

박기찬 논문 정리

Cigarette smoking and thyroid cancer risk : a cohort study

□ Cigarette smoking and thyroid cancer risk: a cohort study.

(PMID:30111870)

Abstract


Citations

Related Articles

Data

BioEntities

External Links

[Cho A¹](#), [Chang Y²](#) , [Ahn J¹](#), [Shin H¹](#), [Ryu S³](#) 

[Affiliations](#) ▶

[British Journal of Cancer](#) [15 Aug 2018]

Type: Journal Article

DOI: [10.1038/s41416-018-0224-5](https://doi.org/10.1038/s41416-018-0224-5) 

<https://europepmc.org/abstract/med/30111870>

도입

요즘 갑상선 암에 걸리는 환자가 많아지고 있는데, 아직 까지 흡연과 갑상선 암사이의 관계에는 많은 논란이 있다. 메타 분석에서 다시 한번 흡연과 갑상선의 관계를 연구하였지만, case-control 연구에서만 흡연이 갑상선 암의 발생률을 낮춘다는 결과가 있었기에 잘 계획되고 표본수가 많은 코호트 연구의 필요성이 생겨났다.

또한 BMI와 TSH(thyroid-stimulating Hormone)이 갑상선 암과 관련이 있다는 선행 연구가 있기 때문에 해당 변수들을 조절변수(mediate variable)로 사용하였다.

대상

1.강북 삼성 병원에서 2006년~2008년 사이에 건강 검진을 받은 18세 이상의 사람을 2014년 까지 추적함 (98,533명).

2.측정 하는 변수들에 결측치가 있는 대상들은 모두 제외. 최종적으로 (96,855명)

(환자들의 기저 질환을 조사하고, 기존에 암이 있거나 하는 환자들을 제외 하였다 라는 언급은 없음)

변수(최종 사용 변수)

1.흡연 상태(Never, former, current)

2.음주 여부(0, <20g/day, >=20g/day for men / 0, 10<g/day, >=10g/day for women)

3.운동 습관(>=3 times vs <3 time per week)

4.BMI(>=25 or < 25)

5.수축기 혈압, 이완기 혈압(수축기 혈압 140 이상 or 이완기 혈압 90 이상을 고혈압으로 분류)

6.교육수준(>=college graduate vs college graduate or below)

7.TSH(Thyroid stimulating hormone) (2018년 논문에서 담배가 TSH를 낮춘다는 연구를함)

8. LDL-C(나쁜 콜레스트롤), HDL-C(좋은 콜레스트롤)

9.HOMA-IR(인슐린 저항성, >2.5 vs < 2.5)

10.hsCRP(고감도C단응담백. 염증이 있을 때 체내에서 발생하는 물질. >= 1.0 vs < 1.0)

분석

1. Baseline characteristic of study를 시행. (기본적으로 남, 여를 구분하여 시행한다. 아래는 남자 대상)

Table 1. Baseline characteristics of study participants by smoking status among men

