

흡연 유무에 따른 천식 연관성

[가설] 흡연 유무에 따른 천식 연관성

1. 배경 및 연구 동기

최근 알레르기 질환에 대한 연구를 진행하는 과정에서 보다 구체적이고 심층적인 질병 분석의 필요성을 느끼게 되었고, 이에 따라 **천식**에 초점을 맞추어 연구를 진행하게 됨. 천식은 폐속 기관지가 매우 예민해져 좁아지는 상태로, 숨이 차고 쌉쌉거리는 숨소리와 심한 기침 등의 증상을 유발하는 질환임. 이는 환자의 삶의 질에 큰 영향을 미칠 수 있으며, 특히 고령층에서는 더욱 심각한 문제로 대두되고 있음.

천식과 관련된 최신 통계 자료에 따르면, **코로나19 팬데믹 이후 천식 환자가 급증한 것으로 확인됨**. 2023년 1월부터 8월까지 천식 환자 수는 약 **142만 3,451명**으로, 이는 2022년 한 해 동안 기록된 **86만 7,642명** 대비 39% 증가한 수치를 확인 가능함. 이 중에서도 **50~80대 고령층에서 남녀 구분 없이 약 45%**의 환자 증가가 확인됨. 이러한 통계는 천식 발병의 주요 원인과 연관 요인에 대한 구체적인 이해와 연구의 필요성을 강조함.

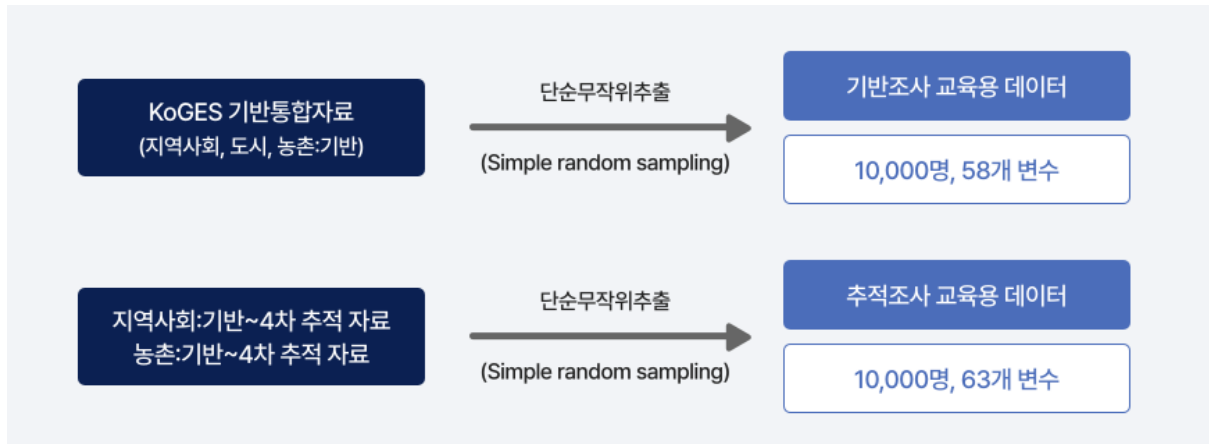
특히, 천식은 단순한 환경적 요인뿐 아니라 흡연, 음주, 간접흡연 등과 같은 생활습관 요인과도 밀접하게 연결되어 있음. 따라서 천식의 발병 원인을 파악하고 주요 요인을 분석함으로써 예방 및 관리 방안 마련에 기여하고자 본 연구를 기획함.

건강이 최고

코로나19 끝나자 '천식' 왔다... 고령 환자 45% 급증



2. 데이터 정보



[출처: 국립보건연구원]

본 연구는 **국립보건연구원의 교육용 데이터**를 활용하여 진행함. 해당 데이터는 KoGES 역학 자료를 기반으로 연구자의 자료 이해도를 높이기 위해 제공된 교육용 데이터로, 10,000명(기반조사) 및 1,000명(추적조사)을 단순 무작위 추출하여 생성됨. 데이터는 일부 변형되어 제공되며, 약 60개의 변수로 구성되어 있음. 천식과 관련된 주요 원인을 분석하기 위해 다음과 같은 변수를 활용함.

종속변수 (Y)

1. 천식 진단 경험 여부 (AS1_08_DISEASE)
2. 과거 만성폐질환 진단 여부 (AS2_07_DISEASE)
3. 알레르기 질환 치료 여부 (AS3_TREAT)

독립변수 (X)

1. 흡연 여부 (AS1_03_DRSM, AS2_03_DRSM)
2. 간접흡연 여부 (AS1_Psm, AS2_Psm)
3. 음주 여부 (AS1_Drink, AS2_Drink)
4. 연령 (AGE)
5. 성별 (AS1_Sex)

이 데이터는 천식의 발병 원인과 생활습관 요인 간의 관계를 이해하기에 적합하며, 나이 및 성별에 따른 천식의 위험도를 분석함으로써 고위험군에 대한 효과적인 예방 방안을 제시할 수 있는 기초 자료로 활용됨.

