

연구보고서 2019-20-011

www.nhimc.or.kr

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발 (6차년도)

조경희 · 최준호 · 김동욱 · 소식영 · 조현순 · 김보현 · 최일주
김 열 · 지용호 · 배철영 · 임유리 · 배진식 · 김보미

NHIS

2019 NHIS Ilsan Hospital
National Health Insurance Service Ilsan Hospital



국민건강보험 일산병원 연구소

연구보고서

2019-20-011

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발 (6차년도)

조경희·최준호·김동욱·소석영·조현순·김보현·최일주
김 열·지용호·배철영·임유리·배진식·김보미



국민건강보험 일산병원 연구소

[저 자]

책임 연구자 : 국민건강보험 일산병원 가정의학과	조경희
공동 연구원 : 국민건강보험 일산병원 가정의학과	최준호
국민건강보험공단 빅데이터실	김동욱
국민건강보험 일산병원 의료정보팀	소석영
국립암센터	조현순
국립암센터	김보현
국립암센터	최일주
국립암센터	김열
서울대학교	지용호
메디에이지	배철영
메디에이지	임유리
EDGC	배진식
보조 연구원 : 국민건강보험 일산병원 임상연구지원팀	김보미

연구관리번호	IRB 번호
NHIS-2019-1-244	NHIMC-2019-01-010

본 연구보고서에 실린 내용은 국민건강보험공단 일산병원의 공식적인 견해와 다를 수 있음을 밝혀둡니다.

머리말

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발은 수요자 중심의 맞춤형 서비스 제공의 필요에 따라 2014년부터 시작되었다. 전 국민이 대상인 국민건강보험공단 수검자 개개인을 대상으로 하여 이용자에게 맞춤형 건강, 질병정보 제공을 통하여 질병 발생의 예방과 건강 관리에 도움이 되게 하고, 무분별하게 범람하는 많은 건강정보로부터 의료 이용자를 보호하고, 신뢰할 수 있는 건강·질병정보 제공을 하고자 시작되었다.

2-3차 년도는 (2015년- 2016년) 부적절한 진단 기준, 불필요한 동일 설문지의 반복, 검진결과와 관련 없는 개별화되지 않은 획일적 내용, 중복되는 처방메시지 및 독립적으로 개발된 각 프로그램의 분절화 등 여러 가지 문제점을 극복하기 위하여 통합정보시스템을 개발하였다.

4-5차 년도(2017년-2018년) 는 고혈압, 당뇨, 고지혈증 등 대표적 위험 요인 및 음주, 흡연, 운동 등 생활습관 교정을 위한 개별 메시지 개발과 함께, 뇌졸중, 심뇌혈관 질환 및 일부 암의 예측 모델을 개발 하였다. 대사증후군 나이 모델 개발과 함께 치매 예측 연구는 치매특별등급의 현황과 문제점을 알아보았다. 국가건강검진 DB를 통하여 개인별 맞춤형 건강 처방 메시지가 유기적으로 연계되도록 하여 세분화된 건강상태에 따라 적합한 처방메시지를 제공하고, 기존 건강 나이의 분석과 함께 몇몇 단위 질병별 예측 프로그램을 개발하여 맞춤형 개인건강기록(PHR)가 구성될 수 있도록 제시하였다.

6차 년도 『2019년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』에서는 국가적 극복 과제인 암을 중심으로 하여 위암 예측모델을 개발 및 고도화하였고, 간암 예측모형 위험 인자를 확인하였다. 심뇌혈관의 가장 중요한 위험요인인 대사증후군과 유전자의 관련성을 분석하였다. 병원 임상 기록 Data 이용을 위하여 범국가적으로 구축되고 있는 CDM data를 이용하여 위암과 헬리코박터 균의 상관성을 확인하여 CDM data 검증을 하고자 하였다.

머리말

본 연구에서는 국민건강보험공단 일산병원 가정의학과 조경희교수, 최준호전임의, 국민건강보험공단 빅데이터실 김동욱 센터장을 중심으로 하여 국립 암센터의 조현순교수, 김열교수, 최일주교수, 김보현 교수, 서울대학교 지용호 연구원, 메디에이지 배철영교수, 임유리 실장, EDGC 배진식 소장와 함께 일산병원 의무기록팀 신동교 팀장의 연구 도움을 받았다. 끝으로 본 보고서의 내용은 저자들의 의견이며 보고서 내용상의 하자 역시 저자들의 책임이며 국민건강보험공단 일산병원 연구소의 공식적인 견해는 아님을 밝혀둔다.

2020년 1월

국민건강보험공단 일산병원장

일산병원 연구소 소장

김성우
최운정



목차

요 약	1
제1장 서 론	9
제1절 연구 배경 및 필요성	11
제2절 연구의 목적	12
제2장 이론적 고찰	13
제1절 위암 질환 예측	15
제2절 간암 질환 예측	18
제3절 대사증후군의 유전형과 표현형 융합분석	22
제4절 위암과 헬리코박터 관련성	24
제3장 연구 내용 및 방법	27
제1절 연구 내용	29
제2절 연구 방법	30
제3절 각 세부 과제별 연구 방법	32
제4장 분석결과	43
제1절 위암 예측 모형 연구	45
제2절 간암 예측 모델 개발 연구	54
제3절 대사증후군의 유전형과 표현형 융합 분석	58
제4절 위암과 헬리코박터 관련성 분석을 통해 CDM모델의 정합성 검정	73
제5장 결론 및 제언	75
제1절 결 론	77
제2절 제 언	79
참고문헌	83

표목차

<표 2-1> 가족력관련 위암 위험요인이 포함된 관련 연구	17
<표 2-2> 성별 주요 암 발생 현황, 2016년도 국가암등록통계자료	18
<표 2-3> 연령 군별 주요 암발생률: 남녀전체, 2016, 국가암등록통계자료	18
<표 2-4> multi variable cox model for HCC occurrence and CAMD score ...	19
<표 2-5> Factors associated with HCC by cox proportional hazards regression	20
<표 2-6> Cox models evaluating risk factors for the ADREESS-HCC model	20
<표 2-7> 이론적 고찰을 통해 확인된 간암 위험요인	21
<표 2-8> 각 연구에서 사용된 헬리코박터균 관련 위암 발생 위험 인자	25
<표 3-1> 맞춤형 DB 세부특성	30
<표 3-1> 맞춤형 DB 세부특성	33
<표 3-2> 가족력 분류 체계 비교	35
<표 3-3> 출생시점별 가족력 데이터베이스의 연계율(공단빅데이터실 추출 자료)	36
<표 3-4> modified ATP III criteria	39
<표 3-5> 대사증후군 유전적 위험도 분석을 위한 유전자 항목과 종류	40
<표 3-6> 맞춤 가이드(영양과 운동)을 위한 유전자 분석 항목과 종류	40
<표 4-1> 위암 예측 모델 분석 대상자의 특성(2009-2010)	45
<표 4-2> 위암검진 수검자의 가족력	47
<표 4-3> 위암발생자의 가족력	47

<표 4-4> 본인 위암 여부와 가족 위암 발생의 가족관계별, 암유형별 분포(2009-2010)	48
<표 4-4-1> 본인 위암 여부와 가족 위암 발생의 가족관계별, 암유형별 분포(2009-2010)	49
<표 4-5> 본인 위암 여부와 가족 위암 발생의 가족관계별 분포(2009-2010)	49
<표 4-6> 가족력 DB와 문진에서의 가족력 정보와의 비교	50
<표 4-7> 가족력 DB와 문진에서의 가족력 정보와의 비교 (missing 제외)	51
<표 4-8> 위암, 위암 관련 암, 기타 암 가족력에 따른 본인 위암 위험도의 단 변수 회귀 분석	51
<표 4-9> 위암-위암관련암-기타암으로 우선순위를 부여했을 때, 위암 및 위암 관련 암, 기타 암 가족력에 따른 본인 위암 위험도의 단변수 분석	53
<표 4-10> 본인 암 발생을 종속변수로 가족력 등 위암 위험 요인들과의 관련성	53
<표 4-11> 일반 검진자 대상 분석	55
<표4-12> 일반검진자 대상 분석(생활 습관)	56
<표4-13> 간암검진자 대상 분석	57
<표4-14> 간암검진자 대상 분석(생활습관)	58
<표 4-15> 대사증후군 여부에 따른 연구대상자 특성	59
<표 4-16> 유전적 위험도에 따른 연구대상자 분포	60
<표 4-17> 대사증후군 위험요인에 따른 연구대상자 분포	61
<표 4-18> 대사증후군 나이에 따른 연구대상자 분포	62
<표 4-19> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포(빈도)	63
<표 4-20> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포	63
<표 4-21> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포	63
<표 4-22> 유전적 위험도와 대사증후군 기준 타입 별 유전적 위험도 평균 수치	64

<표 4-23> 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포(빈도)	64
<표 4-24> 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포	65
<표 4-25> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포	65
<표 4-26> 유전적 위험도와 대사증후군 나이 기준 타입 별 유전적 위험도 평균 수치	65
<표 4-27> 유전적 위험도 '낮음' - 생활습관에 따른 대사증후군 발생률	66
<표 4-28> 유전적 위험도 '중간' - 생활습관에 따른 대사증후군 발생률	67
<표 4-29> 유전적 위험도 '높음' - 생활습관에 따른 대사증후군 발생률	67
<표 4-30> 재조정된 유전적 위험도에 따른 연구대상자 분포	68
<표 4-31> 재조정된 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포(빈도)	69
<표 4-32> 재조정된 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포 ..	69
<표 4-33> 재조정된 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포 ..	69
<표 4-34> 재조정된 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포(빈도)	70
<표 4-35> 재조정된 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포	70
<표 4-36> 재조정된 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포	70
<표 4-37> 재조정된 유전적 위험도 '낮음' - 생활습관에 따른 대사증후군 발생률	71
<표 4-38> 재조정된 유전적 위험도 '중간' - 생활습관에 따른 대사증후군 발생률	71
<표 4-39> 재조정된 유전적 위험도 '높음' - 생활습관에 따른 대사증후군 발생률 ..	72
<표 4-40> ulcer 보정 CDM 결과	74
<표 4-41> ulcer, atrophic, gastritis, age, sex 보정 CDM 결과	74

그림목차

[그림 2-1] 주요 암 종별 발생분율, 2016	15
[그림 2-2] 성별 주요 암 종별 발생분율, 2016	16
[그림 2-3] CDM database 구성도	32
[그림 3-1] 위암예측모형 연구 설계 도식도	34
[그림 3-2] 위암예측모형 연구 대상자 선정 과정	34
[그림 3-3] 일반검진 연구 대상자 선정	36
[그림 3-4] 간암검진 연구 대상자 추출	37
[그림 3-5] 간암 건진 대상자 선정	37
[그림 4-1] 대사증후군 유전적 위험도에 따른 연구대상자 분포	60
[그림 4-2] 대사증후군 위험요인에 따른 연구대상자 분포	61
[그림 4-3] 대사증후군 나이에 따른 연구대상자 분포	62
[그림 4-4] CDM 연구 대상자 선정	73

요약

요약

1. 연구 배경 및 목적

본 연구는 2014년도부터 시작 된 『개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』 연구의 6차년 도 연구이다. 맞춤형 DB 및 암 검진 자료를 연계하여 질병 예측 프로그램을 개발하고, 대사 증후군의 표현형과 유전형의 상관성 분석을 통해 현재 건강인 이용자에게 개인별 맞춤 건강 정보를 제공하고 이를 통해 건강생활 습관을 유도하여 질병 발생의 예방과 삶의 질 향상에 도움이 되고자 한다.

2. 연구 내용 및 방법

본 연구에서는 각 연구에 맞추어 건강보험공단 빅데이터의 맞춤형 자료를 이용하여 진행하였으며 대사증후군 분석연구에서는 개인 유전자 정보를, 헬리코박터균과 위암 상관성 관련 연구에서는 CDM(Common data model)을 이용하였다.

맞춤형 자료를 사용한 건강검진결과와 산정특례 등록 자료와 연계하여 위암예측 프로그램을 개발하였고 가족력에 대해 행정안전부의 가족력 데이터를 접목하여 예측 프로그램을 고도화 하였다. 간암은 2009,2010년 구성된 전수 연계조사 자료와 함께 간암 검진자만을 대상으로 하여 국립암센터와 공동 연구로 공단데이터를 근거로 한 간암 예측 연구를 진행하였다. 위암과 헬리코박터균의 관련성을 연구한 CDM 모델 구축은 일산병원 의무기록팀과 함께 진행하여 CDM 데이터 정합성 검증을 하였다. 대사증후군의 표현형과 유전형의 상관성 분석 연구는 2019년 일산병원 검진센터 방문자 대상자 중 동의를 받아 대상 군을 선정하였다. 일반국가검진 검사 및 설문지 데이터에 유전데이터를 접목하여 대사증후군 발병과 생활습관 및 유전체의 관련성에 대하여 연구하였다.

3. 연구 결과

위암 가족력 모형 연구에서는 2009-2010년 동안 국민건강보험공단 맞춤형 DB의 건강검진결과 및 위암검진 수검자 대상으로 2017년까지 위암발생여부를 산정특례 등록 자료를 연계하여 분석하였다. 고도화된 위암 예측 모형 개발을 위해 행정안전부 자료와 연계하여 구축한 가족력 데이터베이스를 활용하였고 추적기간 동안 위암 발생과의 관련성을 분석하였다. 연구 대상의 특성은 이전연도 연구를 참고하여 성별, 나이, 체질량 지수, 수입, 동반질환 (당뇨, 고혈압, peptic ulcer, atrophic gastritis), 흡연, 음주, 운동 등을 모델에 포함하였다. 추적기간 7년간 위암 발생을 확인한 결과 전체 5,017,985명 중 41,821명(0.8%), 남성 2,178,293명중 28,235명 (1.3%) 에서 위암 발생이 확인되었으며, 여성 2,830,903명중 13,586명 (0.5%) 에서 위암 발생이 확인되었다. 부모의 위암, 위암 관련암, 기타 암과 본인 위암여부와 관련성은 없거나 역 관련성을 보였고 자녀 및 형제가 모두 가족력이 있는 경우 본인 위암 발생의 분율이 높게 나타났다. 위암 가족력 정보와 본인위암 여부를 교차분석한 기술통계에서 아버지와 어머니, 1촌의 위암 가족력, 위암 관련 암 가족력, 기타 암 가족력에서 관련성과 본인 위암과의 관련성이 없는 것으로 나타났다. 이번 연구에서 수집한 가족력 DB정보와, 2009-2010년 문진을 통해 수집한 가족력 정보를 결측을 제외한 상태에서 비교하였을 때, 가족력 DB에서 위암 가족력이 없다고 한 대상자중 문진항목에서도 위암 가족력이 있다고 대답한 분율이 22.7%, 두 자료원에서 공히 위암 가족력이 있다고 한 분율이 44.5%, 위암 가족력은 28.0%, 기타 암 가족력은 38.65%로 많은 차이가 나타났다. 본인 암발생을 종속변수로 가족력 등 위암 위험 요인들과의 관련성을 다변수 회귀분석한 결과, BMI, income, diabetes, HTN, endoscopy, smoking, active physical activity 등을 통제된 이후에는 위암 가족력에 따른 본인 위암 발생은 1.14배 유의하게 높은 것으로 나타났다.

간암 질환 예측 연구에서는 일반 검진자를 대상으로 간암과 나이, 성별, 수입, 비만도, 허리둘레, 혈압, 당뇨, 고혈압, 사망률, 혈당, AST, ALT, 콜레스테롤 그리고 흡연, 음주, 운동 등 생활 습관과의 관련성을 보았다. 간암은 남성, 나이는 70대에서 가장 많이 발생하였고, 체질량지수, 허리둘레, 혈당, 고혈압과 AST, ALT 등과 상관성을 보였다. 생활습관의 경우 흡연력 존재, 음주력이 높은 경우, 운동량이 더 많은 경우에서 간암발생이 높게 나타났다. 간암 검진자를 대상으로 하여 간초음파 관찰 소견, AFP, HBV, HCV, 생활습관(음주, 흡연, 운동), 혈압, 혈당, 체질량 지수, 총콜레스테롤, 중성지방, 고밀도

콜레스테롤, 저밀도 콜레스테롤, 크레아티닌, 콩팥청소율, AST, ALT과 간암 발생과 관련성을 보았다. HBV, HCV 양성인 군에서 간암발생이 높았고, 남성에서 간암 발생이 증가하였으며, 나이는 60대에 가장 높은 발병을 보였다. 체중, 혈압, creatinine, 간수치인 AST,ALT,rGTP는 간암 발생율을 증가시키는 경향을 보였으며 콜레스테롤 수치인 저밀도 콜레스테롤, 중성지방, 콜레스테롤은 간암 군에서 더 낮게 나타났다. 고밀도 콜레스테롤이 낮은 경우 간암발생이 높았다. 흡연, 음주, 간암 발생을 증가시키는 경향을 보였으나, 운동은 간암발생과 관련성이 없었다.

대사증후군의 유전형과 표현형 융합분석 연구에서는 대사증후군 유병률은 249명 중 66명으로 26.51%로 나타났다. 대사증후군 관련 유전자를 대상으로 나이, 과거력, 신체 측정값(체질량지수, 허리둘레, 수축기혈압, 이완기혈압), 혈액 검사 측정값(공복혈당, 중성지방, 고밀도콜레스테롤수치), 생활습관 항목(운동, 흡연, 음주), 대사증후군 발생률과의 관계를 분석하였다. 유전적 위험도가 '낮음'(27.27%)이나 '중간'(23.85%)보다는 '높음'(50.00%)에서 대사증후군 발생률이 더 높았고 유전적 위험도가 '낮음'(33.33%)이나 '중간'(30.28%)보다는 '높음'(50.00%)에서 대사증후군 나이가 '나쁨'이 더 많았다. 유전적 위험도 '높음' 타입을 기준으로 정상군(1.67%), 위험군(2.44%), 대사증후군(6.06%)으로 갈수록 유전적 위험도 '높음' 비율이 점점 높아지는 양상을 보였고 유전적 위험도 '높음' 타입을 기준으로 대사증후군 나이가 ' 좋음'(0%), 대사증후군 나이가 '평균'(3.05%), 대사증후군 나이가 '나쁨'(4.94%)으로 갈수록 유전적 위험도 '높음' 비율이 점점 높아지는 양상을 보였다. 그리고 대사증후군 타입 별 유전적 위험도 평균 수치는 정상군(6.42), 위험군(6.46), 대사증후군(6.58)으로 높아지는 경향을 보여주고 대사증후군 나이 타입 별 유전적 위험도 평균 수치는 대사증후군 나이 ' 좋음'(6.41), 대사증후군 나이 '평균'(6.49), 대사증후군 나이 '나쁨'(6.49)으로 ' 좋음'보다는 '평균'과 '나쁨' 타입의 유전적 위험도 평균 수치가 더 높은 경향을 보여주었다. 또한 유전적 위험도 '낮음' 타입에서 흡연을 하는 집단, 음주를 하는 집단, 저활동군 집단의 대사증후군 발생률이 더 높았다.

위암과 헬리코박터 관련성 분석 연구에서는 CDM모델의 적합성 검증한 결과 2010년 이전 일산병원에서 위내시경을 실시하여 헬리코박터균 양성, 위궤양, 십이지장궤양, 위축성위암 진단받은 대상자 총 6420명을 대상으로 분석한 결과 성별, 나이,위/십이지장궤양,위축성위암을 보정하여 CDM 분석 결과는 헬리코박터 양성군이 39명, 헬리코박터 음성군이 23명이었으며 헬리코박터균 없는 군에 비해 있는 군의 Harzard ratio는 1.48으로 높았으나 이는 통계적으로 유의하지 않았다.

4. 결론 및 제언

전체 대상자의 본인 위암 여부와 가족위암 발생의 가족관계별 분포를 분석한 기술통계 결과에서 부모의 위암 가족력과 본인의 위암 여부의 관련성이 역 상관성을 보이는 것처럼 나온 결과는 이번 연구에서 후향적으로 추적한 가족력데이터베이스의 한계로 설명할 수 있을 것으로 판단된다. 즉 연구에 포함된 위암검진 수진자의 대부분은 40대 이상 중년들이며, 이들의 출생시점인 1970년대 이전의 경우 2000년대 이후에 비해 출산율이 높았고 외동인 가족이 많지 않아 형제 자매가 있을 확률이 높았다. 부모세대의 연계율은 매우 낮은 반면에, 형제 자매의 가족력은 자료가 집계가 잘되어 그로 인한 가족력의 위험이 높아 보일 수 있으므로, 부모에 의한 가족력과의 직접 비교를 하기에는 문제가 있다. 추후 연구에서는 다른 암종에서도 비슷한 경향의 문제점이 드러나는지를 확인할 필요가 있으며, 연계율이 향상된 더욱 정교하게 구축된 데이터가 필요하다.

대사증후군 유전자 분석 연구에서는 총 249명의 연구대상자 정보를 통해 대사증후군의 표현형과 유전자형을 융합 분석할 수 있었다. 본 연구는 유전적 요인과 생활습관이 대사증후군 발생과 관계가 있다는 것을 파악했다는 것에 그 의의가 있다고 할 수 있다. 약 3개월간 모집된 자발적 동의자를 대상으로 분석하였기에 모집할 수 있는 인원의 한계로 인하여 적은 인원으로 분석이 진행 되었으며 통계적 유의성을 입증하는데 한계점이 있었다. 이에 대해 연구 대상자 모집 기간 및 모집단 수를 늘려 통계적 유의성을 보완하는 연구를 통하여 추후 연구 진행이 필요할 것으로 사료된다. 또한 기존에 분류되어 있는 대사증후군 유전적 위험도 타입 분류 기준에 대해 대사증후군의 유전적 위험군을 잘 선정할 수 있는 고도화 연구가 필요할 것으로 판단된다. 모집단의 추가와 함께 대사증후군의 유전적 위험도 분류가 더욱 정교하게 수정 및 보완이 진행된다면 생활습관과 유전적 요인의 대사증후군 발생과 관련한 좀 더 고도화된 분석 결과를 기대해 볼 수가 있을 것이다.

간암 질환 예측 연구는 가장 중요한 요소인 나이, B형간염, C형간염 외에도 국가 간암 검진에서 사용되는 간초음파 관찰 소견, AFP와 일반검진의 생활습관(음주, 흡연, 운동)을 이용하여 간암 예측을 진행하였다. 실제적인 간암 위험 예측 모델을 구축하기 위해서는 가장 중요한 인자인 B형간염, C형간염 환자들의 전수 자료를 통한 고도화된 분석이 필요하겠다. 또한 표준화 된 초음파 결과 입력의 이용 및 AFP 정량 검사 결과의 이용을 통하여 예측력을 증가시켜 간암 발생 고위험군을 정교하게 선정한 후속

연구가 필요하다.

