



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

한국인에서 체질량지수와 폐암의 관련성 분석

연세대학교 보건대학원
역학건강증진학과 건강증진교육
김 정 훈

한국인에서 체질량지수와 폐암의 관련성 분석

지도 지 선 하 교수

이 논문을 보건학석사 학위논문으로 제출함



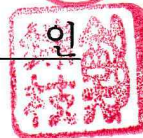
2020년 6월 일

연세대학교 보건대학원

건강증진교육 전공

김 정 훈

김정훈의 보건학 석사학위 논문을 인준함

심사위원 지 선 하 
심사위원 김 의 권 
심사위원 정 금 지 

연세대학교 보건대학원

2020년 6월 일

감사의 글

대학원이라는 곳은 학문을 배우고 지식의 수준을 높여 좀 더 높은 곳을 가기 위한 계단이라고 생각합니다. 그 외에 여러 이유로 많은 사람들이 자신이 원하는 것을 얻고자 대학원을 진학하지만, 대학원 생활이 사회초년생인 저에게는 보다 많은 것을 가져다주었습니다. 뿐만 아니라 대학원 내에 국민건강증진연구소에서의 경험은 가치 있는 사람이 될 수 있도록 많은 변화를 가져다주었습니다. 없었던 목표가 생기고 열정이 생겼으며 목표를 향해 나아 갈 수 있는 용기를 얻었습니다. 그 덕에 최대의 난제였던 즐거움을 느끼며 할 수 있는 일이 무엇인지 알게 되었습니다.

대학원 기간 동안 많은 우여곡절이 있었지만 위기를 극복 하고 졸업까지 달려올 수 있었던 건 항상 믿음을 주시고 배려 해주시며 높은 위치에 계심에도 항상 겸손한 성품을 지니고 계신 지선하 교수님, 저에게 부족했던 부분과 지녀야 할 능력을 길러주신 김희진 교수님, 때로는 칭찬으로 때로는 꾸중으로 똑바로 나아가게끔 버팀목이 되어주신 정금지 교수님, 학문과 연구에 대한 열정으로 연구의 재미를 알려주신 이지영 교수님 그리고 4층 박사 선생님들과 3층 선생님들이 계셔서 가능했습니다. 이제 졸업을 앞두고 곧 학교와 연구소를 떠나 다른 곳으로 가게 되겠지만 멀리서라도 보건대학원의 무궁한 발전과 국민건강증진연구소 식구들이 항상 건강하고 모든 일이 잘 될 수 있기를 간절히 기도 하겠습니다.

마지막으로 한 가정의 가장으로써 모든 어려운 일을 감수하시어 부족함 없이 길러 주신 아버지, 가정의 평화와 행복이 유지되도록 노력하시는 어머니, 정신적 멘토로써 항상 모범이 되고 바른길로 갈 수 있도록 안내자 역할을 해주는 형에게도 이 지면을 빌려 감사의 마음을 전합니다.

김정훈 올림

차 례

국문요약	VI
I. 서론	
1. 연구의 배경 및 필요성	1
2. 연구목적	4
II. 문헌고찰	
1. 비만과 폐암의 선행연구	5
2. 체질량지수와 폐암 유발물질의 생물학적 관련성	7
3. 멘텔리언 무작위할당의 개념과 선행연구	9
III. 연구방법	
1. 연구자료	11
2. 연구대상	11
3. 분석방법	12
가. 변수선택	12
나. 변수정의 및 클리닝	12
다. 분석 프로그램 및 통계적 방법	13
IV. 연구결과	
1. 대상자의 일반적 특성	14
가. 폐암여부에 따른 전체코호트와 부표본의 체질량지수 차이	18

2. 전체코호트(Whole cohort), 부표본(Subsample) 로지스틱 분석	20
가. 전체코호트의 체질량지수에 따른 조직학적 분류된 폐암 발생 위험	20
나. 부표본의 체질량지수에 따른 조직학적 분류된 폐암 발생 위험	24
다. 전체코호트의 체질량지수와 follow-up year에 따른 폐암 발생 위험	28
라. 부표본의 체질량지수와 follow-up year에 따른 폐암 발생 위험 ..	30
마. 전체코호트에서 체질량지수와 흡연 상태 및 follow-up year에 따른 폐암 발생 위험	32
바. 부표본에서 체질량지수와 흡연 상태 및 follow-up year에 따른 폐암 발생 위험	34
3. 멘델리언 무작위할당 분석	37
가. 체질량지수 관련 SNP	37
나. SNP 검증	38
V. 고찰	42
1. 제한점	44
가. 관찰연구	44
나. 멘델리언 무작위할당 연구	45
2. 강점	45
VI. 결론	46

참고문헌 47

Abstract 54

표 차 례

Table 1. Selected characteristics of participants in KCPS-II	16
Table 2. Hazard ratio for lung cancer by BMI in Korean, whole cohort of KCPS-II	21
Table 3. Hazard ratio for lung cancer by BMI in Korean, subsample of KCPS-II	25
Table 4. Hazard ratio for Lung cancer by BMI stratified by follow-up year in whole cohort of KCPS-II	29
Table 5. Hazard ratio for Lung cancer by BMI stratified by follow-up year in subsample of KCPS-II	31
Table 6. Hazard ratio for Lung cancer by BMI stratified by smoking state, whole cohort of KCPS-II	33
Table 7. Hazard ratio for Lung cancer by BMI stratified by smoking state, subsample of KCPS-II	35
Table 8. GWAS Result of Body Mass Index	39
Table 9. MR analyses of lung cancer by estimated BMI in subcohort of KCPS-II	41

그림 차례

Figure 1. Distribution of bmi according to lung cancer in each cohort, p-value	19
Figure 2-a. Hazard ratio for lung cancer by BMI in Korean, whole cohort	22
Figure 2-b. Hazard ratio for Adenocarcinoma cancer by BMI in Korean, whole cohort	22
Figure 2-c. Hazard ratio for Squamous cell carcinoma cancer by BMI in Korean, whole cohort	23
Figure 2-d. Hazard ratio for Small cell cancer by BMI in Korean, whole cohort	23
Figure 3-a. Hazard ratio for Lung cancer by BMI in Korean, subsample	26
Figure 3-b. Hazard ratio for Adenocarcinoma cancer by BMI in Korean, subsample	26
Figure 3-c. Hazard ratio for Squamous cell carcinoma cancer by BMI in Korean, subsample	27
Figure 4. Manhattan plot for BMI(SNP of BMI is adjusted age, sex)	37

국 문 요 약

한국인에서 체질량지수와 폐암의 관련성 분석

연구 배경

선행된 관찰연구에 따르면 낮은 체질량지수가 폐암 발생을 높이는 결과를 보여주고 있지만, 반대로 멘델리언 무작위할당 연구에서는 높은 체질량지수가 폐암 발생을 높이고 조직학적 특성에 따라 편평상피세포암과 소세포암에서 유의한 관련성이 있다고 알려져 있다. 따라서 본 연구는 체질량지수가 폐암 발생에 미치는 영향을 알아내고자 한국인을 대상으로 관찰연구와 멘델리언 무작위할당 연구를 통해 비교분석 하고자 하였다.

대상 및 방법

본 연구는 KCPS-II (Korean Cancer Prevention Study-II, 한국인 암 예방연구-II)를 이용하여 관찰연구 대상자인 159,844명과 유전 정보가 있는 16,959명에 대하여 체질량지수와 폐암과의 관련성을 알아보았다. 관찰연구 대상자는 콕스 비례모형을 통해 분석하였고, 유전데이터를 포함하는 대상자는 멘델리언 무작위할당 연구를 시행하였다.

연구 결과

KCPS-II 전체코호트의 관찰연구에서는 체질량지수 23~25kg/m²을 기준집단으로 하였을 때 20kg/m² 미만 인 집단에서 폐암에 걸릴 위험이 1.66배 높았고 (HR=1.66, 95% CI=1.16-2.39), 추적기간년수(follow-up year) 4년 이상 체질량지수 20kg/m² 미만 인 흡연집단에서 폐암을 유발하는 유의한 양의 관련성을 확인 하였다(HR=2.58, 95% CI=1.40-4.76). 하지만 멘델리언 무작위할당 연구에서는 체질량지수와 폐암과의 관련성이 없었다.

결론

관찰연구와 멘델리언 무작위할당 분석의 결과에 따라 체질량지수와 폐암과의 직접적인 관련성은 없었으며, 관찰연구의 전체 집단에서 낮은 체질량지수가 편평상피세포암의 발생을 높이는 것과 낮은 체질량지수를 가지는 흡연자에서 높은 폐암 발생을 보이기 때문에 체질량지수와 폐암 발생의 음의 관련성은 체질량지수가 아닌 흡연효과로 인한 체중감소로부터 영향을 받은 결과일 것이다.

핵심어 : 체질량지수, 비만, 흡연, 폐암, 멘델리언 무작위할당 연구

I. 서론

1. 연구 배경 및 필요성

전 세계적으로 폐암은 남성에서 사망률이 가장 높고 여성에서는 두 번째로 높은 사망률을 보이는 악성 종양이다. 폐암은 암세포의 크기와 형태에 따라 크게 ‘비소세포폐암(non-small cell lung cancer)’ 과 ‘소세포폐암(small cell lung cancer)’ 으로 분류 할 수 있으며, 비소세포폐암은 다시 ‘선암(adenocarcinoma)’ 과 ‘편평상피세포암 (squamous cell carcinoma)’ , ‘대세포암(large cell carcinoma)’ 등으로 구분 할 수 있다(Novaes et al, 2008).

폐암의 가장 흔한 원인은 흡연으로 알려져 있으며, 그 밖에 라듐으로부터 생성되는 라돈, 실외 및 실내 대기오염 노출로부터 유발하는 것으로 알려져 있다(Mao et al, 2016).

이처럼 폐암의 발생은 다양한 노출로부터 기인하는데, 일부 연구에서는 체질량지수(BMI; body mass index)와 같은 비만관련 요인들이 폐암과 음의 상관관계를 보여주고 있다(Yu et al, 2018).

비만은 유방, 대장, 췌장, 신장암 등 대부분의 암에서 발생 위험을 올리는 양의 관련이 있는 것이 일반적이다.(Basen-Engquist et al, 2011).

최근 메타분석에 따르면 다른 암과 비교하여 폐암에서는 비만과 음의 관련성이 있다는 연구 결과를 보여주며 조직학적 특성에 따라서도 차이를 보여주고 있다(Duan et al, 2015). 특히, 흡연자 사이에서의 높은 체질량지수는 폐

암 발생 위험을 줄이며(Smith et al, 2012) 저밀도 지방단백질(LDL; low density lipoprotein) 또한 폐암과 음의 관련성을 보여주었다(Kucharska-newton et al, 2009). 관련성에 대한 연구의 한계점으로 흡연시 체중감소를 예로 들고 있지만 몇몇 연구에서는 비흡연자 사이에서의 결과도 동일하게 나올 뿐만 아니라 흡연 시 체중감소를 통제하였음에도 불구하고 체질량지수와 폐암의 음의 관련성을 입증하였다(Dewi et al, 2016).

그러나 위와 같이 관찰연구를 통해 체질량지수와 폐암 발생간의 관련성이 알려졌음에도 인과성을 설명하는 데는 여전히 논란이 많다.

멘델리언 무작위할당(MR; mendelian randomization)이란, 인간 세포에서의 감수분열이 일어나는 동안 유전자의 자연적이고 무작위적인 할당에 의존하고 있는 방법이다(Smith et al, 2003).

개인마다 태어날 때 질병과 관련된 유전자를 물려받을 수도 아닐 수도 있다. 이는 부모로부터 무작위적으로 할당되는 것이기 때문에 멘델리언 무작위 할당 연구는 무작위 대조시험(RCT; randomized controlled trial)과 동일한 개념으로 위험인자가 무작위적으로 개인에게 할당된, 혼란변수를 통제된 상태에서 연구를 진행한다는 것을 전제로 한다. 따라서 무작위할당을 통해 혼란변수는 통제되었다고 설명 할 수 있기 때문에 독립변수와 종속변수의 인과적인 특성을 설명 할 수 있다.(Emdin et al, 2017).

