분류 |
| 1 = 사망(Death) |
| 2 = 생명의 위협(Life Threatening) |
| 3 = 입원 또는 입원기간의 연장 (Caused/Prologned Hospitalization) |
| 4 = 중대한 불구나 기능저하 (Disabling/Incapacitating) |
| 5 = 선천적 기형 초래 (Congenital Anomaly/Birth Defect) |
| 6 = 분류되지 않은 의학적으로 중대한 사건(Medically Important\_Unclassified) |
| 7 = 중대하지 않음 (Medically Not Important) |

* + 1. event\_at\_last\_observation

이미 지정한 **최종 관찰 시점에서의 이상사례 지속 및 회복 여부(AEatLastObs) 개체의 최종 관찰결과**를 설정하는 개체 레이블이다. 아래 표의 항목 중 하나의 값을 갖는다. 이때 ‘마지막 관찰 시점’은 이상사례 상세서술 상에서 시간 상으로 가장 늦은 이상사례 관찰 시점을 의미한다. 따라서, 가령 이상사례가 회복되었다가 다시 재발한 경우까지 보고된 경우에는 ‘회복됨’이 아니라 ‘회복되지 않았음’으로 분류되어야 한다.

|  |  |
| --- | --- |
| E2B R3 서식에 따른 이상사례 최종 관찰 결과 분류 | 분류별 상황 및 예시 |
| 1 = 회복됨 (Recovered) | 마지막 관찰 시점에서 해당 이상반응이 더 이상 관찰되지 않으며 이상반응 발생 이전의 건강 상태를 되찾은 경우  예시) “환자는 복막염을 회복하였음” |
| 2 = 회복중 (Recovering) | 마지막 관찰 시점에서 온전히 환자가 회복되지는 않았으나 회복되는 과정에 있다고 판단되는 경우 |
| 3 = 회복되지 않았음 (Not recovered) | 마지막 관찰 시점에서 해당 이상반응이 여전히 관찰되며, 회복되거나 회복되는 과정에 있다고 판단되지 않는 경우  예시) ‘2017년 10월 두드러기 재발함’ |
| 4 = 회복했지만 후유증 있음 (Recovered with sequelae) | 마지막 관찰 시점에서 해당 이상반응이 더 이상 관찰되지 않으나 이상반응에 의한 신체의 손상이 회복되지 않아 관찰되는 경우 |

* + 1. startcontinue\_or\_end

이미 지정한 **이상사례 개체가 DateStartorContinue 개체와 관계를 맺으나 실제로 해당 날짜에 증상이 종료/중지한 경우** 활용하는 개체 레이블이다. 그러한 경우 ‘end’로 startcontinue\_or\_end 개체 레이블을 지정한다.

* 1. Drug
     1. compound identifier & product identifier

**DrugCompound entity의 성분명 식별자와 DrugProduct entity의 상품명 식별자 코드**를 지정하는 객체 레이블이다. Drug entity를 설정하는 경우에는 필수적으로 확정해야 한다. compound identifier & product identifier 각각 식품의약품안전처가 제공하는 ‘2020년 하반기 성분코드집’과 ‘2020년 하반기 의약품코드집’을 사용한다. compound identifier의 경우 M으로 시작하는 총 8글자의 식별자(ex. 통합성분코드 ‘CM050288’ – 디클로페낙)를 서술형으로 입력하며, product identifier의 경우 총 10글자의 식별자(ex. 통합품목기준코드 ‘C195600001’ – 판피린정)를 서술형으로 입력한다. 만약, 숫자를 포함하여 의약품명(상품명)의 drug entity를 지정한 경우, 의약품명+숫자가 가장 유사한 의약품명의 코드를 레이블로 지정한다.

단, 상기의 코드집에 등록되지 않은 개별 한방병원 또는 한의원 조제 한약의 경우, C000000000로 지정한다.

* + 1. not\_occurred (not\_administered)

이상사례 상세서술 내에서 언급된 **Drug 개체의 실제 투여 여부**를 표현하는 개체 레이블이다. 별도로 지정하지 않는 경우에 대한 상세 설명은 ‘D. 기타 고려사항: not\_occured/not\_concerned를 지정하지 않는 경우’를 참고한다.

* + 1. not\_concerned

이상사례 상세서술 내에서 언급된 **Drug 개체가 이상사례 보고자료 내 의심 약물인지**를 지정하는 개체 레이블이다. 상세서술 내에서 서로 다른 두 개 이상의 약제가 언급된 경우 이상사례 및 과거질병를 구분하기 위해 해당 상세서술 내에서 ‘의심약제’가 무엇인지 확정해야 한다. 어노테이터가 이상사례, 과거질병 개체를 구분하여 지정하는데 기준으로 삼은 약제 1개를 지정하여야 한다. 별도로 지정하지 않는 경우에 대한 상세 설명은 ‘D. 기타 고려사항: not\_occured/not\_concerned를 지정하지 않는 경우’를 참고한다.

* + 1. startcontinue\_or\_end

이미 지정한 **Drug 개체가 DateStartorContinue 개체와 관계를 맺으나 실제로 해당 날짜에 투약을 종료/중지한 경우** 활용하는 개체 레이블이다. 그러한 경우 ‘end’로 startcontinue\_or\_end 개체 레이블을 지정한다.

