

발 간 등 록 번 호

G000EV1-2022-116

생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역

연 구 진

연구책임자 이태진 교수

연 구 원 김새롬 연구교수

문주현 박사과정

주제어 가명정보결합, 노동패널, 청구자료, 생애주기, 의료이용



SNU Health

서울대학교 보건대학원



건강보험심사평가원

HEALTH INSURANCE REVIEW & ASSESSMENT SERVICE

제출문

건강보험심사평가원장 귀하

이 보고서를 '생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역' 과제의 최종보고서로 제출합니다.

2022. 12.

연구책임자: 서울대학교 보건대학원 교 수 이태진

연 구 원: 서울대학교 보건대학원 연구교수 김새롬

서울대학교 보건대학원 박사과정 문주현

※ 본 연구는 국민건강보험공단의 자료(연구관리번호 NHIS-2022-9-004)를 활용했습니다. 보고서는 위탁연구 과제를 수행한 연구자의 개인적인 의견이므로 건강보험심사평가원 및 국민건강보험공단의 공식적인 견해와는 다를 수 있습니다.



제1장 서론	1
---------------	----------

1. 연구 배경과 목적	3
가. 연구 배경	3
나. 연구 목적	6
2. 연구 내용과 방법	7
가. 연구 내용	7
나. 연구 방법	11

제2장 선행문헌 고찰	13
--------------------	-----------

1. 이론적 배경	15
가. 생애주기 개념의 발전	15
나. 생애주기 개념의 활용	17
2. 인구집단의 건강과 의료이용에 관한 국내연구	19
가. 청구자료 및 임상자료를 활용한 연구	19
나. 노동패널자료를 활용한 연구	21
3. 한국의 건강 및 의료이용	24
가. 건강보험 빅데이터	24
나. 연도별, 연령별, 성별 질병 및 사망 현황	26

제3장 가명정보결합 분석결과	31
------------------------	-----------

1. 가명정보결합 자료의 구조	32
가. 한국노동패널조사	32
나. 건강보험 빅데이터	34

다. 가명정보결합 자료	34
2. 가명정보결합 자료의 특성	36
가. 결합대상자의 인구학적 특성	36
나. 연도별 일반적 특성	38
다. 건강 및 의료이용 관련 특성	47
라. 의료비 지출	52
3. 생애주기별 인구집단 특성에 따른 의료이용과 의료비 지출	55

제4장 결론	61
---------------	-----------

1. 연구결과	63
2. 연구의의	65
3. 정책적 제언	66

■ 참고 문헌	69
---------------	----

표목차

〈표 1〉 2019-2020 OECD대비 한국의 주요 건강지표 현황	3
〈표 2〉 2019-2020 가구유형별 의료이용 현황	4
〈표 3〉 성연령별 사망자 수 현황(2010년, 2015년, 2020년)	29
〈표 4〉 자료원별 결합률과 개체수	35
〈표 5〉 가명정보결합자료의 주요 변수	35
〈표 6〉 2011 - 2015 결합대상자의 일반적 특성	38
〈표 7〉 2016 - 2020 결합대상자의 일반적 특성	41
〈표 8〉 2011 - 2020 결합자료의 노동시간 특성	44
〈표 9〉 2011 - 2020 결합자료의 가구 근로소득 특성	44
〈표 10〉 2011 - 2020 결합자료의 연간 가구 총 근로소득 10분위 경계값	45
〈표 11〉 2011 - 2020 결합자료의 건강보험료 특성	45
〈표 12〉 2011 - 2020 결합자료의 보건의료비 지출 특성	46
〈표 13〉 2011 - 2020 노인장기요양보험등급 신청 및 인정 현황	48
〈표 14〉 2011 - 2020 집계방식별 사망자 수 비교	48
〈표 15〉 연령대별 2011년, 2020년 외래의료서비스 이용 시 다빈도 주상병 Top 5	49
〈표 16〉 연령대별 2011년, 2020년 입원의료서비스 이용 시 다빈도 주상병 Top 5	50
〈표 17〉 연도별 의료서비스 유형별 이용 여부	51
〈표 18〉 연도별 종사상 지위에 따른 외래서비스 이용 비율 추이	52
〈표 19〉 연도별 종사상 지위에 따른 입원서비스 이용 비율 추이	52
〈표 20〉 2018-2020 결합자료의 의료이용 유형별 이용여부 영향요인	55
〈표 21〉 2018-2020 여성의 의료이용 유형별 이용여부 영향요인	56
〈표 22〉 2018-2020 남성의 의료이용 유형별 이용여부 영향요인	57
〈표 23〉 2018-2020 결합자료의 의료이용 유형별 요양급여총액 영향요인	59

