

〈건강검진코호트DB 결과 정리〉

- 2월 20일 Version

- 작성자 : 이은경

〈What TO DO〉

: Disease onset date - Cohort entry date간 FU_Duration 분포 파악

: Cohort entry date - Outcome 발생 시점 간 FU_Duration 분포 파악

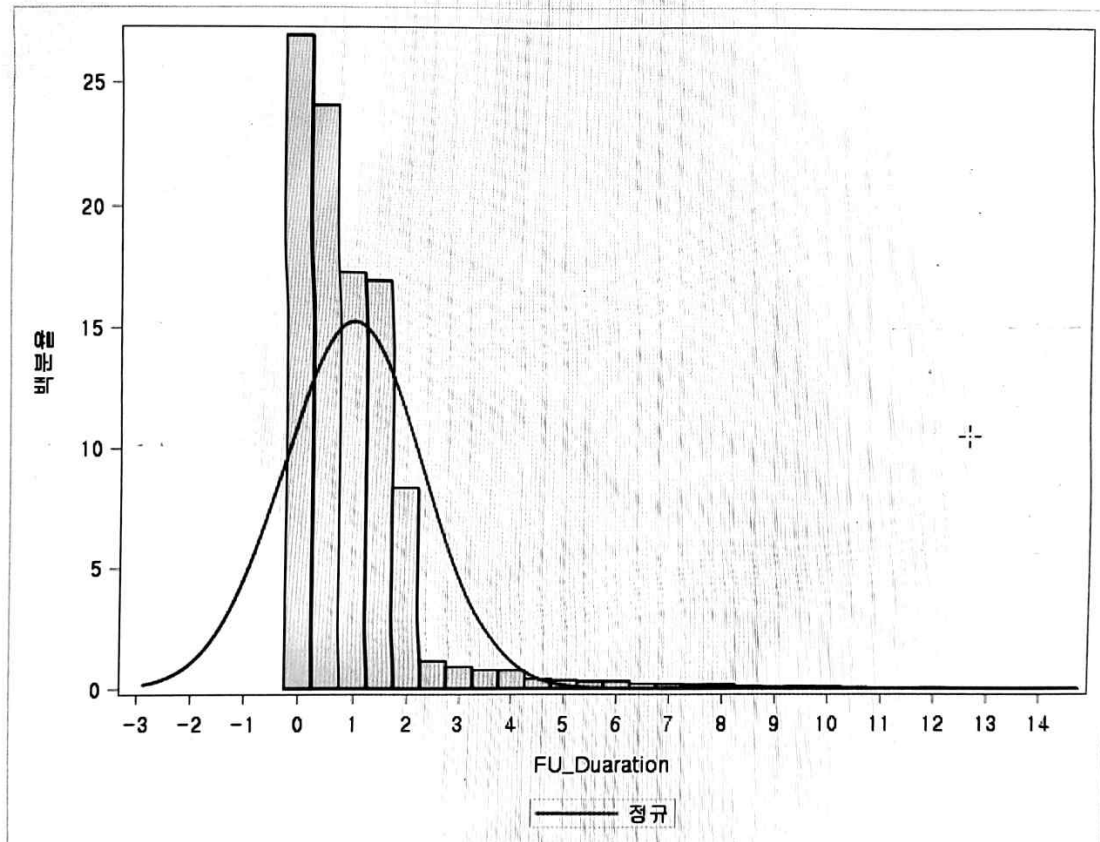
〈Share & Result〉

1) Disease onset date와 Cohort entry date간 Duration 정의한 후, 분포 파악 (단위는 year)

[요약 통계량]

Duration	N수	평균값	최솟값	1분위수	중앙값	3분위수	최댓값
	168,339	1.010	0	0.183	0.728	1.44	14.543

[Histogram]