3. 분석 방법

1) 분석 대상자 선정 방법 (유병률 및 제외된 대상자 수와 최종 분석 대상자 수)

- 2기에서 많은 결측치 발견

⇒ 1~2기를 제외한 3기~10기의 데이터만 활용하여 분석 진행

- 3~10기 총 변수 7515

 fcf 7515

- 날짜 유형 변경 및 계산 진행

```
> fcf[1:10, ] %>% select(AS3_EDATE3, edate3, edate3_n)
  AS3_EDATE3   edate3 edate3_n
1    200509 2005-09-01    13027
2    200505 2005-05-01    12904
3    200505 2005-05-01    12904
4    200505 2005-05-01    12904
5    200505 2005-05-01    12904
6    200505 2005-05-01    12904
7    200505 2005-05-01    12904
8    200505 2005-05-01    12904
9    200505 2005-05-01    12904
10   200505 2005-05-01    12904
> fcf[1:10, ] %>% select(AS10_EDATE10, edate10, edate10_n)
  AS10_EDATE10   edate10 edate10_n
1    201903 2019-03-01    17956
2    201904 2019-04-01    17987
3    201904 2019-04-01    17987
4    201908 2019-08-01    18109
5    201908 2019-08-01    18109
6    201908 2019-08-01    18109
7    201908 2019-08-01    18109
8    201904 2019-04-01    17987
9    201905 2019-05-01    18017
10         NA      <NA>         NA
```

- 기반조사 기준 천식 유병자 제외 : 7504명

천식, 만성폐질환, 알레르기 변수 세 개 활용(천식약은 3000개 이상의 결측으로 제외)

```

> sum(is.na(fcf$AS3_PDFAS))#결측 10명- 천식
[1] 10
> sum(is.na(fcf$AS3_PDFCL)) #결측 10명- 만성폐질환
[1] 10
> sum(is.na(fcf$AS3_TRTAL)) #결측 10명- 알레르기 질환
[1] 10
> #sum(is.na(fcf$AS3_DRUGAS)) #결측 3545명 -천식약 복용 **제외**

```

- 기반조사 천식 유병자 제외 : 7429명

```

> freq(fcf2$atm1) #유병율: 1.00% ( 75 )
fcf2$atm1
      Frequency  Percent
0             7429   99.0005
1              75    0.9995
Total          7504 100.0000

```

- 최종 분석 대상자 : 7429명

```

#기반조사 당시 천식 유병 27명 제외
fcf3<-subset(fcf2, atm1==0) # N= 7429

```

2) 사용 변수 종류 및 변수 처리 방법

- 흡연, 음주, 간접흡연 결측 제거 : 총 2명의 결측 제거

```

sum(is.na(fcf3$AS3_SMOKE)) #결측수 2명
[1] 2
fcf4<- fcf3 %>% filter(! (is.na(AS3_SMOKE))) #N=7427
#간접흡연변수에 결측있는 대상자수
sum(is.na(fcf3$AS3_PSM)) #결측수 0명
[1] 0
fcf5<- fcf4 %>% filter(! (is.na(AS3_PSM)))
#음주변수 에 결측있는 대상자수
sum(is.na(fcf3$AS3_DRINK)) #결측수0명
[1] 0

```

- 천식, 만성폐질환, 알레르기를 동시에 반응하는 새로운 변수 sdbp3~sdbp10생성
- 한 번이라도 질병에 걸렸을 경우 1로 분류

```
# 각 추적 차수별 천식 발생 여부 변수 생성 -> 추적관찰 기간동안 천식 진단 여부 변수 결측 제거
fcf7 <- fcf6 %>% mutate(sdbp3 = ifelse((AS3_PDFAS == 2 | AS3_PDFCL ==2 | AS3_TRTAL ==2), 1, 0),
  sdbp4 = ifelse((AS4_ASTH == 2 | AS4_CLD ==2 | AS4_TREATD4 ==2 ), 1, 0),
  sdbp5 = ifelse((AS5_TREATD11==2 | AS5_TREATD27 ==2 | AS5_ALLER==2), 1, 0),
  sdbp6 = ifelse((AS6_TREATD11==2 | AS6_TREATD32==2 | AS6_ALLER==2), 1, 0),
  sdbp7 = ifelse((AS7_TREATD11==2 | AS7_CLD==2 | AS7_ALLER==2), 1, 0),
  sdbp8 = ifelse((AS8_TREATD11==2 | AS8_TREATD32==2 | AS8_ALLER==2), 1, 0),
  sdbp9 = ifelse((AS9_TREATD11==2 | AS9_TREATD32 ==2 | AS9_ALLER==2), 1, 0),
  sdbp10 = ifelse((AS10_TREATD11 == 2 | AS10_TREATD32 == 2 | AS10_ALLER == 2), 1, 0))
```

- 5기 데이터의 “만성폐질환”이 양성자료만 존재 -> NA결측 수 증가

▶ 현재 치료질환_만성 폐 질환 치료 여부	1=아니오, 2=예	범주형	●	●	x	x
----------------------------	------------	-----	---	---	---	---

- 5기의 데이터는 모든 NA를 전부 0으로 변경하여 진행

```
> fcf7$sdbp5[is.na(fcf7$sdbp5)]<-0
> freq(fcf7$sdbp5)
fcf7$sdbp5
      Frequency Percent
0           7347  98.923
1             80   1.077
Total        7427 100.000
```

