

< 건강검진코호트DB Meeting 내용 정리 >

- 4월 21일 Version -

< What To Do >

: Smoking Status 'Never'에서 'Not Response' 분리한 후. Sub-group Analysis 다시 check

< Share & Result >

1) 기존 흡연상태 변수 'SMK-STAT-TYPE-RSPS-CD' 범주 파악

: 'SMK-STAT-RSPS-CD' 변수가 가지는 범주가 ' ' (빈칸) / ' .' / ' 1' (Never)
/ ' 2' (Past) / ' 3' (Current) 이다.

서로 다르게 간주됨 ('.'을 Missing으로 인식하지
않음)

⇒ Study Population 168,338명에서 '흡연상태' 변수 값이 ' ' (빈칸) 인 객체는 359명,

' .' 인 객체는 2명이다 - (1)

반면, Study Population 168,338명 중 Cohort entry Date 이전 검진 기록이 잡히지 않는 객체는 총 4,303명 - (2)

: 이때, (1), (2) 는 다른 경우라고 생각된다. (1)은 'Not Response'로 간주할 수 있으나, (2)는 무응답인지. 아니면 Never / Past / Current Smoker 인지 알 수 없음.

(2) Case를 무응답으로 처리하지 않고 결측으로 처리하면 Coxph model 적합 때 해당 객체의 정보 이용되지 않음.

Question)

(2)의 경우를 어떻게 처리하는 것이 Best 인가? (일단, 'Never Smoker'와 'Non-Response'는 분리하는 것이 맞다고

판단됨)