1. **Target Trial Emulation이란 무엇인가요?**

Target Trial Emulation은 관찰 데이터(예: EMR, registry data)를 사용하여 가상의 임상시험(Target Trial)을 설계하고, 그 설계를 관찰 연구에 모사하여(emulation) 의학적 질문에 답하는 체계적인 연구 방법입니다.

즉, "만약 이 약의 효과를 확인하기 위해 잘 통제된 무작위배정 임상시험(Randomized Controlled Trial)을 진행했다면, 어떤 디자인이었을까?"라는 질문을 먼저 던지고, 그 임상시험의 규칙을 관찰 데이터에 적용하는 것입니다.

**1.1 Target Trial Emulation을 적용하는 이유**

무작위배정 임상시험(RCT)이 윤리적, 비용적, 또는 현실적인 이유로 불가능할 때, 우리는 전자의무기록과 같은 관찰 데이터에 의존하게 됩니다.

하지만 관찰 데이터에는 흔히 다음과 같은 문제들이 발생하여 연구 결과의 신뢰도를 떨어뜨릴 수 있습니다:

1. **교란(Confounding):** 특정 약물로 치료를 받는 사람과 그렇지 않은 사람이 처음부터 상태가 다르기 때문에, 도출된 결과가 약물 때문인지, 아니면 환자의 상태 차이 때문인지 알기 어렵습니다.
2. **선택 편향 (Selection Bias):** 특히 불멸 시간 편향(Immortal Time Bias)\*과 같은 편향이 발생하기 쉽습니다. 이 편향은 관찰 데이터로부터 연구 참여자를 선정하는 과정에서 베이스라인 이후의 정보를 사용하는 경우 발생하여 결과를 왜곡시킵니다.

TTE는 이러한 흔한 연구 설계상의 결함을 피하도록 강제하고, 이를 통해 관찰 연구 결과가 RCT의 결과와 더 잘 일치하도록 돕습니다.

\* 불멸 시간 편향(Immortal Time Bias, ITB)은 관찰 연구, 특히 EMR을 사용하여 시간이 지남에 따른 치료 효과를 평가할 때 흔히 발생하는 선택 편향(selection bias)입니다.

이 편향은 연구 대상자가 특정 치료를 받기 시작한 시점부터 연구 기준선(Time Zero)까지의 기간 동안, 해당 대상자에게 관심 결과(outcome)가 발생할 수 없는 '불멸의 시간(Immortal Time)'이 존재할 때 발생합니다. 이 시간을 적절히 처리하지 못하고 분석에 포함하면, 연구 결과가 인위적으로 왜곡됩니다. 관찰 연구에서는 모든 대상자가 무작위배정 대조군 임상시험(RCT)처럼 동시에 치료를 시작하지 않기 때문에 발생합니다.

1. **TTE 설계의 핵심 고려 사항**

임상시험을 설계할 때 연구자가 반드시 고려해야 할 핵심 요소들은 아래와 같이 정리할 수 있습니다.