멘델리언 무작위할당 연구는 질병과 관련된 유전인자로 단일염기다형성(Single Nucleotide polymorphism, SNP)을 사용하여 설명한다. SNP란 DNA(deoxyribonucleic acid) 한 쌍의 염기서열(ex, A=T)에서 하나의 이상의 염기(A,T,G,C) 차이를 보이는 유전적 변이를 말한다. 즉, 인구집단에서 1% 이상의 빈도로 존재하는 2개 이상의 대립 염기서열이 발생하는 자리를 SNP라고 한다(Ricki Lewis, 2009). 개인간 SNP의 차이에 따라 유전자 기능이 달라

질 수 있으며 서로 상호 작용하여 다른 결과를 만들고 질병에 대한 다른 감수성 차이를 보인다.

기존 역학연구는 체질량지수와 폐암간의 음의 관련성을 보였고 특히, 흡연자 집단에서 그리고 폐암 중 선암과 대세포암에서 유의한 결과를 보여주었다 (Sanikini et al, 2018). 그러나 역학연구와 비교하여 멘델리언 무작위할당 연구에서는 역학연구에서 유의하지 않았던 편평상피세포암과 소세포암에서 유의한 양의 관련성을 보여주었다(Carreras-Torres et al, 2017).

관찰연구에서 분석결과에 영향을 줄 수 있는 흡연에 의한 체중감소를 통제하여 지속적인 음의 관련성이 나오더라도 아직도 많은 제한점과 해석을 명확히 하기에는 어려움이 많다. 따라서 관찰연구가 가지는 한계점과 문제가 되는 요인들을 최소화하는 노력이 필요하고 결과를 명확하게 하기 위해 흡연에서 오는 영향과 환경적인 요소들을 통제할 뿐만 아니라 유전처럼 태어나면서 정해지는 생물학적 차이를 고려해주고 다양한 분석 방법을 통해서 비교연구가 이루어져야한다. 그러나 관찰연구는 서양인 뿐만 아니라 동양에서도 연구된 사례가 있는 반면에, 멘델리언 무작위할당 연구는 대부분 서양인을 대상으로 연구되었다. 게다가, 독립변수와 종속변수를 포함하는 데이터가 하나의 집단이 아닌 여러 집단에서 측정된 값이기 때문에 결과 해석 측면에서 인종, 지역, 환경 등 인구사회학적특성에서 오는 영향을 통제하기에는 한계가 있다. 따라서, 체질량지수와 폐암간의 관련성을 보는 관찰연구와 비교할 수 있는 멘델리언 무작위할당 연구가 서양인을 대상으로만 이루어지고 있다는 점과 선행연구 수가 부족하고 인구사회학적특성의 영향을 최대한 통제해야하기 때문에 본 연구는 체질량지수와 폐암과의 관련성을 동양인인 한국인을 대상으로 한 KCPS-II (Korean Cancer Prevention Study-II, 한국인 암 예방연구-II) 자료를 이용하여 관찰연구와 멘델리언 무작위할당의 비교분석을 하고자 한다.

2. 연구 목적

KCPS-II (Korean Cancer Prevention Study-II, 한국인 암 예방연구-II)의 자료를 이용하여 콕스 비례위험모형을 통한 체질량지수와 폐암과의 관련성을 선행연구와 비교하고 멘델리언 무작위할당 분석으로 인과성을 확인하여 관찰연구와 멘델리언 무작위할당 연구의 결과를 비교하고자 한다.

본 연구의 목적은 다음과 같다.

첫 째, 체질량지수와 조직학적 특성에 따른 폐암 간의 역학적 관련성을 확인한다.

둘 째, 흡연 상태별 체질량지수와 폐암과의 관련성을 확인한다.

셋 째, 전장유전체상관성분석(GWAS: Genome wide association studies)을 이용하여 체질량지수와 유의한 관련성이 있는 단일염기다형성 (Single Nucleotide polymorphism, SNP)를 확인한다.

넷째, 체질량지수와 유의한 관련성이 있는 SNP를 활용한 MR분석을 통해 체질량지수와 폐암과의 인과성을 확인한다.

II. 문헌고찰

1. 비만과 폐암의 선행연구

기존 관찰연구에서는 비만이 여러 암(대장암, 유방암, 전립선암, 신장암 등)의 위험을 증가시킨다고 말한다(Bhaskaran et al, 2014). 그러나 이와는 대조적으로 비만을 대신할 수 있는 지표인 체질량지수에 관한 많은 전향적 코호트 연구에서는 체질량지수가 폐암과 음의 관련성이 있음을 발표하였고 일부 연구에서는 체질량지수가 비흡연자, 과거흡연자에서는 폐암을 일으키는 요인으로 양의 관련성이 있다고 보고하였다(Rauscher et al, 2000). 실제로 체질량지수와 폐암의 관련성 연구는 흡연 습관으로 인해 변화되는 요인(흡연시 체질량지수 감소)때문에 체질량지수와 폐암의 관련성을 평가하는 것에 제한점이 있다. 심지어 허리둘레는 체질량지수와 마찬가지로 비만지표로 이용되지만 폐암의 위험을 증가시키는 요인으로 체질량지수와는 반대의 결과를 보여주었다(Dewi et al, 2016). 이 처럼 흡연에 의한 체중감소 영향과 비만의 지표차이에 따라 관련성 방향이 다르게 되는 것이 연구의 해석 측면에서 많은 혼란을 가져온다(Hidayat et al, 2016).

연구의 한계점을 고려한 폐암과 관련된 전향적 코호트 연구만을 포함하여 수행한 메타 분석 연구가 최근에 발표되었는데(Gao et al, 2019), 이 연구에 따르면 흡연상태에 따른 하위집단으로 나누어 체질량지수와 폐암의 관련성을 보았을 때, 비흡연자에서는 음의 관련성을 보였지만 유의하지 않았고 과거 흡연자와 현재 흡연자를 포함한 흡연경험이 있는 집단에서는 체질량지수가 증가함에 따라 폐암 발생의 위험이 낮아지는 유의한 결과를 보였다. 그러나 결과의 한계를 지적하고 있는 흡연 시 체중감소를 고려한 연구만을 메타 분석한 결과에서는 통계적으로 유의하지 않았다. 조직학적 특성에 따라 선암과 소세

포함으로 분류하여 메타 분석을 수행한 결과에서는 선암에서 음의 관련성과 통계적 유의함을 보였고 소세포암은 통계적으로 유의하지 않는 결과를 보였다 (Gao et al, 2019). 허리둘레로 폐암과의 관련성을 본 연구에서는 흡연여부를 고려하지 않고 보았을 때와 흡연자, 과거흡연자, 비흡연자로 층화 분석한 결과 모두에서 유의한 양의 관련성을 보여주었다(Gao et al, 2019). 체질량 지수와 허리둘레에 다른 결과는 생물학적 기전으로 설명이 가능하다. 허리둘레와 폐암의 양의 관련성의 증거로는 폐암 환자와 정상인 사이에서의 생물학적 차이로 폐암 환자에서 높은 혈청 페리틴 수치와 인슐린 저항성을 가지고 있다는 것인데, 실제로 높은 허리둘레를 보유하고 있는 환자들에서 정상인과 비교하였을 때 높은 혈청 페리틴 수치와 인슐린 저항성을 가지고 있었다(Jehn et al, 2004)(Petridou et al, 2011).

2. 체질량지수와 폐암 유발물질의 생물학적 관련성

체질량지수와 폐암의 관련성 분석에 대한 선행된 몇 가지 생물학적 연구가 관련성 결과를 보충해주고 있다. 노출시 DNA(deoxyribonucleic acid)손상을 야기시킨다는 PAHs(polycyclic aromatic - hydrocarbons)는 2개의 벤젠 고리로 구성된 방향족 유기화합물로 화석 연료와 같은 유기 물질의 불완전 연소의 결과로 형성되는 광범위한 환경오염 물질이다(Garcia-Suastegui et al, 2011). PAHs는 일반적으로 직업적, 환경적, 흡연, 음식으로부터 노출 될 수 있으며 일부 역학 연구에서는 PAHs의 노출에 따라 폐암, 피부암, 방광암 등 발암 가능성을 주장하고 있다(Armstrong et al, 2002). 뿐만 아니라 PAHs의 장기간 노출에 따라 호흡기로 들어오는 경우 폐암을 유발하고 음식의 경우 피부암 등 노출 부위에 따라 유발 양상도 달라지는 것을 증명 하였다(Lynch et al, 2013). 특히, PHAs 중에서도 Benzo (a) pyrene이 가장 널리 연구되었으며 흡입 시 폐 종양을 유도하는 물질로 잘 알려져 있다(Hecht, 1999)

선행 연구에 따르면 체질량지수는 백혈구에서 측정된 DNA를 손상시키는 발암 물질의 부가물 수준에 영향을 준다고 한다(Godschalk et al, 2002). 그 이유는 암을 유발하는 방향족 부가물의 생성 전단계인 화학 구성물들이 지용성이어서 신체 내 지방조직에 갇혀 부가물을 형성 하는 것을 억제하기 때문이라고 한다(Godschalk et al, 2002). 또한 체질량지수와 폐암과의 관련성은 흡연 집단에서 뚜렷하게 나타나고 있는데 흡연을 많이 하는 집단이 흡연을 적게 하는 집단과 비교하여 발암물질 부가물의 반감기가 더 길다는 근거를 예시로 들 수 있으며(Godschalk et al, 2002), PHAs의 물질로 알려진 산화적 DNA손상의 지표 8-OHDG(8-hydroxyde oxyguanosine)의 수준 또한 흡연자 사이에서 체질량지수와 음의 관련성을 보여주고 있었다(Li et al, 2015).

유전학적으로는 체질량지수와 관련있는 비만 유전자 마커인 FTO(Fat mass and obesity-associated protein)의 유전자(rs9939609)가 폐암 위험을 감소시키는 데에 관련성이 있다고 한다(Brennan et al, 2009). 최근에는 DNA 메틸화와 체질량지수 사이에서 관련성 분석이 수행되었고 혈액과 지방조직의 HIF3A(Hypoxia-inducible factor 3 alpha)유전좌위에서의 발현의 차이가 체질량지수에 영향을 준다고 발표하였다. HIF3A는 세포 생존 및 세포사멸 이외에 혈관 신생과 관련된 많은 유전자를 조절할 수 있는 유전자위로 HIF3A 발현차가 폐암 발생에 영향을 줄 수 있음을 제안하였다(Dick et al, 2014).

3. 멘델리언 무작위할당의 개념과 선행연구

멘델리언 무작위할당 연구는 일반적인 전통적인 역학 연구에서의 혼란변수 개입을 통제할 방법으로써 독립변수와 종속변수의 인과성을 보기위한 방법으로 알려져 있다(Smith et al, 2003). 멘델리언 무작위할당 연구는 경제학에서 사용하는 2단계 최소제곱법(2 stage least squares, 2sls)의 방법을 토대로 구상되었다. 2단계 최소제곱법은 회귀분석에서 종속변수의 잔차 즉, 오차항이 독립변수와 상관관계가 없을 것이라는 가정이 틀린 경우 종속변수의 오차항과 관련없는 도구변수(instrumental variable)를 사용하여 2단계를 걸쳐 선형회귀분석을 하는 통계적 방법으로 분석은 두 단계에 걸쳐 진행된다. 첫째, 도구변수를 사용하여 문제가 되는 독립변수의 값을 추정하고 둘째, 도구변수로부터 추정된 독립변수 값을 이용하여 원래 식의 종속변수 값을 추정하는 것이다(Bollen, 1996).

2단계 최소제곱법을 기반으로 한 멘델리언 무작위할당 연구 분석은 도구변수로 SNP의 정보를 사용하여 추정한다. 개인의 SNP 정보는 부모로부터 감수분열을 통한 DNA 재조합을 통해 물려받기 때문에 자연적으로 유전정보를 무작위적으로 가지고 있다는 점에서 임상시험과 유사하다(Swanson et al, 2017). 유전변수는 출생당시부터 변하지 않는 값이기 때문에 혼란변수의 영향을 받지 않아 도구변수를 이용한 2단계 최소제곱법 사용에 적합하다. 2단계 최소제곱법과 마찬가지로 멘델리언 무작위할당 분석은 두 가지 단계로 진행되는데 첫째, SNP의 정보를 도구변수로 하여 원래 회귀모형의 독립변수를 추정하고 둘째, 추정된 독립변수 값을 회귀모형에 적용시켜 종속변수를 추정한다.