헬리코박터균과 위암의 상관성을 분석한 CDM data 정합성 검증한 연구는 최근 빅데이터 연구가 시작된 이후 본격적으로 진행된 국민건강보험 일산병원 자체 연구로서 의의가 크다. 지속적인 CDM 데이터 검증을 통해 현재 CDM 데이터의 한계점을 파악하여 데이터 표준화 및 의료기관 간의 데이터 공유에 이바지하게 될 것이다.

6차년도 (2019년) 맞춤형건강관리 연구에서는 각 부분의 연구의 고도화를 시도하였고, 일반 수검자 및 암검진자들에게 만성질환 예방과 건강 증진에 도움이 될 수 있는 지표가 되도록 시도하였다, 또한 병원 임상기록 이용을 위하여, 다기관 공동 연구의 도구로써 CDM을 이용한 연구의 데이터 검증을 시도해 보았고, 미래에 본격적으로 도입 될 정밀의학의 유전자 분석을 사용하여 미래 질병 예측의 방향을 제시하고자 하였다.

제 1 장

서 론

제1절 연구 배경 및 필요성	11
제2절 연구의 목적	12

제1장 서론

제1절 연구 배경 및 필요성

현재 세계적 연구의 트렌드는 일개 병원 데이터를 이용하여 개별화된 방법으로 연구를 진행하는 것이 아니라 표준화된 방법으로 다량의 정보를 수집하여 빅데이터를 통해 분석을 진행하는 것이다. 국민건강보험공단에서도 이러한 변화의 흐름에 맞추어 보험공단의 빅데이터를 이용하여 맞춤형 서비스를 건강iN(hi.nhis.or.kr)을 통해 구축해 다른 추가적 검사를 진행하지 않고 국가검진 데이터만으로 일반인들에게 질병예측에 관한 정보를 전달하여 국민 건강을 향상시키고자 노력해왔다.

본 연구는 개인건강기록 서비스(PHR)의 기초 연구로 시작하여 2014년, 2015년, 2016년, 2017년, 2018년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발에 이은 후속 6차년도 연구이다. 올해 연구에서는 국민건강보험공단의 빅데이터실에서 보유하고 있는 맞춤형 DB를 통하여 수요자 중심의 맞춤형 서비스를 개발하는 것으로 위, 간암 예측 프로그램의 업데이트와 추가 개발을 진행하였다. 또한 국가 검진 수검자를 대상으로 대사증후군의 유전자 검사를 통하여, 표현형과 유전형의 상관성을 분석하여 추후 정밀의료 시대를 맞이하여 유전자 연구의 도입을 위한 기초 연구를 시도해 보았다.

『2014년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』에서는 과거 건강iN 서비스의 여러 문제점들을 제시하였고, 이를 바탕으로 『2015년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』에서는 의료 이용자들에게 개인의 건강검진 정보를 통한 통합 처방 처방메세지를 제공함으로써 통합된 개인 건강기록서비스(PHR, personal health record)를 구축하였고 또한 새롭게 뇌졸중 위험 예측 모델 과 심장질환 위험 예측 모델을 개발하여 건강iN 서비스의 'My Health Bank'의 이름으로 탑재하여, 의료 이용자들의 건강위험요소를 파악하여 생활습관을 교정할 수 있도록 하였다. 『2016년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』에서는 이전 개발한 당뇨 위험 예측 모델 개발 및 타당도 확인, 뇌졸중 및 심장질환 위험 예측 프로그램 타당도 및

업데이트, 건강나이 모델 개발 및 검증 한 바 있다. 『2017년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』에서는 위암, 폐암질환 예측모형 개발, 당뇨 질환 위험 예측 모델 고도화 및 건강인 웹페이지 탑재 알고리즘 개발, 심뇌혈관 질환 예측 모델 고도화, 각 생체지표의 연령별, 성별 국가 현황과 생활습관 항목과의 연관성 조사 및 건강노화나 이 고도화 개발을 진행하였다. 『2018년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』에서는 국가적 극복 과제인 암을 중심으로 하여 위암 및 폐암의 예측모델을 개발 및 고도화하였고, 심뇌혈관의 가장 중요한 위험요인인 대사증후군의 현황과 대사증후군 나이를 연구하여 개인의 대사증후군 정도를 쉽게 계량화 하는 방안을 모색해 보았다. 또한 미래 극복과제인 치매에 대하여 산정특례를 이용하여 위험요인 현황을 알아보았다.

건강in과 같은 개인 건강기록서비스(PHR)는 개인이 어플리케이션만으로 본인의 건강 정보를 확인하고 행동변화를 유도할 수 있고, 대사증후군 관련 유전자 분석을 통해 본인의 취약한 유전인자를 파악하면 이를 보충하기 위해 특정화된 생활습관에 집중하여 건강 관리를 할 수 있게 한다. 이는 궁극적으로는 국민 건강의 향상을 가져오게 되며, 국가적으로는 질병의 예측을 통한 예방적 의료의 강화와 함께 전체 의료비가 절감하게 될 것이다.

제2절 연구의 목적

본 연구는 『2014년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』 연구에 이은 후속 6차년도 연구로써, 맞춤형 DB를 통하여 질병 예측 프로그램을 추가 업데이트 하고, 대사증후군의 유전형 분석을 통한 대사나이의 정합성 검증, CDM tool을 이용한 위암과 헬리코박터의 연관성을 분석하여, 이용자의 건강, 질병정보 제공을 통하여 질병 발생의 예방과 질병 관리에 도움이 되고자 한다.

- 1) 위암 질환 예측 모델 가족력 연계 분석
- 2) 간암 질환 예측 연구
- 3) 대사증후군 표현형과 유전형 관련성 검증 및 연계 이용 방향 제시
- 4) 위암과 헬리코박터 관련성 분석을 통해 CDM모델의 정합성 검증

제2장

이론적 고찰

제1절 위암 질환 예측	15
제2절 간암 질환 예측	18
제3절 대사증후군의 유전형과 표현형 융합분석	22
제4절 위암과 헬리코박터 관련성	24

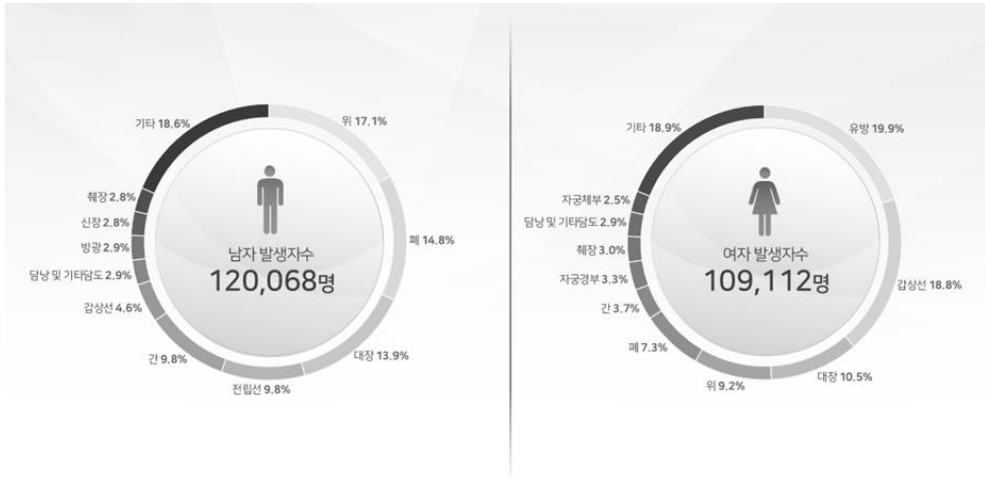
제2장 이론적 고찰

제1절 위암 질환 예측

위암 발생률은 전 세계적 추세로는 감소하고 있지만 암사망률에 있어 3위를 기록하고 있는 암이며, 국내에서도 가장 발생률이 높은 악성종양 중 하나이다. 2016년 기준으로 남녀 암 발생분율 중 13.3%에 해당하며, 남성에서 17.1%, 여성에서 9.2%의 분율을 보이고 있다. 암종별 5년 생존율 추이가 집계된 93년 이래로 2016년까지 꾸준히 증가하여 2016년 기준으로 전체에서 76.0% (남성 76.9%, 여성 73.8%)의 생존율을 보이고 있다². 이와 같이 위암은 조기 발견으로 위한 사망률 감소와 생존율 증가가 큰 암종이므로 예측 모형을 통한 조기 발견과 예방이 더욱 중요하다. 일반적으로 잘 알려진 위암의 위험 요인은 성별, 나이, 인종, 지역, 유전적 요인, 식이, 헬리코박터균 감염, 위림프종, 위 절제술 이력, 악성 빈혈, 메네트리에병력, 흡연을 비롯한 환경적 요인, 화학적 발암원 등이 복합적으로 작용하는 것으로 보고됐다³.



[그림 2-1] 주요 암 종별 발생분율, 2016



[그림 2-2] 성별 주요 암 종별 발생분율, 2016

유전적 요인이 종양 발생의 높은 분율을 기여하는 유전적 종양에서는 조기 발견에 의한 완치 가능성이 다른 암종들에 비해 높다. 유전성 종양의 특징은 일반적으로 비유전성 종양에 비해 발생 연령이 젊고 발생장기에서 다발적으로 발생한다고 알려져 있다. 또한 유전성 종양의 유전자가 규명된 경우에는 환자의 가족을 대상으로 유전자 검사를 통해 유전자를 가진 경우(carrier)보다 적극적인 경과 관찰 또는 예방적 수술 내지 장기 적출을 시행하기도 한다. 유전성 위암은 서양에서 전체 위암발생의 3~5%를 차지한다고 보고되었다. 일본인 인구집단 대상으로 위암 발생 예측 모형을 개발한 Charvat 등의 연구에서는 일본 보건소 코호트로 지역별로 보건소 또는 주요 병원서 19,028명을 1993년부터 2009년까지 추적하였고 국가 암등록 자료와 연계하여 가족력 정보는 자가 설문조사하였고 위축위염과 Helicobacter pylori 유무에 따라 임상적 분류 카테고리를 ABCD4가지로 층화하여 가족력, 생활습관과의 상호작용을 분석하였다. 연구 결과 단일 위험 요인별 상대 위험도에서 성, 연령, 헬리코박터균 보유 여부가 유의 했고 가족력은 유의하지 않았으나, 임상적으로 위험한 군에서 가족력이 있고 불건강한 생활습관을 가지고 있고 고령일 때 위험도가 유의하게 올라가는 경향을 보였다. Song등의 연구에서는 스웨덴인구집단 대상으로 가족력에 따라 위점막 이상과 위암발생의 위험을 살펴보았다. 1931년 이후로 태어난 사람 중 1979년에서 2014년 biopsy를 받은 사람들을 연구 대상으로 하였으며, 가족력은 Swedish Multi-Generation Register 자료에 등록된 정보를 이용하여 1촌 과 2촌 가족정보를 획득하였다. 전체 위암 non-cardia 위암, cardia위암가족력

을 가지고 있는 군에서 모두 유의하게 가족력이 있을때 위험도가 높았고 그중에서도 형제-자매에서 높게 나타났다. 이론적 고찰 과정에서 분류한 문헌들로 정리한 가족력 포함 위암 관련 요인은 <표2-1>과 같다.

여러 연구에서 가족력 자체가 유전적인 배경으로서 작용되는지 또는 가족 구성원들 간에 유사한 생활습관 및 식습관 등을 공유하면서 생기는 환경적인 영향 때문인지에 대하여 다양한 결과들이 보고되었다. 환경적 영향이 유전적 영향보다 크다는 주장도 있고, 같은 유전학적 배경을 지니지만 서로 다른 환경에 있는 이민 인구의 암유병률이 모국에서의 유병률과 차이를 보이기도 하였다. 이번 연구에서는 위암의 가족력에 의한 위험도를 산출하였고, 향후에 다른 요인들과의 비교를 통하여 어떤 요인이 위암 발생에 가장 많은 기여를 하는지, 어떠한 위험요인에 대한 개입이 선행되어야 하는지에 대한 연구가 필요하다.

<표 2-1> 가족력관련 위암 위험요인이 포함된 관련 연구

	Charvat et al (2016)	Goshayeshi et al. (2008)	Song et al., (2018)	Eto et al., 2006 (case control)	Chen et al. 2004 (case control)	Hong et al 2006 (case control)	Shin et al 2008 (case control)	gong et al 2014 (case control)
Age	√	√		√	√	√	√	√
gender	√	√	√	√	√	√	√	√
height								
DM								
BMI					√			
Smoking	√					√		√
Alcohol						√		√
Salted food	√							
Exercise								
Education /income		√			√		√	
diet (fruit and vegetable consumption/spicy food)							√	
Age at GC Diagnosis		√						
H. pylori Status						√	√	
Family History	√	√	√	√	√	√	√	√

제2절 간암 질환 예측

2016년도 국가암등록통계자료에 의하면 10만명 당 남성의 경우 11,774명 9.8%로 5위, 여성의 경우 3,997명(3.7%)로 6위를 차지하며 의료비 면에서도 2위를 차지할 정도로 높다.<표 2-2>

연령군별 주요 암발생률을 남녀 전체로 본 경우에도 특이 65세 이상에서 5위를 차지할 정도로 발생률이 높은 암에 속해 고령에서 더욱 집중해서 관리해야 할 암으로 판단된다.<표 2-3>

<표 2-2> 성별 주요 암 발생 현황, 2016년도 국가암등록통계자료

순위	남자					여자				
	암종	발생자수	분율	조 발생률	표준화 발생률*	암종	발생자수	분율	조 발생률	표준화 발생률*
	모든 암	120,068	100.0	470.3	307.6	모든 암	109,112	100.0	426.5	281.2
	갑상선암 제외	114,530	-	446.6	288.6	갑상선암 제외	88,599	-	346.3	211.4
1	위	20,509	17.1	80.3	50.9	유방	21,747	19.9	85.0	62.5
2	폐	17,790	14.8	68.7	42.9	갑상선	20,513	18.8	80.2	68.8
3	대장	16,672	13.9	65.3	41.6	대장	11,455	10.5	44.8	23.3
4	전립선	11,800	9.8	46.2	28.2	위	9,955	9.2	38.1	22.0
5	간	11,774	9.8	46.1	29.2	폐	7,990	7.3	31.2	15.8
6	갑상선	5,538	4.6	21.7	19.0	간	3,997	3.7	15.6	7.9
7	담낭 및 기타담도	3,490	2.9	13.7	8.4	자궁경부	3,566	3.3	13.9	10.8
8	방광	3,468	2.9	13.7	8.5	췌장	3,271	3.0	12.8	6.1
9	신장	3,410	2.8	13.4	9.3	담낭 및 기타담도	3,195	2.9	12.5	5.6
10	췌장	3,384	2.8	13.3	8.3	자궁체부	2,771	2.5	10.8	7.5

<표 2-3> 연령 군별 주요 암발생률: 남녀전체, 2016, 국가암등록통계자료

(조발생률, 명/10만 명)

순위	0-14세	15-34세	35-64세	65세 이상
	모든 암 (15.4)	모든 암 (67.1)	모든 암 (474.6)	모든 암 (1559.9)
1	백혈병 (4.8)	갑상선 (31.3)	갑상선 (80.2)	폐 (256.7)
2	뇌 및 중추신경계 (2.2)	유방 (5.7)	유방 (73.3)	대장 (228.9)
3	비호지킨 림프종 (2.1)	백혈병 (3.3)	위 (61.8)	위 (228.2)
4	신장 (0.5)	자궁경부 (2.7)	대장 (51.5)	전립선 (132.1)
5	갑상선 (0.3)	비호지킨 림프종 (2.6)	폐 (34.5)	간 (111.9)

간암은 증상을 유발하지 않는 경우가 많아 발견 시 치료하기 어려운 병기에 발견되는 경우가 많다. 초기에 발견되는 경우, 간절제술, 간이식, 고주파열치료술 등 근치적인 치료를 받을 수 있으나, 진행성 간암으로 진단을 받게 되는 경우 전신항암요법 등 비근치적 치료를 받게 되는 경우가 많아, 간암 고위험군에서 선별 검사를 통한 조기 진단이 매우 중요하다.

우리나라에서는 1990년~1998년까지 만성간질환을 가진 환자 4339명을 추적 관찰하여 나이, HBV, HCV, 간경변증, 음주력, ALT, AFP 등이 간암과 관련성이 있음을 밝혔다. 간암 발생의 위험인자에 따른 간암 발생 예측 모형 등을 개발하기도 하였다.

기존 연구로는 18세 이상의 태국과 홍콩 health care system 자료로 development set(태국 23,851명), validation set(홍콩19,321명)을 구성하여 항바이러스제 치료를 받은 B형간염 환자를 대상으로 HCC 발생 risk score인 CAMD score를 측정된 결과 40세 이하이며 간경화가 동반된 경우 10 점, 60세 이상 군에서 10점, 남성이 2점, 당뇨병자에서 1점으로 발생위험이 높게 측정되었다. <표 2-4>

<표 2-4> multi variable cox model for HCC occurrence and CAMD score

Variables	Adjusted hazard ratio	CAMD score
Cirrhosis		
No cirrhosis	Reference	0
Cirrhosis with age <40 years	18.8 (95% CI 9.2-38.7)	10
Cirrhosis with age ≥40 years	4.6 (95% CI 3.8-5.6)	6
Age (years)		
<40	Reference	0
40-49	4.5 (95% CI 2.4-8.5)	5
50-59	9.0 (95% CI 4.8-16.8)	8
≥60	15.9 (95% CI 8.5-29.7)	10
Male sex		
Female sex	Reference	0
Male sex	1.8 (95% CI 1.4-2.2)	2
Diabetes mellitus		
Not diabetic	Reference	0
Diabetic	1.3 (95% CI 1.1-1.6)	1

2000년 1월 1일~2009년 12월 31일 Liver Centre에서 간경화로 예상되는 인원 대상으로 cohort연구를 진행하였다. univariable analysis 결과 steatohepatitis 4.14, HCV 6.74, HBV 6.50로 age 60세 이상 2.94, FIB-4(fibrosis-4 score) > 5.5인 경우 4.10,로 HR가 계산되었고, multivariable analysis 결과 HCV 3.83, HBV 3.66, 60세 초과 3.77으로 Harzard ratio가 계산되었다.<표 2-5>

<표 2-5> Factors associated with HCC by cox proportional hazards regression

Covariate	Univariable HR (95% CI)	p value	Multivariable HR (95% CI)	p value
Etiology				
Autoimmune	1		1	
HCV-SVR	1.74 (0.65-4.67)	0.27	1.04 (0.38-2.83)	0.94
Other	2.46 (0.95-6.34)	0.06	1.65 (0.63-4.27)	0.31
Steatohepatitis	4.14 (1.80-9.51)	0.001	2.12 (0.91-4.92)	0.08
HCV	6.74 (3.13-14.50)	<0.001	3.83 (1.76-8.33)	0.001
HBV	6.50 (2.99-14.16)	<0.001	3.66 (1.66-8.07)	0.001
Age	1.03 (1.02-1.05)	<0.001	1.04 (1.03-1.06)	<0.001
Age category				
<45	1		1	
45-60	1.91 (1.27-2.87)	0.002	1.94 (1.28-2.93)	0.002
>60	2.94 (1.94-4.46)	<0.001	3.77 (2.45-5.80)	<0.001
Male Sex	2.84 (2.04-3.94)	<0.001	2.92 (2.07-4.12)	<0.001
Platelets ($\times 10^9/L$)	0.99 (0.99-0.99)	<0.001	0.99 (0.99-1.00)	<0.001
Platelet Category ($\times 10^9/L$)				
<80	1		1	
80-139	0.78 (0.57-1.07)	0.12	0.78 (0.57-1.07)	0.13
140-200	0.40 (0.27-0.59)	<0.001	0.40 (0.27-0.60)	<0.001
>200	0.25 (0.16-0.41)	<0.001	0.31 (0.19-0.50)	<0.001
APRI				
<1.5	1			NS
1.5-4	1.69 (1.26-2.28)	<0.001		
>4	1.57 (1.09-2.26)	0.01		
FIB-4				
0-1.5	1			NS
1.5-5.5	2.55 (1.44-4.53)	0.001		
>5.5	4.10 (2.29-7.34)	<0.001		
MELD >8	1.61 (1.23-2.11)	0.001		NS

AST, aspartate aminotransferase; APRI, AST to platelet ratio index; CI, confidence interval; FIB-4, Fibrosis-4 score; HBV, Hepatitis B virus; HCV, Hepatitis C virus; HR, Hazard Ratio; MELD, model for end-stage liver disease; SVR, sustained virologic response.

미국에서는 2002년~2011년 national liver transplantation 대기자 자료를 이용하여 34,932명에 대한 cohort 연구로 여러 변수 (age, etiology of cirrhosis, DM, sex, race, BMI, TIPS과거력, variceal 출혈력, bilirubin, INR, Cr)에 대하여 Cox proportional hazard regression로 ADDRESS-HCC모델 개발하였다. 그 결과 alcohol/metabolic liver disease의 경우 HR이 2.28, viral이 4.72로 나타났다. 그 외 당뇨 1.31, 남성 2.03, CTP(Child turcotte-pugh) 1.17, MELD(model of end-stage liver disease score) 1.02, bilirubin 1.03, INR 1.36, 크레아티닌 1.16으로 Harzard ratio를 상승시키는 것으로 나타났다 <표 2-6>

<표 2-6> Cox models evaluating risk factors for the ADREESS-HCC model

Variable	Univariate Model HR (95% CI)	P	Multivariate Model HR (95% CI)	P
Age (per y increase)	1.04 (1.03-1.04)	<.001	1.05 (1.04-1.06)	<.001
Etiology of cirrhosis^d				
Autoimmune	Reference	-	Reference	-
Alcohol/metabolic	2.28 (1.69-3.07)	<.001	1.43 (1.04-1.97)	.030
Viral	4.72 (3.55-6.26)	<.001	3.38 (2.50-4.55)	<.001
Diabetes present	1.31 (1.11-1.54)	.001	1.25 (1.06-1.48)	.008
Male sex	2.03 (1.76-2.33)	<.001	1.99 (1.71-2.30)	<.001
CTP score (per unit increase)	1.17 (1.13-1.21)	<.001	1.17 (1.13-1.21)	<.001
Nonwhite race	1.26 (1.12-1.43)	<.001	1.20 (1.06-1.37)	.005
MELD score (per unit increase)	1.02 (1.01-1.04)	.007	-	-
Body mass index (per 5-unit increase)	1.01 (0.96-1.07)	.681	-	-
Presence of TIPS	0.88 (0.69-1.12)	.315	-	-
Previous variceal hemorrhage	0.93 (0.67-1.29)	.643	-	-
Bilirubin (per unit increase)	1.03 (1.01-1.05)	<.001	-	-
INR (per unit increase)	1.36 (1.22-1.52)	<.001	-	-
Creatinine (log)	1.16 (1.00-1.33)	.040	-	-

Abbreviations: 95% CI, 95% confidence interval; ADDRESS, age, diabetes, race, etiology cirrhosis, sex, and severity; CTP, Child-Turcotte-Pugh; HCC, hepatocellular carcinoma; HR, hazards ratio; INR, international normalized ratio; MELD, model of end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt.