|  |  |  |
| --- | --- | --- |
| **핵심 질문** | **반드시 고려해야 할 사항** | **흔히 발견되는 문제점** |
| **1. 누가 연구에 포함되는가?**  (Eligibility) | **Time Zero 원칙:** 연구 참가자의 연구 시작 시점을 명확히 정의하고, 이 시점 이전에 얻은 정보만을 사용하여 적격 기준을 정의해야 합니다. | **연구 시작 시점 이후의 정보를 사용:** 연구 참여 자격을 정의할 때 미래의 정보(예: 최소 1회 이상의 추적 관찰 기록)를 사용하는 경우가 흔합니다. 이는 명백한 편향을 유발합니다. |
| **2. 치료 전략은 무엇이고 언제 시작되는가?**  (Treatment strategies) | **치료 전략 정의:** 비교하고자 하는 치료 전략(예: 특정 약물 사용 vs. 사용하지 않음)과 약물 중단 또는 변경에 대한 프로토콜을 명확히 정의해야 합니다. | **시작 시점의 불분명함:** 관찰 연구에서는 모든 환자가 동시에 치료를 시작하지 않기 때문에, 현실적인 '유예 기간(Grace Period)'을 설정하여 치료 시작을 허용해야 합니다. 이를 명시하지 않으면 인과적 질문이 불분명해집니다. |
| **3. 무작위 배정을 어떻게 모사하는가?**  (Assignment procedures) | **교란 변수 선택:** 무작위 배정을 흉내 내기 위해 통계적 조정이 필요하며, 교란 변수는 임상 지식과 이론을 바탕으로 선택해야 합니다 (통계적 연관성만으로 선택해서는 안 됨). | **통계적 선택에 의존:** 교란 요인 선택 시 오로지 통계적 기준(예: P-값 기준)에만 의존하거나, 아예 교란을 조정하지 않는 경우가 많습니다. |
| **4. 우리가 알고 싶은 인과적 효과는 무엇인가?**  (Causal contrasts of interest) | **인과 대비(Causal Contrast) 명시:** 연구자가 얻고자 하는 결과(답변)가 ITT 효과인지 Per-protocol 효과인지 분석 전에 명확해야 합니다. | **인과 대비 정의 부족:** 대부분의 관찰 연구에서 ITT 또는 Per-protocol 효과 중 어떤 것을 추정할 것인지 명확히 정의하지 않는 문제가 발견되었습니다. |

1. **연구기간의 설정: Time Zero 및 Grace Period**

**3.1 Time Zero (시간 0점)의 정의와 역할**

Time Zero는 대상자가 적격성 기준(Eligibility criteria)을 충족하고, 치료 전략이 배정되며, 추적 관찰이 시작되는 시점으로 정의됩니다. 이는 무작위배정 대조군 시험(RCT)에서 무작위배정(Randomization)이 일어나는 시점의 관찰 연구에서의 유사점(observational analog)입니다.

**3.2 Grace Period (유예 기간)의 역할과 분석적 처리**

유예 기간은 실제 임상 환경의 제약 조건을 반영하기 위해 설정됩니다. 예를 들어, 환자가 치료를 시작하기로 결정한 후에도 임상 검사나 행정적 절차를 완료하는 데 몇 주가 걸릴 수 있으므로, 치료 전략을 ‘Time Zero로부터 일정 기간(예: 1~24개월) 내에 치료를 시작하는 것'으로 정의하는 것입니다.

유예 기간을 설정하면 전략이 현실성을 유지하고, 관찰 데이터에서 표적 시험을 모방하는 데 사용할 수 있는 인원의 수를 늘릴 수 있습니다. 하지만 유예 기간을 사용할 경우, 분석에 복잡성이 더해질 수 있는데, 예를 들어, 유예 기간 중 치료를 시작하지 않고 사망한 경우, 이를 단순히 대조군(치료 미시작군)으로만 배정하면 해당 군의 위험 추정치를 인위적으로 부풀려 편향을 유발할 수 있습니다.

이러한 문제를 해결하기 위해, 연구 시작 시점(Time Zero)부터 추적 관찰을 시작하되, 특정 분석 기법을 사용하여 편향을 방지해야 합니다.

1. **교란 변수의 선택**

**4.1 교란 변수 선택의 기본 원칙 (임상 지식 기반)**

무작위 배정 절차(Assignment procedures)를 성공적으로 모방하기 위해서는, 교란 변수를 선택할 때 통계적 연관성에만 의존해서는 안 되며, 반드시 임상적 지식(Subject Knowledge)과 이론을 바탕으로 해야 합니다.

4.1.1 이론 기반 선택 (DAGs 활용)

• 이론 및 문헌 검토: 교란 변수는 임상 지식 및 문헌 검토를 통해 선택합니다.