최근 멘델리언 무작위할당 연구에 따르면 체질량지수가 폐암의 위험을 증가시킨다고 하며 폐암의 조직학적 분류에 따라서는 편평상피세포암, 소세포암에

서 유의한 양의 관련성을 보여주었고 선암에서는 유의한 관련성을 보여주지 않았다(Carreras-Torres et al, 2017). 또한 흡연자와 비흡연자로 나누어서 체질량지수와 폐암과의 관계를 보았을 때도 통계적으로 유의한 양의 관련성을 보였다(Carreras-Torres et al, 2017).

Ⅲ. 연구방법

1. 연구 자료

KCPS-II는 2004년 4월에 시작되어 2005년 12월에 서울시 지원의 한국인 대사 증후군 연구로부터 시작되었다.

KCPS-II 바이오뱅크는 2004년 2개의 병원(세브란스 병원, 분당 차병원)으로부터 시작하여 2006년을 기점으로 11개의 병원(한국의학연구소, 이대목동 병원, 서울대학교 병원, 고려대학교 병원, 경희대학교 병원, 카톨릭대부천 성모병원, 서울성모병원, 서울의료원, 서울아산병원, 한양대학교 구리병원)으로 확장되었다. 2004년에서 2013년 사이에 11개의 건강증진센터에서 사전 동의를 받은 대상자 수는 159,844명이며 2013년 이후로 해당 대상자에 대해 지속적으로 추적 관찰 중이다(Jee et al, 2018). 한국 세브란스 의료윤리위원회는 이 연구를 승인했다(no. 4-2011-0277).

2. 연구 대상

2004년에서 2013년 사이에 11개의 건강증진센터에서 사전 동의를 받은 전체 코호트 대상자(Whole cohort)는 159,844명이며 그 중에 부표본(subsample)으로 유전데이터를 가지고 있는 대상자는 16,959명이다. 본 연구에서는 관찰연구와 멘델리언 무작위할당 연구를 비교 분석하기 위하여 관찰연구에서는 전체 코호트와 부표본을 이용한 결과의 차이를 보았고 부표본을 이용하여 멘델리언 무작위할당 분석을 통해 관찰연구의 결과와 비교하였다.

3. 분석 방법

가. 변수 선택

관찰연구를 위한 인구사회학적 요인으로 성, 연령, 건강행태 요인으로 흡연상태, 음주여부, 운동여부, 임상적 요인으로 체질량지수, 허리둘레, 수축기 혈압, 공복혈당, 당뇨여부, 고혈압여부, 폐암(전체 폐암, 선암, 편평상피세포암, 소세포암)을 보았다.

나. 변수 정의 및 클리닝

당뇨 여부는 당뇨병 치료제 사용여부, 공복혈당 126mg/dL 이상인 집단을 기준으로 정의 하였고 고혈압 여부는 고혈압 약 사용여부, sbp140mmHg 이상인 집단을 기준으로 정의하였다.

이상치를 제거 하기위해 체질량지수 $10\text{kg}/\text{m}^2$ 미만 $50\text{kg}/\text{m}^2$ 초과, 허리둘레 125cm 초과, 수축기 혈압 220mmHg 초과 공복혈당 400mg/dL초과인 집단과 나이, 성별, 체질량지수, 허리둘레, 수축기 혈압, 공복혈당, 흡연상태, 음주여부, 운동여부가 결측인 집단은 제외시켰다. 최종적으로 전체 코호트(whole cohort)는 137,385명을 대상으로 하였고 부표본(subsample)은 멘델리언 무작위할당 연구에서 SNP를 추출하기 위해 유전데이터 모두 가지고 있는 16,959명을 대상으로 분석을 진행하였으며 관찰연구에서는 전체코호트와 동일한 기준으로 적용시켜 14,443명을 대상으로 분석하였다.

다. 분석 프로그램 및 통계적 방법

SAS 9.4 버전을 이용하여 전체코호트 및 부표본 대한 일반적 특성과 그래프, 콕스 분석을 하였고 체질량지수와 SNP의 관련성은 Plink-1.07-dos, 버전을 이용하여 선형회귀분석 모델을 통해 연령과 성별을 보정한 뒤 $MAF < 0.01$, $HWE < 0.0001$ 를 기준으로 하여 QC(quality control)하여 정보를 추출하였다. Haploview 4.2 버전을 이용하여 X축을 체질량지수 Y축을 $-\log_{10} P$ 로 설정하여 맨하탄 플롯(Manhattan plot)을 확인하고 R을 이용하여 Clumping 으로 $r^2 = 0.001$ 이상인 연관비평형(linkage disequilibrium; LD) 들은 제외시켰다.

최종적으로 추출된 SNP들은 SAS 9.4 버전에서 로지스틱을 통해 각 SNP의 체질량지수에 대한 β 계수에 risk allele의 유전자형(Homozygous) 점수인 0,1,2를 각각 곱하여 가중치를 적용시킨 유전자 점수(WGRS; Weighted genetic risk score)를 계산하였다.

멘델리언 무작위할당 분석은 2단계로 진행하였다. 첫 번째는 선형회귀분석을 통하여 SNP에 따른 체질량지수 추정치 값을 구하고 SNP와 체질량지수의 관련성 평가는 F통계량을 기준으로 판단하였다. 두 번째는 첫 번째에서 추정된 체질량지수 값으로 폐암의 조직학적 분류에 따라 서로 관련성이 있는지 파악하였다. 멘델리언 무작위할당 분석 프로그램으로 STATA/IC 13 버전을 사용하였다.

IV. 연구결과

1. 대상자의 일반적 특성

2004년부터 2013년까지의 전향적 코호트 대상자로 구축된 한국인 암 예방 연구-II(KCPS-II)로부터 전체 코호트(Whole cohort)와 부표본(subsample)의 일반적 특성을 알아보기 위해 남, 녀로 층화하여 두 집단의 남녀 간 차이를 살펴보았다(표 1). 두 집단의 인구사회학적, 건강행동, 임상적 요인을 살펴보았으며 인구사회학적 특성은 나이와 성별을 보았고 건강행동은 흡연상태, 음주 여부 및 운동여부, 임상적 요인은 체질량지수, 허리둘레, 수축기혈압, 공복혈당, 폐암여부를 보았다.

전체 코호트의 대상자는 총 137,385명이었으며 그 중 남성 85,223명, 여성 52,162명 이었다. 부표본의 대상자는 총 14,433명이었으며 그 중 남성 9,494명, 여성 4,939명 이었다.

인구사회학적요인으로 나이에 있어 코호트 간 남성 집단의 차이와 여성 집단의 차이를 보았다. 남성 집단에서의 나이(전체코호트 41.4세, 부표본 45.7세, 이하 동일 순서)와 여성 집단에서의 나이(39.84세, 44.10세) 모두 부표본에서 높았으며 성별에 따른 코호트 간 나이의 차이가 있었다($p < 0.0001$).

건강행동요인에서는 흡연, 음주, 운동을 보았다. 코호트 간 남성 집단에서의 음주(87.63%, 85.74%), 운동(33.75%, 31.50%)과 여성 집단에서의 음주(52.94%, 44.20%), 운동(53.33%, 46.71%) 모두 성별에 따른 코호트 간 차이를 보였지만($p < 0.0001$) 전체 코호트와 부표본의 여성에서 흡연은 유의한 차이를 보이지 않았다.

임상적 요인으로는 체질량지수, 허리둘레, 수축기혈압, 공복혈당, 당뇨병 여부, 고혈압, 폐암여부 및 유형을 보았다. 남성 집단에서의 체질량지수 (전체 코호트 24.40kg/m², 부표본 24.53kg/m², 이하 동일 순서), 허리둘레(84.71cm, 85.52cm), 수축기 혈압(121.26mmHg, 122.54mmHg), 공복혈당(92.91mg/dL, 95.97mg/dL), 당뇨병(5.14%, 8.68%), 고혈압(15.01%, 20.92%), 전체 폐암 (0.34%, 2.76%), 선암(0.18%, 1.31%), 편평상피세포암(0.09%, 0.70)과 여성 집단에서의 체질량지수(22.03kg/m², 22.57kg/m²), 허리둘레(73.80cm, 75.42cm), 수축기 혈압(111.85mmHg, 114.18mmHg), 공복혈당(87.54mg/1, 88.84mg/1), 당뇨병(2.19%, 3.50%), 고혈압(8.16%, 12.57%), 전체 폐암 (0.18%, 1.74%), 선암(0.16%, 1.42%), 편평상피세포암(0.00%, 0.02%) 모두 성별에 따른 코호트 간 차이를 보였다(p<0.0001).

Table1. Selected characteristics of participants in KCPS-II

Characteristic	Men			Women		
	whole cohort (n=85,223)	subcohort (n=9,494)	p-value	whole cohort (n=52,162)	subcohort (n=4,939)	p-value
	Mean(SD)			Mean(SD)		
Age	41.42(9.56)	45.92(10.67)	<.0001	39.84(10.89)	44.10(10.86)	<.0001
BMI (kg/m ²)	24.40(2.89)	24.53(2.83)	0.0004	22.03(3.05)	22.57(3.11)	<.0001
Waist (cm)	84.71(7.78)	85.52(7.69)	<.0001	73.80(8.26)	75.42(8.49)	<.0001
SBP (mmHg)	121.26(13.01)	122.54(13.68)	<.0001	111.85(14.21)	114.18(15.38)	<.0001
FBS (mg/ℓ)	92.91(20.10)	95.97(22.72)	<.0001	87.54(14.61)	88.84(15.18)	<.0001
	N(%)			N(%)		
Diabetes(yes)	4,377(5.14)	824(8.68)	<.0001	1,144(2.19)	173(3.50)	<.0001
Hypertension(yes)	12,790(15.01)	1,986(20.92)	<.0001	4,254(8.16)	621(12.57)	<.0001
Smoking status			<.0001			0.7070
Never smoker	25,357(29.75)	2,315(24.38)		48,075(92.16)	4,548(92.08)	
Formal smoker	22,034(25.85)	2,665(28.07)		2,005(3.84)	188(3.81)	
Current smoker	37,832(44.39)	4,514(47.55)		2,082(3.99)	203(4.11)	

Table 1. (continue)