Characteristics	Smoking status			P for trend	
	Overall	Never smoker	Ex-smoker		Current smoker
Number	56,060	17,005	16,392	22,663	<0.001
Age (years) ^a	40.4 (7.7)	39.8 (8.1)	42.2 (8.0)	39.7 (6.8)	<0.001
BMI (kg/m ²)	24.5 (2.8)	24.2 (2.8)	24.6 (2.7)	24.6 (2.9)	<0.001
Obesity (%)	40.0	36.7	41.1	41.6	<0.001
Alcohol intake (%) ^b	19.7	10.3	19.3	27.0	<0.001
Vigorous exercise (%) ^c	20.4	21.2	27.0	15.1	<0.001
High education level (%) ^d	97.9	97.1	96.8	99.3	0.212
Diabetes (%)	4.7	3.6	5.4	5.0	<0.001
Hypertension (%)	20.4	19.6	24.2	18.4	<0.001
Systolic BP (mmHg) ^a	118.5 (12.9)	118.7 (13.0)	119.5 (13.1)	117.6 (12.6)	<0.001
Diastolic BP (mmHg) ^a	76.8 (8.8)	76.7 (8.8)	77.5 (8.8)	76.4 (8.8)	<0.001
Glucose (mg/dL) ^a	97.8 (15.9)	97.1 (15.0)	98.6 (15.4)	97.7 (17.0)	0.001
Total cholesterol (mg/dL) ^a	195.2 (33.0)	193.1 (32.3)	196.5 (33.1)	195.8 (33.4)	<0.001
LDL-C (mg/dL) ^a	115.7 (29.1)	114.7 (28.4)	116.0 (29.0)	116.3 (29.7)	<0.001
HDL-C (mg/dL) ^a	49.4 (10.5)	49.8 (10.5)	50.1 (10.6)	48.7 (10.3)	<0.001
Triglycerides (mg/dL) ^e	124 (88–177)	111 (80–157)	121 (87–172)	137 (97–195)	<0.001
HOMA-IR ^e	1.93 (1.47–2.55)	1.92 (1.47–2.53)	1.97 (1.51–2.58)	1.91 (1.44–2.55)	0.054
hsCRP (mg/L) ^e	0.6 (0.3–1.1)	0.5 (0.3–1.0)	0.5 (0.3–1.1)	0.6 (0.3–1.2)	<0.001

Data are presented as the mean (standard deviation), median (interquartile range), or percentage. BMI body mass index, BP blood pressure, LDL-C low-density lipoprotein-cholesterol, HDL-C high-density lipoprotein-cholesterol, HOMA-IR homeostasis model assessment of insulin resistance, hsCRP high sensitivity C-reactive protein. ^aMean (standard deviation) ^b≥20 g of ethanol per day ^c≥3 times per week ^d≥college graduate. ^eMedian (interquartile range)

분석

2. Incidence rate는 인년을 사용하여 계산

2. 인년(person-years)

인년

: 각 개인의 서로 다른 관찰기간의 합.

ex) 1명을 1년간 관찰하면 = 1인년 , 5명을 1년간 관찰하면 = 5인년 , 5명을 5년간 관찰하면 = 25인년

1000인년당 발생률 =

특정 기간 동안 그 집단에서 발생하는 새로운 질병의 사례 수

----- x 1000

총 인년(해당 기간 동안 각 대상자의 관찰 기간의 합)

3. 모수적 비례 오즈 모형을 사용함.

(해당 논문에서는 1년 단위의 자료 밖에 없기 때문에, 정확히 갑상선 암이 언제 발생했는 지를 알 수 없어, 이를 보완하고자 모수적 모형을 사용하였고, 해당 모수(기저함수)는 자유도가 4인 cubic spline을 통해 parameterize하였다. 일반적으로는 그냥 비모수적 비례 오즈 모형을 사용하면 된다.)

분석

4. model 1 : age를 보정변수로 사용

model 2 : 건강 검진을 받은 지역, 첫 건강검진을 시행 받은 년도, 알코올 섭취량(≥ 20 vs < 20),
regular exercise(≥ 3 vs < 3), 교육 수준(\geq college graduate vs $<$ college

graduate),

model 3 : BMI, TSH(thyroid stimulating hormone) 및 기타 변수들

5. 추가적인 분석으로 time varying covariate을 보정해줌.

(건강 검진을 받은 시점에 따라 흡연 상태, pack-years 등 시점에 따라 바뀌는 변수들 보정)

6. sensitivity analysis를 시행.

1. 본래 기존에 갑상선 암을 갑상선과 관련된 모든 암에 대한 정의로 하였다면, 갑상선 암의 세부적인 종류에 대해서도 흡연 여부가 유의한 결과를 나타 내는지 확인(다음 슬라이드에 Sensitivity analysis 설명)

2. 3-year wash-out 기간을 두고 분석을 시행.