- 결측치 NA 제거 : 그대로 결측 제거하기엔 결측 수가 너무 많음

```
freq(fcf7$sdbp3)
freq(fcf7$sdbp4)
summary(freqlist(~sdbp4 + AS4_ASTH + AS4_CLD + AS4_TREATD4, data=fcf7)) #결측 1246
freq(fcf7$sdbp5)
summary(freqlist(~sdbp5+ AS5_TREATD11+ AS5_TREATD27 + AS5_ALLER , data=fcf7)) #결측80
freq(fcf7$sdbp6)
summary(freqlist(~sdbp6+ AS6_TREATD11+ AS6_TREATD32 + AS6_ALLER, data=fcf7)) #결측 4356
freq(fcf7$sdbp7)
summary(freqlist(~sdbp7+ AS7_TREATD11+ AS7_CLD + AS7_ALLER, data=fcf7)) #결측4703
freq(fcf7$sdbp8)
summary(freqlist(~sdbp8+ AS8_TREATD11+ AS8_TREATD32 + AS8_ALLER, data=fcf7)) #결측4497
freq(fcf7$sdbp9)
summary(freqlist(~sdbp9+ AS9_TREATD11+ AS9_TREATD32 + AS9_ALLER, data=fcf7)) #결측4544
freq(fcf7$sdbp10)
summary(freqlist(~sdbp10 +AS10_TREATD11+ AS10_TREATD32 + AS10_ALLER, data=fcf7)) #4701
#p.91 NA->0처리?
```

- 이들 중 처음부터 조사에 참여 안 한 대상자들은 통계 분석 대상으로 포함 시킴

```

#sdbp5 NA만 0처리
fcf7$sdbp5[is.na(fcf7$sdbp5)]<-0
freq(fcf7$sdbp5)

fcf7_1<-fcf7 %>% filter(!is.na(sdbp3) & !is.na(edate3_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp3 + AS3_PDFAS + AS3_PDFCL + AS3_TRTAL + AS3_EDATE3 + edate3, data=fcf7_1)) #확인

freq(fcf7_1$sdbp4)#NA = 1246
summary(freqlist(~sdbp4 + AS4_ASTH + AS4_CLD + AS4_TREATD4 + AS4_EDATE4 + edate4, data=fcf7_1)) #결측 12
fcf7_2<-fcf7_1 %>% filter(!is.na(sdbp4) & !is.na(edate4_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp4 + AS4_ASTH + AS4_CLD + AS4_TREATD4 + AS4_EDATE4 + edate4, data=fcf7_2)) #확인

freq(fcf7_2$sdbp5) #NA=0
summary(freqlist(~sdbp5+ AS5_TREATD11+ AS5_TREATD27 + AS5_ALLER+AS5_EDATE5 + edate5, data=fcf7_2)) #0
fcf7_3<-fcf7_2 %>% filter(!is.na(sdbp5) & !is.na(edate5_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp5 + AS5_TREATD11+ AS5_TREATD27 + AS5_ALLER+AS5_EDATE5 + edate5, data=fcf7_3)) #확인

freq(fcf7_3$sdbp6) #NA=4350
summary(freqlist(~sdbp6+ AS6_TREATD11+ AS6_TREATD32 + AS6_ALLER+ AS6_EDATE6 + edate6, data=fcf7_3)) #결측 355
fcf7_4<-fcf7_3 %>% filter(!is.na(sdbp6) & !is.na(edate6_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp6+ AS6_TREATD11+ AS6_TREATD27 + AS6_ALLER+ AS6_EDATE6 + edate6, data=fcf7_4)) #확인

freq(fcf7_4$sdbp7) #NA=2042
summary(freqlist(~sdbp7+ AS7_TREATD11+ AS7_CLD + AS7_ALLER+ AS7_EDATE7 + edate7, data=fcf7_4)) #결측 243
fcf7_5<-fcf7_4 %>% filter(!is.na(sdbp7) & !is.na(edate7_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp7+ AS7_TREATD11+ AS7_TREATD27 + AS7_ALLER+ AS7_EDATE7 + edate7, data=fcf7_5)) #확인

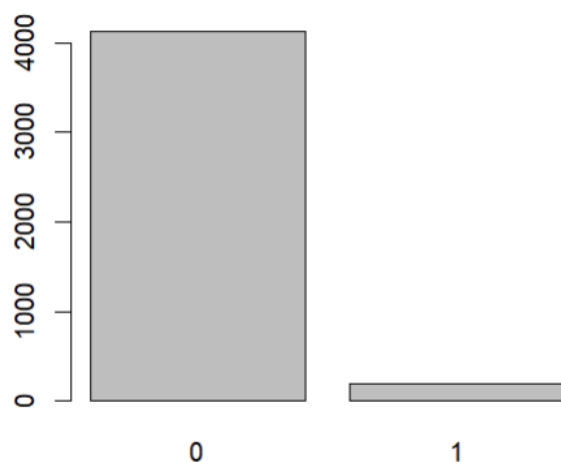
freq(fcf7_5$sdbp8) #NA=1597
summary(freqlist(~sdbp8+ AS8_TREATD11+ AS8_TREATD32 + AS8_ALLER+ AS8_EDATE8 + edate8, data=fcf7_5)) #결측 108
fcf7_6<-fcf7_5 %>% filter(!is.na(sdbp8) & !is.na(edate8_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp8+ AS8_TREATD11+ AS8_TREATD32 + AS8_ALLER+ AS8_EDATE8 + edate8, data=fcf7_6)) #확인

freq(fcf7_6$sdbp9) #NA=1550
summary(freqlist(~sdbp9+ AS9_TREATD11+ AS9_TREATD32 + AS9_ALLER+ AS9_EDATE9 + edate9, data=fcf7_6)) #결측 32
fcf7_7<-fcf7_6 %>% filter(!is.na(sdbp9) & !is.na(edate9_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp9+ AS9_TREATD11+ AS9_TREATD32 + AS9_ALLER+ AS9_EDATE9 + edate9, data=fcf7_7)) #확인

freq(fcf7_7$sdbp10) #NA=1683
summary(freqlist(~sdbp10 +AS10_TREATD11+ AS10_TREATD32 + AS10_ALLER+ AS10_EDATE10 + edate10, data=fcf7_7)) #결
fcf7_8<-fcf7_7 %>% filter(!is.na(sdbp10) & !is.na(edate10_n))) #조사 안되서 모르는 것이 아닌 조사되었는데도 모르는
summary(freqlist(~sdbp10+ AS10_TREATD11+ AS10_TREATD32+ AS10_ALLER+ AS10_EDATE10 + edate10, data=fcf7_8)) #확인