<표 2-7> 이론적 고찰을 통해 확인된 간암 위험요인

	Kim et al (2018, 한국)	Hou et al (2018, 태국, 중국)	Yan et al (2011, 태국)	Loannou et al (2018)	Tomoda et al (2017, 일본)	Ferrina et al (2014, 미국)	Dural et al (2018, 케냐)	Takeno et al (2012, 일본)	Wen et al (2012, 태국)	Hung et al (2015, 태국)	Li et al (2018, 태국)
population	HBV	HBV	HBV	HCV	CLD	cirrhosis	cirrhosis	general	general	general	DM
Age											
Sex											
Race											
BMI											
Smoking											
Alcohol											
Physical activity											
FHx											
Etiology	HBV		HBsAg, HBV DNA								
	HCV										
	Autoimmune										
	Other (alcoholic)					metabolic	steatohepatitis		CVD		
Esophageal varix											
Cirrhosis											
DM											
AOT											
ALT											
Platelet											
Glucose											
Albumin											
AFP											
간기능 score											
DM med											
Dyslipidemia med											
기타						INR, Cr, bilirubin					

세계적으로 간암의 위험인자에 대한 연구가 많이 이루어져 왔는데, 지역적 분포에 따라 발생률이 차이가 있어서 우리나라를 기준으로 진행한 간암 발생 위험 예측 모델이 필요하다. 동아시아 국가 내에서도 중국의 경우 간암 발생의 80%가 B형 간염 바이러스 감염과 관련성이 높으며, 일본은 간암 발생의 75%가 C형 간염 바이러스 감염과 관련이 있는 것으로 알려져 있다. 우리나라도 우리 국내 역학에 맞는 개별화된 간암의 예측 모형이 필요하다. 수검률이 높은 국민건강보험공단 건강검진 수검자의 검진결과를 활용하여 모형을 구축할 경우, 우리나라를 대표하는 빅데이터를 통하여 간암 위험인자의 확인과 함께 간암예측 모델을 개발할 수 있다.

제3절 대사증후군의 유전형과 표현형 융합분석

1988년 Reaven GM에 의해 처음 “Syndrome X”로 명명된 대사증후군(Metabolic syndrome)은 복부비만, 높은 혈압, 이상지질혈증, 높은 혈당 등이 동반되어 나타나는 대사적 이상 상태를 이야기 한다. 다시 말해 뇌심혈관질환 및 당뇨병의 위험을 높이는 체지방 증가, 혈압 상승, 혈당 상승, 혈중 지질 이상 등의 이상 상태들의 집합을 말한다.

미국의 NHANES III(National Health and Nutrition Examination Survey)에 따르면 대사증후군의 유병률이 20세 이상 성인 남자는 1988년에서 1994년 사이 25.6%에서 2007년에서 2012년 사이 33.4%로 증가하였다. 또한 성인 여자에서는 1988년에서 1994년 사이 25.0%에서 2007년에서 2012년 사이 34.9%로 증가한 것으로 나타났다.

대사 증후군의 발병 원인은 잘 알려져 있지 않다. 일반적으로 인슐린 저항성(insulin resistance)이 근본적인 원인으로 작용한다고 추정하고 있지만 이 역시 대사 증후군의 명확한 발병을 명확하게 설명하지는 못한다.

인슐린 저항성이란 인슐린이 분비됨에도 불구하고 인슐린의 작용이 감소된 상태를 말한다. 즉 근육, 지방 및 간세포가 인슐린에 잘 반응하지 않아 혈액에서 포도당을 쉽게 흡수 할 수없는 경우이다. 근육과 간 등에서 혈당을 이용하지 못해, 결과적으로 췌장은 포도당이 세포로 들어가는 것을 돕기 위해 더 많은 인슐린을 만들게 된다. 이로 인해 당뇨병 전 단계 또는 당뇨병이 유발되며 높은 인슐린에 의해 염분과 수분이 증가하여 고혈압이 생기기도 한다. 또한 증가된 인슐린은 지방이 쌓이는 것을 유도하여 비만을 촉발하고, 중성지방의 혈중 농도를 높여 이상지질혈증이 나타난다.

다시 말해 인슐린 저항성이란 혈당을 낮추는 인슐린에 대한 몸의 반응이 감소하여 근육 및 지방세포가 포도당을 잘 섭취하지 못하게 되고 이를 극복하고자 더욱 많은 인슐린이 분비되어 여러 가지 문제를 일으키는 것을 말한다

2017년 우리나라 통계청 사망원인 순위 2,3위는 심혈관 질환(10.8%), 뇌혈관질환(8.0%)로 나타났다. 이런 심뇌혈관 질환은 전세계적으로도 1,2위의 사망원인으로서 사회경제적인 측면에서도 부담을 주는 질환으로 이런 심뇌혈관 질환을 막기 위한 여러 연구들이 이루어지고 있다. 심뇌혈관 질환을 예방하게 하는 주요한 기전은 이상지질혈증으로 인해 생기는 죽상경화증이다. 이상지질혈증을 일으키는 원인은 체중증가, 스트레스 음주, 흡연, 신체활동 부족으로 알려져 있다. 이상지질혈증을 조절하기 위한 생활습관 개선의 일환으로 신체활동에 대한 연구가 이루어져 실제로 여러 신체활동과 이상지질혈

중의 상관성이 밝혀지고 있고 이는 심혈관 질환에 의한 사망률과 이환률을 낮추고, 여러 질환의 사망률과 이환률을 낮추는 것으로 알려져 왔다.

국내에서도 생활습관과 대사증후군의 발병에 관련된 여러 연구가 진행되어 왔다. 대사증후군의 위험인자로는 복부비만, 인슐린저항, 열량섭취의 과다, 부족한 신체활동이 있으며 이 외에도 환경적, 사회경제적 및 심리적 요인이 기여하는 것으로 알려져 있다. 이 중 인슐린저항은 당뇨병의 위험인자로 인슐린 저항성 정도에 따른 대사증후군의 상대적 위험도(insulin, HOMA-IR을 사분위하여 대사증후군 상대위험 비교한 연구 결과 결론적으로 인슐린 저항성 정도가 상승시 대사증후군의 위험도가 증가하는 경과를 가져왔다. 이는 대사 증후군이 다양한 대사경로와 연관되어 기타 다른 질환의 발생위험을 높인다는 증거이다. 또 지속적인 과도한 열량섭취와 신체활동 부족은 이소성 지방조직의 축적을 증가시키며 인슐린 민감도를 악화시킬 수 있다 할 수 있겠다. 국민보험공단자료를 바탕으로 2002,2003년과 2009,2010년의 생활습관의 개선여부, 대사증후군의 유병률을 비교한 결과 흡연, 음주, 운동 등 생활습관이 개선된 군에서 모두 통계적으로 유의하게 대사증후군 위험도가 줄어든 것으로 밝혀졌다.

대사증후군이라는 임상군이 정의된 이래 생활습관과의 관계를 밝히는 다양한 연구가 진행되었다. 이들 연구의 대부분이 생활습관과 대사증후군의 관계를 밝히는 연구이며 대사증후군의 표현형과 유전자형을 융합 분석한 연구는 거의 진행된 바가 없다.

대사증후군의 표현형과 유전자형을 융합 분석한 몇몇 연구들을 살펴보면 CDH13과 adiponectin level이 대사증후군에 미치는 영향을 연구한 결과 대사증후군 및 CVD(Cardiovascular disease) 발생 위험도의 유의한 연관성을 밝혀냈다. 그리고 2010년 진행된 ENPP1 K121Q 유전자 다형성과 대사증후군과의 관련성 연구에서는 유전자 다형성 유무에 따라 나누었을 때 두 집단 간에 대사증후군 구성요소가 의미있는 차이를 보였으며 K121Q 유전자 다형성은 대사증후군과 관련이 높음을 발견 하였다. 또한 체지방감소 유전자와 대사증후군이 관련이 있음을 발견한 연구와 SIRT2 유전자에서 대사증후군의 genotype과 phenotype의 연관성을 발견하였으며, 이들의 통계적인 연관성을 밝힌 한국인을 대상으로한 연구가 있다.

그동안 진행되었던 많은 연구들은 대사증후군의 표현형 정보(phenotype)를 토대로 생활습관 개선이나 관리의 필요성 등을 제안 하거나, 대사증후군과 특정 유전자의 연관성을 보는 연구가 대부분이었다. 특히 대사증후군과 유전자의 연관성을 보는 연구는 대사증후군과 관련이 있는 모든 유전자를 분석한 연구는 현재까지 진행된 연구가 없었다.

본 연구에서는 표현형 정보와 개인생활습관의 차이 또는 표현형정보와 일부 유전자의 연관성을 확인한 기존 연구들에서 보인 한계점을 보완하기 위해, 각 개인이 가지고 있는 고유의 유전자(genotype) 정보와 생활습관 데이터를 융합 분석해 보았다.

즉 대사증후군과 관련이 있는 표현형 정보와 유전형 정보를 융합 분석한 결과와 생활습관이 미치는 영향을 통합 분석하여 개인별 맞춤형 건강의학 정보를 제시하고자 하였다.

다시 말해 각 개인이 같은 표현형 정보를 가지고 있어도 유전적 차이나 생활습관에 따라 대사 결과가 다르게 나타난다고 생각하고 대사증후군의 유전형과 표현형을 융합 분석하고 생활습관이 미치는 영향까지 확인 하였다.

결과적으로 본 연구를 통해 개인의 생활습관과 유전적 요인이 대사증후군 발생과 얼마나 관계가 있는지 대사증후군의 명확한 원인을 파악하여, 개인에게 맞는 상대적인 기준을 제시하기 위해 대사증후군의 표현형과 유전자형을 융합 분석한 연구를 실시하였다.

제4절 위암과 헬리코박터 관련성

위암의 발생의 위험요인으로는 성별, 나이, 인종, 지리, 헬리코박터균(H. pylori) 감염, 위림프종, 식이요인, 흡연, 과음, 위 절제술 이력, 악성 빈혈, 메네트리에병, 혈액형, 가족력 등이 있다. 싱가포르에서는 1993-1998년 동안 등록된 45-74세의 싱가포르계 중국인 남녀 63,257명을 대상으로 여러 가지 생활 습관 요소와 위선암 발생 위험 사이의 관련성을 조사하는 연구를 수행하였다. 위암 위험요인으로 알려진 흡연, 음주, 비만, 신체 활동, 염분 섭취, 식이 습관을 사용하여 종합 생활 습관 점수를 계산하고, 위선암 위험과 관련하여 종합 생활 습관 점수와 헬리코박터 파일로리 감염상태의 연관성을 조사하였다. 헬리코박터 파일로리에 노출된 집단에서 건강한 생활 습관 요소와 위선암 위험 사이의 역상관 관계를 확인하였다. 대만에서는 국민 건강 보험 연구 데이터베이스 (National Health Insurance Research Database)를 기반으로 1997년 1월 1일부터 2004년 12월 31일 사이에 처음으로 소화성 궤양 진단을 받은 278,898명의 환자를 대상으로 연구를 진행하였다. 나이, 성별, 위궤양, 소화성 궤양 합병증, 헬리코박터 균 제균, 비스테로이드성 항염증제 사용, 감시 내시경 검사 및 동반질병과 같이 위암 위험과 관련이 있는 인자들을 사용하여 소화성 궤양 환자를 위한 맞춤형 위험 예측 모델 개발하였다.

국내 관련 연구로 2008년 1월부터 2013년 12월 동안 강남 세브란스 병원 건강 증진 센터에서 위 내시경 검사를 받은 사람 60,261명을 대상으로 cross-sectional 조사를 진행

하였다. 나이, 성별, BMI, 음주량, 식습관, 당뇨유무, 흡연, 헬리코박터균, 위축성 위염, 장상피화생을 위험 인자로 고려하여 위암과의 연관성을 확인하였다. 국립암센터 암방울 팀은 건강보험공단에서 연구목적으로 공개한 국민건강보험공단 1차 검진, 문진 자료를 기반으로 1996-1997년 수검자 대상으로 나이, 비만도, 가족력, 식사 횟수, 염분 섭취량, 식이 습관, 육류 섭취량, 음주량, 흡연량, 육체활동을 예측인자로 고려하여 위암 예측모형을 구축하여 2015년에 발표하였다. 해당 연구는 남성, 여성 모델의 좋은 설명력을 보여 주었으나, 10여년 전 자료를 사용한 연구이므로 현시점에서 적용하기에는 한계가 있었다. 우리나라는 세계적으로 높은 국가보험 가입률을 유지하고 있으며 국민건강보험의 가입자 중 국가 암 검진 대상자는 2004년 34.4%선별율, 2005년은 50%에 가까운 선별율 보여 공단 빅데이터 자료는 연구 가치를 가진다. 또한 한국인의 특성을 반영하는 대표성과 신뢰성이 높아 유용하게 활용이 가능하며, 보험의 가입자는 정기적으로 건강검진을 받아 추적 관찰 면에서도 용이한 점이 있다.

<표 2-8> 각 연구에서 사용된 헬리코박터균 관련 위암 발생 위험 인자

		Lee (대만코호트) et al, 2015	Park Et al 2016	Hong et al 2006 (case control)	Shin et al 2008 (case control)
Age					
Gender					
Height					
DM					
BMI					
Smoking					
Alcohol					
physical activity					
Education					
가족력	Family Hx of cancer				
식이	Meat consumption				
	Meal regularity				
	Meal preference				
	Salt preference				
Occupation					
Pernicious anemia					
Peptic ulcer disease					
Prior Hx of gastric surgery					
NSAID Use					
Endoscopy Surveillance					
Atrophic gastritis			AG,IM		
H.Pylori infection/eradication					

5차년도 (2018년) 위암예측모형 고도화 연구를 진행하였고, 일반검진 문진자료 뿐 아니라 암검진 문진자료와 위암검진결과 기록지를 사용하여 위암 발생에 영향을 미치는 암종별 가족력/과거력, 동반 질환 등 위험 요인들을 추가적으로 고려하여 고도화된 모형 개발하자 하였다. 연령이 높을수록, 당뇨가 있는 경우, 위암검진결과 비정상인 경우, 위암가족력이 있는 경우 흡연자일 경우 5년 위암 발생 고위험군으로 나타났다. 하지만 청구 자료의 한계로 인한 Peptic ulcer, atrophic gastritis, H.pylori 등의 파악의 어려움 및 검진 자료의 실제 임상자료의 정확성 부족으로 첨부하지 못하였다. 이에 대한 보완하여 변수 재구성 및 희소 사건을 고려하는 모형으로 추후 연구 진행이 필요하다고 판단하였고, 이를 보완 위해 병원간 데이터 공유 시스템인 CDM(common data model)을 통해 고도화된 모형 개발 및 CDM 검증을 시도하였다. 이는 좀 더 정교하고 실질적으로 위암예측이 가능하게 하여 국민 건강생활습관을 제시할 수 있는 근거가 될 것이다.

제3장

연구 내용 및 방법

제1절 연구 내용	29
제2절 연구 방법	30
제3절 각 세부 과제별 연구 방법	32

제3장

연구 내용 및 방법

제1절 연구 내용

1. 위암 질환 예측

위암 가족력 모형 연구에서는 2009-2010년 동안 국민건강보험공단 맞춤형 DB의 건강 검진결과와, 위암검진을 모두 수검 받은 대상자들을 2017년까지 추적하여 발생한 위암의 산정특례 등록자료를 연계하였고, 행정안전부와 연계하여 구축한 가족력 데이터베이스와 추적기간 동안 발생한 대상자들의 위암의 관련성을 분석하였다.

2. 간암 질환 예측

간암 질환 예측 연구에서는 일반검진 자 대상으로 간암과 나이, 성별, 수입, 비만도, 허리둘레, 혈압, 당뇨, 고혈압, 사망률, 혈당, AST, ALT, 콜레스테롤 그리고 흡연, 음주, 운동 등 생활 습관과의 관련성을 보았다. 국가 간암검진자 대상으로 진행되는 AFP 혈청 검사와 복부 초음파 결과 검사소견을 이용하여 위험인자별 콕스 회귀 모형을 적용한 간암 질환 예측을 하였다.

3. 대사증후군의 유전형과 표현형 융합분석

본 연구를 통해 개인의 생활습관과 유전적 요인이 대사증후군 발생과 관계가 있는지 파악하여, 개인에게 맞는 상대적인 기준을 제시하기 위해 대사증후군의 표현형과 유전자형을 융합 분석한 연구를 실시하였다.

4. 위암과 헬리코박터 관련성 분석을 통해 CDM모델의 정합성 검증

CDM data를 이용하여 일산병원 내에서 헬리코박터 진단받거나 약제를 쓴 군에서 대조군과 비교하여 위암의 발생 위험도를 비교하였다.

제2절 연구 방법

1. 자료 개요

(1) 국민건강보험 청구 DB

위암 및 간암 연구에서 사용한 자료는 모두 맞춤형 자료이다. 맞춤형 자료란 국민건강보험공단이 수집, 보유, 관리하는 모든 수검자의 건강보험 및 장기요양보험 자료에서 연구에 맞는 변수를 요청하여 받은 DB이다. DB 항목 중 자격 및 보험료, 사망, 명세서(20t), 진료내역 (T30), 처방교부상세 (T60), 일반검진 문진표 및 결과, 암검진 문진표상 동반질환/과거력 등을 대상으로 하였다.

<표 3-1> 맞춤형 DB 세부특성

구분	세부 특성
자격DB	<ul style="list-style-type: none"> 대상 : 건강보험가입자 및 의료급여수급권자(외국인 제외) 변수 : 성, 연령대, 지역, 가입자 구분, 소득분위 등 대상자의 사회경제적 변수 및 장애, 사망관련 변수로 구성
진료DB	<ul style="list-style-type: none"> 내용 : 대상자가 요양기관에 방문하여 진료 등을 받은 내역에 대해 요양기관으로부터 요양급여가 청구된 자료 구성 : 의과_보건의관(T1), 치과_한방(T2), 약국(T3)자료에 대한 명세서(20t), 진료내역(30t), 상병내역(40t), 처방전교부상세내역(60t)의 10개 세부DB로 구성 변수 : 명세서 공통, 진료, 상병, 처방 관련 등의 변수로 구성
건강검진DB	<ul style="list-style-type: none"> 내용 : 건강검진 주요 결과 및 문진에 의한 생활습관 및 행태관련 자료 ... 1차 일반건강검진 자료, 2008년부터 생애전환기건강진단 자료 포함 구성 : 2002~2008년, 2009~2017년 건강검진DB 별도 구성 ... 검진제도 개편(2009년)으로 주요 검진 및 문진항목 변경
요양기관DB	<ul style="list-style-type: none"> 내용 : 요양기관의 종별, 설립구분별, 지역(시도)별 현황 및 시설, 장비, 인력관련 자료

(2) 암검진 DB(간암검진)

암검진 서식지는 수검자가 암검진 받기 전에 작성하는 ‘암검진 문진표’, 암검진을 실시한 의료기관에서 암검진 결과를 기록하고 보관하는 서식인 ‘암검진 결과 기록지’. 의료기관에서 수검자에게 암검진 결과를 통보하는 ‘암검진 결과 통보서’를 의미한다.

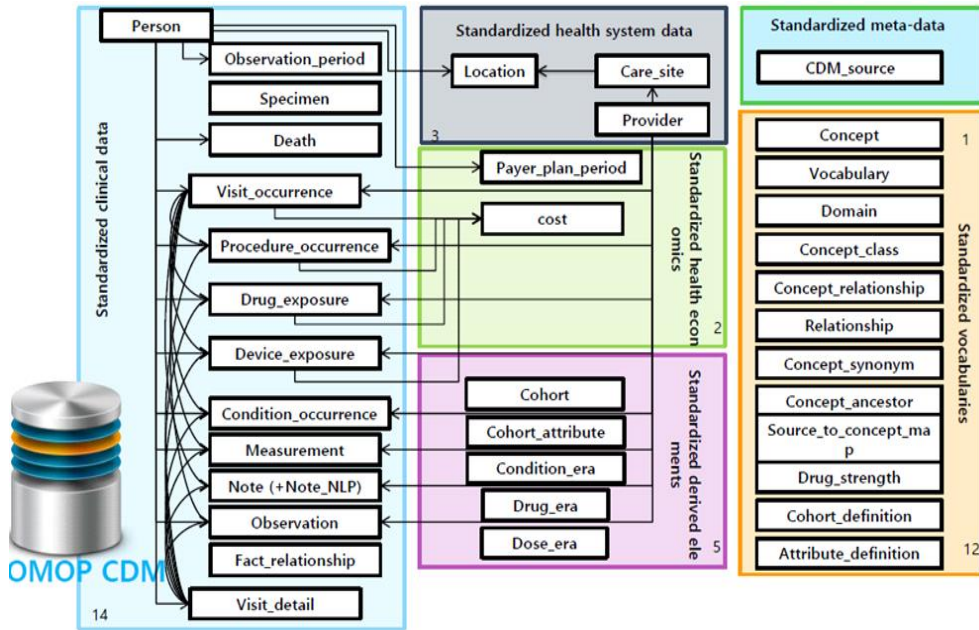
기존의 암검진 서식지는 ‘일반건강검진 및 암검진 실시기준(보건복지가족부 고시 제 2007-129호)’ 제 13조에 의한 건강검진운영세칙에 규정되었으나 2014년 개정 서식지는 암관리법 제 11조 및 같은 법 시행규칙 제 4조의 규정에 따라 ‘암검진실시기준(보건복지부 고시 제 2014-202호)’에 따른다. 개정시기에 따라 검진기록지가 변경되므로 이에 따라 정합성 맞추는 작업이 추후 필요하겠다.