* 1. Others
     1. test name identifier (MedDRA)

이상사례 상세서술 내에서 언급된 **TestName 개체가 MedDRA 용어 체계 내 어떤 필드에 해당하는지** 지정하는 개체 레이블이다. TestName entity를 설정하는 경우 필수적으로 확정해야 한다. 서술형으로 지정된 8자리 MedDRA 식별자(ex. ‘10011328’ – Creatine)를 입력한다. 어노테이션 방식은 기본적으로 Pathological Finding entity의 identifier 설정 방식과 동일하다. 다만 test name identifier는 SOC(System-Organ Class) 상으로 ‘Investigations’에 속하는 LLT 및 PT 용어에 제한된다.

* + 1. non-drug treatment (MedDRA)

이상사례 상세서술 내에서 언급된 **NonDrugTreatment 개체가 MedDRA 용어 체계 내 어떤 필드에 해당하는지** 지정하는 개체 레이블이다. NonDrugTreatment entity를 설정하는 경우 필수적으로 확정해야 한다. 서술형으로 지정된 8자리 MedDRA 식별자(ex. ‘10065608’ – 경피적 관상 동맥 중재)를 입력한다. 어노테이션 방식은 기본적으로 Pathological Finding entity의 identifier 설정 방식과 동일하다.

* + 1. action taken with drug (E2BR3)

이상사례 상세서술 내에서 언급된 **ActionTakenwDrug 개체가 E2BR3에서 제공하는 의약품 조치 분류에서 어떤 필드에 해당하는지** 지정하는 개체 레이블이다. ActionTakenwDrug entity를 설정하는 경우 필수적으로 확정해야 한다.

|  |
| --- |
| E2BR3 서식에 따른 의약품 조치 분류 |
| 1 = 투약 중지 |
| 2 = 투약량 감소 |
| 3 = 투약량 증가 |
| 4 = 투약량 변경하지 않음 |
| 5 = 모름 |

* + 1. WHO-UMC results of assessment (WHO-UMC)

이상사례 상세서술 내에서 언급된 **CasualityAssessment 개체가 WHO-UMC에서 제공하는 인과성 평가 결과에서 어떤 필드에 해당하는지** 지정하는 개체 레이블이다. ActionTakenwDrug entity를 설정하는 경우 필수적으로 확정해야 한다. 아래 표와 같은 분류별 상황에 따라 레이블을 부여하지만 기본적으로 보고자나 원보고자가 상세서술에서 선택한 표현에 기반하여 레이블을 부여하도록 한다.

|  |  |
| --- | --- |
| WHO-UMC 기준에 따른 인과성 평가 결과 | 분류별 상황 및 예시 |
| 1 = 확실함(Certain) | 의약품 투여 시점 및 이상반응의 발생 시점의 선후 관계나 임상정보를 미루어보아 이상사례일 가능성이 높으며, 다른 요인에 의한 증상이라고 생각되지 않는 경우 |
| 2 = 상당히확실함(Probable) | 의약품 투여 시점 및 이상반응의 발생 시점의 선후 관계나 임상정보를 미루어보아 이상사례일 가능성이 높으며, 다른 요인에 의한 증상일 가능성이 낮은 경우 |
| 3 = 가능함(Possible) | 의약품 투여 시점 및 이상반응의 발생 시점의 선후 관계나 임상정보를 미루어보아 이상사례일 가능성이 있으나, 다른 요인에 의한 증상일 가능성이 배제되지 않은 경우 |
| 4 = 가능성적음(Unlikely) | 인과성을 완전히 배제할 수는 없으나 (병용약물이나 기저 질환에 의한) 다른 인과관계가 더 설득력이 높은 경우 |
| 5 = 평가곤란(Conditional/unclassified) | 안전성 평가를 위한 임상정보가 부족하여서 평가가 어렵다고 판단한 경우 |
| 6 = 평가불가(Unassessable/unclassified) | 보고자 및 원보고자 평가를 수행하지 않았다고 서술한 경우 |

* 1. 기타 고려사항: not\_occured/not\_concerned를 지정하지 않는 경우
     1. 문서 내에서 not\_occured 또는 not\_concerned 레이블을 지정한 개체가 하나도 없는 경우
        + Drug entity 중 not\_occurred로 지정한 개체가 없는 경우: Drug entity로 annotation한 모든 의약품을 투약한(occurred) 의약품으로 간주
        + Drug entity 중 not\_concerned로 지정한 개체가 없는 경우: Drug entity로 annotation한 모든 의약품을 모두 의심약제(concerned)로 간주
        + PathologicalFinding entity 중 not\_occurred로 지정한 개체가 없는 경우: PathologicalFinding으로 annotation한 모든 이상반응/질병을 모두 실제 발생한(occurred) 이상반응/질병으로 간주
     2. Drug/PathologicalFinding 개체에서 한 개체라도 not\_occurred, not\_concerned 레이블 지정을 한 경우
        + not\_occurred, not\_concerned로 지정하지 않은 개체는 모두 occurred, concerned로 간주한다.
* 등장한 10개 이상사례 개체 중 2개를 not\_occurred로 처리할 경우 나머지 8개 이상사례 개체는 자동으로 occurred로 인식
* 등장한 3개 약물 개체 중 1개를 concerned로 처리할 경우, 나머지 2개 약물 개체는 자동으로 not\_concerned로 인식
* 등장한 이상사례 4개는 not\_occurrred 레이블을 따로 처리하지 않고, 등장한 2개 약물 개체 중 1개만 occurred 처리할 경우, 이상사례 4개는 occurred로, 레이블 처리하지 않은 나머지 1개 약물 개체는 not\_occurred로 인식