• DAGs (Directed Acyclic Graphs) 사용: DAGs를 사용하여 교란 요인을 시각적으로 명시할 것이 권장됩니다. 이는 통계적 선택에만 의존하는 것보다 편향을 줄이는 데 도움이 됩니다.



SES= socioeconomic status DMARDs= disease modifying antirheumatic drugs, HRT= hormone replacement therapy, BMD= bone mineral density, MOF= major osteoporotic fracture

**Figure 1**. 교란 요인 선택을 위한 DAG 활용의 예시 (Tronstad I et al. *RMD Open*. 2024;10(1):e003919.)

4.1.2 시간성(Temporality) 유지

• 기준선(Time Zero) 정보: 교란 변수는 연구 기준선(baseline) 시점 또는 그 이전에 측정된 요인이어야 합니다.

• 시점 일치: 교란 변수로 간주되는 검사 수치(laboratory values) 등은 치료 배정 시점 이전에 평가되어야 하며, 일관된 기간(look-back window) 내에서 확인되어야 합니다.

**4.2 류마티스 관절염 연구에서 고려해야 할 구체적인 교란 변수**

류마티스 관절염 환자를 대상으로 DMARDs (Disease-Modifying Anti-Rheumatic Drugs)의 효과를 비교하는 연구(예: MTX+HCQ(±SSZ) vs. MTX+LEF 병용 요법의 유지율 비교 또는 2차 DMARDs 추가 요법의 심혈관 질환 위험 비교)에서 조정이 필요한 주요 교란 변수들은 다음과 같습니다.

|  |  |  |
| --- | --- | --- |
| **교란 변수 범주** | **구체적인 변수 (예시)** | **조정 필요성 및 근거** |
| **인구 통계 및 사회 경제적 요인** | 나이(Age), 성별(Gender), 인종 및 민족(Race and ethnicity), 건강 보험 적용 범위(Health insurance coverage) | 이들은 치료 결정이나 결과에 영향을 미칠 수 있습니다. |
| **RA 질병 활동도 및 중증도** | DAS28-ESR/CRP (Disease Activity Score), CDAI (Clinical Disease Activity Index), SDAI, ESR, CRP, 질병 기간(Disease duration), 관절 부종 및 압통 수(Swollen/Tender joint counts), 환자 및 의사의 전반적인 평가(Patient/Physician global assessment). | RA 치료는 치료 목표 달성 전략(treat-to-target strategy)에 기반하므로, 기준 시점의 질병 활동도는 치료 선택(용량, 종류)에 직접적인 영향을 미치는 주요 교란 요인입니다. |
| **면역 및 기능 상태** | 류마티스 인자(RF) 및 ACPA 상태 (Anti-citrullinated protein antibody), 신체 기능 상태 (HAQ-DI: Health Assessment Questionnaire - Disability Index) 및 통증(pain). | 환자의 혈청학적 상태와 기능 수준은 예후 및 치료 반응과 관련이 있습니다. |
| **동반 질환 및 약물** | **동반 질환의 유무 또는 개수**, 고혈압(Hypertension), 당뇨병(Diabetes mellitus), 심혈관 동반 질환(Atherosclerotic CVD, Atrial fibrillation). | 특히 심혈관 결과(MACE)를 연구할 때, 기존의 심혈관 위험 요인은 교란 변수로 포함되어야 합니다. 동반 질환 유무는 치료 중단(discontinuation)의 위험 요소이기도 합니다. |
| **이전 치료 경험** | 이전 DMARD 치료 경험 (수 및 유형), 스테로이드 및 NSAIDs 동반 사용 여부, MTX 투여 경로 및 용량. | 이전 치료 실패 이력은 이후 치료제의 효과와 관련이 깊어 반드시 조정해야 합니다. |
| **생화학적 수치** | eGFR (estimated glomerular filtration rate), 총 콜레스테롤(Total Cholesterol), 간 전이효소(Liver transaminases) 등(eligibility criteria에도 사용). | 이러한 수치는 치료 결정(예: 특정 약물의 금기 여부)과 결과에 영향을 미칠 수 있습니다. |

**4.3 통계적 조정 방법**

교란 변수를 선택한 후, 관찰 데이터에서 무작위화를 모방하기 위해 다양한 통계적 조정 방법이 사용됩니다.