(3) 가족관계 DB

가족(친족)력 데이터베이스는 건강보험공단에서 건강보험 자격관리를 위해 세대주와 보유하고 있는 세대원 관계 자료를 행정안전부 데이터와 연계한 데이터베이스이다. 세대주와 세대원의 1대1 관계 구축 후, 이 관계의 역 혹은 확장 관계 반복 생산을 통해 한 개인의 누적적인 친족 관계가 구성된다. 현재 한국 법적제도상 기술되어 있는 친족 및 특수관계인의 정의를 최대한 반영하면서, 구축 시간 및 관계 확장에 서의 오차 등을 고려해 4촌 이내 친족과 반영하면서, 구축시간 확장에 따른 오차 등을 고려하여 4촌 이내 친족과 일부 기타 친족을 범위로 구축된 DB를 이용하였다.

(4) CDM DB

OHDSI는 대규모 분석을 통해 보건의료 관찰데이터의 가치를 이끌어내기 위해 노력하는 여러 이해당사자 및 학제간의 협업 프로그램으로 모든 솔루션을 오픈소스로 제공하고 있다. 중앙조정본부(central coordinating center)는 미국 컬럼비아 대학에 있으며 2018년 현재 OHDSI 연구 네트워크에 전 세계 6억 명이 넘는 환자 자료가 등록되어 있다. 적재 자료(개원~)는 OMOP CDM(환자 중심으로 구축된 표준화된 임상 데이터모델), 적재 주기는 1회/1년(매 적재시 개원부터 종료일까지 전수 재추출), CDM database는 병원 자료를 기반으로 하는 테이블과 이를 이용해 2차적으로 구축된 standardized derived tables, 그리고 표준용어 테이블로 구성되어 있다. ATLAS를 이용하여 CDM data 추출하고 통계 분석이 가능한 DB이다. 국민건강보험 일산병원은 개원시기부터 축적된 환자의 임상 data를 OMOP CDM으로 변환하였고 OHDSI 네트워크에 가입한 상태로 이 CDM DB를 이용하여 자료를 분석하였다.



[그림 2-3] CDM database 구성도

제3절 각 세부 과제별 연구 방법

1. 위암 질환 예측 연구

(1) 자료 개요

이 연구에서 사용한 자료는 맞춤형 자료이다. 맞춤형 자료란 국민건강보험공단이 수집, 보유, 관리하는 모든 수검자의 건강보험 및 장기요양보험 자료에서 연구에 맞는 변수를 요청하여 받은 DB이다.

<표 3-1> 맞춤형 DB 세부특성

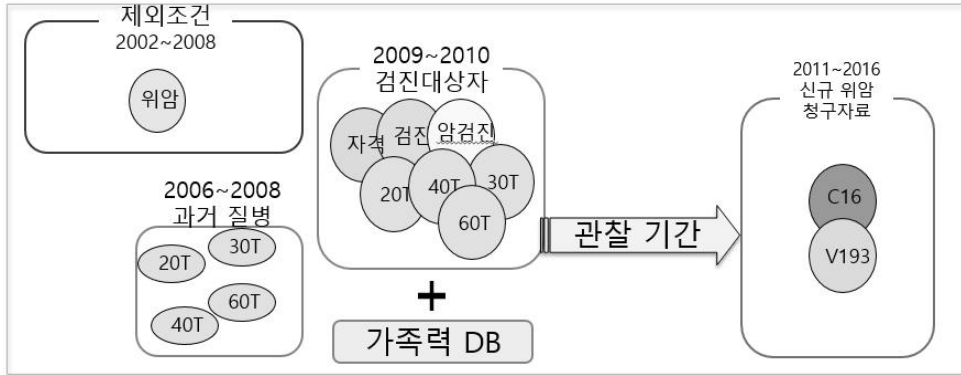
구분	세부 특성
자격DB	<ul style="list-style-type: none"> 대상 : 건강보험가입자 및 의료급여수급권자(외국인 제외) 변수 : 성, 연령대, 지역, 가입자 구분, 소득분위 등 대상자의 사회경제적 변수 및 장애, 사망관련 변수로 구성
진료DB	<ul style="list-style-type: none"> 내용 : 대상자가 요양기관에 방문하여 진료 등을 받은 내역에 대해 요양기관으로부터 요양급여가 청구된 자료 구성 : 의과.보건기관(T1), 치과.한방(T2), 약국(T3)자료에 대한 명세서(20T), 진료내역(30T), 상병내역(40T), 처방전교부상세내역(60T)의 10개 세부DB로 구성 변수 : 명세서 공통, 진료, 상병, 처방 관련 등의 변수로 구성
건강검진DB	<ul style="list-style-type: none"> 내용 : 건강검진 주요 결과 및 문진에 의한 생활습관 및 행태관련 자료 ... 1차 일반건강검진 자료, 2008년부터 생애전환기건강진단 자료 포함 구성 : 2002~2008년, 2009~2017년 건강검진DB 별도 구성 ... 검진제도 개편(2009년)으로 주요 검진 및 문진항목 변경
요양기관DB	<ul style="list-style-type: none"> 내용 : 요양기관의 종별, 설립구분별, 지역(시도)별 현황 및 시설, 장비, 인력관련 자료

(2) 연구 설계 및 대상자

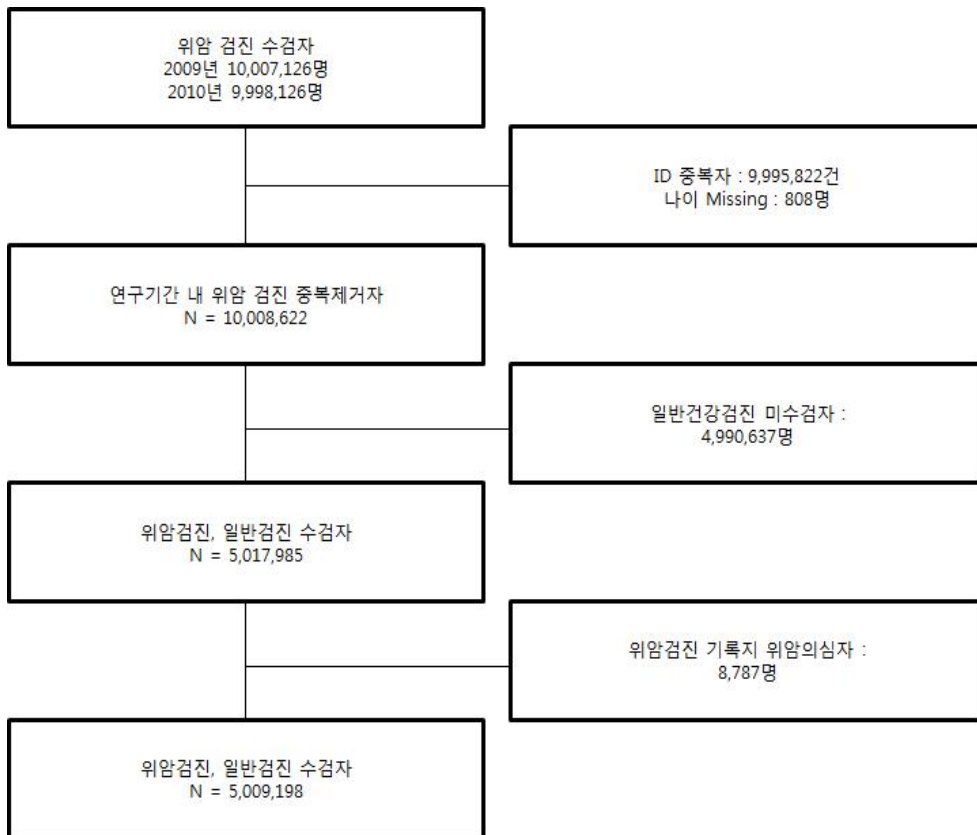
2019년 위암 연구에서는 2018년 연구 모형을 확장하여 가족력 (1촌, 2촌) 정보를 고려하여 예측 모형을 고도화했다. 위암의 경우 40세 이상의 2009년부터 2010년까지 위암과 전혀 관련이 없는 사람들을 선택하기 위해 2002년부터 2008년까지 위암 관련 된 코드가 있던 모든 대상자들을 제외하였고 자격데이터 상의 결측 또는 오류가 있는 대상자들도 제외하였다. 추가적으로 위암검진을 받지마자 위암기록지에 위암의 심으로 판명된 대상자들도 분석에서 제외하였다. 연구에 사용한 Dataset 은 전년도 연구에서 활용한 연구자료로써 자격 및 보험료, 사망, 명세서(20T), 진료내역(30T), 처방교부상세(60T), 일반검진 문진표 및 결과, 암검진 문진표상 동반질환/과거력/가족력, 위암검진 결과기록지를 구축하였으며, 최종적으로 선정된 5,009,198명의 가족력 데이터를 행정안전부 가계자료와 연계하였다.

그림 3-1은 위암연구 대상자 선별 과정의 도식화이다. 최초 위암 검진 수검자는 2009과 2010년에 각각 10,007,126명, 9,998,126명이었다. 이 중 ID 중복자 9,995,822명, 연령 data missing 대상자 808명, 연구기간 내 위암 검진 중복 대상자를 제거하여 10,008,622명 선별하고 추가적으로 일반건강검진 미수검자 4,990,637명, 위 건강검

진 기록지상 위암 의심자 8,787명을 제외하여 최종적으로 5,009,19명의 대상자가 연구에 포함되었다.



[그림 3-1] 위암예측모형 연구 설계 도식도



[그림 3-2] 위암예측모형 연구 대상자 선정 과정

(3) 가족력 데이터베이스 구축

가족(친족)력 데이터베이스는 한국의 제도에 기술되고 있는 친족 및 특수관계인의 정의를 사용하고, 구축 시간 및 관계 확장에서의 오차 등을 고려해 4촌 이내 친족으로 반영하면서, 구축시간 확장에 따른 오차 등을 고려하여 4촌 이내 친족과 일부 기타 친족을 범위로 구축되었다.

본 연구에서는 대상자 수를 고려하여 1촌, 2촌 가족력까지 연구범위에 포함시켰다. 가족력 분석에 포함된 연구대상자의 출생시점별 가족력 데이터베이스와의 연계율에서 전반적으로 최근 출생연도 대상자에 비해 과거 출생 대상자들의 가족력 연계율이 낮게 나타났다. 그중에서도 조부모 및 외조부의 경우 1980년대생과의 가족력부터 급격하게 연계율이 낮아졌으며, 1970년대, 1960년대생으로 갈수록 조부모와의 연계율이 급격하게 떨어졌다. 이는 1970, 1960년대생 대상자의 조부모의 경우 이미 돌아가셨을 확률이 상대적으로 높고 당시 기록이 보관되지 않았을 가능성이 높다. 이는 가족력데이터베이스 이용 연구의 큰 한계로 노인성 질환의 가족력을 분석하는데 있어서 분석의 문제점이 된다. 가족력데이터베이스 연계에 있어 또 하나의 특징은 부모의 연계율에 비해 형제자매의 연계율이 전반적으로 떨어지는 경향을 보이며, 이는 외동인 대상자의 가족력이 정확하게 반영하지 않았기 때문으로 해석할 수 있다 (표 3-3).

<표 3-2> 가족력 분류 체계 비교

대분류	소분류	가족 (친족) DB	민법의 친족	국세기본법의 특수 관계	상속세 및 증여세법의 특수 관계
배우자	배우자	포함	포함	포함	포함
혈족	혈족	4촌 이내	8촌 이내	6촌 이내	6촌 이내
인척	혈족의 배우자	4촌 이내	8촌 이내	6촌 이내	6촌 이내
	배우자의 혈족	4촌 이내	4촌 이내	4촌 이내	4촌 이내
	배우자의 혈족의 배우자	4촌 이내	4촌 이내	4촌 이내	4촌 이내
	혈족의 배우자의 혈족	자녀·형제 자매의 배우자의 부모·형제 자매	1990년 제외	미포함	직계비속의 배우자의 2촌 이내 혈족과 그 배우자

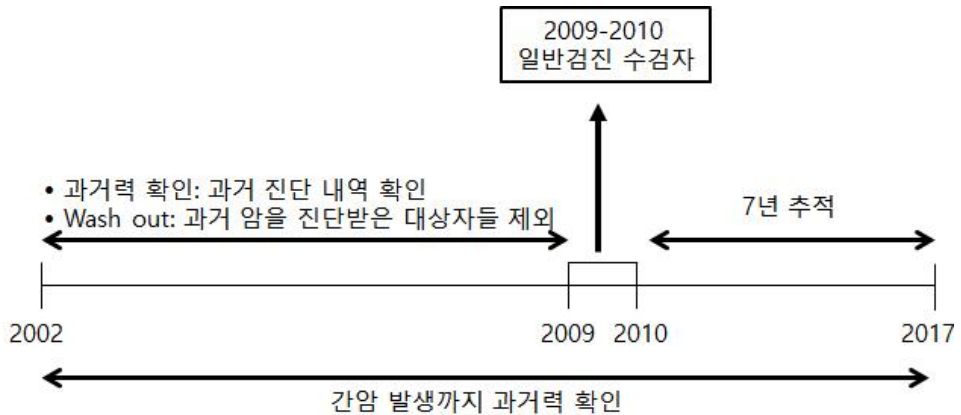
<표 3-3> 출생시점별 가족력 데이터베이스의 연계율(공단빅데이터실 추출 자료)

	친족전체	부	모	형제자매	조부모	외조부모	배우자
2010년대생	99.7%	96%	96%	72%	97%	93%	.
2000년대생	99.9%	98%	98%	86%	92%	86%	0%
1990년대생	99.9%	98%	98%	91%	71%	53%	3%
1980년대생	99.8%	93%	96%	91%	48%	19%	49%
1970년대생	99.5%	74%	84%	85%	28%	7%	80%
1960년대생	98.5%	31%	45%	37%	9%	2%	90%
1950년대 이전생	97.9%	7%	17%	7%	1%	0%	86%

2. 간암 질환 예측 연구

(1) 일반검진자 대상 간암 질환 예측을 위한 데이터셋 구축

2009년~2010년까지 일반검진 대상자 전수를 대상으로 2002~2010년 과거 간암을 진단받은 대상자는 제외 하여 총 9,121,523 명을 대상으로 하여, 2017년까지 7년간 추적하여 간암 발생 여부 조사하였다.

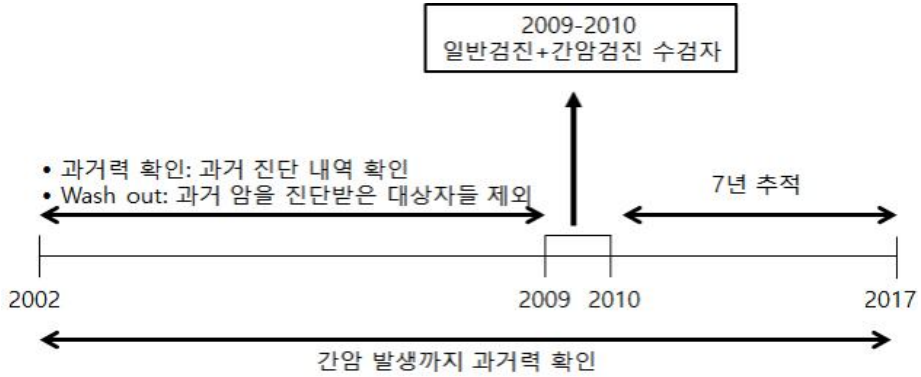


[그림 3-3] 일반검진 연구 대상자 선정

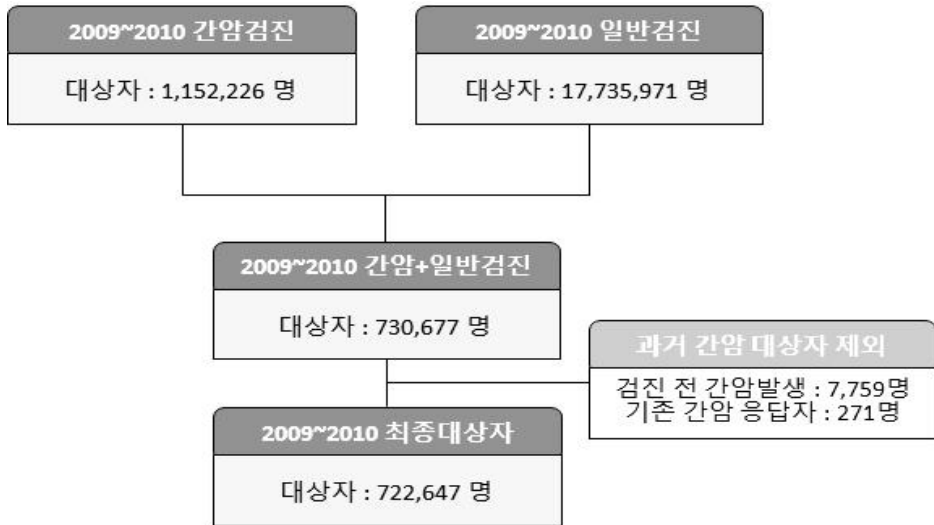
(2) 간암검진자 대상 간암 질환 예측을 위한 데이터셋 구축

2009~2010년 간암검진자 전수는 1,152,226명, 2009~2010년 일반검진 전수는 17,735,971명이었고 두 대상자 모두 존재하는 교집합 부분을 730,677명 있었다. 과거 간암 검진 대상자는 제외하였고 이 중 검진 전 간암발생 7,759명, 기존 간암 응답자

271명을 제외하였고 최종적으로 722,647명으로 선정하였다. 이 군에 있어서 2017년 까지 7년간 추적하여 간암 발생 여부 조사하였다.



[그림 3-4] 간암검진 연구 대상자 추출



[그림 3-5] 간암 건진 대상자 선정

(3) 예측 모델 개발에 포함될 변수 선정

해외 연구 사례에 근거한 위험 요소들은 원인의 경우 일반 특성인 나이, 성별, 체질량 지수, etiology인 HBV, HCV, CLD, autoimmune hepatitis, 알콜성 간염, 기타 대사질환에 의한 간염, 동반질환인 간경화, 식도정맥류, 당뇨, 고지혈증 진단

그 외 혈청 검사인 AST, ALT, 혈소관, 알부민 등이 존재하였다<표 2-7>.

① 일반검진 자 대상 위험인자 선정

나이, 성별, 수입, 체질량지수, 허리둘레, 혈압, 당뇨, 고혈압, 사망률, 혈당, AST, ALT, 콜레스테롤 그리고 흡연, 음주, 운동 등 생활 습관과의 관련성을 보았다.

② 간암 검진자를 대상 위험인자 선정

나이, 성별, 체질량지수, 혈압, 혈당, 체질량 지수, 총콜레스테롤, 중성지방, 고밀도 콜레스테롤, 저밀도 콜레스테롤, 크레아티닌, 콩팥청소율, 생활습관(음주, 흡연, 운동), AST, ALT 및 HBV, HCV, AFP, 간초음파 관찰 소견과 간암 발생과 관련성을 보았다.

3. 대사증후군의 유전형과 표현형 융합분석

(1) 연구 대상자 및 위험요인 측정

연구 대상자는 검진 및 문진 내역이 있는 검진 대상자 중 유전자 분석 관련 연구 참여에 동의한 249명을 대상으로 선정하였다. 국민 일반 건강 검진 문진표를 통해 고혈압, 당뇨병, 이상지질혈증 약물치료여부를 확인하였다. 건강검진의 신체 측정 자료를 통해 체질량지수, 허리둘레, 수축기 혈압, 이완기 혈압의 측정값을 이용하였으며, 건강검진의 혈액 검사 자료를 통해 공복혈당, 중성지방, 고밀도콜레스테롤의 결과치를 이용하였고 생활습관에 대해 흡연력, 음주력, 신체활동여부 및 강도에 대하여 확인하였다.

대사증후군 위험률 관련 요인으로 나이, 체질량 지수, 수축기혈압, 이완기혈압, 공복혈당, 중성지방, 허리둘레, 고밀도 콜레스테롤, 흡연력, 음주력, 운동량을 선정하였다. 나이는 40세 미만과, 40대, 50대, 60대, 70세 이상으로 분류하였으며 흡연력은 설문에서 비흡연자 (non-smoker), 과거흡연자 (ex-smoker)와 현재흡연자(current smoker)로 분류하여 정의하였다. 흡연력 유무의 경우, 과거흡연자는 제외하고 비흡연자와 현재흡연자를 사용하였다. 음주는 설문에서 음주여부를 기준으로 분류하였다. 운동량은 설문지상 1주일간 총 활동량(격렬한 신체활동 주당 횟수 × 8.0MET × 60분 + 중등강도 신체활동 주당 횟수 × 4.0MET × 60분 + 걷기 주당 횟수 × 3.3MET × 30분)으로 MET (Metabolicsquivalents)18) × min을 계산하였다. 이를 WHO IPAQ 지침서에 근거하여 저활동군, 중등도활동군, 고활동군으로 구분하였다. 운동 유무의 경우, 저활동군은 활동군(무), 중등도활동군과 고활동군을 합쳐 활동군(유)로 분류하였다.

(2) 대사증후군의 정의

대사증후군의 정의는 미국심장협회(AHA)와 국가심폐혈액연구소(NHLBI)에서 2005년에 종적 특성을 고려하여 발표한 modified ATP III criteria를 기준으로 하였다.

<표 3-4> modified ATP III criteria

Elevated waist circumference*	≥ 90 cm in men ≥ 85 cm in women
Elevated triglycerides	≥ 150 mg/dL (1.7 mmol/L) or On drug treatment for elevated triglycerides†
Reduced HDL-C	≤ 40 mg/dL (1.03 mmol/L) in men ≤ 50 mg/dL (1.3 mmol/L) in women
Elevated blood pressure	≥ 130 mm Hg systolic blood pressure / ≥ 85 mm Hg diastolic blood pressure or On antihypertensive drug treatment in a patient with a history
Elevated fasting glucose	≥ 100 mg/dL or On drug treatment for elevated glucose

<표 3-4>에 있는 modified ATP III criteria 위험요인 5개를 기준으로 0개 보유한 타입은 '정상군', 1-2개 보유한 타입은 '위험군', 3개 이상 보유한 타입은 '대사증후군'으로 분류하였다.

(3) 대사증후군 유전적 위험도 분류

국가건강검진 실시하는 대상자 중 본 연구 목적의 유전자 연구 참여에 동의한 자로부터 채취한 혈액에서 수집된 인체유래물로부터 DNA를 추출하여 약 70만개의 SNP marker를 분석할 수 있는 SNP chip을 이용한 DNA microarray방법을 사용하여 유전체를 분석하였다. 분석된 유전자를 대상으로 나이, 과거력, 생활습관 항목(운동, 흡연, 음주), 대사증후군 유무, 대사증후군 내 인자와의 위험성을 파악하고자 하였다.

