

Simulants: 합성 임상시험 데이터로 새로운 인사이트 발견 및 의료 혁신 가속화

개요:

매일, 임상 시험의 데이터가 다음 의학적 혁신을 찾기 위해 전자 데이터 수집(EDC) 시스템에 입력됩니다. 이 데이터에는 임상 시험 프로토콜의 일부로 수집된 공변량, 엔드포인트, 그리고 건강 데이터가 포함되어 있어 새로운 치료제의 안전성과 효과에 대한 풍부한 인사이트를 제공하지만, 개별 조직 및 각각의 시스템 내에 갇혀 있어 접근하기 어렵습니다.

과거 임상 시험 데이터를 활용하는 것은 이미 효과가 입증된 점과 그렇지 않은 점을 통해 연구 및 개발을 가속화할 수 있는 통찰을 제공함으로써 도움이 될 수 있습니다. 하지만, 환자 식별 정보(PII)나 지적 재산(IP)을 포함한 다른 민감한 데이터와 마찬가지로, 임상시험 데이터는 규제 요건, 기술 보호 프로토콜, 임상시험 데이터의 높은 소유 가치, 그리고 환자 동의의 문제와 환자 신뢰 유지에서 비롯된 의뢰사의 엄격한 프라이버시 요건에 의해 접근성이 제한적입니다. 임상 데이터 공유 서약, 정책 및 프로토콜이 있음에도 불구하고 이러한 방해 요소는 비식별화된 환자 레벨 임상시험 데이터의 공유를 거의 불가능하게 만들 정도로 큰 영향을 미치고 있습니다.

생성형 AI의 등장으로 이제 개인정보 보호를 유지하는 동시에 원본 데이터 소스의 속성과 패턴을 모방하는 합성 데이터를 생성할 수 있게 되었습니다. 실제 사람 또는 사건에서 수집된 것이 아닌 인공적으로 또는 알고리즘을 바탕으로 생성된 합성 데이터는 가장 민감하고 중요한 데이터를 공유 및 액세스할 수 있는 혁신적인 방법을 제공합니다.

합성 데이터를 사용하는 이유

합성 임상시험 데이터는 임상 개발자와 연구자가 환자의 개인 정보와 IP를 보호하면서 교차 의뢰자 과거 임상시험 데이터와 상호작용할 수 있는 고유한 기능을 제공합니다.

KEY POINT:

합성 데이터는 개인 정보를 보호하면서 기본 데이터셋의 특성을 유지하여 민감한 데이터에 접근하는 실용적이고 효율적이며 안전한 방법을 제공합니다.

최초의 합성 데이터 솔루션인 Simulants는 임상시험 과거 데이터를 기반으로 합성 데이터를 생성합니다. 이는 임상시험 프로토콜을 통해 수집된 공변량, 엔드포인트 및 변수가 포함된 데이터입니다. 이를 통해 높은 신뢰도의 데이터셋을 보장합니다.

이제 임상 개발자는 Simulants를 활용해 (i) 과거 임상시험 데이터에 포함된 중요 인사이트를 확보하고, (ii) 대표성 있는 균형 잡힌 합성 임상시험 데이터를 생성해 현재 임상시험 데이터의 편향을 해소할 수 있습니다.

환자 개인정보 보호

합성 데이터 사용 시, 연구에 대한 환자의 명시적 동의를 존중함으로써 의학 연구 커뮤니티에서 신뢰와 윤리적 데이터 처리를 촉진합니다.

의뢰자 IP 보호

합성 데이터 생성은 연구자를 위한 실용적이고 접근성 높은 솔루션을 제공함으로써 광범위한 데이터 공유 계약이나 복잡한 데이터 처리 절차 없이도 고품질의 데이터를 활용할 수 있게 해줍니다.

의료 혁신 가속화

데이터를 이용해 임상시험 설계를 개선하고, 선정/제외 기준을 최적화하며, 과거 임상시험에서 관련 엔드포인트를 식별하여 보다 효율적인 신약 개발 프로세스를 구현합니다.