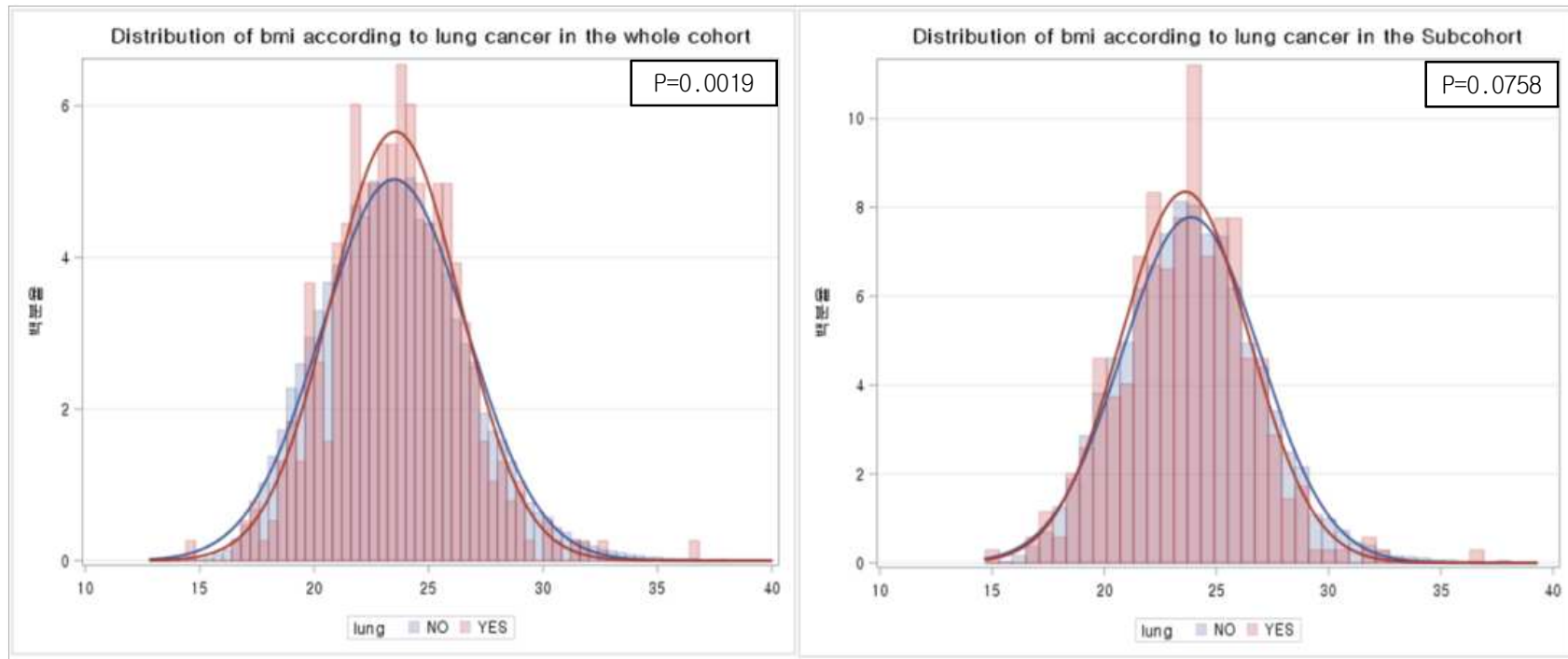
Characteristic	Men			Women		
	whole cohort (n=85,223)	subcohort (n=9,494)	p-value	whole cohort (n=52,162)	subcohort (n=4,939)	p-value
	N(%)			N(%)		
Alcohol status (yes)	74,678(87.63)	8,140(85.74)	<.0001	27,612(52.94)	2,183(44.20)	<.0001
Exercise status (yes)	28,767(33.75)	2,991(31.50)	<.0001	27,819(53.33)	2,307(46.71)	<.0001
Lung cancer	287(0.34)	262(2.76)	<.0001	95(0.18)	86(1.74)	<.0001
Adenocacinoma	157(0.18)	123(1.31)	<.0001	85(0.16)	70(1.42)	<.0001
Squamous cell carcinoma	74(0.09)	65(0.70)	<.0001	1(0.00)	1(0.02)	<.0001
Small cell carcinoma	20(0.02)			0(0.00)		
Other lung cancer	36(0.04)			9(0.02)		

* We deleted range on BMI to below 10kg/m² and to excess 50kg/m², waist to excess 125cm, SBP to excess 220mmHg, FBS to excess 400mg/dL

가. 폐암여부에 따른 전체코호트와 부표본의 체질량지수 차이

그림 1은 전체 코호트와 부표본에서 전체 대상자의 폐암에 따른 체질량지수의 차이를 본 그래프이다. 전체 코호트 내에서 폐암인 집단의 평균 체질량지수는 $23.55(2.82)\text{kg}/\text{m}^2$ 이었으며 폐암이 아닌 집단의 평균 체질량지수는 $23.50(3.17)$ 이었고 T-test 결과 $p=0.0019$ 로 통계적으로 유의하여 폐암 여부에 따라 두 집단 간의 체질량지수의 차이를 있음을 보여주었다. 부표본에서는 폐암인 집단의 평균 체질량지수가 $23.60(2.87)\text{kg}/\text{m}^2$ 이었으며 폐암이 아닌 집단의 평균 체질량지수는 $23.86(3.08)\text{kg}/\text{m}^2$ 이었고 T-test 결과 $p=0.0758$ 으로 통계적으로 유의하지 않아 폐암 여부에 따라 두 집단 간의 체질량지수의 차이가 없음을 보여주었다.

Figure 1. Distribution of BMI according to lung cancer in each cohort



2. 전체코호트(Whole cohort), 부표본(Subsample) 로지스틱 분석

가. 전체코호트의 체질량지수에 따른 조직학적 분류된 폐암 발생 위험

전체코호트에서 137,385명을 대상으로 나이, 성별, 흡연상태, 음주여부, 운동여부를 보정하여 콕스 분석을 하였다(표 2). 체질량지수 23kg/m²이상 25kg/m²미만인 집단을 기준으로 하였을 때 기준집단 대비 체질량 지수 20kg/m²미만인 집단에서 폐암에 걸릴 위험이 1.66배 높았고(HR=1.66, 95% CI=1.16-2.39), 기준집단 대비 체질량지수 20kg/m²이상 23kg/m²미만인 집단에서 편평상피세포암에 걸릴 위험이 2.42배 높았다(HR=2.42, 95% CI=1.26-4.65). 기준 집단 대비 체질량지수가 낮은 집단은 대체적으로 높은 질병 발생 위험을 보여주었고 높은 집단은 낮은 발생 위험을 보여주었다(그림 2).

Table 2. Hazard ratio for lung cancer by BMI in Korean, whole cohort of KCPS-II

Cancer type		BMI<20	20≤BMI<23	23≤BMI<25	25≤BMI<30	30≤BMI
Lung cancer	No. cancer	42/18758	128/43415	109/33530	106/37785	7/3929
	HR (95% CI)	1.66 (1.16-2.39)	1.16 (0.89-1.50)	1 (ref)	0.84 (0.64-1.01)	0.74 (0.34-1.58)
Adenocarcinoma	No. cancer	25/18758	76/43415	82/33530	68/37785	4/3929
	HR (95% CI)	1.28 (0.81-2.03)	0.96 (0.69-1.33)	1 (ref)	0.74 (0.53-1.03)	0.58 (0.21-1.59)
Squamous cell carcinoma	No. cancer	7/18758	31/43377	15/33530	26/37785	2/3929
	HR (95% CI)	2.19 (0.86-5.53)	2.42 (1.26-4.65)	1 (ref)	1.66 (0.85-3.27)	2.12 (0.48-9.41)
Small cell carcinoma	No. cancer	2/18758	4/43377	8/33530	8/37785	0/3929
	HR (95% CI)	1.12 (0.23-5.43)	0.57 (0.17-1.90)	1 (ref)	0.68 (0.23-1.96)	NA

* This model is adjusted for age, sex, smoking status, alcohol status, exercise status, diabetes status

Figure 2-a. Hazard ratio for lung cancer by BMI in Korean, whole cohort

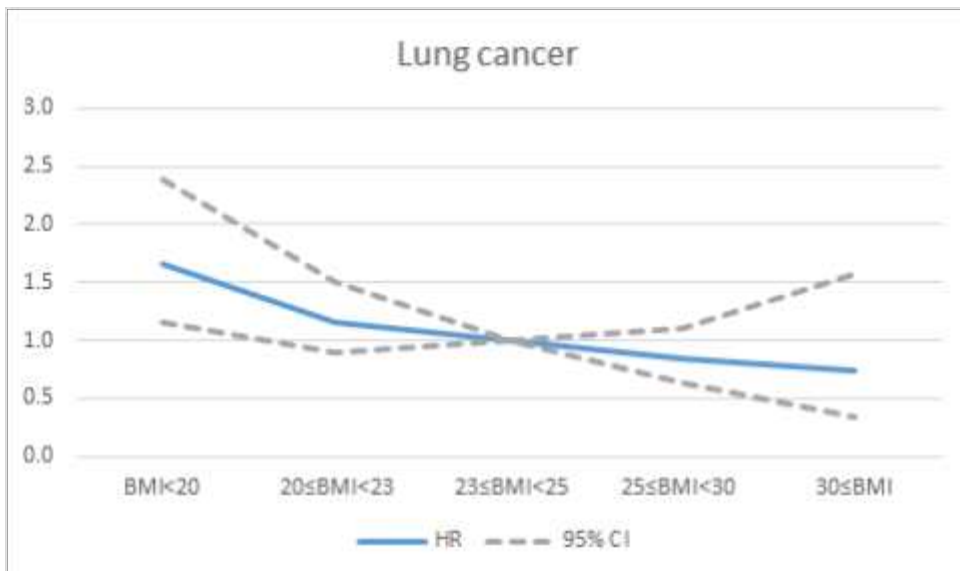


Figure 2-b. Hazard ratio for Adenocarcinoma cancer by BMI in Korean, whole cohort

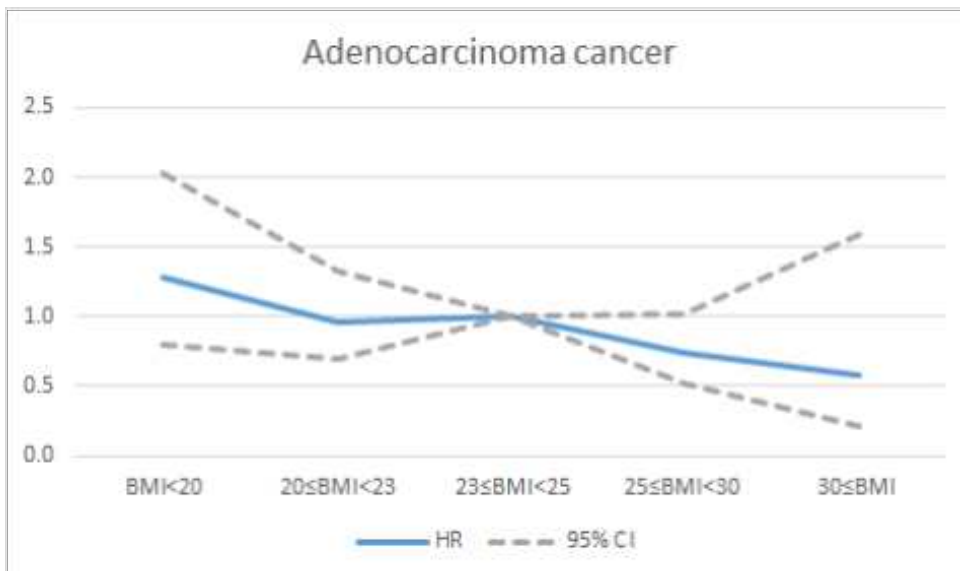


Figure 2-c. Hazard ratio for Squamous cell carcinoma cancer by BMI in Korean, whole cohort

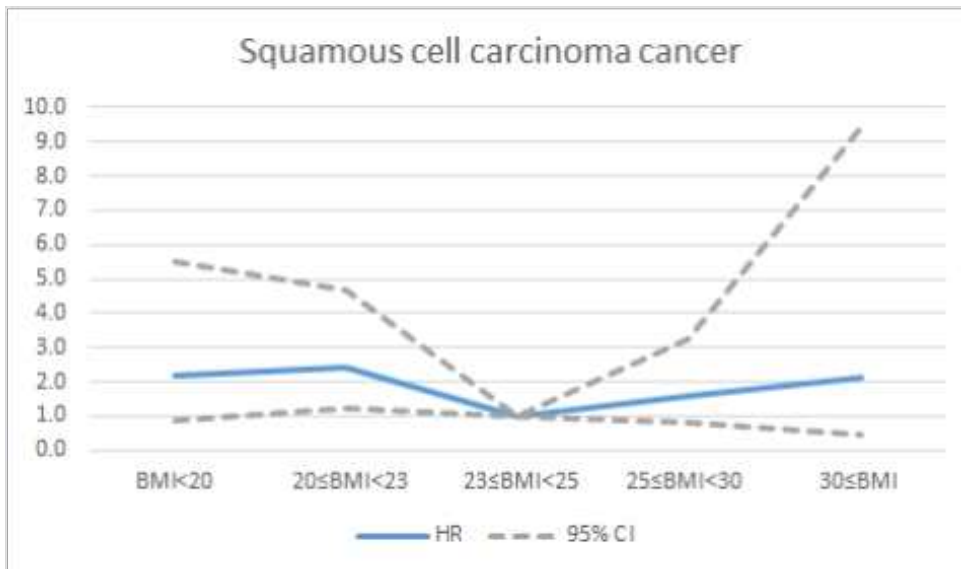
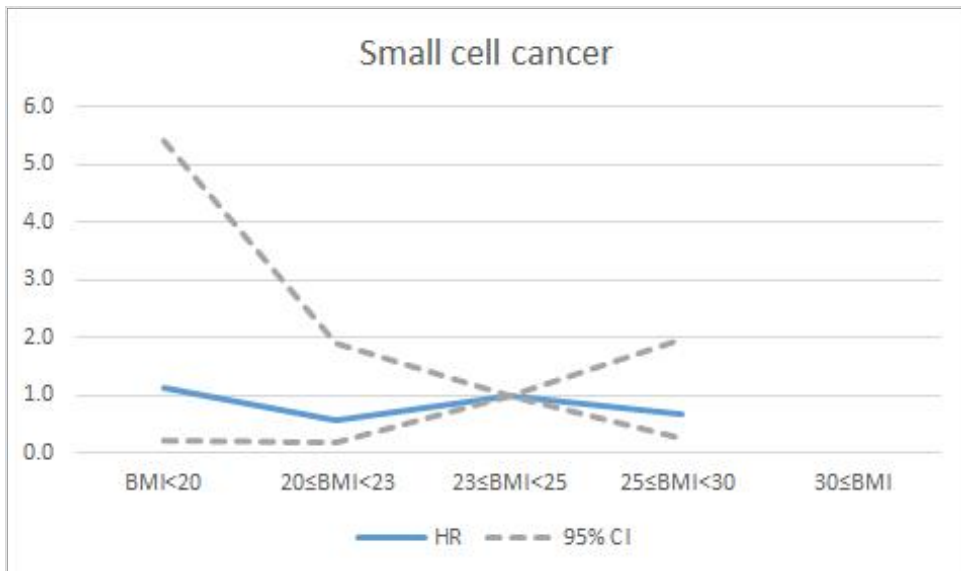