2) Outcome 정의한 후 얻은 table 일부

	PERSON_ID	Death	MI	Cancer	Heart_failure	TIA_Stroke	Outcome_Duration	Outcome_Category
1	10000355	1	0	0	0	0	5,3196440794	Death
2	10000821	1	0	0	0	0	3,9288158795	Death
3	10002153	1	0	0	0	0	5,8535249829	Death
4	10003229	1	0	0	0	0	14,020533881	Death
5	10003681	1	0	0	0	0	7,7754962355	Death
6	10005231	1	0	0	0	0	8,9911019849	Death
7	10006251	1	0	0	0	0	6,5270362765	Death
8	10006326	1	0	0	0	0	6,3436002738	Death
9	10006642	1	0	0	0	0	9,697467488	Death
10	10007049	1	0	0	0	0	15,26899384	Death
11	10007259	1	0	0	0	0	14,069815195	Death
12	10007371	0	0	0	0	0	14,67761807	Last_follow
13	10007707	1	0	0	0	0	7,5099247091	Death
14	10008531	1	0	0	0	0	5,6399726215	Death
15	10008871	1	0	0	0	0	5,550992471	Death
16	10008958	0	0	0	0	0	16,424366872	Last_follow
17	10009574	0	0	0	0	0	12,128678987	Last_follow
18	10010182	1	0	0	0	0	7,9863107461	Death
19	10011912	1	0	0	0	0	7,40862423	Death
20	10012357	1	0	0	0	0	3,6577686516	Death
21	10013746	1	0	0	0	0	2,5160848734	Death
22	10014968	1	0	0	0	0	12,969199179	Death
23	10015045	0	0	0	0	0	16,164271047	Last_follow
24	10015520	1	0	0	0	0	12,892539357	Death
25	10015873	1	0	0	0	0	5,0814510609	Death
26	10016420	1	0	1	0	0	0,038329911	Cancer
27	10017110	1	0	0	0	0	8,8432580424	Death
28	10017760	1	0	0	0	0	0,8925393566	Death
29	10018712	1	0	0	0	0	9,1608487337	Death
30	10019300	1	0	0	0	0	2,5352498289	Death
31	10019620	1	0	0	0	0	13,273100616	Death
32	10020737	1	0	0	0	0	3,0088980151	Death
33	10021625	1	0	0	0	0	4,2984257358	Death
34	10021646	1	0	0	0	0	5,28678987	Death
35	10021672	0	0	0	0	0	16,131416838	Last_follow
36	10023224	1	0	0	0	0	13,396303901	Death

- “Outcome_Duration”은 Cohort entry date부터 Outcome 발생까지 걸린 시간을 의미하며, 한 객체 당 여러 건의 Outcome이 발생한 경우, 가장 과거의 시점만 고려
- “Outcome_category”는 각 객체의 Outcome 종류를 의미하며, 가질 수 있는 값은 “Death”, “MI”, “Cancer”, “TIA_or_Stroke”, “Heart_failure”, “Last_follow”이다. / “Last_follow”는 추적 종료일인 2019.12.31.일까지 어떠한 Outcome도 발생하지 않은 사람을 뜻한다.

2)-① Study population에 포함되는 객체 대상, Cohort entry date와 Outcome 발생까지의 Duration 정의한 후, 분포 파악(Outcome이 발생하지 않은 객체의 경우 last Follow Up date는 2019.12.31.일)
(단위는 year)