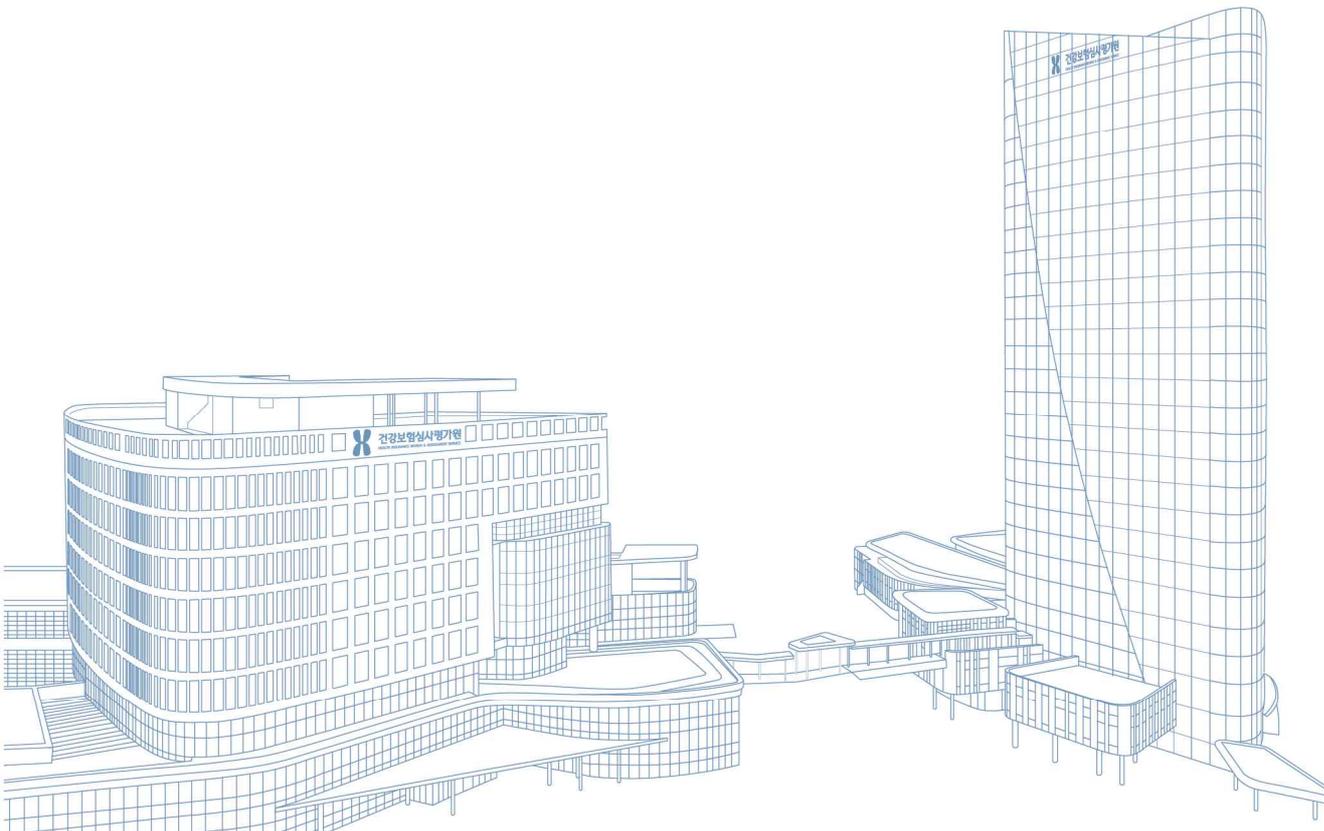



그림목차

[그림 1] 가명정보 결합 절차	9
[그림 2] 가명정보결합 참여 기관과 역할	10
[그림 3] 자료 분석 절차	11
[그림 4] Life course view of noncommunicable disease (NCD) risk	16
[그림 5] 결합 대상 건강보험 빅데이터	25
[그림 6] 2020년 30대 환자수 및 요양급여총액에 따른 다빈도 질환 순위	26
[그림 7] 2020년 60대 환자수 및 요양급여총액에 따른 다빈도 질환 순위	27
[그림 8] 2020년 연령대별 입원 및 외래 환자수에 따른 다빈도 질환 순위와 유형	28
[그림 9] 2020년 연령대별 주요 사망원인 순위	30
[그림 10] 결합자료와 노동패널조사의 2011년과 2020년 연령대 분포	37
[그림 11] 결합자료의 연령대별 분포	37
[그림 12] 자료원별·연도별 건강보험료 및 보건의료비 추이 (단위: 만원)	47
[그림 13] 자료원별·연도별 건강보장 유형	47
[그림 14] 연도별·연령대별 전체 의료이용에 따른 요양급여총액 (단위: 만 원)	53
[그림 15] 연도별·연령대별 외래서비스 이용에 따른 요양급여총액 (단위: 만 원)	53
[그림 16] 연도별·연령대별 입원서비스 이용에 따른 요양급여총액 (단위: 만 원)	54
[그림 17] 연도별·종사상지위별 전체 요양급여총액 (단위: 만 원)	54



요약



요약