Figure 2-d. Hazard ratio for Small cell cancer by BMI in Korean, whole cohort



나. 부표본의 체질량지수에 따른 조직학적 분류된 폐암 발생 위험

부표본에서 14,433명을 대상으로 나이, 성별, 흡연상태, 음주여부, 운동여부를 보정하여 콕스 분석을 하였다(표 3). 체질량지수 $23\text{kg}/\text{m}^2$ 이상 $25\text{kg}/\text{m}^2$ 미만인 집단을 기준으로 하였을 때 기준집단 대비 체질량 지수 $20\text{kg}/\text{m}^2$ 미만인 집단에서 폐암에 걸릴 위험이 1.86배 높았고(HR=1.86, 95% CI=1.26-2.73)이었고, 편평상피세포암이 발생할 위험은 기준 집단 대비 체질량 지수 $20\text{kg}/\text{m}^2$ 미만인 집단에서 질병 발생 위험이 3.28배 (HR=3.28, 95% CI=1.21-8.90), $20\text{kg}/\text{m}^2$ 이상 $23\text{kg}/\text{m}^2$ 미만인 집단에서 2.99배 높았다(HR=2.99, 95% CI=1.38-6.48). 기준 집단 대비 체질량지수가 낮은 집단의 질병 발생 위험은 전체적으로 높았다(그림 3).

Table 3. Hazard ratio for lung cancer by BMI in Korean, subsample of KCPS-II

Cancer type		BMI<20	20≤BMI<23	23≤BMI<25	25≤BMI<30	30≤BMI
Lung cancer	No. cancer	39/1493	101/4253	98/3762	104/4501	6/424
	HR (95% CI)	1.86 (1.26-2.73)	1.16 (0.86-1.55)	1 (ref)	0.89 (0.67-1.20)	0.79 (0.34-1.81)
Adenocarcinoma	No. cancer	23/1493	53/4253	61/3762	52/4501	4/424
	HR (95% CI)	1.43 (0.87-2.36)	0.88 (0.60-1.28)	1 (ref)	0.72 (0.49-1.06)	0.75 (0.27-2.08)
Squamous cell carcinoma	No. cancer	7/1493	24/4253	11/3762	23/4501	1/424
	HR (95% CI)	3.28 (1.21-8.90)	2.99 (1.38-6.48)	1 (ref)	2.02 (0.92-4.41)	1.67 (0.21-13.30)

* This model is adjusted for age, sex, smoking status, alcohol status, exercise status, diabetes status

Figure 3-a. Hazard ratio for Lung cancer by BMI in Korean, subsample

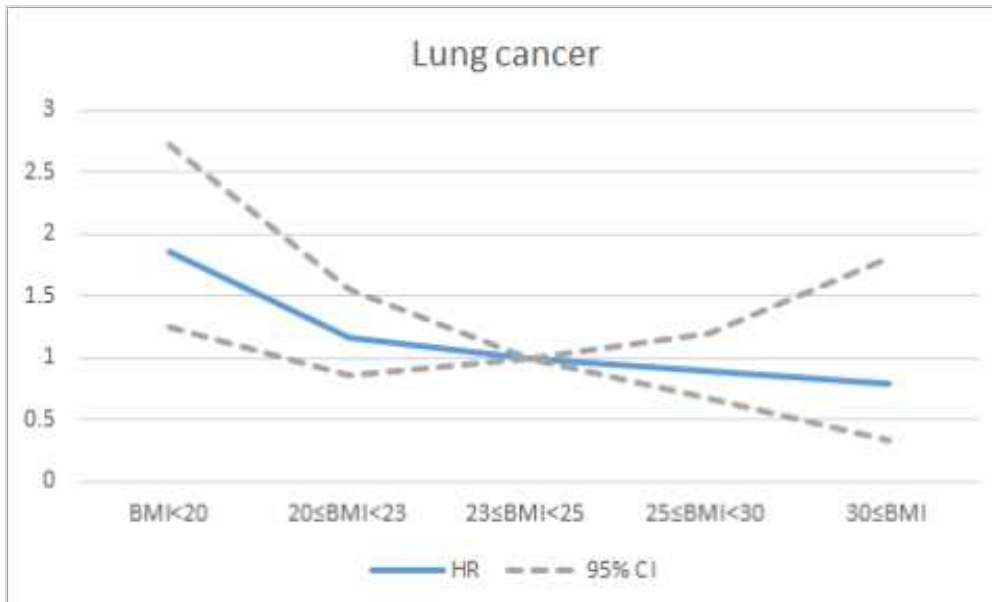


Figure 3-b. Hazard ratio for Adenocarcinoma cancer by BMI in Korean, subsample

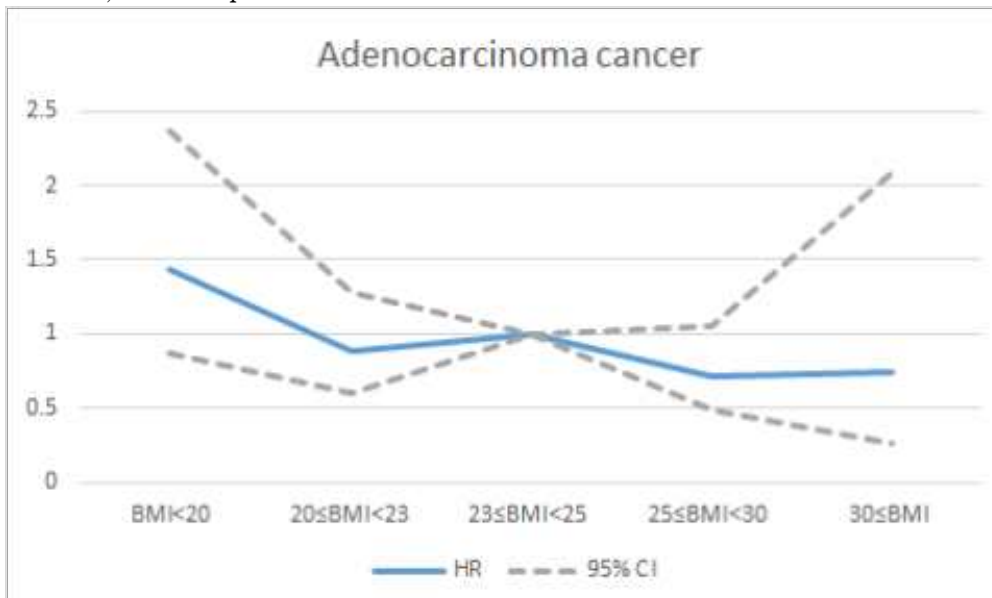
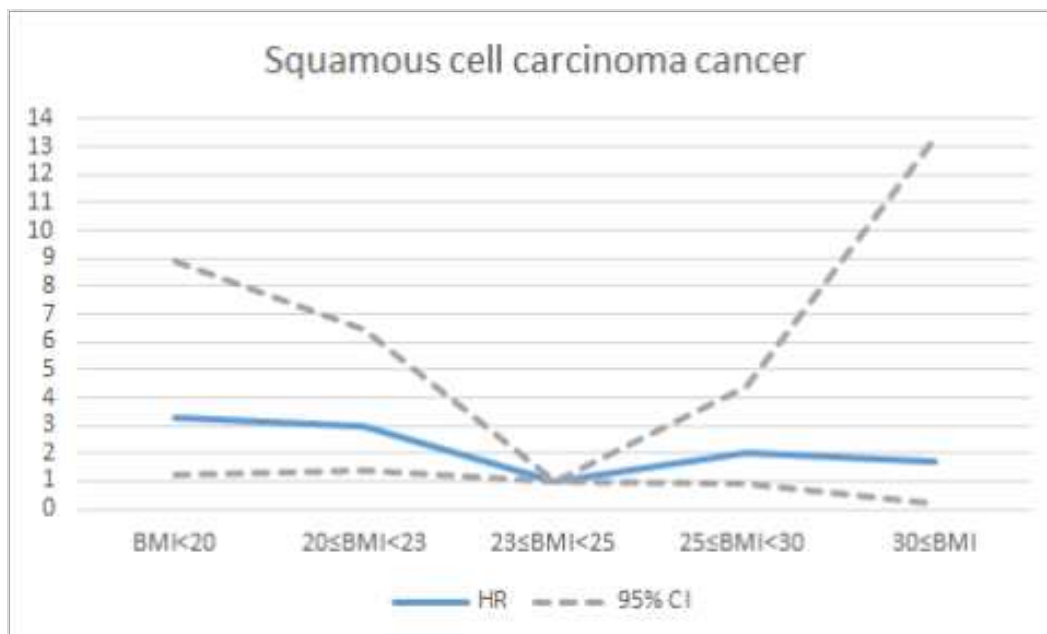


Figure 3-c. Hazard ratio for Squamous cell carcinoma cancer by BMI in Korean, subsample



다. 전체코호트의 체질량지수와 follow-up year¹⁾에 따른 폐암 발생 위험

표 2와 3의 콕스 분석은 폐암의 과거 발생을 제외하고 대상자의 코호트 참여 시점을 기준으로 하여 추적기간년수(follow-up year)를 고려한 방법으로 실행 하였지만 추적기간년수가 짧은 경우 역 인과관계의 영향이 잠재되어 있을 수 있기 때문에 추적기간년수를 코호트 참여시점을 기준으로 4년 이내의 폐암 발생, 4년 이후의 폐암 발생으로 층화분석 하였다(표 4).

전체 코호트를 추적기간년수에 따라 층화 분석한 결과 4년 이후의 폐암 발생 집단에서 체질량지수의 기준집단 대비 체질량지수 $20\text{kg}/\text{m}^2$ 미만인 그룹에서 폐암에 걸릴 위험이 1.82배 높았다(HR=1.82, 95% CI=1.19-2.78).

1) 연구 등록시점에서 진단시점까지의 기간

Table 4. Hazard ratio for Lung cancer by BMI stratified by follow-up year in whole cohort of KCPS-II

Follow-up year		BMI<20	20≤BMI<23	23≤BMI<25	25≤BMI<30	30≤BMI	
Lung cancer	0 to 4 year	No. cancer	10/18726	39/43326	31/33442	32/37700	3/3925
		HR (95% CI)	1.40 (0.68-2.88)	1.32 (0.82-2.12)	1 (ref)	0.88 (0.54-1.45)	1.10 (0.33-3.59)
	over 4 year	No. cancer	31/18747	80/43367	76/33487	74/37742	4/3926
		HR (95% CI)	1.82 (1.19-2.78)	1.12 (0.82-1.53)	1 (ref)	0.84 (0.61-1.16)	0.60 (0.22-1.65)

* This model is adjusted for age, sex, smoking status, alcohol status, exercise status, diabetes status

라. 부표본의 체질량지수와 follow-up year에 따른 폐암 발생 위험

표 4와 마찬가지로 부표본 집단에 대해 추적기간년수를 고려하여 4년 이내의 폐암발생, 4년 이후의 폐암 발생으로 층화분석을 하였다(표 5).