```

- 최종 분석 변수 : N=4331



```

> freq(fcf8$atm_d) #
fcf8$atm_d
Frequency Percent
0          4131  95.382
1           200   4.618
Total       4331 100.000

```

- 총 추적기간동안 발생 여부인 새로운 변수 atm_d생성

```
# 총 추적 기간 동안 천식 발생 여부 변수 생성
fcf8 <- fcf7_8 %>% mutate(atm_d = ifelse((sdbp4 == 1 |
                                         sdbp5 == 1 | sdbp6 == 1 |
                                         sdbp7 == 1 | sdbp8 == 1 |
                                         sdbp9 == 1 | sdbp10 == 1 ),1 , 0))

> fcf8[100:105,] %>% select(sdbp4 ,sdbp5 ,sdbp6,sdbp7,sdbp8, sdbp9, sdbp10, atm_d)
  sdbp4 sdbp5 sdbp6 sdbp7 sdbp8 sdbp9 sdbp10 atm_d
100    0    0    0    NA    0    0    0    NA
101    0    0    0    0    0    0    0    0
102   NA    1    1    NA    1    NA    NA    1
103    0    0    0    0    0    0    0    0
104    0    0    0    0   NA   NA   NA   NA
105   NA    0    0    0    0    0    0    NA
```

- 보정변수 범주화 / 흡연, 간접흡연, 음주 범주화 / 성별은 따로 범주화 진행 안함

```
> freq(fcf9$drink_yn)
fcf9$drink_yn
  Frequency Percent
0         2409    55.62
1         1922    44.38
Total       4331   100.00

> #smoke_yn
> fcf10<-fcf9 %>% mutate(smoke_yn=ifelse(AS3_SMOKE==3,1,0))
> freq(fcf10$smoke_yn)
fcf10$smoke_yn
  Frequency Percent
0         3456    79.8
1          875    20.2
Total       4331   100.0

> #psm_yn
> fcf11<-fcf10 %>% mutate(psm_yn=ifelse(AS3_PSM==2,1,0))
> freq(fcf11$psm_yn)
fcf11$psm_yn
  Frequency Percent
0         3114    71.9
1         1217    28.1
Total       4331   100.0
```

- 관찰기간 생성 : 연구종료 시점 - 연구 시작시점

```
end start f_time_d
1 17956 13027 4929
2 17987 12904 5083
3 17987 12904 5083
4 18109 12904 5205
5 18109 12904 5205
6 16770 12904 3866
7 18109 12904 5205
8 17987 12904 5083
9 18017 12904 5113
10 15918 12904 3014
```


	sdbp4	sdbp5	sdbp6	sdbp7	sdbp8	sdbp9	sdbp10	edate4_n	edate5_n	edate6_n	edate7_n	edate8_n	edate9_n	edate10_n	max_date	end_date	end	start	f_time_d
1	0	0	0	0	0	0	0	13604	14365	15095	15796	16526	17226	17956	17956	17956	17956	13027	4929
2	0	0	0	0	0	0	0	13634	14396	15095	15826	16709	17348	17987	17987	17987	17987	12904	5083
3	0	0	0	0	0	0	0	13634	14396	15065	15826	16709	17348	17987	17987	17987	17987	12904	5083
4	0	0	0	0	0	0	0	13634	14549	15034	15826	16679	17318	18109	18109	18109	18109	12904	5205
5	0	0	0	0	0	0	0	13634	NA	15034	15918	16679	17318	18109	18109	18109	18109	12904	5205
6	0	0	0	0	0	NA	0	13634	14549	15126	15826	16770	NA	18109	18109	16770	16770	12904	3866
7	0	0	0	0	0	0	0	13634	14549	15034	15826	16770	17379	18109	18109	18109	18109	12904	5205
8	0	0	0	0	0	0	0	13634	14426	15034	15826	16617	17257	17987	17987	17987	17987	12904	5083
9	0	0	0	0	0	0	0	13634	14426	15034	15826	16679	17318	18017	18017	18017	18017	12904	5113
10	0	0	0	0	NA	NA	NA	13634	14426	15095	15918	NA	NA	NA	15918	15918	15918	12904	3014