<표 3-5> 대사증후군 유전적 위험도 분석을 위한 유전자 항목과 종류

항목	유전자	rs number
비만	<i>FTO</i>	rs9939609
	<i>MC4R</i>	rs17782313
	<i>GNPDA2</i>	rs10938397
	<i>NEGR1</i>	rs3101336
	<i>FAIM2</i>	rs7138803
혈압	<i>CSK</i>	rs1378942
	<i>ATP2B1</i>	rs17249754
	<i>CYP17A1</i>	rs1004467
	<i>GUCY1A1</i>	rs13139571
	<i>NPR3-C5orf23</i>	rs1173771
혈당	<i>CDKN2A-CDKN2B</i>	rs10811661
	<i>GCKR</i>	rs780094
	<i>MTNR1B</i>	rs10830963
	<i>DGKB-TMEM195</i>	rs2191349
	<i>SLC30A8</i>	rs13266634
지질	<i>HMGCR</i>	rs3846662
	<i>ABCA1</i>	rs2230806
	<i>LPL</i>	rs325
	<i>APOA5</i>	rs651821
	<i>MLXIPL</i>	rs3812316
체지방 분해율	<i>ADIPOQ</i>	rs822396
	<i>CDH13</i>	rs4783244

<표 3-6> 맞춤 가이드(영양과 운동)을 위한 유전자 분석 항목과 종류

항목	유전자	rs number
기초대사량	<i>LEPR</i>	rs1805094
근력	<i>ACTN3</i>	rs1815739
근지구력	<i>PPARD</i>	rs2016520
	<i>NFIA-AS2</i>	rs1572312
심폐지구력	<i>PGC1A</i>	rs8192678
	<i>CRP</i>	rs1205

기존에 분석 항목에 대해 유의한 상관관계가 보고된 논문을 바탕으로 각 항목별 유전적 위험도를 wild type인 경우 낮음, hetero type인 경우 보통, variant homo인 경우에는 높음으로 표시하였다. 각 항목별 유전자의 위험도가 낮음의 경우 1점, 보통의 경우 2점, 높음의 경우 3점을 계산하여 비만유전자 위험도와 혈압유전자 위험도, 혈당유전자 위험도, 지질유전자 위험도를 합하여 대사증후군 유전적 위험도를 계산하였다. 대사증후군 유전적 위험도가 4-6점인 경우 유전적 위험도 ‘낮음’ 타입, 7-9점인 경우 유전적 위험도 ‘중간’ 타입, 10-12점인 경우 유전적 위험도 ‘높음’ 타입으로 분류하였다. 또한 유전적 위험도를 기준으로 유전적 위험도 타입을 재조정하는 추가 분석을 하였다. 대사증후군 유전적 위험도가 4-5점인 경우 재조정된 유전적 위험도 ‘낮음’ 타입, 6-7점인 경우 재조정된 유전적 위험도 ‘중간’ 타입, 8-12점인 경우 재조정된 유전적 위험도 ‘높음’ 타입으로 분류하였다.

(4) 대사증후군 나이 분류

대사증후군 나이에서 주민등록나이를 뺀 값을 기준으로 대사증후군 나이 타입을 나타내었다. -5세 이하인 경우 대사증후군 나이 ‘좋음’ 타입, -5세에서 +5세 사이는 대사증후군 나이 ‘평균’ 타입, +5세 이상인 경우 대사증후군 나이 ‘나쁨’ 타입으로 분류하였다.

(5) 통계적 분석 방법

연구대상자 특성 파악의 경우 대사증후군과 비대사증후군을 나누어 연속 변수는 평균과 표준 편차, 명목 변수는 빈도와 백분율(%)을 이용하여 정리하였다. 유전적 위험도와 대사증후군 융합분석은 연구대상자 빈도와 row 기준, column 기준으로 백분율을 비교하였고 타입 별 유전적 위험도 수치를 비교하여 ANOVA를 시행하였다. 유전적 위험도 타입 별 생활습관에 따른 대사증후군 발생률은 생활습관(흡연, 음주, 운동) 유무에 따라 빈도와 백분율을 이용하였으며 카이제곱 독립 검정과 Fisher's exact 검정을 시행하였다. 이상의 분석은 파이썬(Python) 3.7.4 프로그램을 사용하여 수행하였으며, P값이 0.05 미만일 경우 통계적으로 의미 있는 결과로 판단하였다.

4. 위암과 헬리코박터 관련성 분석을 통해 CDM모델의 적합성 검증

(1) 연구 대상자

2010년 이전에 위내시경 실시하여 헬리코박터균 양성, 위축성 위염, 위궤양, 십이

지장궤양 있는 환자 중에서 2011년 이후 위암 확인된 군 선정하였다. 질병 분류 코드는 helicobacter양성(B980),atrophic gastritis(K294), pepticulcer(K27), astriculcer(K25), duodenal ulcer(K26)로 하였다.

(2) 통계 분석 방법

CDM을 사용하여 initial event cohort 정의하여 atals로 분석 진행하였다

제4장

분석 결과

제1절 위암 질환 예측 연구	45
제2절 간암 예측 모델 개발 연구	54
제3절 대사증후군의 유전형과 표현형 융합분석	58
제4절 위암과 헬리코박터 관련성 분석을 통해 CDM모델의 적합성 검정	73

제4장 분석결과

제1절 위암 예측 모형 연구

1. 연구 대상의 특성

연구 대상의 특성은 이전연도 연구를 참고하여 성별, 나이, 체질량 지수, 수입, 동반질환 (당뇨, 고혈압, peptic ulcer, atrophic gastritis), 흡연, 음주, 운동 등을 모형에 포함하였다. 추적기간 6년간 위암 발생을 확인한 결과 전체 5,017,985명 중 41,821명(0.8%), 남성 2,178,293명중 28,235명 (1.3%) 에서 위암 발생이 확인되었으며, 여성 2,830,903명 중 13,586명 (0.5%) 에서 위암 발생이 확인되었다.

<표 4-1> 위암 예측 모형 분석 대상자의 특성(2009-2010)

General Characteristics	Overall (n=5,009,196)		Male (n=2,178,293)		Female (n=2,830,903)	
	N	%	N	%	N	%
Gastric Cancer						
No	4,967,375	99.2	2,150,058	98.7	2,817,317	99.5
Yes	41,821	0.8	28,235	1.3	13,586	0.5
Sex						
Male	2,178,283	43.5				
Female	2,830,903	56.5				
Age						
40-49	1,646,936	32.9	715,975	32.9	930,961	32.9
50-59	1,558,208	31.1	663,660	30.5	894,548	31.6
60-69	1,150,836	23.0	523,342	24.0	627,494	22.2
70-80	573,221	11.4	243,389	11.2	329,832	11.7
≥80	79,995	1.6	31,927	1.5	48,068	1.7

General Characteristics	Overall		Male		Female	
	(n=5,009,196)		(n=2,178,293)		(n=2,830,903)	
	N	%	N	%	N	%
BMI						
<18.5	102,492	2.1	42,253	1.9	60,239	2.1
18.5-22.9	1,690,188	33.7	653,887	30.0	1,036,301	36.6
23.0-24.9	1,256,862	25.1	584,111	26.8	672,251	23.8
25.0-29.9	1,505,812	30.1	728,597	33.4	777,215	27.5
≥30	162,436	3.2	56,256	2.6	106,180	3.8
Missing	291,406	5.8	113,189	5.2	178,217	6.3
Income (%)						
의료급여	974,348	19.5	436,608	20.0	537,740	19.0
20%	415,823	8.3	140,104	6.4	275,719	9.7
40%	737,769	14.7	299,303	13.7	438,466	15.5
60%	759,142	15.2	332,978	15.3	426,164	15.1
80%	908,760	18.1	410,503	18.8	498,257	17.6
100%	1,213,354	24.2	558,797	25.7	654,557	23.1
동반 질환						
당뇨 (Diabetes)						
No	3,720,969	74.3	1,591,625	73.1	2,129,344	75.2
Yes	1,228,277	25.7	586,668	26.9	701,559	24.8
고혈압 (Hypertension)						
No	3,043,711	60.8	1,302,298	59.8	1,741,413	61.5
Yes	1,965,485	39.2	875,995	40.2	1,089,490	38.5
소화성 위궤양 (Peptic ulcer)						
No	4,523,306	90.3	1,988,219	91.3	2,535,087	89.6
Yes	485,890	9.7	190,074	8.7	295,816	10.5
위축성 위염 (Atrophic Gastritis)						
No	4,729,134	94.4	2,074,473	95.2	2,654,661	93.8
Yes	280,062	5.6	103,820	4.8	176,242	6.2
헬리코박터 파일로리 (Helicobacter pylori)						
No	5,006,999	99.6	2,177,029	99.94	2,829,970	99.97
Yes	2,197	0.04	1,264	0.06	933	0.03

2. 위암검진 문진표상의 가족력

5차년도 (2018년) 연구결과 중 위암발생자 41,821명의 가족력은 위암 116,168명 (2.32%), 없음 90,795명(1.81%), 위암관련암 82,213명(1.64%) 기타암 140,267명(2.8%), 결측치 4,579,753명(91.43%)이었다. 남성의 가족력은 위암 47,603명(2.19%), 위암관련암 27,691명(1.27%), 기타암 54,744명(2.51%), 결측치는 91.97% 였고, 여성의 가족력은 위암 68,565명(2.42%), 위암관련암 54,522명(1.93%), 기타암 85,523명(3.02%), 결측치 91%였다. 위암발생자들의 가족력 조사결과 암의 가족력이 없는 경우 보다 위암의 가족력이 있는 경우가 더 많았다

<표4-2> 위암검진 수검자의 가족력

General Characteristics	Overall		Male		Female	
	(n=5,017,985)		(n=2,178,293)		(n=2,830,293)	
	N	%	N	%	N	%
암 검진문진표 (가족력)						
없음	90,795	1.81	44,838	2.06	45,957	1.62
위암	116,168	2.32	47,603	2.19	68,565	2.42
위암관련암	82,213	1.64	27,691	1.27	54,522	1.93
기타암	140,267	2.8	54,744	2.51	85,523	3.02
missing	4,579,753	91.43	2,003,417	91.97	2,576,336	91.00

<표4-3> 위암발생자의 가족력

General Characteristics	Overall		Male		Female	
	(n=41,821)		(n=28,235)		(n=13,686)	
	N	%	N	%	N	%
암 검진문진표 (가족력)						
없음	810	1.94	564	2	246	1.81
위암	1,221	2.92	752	2.66	469	3.45
위암관련암	544	1.3	320	1.14	224	1.65
기타암	967	2.31	619	2.19	348	2.56
missing	38,279	91.53	25980	92.01	12,299	90.53

3. 연구 대상의 본인 위암 여부와 가족위암 발생의 가족관계별 분포

표 4-4는 위암, 위암 관련 암, 기타 암 가족력 유무를 부모, 1촌·2촌로 층화하여 조합 가능한 11가지 경우의 수와 본인 위암 여부와의 관련성을 기술하였다. 부모의 위암, 위암 관련암, 기타 암과 본인 위암여부와의 관련성은 없거나 역 관련성을 보였고 자녀 및 형제가 모두 가족력이 있는 경우 본인 위암 발생의 분율이 높게 나타났다.

<표 4-4> 본인 위암 여부와 가족 위암 발생의 가족관계별, 암유형별 분포(2009-2010)

N (%)	본인위암 여부 (5,009,196)	
	본인 위암 없음 (4,967,375)	본인 위암 있음 (41,821)
가족 관계 (부모)	(있음 : 위암*, 위암 관련 암**, 기타 암*** 중 하나라도 있는 경우)	
0	부모 모두 없음	4691379 (99.15) 40205 (0.85)
1	부있고 모 없음	147,284 (99.51) 702 (0.49)
2	모 있고 부 없음	116,501 (99.31) 815 (0.69)
3	부모 둘 다 있음	16,711 (99.41) 99 (0.59)
4	부 위암 있음 + 모 없음	30,982 (99.50) 156 (0.50)
5	부 위암관련암/기타암 있음 + 모 없음	111,802 (99.51) 546 (0.49)
6	모 위암 있음 + 부 없음	16,457 (99.03) 162 (0.97)
7	모 위암관련암/기타암 있음 + 부 없음	100,044 (99.35) 653 (0.65)
8	부 위암 있고 모 위암관련암/기타암 있음	3,002 (99.47) 16 (0.53)
9	부위암관련암/기타암 있음 + 모 위암 있음	1,871 (99.10) 17 (0.90)
10	부모 모두 위암관련암/기타암 있음	11,279 (99.44) 63 (0.56)
11	부모 모두 위암 있음	559 (99.47) 3 (0.53)
1촌 2촌	(있음 : 위암*, 위암 관련 암**, 기타 암*** 중 하나라도 있는 경우)	
0	1촌2촌 모두 없음	3,475,764 (99.2) 28,086 (0.8)
1	1촌이었고 2촌없음	469,279 (99.3) 3,208 (0.7)
2	2촌이었고 1촌 없음	835,069 (98.9) 8,981 (1.1)
3	1촌2촌둘 다 있음	163,813 (99.1) 1,546 (0.9)
4	1촌 위암 있음 + 2촌없음	90,334 (99.3) 600 (0.7)
5	1촌 위암관련암/기타암 있음 + 2촌없음	403,295 (9.4) 2,608 (0.6)
6	2촌위암 있음 + 1촌 없음	168,063 (98.9) 1,937 (1.1)
7	2촌위암관련암/기타암 있음 + 1촌 없음	667,006 (99.0) 7,044 (1.0)
8	1촌 위암 있고 2촌위암관련암/기타암 있음	22,635 (99.1) 216 (0.9)
9	1촌 위암관련암/기타암 있음 + 2촌위암 있음	26,458 (98.9) 282 (1.1)
10	1촌2촌 모두 위암관련암/기타암 있음	108,918 (99.1) 977 (0.9)
11	1촌2촌 모두 위암 있음	5,802 (98.8) 71 (1.2)

*위암: C16 + V193

**위암 관련 암: 유방암 (C50+V196), 대장암 (C18~C20+V193), 자궁경부암 (C53+V193), 간암 (C22+V193), 폐암 (C34+V193)

***기타 암: 나머지 암들 (C+V193)

<표 4-4-1> 본인 위암 여부와 가족 위암 발생의 가족관계별, 암유형별 분포(2009-2010)

N (%)	본인위암 여부 (5,009,196)	
	본인 위암 없음 (4,967,375)	본인 위암 있음 (41,821)
부모 위암 없음	4,691,379 (99.15)	40,205 (0.85)
부모 위암 있음	275,996 (99.42)	1,616 (0.58)

전체 대상자의 본인 위암 여부와 가족위암 발생의 가족관계별 분포는 <표4-5>와 같았다. 위암 가족력 정보와 본인위암 여부를 교차 분석한 기술통계 상으로는 아버지와 어머니, 1촌의 위암 가족력, 위암 관련 암 가족력, 기타 암 가족력에서 관련성과 본인 위암과의 관련성이 없는 것으로 나타났다.

<표 4-5> 본인 위암 여부와 가족 위암 발생의 가족관계별 분포(2009-2010)

가족관계		본인위암 여부 (5,009,196)		P-value
		없음 (4,697,375)	있음 (41,821)	
아버지	없음	4,807,880 (99.2)	41,020 (0.8)	<.0001
	위암	34,543 (99.5)	175 (0.5)	
	위암 관련 암	65,868 (99.5)	334 (0.5)	
	기타 암	59,084 (99.5)	292 (0.5)	
어머니	없음	4,834,163 (99.2)	40,907 (0.8)	<.0001
	위암	18,887 (99.0)	182 (1.0)	
	위암 관련 암	49,975 (99.4)	317 (0.6)	
	기타 암	64,360 (99.4)	415 (0.6)	
1촌	없음	4,310,833 (99.1)	37,067 (0.9)	<.0001
	위암	118,771 (99.3)	887 (0.7)	
	위암 관련 암	253,055 (99.3)	1,747 (0.7)	
	기타 암	284,716 (99.3)	2,120 (0.7)	
2촌	없음	3,968,493 (99.2)	31,294 (0.8)	<.0001
	위암	200,323 (98.9)	2,290 (1.1)	
	위암 관련 암	423,491 (98.9)	4,506 (1.0)	
	기타 암	375,068 (99.0)	3,731 (1.0)	
친가할아버지	없음	4,965,728 (99.2)	41,815 (0.8)	0.1565
	위암	278 (100.0)	0 (0.0)	
	위암 관련 암	645 (99.7)	2 (0.3)	
	기타 암	724 (99.5)	4 (0.5)	
외가할아버지	없음	4,966,794 (99.2)	41,818 (0.8)	0.2700
	위암	112 (100.0)	0 (0.0)	
	위암 관련 암	212 (98.6)	3 (1.4)	
	기타 암	257 (100.0)	0 (0.0)	

가족관계		본인위암 여부 (5,009,196)		P-value
		없음 (4,697,375)	있음 (41,821)	
친가할머니	없음	4,962,557 (99.2)	41,804 (0.8)	0.0028
	위암	609 (99.7)	2 (0.3)	
	위암 관련 암	1,670 (99.8)	4 (0.2)	
	기타 암	2,539 (99.6)	11 (0.4)	
외가할머니	없음	4,966,165 (99.2)	41,818 (0.8)	0.0633
	위암	168 (98.8)	2 (1.2)	
	위암 관련 암	427 (99.8)	1 (0.2)	
	기타 암	615 (100.0)	0 (0.0)	

표4-6는 이번 연구에서 수집한 가족력 DB정보와 2009-2010년 문진을 통해 수집한 가족력 정보를 비교한 표이다. 결측치를 제외하지 않고 총 5,009,186명으로 분석을 하였으며 보다 정확한 비교를 위해 결측을 제외한 상태에서 표4-5에서 추가 분석을 진행했다.

<표 4-6> 가족력 DB와 문진에서의 가족력 정보와의 비교

가족력 DB	문진 항목: 가족력 여부					Total
	암 없음 (90,795)	위암 있음 (116,168)	위암관련암 있음 (82,213)	기타 암 있음 (140,267)	Missing (4,579,753)	
가족력 없음	64,811 (1.9)	77,470 (2.2)	51,655 (1.5)	91,693 (2.6)	3,218,221 (91.9)	3,503,850
위암 가족력 있음	5,231 (1.7)	14,22 (4.5)	4,796 (1.5)	7,748 (2.5)	284,398 (89.9)	316,398
위암관련암 가족력 있음	10,924 (1.7)	12,875 (2.0)	17,992 (2.8)	22,438 (3.5)	568,949 (89.9)	633,178
기타암 가족력 있음	9,829 (1.8)	11,598 (2.1)	7,770 (1.4)	18,338 (3.3)	508,185 (91.4)	555,770
	90,795	16,168	82,213	140,267	4,579,753	5,009,196

표 4-7은 이번 연구에서 수집한 가족력 DB정보와, 2009-2010년 문진을 통해 수집한 가족력 정보를 결측을 제외한 상태에서 비교한 표이다. 가족력 DB에서 위암 가족력이 없다고 한 대상자중 문진항목에서도 위암 가족력이 있다고 대답한 비율이 22.7%, 두 자료원에서 공히 위암 가족력이 있다고 한 비율이 44.5%, 위암 가족력은 28.0%, 기타 암 가족력은 38.65%로 나타났다.

<표 4-7> 가족력 DB와 문진에서의 가족력 정보와의 비교 (missing 제외)

	문진 항목: 가족력 여부				Total
	암 없음 (90,795)	위암 있음 (116,168)	위암관련암 있음 (82,213)	기타 암 있음 (140,267)	
가족력 없음	64,811 (22.7)	77,470 (27.1)	51,655 (18.1)	91,693 (32.1)	285,629
위암 가족력 있음	5,231 (16.3)	14,225 (44.5)	4,796 (15.0)	7,748 (24.2)	32,000
위암관련암 가족력 있음	10,924 (17.0)	12,875 (20.0)	17,992 (28.0)	22,438 (34.9)	64,229
기타암 가족력 있음	9,829 (20.7)	11,598 (24.4)	7,770 (16.3)	18,338 (38.6)	47,535
	90,795	16,168	82,213	140,267	429,443

표 4-8은 위암 및 위암 관련 암, 기타 암 가족력에 따른 본인 위암 위험도를 나타낸 표이다. 위암 가족력 유형은 부모·자녀, 형제·자매를 묶음으로 2가지의 가족력 유형으로 분류하였다. 2가지 유형의 가족력에서 위암, 위암 관련암, 기타 암을 2가지 이상 동시에 가지고 있을 때 위암-위암관련암-기타 암 순으로 우선순위를 부여하여 우선 순위 암의 가족력을 반영하였다.

부모·자녀, 형제·자매가 모두 위암이 있을 때, 본인이 위암에 걸릴 상대 위험도는 부모·자녀, 형제·자매가 모두 위암이 없을 때에 비해 1.51배 높게 나타났다. 부모·자녀, 형제·자매가 모두 위암 관련 암이 있을때의 본인 위암 위험도는 1.13배 높게 나타났다.

<표 4-8> 위암, 위암 관련 암, 기타 암 가족력에 따른 본인 위암 위험도의 단 변수 회귀 분석

	부모,자녀	형제,자매	본인 위암여부		HR (95% CI)	P-value
			없음	있음		
0	없음	없음	3,475,764 (70.0)	28086 (67.2)	1.0	<.0001
1	위암 있음	없음	90,334 (1.8)	600 (1.4)	0.82 (0.76-0.89)	
2	위암관련암 있음	없음	191,929 (3.9)	1198 (2.9)	0.77 (0.73-0.82)	
3	기타암 있음	없음	210,446 (4.2)	1410 (3.4)	0.83 (0.79-0.88)	

			본인 위암여부		HR	P-value
	부모,자녀	형제,자매	없음	있음	(95% CI)	
4	없음	위암 있음	168,063 (3.4)	1937 (4.6)	1.42 (1.36-1.49)	
5	위암 있음	위암 있음	5,802 (0.1)	71 (0.2)	1.51 (1.20-1.91)	
6	위암관련암 있음	위암 있음	11,741 (0.2)	120 (0.3)	1.26 (1.06-1.51)	
7	기타암 있음	위암 있음	14,717 (0.3)	162 (0.4)	1.36 (1.16-1.58)	
8	없음	위암관련암 있음	355,292 (7.2)	3834 (9.2)	1.33 (1.29-1.38)	
9	위암 있음	위암관련암 있음	11,560 (0.2)	125 (0.3)	1.34 (1.12-1.59)	
10	위암관련암 있음	위암관련암 있음	25,839 (0.5)	236 (0.6)	1.13 (0.99-1.28)	
11	기타암 있음	위암관련암 있음	30,800 (0.6)	311 (0.7)	1.25 (1.11-1.39)	
12	없음	기타암 있음	311,714 (6.3)	3210 (7.7)	1.27 (1.23-1.32)	
13	위암 있음	기타암 있음	11,075 (0.2)	91 (0.2)	1.02 (0.83-1.25)	
14	위암관련암 있음	기타암 있음	23,546 (0.5)	193 (0.5)	1.02 (0.88-1.17)	
15	기타암 있음	기타암 있음	28,733 (0.6)	237 (0.6)	1.02 (0.90-1.16)	

표 4-9은 부모·자녀, 또는 형제·자매에서 위암, 위암관련암, 기타 암을 2가지 이상 동시에 가지고 있을 때 위암-위암관련암-기타 암 순으로 우선순위를 지정했을 때 나오는 4가지 가족력에 대해

가족력 없음을 기준으로 각 암별 가족력에 따른 본인 위암의 위험도가 어떻게 달라지는지 분석하였다.