부표본을 추적기간년수에 따라 층화 분석한 결과 4년 이후의 폐암 발생 집단에서 체질량지수의 기준집단 대비 체질량지수 $20\text{kg}/\text{m}^2$ 미만인 그룹에서 폐암에 걸릴 위험이 2.05배 높았다(HR=2.05, 95% CI=1.30-3.23).

Table 5. Hazard ratio for Lung cancer by BMI stratified by follow-up year in subsample of KCPS-II

Follow-up year		BMI<20	20≤BMI<23	23≤BMI<25	25≤BMI<30	30≤BMI	
Lung cancer	0 to 4 year	No. cancer	10/1464	34/4186	26/3690	31/4428	3/421
		HR (95% CI)	1.64 (0.78-3.43)	1.40 (0.84-2.34)	1 (ref)	1.01 (0.60-1.71)	1.32 (0.40-4.37)
	over 4 year	No. cancer	28/1482	60/4212	60/3724	62/4459	3/421
		HR (95% CI)	2.05 (1.30-3.23)	1.09 (0.76-1.56)	1 (ref)	0.87 (0.61-1.24)	0.58 (0.18-1.84)

* This model is adjusted for age, sex, smoking status, alcohol status, exercise status, diabetes status

마. 전체코호트에서 체질량지수와 흡연 상태 및 follow-up year에 따른 폐암 발생 위험

선행연구에 따르면 체질량지수가 높은 경우 지방조직에 발암물질을 유발하는 인자들이 간혀 장기에 영향을 덜 주기 때문에 체질량지수와 흡연이 관련성이 있다는 결과를 알아내었다(Godschalk, 2002). 따라서 관찰연구에서의 흡연여부에 따라 다른 체질량지수와 폐암과의 관련성을 살펴보기 위해 흡연 상태에 따라 체질량지수와 폐암이 어떤 관련성이 있는지 분석하였다(표 6).

전체 코호트를 대상으로 흡연 상태에 따른 체질량지수와 폐암 발생을 분석한 결과, 현재흡연자의 추적기간년수 4년 이상인 집단에서 체질량지수 기준집단 대비 체질량지수 $20\text{kg}/\text{m}^2$ 미만인 집단에서 폐암이 발생할 위험이 2.58배 높았다(HR=2.58, 95% CI=1.40-4.76).

Table 6. Hazard ratio for Lung cancer by BMI stratified by smoking status in whole cohort of KCPS-II

smoke status	follow-up year		BMI<20	20≤BMI<23	23≤BMI<25	25≤BMI<30	30≤BMI
non-smoker	0 to 4 year	No. cancer	7/15723	21/32632	20/22829	20/23826	2/2299
		HR (95% CI)	1.47 (0.61-3.54)	1.08 (0.58-2.00)	1 (ref)	0.88 (0.47-1.63)	1.15 (0.27-4.92)
	over 4 year	No. cancer	14/15730	47/32658	49/22858	48/23854	3/2300
		HR (95% CI)	1.17 (0.64-2.15)	0.97 (0.65-1.46)	1 (ref)	0.89 (0.60-1.32)	0.75 (0.23-2.41)
Smoker	0 to 4 year	No. cancer	3/3003	18/10694	11/10613	12/13874	1/1626
		HR (95% CI)	1.03 (0.29-3.73)	1.66 (0.79-3.53)	1 (ref)	0.99 (0.44-2.25)	1.21 (0.16-9.42)
	over 4 year	No. cancer	17/3017	33/10709	27/10629	26/13888	1/1626
		HR (95% CI)	2.58 (1.40-4.76)	1.29 (0.78-2.15)	1 (ref)	0.85 (0.50-1.46)	0.45 (0.06-3.28)

* This model is adjusted for age, sex, smoking status, alcohol status, exercise status, diabetes status

* Non-smoker is include former smoker

바. 부표본에서 체질량지수와 흡연 상태 및 follow-up year에 따른 폐암 발생 위험

부표본 집단을 대상으로 흡연 상태에 따른 체질량지수와 폐암 발생 관계를 분석한 결과, 현재흡연자의 추적기간년수 4년 이상인 집단에서 체질량지수 기준집단 대비 체질량지수 $20\text{kg}/\text{m}^2$ 미만인 집단에서 폐암이 발생할 위험이 2.58 배 높았다(HR=2.73, 95% CI=1.43-5.23)(표 7).

Table 7. Hazard ratio for Lung cancer by BMI stratified by smoking status in subsample of KCPS-II

smoke status	follow-up year		BMI<20	20≤BMI<23	23≤BMI<25	25≤BMI<30	30≤BMI
non-smoker	0 to 4 year	No. cancer	7/1187	17/3001	18/2398	20/2722	2/256
		HR (95% CI)	1.56 (0.63-3.86)	0.97 (0.50-1.90)	1 (ref)	0.95 (0.50-1.79)	1.21 (0.28-5.27)
	over 4 year	No. cancer	12/1192	34/3018	37/2417	40/2742	2/256
		HR (95% CI)	1.30 (0.67-2.55)	0.94 (0.59-1.50)	1 (ref)	0.94 (0.60-1.47)	0.63 (0.15-2.63)
Smoker	0 to 4 year	No. cancer	3/277	17/1185	8/1292	11/1706	1/165
		HR (95% CI)	1.41 (0.37-5.34)	2.24 (0.96-5.22)	1 (ref)	1.25 (0.50-3.13)	2.00 (0.25-16.24)
	over 4 year	No. cancer	16/290	26/1194	23/1307	22/1717	1/165
		HR (95% CI)	2.73 (1.43-5.23)	1.27 (0.72-2.22)	1 (ref)	0.83 (0.46-1.49)	0.58 (0.08-4.34)

* This model is adjusted for age, sex, smoking status, alcohol status, exercise status, diabetes status

* Non-smoker is include former smoker

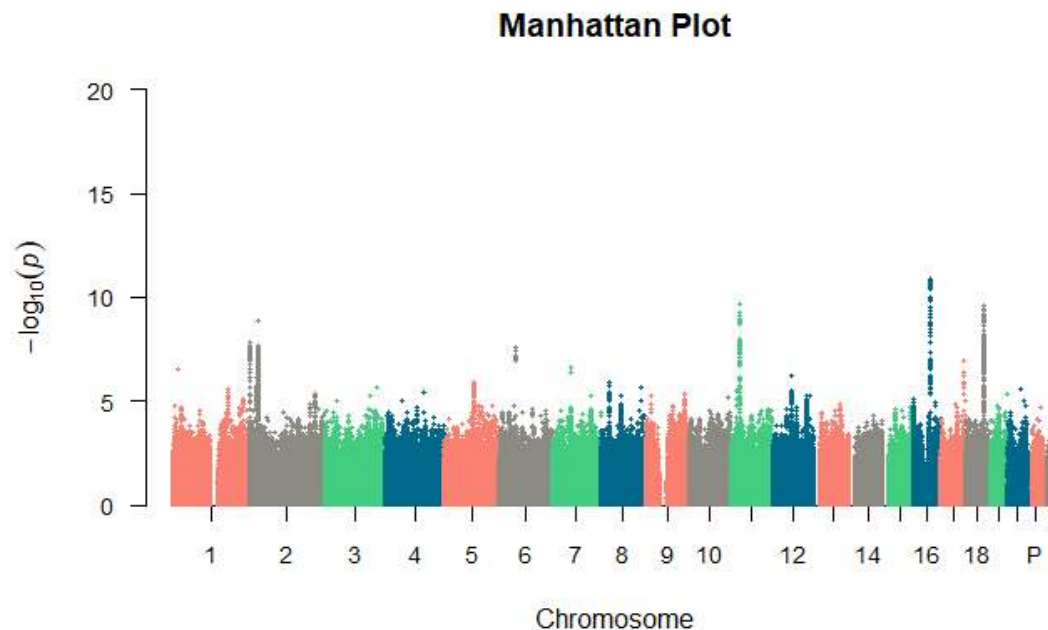
이 연구의 관찰연구에서는 선행연구와 마찬가지로 체질량지수 기준 집단 대비 체질량지수가 낮은 집단에서 폐암 발생 위험에 영향을 주는 것으로 나타났으며 비흡연자에서는 관련성이 없었고 흡연자의 경우 통계적으로 유의한 관련성을 보였다. 특히, 흡연자에서 낮은 체질량지수가 흡연 노출 기간이 길어짐에 따라 폐암 발생의 위험이 더 크고 유의하였지만 기준집단보다 높은 체질량지수에서는 통계적으로 유의한 결과가 없었다.

3. 멘델리언 무작위할당 분석

가. 체질량지수 관련 SNP

KCPS-II 대상자의 전체 코호트 159,844명 중 유전 자료를 가지고 있는 부표본 16,959명을 대상으로 하여 체질량지수 SNP를 추출하였다. 품질관리(QC, Quality control)는 대립 유전자형 빈도(MAF, Minor allele frequency) <0.01 , 하디-웨인버그 평형(HWE, Hardy-Weinberg equilibrium) <0.0001 을 기준으로 하여 진행하여 맨하탄 플롯(Manhattan plot)으로 SNP를 확인하였다(그림 4).

Figure.4 Manhattan plot for BMI(SNP of BMI is adjusted age, sex)



멘하탄 플롯 이후의 추가적인 QC(clumping 0.001 이하, $p < 10^{-8}$)를 하여 체질량지수와 연관된 총 6개의 SNP를 추출하였다(rs6744646, rs201259242, rs62405419, rs12419948, rs9923147, rs72982988)(표 8).

체질량지수 연관 SNP로 잘 알려져 있는 FTO(rs9923147) 유전자 이외에도 기준에 선행연구 된 SNP가 발견되었고 총 6개의 SNP중 rs201259242(CENPO)에서 F통계량이 가장 높았다.

나. SNP 검증

선행되어 온 문헌을 통하여 이 연구에서 추출한 SNP에 대한 유효성 검증을 하였다. 문헌검색 사이트 Pubmed와 GWAS catalog를 통해 rs6744646에 관한 문헌을 찾을 수 있었다. rs6744646는 UK biobank에서 GWAS를 통해 유전자간 부위(Intergenic region)에서 발굴되었고 체질량지수와 관련이 있었으며 통계적 유의수준은 1.20×10^{-110} 이었다(Zhu, 2020). 반면에 rs9923147은 FTO 유전자 자리에 있는 SNP지만 체질량지수와 같이 비만을 나타내는 지표에 대한 연구는 아직 이루어진 바가 없다. 하지만 FTO유전자는 비만과 매우 밀접하고 연관 있는 유전자로 잘 알려져 있다(Yang, 2012).

Table 8. GWAS Result of Body Mass Index.

Gene	Chromosome	SNP	Minor Allele	BETA coefficient	Standard error	P-value	MAF	F-statistic
TMEM18 AC093326.1	2	rs6744646	A	-0.3075	0.05596	3.95E-08	0.08845	28.08
CENPO	2	rs201259242	C	0.2095	0.03801	3.60E-08	0.2308	32.23
TFAP2B	6	rs62405419	T	0.1995	0.03621	3.68E-08	0.2864	29.26
BDNF-AS	11	rs12419948	A	-0.1834	0.03215	1.18E-08	0.4591	28.4
FTO	16	rs9923147	T	0.2359	0.04323	4.94E-08	0.164	27.81
	18	rs72982988	A	0.2278	0.04176	4.95E-08	0.2049	23.71

GWAS에서 추출된 체질량지수 관련 SNP는 risk allele 빈도에 따라 유전자형 (Homozygous)을 0, 1, 2로 정의하였고 각 대상자에서 SNP의 유전자형을 더해 GRS(Genetic Risk Score)를 만들었다(Table 9). GRS는 다음과 같이 총 3 가지 방법을 통해서 생성되었다.