- 관찰기간 (년) 생성

```
> fd1_n[1:10,] %>% select(DIST_ID, atm_d, start, end, f_time_d, f_time_y)
      DIST_ID atm_d start  end f_time_d f_time_y
1 NIH23M5510008    0 13027 17956    4929    13.5
2 NIH23M5231052    0 12904 17987    5083    13.9
3 NIH23M5681460    0 12904 17987    5083    13.9
4 NIH23M5246270    0 12904 18109    5205    14.3
5 NIH23M5361589    0 12904 18109    5205    14.3
6 NIH23M5111272    0 12904 16770    3866    10.6
7 NIH23M5718245    0 12904 18109    5205    14.3
8 NIH23M5831186    0 12904 17987    5083    13.9
9 NIH23M5737829    0 12904 18017    5113    14.0
10 NIH23M5637754    0 12904 15918    3014     8.3
```

3) 통계 분석 방법

- crosstable과 chisquare를 활용한

천식유무 별 흡연,간접흡연,음주,성별 분류

[음주]

```
=====
              fd1_n$atm_d
fd1_n$drink_yn    0      1  Total
-----
0              2295    114   2409
              0.003    0.068
              0.953    0.047   0.556
              0.556    0.570
              0.530    0.026
-----
1              1836     86   1922
              0.004    0.086
              0.955    0.045   0.444
              0.444    0.430
              0.424    0.020
-----
Total              4131    200   4331
              0.954    0.046
=====
```

```
> chisq.test(fd1_n$drink_yn, fd1_n$atm_d)
```

Pearson's Chi-squared test with Yates' continuity correction

data: fd1_n\$drink_yn and fd1_n\$atm_d
X-squared = 0.10804, df = 1, p-value = 0.7424

[흡연]

fd1_n\$smoke_yn	fd1_n\$atm_d		Total
	0	1	
0	3295	161	3456
	0.001	0.012	
	0.953	0.047	0.798
	0.798	0.805	
	0.761	0.037	
1	836	39	875
	0.002	0.049	
	0.955	0.045	0.202
	0.202	0.195	
	0.193	0.009	
Total	4131	200	4331
	0.954	0.046	

```
> chisq.test(fd1_n$smoke_yn,fd1_n$atm_d)
```

Pearson's Chi-squared test with Yates' continuity correction

data: fd1_n\$smoke_yn and fd1_n\$atm_d
X-squared = 0.026712, df = 1, p-value = 0.8702

[간접흡연]

fd1_n\$psm_yn	fd1_n\$atm_d		Total
	0	1	
0	2978	136	3114
	0.020	0.423	
	0.956	0.044	0.719
	0.721	0.680	
	0.688	0.031	
1	1153	64	1217
	0.052	1.083	
	0.947	0.053	0.281
	0.279	0.320	
	0.266	0.015	
Total	4131	200	4331
	0.954	0.046	

```
> chisq.test(fd1_n$psm_yn,fd1_n$atm_d)
```

Pearson's Chi-squared test with Yates' continuity correction

data: fd1_n\$psm_yn and fd1_n\$atm_d
X-squared = 1.3829, df = 1, p-value = 0.2396

[성별]

```
=====
fd1_n$AS3_SEX      fd1_n$atm_d
                   0      1      Total
-----
1                   1863      83      1946
                   0.025      0.524
                   0.957      0.043      0.449
                   0.451      0.415
                   0.430      0.019
-----
2                   2268      117      2385
                   0.021      0.428
                   0.951      0.049      0.551
                   0.549      0.585
                   0.524      0.027
-----
Total               4131      200      4331
                   0.954      0.046
=====
> chisq.test(fd1_n$AS3_SEX,fd1_n$atm_d)

Pearson's Chi-squared test with Yates' continuity correction

data: fd1_n$AS3_SEX and fd1_n$atm_d
X-squared = 0.85798, df = 1, p-value = 0.3543
```

● 직접흡연과 간접흡연을 활용한 log-rank 구현

```
> # log-rank test
> log_rank <- survdiff(y ~ smoke_yn, data = fd1_n)
> log_rank$pvalue
[1] 0.9166884
> # log-rank test
> log_rank <- survdiff(y ~ psm_yn, data = fd1_n)
> log_rank$pvalue
[1] 0.3448442
```

직접흡연의 경우 p-값이 0.9167로 0.05보다 높아, 흡연 여부에 따라 생존율에 유의미한 차이가 존재하지 않으며

간접흡연 또한 p-값이 0.3448로 0.05보다 높아, 흡연 여부에 따라 생존율에 유의미한 차이가 존재하지 않는다

➔ 이는 직접흡연, 간접흡연 여부 둘다 천식 발생에 큰 영향을 미치지 않는다는 것을 확인 할 수 있음

3. 결과

1) 표 1과 해석 (대상자 특성 및 고려한 변수들)