제1장 서론

- 인구고령화와 평균적인 건강수준 향상과 함께 지난 20여 년간 국내 의료이용의 양적, 재정적 규모는 꾸준히 증가하면서 한국도 질병예방 및 건강증진을 강조하고 인구집단 내 건강격차를 완화하기 위해 생애주기 접근에 기반한 보건의료정책과 개입전략을 추진하고 있음
- 하지만 정책의 개발과 평가를 위해 중재연구에 준하는 무작위대조실험이나 추적조사를 매년 수행하기란 불가능하고 기존 자료를 활용한 연구가 수행한 경우 개별 자료원이 갖는 표본추출, 변수구성, 정보관리 방법이 정책목적에 부합하지 않는 경우 분석은 다소 제한적임
- 본 연구는 2022년 제2기 가명정보 결합 선도사례의 '건강(의료형평성) 분야'로 선정되어 가명정보결합 기법을 활용해 공익적 목적으로 수집 및 관리되던 세 가지 자료를 결합하고 생애주기별 인구집단의 특성과 경험에 따라 의료이용이 달라지는지 파악하는 것을 목적으로 함

제2장 선행문헌 고찰

- 생애주기 접근은 태아가 성장하는 주산기 발달부터 사망에 이르기까지 인구집단 건강이 영양, 교육, 결혼, 직업, 거주환경 등 사회적 결정요인에 영향을 받는다는 것을 고려한 개인의 전 생애과정에서 경험하는 건강행태, 외상, 기능저하에 따라 건강수준이 달라진다는 것을 의미함
- 특정 생애주기 인간의 의사결정이나 과업달성 여부에 따른 행동변화를 설명하는 연구가 활발해지면서 노인인구의 노화를 질병이 아닌 발달과정의 일부로 취급하는 사회적 인식이 보편화되고, 사회규범과 생애경험의 상호작용이 개인

