

gen - dummies (input options : codeBook, study, Balays)

→ como, atc, outcomes, covid19

과거 twjhe 에서 ← 현재 t에서 이미 search window가 적용되었기 때문에 추가 search window 적용의 필요x
first visit 값에 따라 search window 기준 적용이 필요하다.



현재 outcomes의 경우 df-rule, r.v 적용 라정이 없음

patient-info (전역변수) : Jid 데이터와 first-visit 정보

→ final 2 받음.

① for 문 이용해서 twjhe 200 에 search window 적용. (merge 사용 불가)

변수를 코드값이 기록된 채로 생성

df-rule.	JID MID	Exposure itself	Exposure and Outcome -	Outcome -	...
twjhe 200 기준	1 2313	TRUE	TRUE	TRUE	
Mid를 통한	1 246	FALSE	TRUE	TRUE	
200, 400, 500 적용	1 521	FALSE	FALSE	FALSE	

Exposure-rule confounder-rule

② codeBook 한 줄씩 읽기 → 여러 기준으로 질병/약을 다르게 합쳐 frequency 필요.

code-val = N17

uniqueCode = ID2001.N17 → 중복을 없애고 uniqueCode 3 변수 생성.

code-len = 3 → 코드 길이는 항상 다름.

r.v = "exposure and outcome related" → 해당 스타디움에서 이 코드의 역할
(df-rule 에서 규칙 적용에 사용)
como 와 atc 만 주재

- 1) final 에 all zero 면서 uniqueCode 이름을 주는 변수 생성
- 2) r.v 규칙 중 true mid 중 N17 기록된 Jid 찾기 → searchJID에 저장
- 3) patient-info 에서 searchJID에 포함되는 모든 uniqueCode 변수가 1 이 되도록 변경.
- 4) 반복 후 final return.

③ 원하는 데이터셋에서 찾는 라벨 자세히

* 현재 약 코드는 모두 ATC 이다. 해당 ATC의 DIV-CD (제통 코드) 혹은 GNL-CD (주성분 코드) 를 찾는 라벨이 추가될 필요하다.

(→ rule 변수에 저장 (ATC-mapping-df 에서 code_val 에 해당하는 주성분 코드만 생성)
영정보 저장)

SF-split[[1]][1] = col9-twjhe200-trans-dn / MAIN-SICK

SF-split[[1]][2] = col9-twjhe200-trans-dn / SUB-SICK

SF-split[[1]][3] = NULL.

df-name = col9-twjhe200-trans-dn, var-name = MAIN-SICK

df.all → get, df.all 에서 r.v가 true 인 mid를 가지고 와서 df 생성
(r.v가 없다면 df 와 df.all 은 같다.)

① ATC는 code_val 의 길이다 상단영역이 늘 8자리인 값을 찾아야 하므로
예외처리 rule 이 true 인 주성분 코드를 mapped_var1 에 저장하며
df-name 의 var-name 에서 search.

② 이외의 코드 타입은 code_len 에 해당하는 길이 만큼

df-name 의 var-name 을 걸어야 한다. 그 후 search

→ 두 경우 모두, 찾은 JID 를 search JID 에 저장하고

final 에서 search JID 는 12 번정

→ [2] 3 같은 라벨 반복