	천식진단여부		P
	NO	YES	
Smoke_yn			0.8702
NO	3295(79.8)	161(80.5)	
YES	836(20.2)	39(19.5)	
Psm_yn			0.2396
NO	2978(72.1)	136(68.0)	
YES	1153(27.9)	64(32.0)	
Drink_yn			0.7424
NO	2295(55.6)	114(57.0)	
YES	1836(44.4)	86(43.0)	
SEX			0.3543
남자	1863(45.1)	83(41.5)	
여자	2268(54.9)	117(58.5)	

[흡연]

흡연을 안하는 사람들 중 안걸린 사람이 79.8 걸린사람이 80.5로 둘 차이가 거의 없지만 흡연을 안하는 사람이 더 높으며

흡연을 하지만 안걸린 사람의 경우 20.2, 걸린사람의 경우 19.5로 흡연을 하는 사람 또한 차이가 거의 없지만 흡연을 안하는 사람이 더 높다

이에 p-value값은 0.8702로 0.05보다 큰 값을 가짐으로써 통계적으로 유의미하지 않다

[간접흡연]

간접흡연을 안한 사람들의 경우 걸린사람이 68이고 안걸린사람이 72.1로 안걸린 사람이 더 높으며

간접흡연을 할 경우 천식에 안걸린 사람이 27.9걸린사람이 32.0으로 걸린사람의 비율이 더 높다

이에 p-value값은 0.2396으로 0.05보다 큰 값을 가짐으로 통계적으로 유의미하지 않다

[간접흡연 vs 직접흡연]

둘다 천식에는 큰 영향을 미치진 않지만 간접흡연과 직접흡연을 비교 했을 때

간접흡연을 할 경우 직접흡연보다 천식에 걸릴 확률이 더 높게 나온다는 것을 확인 할 수 있다

[음주]

음주를 안하는 사람들 중 천식에 안 걸린 사람이 55.6, 천식에 걸린 사람이 57.0으로 천식에 걸린 사람의 비율이 더 높으며

음주를 하는 사람들 중 천식에 안 걸린 사람이 44.9, 천식에 걸린 사람이 43.0으로 천식에 안 걸린 사람이 비율이 더 높다

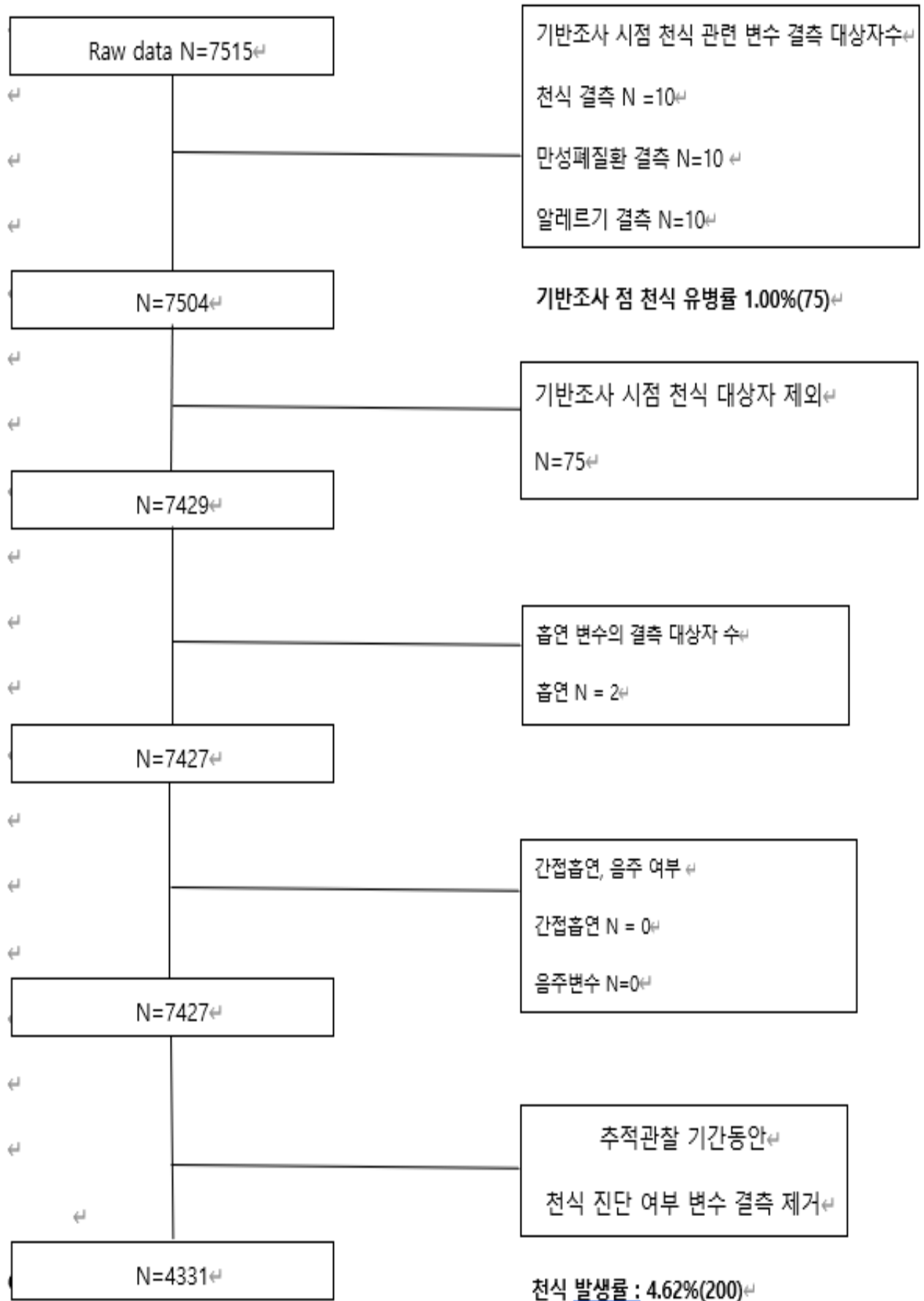
이에 p-value가 0.7424으로 0.05보다 높으며 통계적으로 유의미 하지 않다