의 신체적, 정신적, 사회적 건강에 미친 영향에 대한 이론이 발전함

- 건강보험 청구자료나 의료패널을 활용해 생애주기별 인구집단의 건강을 분석한 연구는 다수 진행되었지만 사회경제적 요인이나 생애경험에 관련한 변수의 부족으로 인해 특정시기 다빈도 질환의 유병률 및 사망률, 그리고 의료비 지출에 관한 분석에 집중되어 있음
- 한국노동패널조사 자료를 활용한 중고령자 대상 연구는 20년간 청년 및 여성 대상 연구에 비해 빈번하게 연구되었으며 자료원의 한계를 보정하기 위해 부가조사와 국가통계자료를 결합한 시도가 많았지만 건강과 의료이용에 관한 연구는 125건 중 6건이었음
- 한국노동패널 조사 자료를 활용한 연구 중 여성에 관한 150건의 연구 중에서 건강에 관한 연구는 단 3건이며, 결혼과 출산에 관한 연구는 27건으로 의료비나 건강보험 가입에 관한 연구는 제한적이었음
- 국내외에선 건강의 사회적 결정요인에 따른 건강불평등을 규명하고 건강불평등을 완화하기 위한 다양한 연구, 정책, 사업을 설계 및 평가하는 방향으로 확장하고 있음

제3장 가명정보결합과 분석결과

- 가명정보 결합 자료는 한국노동연구원, 건강보험심사평가원, 국민건강보험공단에서 제공받음
 - 한국노동연구원이 생성 및 관리하는 한국노동패널조사 자료는 가구와 개인의 노동시장 특성을 조사하는 것이 목적이며, 고용 및 노동 특성뿐만 아니라 개인의 인구학적 특성, 교육수준과 혼인상태, 소비지출 행태를 매년 조사하고, 그 외 시기별, 주제별 부가조사를 비정기적으로 시행해 사회적 결정요인에 대한 포괄적 분석이 가능한 자료임
 - 건강보험심사평가원은 환자의 상병명, 일반정보, 처치 및 수술료의 진료내역, 청구비용, 의약품, 요양기관 정보 등 모든 의료이용 건수별 정보를 수집 및

- 관리하는 기관임. 의료이용 내역이 담긴 청구자료는 비급여 서비스를 제외한 건강보험가입자 및 의료급여 수급권자의 모든 의료이용 정보를 포함하고 있음
- 국민건강보험공단은 건강보험 가입자의 자격과 보험료 징수를 관리하고 국민 건강보험과 노인 장기요양보험의 요양급여비용을 지급하는 기관으로 대상자를 식별할 수 없는 방식으로 가공한 정보(자격정보, 보험료분위, 건강검진 및 노인 장기요양보험 관련 이용 및 평가 정보)를 공익적 목적의 연구에 제공하고 있음
 - 가명정보 결합은 2011년부터 2020년까지 건강보험심사평가원 56,870,039개, 국민건강보험공단 37,268,095개, 한국노동연구원 34,840개의 레코드를 Key로 변환하여 결합함
 - 결합결과, 한국노동연구원 자료에서 총 21,162개(결합률 60.74%)로 건강보험심사평가원 19,881개, 국민건강보험공단 16,591개의 레코드가 결합되어 3개 기관에 모두 일치하는 레코드수는 최종 15,310개(Key)임
 - 가명정보 결합자료에서 직업력 변화로 인해 중복 생성된 경우, 건강보험 자격 정보가 결측인 경우, 사망일 이후 정보가 남아있는 경우를 제거하고 2011년부터 2020년까지의 불균형 패널을 생성함
 - 결합자료의 52.4%는 남성, 65세 이상 고령자는 19.9%이며, 고졸자 31.1%이며, 종사상지위별로 미취업자 38.1%가 가장 많았고, 수도권 거주자(서울, 경기, 인천)는 51.4%, 1인 가구는 23.9%임
 - 경제활동참여자의 주당 근로시간은 평균 40.7시간에서 36.5시간(표준편차 15.2~18.9)으로, 초과근로시간은 9.2시간에서 6.9시간(표준편차 4.8~7.7)으로 감소함
 - 10세 단위 연령대별 외래와 입원 의료서비스의 다빈도 상병분석결과, 외래의료서비스의 경우 50대 이하 청장년층은 주로 치과 관련 질환과 임신부 산전관리였고, 50대 이상 중고령층은 등통증(M54), 요추탈구(S33), 척추병증(M48), 무릎관절증(M17)처럼 근골격계 질환이나 고혈압(I10), 당뇨(E11), 만성신장병(N18)이 다빈도 상병이었음