위암 가족력이 없는 대상자에 비해 위암 가족력이 있을 때 본인 위암 위험은 1.23배, 본인의 위암 관련암 이 있을때는 1.14배, 기타암 가족력이 있을때는 1.09배 본인 위암 위험도가 높았다.

<표 4-9> 위암-위암관련암-기타암으로 우선순위를 부여했을 때, 위암 및 위암 관련 암, 기타 암 가족력에 따른 본인 위암 위험도의 단변수 분석

	본인 위암 여부		HR (95% CI)
	X	O	
가족력 없음	3,475,764 (99.2)	28,086 (0.8)	1.0
위암 가족력 있음	313,292 (99.0)	3,106 (1.0)	1.23 (1.18-1.27)
위암관련암 가족력 있음	627,406 (99.1)	5,772 (0.9)	1.14 (1.11-1.17)
기타암 가족력 있음	550,913 (99.1)	4,857 (0.9)	1.09 (1.06-1.12)

표 4-10은 본인 암발생을 종속변수로 가족력 등 위암 위험 요인들과의 관련성을 다변수 회귀분석한 결과이다. BMI, income, diabetes, HTN, endoscopy, smoking, active physical activity 등을 통제했음에도 위암 가족력에 따른 본인 위암 발생은 1.14배 유의하게 높은 것으로 나타났다.

<표 4-10> 본인 암 발생을 종속변수로 가족력 등 위암 위험 요인들과의 관련성

Variables	HR (CI)	P-value
Age		
40-49	1.0	
50-59	2.50 (2.35-2.67)	<.0001
60-69	4.72 (4.42-5.04)	<.0001
≥70	6.21 (5.76-6.69)	<.0001
BMI		
<18.5	1.0	
18.5-22.9	1.09 (0.98-1.21)	0.1131
23.0-24.9	0.97 (0.92-1.02)	0.202
≥25.0	0.96 (0.91-1.01)	0.1129
Income		
의료급여	1.0	
20%	1.07 (0.97-1.17)	0.1711
40%	1.08 (0.99-1.18)	0.0756
60%	1.12 (1.03-1.22)	0.0114
80%	1.12 (1.03-1.22)	0.0101
100%	1.12 (1.03-1.22)	0.0087
hypertension	0.96 (0.92-1.01)	0.082
Endoscopy	1.11 (1.06-1.16)	<.0001

Variables	HR (CI)	P-value
Family history	1.0	
위암 가족력 있음	1.14 (1.05-1.23)	0.0009
위암관련암 가족력 있음	0.99 (0.93-1.05)	0.7845
기타암 가족력 있음	1.03 (0.97-1.10)	0.3113
Smoking		
Non-smoking		
Current smoking	1.007 (1.006-1.008)	<.0001
Physical activity (MET score)		
<600	1	
600-3000	1.02 (0.98-1.06)	0.4286
3000<	1.01 (0.94-1.09)	0.7631

제2절 간암 예측 모델 개발 연구

1. 일반검진자 대상 간암 예측

남성에서 더 발생하며, 나이는 70대에서 가장 많이 발생하였다. 경제수준은 발병과 큰 관련성이 없었으며, 저체중과 비만이 있는 경우 발병률이 높고 허리둘레가 증가한 경우 높았다. 당뇨, 고혈압이 있는 경우 간암발생이 높았으며 간암 있는 군에서 혈압, 혈당 높게 측정되었다. 간수치 중 AST,ALT,r GT 모두 간암군에서 높았으며 지방간을 예측하는 fatty liver index 역시 간암군에서 높았다. 콜레스테롤 지표는 HDL 포함하여 LDL, Triglyceride, cholesterol 모두 낮게 나타났다. <표 4-11 >

생활습관의 경우 흡연력있고 음주력이 높은 경우 간암발생이 높았다. 운동량이 더 많은 군에서 간암발생이 높게 나타났다.<표-4-12>

<표 4-11> 일반 검진자 대상 분석

N(%)	간암여부(N=9,121,523)		p-value
	No 9,065,890 (99.4)	Yes 55,633 (0.6)	
Sex			
Male	4,451,210 (99.1)	40,269 (0.9)	<.0001
Female	4,613,841 (99.7)	15,346 (0.3)	
Age group			<.0001
0-39	876 (97.9)	19 (2.1)	
40-49	3,492,720 (99.8)	7,278 (0.2)	
50-59	2,691,011 (99.45)	14,933 (0.55)	
60-69	1,780,390 (99.01)	17,866 (0.99)	
70-79	940,223 (98.6)	13,452 (1.4)	
≥80	160,670 (98.7)	2,085 (1.3)	
Medical cost			<.0001
Medical aid			
1-4	293,798 (99.4)	1,722 (0.6)	
5-8	1,314,211 (99.4)	8,192 (0.6)	
9-12	1,253,846 (99.4)	7,699 (0.6)	
13-16	1,468,007 (99.4)	9,638 (0.6)	
≥16	1,927,903 (99.4)	12,140 (0.6)	
≥16	2,808,125 (99.4)	16,242 (0.6)	
BMI			<.0001
Mean (SD)	23.8 (3.0)	24.1 (3.2)	
<18.5	241,729 (99.4)	1,507 (0.6)	<.0001
18.5-22.9	3,468,131 (99.5)	19,000 (0.5)	
23-24.9	2,364,754 (99.4)	14,262 (0.6)	
25-29.9	2,709,636 (99.3)	18,693 (0.7)	
≥30	281,640 (99.2)	2,171 (0.8)	
DM			<.0001
No	8,616,539 (99.4)	50,422 (0.6)	
Yes	449,140 (98.85)	5,211 (1.15)	
HTN			<.0001
No	8,187,370 (99.4)	48,335 (0.6)	
Yes	878,329 (99.2)	7,298 (0.8)	
Death			<.0001
No	8,543,297 (99.8)	20,844 (0.2)	
Yes	522,593 (93.8)	34,789 (6.2)	
DBP	76.5 (10.2)	78.0 (10.2)	
SBP	123.4 (15.4)	127.4 (15.9)	<.0001
waist	80.7 (8.4)	84.0 (8.6)	<.0001
FBS	99.2 (24.4)	106.6 (32.9)	<.0001
AST	24.8 (10.5)	39.1 (24.3)	<.0001
ALT	23.5 (14.8)	35.6 (25.9)	<.0001
rGT	32.7 (32.3)	64.7 (59.0)	<.0001
Total Cholesterol	182.4 (24.0)	173.5 (26.8)	<.0001
Triglyceride	128.7 (98.3)	120.2 (94.2)	<.0001
HDL	54.7 (30.9)	53.2 (27.6)	<.0001
LDL	105.5 (52.0)	99.3 (45.4)	<.0001
*Fatty Liver Index	26.4(22.4)	35.6(23.7)	<.0001

(*The fatty liver index (FLI), which is an algorithm based on waist circumference, body mass index (BMI), triglyceride, and gamma-glutamyl-transferase (GGT))

<표4-12> 일반검진자 대상 분석(생활 습관)

간암여부(N=9,121,523)			
Smoking status			
Non-smoker	5,837,090 (99.5)	27,540 (0.5)	<.0001
Former	1,412,528 (99.1)	12,187 (0.9)	
Current	1,779,484 (99.1)	15,707 (0.9)	
Alcohol			
no	909,800 (99.4)	5,418 (0.6)	<.0001
low	3,135,491 (99.4)	19,256 (0.6)	
medium	237,473 (99.3)	1,767 (0.7)	
high	247,238 (98.8)	3,020 (1.2)	
Physical activity (MET Score)			
0-599	4,446,215 (99.4)	27,818 (0.6)	<.0001
600-2999	3,844,775 (99.4)	22,368 (0.6)	
3000≥	774,900 (99.3)	5,447 (0.7)	

2. 국가 간암검진자 대상 간암 예측

나이는 60대에 가장 높은 발병을 보였으며 HBV 음성 229,072명 중 1,048명(0.46%)가 발생한 한편 HBV 양성 6,218명 중 295명(4.74%)에서 간암이 발생하였고, HCV 음성 13,578명 중 177명(1.3%), HCV 양성 633명 중 43명(6.79%)로 나타났다 남성에서 간암 발생이 증가하였으며, 체중, 혈압, creatinine, 간수치인 AST,ALT,rGT는 간암 발생을 증가시키는 경향을 보였으며 콜레스테롤 수치인 저밀도콜레스테롤, 중성지방, 콜레스테롤은 간암 군에서 더 낮게 나타났다. 고밀도 콜레스테롤 역시 낮은 경우 간암발생이 높았다. <표 4-13 > 흡연, 음주, 간암 발생을 증가시키는 경향을 보였으며, 운동은 간암발생과 관련성이 없었다. <표 4-14>

<표4-13> 간암검진자 대상 분석

		Total (n=722,647)	non-HCC (n=702,781(97.2%))	HCC (n=20,556(2.8%))
sex, n(%)	male	374,846	359,700(95.82)	15,681(4.18)
	female	347,801	343,081(98.6)	4,875(1.4)
age, mean±sd		56.77±11.23 (n=722,647)	56.69±11.27 (n=702,776)	59.30±9.54 (n=19,871)
age, n(%)	< 40	7	7(100)	0(0)
	40~50	218,059	214935(98.57)	3124(1.43)
	50~60	226,848	219830(96.91)	7018(3.09)
	60~70	163,207	156768(96.05)	6439(3.95)
	70~80	93,200	90247(96.83)	2953(3.17)
	80 over	21,326	20989(98.42)	337(1.58)
ALT(mean±sd)		23.53±29.23 (n=243,342)	23.48±29.17 (n=241,737)	31.22±36.43 (n=1,605)
HBV, n(%)	negative	229,072	228024(99.54)	1048(0.46)
	positive	6,218	5923(95.26)	295(4.74)
HCV, n(%)	negative	13,578	13401(98.7)	177(1.3)
	positive	633	590(93.21)	43(6.79)
간초음파.관찰소견	정상	162,324	160388(98.81)	1936(1.19)
	거친에코상	52,032	46759(89.87)	5273(10.13)
	간경변	19,685	14358(72.94)	5327(27.06)
	이형결절	1,988	1437(72.28)	551(27.72)
	양성질환	207,628	202726(97.64)	4902(2.36)
	간암의심	2,217	779(35.14)	1438(64.86)
	기타	121,720	114680(94.22)	7040(5.78)
AFP	negative	117,791	113968(96.75)	3823(3.25)
	positive	1,516	912(60.16)	604(39.84)
BMI		24.28±3.11	24.28±3.11	24.29±3.1
SBP		124.32±15.18	124.26±15.16	126.05±15.38
DBP		77.1±10.13	77.08±10.13	77.44±10.1
Hb		13.95±1.57	13.95±1.57	14.05±1.59
FBS		102.23±27.4	102.03±27.1	107.57±34.04
AST		33.68±39.07	32.83±38.1	56.63±54.61
ALT		33.95±43.47	33.32±42.78	50.73±56.5
rGT		56.79±90.91	54.71±87.31	112.32±149.17
Total cholesterol		192.95±42.17	193.6±41.89	175.48±45.78
Triglyceride		137.78±112.37	138.74±112.75	112.19±98.08
HDL		55.23±31.45	55.26±31.49	54.47±30.44
LDL		113.91±74.23	114.34±74.55	102.42±64.09
Cr		1.03±1.14	1.03±1.14	1.07±1.22
GFR		82.03±25.6	82.09±25.52	80.42±27.58

<표4-14> 간암건진자 대상 분석(생활습관)

		Total (n=722,647)	non-HCC (n=702,781(97.2%))	HCC (n=20,556(2.8%))
smoking	Non	268300	260882(97.24)	7418(2.76)
	Former	99061	94642(95.54)	4419(4.46)
	Current	103882	98731(95.04)	5151(4.96)
alcohol	non	259438	250340(96.49)	9098(3.51)
	1	75206	72873(96.9)	2333(3.1)
	2	50661	48879(96.48)	1782(3.52)
	3	36993	35472(95.89)	1521(4.11)
	4	14217	13618(95.79)	599(4.21)
	5	12228	11697(95.66)	531(4.34)
	6	7066	6737(95.34)	329(4.66)
	daily	13887	13159(94.76)	728(5.24)
	exercise	0-599	216316	208368(96.33)
600-2999		205050	197841(96.48)	7209(3.52)
3000≥		46957	45249(96.36)	1708(3.64)

제3절 대사증후군의 유전형과 표현형 융합분석

1. 연구대상자의 임상적 특성과 생활습관

(1) 대사증후군 발생 여부에 따른 연구대상자 특성 파악

전체 연구대상자 총 249명 중 대사증후군은 66명으로 유병률은 26.51%였다. 전체적으로 나이가 40~49세와 50~59세에 해당하는 비율이 81.93%로 높았으며 성별은 여자(56.22%)가 남자(43.78%)보다 더 많았다. 그러나 대사증후군에 타입에서는 남자(57.58%)가 여자(42.42%)보다 더 많았다. 대사증후군 타입의 경우 비대사증후군 타입에 비해 연속변수인 체질량지수, 허리둘레, 수축기혈압, 이완기혈압, 공복혈당, 중성지방 평균이 더 높았으며 고밀도콜레스테롤의 경우 더 낮게 나왔다. 또한 대사증후군 타입의 경우 현재흡연인 사람이 13.64%로 비대사증후군 중 현재흡연인 사람 13.11%에 비해 더 많았다. 음주의 경우 저위험과 중위험, 고위험 음주를 합하여

보면 대사증후군 중 72.73%를 차지하고 비대사증후군 중 62.84%를 차지한다. 그리고 운동의 경우 대사증후군 타입에서 저활동군은 36.36%로 비대사증후군 중 저활동군인 39.34% 보다 낮았다. <표 4-15>

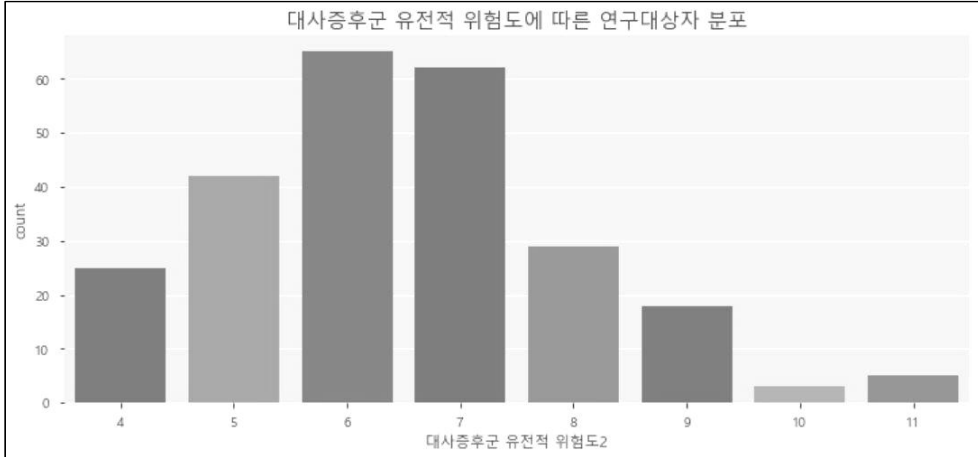
<표 4-15> 대사증후군 여부에 따른 연구대상자 특성

변수	대사증후군 여부	비대사증후군 (n=183)	대사증후군 (n=66)	Total (n=249)
나이	<40	16 (8.74%)	1 (1.52%)	17 (6.83%)
	40-49	88 (48.09%)	29 (43.94%)	117 (46.99%)
	50-59	59 (32.24%)	28 (42.42%)	87 (34.94%)
	60-69	19 (10.38%)	8 (12.12%)	27 (10.84%)
	≥70	1 (0.55%)	0 (0%)	1 (0.40%)
성별	남	71 (38.80%)	38 (57.58%)	109 (43.78%)
	여	112 (61.20%)	28 (42.42%)	140 (56.22%)
체질량지수		22.65 ± 2.64	26.70 ± 2.91	23.72 ± 3.25
허리둘레		78.87 ± 8.05	90.20 ± 7.09	81.87 ± 9.26
수축기혈압		122.40 ± 11.09	134.44 ± 10.13	125.59 ± 12.06
이완기혈압		75.17 ± 8.41	82.17 ± 8.81	77.03 ± 9.04
공복혈당		95.76 ± 17.25	110.45 ± 20.78	99.65 ± 19.33
고밀도콜레스테롤		59.63 ± 14.47	47.70 ± 8.71	56.47 ± 14.19
중성지방		120.85 ± 75.60	221.86 ± 121.56	147.63 ± 100.32
흡연	안피움	125 (68.31%)	40 (60.60%)	165 (66.27%)
	과거흡연	34 (18.58%)	17 (25.76%)	51 (20.48%)
	현재흡연	24 (13.11%)	9 (13.64%)	33 (13.25%)
음주	안마심	68 (37.16%)	18 (27.27%)	86 (34.54%)
	저위험	64 (34.97%)	17 (25.76%)	81 (32.53%)
	중위험	15 (8.20%)	9 (13.64%)	24 (9.64%)
	고위험	36 (19.67%)	22 (33.33%)	58 (23.29%)
운동	저활동군	72 (39.34%)	24 (36.36%)	96 (38.56%)
	중등도활동군	105 (57.38%)	39 (59.10%)	144 (57.83%)
	고활동군	6 (3.28%)	3 (4.54%)	9 (3.61%)

(2) 유전적 위험도에 따른 연구대상자 분포

대사증후군 유전적 위험도에 따라 연구대상자의 분포를 <그림 4-2>를 통해 살펴 보았다. 유전적 위험도 '낮음' 타입이 132명으로 53.01%, 유전적 위험도 '중간' 타입이

109명으로 43.78%, 유전적 위험도 '높음' 타입이 8명으로 3.21%를 차지하였다. <표 4-16>



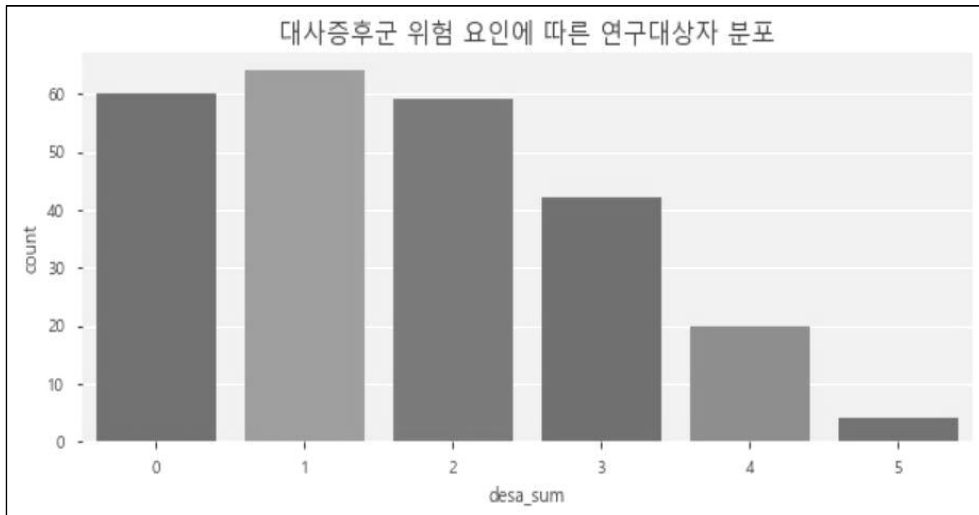
[그림 4-1] 대사증후군 유전적 위험도에 따른 연구대상자 분포

<표 4-16> 유전적 위험도에 따른 연구대상자 분포

유전적 위험도 타입	유전적 위험도	수	비율	수	비율
유전적 위험도 = '낮음'	4	25	10.04%	132	53.01%
	5	42	16.87%		
	6	65	26.10%		
유전적 위험도 = '중간'	7	62	24.90%	109	43.78%
	8	29	11.65%		
	9	18	7.23%		
유전적 위험도 = '높음'	10	3	1.20%	8	3.21%
	11	5	2.01%		
	12	0	0%		

(3) 대사증후군 위험요인에 따른 연구대상자 분포

대사증후군 위험요인에 따라 연구대상자의 분포를 <그림 4-2>을 통해 살펴보았다. 대사증후군 위험요인 수가 0개인 '정상군' 타입이 60명으로 24.10%, 대사증후군 위험요인 수가 1-2개인 '위험군' 타입이 123명으로 49.39%, 대사증후군 위험요인 수가 3개 이상인 '대사증후군' 타입이 66명으로 26.51%를 차지하였다. <표 4-17>



[그림 4-2] 대사증후군 위험요인에 따른 연구대상자 분포

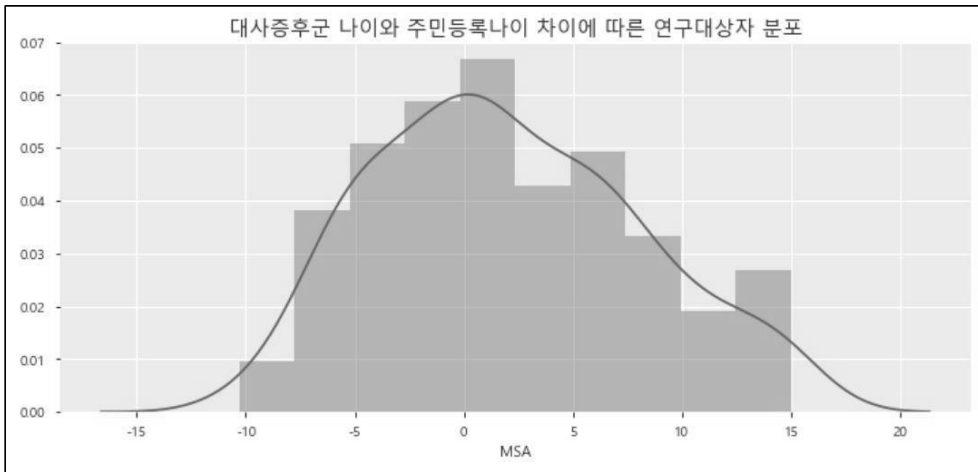
<표 4-17> 대사증후군 위험요인에 따른 연구대상자 분포

대사증후군 타입	위험요인 수	수	비율	수	비율
정상군	0개	60	24.10%	60	24.10%
위험군	1개	64	25.70%	123	49.39%
	2개	59	23.69%		
대사증후군	3개	42	16.87%	66	26.51%
	4개	20	8.03%		
	5개	4	1.61%		

(4) 대사증후군 나이에 따른 연구대상자 분포

대사증후군 나이에 따라 연구대상자의 분포를 <그림 4-3>를 통해 살펴보았다. 대사증후군 나이 ' 좋음' 타입이 37명으로 14.86%, 대사증후군 나이 '평균' 타입이 131명으로 52.61%, 대사증후군 나이 '나쁨' 타입이 81명으로 32.53%를 차지하였다.

