Dataset generation for M&S

Qc Checklist

**1. Dataset generation strategy**

제공받은 파일은 총 3개(복사본 ★ S24008\_Final Resul\_보고서용\_240527.xlsx, G001\_Ocular PK 분석결과\_(16주차)\_2024.06.24 (공유본).xlsx, G001\_Ocular PK 분석결과\_(16주차)\_2024.06.24 (공유본).xlsx)이다. R에서의 data tidying의 편리함을 위해서 xlsx파일을 모두 csv로 변환한 뒤에 R로 raw상태의 파일을 불러왔습니다. 3개의 파일의 내용은 G1~G11에서 시간에 따른 Aflibercept 농도(ng/mL)이였고 총 7개의 SITE(순서대로 aq, vit, iris, retina, choroid, optic nerve, serum)와 좌/우로 농도값이 나누어져 있었다. ID의 경우 명시된 ID값을 그대로 사용하였고, TIME의 경우 시간(TIME), 날(DAY), 주(WEEK)로 나누었고, 농도값이 BQL인 값들은 모두 0(DV =0, MDV=1)으로 처리하였으며 좌우 구분을 위해 LEFT(1: left, 0: right) column을 생성하였고, 편의를 위해서 GROUP, BLQ, DRUG(4: Vehicle, 3: 3TF, 2: Aflibercept, 1: AIO), ROW\_Number column을 추가하였다.

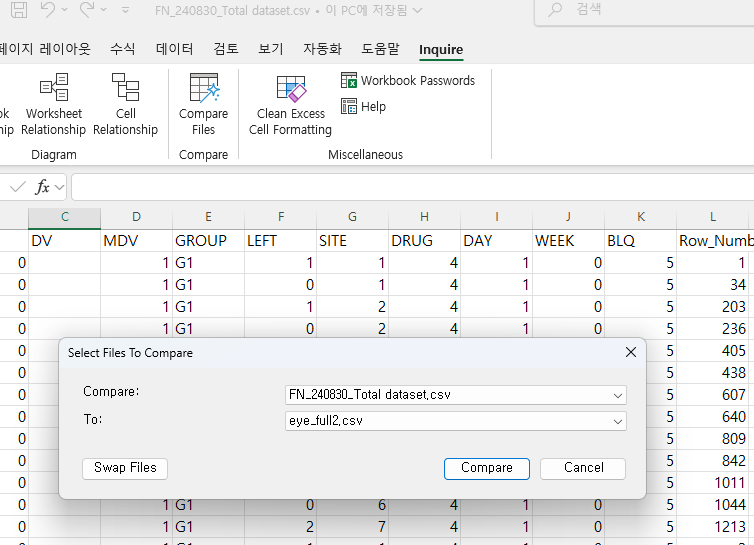
**2. Dataset generation**

Dataset generation은 R (Version 4.3.1), `tidyverse` package (version 2.0.0)을 이용해서 진행하였다. 최종적으로 생성된 dataset은 총 12개의 column, 1547의 row가 생성되었다. 16주차와 24주차 data의 경우 기존의 ID와 겹쳐서 ID에 1000을 더하였다. 나머지 column들의 경우 위에서 명시한 방법으로 생성하였고, ID별로 시간에 따른 순서로 정렬하여 최종적으로 dataset를 생성하였다. Dataset을 생성하는데 사용된 R 코드는 아래에 .txt 파일로 첨부하였다.

**2. Dataset evaluation**

Dataset의 비교는 excel에서 inquire 옵션을 사용하였다. Excel inquire는 excel에서 제공하는 추가 기능으로 서로 다른 excel 파일을 비교할 때 사용하는 추가 기능이다. 서로 다른 2개의 dataset를 비교한 결과 column과 row는 모두 동일하게 관찰되었으나 inquire에서 서로 다른 점을 보고하였다. 첫번째로 측정 위치가 CMT SITE로 서로 다르게 기입이 되었는데 SITE로 모두 동일하게 변경하였다. 두번째로는 임의의 농도 값들이 서로 다르다고 보고되었는데 이는 R에서 처리하는 소수점 차이로 인해 서로 같은 값임에도 불구하고 서로 다른 값이라고 인식한 것으로 보인다. Inquire 상에서 1107개의 서로 다른 row 개수를 보고했는데 Column 이름이 서로 다른 (CMT, SITE)을 제외한 1106는 모두 DV값이 서로 다른 것이었고, 이는 비교하는 dataset을 (filter(MDV==0))과 비교하였을 때 1106의 row로 동일함을 확인하였다. 위의 r코드 생성한 데이터는 소수점이 모두 1자리 수로 보이는 것으로 보아 raw data의 csv로 변환하는 과정과 R tidying 과정에서 소수점관련 에러가 발생한 것으로 보인다. 그 외에는 특별히 다른 점이 관찰되지 않았다. Inquire 결과는 .xlsx 파일로 저장하였다.

**2. Figures**



텍스트, 스크린샷, 소프트웨어, 도표이(가) 표시된 사진

자동 생성된 설명

Excel Inquire과 그 결과 값