- 입원 의료서비스의 다빈도 상병분석 결과, 2011년 대비 2020년에 암질환 관련 신생물(구강암(C05), 위암(C16), 결장암(C18), 직장암(C20), 폐암(C34), 유방암(C50))이 여러 연령대에서 확인됨
 - 외래 의료서비스와 같이 20, 30대에서는 자연분만(O80)과 제왕절개(O82)가 있었고, 중고령층인 60대와 70대의 경우 백내장(H25), 80대 이상은 치매(F00)가 주로 확인됨
- 개인별 연간 외래 및 입원의료서비스의 평균 이용률은 87.9%, 13.3%이며, 종사상 지위에 따른 비율은 미취업자(33.6%, 6.4%), 상용직 근로자(30.0%, 3.3%), 자영업자(12.3%, 1.9%), 임시일용직 근로자(9.7%, 1.3%), 무급가족 종사자(2.4%, 0.4%) 순임
- 10년간 평균 요양급여 총액은 모든 연령대에서 증가하였으며, 40대 이하보다 40대 이후로 갈수록 비용의 증가폭이 가파르게 증가함
 - 연평균 증가율은 20대가 8.0%, 70대 7.7%, 80대 이상 7.4% 순임
- 2018년부터 2020년까지 3개 년도의 자료로 의료서비스 유형별 이용여부와 인구집단의 특성에 따른 차이가 있는지 검토하기 위해 로지스틱 고정효과 패널회귀 분석을 실시함
 - 전체 의료이용과 외래 또는 입원의료서비스는 65세 이상이 64세 이하보다 전체 의료서비스를 이용할 가능성이 1.42배, 1.38배 높고, 교육수준에 따른 차이는 입원 의료서비스 이용에서 3.11배 높았음
 - 종사상지위의 경우 미취업자보다 무급가족종사자와 자영업자 전체 의료서비스를 적게 이용하고 있고 상근직과 임시일용직은 1.14배, 1.09배 많이 이용함
- 의료서비스 유형별 의료이용과 요양급여 총액에 미치는 영향요인을 살펴보기 위해 일반화 선형회귀분석을 수행함
 - 외래 서비스의 경우 요양급여총액이 남성보다 여성이 0.9% 더 높았지만 입원 서비스는 남성보다 여성이 1.1% 낮아 성별 의료비 부담에 차이가 있음
 - 65세 이상 고령인구는 64세 이하보다 전체의료서비스에 대한 요양급여총액이 5.4% 높았음(외래 4.8%, 입원 2.7% 높음, $p < 0.0001$)

- 교육수준이 낮은 집단보다 높은 집단에서 의료비 지출이 1.0~2.0% 더 낮았고 미혼보다 기혼, 이혼·별거·사별의 경우 의료비 지출은 입원 1.0%, 전체 4.2% 높았음($p < 0.01$)
- 노인이나 여성이 많이 포함된 미취업자보다 상근직과 임시일용직, 자영업자가 전체 의료비 지출이 1.8%, 0.8% 0.7% 더 낮았고 외래예선 상근직(12%)과 임시일용직(0.4%), 입원에선 상근직(0.7%)과 자영업자(0.7%)의 의료비 지출이 미취업자보다 적었고 이는 통계적으로 유의함($p < 0.01$)