Simulants:

Simulants는 33,000여 건의 교차 의뢰자 임상시험과 1,000만 명 이상의 환자를 아우르는 표준화된 글로벌 과거 임상시험 데이터로 구성된 메디데이터의 독점 저장소에서 도출된 혁신적인 합성 데이터 솔루션입니다. 이 포괄적인 기본 데이터셋은 각 시험계획서에서 수집된 공변량과 엔드포인트 및 변수를 포함합니다. 합성 데이터에 사용되는 일반적인 데이터셋과 비교할 때, Medidata의 특허 알고리즘에 기반한 Simulants는 고유한 중요 데이터를 포함하는 임상시험 데이터셋에서 고신뢰도의 합성 데이터를 생성하는 데 최적화되어 있습니다. 엄격한 규제 및 개인정보 보호 요건을 충족하는 것으로 입증된 Simulants는 임상 개발자의 데이터 접근성을 넓혀 연구 정확성을 개선하고 인사이트에 기반한 근거를 제공하며 임상시험 설계를 최적화합니다.

Simulants 활용 사례:

질병에 대한 심도 있는 이해

- 시험계획서 및 임상시험 매개변수 최적화
- 가장 큰 의료 수요 부족을 가진 부분 모집단 식별
- 풍부한 환자 여정 구성 및 시간과 모집단 간의 치료 결과 비교, 실제 임상 치료 패턴 평가
- 치료 유효성 및 생체표지자 간 연관성 파악

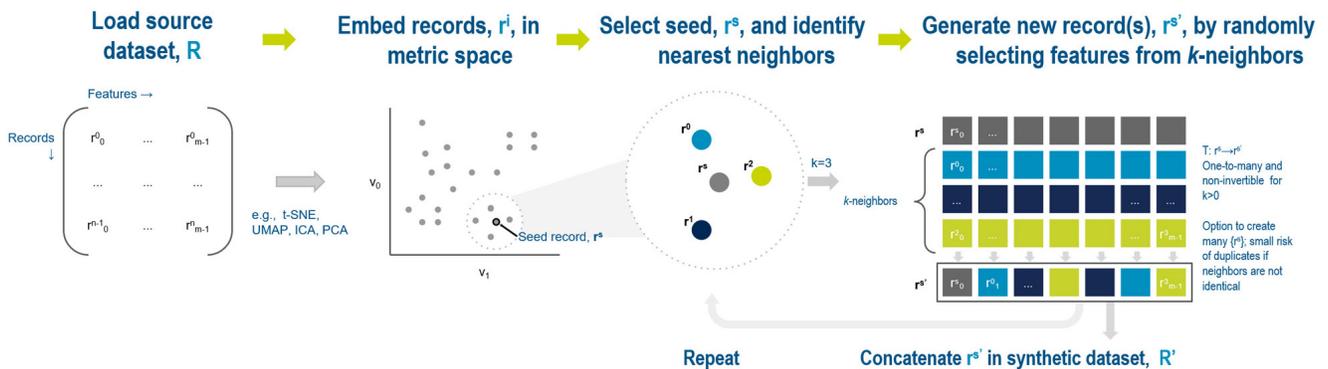
환자 안전성 향상

- 중도 탈락 가능성이 가장 높은 환자 식별, 참여 또는 모집 전략 개발
- 심각한 이상반응(SAE) 발생 위험이 가장 높은 환자 식별
- SAE 조기 징후 식별 및 장기 안전성 결과 모니터링

임상시험 성공 가능성 증대

- 다양한 임상시험 시나리오 모델링을 통한 보다 효과적인 임상시험 설계
- 합성 데이터를 이용한 모델 훈련 또는 검증, 보다 균형 잡힌 데이터셋을 생성하고 데이터 내 표본이 작은 그룹을 '업샘플링'하는 기능 포함
- 임상 개발 의사결정, 규제 제출 또는 지불자 논의를 위한 생성형 비교 근거
- 환자 집단 모집 "불가능 여부" 파악

간소화된 프로세스



발표 자료

- US Patent:** [원본 데이터셋에서 합성 데이터셋을 생성하기 위한 시스템 및 방법](#)
- ICML 2023:** [합성 임상시험 데이터의 생성형 모델링 개선을 위한 해석 가능한 데이터 증강 프레임워크\(우수 논문상\)](#)
- Neurips 2022:** [합성 임상시험 데이터를 위한 소스 데이터 개인정보보호 프레임워크](#)
- Neurips 2022:** [환자 레벨 개인정보를 보호하는 합성 임상시험 데이터](#)
- AMIA 2022:** [Simulants: 환자 레벨 개인정보를 보호하는 방식의 합성을 통한 합성 임상시험 데이터](#)

다쏘시스템 계열사 메디데이터는 생명 과학의 디지털 혁신을 주도하고 있습니다.

더 자세한 정보는 홈페이지(www.medidata.com/kr)를 통해 확인하실 수 있습니다. 문의사항 Medidata.KR.Marketing@3ds.com