4.3.1 기준선 교란 조정

치료 전략 개시 그룹 간의 비교를 위해 기준선 교란 변수를 조정해야 합니다. 이는 성향 점수 매칭(propensity score matching), 계층화(stratification), 회귀 분석(regression), 표준화(standardization) 또는 역 확률 가중치(Inverse Probability Weighting, IPTW) 등의 방법을 통해 수행될 수 있습니다.

4.3.2 시간 가변 교란 조정

연구 대상 치료 전략이 장기간 지속되는 경우, 치료 순응도와 관련된 기준선 및 기준선 이후(post-baseline) 예후 인자를 조정해야 합니다. 치료 전략에 의해 영향을 받는 기준선 후 예후 인자(시간 가변 교란 변수)가 존재하는 경우, Robins의 g-방법(g-methods)(예: 시간 가변 가중치를 사용한 IPTW, g-공식 등)이 일반적으로 필요합니다.

1. **통계 분석 계획**

**5.1 ITT (Intention-to-Treat) 효과와 PP (Per-Protocol) 효과의 분석**

이 두 가지 효과는 가상의 임상시험(Target Trial) 프레임워크를 포함하여 임상시험(RCT) 및 관찰 연구(Observational Studies)에서 인과적 효과(Causal Effects)를 추정하는 데 있어 핵심적인 인과 대비(Causal Contrast)입니다.

5.1.1 ITT 효과의 정의 및 분석 방법

• 정의: 연구 기준선(baseline)에서 치료 전략에 배정된 것의 비교 효과를 추정합니다. ITT 분석은 치료 순응도(adherence)와 관계없이 수행되며, 관찰 데이터의 분석에서도 기준선 교란 변수(baseline confounders)에 대한 적절한 조정만 이루어진다면 인과적 효과를 추정할 수 있습니다.

• 관찰 연구에서의 유사 분석: 관찰 연구에서는 무작위 배정이 없기 때문에, ITT 효과를 직접적으로 에뮬레이션하기 어렵습니다. 관찰 연구에서 ITT 분석의 가장 가까운 유사 분석(observational analog)은 치료 시작자(initiators) 그룹 간의 비교이며, 이때 기준선 교란 변수에 대한 적절한 조정이 필요합니다.

5.1.2 PP 효과의 정의 및 분석 방법

• 정의: 연구 계획서에 명시된 치료 전략을 준수했을 때의 비교 효과를 추정합니다. PP 효과는 생물학적 효과(biological efficacy)에 더 가깝거나, "만약 모든 환자가 프로토콜을 따랐다면 결과가 어땠을까?"라는 이상적인 시나리오에 답합니다.

• 관찰 연구에서의 필요성: 관찰 연구에서 ITT 효과를 정의하기 어려울 때, PP 효과를 추정하여 임상적 관련성을 갖는 질문에 답할 수 있습니다

• 분석 방법: 치료 전략이 장기간 지속되는 경우, PP 효과를 추정하기 위해서는 기준선 및 기준선 후(postbaseline) 교란에 대한 조정이 필수적입니다. 관찰 연구에서 PP 효과를 추정하는 일반적인 단계는 다음과 같습니다.

1. **순응도 가중치 추정:** 시간 가변 순응도 역 확률 가중치(Inverse Probability of Adherence Weights)를 추정합니다.

2. **치료 가중치 추정:** 치료를 결과 변수로, 기준선 공변량(confounders)을 예측 변수로 하는 로지스틱 회귀 모델을 사용하여 IPTW를 추정합니다.