1. Un-weighted GRS: 대립유전자 0, 1, 2를 각 SNP마다 총 더한 값
2. Weighted GRS: 대립유전자 0, 1, 2에 대해 나이, 성별을 보정한 회귀 분석을 통하여 얻어진 대립유전자 1 증가할 때의 평균 추정치 값을 가중치로 적용하여 각 SNP 마다 가중치 값을 총 더한 점수
3. Weighted GRS+: 대립유전자 0, 1, 2에 대해 나이, 성별을 보정한 회귀분석에서 0에서 1증가 할때의 추정치와 0에서 2증가 할때의 추정치 값을 각각 구하여 가중치로 적용하여 각 SNP 마다 가중치 값을 총 더한 점수

GRS를 이용하여 체질량지수를 추정하고 추정한 값으로 폐암과의 로지스틱 분석을 통해 멘델리언 무작위할당 분석을 진행하였다(table 9). 멘델리언 무작위할당 분석 결과 여성 집단의 체질량지수의 증가가 편평상피세포암의 위험을 증가시켰다(Un-weighted GRS(OR=6.163, CI=0.879- 43.222), P=0.072), Weighted GRS(OR=6.355, CI=0.793-50.901), P=0.087), Weighted GRS+(OR=4.412, CI=0.573-33.973), P=0.161)

Table 9. MR analyses of lung cancer by estimated BMI in subcohort of KCPS-II

	GRS type		F-statistic G-X	OR	95% CI	P-value
Lung Cancer	un-weighted	total	150.9	1.12	(0.90-1.39)	0.31
		men	98.03	1.03	(0.80-1.33)	0.82
		women	68.46	1.38	(0.92-2.07)	0.12
	weighted	total	159.67	1.04	(0.87-1.25)	0.67
		men	105.42	1.00	(0.81-1.25)	0.97
		women	67.85	1.15	(0.81-1.63)	0.45
	weighted+	total	168.64	1.02	(0.85-1.23)	0.82
		men	109.81	0.96	(0.78-1.19)	0.70
women		70.99	1.20	(0.85-1.70)	0.30	
Adeno carcinoma cancer	un-weighted	total	150.9	1.06	(0.79-1.42)	0.70
		men	98.03	0.85	(0.57-1.27)	0.43
		women	68.46	1.36	(0.89-2.10)	0.16
	weighted	total	159.67	0.97	(0.73-1.28)	0.81
		men	105.42	0.86	(0.59-1.27)	0.45
		women	67.85	1.11	(0.73-1.69)	0.64
	weighted+	total	168.64	0.95	(0.72-1.25)	0.72
		men	109.81	0.79	(0.55-1.15)	0.22
women		70.99	1.19	(0.79-1.79)	0.41	
Squamous cell carcinoma cancer	un-weighted	total	150.9	1.26	(0.88-1.80)	0.21
		men	98.03	1.20	(0.83-1.73)	0.34
		women	68.46	6.17	(0.88-43.22)	0.07
	weighted	total	159.67	1.19	(0.85-1.66)	0.33
		men	105.42	1.13	(0.80-1.60)	0.48
		women	67.85	6.36	(0.79-50.90)	0.09
	weighted+	total	168.64	1.15	(0.82-1.60)	0.42
		men	109.81	1.01	(0.78-1.54)	0.58
women		70.99	4.41	(0.57-33.97)	0.16	

V. 고찰

본 연구에서 KCPS-II (한국인 암 예방연구) 데이터를 이용한 한국인에서 체질량지수와 폐암과의 관련성을 본 결과, 체질량지수를 $23\text{kg}/\text{m}^2$ 이상 $25\text{kg}/\text{m}^2$ 을 기준집단으로 하였을 때 $20\text{kg}/\text{m}^2$ 미만 인 집단에서 폐암에 걸릴 위험이 1.66 배 높았고(HR=1.66, 95% CI=1.16-2.39) 추적기간년수 4년 이하 4년 이상으로 층화하여 분석한 결과 추적기간 4년 이상인 대상자들 중에 체질량지수 $20\text{kg}/\text{m}^2$ 인 집단에서만 유의한 양의 관련성을 보였다(HR=2.05, 95% CI=1.30-3.23). 또한 4년 이상의 발생 위험도는 체질량지수 $20\text{kg}/\text{m}^2$ 미만 인 흡연자 집단에서 더 높고 유의하게 관찰되었다(HR=2.55, 95% CI=1.40-4.78).

체질량지수가 폐암 발생에 영향을 준다는 연구는 일부 전향적 코호트 연구를 통해 알려진바 있다. 특히 폐암 발생에 미치는 영향은 흡연상태에 따른 현재 흡연자와 비흡연자로 나뉘었을 때 서로 다른 결과를 제시하고 있다. 주로 현재 흡연자에서 체질량지수가 낮을수록 폐암 발생의 위험을 증가시키는 반면, 비흡연자에서는 폐암과의 관련성이 유의하지 않거나 흡연자에서의 결과와 반대의 방향성을 보여주었다(Rauscher, 2000).

체질량지수와 폐암과의 관련성에 대한 해석에서 항상 논란이 되는 것이 있다. 그 중 항상 거론되는 것이 흡연으로 인한 체중감소인데 흡연 시 체중이 감소하면 낮은 체질량지수에서 흡연자가 많이 관찰되기 때문에 연구는 흡연에 의한 체중감소 효과를 통제한 상태에서 진행되어야한다. 하지만 체중감소 효과를 통제하기 위해서는 개인의 흡연전의 체질량지수와 흡연후의 체질량지수 정보가 필요하다. 또한 폐암은 장기간에 걸친 흡연의 영향을 받기 때문에 오

랜 기간 동안 추적한 대규모 코호트 자료가 요구된다. 그러나 실제로 장기간에 걸친 대규모 코호트 자료가 많지 않기 때문에 흡연효과를 통제하는 부분에 있어 연구에 많은 어려움이 있다. 따라서 이러한 문제를 최대한 통제하기 위해서 본 연구에서는 추적기간년수(follow-up year)를 고려하여 KCPS-II(한국인 암 예방연구) 데이터와 유전데이터를 활용하여 관찰연구와 유전분석인 멘델리언 무작위할당 분석을 병행하여 두 분석 간의 결과에 대한 차이를 비교하였다.

체질량지수와 폐암과의 관련성 분석에 대한 메타분석에 따르면 체질량지수와 폐암과의 관찰연구에서 폐암의 조직학적 특성에 따라 서로 다른 폐암 발생의 결과를 보여주고 있다. 체질량지수는 선암, 편평상피세포암 발생에 예방 효과가 있는 반면, 소세포암과는 전혀 관련이 없는 것으로 나타났다(Jing Gao et al, 2019). 선행된 메타분석에서 조직학적 특성에 따른 결과를 확인하고자 본 연구에서도 재현연구를 하였다. 그 결과 체질량지수 $23\text{kg}/\text{m}^2$ 이상 $25\text{kg}/\text{m}^2$ 미만을 기준으로 하였을 때 $20\text{kg}/\text{m}^2$ 미만인 집단에서 편평상피세포암 발생의 위험이 증가하였고($\text{HR}=3.28$, $95\% \text{ CI}=1.21-8.90$)과 $20\text{kg}/\text{m}^2$ 이상 $23\text{kg}/\text{m}^2$ 이하인 집단에서도 유의한 양의 관련성을 볼 수 있었다($\text{HR}=2.99$, $95\% \text{ CI}=1.38-6.48$).

최근 연구된 멘델리언 무작위할당 분석에서는 체질량지수가 높을수록 폐암 발생 위험을 증가시킨다고 한다. 체질량지수가 폐암 발생에 미치는 영향은 조직학적 차이에 따라 다르고 특히, 편평상피세포와 소세포암에서 높은 체질량지수가 폐암 발생의 위험을 증가시켰다(Carreras-Torres et al, 2017). 본 연구의 멘델리언 무작위할당 분석은 폐암의 조직학적 특성에 따라 대체적으로 유의한 관련성은 없었고 여성 집단에서만 체질량지수의 증가가 편평상피세포암 발생을 증가시켰다. 하지만 통계적으로 크게 유의하진 않았다

(Un-weighted GRS(OR=6.16, CI=0.88-43.22), P=0.07), Weighted GRS(OR=6.36, CI=0.80-50.90), P=0.09), Weighted GRS+(OR=4.41, CI=0.57-33.97), P=0.16).

체질량지수와 폐암과의 관련성은 아직까지도 많은 역설적인 문제와 다양한 주장이 있다. 본 연구에서 밝혀진 사실에 따르면 흡연자에서의 낮은 체질량지수가 폐암 발생의 위험을 증가시킨다는 것을 알 수 있었다. 하지만 흡연은 체중감소를 일으키기 때문에 관찰연구에서 추적기간년수를 보정하였다 해도 흡연대상자가 등록 시점에서 이미 흡연을 하고 있는 상태였기 때문에 보정에 대한 한계점이 존재한다. 또한 폐암의 조직학적 특성에 따라 분석한 결과 편평상피세포암에서만 낮은 체질량지수 집단에서 폐암 발생의 위험이 높은 것을 알 수 있었는데, 편평상피세포암은 주로 흡연에 기인하여 발생하는 암이기 때문에 낮은 체질량지수를 가지고 있는 집단에서 폐암 발생이 높게 나타나는 것은 체질량지수가 주는 효과가 아닌 흡연에 의한 효과가 더 크게 작용하여 흡연자 집단에서의 특성인 것처럼 보일 수 있다. 반면에, 멘텔리언 무작위할당 분석에서는 체질량지수가 대체적으로 폐암과의 관련성이 없다는 것으로 보아 체질량지수와 폐암발생 사이의 관련성은 없는 것으로 보인다.

1. 제한점

가. 관찰연구

선행연구에서는 정상 체질량지수 대비 높은 체질량지수의 집단에서 낮은 폐암 발생률을 보였지만 본 연구에서는 선행연구와 결과가 비슷하더라도 통계적으로 유의하지는 않았다. 그리고 폐암발생은 흡연 및 사회적, 환경적 환경의 영향으로부터 장기간 노출에 의해서 영향을 받는데 폐암 발생 대상자의 표본수가 적은 관계로 흡연 기간의 설정 한계와 발생빈도가 적었다. 따라서, 충분

한 표본이 확보가 되지 않아 표본크기 문제로 통계적으로 유의하지 않은 결과가 나올 수 있고 흡연 시 체중감소를 통제하는 부분에서도 어려움이 있었기에 연구의 한계가 있었다. 본 연구의 관찰연구에서는 폐암 발생의 관련성을 추적기간년수를 적용시킴으로써 흡연효과를 통제한 상태에서 진행하고 조직학적 특성에 따라 결과를 살펴보았지만 표본의 부족으로 유의한 관련성이 여전히 잠재되어 있을 수 있다.

나. 멘델리언 무작위할당 연구

관찰연구와 마찬가지로 멘델리언 무작위할당 연구에서도 낮은 폐암의 표본 수는 연구에 있어서 제약이 많았다. 멘델리언 무작위할당의 선행연구에서는 편평상피세포와 소세포암에서 체질량지수의 증가가 폐암 발생을 증가시킨다고 알려졌지만 본 연구에서의 부표본의 경우는 소세포암 자료가 없었고 편평상피세포암의 자료의 수도 적었기 때문에 유의한 관련성이 잠재되어 있을 수 있다. 또한 표본 수의 부족으로 흡연상태에 따른 조직학적 특성을 증화한 결과는 볼 수 없었다.

2. 강점

추적기간년수를 증화하여 분석하였기 때문에 관찰연구에서 역 인과관계를 통제하였고 기존의 멘델리언 무작위할당 연구는 two-sample analysis로 서양인을 대상으로 진행했기 때문에 one-sample로 아시아인을 대상으로 한 연구는 본 연구가 유일하다.