<표 4-18>



[그림 4-3] 대사증후군 나이에 따른 연구대상자 분포

<표 4-18> 대사증후군 나이에 따른 연구대상자 분포

대사증후군 나이 타입	대사증후군 나이 - 주민등록나이	수	비율
좋음	≤ -5세	37	14.86%
평균	-5세 ~ +5세	131	52.61%
나쁨	≥ +5세	81	32.53%

2. 유전적 위험도와 대사증후군 융합분석

(1) 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포

유전적 위험도와 대사증후군 기준을 이용하여 9개의 타입으로 분류하고 연구대상자 분포를 살펴보았다. <표 4-19>

유전적 위험도가 '낮음'에서 '대사증후군' 타입은 36명(27.27%), 유전적 위험도가 '중간'에서 '대사증후군' 타입은 26명(23.85%), 유전적 위험도가 '높음'에서 '대사증후군' 타입은 4명(50.00%)으로 유전적 위험도가 '낮음'이나 '중간'보다는 유전적 위험도가 '높음'에서 대사증후군 발생률이 더 높았다. <표 4-20>

정상군, 위험군, 대사증후군에서 유전적 위험도가 '높음' 타입을 보면 정상군이 1명(1.67%), 위험군이 3명(2.44%), 대사증후군이 4명(6.06%)로 높아지는 양상이었다. <표 4-21>

<표 4-19> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포(빈도)

유전적 위험도 \ 대사증후군 타입	정상군 (n=60)	위험군 (n=123)	대사증후군 (n=66)
낮음 (n=132)	30	66	36
중간 (n=109)	29	54	26
높음 (n=8)	1	3	4

<표 4-20> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포

(row 기준 백분율)

유전적 위험도 \ 대사증후군 타입	정상군 (n=60)	위험군 (n=123)	대사증후군 (n=66)	Total (n=249)
낮음 (n=132)	22.73%	50.00%	27.27%	100%
중간 (n=109)	26.61%	49.54%	23.85%	100%
높음 (n=8)	12.50%	37.50%	50.00%	100%

<표 4-21> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포

(column 기준 백분율)

유전적 위험도 \ 대사증후군 타입	정상군 (n=60)	위험군 (n=123)	대사증후군 (n=66)
낮음 (n=132)	50.00%	53.66%	54.55%
중간 (n=109)	48.33%	43.90%	39.39%
높음 (n=8)	1.67%	2.44%	6.06%
Total (n=249)	100%	100%	100%

(2) 유전적 위험도와 대사증후군 기준 타입 별 유전적 위험도 평균 수치 비교

유전적 위험도와 대사증후군 기준을 이용한 9개의 타입 별 유전적 위험도 평균 수치를 비교하였다. 유전적 위험도 평균을 보면 정상군에서 6.42, 위험군에서 6.46, 대사증후군에서 6.58로 높아지는 양상을 보였다. <표 4-22>

<표 4-22> 유전적 위험도와 대사증후군 기준 타입 별 유전적 위험도 평균 수치

대사증후군 타입 \ 유전적 위험도	정상군 (n=60)	위험군 (n=123)	대사증후군 (n=66)	평균
낮음 (n=132)	5.23	5.26	5.44	5.30
중간 (n=109)	7.48	7.69	7.54	7.60
높음 (n=8)	11.00	10.70	10.50	10.63
평균	6.42	6.46	6.58	

*p-value ≥ 0.05 (NS)

3. 유전적 위험도와 대사증후군 나이 융합분석

(1) 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포

유전적 위험도와 대사증후군 나이 기준을 이용하여 9개의 타입으로 분류하고 연구대상자 분포를 살펴보았다. <표 4-23>

유전적 위험도가 '낮음'에서 대사증후군 나이 '나쁨' 타입은 44명(33.33%), 유전적 위험도가 '중간'에서 대사증후군 나이 '나쁨' 타입은 33명(30.28%), 유전적 위험도가 '높음'에서 대사증후군 나이 '나쁨' 타입은 4명(50.00%)으로 유전적 위험도가 '낮음'이나 '중간'보다는 유전적 위험도가 '높음'에서 대사증후군 나이 '나쁨' 타입이 더 많았다. <표 4-24>

각 대사증후군 나이 그룹에서 유전적 위험도가 '높음' 타입을 보면 대사증후군 나이 '좋음'이 0명(0%), 대사증후군 나이 '평균'이 4명(3.05%), 대사증후군 나이 '나쁨'이 4명(4.94%)으로 높아지는 양상이었다. <표 4-25>

<표 4-23> 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포(빈도)

유전적 위험도 \ 대사증후군 나이	좋음 (n=37)	평균 (n=131)	나쁨 (n=81)
낮음 (n=132)	18	70	44
중간 (n=109)	19	57	33
높음 (n=8)	0	4	4

<표 4-24> 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포
(row 기준 백분율)

유전적 위험도 \ 대사증후군 나이	좋음 (n=37)	평균 (n=131)	나쁨 (n=81)	Total (n=249)
낮음 (n=132)	13.64%	53.03%	33.33%	100%
중간 (n=109)	17.43%	52.29%	30.28%	100%
높음 (n=8)	0%	50.00%	50.00%	100%

<표 4-25> 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포
(column 기준 백분율)

유전적 위험도 \ 대사증후군 나이	좋음 (n=37)	평균 (n=131)	나쁨 (n=81)
낮음 (n=132)	48.65%	53.44%	54.32%
중간 (n=109)	51.35%	43.51%	40.74%
높음 (n=8)	0%	3.05%	4.94%
Total (n=249)	100%	100%	100%

(2) 유전적 위험도와 대사증후군 나이 기준 타입 별 유전적 위험도 평균 수치 비교
유전적 위험도와 대사증후군 나이 기준을 이용한 9개의 타입 별 유전적 위험도 평균 수치를 비교하였다. 유전적 위험도 평균을 보면 대사증후군 나이 '좋음' 타입에서 6.41, 대사증후군 나이 '평균' 타입에서 6.49, 대사증후군 나이 '나쁨' 군에서 6.49로 대사증후군 나이 '좋음' 타입보다는 대사증후군 나이 '평균' 타입, '나쁨' 타입에서 더 높았다. <표 4-26>

<표 4-26> 유전적 위험도와 대사증후군 나이 기준 타입 별 유전적 위험도 평균 수치

*유전적 위험도 \ 대사증후군 나이	좋음 (n=37)	평균 (n=131)	나쁨 (n=81)	평균
낮음 (n=132)	5.11	5.34	5.32	5.30
중간 (n=109)	7.63	7.60	7.58	7.60
높음 (n=8)		10.80	10.50	10.63
평균	6.41	6.49	6.49	

*p-value ≥ 0.05 (NS)

4. 유전적 위험도 타입 별 생활습관에 따른 대사증후군 발생률

(1) 유전적 위험도 '낮음'

유전적 위험도 '낮음' 타입에 대해 생활습관에 따른 대사증후군 발생률을 살펴보았다. 흡연 기준으로 현재흡연이 23.53%로 안피움 22.35%보다 대사증후군 발생률이 높았다. 음주 기준으로는 음주를 하는 타입이 30.93%로 안하는 타입 17.14%보다 대사증후군 발생률이 높았다. 운동에서는 저활동군 타입이 29.63%로 활동군 타입 25.64%보다 대사증후군 발생률이 높았다. <표 4-27>

<표 4-27> 유전적 위험도 '낮음' - 생활습관에 따른 대사증후군 발생률

유전적 위험도='낮음'						
기준		빈도			백분율	
		비대사증후군	대사증후군	Total	비대사증후군	대사증후군
※흡연	안피움	66	19	85	77.65%	22.35%
	현재흡연	13	4	17	76.47%	23.53%
※음주	무	29	6	35	82.86%	17.14%
	유	67	30	97	69.07%	30.93%
※운동	저활동군	38	16	54	70.37%	29.63%
	활동군	58	20	78	74.36%	25.64%

*p-value ≥ 0.05 (NS)

(2) 유전적 위험도 '중간'

유전적 위험도 '중간' 타입에 대해 생활습관에 따른 대사증후군 발생률을 살펴보았다. 흡연 기준으로 현재흡연이 33.33%로 안피움 23.29%보다 대사증후군 발생률이 높았다. 음주 기준으로는 음주를 안하는 타입이 26.32%로 하는 타입 22.54%보다 대사증후군 발생률이 높았다. 운동에서는 활동군 타입이 27.14%로 저활동군 타입 17.95%보다 대사증후군 발생률이 높았다. <표 4-28>

<표 4-28> 유전적 위험도 '중간' - 생활습관에 따른 대사증후군 발생률

유전적 위험도='중간'						
기준		빈도			백분율	
		비대사증후군	대사증후군	Total	비대사증후군	대사증후군
*흡연	안피움	56	17	73	76.71%	23.29%
	현재흡연	10	5	15	66.67%	33.33%
*음주	무	28	10	38	73.68%	26.32%
	유	55	16	71	77.46%	22.54%
*운동	저활동군	32	7	39	82.05%	17.95%
	활동군	51	19	70	72.86%	27.14%

* p-value ≥ 0.05 (NS)

(3) 유전적 위험도 '높음'

유전적 위험도 '높음' 타입에 대해 생활습관에 따른 대사증후군 발생률을 살펴보았다. 흡연 기준으로 안피움이 57.14%로 현재흡연 0%보다 대사증후군 발생률이 높았다. 음주 기준으로는 음주를 하는 타입이 60%로 안하는 타입 33.33%보다 대사증후군 발생률이 높았다. 운동에서는 활동군 타입이 60%로 저활동군 타입 33.33%보다 대사증후군 발생률이 높았다. <표 4-29>

<표 4-29> 유전적 위험도 '높음' - 생활습관에 따른 대사증후군 발생률

유전적 위험도='높음'						
기준		빈도			백분율	
		비대사증후군	대사증후군	Total	비대사증후군	대사증후군
*흡연	안피움	3	4	7	42.86%	57.14%
	현재흡연	1	0	1	100.00%	0.00%
*음주	무	2	1	3	66.67%	33.33%
	유	2	3	5	40.00%	60.00%
*운동	저활동군	2	1	3	66.67%	33.33%
	활동군	2	3	5	40.00%	60.00%

* p-value ≥ 0.05 (NS)

5. 유전적 위험도 재조정을 통한 추가 분석

(1) 재조정된 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포

유전적 위험도를 기준으로 유전적 위험도 타입을 재조정하는 추가 분석을 진행하였다. 유전적 위험도 4-5점인 재조정된 유전적 위험도 '낮음' 타입이 67명으로 26.91%, 유전적 위험도 6-7점인 재조정된 유전적 위험도 '중간' 타입이 127명으로 51.00%, 유전적 위험도 8-12점인 재조정된 유전적 위험도 '높음' 타입이 55명으로 22.09%를 차지하였다. <표 4-30>

재조정된 유전적 위험도와 대사증후군 기준을 이용하여 9개의 타입으로 분류하고 연구대상자 분포를 살펴보았다. <표 4-31>

재조정된 유전적 위험도가 '낮음'에서 '대사증후군' 타입은 15명(22.39%), '중간'에서 '대사증후군' 타입은 37명(29.13%), '높음'에서 '대사증후군' 타입은 14명(25.45%)으로 재조정된 유전적 위험도가 '낮음'보다는 '중간'이나 '높음'에서 대사증후군 발생률이 더 높았다. <표 4-32>

정상군, 위험군, 대사증후군에서 재조정된 유전적 위험도가 '높음' 타입을 보면 정상군이 11명(18.33%), 위험군이 30명(24.39%), 대사증후군이 14명(21.21%)로 정상군보다는 위험군이나 대사증후군에서 재조정된 유전적 위험도 '높음' 타입이 더 많았다. <표 4-33>

<표 4-30> 재조정된 유전적 위험도에 따른 연구대상자 분포

재조정된 유전적 위험도 타입	유전적 위험도	수	비율	수	비율
유전적 위험도 = '낮음'	4	25	10.04%	67	26.91%
	5	42	16.87%		
유전적 위험도 = '중간'	6	65	26.10%	127	51.00%
	7	62	24.90%		
유전적 위험도 = '높음'	8	29	11.65%	55	22.09%
	9	18	7.23%		
	10	3	1.20%		
	11	5	2.01%		
	12	0	0%		

<표 4-31> 재조정된 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포(빈도)

재조정된 유전적 위험도 \ 대사증후군 타입	정상군 (n=60)	위험군 (n=123)	대사증후군 (n=66)
낮음 (n=67)	18	34	15
중간 (n=127)	31	59	37
높음 (n=55)	11	30	14

<표 4-32> 재조정된 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포
(row 기준 백분율)

재조정된 유전적 위험도 \ 대사증후군 타입	정상군 (n=60)	위험군 (n=123)	대사증후군 (n=66)	Total (n=249)
낮음 (n=67)	26.86%	50.75%	22.39%	100%
중간 (n=127)	24.41%	46.46%	29.13%	100%
높음 (n=55)	20.00%	54.55%	25.45%	100%

<표 4-33> 재조정된 유전적 위험도와 대사증후군 기준 타입 별 연구대상자 분포
(column 기준 백분율)

재조정된 유전적 위험도 \ 대사증후군 타입	정상군 (n=60)	위험군 (n=123)	대사증후군 (n=66)
낮음 (n=67)	30.00%	27.64%	22.73%
중간 (n=127)	51.67%	47.97%	56.06%
높음 (n=55)	18.33%	24.39%	21.21%
Total (n=249)	100%	100%	100%

(2) 재조정된 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포

재조정된 유전적 위험도와 대사증후군 나이 기준을 이용하여 9개의 타입으로 분류하고 연구대상자 분포를 살펴보았다. <표 4-34>

재조정된 유전적 위험도가 ‘낮음’에서 대사증후군 나이 ‘나쁨’ 타입은 21명(31.34%), ‘중간’에서 대사증후군 나이 ‘나쁨’ 타입은 44명(34.64%), ‘높음’에서 대사증후군 나이 ‘나쁨’ 타입은 16명(29.09%)으로 재조정된 유전적 위험도에 따른 대사증후군 나이의 경향성은 보이지 않았다. <표 4-35>

각 대사증후군 나이 그룹에서 재조정된 유전적 위험도가 ‘높음’ 타입을 보면 대사증후군 나이 ‘ 좋음’이 8명(21.62%), 대사증후군 나이 ‘평균’이 31명(23.66%), 대사증후

군 나이 '나쁨'이 16명(19.75%)으로 대사증후군 나이에 따른 재조정된 유전적 위험도 '높음' 비율의 경향성은 보이지 않았다. <표 4-36>

<표 4-34> 재조정된 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포(빈도)

대사증후군 나이 재조정된 유전적 위험도	좋음 (n=37)	평균 (n=131)	나쁨 (n=81)
낮음 (n=67)	13	33	21
중간 (n=127)	16	67	44
높음 (n=55)	8	31	16

<표 4-35> 재조정된 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포
(row 기준 백분율)

대사증후군 나이 재조정된 유전적 위험도	좋음 (n=37)	평균 (n=131)	나쁨 (n=81)	Total (n=249)
낮음 (n=67)	19.40%	49.26%	31.34%	100%
중간 (n=127)	12.60%	52.76%	34.64%	100%
높음 (n=55)	14.55%	56.36%	29.09%	100%

<표 4-36> 재조정된 유전적 위험도와 대사증후군 나이 기준 타입 별 연구대상자 분포
(column 기준 백분율)

대사증후군 나이 재조정된 유전적 위험도	좋음 (n=37)	평균 (n=131)	나쁨 (n=81)
낮음 (n=67)	35.14%	25.19%	25.93%
중간 (n=127)	43.24%	51.15%	54.32%
높음 (n=55)	21.62%	23.66%	19.75%
Total (n=249)	100%	100%	100%

(3) 재조정된 유전적 위험도 '낮음' - 생활습관에 따른 대사증후군 발생률

재조정된 유전적 위험도 '낮음' 타입에 대해 생활습관에 따른 대사증후군 발생률을 살펴보았다. 흡연 기준으로 안피움이 19.51%로 현재흡연 8.33%보다 대사증후군 발생률이 높았다. 음주 기준으로는 음주를 하는 타입이 26.42%로 안하는 타입 7.14%보다 대사증후군 발생률이 높았다. 운동에서는 저활동군 타입이 28.57%로 활동군 타입 17.95%보다 대사증후군 발생률이 높았다. <표 4-37>

<표 4-37> 재조정된 유전적 위험도 '낮음' - 생활습관에 따른 대사증후군 발생률

재조정된 유전적 위험도='낮음'						
기준		빈도			백분율	
		비대사증후군	대사증후군	Total	비대사증후군	대사증후군
※흡연	안피움	33	8	41	80.49%	19.51%
	현재흡연	11	1	12	91.67%	8.33%
※음주	무	13	1	14	92.86%	7.14%
	유	39	14	53	73.58%	26.42%
※운동	저활동군	20	8	28	71.43%	28.57%
	활동군	32	7	39	82.05%	17.95%

(4) 재조정된 유전적 위험도 '중간' - 생활습관에 따른 대사증후군 발생률

재조정된 유전적 위험도 '중간' 타입에 대해 생활습관에 따른 대사증후군 발생률을 살펴보았다. 흡연 기준으로 현재흡연이 43.75%로 안피움 24.14%보다 대사증후군 발생률이 높았다. 음주 기준으로는 음주를 하는 타입이 31.76%로 안하는 타입 23.81%보다 대사증후군 발생률이 높았다. 운동에서는 활동군 타입이 32.10%로 저활동군 타입 23.91%보다 대사증후군 발생률이 높았다. <표 4-38>

<표 4-38> 재조정된 유전적 위험도 '중간' - 생활습관에 따른 대사증후군 발생률

재조정된 유전적 위험도='중간'						
기준		빈도			백분율	
		비대사증후군	대사증후군	Total	비대사증후군	대사증후군
※흡연	안피움	66	21	87	75.86%	24.14%
	현재흡연	9	7	16	56.25%	43.75%
※음주	무	32	10	42	76.19%	23.81%
	유	58	27	85	68.24%	31.76%
※운동	저활동군	35	11	46	76.09%	23.91%
	활동군	55	26	81	67.90%	32.10%

(5) 재조정된 유전적 위험도 '높음' - 생활습관에 따른 대사증후군 발생률

재조정된 유전적 위험도 '높음' 타입에 대해 생활습관에 따른 대사증후군 발생률을

살펴보았다. 흡연 기준으로 안피움이 29.73%로 현재흡연 20.00%보다 대사증후군 발생률이 높았다. 음주 기준으로는 음주를 안하는 타입이 30.00%로 하는 타입 22.86%보다 대사증후군 발생률이 높았다. 운동에서는 활동군 타입이 27.27%로 저활동군 타입 22.73%보다 대사증후군 발생률이 높았다. <표 4-39>

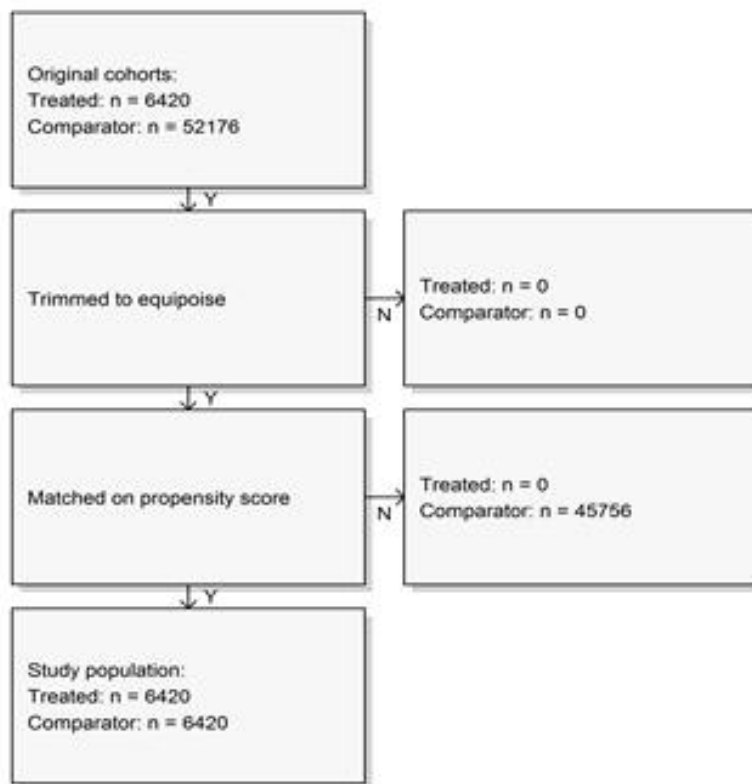
<표 4-39> 재조정된 유전적 위험도 '높음' - 생활습관에 따른 대사증후군 발생률

재조정된 유전적 위험도='높음'						
기준		빈도			백분율	
		비대사증후군	대사증후군	Total	비대사증후군	대사증후군
* 흡연	안피움	26	11	37	70.27%	29.73%
	현재흡연	4	1	5	80.00%	20.00%
* 음주	무	14	6	20	70.00%	30.00%
	유	27	8	35	77.14%	22.86%
* 운동	저활동군	17	5	22	77.27%	22.73%
	활동군	24	9	33	72.73%	27.27%

유전적 위험도 재조정을 통한 추가 분석에서는 이전 분석에 비해 향상된 결과를 보이지 않았다.