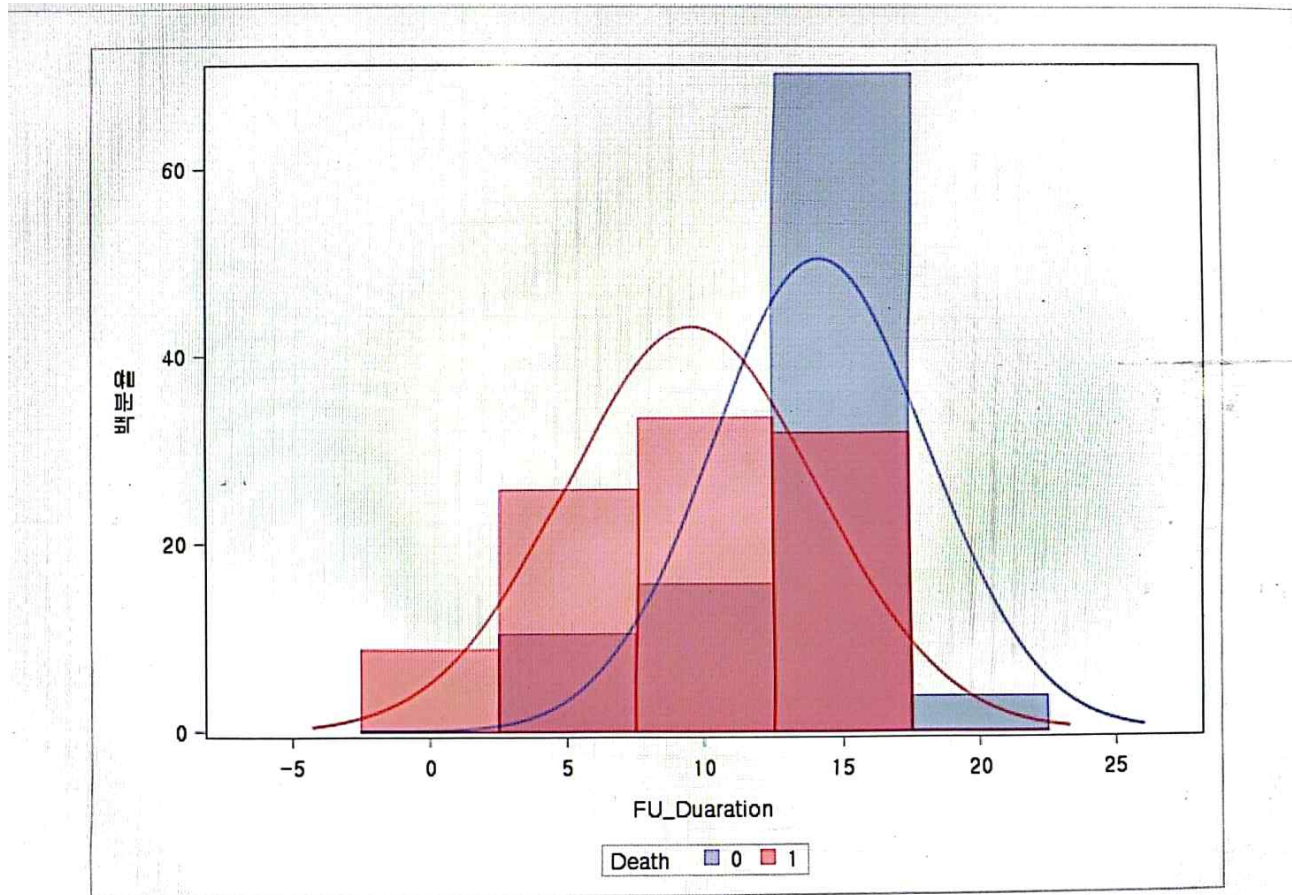
Duration	N수	평균값	최솟값	1분위수	중앙값	3분위수	최댓값
	168,339	13.472	-0.238	10.943	16.024	17.002	17.949

2)-②. Study population에 포함되는 객체 대상, Cohort entry date와 Death Outcome 발생까지 Duration 정의한 후, 분포 파악(Outcome이 발생하지 않은 객체의 경우 last Follow Up date는 2019.12.31.일) (단위는 year)

[요약 통계량]

Death	N수	평균값	최솟값	1분위수	중앙값	3분위수	최댓값
0(발생안함)	143,965	14.230	2.997	12.172	16.175	17.067	17.949
1(발생함)	24,374	9.5069	-0.238	5.798	9.875	13.489	17.620

[Histogram]