3. **최종 모델:** 1단계와 2단계에서 추정한 가중치의 곱을 사용하는 가중치 풀링 로지스틱 회귀 모델(weighted pooled logistic regression model)을 사용하거나, 1단계 가중치만을 사용하고 기준선 공변량을 조정한 가중치 풀링 로지스틱 회귀 모델을 사용합니다. 이 모델을 사용하여 생존 곡선, 특정 시점의 위험 차이, 평균 위험비(HR) 등을 계산할 수 있습니다.

    ◦ 예시: Rivera AS et al.의 연구에서는 클론 생성(cloning) 방법과 가중치 풀링 로지스틱 회귀 모델을 사용하여 PP 효과를 추정했습니다.

**5.2 인과적 효과 추정을 위한 통계적 도구**

5.2.1 역 확률 치료 가중치 (Inverse Probability of Treatment Weighted, IPTW)

IPTW는 기준선 교란 변수(baseline confounders)를 조정하여 무작위 배정(random assignment)을 흉내 내기 위해 사용되는 방법입니다. IPTW는 각 연구 대상자가 실제로 받은 치료 전략을 받을 확률의 역수를 가중치로 사용합니다. 이 가중치를 적용하면, 치료를 받은 그룹과 받지 않은 그룹 간의 기준선 교란 변수 분포가 균형을 이루게 되어 (마치 무작위 배정된 것처럼) 교란 변수에 의한 편향이 제거됩니다.

• 분석 역할:

    ◦ IPTW는 ITT 효과의 관찰 연구 아날로그(analog)를 추정하는 데 사용될 수 있습니다.

    ◦ 관찰 데이터에서 ITT 효과의 가장 가까운 분석은 치료 시작자(initiators) 간의 비교이며, IPTW는 이때 기준선 교란 변수를 적절히 조정하는 데 활용됩니다.

5.2.2 Cloning-Censoring-Weighting (CCW)

클로닝, 중도 절단 및 가중치 부여 방법은 유예 기간이 있는 분석에서 PP 효과를 추정하는 데 사용됩니다. CCW 방법은 불멸 시간 편향을 피하면서 PP 효과를 추정하기 위해 세 가지 주요 단계를 포함합니다.

1. Cloning (복제): 연구 기준선(study baseline)에서 데이터가 2가지 이상의 치료 전략과 일치하는 모든 개인을 정확히 복제하여(Clones) 데이터 세트를 확장합니다.
2. Censoring (중도 절단): 복제된 개인들은 자신의 데이터가 할당된 전략과 일치하지 않게 되는 시점에 중도 절단됩니다.
3. Weighting (가중치 적용): 이 정보 제공적 중도 절단으로 인해 발생할 수 있는 잠재적 편향을 수정하기 위해 시간 가변 인자에 대한 적절한 조정(역 확률 가중치 사용)이 필요합니다.
4. **References**
5. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764.
6. Rivera AS, Pierce JB, Sinha A, et al. Designing target trials using electronic health records: A case study of second-line disease-modifying anti-rheumatic drugs and cardiovascular disease outcomes in patients with rheumatoid arthritis. *PLoS One*. 2024;19(6):e0305467.
7. Zhao SS, Lyu H, Solomon DH, Yoshida K. Improving rheumatoid arthritis comparative effectiveness research through causal inference principles: systematic review using a target trial emulation framework. *Ann Rheum Dis*. 2020;79(7):883-890.
8. Bhavsar SV, Movahedi M, Cesta A, Pope JE, Bombardier C; other OBRI investigators. Retention of triple therapy with methotrexate, sulfasalazine, and hydroxychloroquine compared to combination methotrexate and leflunomide in rheumatoid arthritis. *Joint Bone Spine*. 2024;91(4):105732.
9. Tronstad I, Hoff M, Horn J, et al. Rheumatoid arthritis, disease-modifying antirheumatic drugs and risk of major osteoporotic fracture: prospective data from the HUNT Study, Norway. *RMD Open*. 2024;10(1):e003919.