Sensitivity Analysis(민감도 분석)

의학 분야에서 민감도 분석이란 쉽게 말해서 현재 내린 결론이 분석 방법의 변화, 결측치를 imputation했냐 안했냐, Outlier을 제거 했냐 안했냐에 따라 변화를 하나 안하냐를 보는 것.

1. 예를 들면, Outlier을 제거 하지 않고 우리가 만든 모델의 특정 변수가 유의 했다면, Outlier을 포함하고도 해당 변수가 유의 한지 아닌지를 보는 것

2. 결측치를 제거 했을 때 변수가 유의했다면, 결측치를 Imputation을 했을 때도, 해당 변수가 유의 한지 아닌지를 보는 것.

3. 변수에 대한 정의를 A로 했을 때, 우리가 원하는 변수가 유의 했다면, 변수에 대한 정의를 A*로 했을 때도 변수가 유의한 지 보는 것.

즉 이러한 변화를 줬을 때도, 변수가 유의하였고 결과가 연구자가 원하는 결과였다면 해당 연구는 Robust한 연구 결과라고 보는 것이다.

Table 3. Development of thyroid cancer by smoking status

Smoking status	Person-years	Incident Cases	Incidence density (per 1,000 person-years)	Age adjusted HR (95% CI)	Multivariable-adjusted HR ^a (95% CI)		HR (95% CI) ^b in model using time-dependent variables
					Model 1	Model 2	
Men							
Never smoker	91,699.7	146	1.6	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Ex-smoker	86,462.7	132	1.5	0.94 (0.74–1.19)	0.93 (0.73–1.18)	0.90 (0.71–1.14)	0.95 (0.74–1.21)
Current smoker	119,971.1	112	0.9	0.59 (0.46–0.75)	0.58 (0.45–0.75)	0.56 (0.43–0.72)	0.60 (0.46–0.78)
<i>P</i> for trend				<0.001	<0.001	<0.001	<0.001
Women							
Never smoker	202,568.8	832	4.1	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Ex-smoker	6542.4	19	2.9	0.73 (0.46–1.15)	0.72 (0.46–1.14)	0.72 (0.46–1.14)	1.01 (0.64–1.62)
Current smoker	3808.3	9	2.4	0.60 (0.31–1.15)	0.64 (0.33–1.23)	0.63 (0.33–1.23)	0.85 (0.60–1.19)
<i>P</i> for trend				0.043	0.060	0.059	0.370

BMI body mass index, *HR* hazard ratio, *CI* confidence interval. ^aEstimated from parametric Cox models. Multivariable model 1 was adjusted for age, center, year of screening exam, alcohol intake, regular exercise, educational level; model 2: model 1 plus adjustment for BMI and TSH ^bEstimated from parametric proportional hazard models with smoking status, alcohol intake, and regular exercise as time-dependent variables and baseline age, center, year of screening exam and education level as time-fixed variables.

남자에게서 p-value가 유의함을 볼 수 있으나 여성에게서는 유의하지 않음을 볼 수 있다.

크게 받는다. 수백만명의 대규모 데이터 집합 전적 변화라던가, 결과변수가 수만개일 수 있다. 또한 원인변수나 결과변수 모두에 통계적 검정력이 크게 떨어진다(adjustment or over-adjustment), 비선형적인 관계(non-linear relationship), 적합하지 않은 통계분석 등의