[성별]

남자의 경우 천식에 안걸린 사람이 45.1, 걸린사람이 41.5로 둘이 거의 비슷하지만 안걸린 사람의 비율이 더 높은 반면

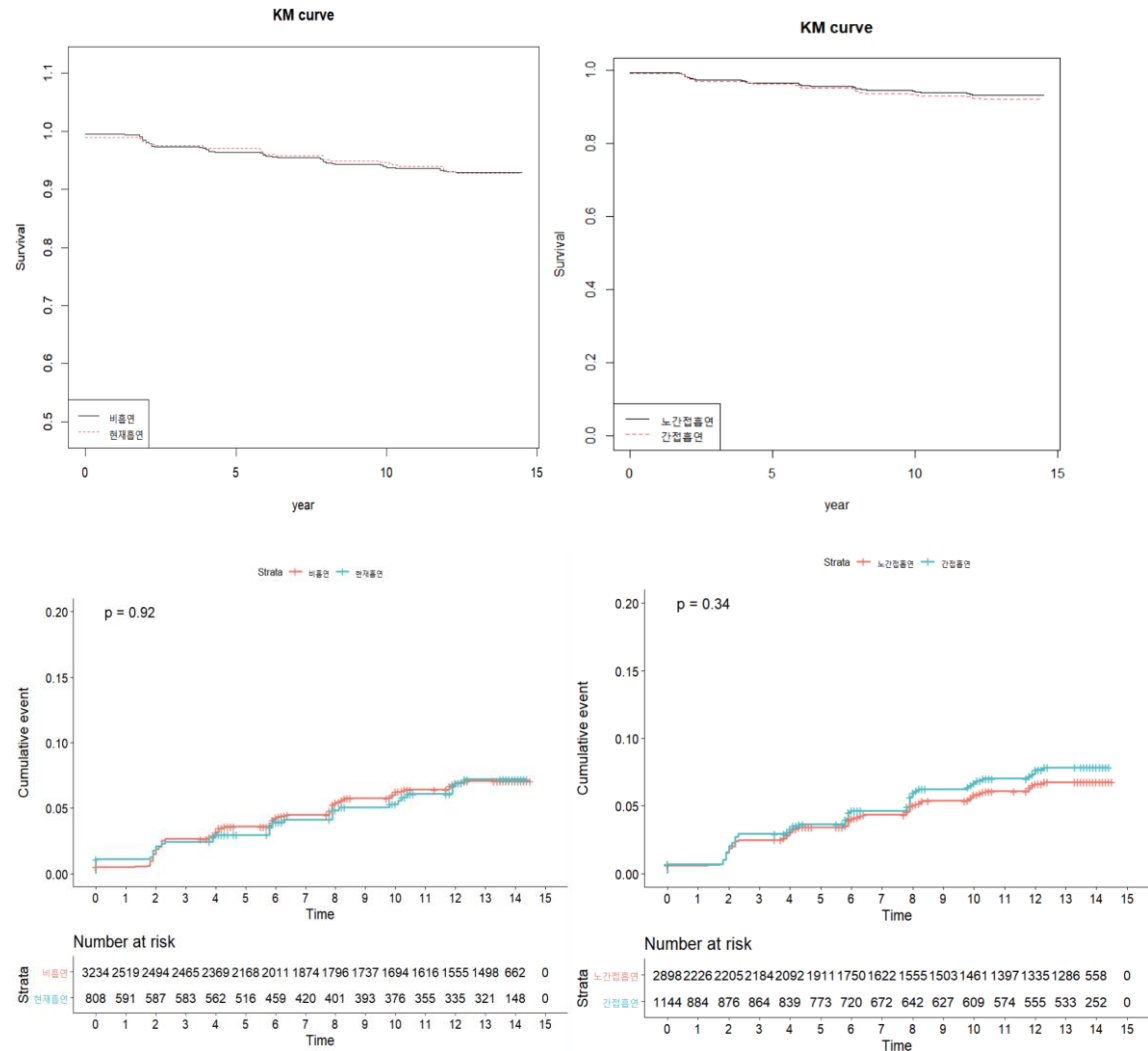
여자의 경우 천식에 안걸린사람이 54.9, 걸린사람이 58.5로 걸린 사람의 비율이 더 높다