제4절 위암과 헬리코박터 관련성 분석을 통해 CDM모델의 정합성 검증

CDM 구축을 위하여 국민보험공단 일산병원에서 위내시경을 시행받은 사람 중 헬리코박터 양성, 위궤양, 십이지장궤양, 위축성위염 진단받은 대상자를 선정하였고 이는 총 6420명이었다. <그림-4-5>



[그림 4-4] CDM 연구 대상자 선정

2011년 이후 위암이 발생한 군에서 헬리코박터 양성은 39명이었고, 위암이 발생하지 않은 군에서 헬리코박터 양성은 24명이었고 이에 관해 Harzard ratio 분석한 결과 1.50이었으나 유의하지는 않았다(p-value 0.138) <표 4-40>. 추가로 성별, 나이, 위/십이지장

궤양, 위축성위염 보정한 결과는 2011년 이후 위암이 발생한 군에서 헬리코박터 양성은 39명이었고, 위암이 발생하지 않은 군에서 헬리코박터 양성은 23명이었고 이에 관해 Harzard ratio 분석한 결과 1.48이었으나 유의하지는 않았다 (p-value 0.151)

<표4-41>.

<표 4-40> ulcer 보정 CDM 결과

Exposure	유	무	HR*	95% lower C.I.	95% upper C.I.	p-value
헬리코박터 여부	39/6420	24/6420	1.50	0.88	2.56	0.138

<표 4-41> ulcer, atrophic, gastritis, age, sex 보정 CDM 결과

Exposure	유	무	HR*	95% lower C.I.	95% upper C.I.	p-value
헬리코박터 여부	39/6419	23/6419	1.48	0.87	2.53	0.151

제5장

결론 및 제언

제1절 결 론	77
제2절 제 언	79

제5장 결론 및 제언

제1절 결론

1. 위암 가족력 모형 연구

단변수 회귀분석 결과 부모의 가족력이 없을 때 형제 자매의 위암가족력이 1.42배로 유의하게 높게 나타난 반면에 형제자매의 가족력이 없을 때 부모의 위암 가족력은 통계적으로 유의하지 않았던 결과는, 연구의 대상자들의 연령에 따른 결과로 보인다. 즉 연구에 포함된 위암검진 수진자의 대부분은 40대이상 중년들이며, 이들의 출생시점인 1970년대 이전의 경우 2000년대 이후에 비해 출산율이 높았고 외동인 가족이 많지 않아 형제 자매가 있을 경우가 높았다. 즉 형제 자매가 많았기에, 형제 자매의 가족력은 보다 집계가 잘되었기에 상대적으로 그로 인한 가족력의 위험이 더 높아 보일 수 있으므로, 부모에 의한 가족력과의 직접 비교를 하기에는 주의가 필요하다.

가족력을 포함한 다변량 회귀분석결과에서 다른 변수들을 고려한 상태에서 위암 가족력이 유의하게 본인 위암 위험도를 높인 것은 의미가 있는 것으로 보인다. 위암관련암 및 기타암과의 관련성이 유의하지 않았던 점은 실제로 위암관련암과 기타암이 위암과의 관련성이 명백하게 높지 않았기 때문으로 보인다.

2. 간암 예측 연구

일반검진자 대상으로 위험인자, 생활습관과 간암의 관련성을 분석한 결과 남성에서 더 발생하며, 나이는 70대에서 가장 많이 발생하였고, 체질량지수, 허리둘레, 혈당, 고혈압, 간혈청검사 결과, fatty liver index 등이 상관성을 보였다. 생활습관의 경우 간암발생군에서 흡연력이 있고 음주량과 운동량이 높게 나타났다.

간암검진자 대상으로 간암검진 결과를 포함하여, 여러 위험인자와 간암과의 관련성을 분석한 결과, 남성, 60대에 가장 높은 발병을 보였다. 간암 군에서 B형 간염 음성 0.46%, 양성 4.74%로 높게 나타났고 C형 간염 경우 음성 1.3%, 양성 6.79%로 나타났다. 체중,

혈압, creatinine, 간혈청검사인 AST,ALT, rGTP 상승은 간암과 상관성을 보였으며 콜레스테롤 수치인 LDL, Triglyceride, total cholesterole은 간암 군에서 더 낮게 나타났다. HDL 역시 낮은 경우 간암발생이 높았다. 생활습관인자인 흡연, 음주이 높아지는 것과 간암이 상관성을 보였으나 운동은 간암발생과 관련성이 없었다

3. 대사증후군의 유전형과 표현형 융합분석

총 249명의 연구대상자 정보를 통해 대사증후군의 표현형과 유전자형을 융합 분석할 수 있었다. 유전적 위험도가 '낮음'(27.27%)이나 '중간'(23.85%)보다는 '높음'(50.00%)에서 대사증후군 발생률이 더 높았고 유전적 위험도가 '낮음'(33.33%)이나 '중간'(30.28%)보다는 '높음'(50.00%)에서 대사증후군 나이가 '나쁨'이 더 많았다. 유전적 위험도 '높음' 타입을 기준으로 정상군(1.67%), 위험군(2.44%), 대사증후군(6.06%)으로 갈수록 유전적 위험도 '높음' 비율이 점점 높아지는 양상을 보였고 유전적 위험도 '높음' 타입을 기준으로 대사증후군 나이가 ' 좋음'(0%), 대사증후군 나이가 '평균'(3.05%), 대사증후군 나이가 '나쁨'(4.94%)으로 갈수록 유전적 위험도 '높음' 비율이 점점 높아지는 양상을 보였다. 그리고 대사증후군 기준 타입 별 유전적 위험도 평균 수치는 정상군(6.42), 위험군(6.46), 대사증후군(6.58)로 통계적으로 유의하지는 않지만 높아지는 경향을 보여주고 대사증후군 나이 기준 타입 별 유전적 위험도 평균 수치는 대사증후군 나이 ' 좋음'(6.41), 대사증후군 나이 '평균'(6.49), 대사증후군 나이 '나쁨'(6.49)로 ' 좋음'보다는 '평균'과 '나쁨' 타입의 유전적 위험도 평균 수치가 더 높은 경향을 보여주었다. 또한 유전적 위험도 '낮음' 타입에서 흡연을 하는 집단, 음주를 하는 집단, 저활동군 집단의 대사증후군 발생률이 더 높았다.

본 연구는 생활습관과 유전적 요인이 대사증후군 발생과 관계가 있다는 것을 파악했다는 것에 그 의의가 있다고 할 수 있다.

4. CDM모델의 정합성 검정

위내시경을 실시하여 헬리코박터균 양성, 위궤양, 십이지장궤양, 위축성위염 진단받은 대상자는 총 6420명이었다. 2011년 이후 위암이 발생한 군에서 헬리코박터 양성은 39명이었고, 위암이 발생하지 않은 군에서 헬리코박터 양성은 24명이었고 이에 관해 Harzard ratio 분석한 결과 1.50이었으나 유의하지는 않았다(p-value 0.138), 추가로 성별, 나이, 위/십이지장궤양, 위축성위염 보정한 결과는 2011년 이후 위암이 발생한 군에서 헬리코박터 양성은 39명이었고, 위암이 발생하지 않은 군에서 헬리코박터 양성은 23명이었고 이에 관해 Harzard ratio 분석한 결과 1.48이었으나 유의하지는 않았다(p-value 0.151)

제2절 제 언

1. 위암 질환 예측 연구

5차년도(2018년) 연구에서는, 1996-1997년 건강검진자료를 기반으로 개발된 남병호 교수 연구팀의 위험 예측 모형을 2004-2005년도 국가 일반검진 수검자 대상으로 2015년 까지 추적하여 10년 위암 발생 위험 예측 모델을 구현하였다. 그러나 나이, 흡연 등을 제외한 다른 위험 요인에서 상대 위험도가 1에 가까워 범주간에 차이가 나타나지 않아 위험을 예측하는데 한계가 있었을 뿐 아니라, 그 외에 위암 위험요인으로 알려진 영양, 헬리코박터 균 감염, 가족력 중 현재 데이터베이스상 에서 구축 가능한 위험요인이 가족력이었기에 올해 연구에서는 위암 가족력 세부 추가를 통해서 예측 모델을 고도화 하고자 했다. 본 연구에서의 가족력 분석은 대상자 수를 고려하여 1촌(부모), 2촌(형제·자매) 가족력까지 연구범위에 포함시켰다.

전체 대상자의 본인 위암 여부와 가족위암 발생의 가족관계별 분포를 분석한 기술통계 결과에서 부모의 위암 가족력과 본인의 위암 여부의 관련성이 역 상관성을 보이는 것처럼 나온 결과와 관련해서는, 이번 연구에서 후향적으로 추적한 가족력데이터베이스의 한계로 설명하여야 할 것으로 보인다.

가족력 분석에 포함된 연구대상자의 출생시점별 가족력 데이터베이스와의 연계율에서 전반적으로 최근 출생연도 대상자에 비해 과거 출생 대상자들의 가족력 연계율이 낮게 나타났다. 국민건강보험공단 가족력 데이터베이스의 경우 2002년부터 2018년 사이에 생존해 있는 대상자들의 가족력만 집계 가능하므로 2002년 이전의 발생 및 사망자의 경우 집계에 한계가 있다. 그중에서도 조부모 및 외조부의 경우 1980년대생 이전의 가족력부터 급격하게 연계율이 낮아졌고 1970년대, 1960년대생으로 갈수록 조부모와의 연계율이 더욱 급격하게 떨어졌다. 본연구에서 조부모, 외조부모 쪽의 가족력 집계율이 낮게 나타난 점 또한 자료의 한계로 해석될 수 있다. 즉, 1970, 1960년대생 대상자의 조부모의 경우 이미 돌아가셨을 확률이 상대적으로 높다. 이는 가족력데이터베이스 이용 연구의 큰 한계로 노인성 질환의 가족력을 분석하는데 있어서는 제한적으로만 분석이 가능하다. 가족력데이터베이스 연계에 있어 또 하나의 특징은 부모의 연계율에 비해 형제자매의 연계율이 전반적으로 떨어지는 경향을 보이며, 이는 외동인 대상자의 가족력이 정확하게 반영하지 않았기 때문으로 해석된다.

즉 본인이 위암이 발생한 경우 대상자 본인도 이미 나이가 많아 부모의 가족력은

이미 오래전에 발생했을 때, 2000년 이전에 발생까지 추적이 제한되었을 것이므로 고령 대상자의 가족력 영향을 온전하게 반영하지 못했을 것으로 사료되며 이는 노인성 질환에 대한 정확한 가족력 분석이 제한되는 건강보험공단자료의 한계에 기인한다.

2. 간암 질환 예측 연구

간암 질환 예측 연구는 가장 중요한 요소인 나이, B형간염, C형간염 외에도 국가 간암 검진에서 사용되는 간초음파 관찰 소견, AFP와 일반검진의 생활습관(음주, 흡연, 운동)을 이용하여 간암 예측을 진행하였다. 실제적인 간암 위험 예측 모델을 구축하기 위해서는 가장 중요한 인자인 B형간염, C형간염자들의 정교한 선정을 통한 고도화 된 분석이 필요 하겠다. 2015년부터 초음파 결과 입력의 표준화 및 AFP 정량 검사가 진행되고 있다. 추후 연구에서는 이를 활용하여 정교한 결과를 획득할 수 있을 것으로 기대된다.

3. 대사증후군의 유전형과 표현형 융합분석

기존 연구들에서 보인 한계점을 보완하기 위해, 대사증후군과 관련이 있는 표현형 정보 및 각 개인이 가지고 있는 고유의 유전자정보와 생활습관 데이터를 융합 분석해 보았다.

약 3개월간 모집된 대상자를 분석하였기에 짧은 기간에 모집할 수 있는 인원수의 한계로 인하여 249명의 적은 인원으로 분석이 진행 되었으며 통계적 유의성을 입증하는 데 한계점이 있었다. 이에 대해 연구 대상자 모집 기간 및 모집단 수를 늘려 통계적 유의성을 보완하는 연구 모형으로 추후 연구 진행이 필요할 것으로 사료된다.

대사증후군의 세 개의 기준에 따른 연구 이외에도, 5개의 대사증후군 관련인자에 대하여 연속형 변수를 이용한 통계분석이 필요하며, 추후 유전도 위험도 점수의 세분화한 분류를 이용한 분석을 진행할 계획이다. 또한 기준에 분류되어 있는 대사증후군 유전적 위험도 타입 분류 기준에 대해 대사증후군의 유전적 위험군을 잘 선정할 수 있는 고도화 연구가 필요할 것으로 사료된다.

추후 모집단의 추가 및 대사증후군의 유전적 위험도 분류가 더욱 정교하게 수정 및 보완이 진행된다면 생활습관과 유전적 요인의 대사증후군 발생과 관련한 좀 더 고도화된 분석 결과를 기대해 볼 수가 있을 것이다.

4. 위암과 헬리코박터 관련성 분석을 통해 CDM모델의 정합성 검증

CDM data를 활용하여 분석을 진행한 결과 hazard ratio 계산이 가능하였지만, 헬리코 박터 균에 대한 정의 등 관련 여러 Data에 대한 initial event cohort 선정이 중요한

역할을 하는 것을 파악할 수 있었다. 또한 이 cohort 구성을 위한 정의에 맞추어 각 병원별 data의 표준화가 이루어져야 CDM(Common Data Model, 공통데이터모델)의 장점인 여러 의료기관에 흩어져 있는 데이터를 공통된 형태로 연구가 가능하다. 추후 코호트 선정의 표준화 작업이 더욱 고도화 되고, 서로간의 연계를 통한 공유가 필요하다. 이번 CDM 시스템의 구축은 임상데이터의 표준화 및 연계 작업의 중요한 시발점이 되며, 추후 여러 공공자료와의 연계를 통하여 연구 역량 강화의 새로운 도약의 계기가 될 것이다.

참고문헌

참고문헌

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136: E359-86.
2. Korea Central Cancer Registry, National Cancer Center. Annual report of cancer statistics in Korea in 2014, Ministry of Health and Welfare, 2016
3. "Stomach Cancer Survival Rates." American Cancer Society, 18 Dec. 2017, www.cancer.org/cancer/stomach-cancer/detection-diagnosis-staging/survival-rates.html.
4. Pandalai PK, Yoon SS. Hereditary diffuse gastric cancer. *Principles of Clinical Cancer Genetics*: Springer, 2010:97-107.
5. McLean MH, El-Omar EM. Genetics of gastric cancer. *Nat Rev Gastroenterol Hepatol* 2014;11: 664-74.
6. Charvat, H., Sasazuki, S., Inoue, M., Iwasaki, M., Sawada, N., Shimazu, T., ... & JPHC Study Group. (2016). Prediction of the 10-year probability of gastric cancer occurrence in the Japanese population: the JPHC study cohort II. *International journal of cancer*, 138(2), 320-331.
7. Song, H., Ekhedden, I. G., Ploner, A., Ericsson, J., Nyren, O., & Ye, W. (2017). Family history of gastric mucosal abnormality and the risk of gastric cancer: a population-based observational study. *International journal of epidemiology*, 47(2), 440-449.
8. Wynder, E. L., & Gori, G. B. (Eds.). (1977). *Contribution of the environment to cancer incidence: an epidemiologic exercise*.
9. Doll, R., & Peto, R. (1981). The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *JNCI: Journal of the National Cancer Institute*, 66(6), 1192-1308.
10. Haenszel, W., & Kurihara, M. (1968). Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *Journal of the National Cancer Institute*, 40(1), 43-68.

11. 국가암등록통계 2016 <https://ncc.re.kr/cancerStatsList.ncc?sea>
12. Lau WY, Leow CK, Li AK. Hepatocellular carcinoma [Review]. Br J Hosp Med 1997;57:101-104.
13. Haydon GH, Hayes PC. Screening for hepatocellular carcinoma [Review]. Eur J Gastroenterol Hepatol 1996 ;8:856-860.
14. 안상훈, 한광협, 윤영훈, 홍성필, 백용한, 전재윤, 등. 한국인에서 간암발생의 위험요인 분석. 대한내과학회지 2000;60:123-130.
15. 정재연, 한광협, 김동기, 안상훈, 송기준, 백용한, 등. 한국인에서 간세포암종 발생의 위험요인 및 개인별 간세포암종발생 예측모형. 대한간학회지 2001;7:449-458.
16. Hsu et al J. Hepatol. 2018, 69, 278-285.
17. Suraj et al .Journal of Hepatology 2018 vol. 68 92-99.
18. Fleming et al. Cancer 2014;120:3485-93
19. Beasley RP. Hepatitis B-virus: the major etiology of hepatocellular carcinoma. Cancer 1988;61:1942-1956.
20. Di Bisceglie AM. Hepatitis C and hepatocellular carcinoma. Semin Liver Dis 1995;15:64-69.
21. Reaven GM, et al. Role of insulin resistance in human disease. Diabetes 1988;37(12):1595-1607.
22. A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. 2018 Guideline on the Management of Blood Cholesterol. Journal of the American College of Cardiology 2018(11);1-121.
- 23.ustin Xavier Moore, MPH, et al. Metabolic Syndrome Prevalence by Race/Ethnicity and Sex in the United States, National Health and Nutrition Examination Survey, 1988-2012. PREVENTING CHRONIC DISEASE PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY 2017;14(3):1-16
24. 질병관리본부 국가건강정보포털, 건강/질병검색 [Cited September 17, 2019]; Available from: <http://health.cdc.go.kr/health/HealthInfoArea/HealthInfo/View.do?idx=490>
25. 서울대학교병원 건강정보, N의학정보 [Cited September 17, 2019]; Available from: <http://www.snuh.org/health/nMedInfo/nView.do>

26. National Institutes of Health USA, What is insulin resistance? [Cited September 17, 2019]; Available from: <https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes/prediabetes-insulin-resistance>
27. National Education Cholesterol Program Adult Treatment Panel III(NCEP ATP III). 2001.
28. Ford, Earl S., Wayne H. Giles, and William H. Dietz. "Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey." *Jama* 287.3 (2002): 356-359.
29. Markku Laakso, et al. The Metabolic Syndrome in Men study: a resource for studies of metabolic and cardiovascular diseases. *Journal of Lipid Research*. 2017;58(1):481-493.
30. Kohro T, Furui Y, Mitsutake N, et al. The Japanese national health screening and intervention program aimed at preventing worsening of the metabolic syndrome. *Int Heart J*. 2008;49(2):193-203.
31. Park SY, Yang YJ, Kim Y. Effects of nutrition education using a ubiquitous healthcare (u-Health) service on metabolic syndrome in male workers. *Korean J Nutr*. 2011;44(3):231-242.
32. Ingram DK, Nakamura E, Smucny D, Roth GS, Lane MA. Strategy for identifying biomarkers of aging in long-lived species. *Exp Gerontol*. 2001;36(7):1025-1034.
33. Borkan GA, Norris AH. Assessment of biological age using a profile of physical parameters. *J Gerontol*. 1980;35(2):177-184.
34. Ming-Sheng Teng, et al. Association of CDH13 Genotypes/Haplotypes with Circulating Adiponectin Levels, Metabolic Syndrome, and Related Metabolic Phenotypes: The Role of the Suppression Effect. *PLOS ONE*. 2015(4):1-13.
35. 김현주, et al. ENPP1 K121Q 유전자 다형성과 대사증후군과의 관련성. *Korean J Fam Med*. 2010;31:913-922.
36. Sanghoon Moon. Multiple genotype-phenotype association study reveals intronic variant pair on SIRT2 associated with metabolic syndrome in a Korean population. *Human Genomics*. 2018(11):1-10.

37. ATP III의 진단기준에 따른 대사증후군의 유병률과 관련위험요인 분석-일개 대학병원 강증진센터 수진자를 대상으로. 가정의학회지, 2003.
38. Philipp Mitteroecker. Multivariate Analysis of Genotype-Phenotype Association. *Genetics*. 2016;202(4):1345-1363
39. Bae CY, Kang YG, Suh YS, Han JH, Kim SS, Shim KW. A model for estimating body shape biological age based on clinical parameters associated with body composition. *Clin Interv Aging*. 2013;8:11-18.
40. Bae CY, Kang YG, Piao MH, et al. Models for estimating the biological age of five organs using clinical biomarkers that are commonly measured in clinical practice settings. *Maturitas*. 2013;75(3):253-260.
41. Bae CY, Kang YG, Kim S, et al. Development of models for predicting biological age (BA) with physical, biochemical, and hormonal parameters. *Arch Gerontol Geriatr*. 2008;47(2):253-265.
42. "Stomach Cancer Survival Rates." American Cancer Society, 18 Dec. 2017, www.cancer.org/cancer/stomach-cancer/detection-diagnosis-staging/survival-rates.html.
43. Zhensheng Wang et al. Composite protective lifestyle factors and risk of developing gastric adenocarcinoma: the Singapore Chinese Health Study. *British Journal of Cancer* (2017) 116, 679-687.
44. Teng-Yu Lee et al. A Tool to Predict Risk for Gastric Cancer in Patients With Peptic Ulcer Disease on the Basis of a Nationwide Cohort. *Clinical Gastroenterology and Hepatology* 2015;13:287-293.
45. 이동숙 (2008). "<위암발병_예측모형.pdf>." *Asian Oncology Nursing* 8(부록1호).
46. Yoo Mi Park et al. Clinical risk assessment for gastric cancer in asymptomatic population after a health check-up: An individualized consideration of the risk factors. *Medicine* (2016) 95:44(e5351).
47. Bang Woll Eom et al. Prediction Model for Gastric Cancer Incidence in Korean Population. *PLOS ONE*. 10(7): e0132613.

연구보고서 2019-20-011

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발 (6차년도)

발행일	2020년 1월 15일
발행인	김성우
편집인	최윤정
발행처	국민건강보험공단 일산병원 연구소
주소	경기도 고양시 일산동구 일산로 100
전화	031) 900-6982~6987
팩스	031) 900-6999
인쇄처	백석기획(031-903-9979)



국민건강보험 일산병원 연구소



(우)10444 경기도 고양시 일산동구 일산로 100 (백석1동 1232번지)
대표전화 031-900-0114 / 팩스 031-900-6999
www.nhimc.or.kr

2019 NHIS Ilsan Hospital Institute of Health Insurance & Clinical Research

N a t i o n a l H e a l t h I n s u r a n c e S e r v i c e I l s a n H o s p i t a l