VI. 결론

본 연구는 한국인 암 예방연구(KCPS-II) 데이터를 이용하여 관찰연구와 멘델리언 무작위할당 분석을 시행하였다. 관찰연구에서는 흡연자에서의 낮은 체질량지수가 폐암 발생 위험을 높인다는 것을 알 수 있었지만, 인과성을 보는 멘델리언 무작위할당 분석에서는 통계적으로 유의한 관련성을 볼 수 없었다. 따라서, 전체적으로 체질량지수와 폐암과의 관련성은 없으며, 관찰연구의 낮은 체질량지수 집단에서 편평상피세포암 발생과 흡연자 집단에서 폐암 발생은 체질량지수 차이에 기인한 것이 아니라 흡연자 집단 내에서 나타내는 집단적 특성의 결과로 해석할 수 있다.

하지만 아직까지도 체질량지수와 폐암과의 관련성 분석은 흡연으로 인한 체중감소와 같은 여러 제한점과 한계점이 존재하고 있기 때문에 기존 관찰연구가 해결하지 못하는 문제점을 다양한 분석방법 그리고 유전기술의 발전과 유전연구를 통해 좀 더 인과성에 근접하는 체계적인 분석이 이루어져야 한다.

참고문헌

- Novaes FT, Cataneo DC, Ruiz Junior RL, Defaveri J, Michelin OC, Antonio José Maria Cataneo. Lung Cancer: Histology, Staging, Treatment and Survival. *J Bras Pneumol* 2008;34(8):595-600.
- Mao Y, Yang D1, He J, Krasna MJ. Epidemiology of Lung Cancer. *Surg Oncol Clin N Am* 2016;25(3):439-45.
- Yu D, Zheng W, Johansson M, Lan Q, Park Y, White E, Matthews CE, Sawada N, Gao YT, Robien K, Sinha R, Langhammer A, Kaaks R, Giovannucci EL, Liao LM, Xiang YB, Lazovich D, Peters U, Zhang X, Bueno-de-Mesquita B, Willett WC, Tsugane S, Takata Y, Smith-Warner SA, Blot W, Shu XQ. Overall and Central Obesity and Risk of Lung Cancer: A Pooled Analysis. *J Natl Cancer Inst* 2018;110(8):831-842.
- Basen-Engquist K, Chang M. Obesity and cancer risk: recent review and evidence. *Curr Oncol Rep* 2011;13(1):71-6.
- Duan P, Hu C, Quan C, Yi X, Zhou W, Yuan M, Yu T, Kourouma A, Yang K. Body mass index and risk of lung cancer: systematic review and doseresponse meta-analysis. *Sci Rep* 2015;5:16938.

Smith L, Brinton LA, Spitz MR, et al. Body mass index and risk of lung cancer among never, former, and current smokers. *J Natl Cancer Inst* 2012;104(10): 778-789.

Kucharska-newton AM, Rosamond WD, Schroeder JC, Marie A, Coresh J, Folsom AR. HDL-cholesterol and the incidence of lung cancer in the Atherosclerosis Risk in Communities (ARIC) study. *Lung Cancer* 2009; 61: 292±300.

Dewi NU, Boshuizen HC, Johansson M, Vineis P, Kampman E, Steffen A, Tjønneland A, Halkjær J, Overvad K, Severi G, Fagherazzi G, Boutron-Ruault MC, Kaaks R, Li K, Boeing H, Trichopoulou A, Bamia C, Klinaki E, Tumino R, Palli D, Mattiello A, Tagliabue G, Peeters PH, Vermeulen R, Weiderpass E, Gram IT, Huerta JM, Agudo A, Sánchez MJ, Ardanaz E, Dorransoro M, Quirós JR, Sonestedt E, Johansson M, Grankvist K, Key T, Khaw KT, Wareham N, Cross AJ, Norat T, Riboli E, Fanidi A, Muller D, Bueno-de-Mesquita HB. Anthropometry and the risk of lung cancer in EPIC. *Am J Epidemiol* 2016;184(2):129-139.

Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1-22.

Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. JAMA 2017
21;318(19):1925-1926.

Ricki Lewis. LIFE. 6th Ed. Mc Graw Hill 2009.

Sanikini H, Yuan JM, Butler LM, Koh WP, Gao YT, Steffen A, Johansson M,
Vineis P, Goodman GE, Barnett MJ, J. Hung R, Chen C & Stücker
I. Body mass index and lung cancer risk: a pooled analysis
based on nested case-control studies from four cohort studies.
BMC Cancer 2018;18(1):220.

Carreras-Torres R, Johansson M, Haycock PC, Wade KH, Relton CL, Martin
RM, Smith GD, Albanes D, Aldrich M, Andrew A, Arnold SM,
Bickeböller H, Bojesen SE, Brunnström H, Manjer J, Brüske I,
Caporaso NE, Chen C, Christiani DC, Christian WJ, Doherty JA,
Duell EJ, Field JK, Davies MPA, Marcus MW, Goodman GE,
Grankvist K, Haugen A, Hong YC, Kiemeny LK, Heijden EHF, M,
Kraft P, Johansson MB, Lam S, Landi MT, Lazarus P, Marchand LL,
Liu G, Melander O, Park SL, Rennert G, Risch A, Haura EB, Scelo
G, Zaridze D, Mukeriya A, Savić M, Lissowska J, Swiatkowska B,
Janout V, Holcatova I, Mates D, Schabath MB, Shen H, Tardon A,
Teare MD, Woll P, Tsao MS, Xifeng Wu, Yuan JM, Hung RJ, Amos
CI, McKay J, Brennan P. Obesity, metabolic factors and risk of
different histological types of lung cancer: A Mendelian
randomization study. PLoS One 2017;12(6):e0177875.

Jee YH, Emberson J, Jung KJ, Lee SJ, Lee S, Back JH, Hong S, Kimm H, Sherliker P, Jee SH, Lewington S. Cohort Profile: The Korean Cancer Prevention Study-II (KCPS-II) Biobank. *Int J Epidemiol* 2018;47(2):385-386f.

Bhaskaran K, Douglas I, Forbes H, Dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *Lancet* 2014;384, 755-765.

Rauscher GH, Mayne ST, Janerich DT. Relation between body mass index and lung cancer risk in men and women never and former smokers. *Am J Epidemiol* 2000;152, 506-513.

Dewi NU, Boshuizen HC, Johansson M, Vineis P, Kampman E, Steffen A, Tjønneland A, Halkjær J, Overvad K, Severi G, Fagherazzi G, Boutron-Ruault MC, Kaaks R, Li K, Boeing H, Trichopoulou A, Bamia C, Klinaki E, Tumino R, Palli D, Mattiello A, Tagliabue G, Peeters PH, Vermeulen R, Weiderpass E, Gram IT, Huerta JM, Agudo A, Sánchez MJ, Ardanaz E, Dorronsoro M, Quirós JR, Sonestedt E, Johansson M, Grankvist K, Key T, Khaw KT, Wareham N, Cross AJ, Norat T, Riboli E, Fanidi A, Muller D, Bueno-de-Mesquita HB. Anthropometry and the risk of lung cancer in EPIC. *Am J Epidemiol* 2016;184, 129-139.

Hidayat K, Du X, Chen G, Shi M, and Shi B. Abdominal obesity and lung cancer risk: systematic review and meta-analysis of prospective studies. *Nutrients* 2016;8, e810.

Gao J, Lin X, He Y, Fu Y, Wu Y, Liao J, Wu Y, Lian X. The Comparison of Different Obesity Indexes and the Risk of Lung Cancer: A Meta-Analysis of Prospective Cohort Studies. *Nutr Cancer* 2019;71(6):908-921.

Jehn M, Clark JM, Guallar E. Serum Ferritin and Risk of the Metabolic Syndrome in U.S. Adults. *Diabetes Care* 2004;27(10):2422-8.

Petridou ET, Sergentanis TN, Antonopoulos CN, Dessypris N, Matsoukis IL, Aronis K, Efremidis A, Syrigos C, Mantzoros CS. Insulin Resistance: An Independent Risk Factor for Lung Cancer? *Metabolism*. 2011;60(8):1100-6.

Garcia-Suastegui WA, Huerta-Chagoya A, Carrasco-Colin KL, Pratt MM, John K, Petrosyan P, Rubio J, Poirier MC, Gonsebatt ME. Seasonal variations in the levels of PAH-DNA adducts in young adults living in Mexico City. *Mutagenesis* 2011;26:385-91.

Armstrong B, Hutchinson E, Fletcher T. Cancer risk following exposure to polycyclic aromatic hydrocarbons (PAHs): a meta-analysis. *Environ Health Perspect* 2004 Jun;112(9):970-8.

Lynch SM, Rebbeck TR. Bridging the gap between biologic, individual, and macroenvironmental factors in cancer: a multilevel approach. *Cancer Epidemiol Biomarkers Prev* 2013;22(4):485-95.

Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst* 1999;91(14):1194-210.

Godschalk RWL, Feldker DEM, Borm PJA, Wouters EFM, Van Schooten FJ. 2002. Body mass index modulates aromatic DNA adduct levels and their persistence in smokers. *Cancer Epidemiol Biomarkers Prev* 2002;11(8):790-3.

Li X, Bai Y, Wang S, Nyamathira SM, Zhang X, Zhang W, Wang T, Deng Q, He M, Zhang X, Zhang W, Wang T, Deng Q, He M, Zhang X, Wu T, Guo H. Association of body mass index with chromosome damage levels and lung cancer risk among males. *Sci Rep* 2015;5:9458.

Brennan P, McKay J, Moore L, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Chow WH, Rothman N, Chabrier A, Gaborieau V, Timpson N, Hung RJ, Smith GD. Obesity and cancer: Mendelian randomization approach utilizing the FTO genotype. *Int J Epidemiol* 2009;38(4):971-5.

Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aïssi D, Wahl S, Meduri E, Morange PE, Gagnon F, Grallert H, Waldenberger M, Peters A, Erdmann J, Hengstenberg C, Cambien F, Goodall AH, Ouwehand WH, Schunkert H, Thompson JR, Spector TD, Gieger C, Trégouët DA, Deloukas P, Samani NJ. DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014;383(9933):1990-8.

Smith GD, Ebrahim S. 'Mendelian Randomization': Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease?. *Int J Epidemiol* 2003;32(1):1-22.

Bollen KA. (1996). An alternative two stage least squares (2SLS) estimator for latent variable equations. *Psychometrika* 2006;volume 61, pages109-121.

Swanson SA, Tiemeier H, Ikram MA, Hernán MA. Nature as a Trialist?: Deconstructing the Analogy Between Mendelian Randomization and Randomized Trials. *Epidemiology* 2017;28(5):653-659.

Zhu Z. Shared Genetic and Experimental Links Between Obesity-Related Traits and Asthma Subtypes in UK Biobank. *J Allergy Clin Immunol* 2020;145(2):537-549.

Yang J. FTO genotype is associated with phenotypic variability of body mass index. *Nature* 2012;490(7419): 267-272.

Abstract

Association between Body Mass Index and Lung Cancer in Korean

JeongHoon Kim

Dept. of Health Promotion & Education

Graduate School of Public Health, Yonsei University

Background Previous observational studies have shown that low body mass index increases lung cancer incidence, but in contrast, in the Mendelian randomization study, high body mass index increases lung cancer incidence and significantly correlates with squamous cell carcinoma and small cell cancer according to histological characteristics. Therefore, this study aims to find out the effect of body mass index on lung cancer incidence by comparative analysis through observational studies and Mendelian Randomization study in Koreans.

Subject and Method Using KCPS-II (Korean Cancer Prevention Study-II), the relationship between body mass index and lung cancer was

examined for 159,844 patients and 16,959 patients with genetic information. The patients of the observational study was analyzed through the Cox proportional risk model, and the patients who included the genetic data conducted a Mendelian Randomization study.

Results In an observational study of the KCPS-II total cohort, the risk of lung cancer was 1.66 times higher in the group with a body mass index less than 20 kg/m (HR=1.66, 95% CI=1.16-2.39) and in smoker over 4 follow-up year, The results of high lung cancer incidence in the group with a body mass index less than 20 kg/m² were confirmed(HR=2.58, 95% CI=1.40-4.76). However, in the Mendelian randomized study, there was no association between body mass index and lung cancer.

Conclusion According to observational studies and Mendelian Randomization study results, there is almost no direct association between body mass index and lung cancer, and what high risk of lung cancer in low body mass index group in the observational study seems like a result of cause of weight loss from smoking effect.