결과

Pack-years	Person-years	Incident cases	Incidence density (per 1000 person-years)	Age-adjusted HR (95% CI)	Multivariable-adjusted HR ^a (95% CI)		HR (95% CI) ^b in model using time-dependent variables
					Model 1	Model 2	
Men							
0	91,699.7	146	1.6	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
<10	36,712.5	34	0.9	0.61 (0.42–0.89)	0.60 (0.41–0.88)	0.59 (0.40–0.86)	0.64 (0.48–0.85)
10–19.9	50,912.8	60	1.2	0.76 (0.56–1.02)	0.73 (0.54–0.99)	0.69 (0.51–0.94)	0.79 (0.59–1.04)
≥20	42,056.5	32	0.8	0.45 (0.31–0.67)	0.44 (0.29–0.65)	0.40 (0.27–0.60)	0.42 (0.27–0.62)
<i>P</i> for trend				<0.001	<0.001	<0.001	<0.001
Women							
0	202,568.8	832	4.1	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
<5	3067.8	4	1.3	0.33 (0.12–0.88)	0.33 (0.13–0.90)	0.33 (0.12–0.88)	0.83 (0.59–1.15)
5–9.9	1331.7	4	3.0	0.76 (0.29–2.04)	0.91 (0.30–2.17)	0.81 (0.30–2.18)	0.49 (0.15–1.53)
≥10	905.1	5	5.5	1.32 (0.55–3.17)	1.39 (0.57–3.38)	1.40 (0.57–3.40)	1.56 (0.64–3.79)
<i>P</i> for trend				0.410	0.504	0.500	0.421

BMI body mass index, *HR* hazard ratio, *CI* confidence interval. ^aEstimated from parametric Cox models. Multivariable model 1 was adjusted for age, center, year of screening exam, alcohol intake, regular exercise, educational level; model 2: model 1 plus adjustment for BMI and TSH ^bEstimated from parametric proportional hazard models with pack-years, alcohol intake, and regular exercise as time-dependent variables and baseline age, center, year of screening exam, and education level as time-fixed variables.

남자에게서 p-value가 유의함을 볼 수 있으나 여성에게서는 유의하지 않음을 볼 수 있다.

크게 받는다. 수백만명의 대규모 데이터가 전적 변이라던가, 결과변수가 수만개일 수 있다. 또한 원인변수나 결과변수 중에서도 통계적 검정력이 크게 떨어지는 adjustment or over-adjustment), 비선형적인 관계(non-linear relationship), 적합하지 않은 통계분석 등의

논의.

본래 TSH, BMI가 갑상선 암의 발생에 영향을 준다고 알려져 있고, 담배가 TSH를 낮춰주고 또 본래 담배를 피는 사람들의 BMI가 낮을 수 있기 때문에 담배를 피는 사람이 갑상선 암에 덜 걸릴 수도 있다. 하지만 이 연구에서는 TSH와 BMI를 보정하였음에도 담배가 갑상선 암의 발생을 낮추는 것으로 나타났다.

그래서 가능한 또 다른 경우는 담배가 항 에스트로겐 효과를 가지고 있다는 것이다. 남성에 비해 여성이 갑상선암에 많이 걸리기에 에스트로겐이 갑상선 암의 원인으로 지목되어 왔다. 즉 담배가 TSH를 낮추고, 항 에스트로겐 효과를 가지고 있기에 담배를 피는 사람이 갑상선 암에 덜 걸리는 것이다.

또한 니코틴이 갑상선 암을 낮추는 원인일 수도 있다. 면역 세포에서의 니코틴 수용체의 활성화가 자가면역 체계를 병원성 Th1, Th2(pathogenic Th1, Th2)으로부터 protective Th2 response로 바꿔주기 때문이다..?????????

등등..이 뒤로부터는 생물학적 설명

그래서 위의 해당 연구에 대한 필요성이 있다..

(해당 부분은 연구자가 개인의 Domain 지식을 가지고 왜 담배가 갑상선 암의 발생률을 낮추는 지 설명 하는 부분)

크게 받는다. 수백만명의 대규모 데이터적 변이라던가, 결과변수가 수만개 수 있다. 또한 원인변수나 결과변수 무에도 통계적 검정력이 크게 떨어지 adjustment or over-adjustment), 의 비선형적인 관계(non-linear relationship, 적합하지 않은 통계분석 등의