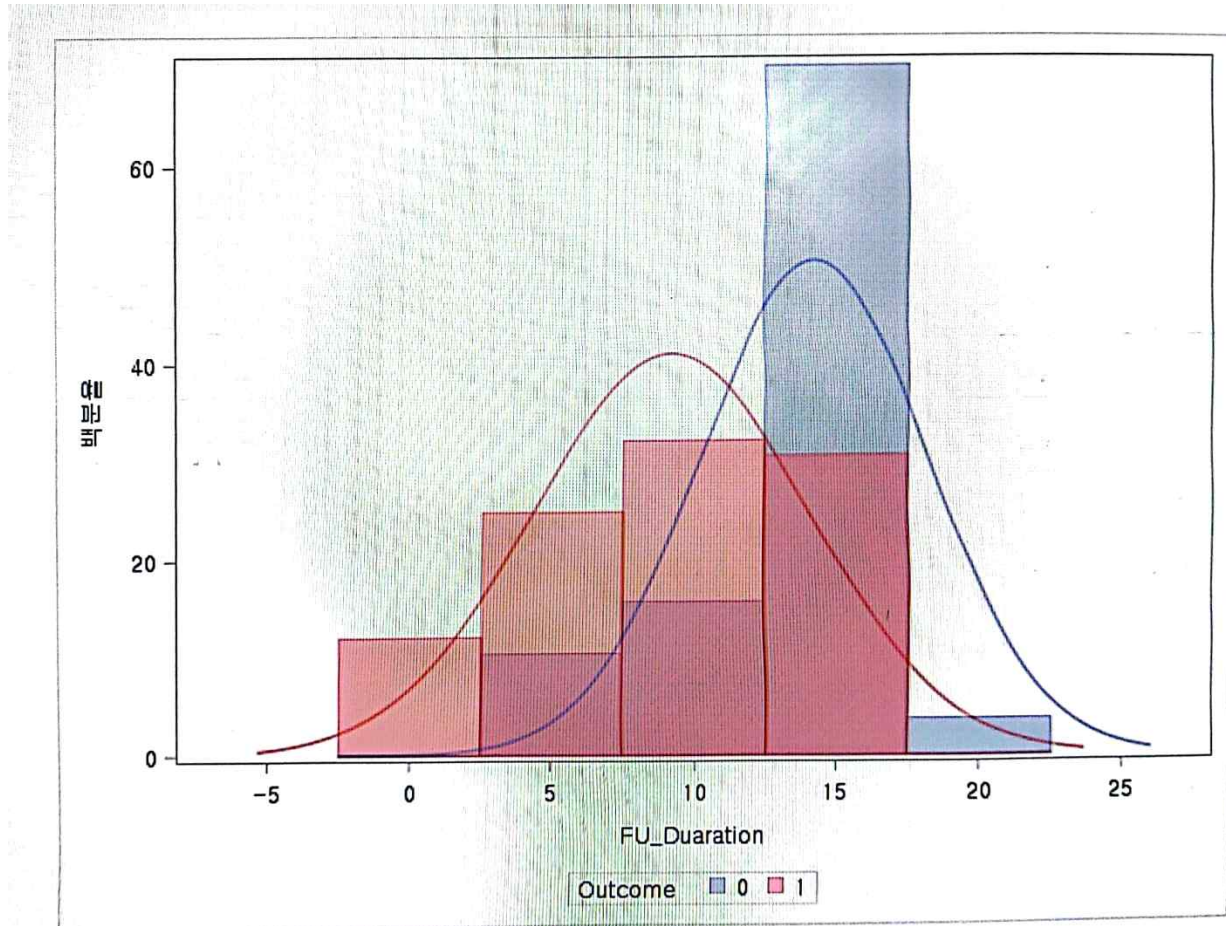


2)-③. Study population에 포함되는 객체 대상, Cohort entry date와 고려하는 모든 Outcome 발생까지 Duration 정의한 후, 분포 파악(Outcome이 발생하지 않은 객체의 경우 last Follow Up date는 2019.12.31.일) - Outcome=1은 고려하는 모든 Outcome 중 하나라도 발생함을 의미. (단위는 year)

[요약 통계량]

Outcome	N수	평균값	최솟값	1분위수	중앙값	3분위수	최댓값
0(발생안함)	143,251	14.227	2.997	12.167	16.175	17.067	17.949
1(발생함)	25,088	9.164	-0.238	5.258	9.587	13.355	17.6208

[Histogram]



2)-④. 고려하는 Outcome 중 하나라도 발생한 객체들만을 대상으로, Cohort entry date와 Outcome 발생 시점까지의 Duration의 Histogram 생성 - 이때, “Outcome_category” 변수 기준으로 grouping

[Histogram]

