

<건강검진코호트DB - 음주 / 흡연 / 운동 관련 Covariate Define>

- 24.06.13일 Version

- 작성자 : 이은경

<Back Ground>

- 기존 방식으로 정의한 음주 / 흡연 / 운동 관련 covariate 들의 missing 비율이 10% 이상으로 무시할 수 없는 수치임을 확인.
- literature review 통해 음주 / 흡연 / 운동 관련 covariate 정의 방안 재 수립

<Share & Result>

- 운동 빈도 / MET minutes → *“Proper exercise”* binary 변수로 대체
(출처) Incidence and risk factors for Progression to diabetes mellitus : A retrospective cohort study
 - 외부적 정의
 - (1) 20분 이상의 고강도 운동을 일주일에 3번 이상 수행
 - (2) 30분 이상의 중강도 운동을 일주일에 5번 이상 수행

→ OR 조건 이용해 정의

- 내부적 정의
 - 외부적 정의 (1) 조건을 MET minutes 으로 환산하면 $20 \times 6 \times 3 = 360$ MET minutes이다.
 - 외부적 정의 (2) 조건을 MET minutes 으로 환산하면 $30 \times 3 \times 5 = 450$ MET minutes이다.

∴ 기존에 정의한 MET minutes 값이 450 이상이면 Proper exercise 값을 1로, 아니면 0으로 정의

- 흡연 상태 / 흡연 기간 / 흡연량 / PACK YEAR 변수
(출처) Incidence and risk factors for Progression to diabetes mellitus : A retrospective cohort study
 - 선행 연구에서는 “Current smoker”만 고려하였으나, 기존 Table1 결과 살펴보았을 때 past smoker 관련 정보를 무시할 수 없다고 판단

∴ 기존 흡연 관련 4가지 변수 중, “흡연 상태” 변수만 이용하기로 결정(나머지 변수는 KEEP)

- 음주 빈도 → *“Heavy alcohol drinker”* binary 변수로 대체
(출처) U-Shaped Associations Between Body Weight Changes and Major Cardiovascular Events in Type 2 Diabetes Mellitus: A Longitudinal Follow-up Study of a Nationwide Cohort of Over 1.5 Million
 - 외부적 정의
 - $\geq 30\text{g/day}$ drinker
 - 내부적 정의
 - $\geq 3.5\text{잔/day}$ (TM1_DRKQTY_RSPS_CD 변수 사용)
 - 기존에 정의한 음주 빈도 변수 값이 0 초과인 객체 대상으로 변수 정의

∴ 음주 빈도가 0 초과인 객체 대상, $\geq 3.5(\text{잔})/\text{day}$ 이면 Heavy alcohol drinker 값을 1로, 아니면 0으로 정의