

〈건강검진코호트DB 결과 정리〉

- 2월 17 ~ 19일 Version

- 작성자 : 이은경

〈What TO DO〉

: Cohort entry date 정의하는 code에서 잘못된 부분 발견

→ 수정 후, Study population 재정의 / insulin 처방 이력 재파악

: Exposure 정의한 후, 요약 통계량 / 히스토그램 생성

〈Share & Result〉

1) 잘못된 Code 수정

-- Cohort entry date 정의 때 검진DB에 BMI가 결측인 기록이 있으면 안되는 부분 고려하지 않음
(BLDS만 고려)

1)-①. Cohort entry date가 정의되는 객체 수 : 197,990명 → 197,961명

1)-②. Study population에 포함되는 객체 수 : 168,362명 → 168,339명

-- 변화가 큰 것으로 보이지는 않음.

1)-③. Cohort entry date 정의되는 객체 대상으로,

Cohort entry date 이전에 insulin 처방 이력이 있는 객체는 총 197,961명 중 948명

Cohort entry date 이후에 insulin 처방 이력이 있는 객체는 총 197,961명 중 11,045명

2) Exposure 정의

: 모든 검진기록(검진DB + 생애전환기검진DB) 가져와, Study population의 PERSON_ID / Cohort entry date가 동일한 기록의 BMI 값 가져옴.

→ 이때, 총 168,339명 중 223명은 기록이 2건씩 존재함. / 이유는 검진DB, 생애전환기검진DB 둘 다에 기록되어 있기 때문이라 짐작(해당 기록의 검진 일자, BMI 값이 객체 별로 모두 다 동일함을 확인)

∴ 객체 별 Unique한 기록만 가져옴

3) Exposure의 요약 통계량 / Histogram 생성

(기준) BMI 18.5 미만 underweighted / BMI ≥ 18.5 and < 25 moderate

BMI ≥ 25 and < 30 overweighted / BMI ≥ 30 obesity

3)-①. 요약 통계량

Category	N수	최솟값	평균값	1분위수	중위수	3분위수	최댓값
Underweight	2751	10.818	17.435	17.0391	17.7070	18.1367	18.488
Moderate	90714	18.5	22.727	21.6367	22.9883	24.0273	24.988
Overweight	66841	25.0	26.841	25.7578	26.6367	27.7383	29.976
Obesity	8033	30.0	31.861	30.4766	31.2188	32/4453	103.359

3)-②. Histogram

: Category 별 Histogram / 변수 분포가 정규분포라 가정하고 그린 Density도 겹쳐서 그림.