이에 p-value가 0.3543으로 0.05로 높아 통계적으로 유의미 하지 않다



2) 발생율과 카플란마이어 그래프 해석

발생율이 4.62%이며 카플란마이어 그래프 결과는 아래와 같음



⇒ 직접흡연과 간접흡연 두 그래프 모두 차이가 거의 없어 둘 다 0.05보다 높은 수치로 유의미하지 않으며 직접흡연과 간접흡연 결과 직접 흡연이 0.94로 간접흡연 0.34보다 높은 수치를 보이고 있음

3) 콕스비례위험모형 최종 분석 결과와 해석

종속변수 : atm_d	95신뢰구간	z
--------------	--------	---

Smoke_yn	(0.7348-1.691)	0.610
Drink_yn	(0.6933-1.311)	0.769
AS3_SEX	(0.7817-1.596)	0.544

[흡연] 흡연이 1씩 증가할 때마다 1.17배 올라가며 신뢰구간에서 1을 포함하고 있어 유의미하지 않다

[음주] 음주가 1씩 증가할 때마다 0.95배 감소하며 신뢰구간에서 1을 포함하고 있어 유의미하지 않다

[성별] 성별이 1씩 증가할 때마다 1.12배 증가하며 신뢰구간에서 1을 포함하고 있어 유의미하지 않다.

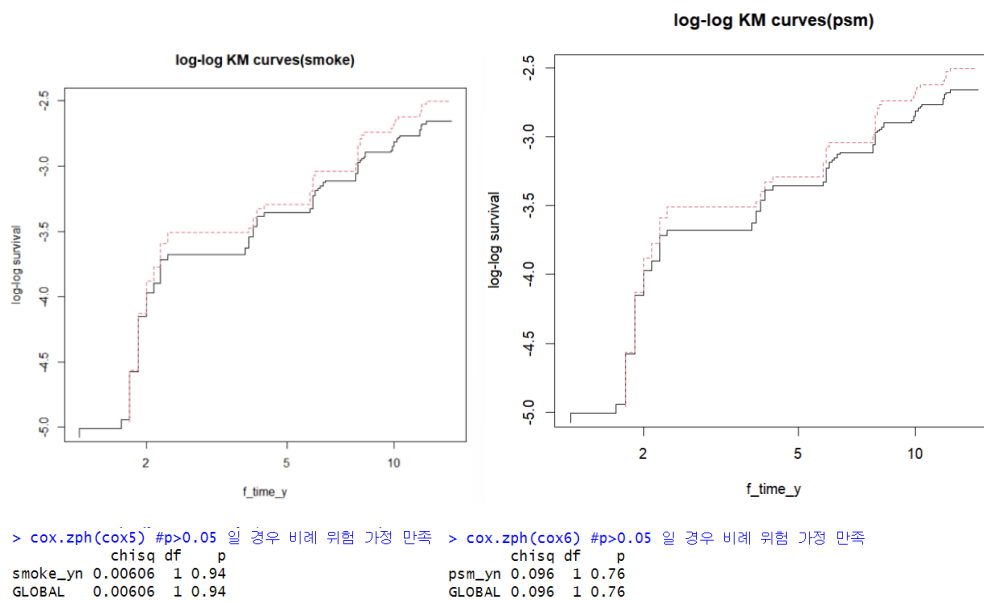
종속변수 : atm_d	95신뢰구간	z
Psm_yn	(0.8635-1.575)	0.316
Drink_yn	(0.6916-1.305)	0.751
AS3_SEX	(0.7825-1.479)	0.653

[흡연] 흡연이 1씩 증가할 때마다 1.17배 증가하며 신뢰구간에서 1을 포함하고 있어 유의미하지 않다

[음주] 음주가 1씩 증가할 때마다 0.94감소하며 신뢰구간에 1을 포함하고 있어 유의미하지 않다.

[성별] 성별이 1씩 증가할 때마다 1.07배 증가하며 신뢰구간에 1을 포함하고 있어 유의미하지 않다

- 비례위험 가정도



➔ 직접흡연과 간접흡연 각각 0.94와 0.76으로 0.05보다 큰 값으로 나와 비례 위험 가정에 만족함

4. 고찰 및 결론

[결론]

흡연과 천식은 연관성이 거의 없으며 천식과 흡연의 연관성은 통계적으로 유의미하지 않음

- 직접흡연과 비교 간접흡연 비교 결과 : 직접흡연보다 간접흡연이 천식 발생율이 더 높다
- 천식에 걸린 사람이 안걸린 사람과의 명수 차이가 약 5배 정도 차이가 남
즉 천식에 걸린 사람이 상대적으로 수가 너무 적어 확률이 낮게 나올 가능성 있음

[고찰]

“서울대학병원 의료정보”에 따르면 천식의 주요 발생원인인 알레르기는 집먼지, 진드기, 꽃가루 등 비듬이나 바퀴벌레, 식품 약물들로 인해서 발생된다고 함.

따라서 천식을 악화하는 원인으로 담배연기, 실내오염, 대기오염 등 다양한 원인요소들이 많

기에 흡연과 천식만 가지고 연관성을 분석 결과가 부족하게 나왔을 것으로 보임