제4장 결론

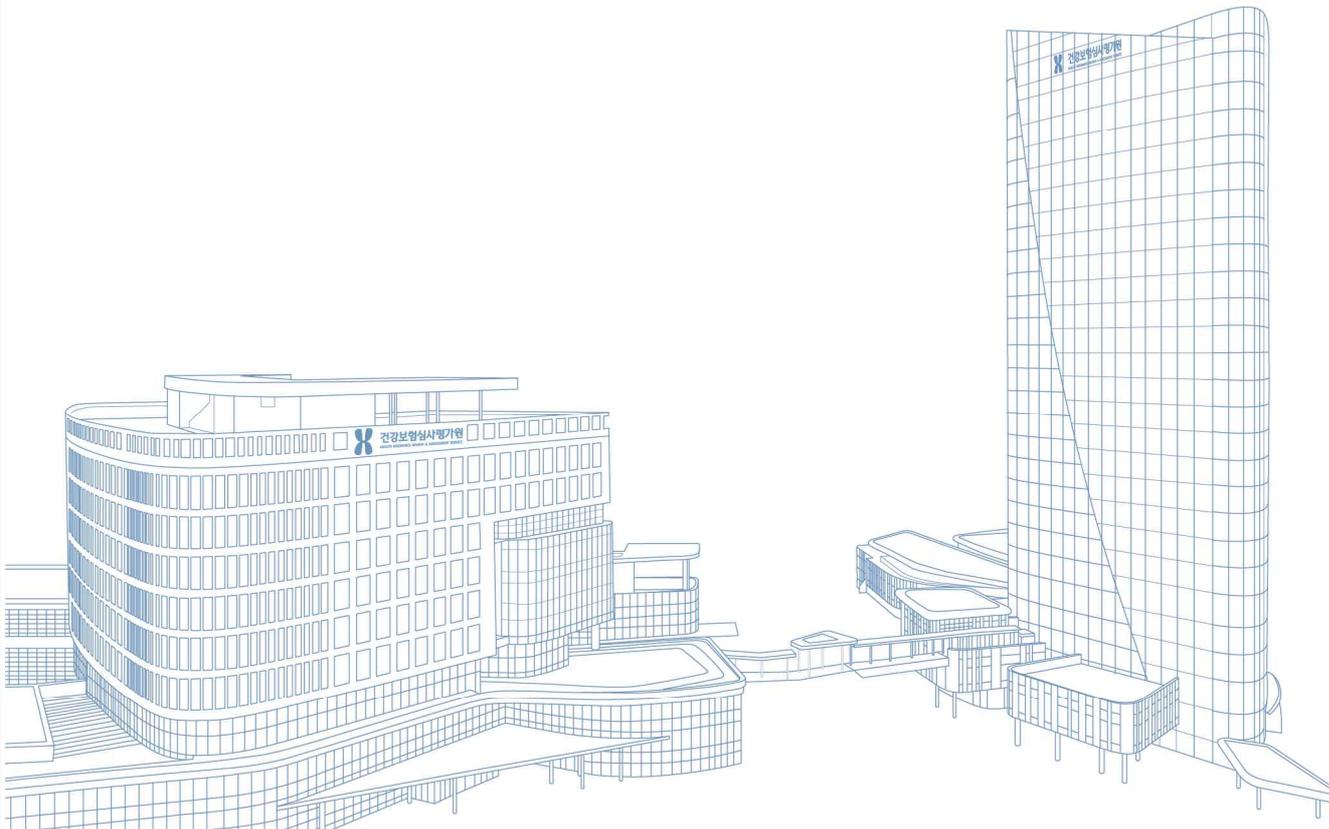
- 본 연구에선 결합과정과 결합된 자료의 특성을 기술하고 생애주기 인구집단의 특성에 따른 의료이용을 파악하는 것을 목표로 함
 - 연도별 연령대별 다빈도 상병 5순위에서 전 연령대에서 근골격계 질환이 많았고 이는 건강보험심사평가원에서 제공하는 다빈도 상병통계와 유사함
 - 근골격계 질환 이외의 다빈도 상병은 소수 인원이 특정상병으로 의료서비스를 반복적으로 이용하는 양상이 나타남
 - 인구학적 특성에 따른 의료비 지출은 교육수준이 높을수록 의료비 지출이 적었고 미혼인 경우에 비해 기혼, 이혼·별거·사별일 때 의료비 지출이 더 많고 종사상지위에 따라 미취업자에 비해 취업자의 의료비지출이 더 적은 것으로 나타났지만 여러 특성이 교차하면서 발생하는 효과는 분석되지 못함
- 개인과 시간의 특성을 통제한 고정효과 분석결과, 한국인의 전반적인 의료이용 접근성에 비해 의료비 부담은 증가하였고 영향요인과 그 효과크기는 분석대상 인구 전체와 성별 층화분석과의 차이가 있었음
 - 의료서비스의 이용과 요양급여 총액은 성별, 연령, 종사상지위에 따른 차이가 있었으나 건강상태의 변화에 따른 개별 대상자의 의료적 필요는 고려하지 못한 한계가 있음
- 결합자료의 인구학적 특성과 의료이용 현황을 국가통계 및 공공데이터와 비교해

가명정보결합자료의 선도 사례로서 단일자료로 분석할 수 없는 사회경제적 요소에 따른 의료이용 실태를 분석하여 가명정보결합을 활용한 공공데이터 연계가 기존 자료원의 한계를 보완할 수 있음을 확인함에 그 의의가 있음

○ 그리고 가명정보 결합자료를 활용한 연구는 자료제공 및 결합과정에 소요되기 때문에 시간적 제약이 상당하므로 향후 가명정보 결합 연구를 추진하기 위한 정책적 제언은 다음과 같음

- 빅데이터 기반 연구에서는 주목받지 못했던 의료이용의 원인들을 폭넓게 탐색할 수 있는 가명정보결합자료를 활용한 정책연구가 활성화 되어야 함
- 생애주기와 생애사건을 경험하는 인구집단의 서로 다른 특성이 교차하면서 발생할 수 있는 건강 및 의료이용 격차에 대한 심층분석이 필요함
- 자료제공기관에서는 연구윤리를 준수한 자료제공과 원활할 처리를 위하여 기관별 내규와 업무분장의 명확화, 담당자의 숙련도 향상을 위한 내외적 인프라 구축이 필요함
- 자료제공기관의 자체적인 규정과 관계기관간의 법률 조항의 해석을 위해 자료원을 검증하는 기준을 일치시키기 위하여 학계와 공공기관의 협의가 필요함
- 가명정보결합을 활용한 자료의 접근, 연계, 활용과정에서 투명성과 효율성을 개선하여야 함
- 보건의료정책 시행과정에서 수집·관리하는 자료에서 관찰할 수 없는 정보를 확보함으로써 일반인구의 건강 관련 행태 및 의사결정, 이외 사회·경제·문화적 특성들을 포함한 보건학적 연구의 활용도가 높아 질것으로 기대됨

서론



제1장

서론

1. 연구 배경과 목적

가. 연구 배경

1) 한국 인구집단의 건강과 의료이용

- 인구고령화와 평균적인 건강수준 향상과 함께 지난 20여 년간 국내 의료이용의 양적, 재정적 규모는 꾸준히 증가함
- 2019년 기준 한국의 기대수명은 83.3세이고 회피가능한 사망률은 인구 10만 명 당 144명으로 OECD 평균에 비해 낮지만, 자살률은 인구 10만 명 당 24.6명으로 OECD 평균인 11.3명의 2배 이상으로 질병과 상해로 인한 사망 외에 건강에 영향을 미치는 사회적 결정요인에 관한 관심이 증가하고 있음
- 한국의 연간 1인당 외래진료 횟수는 17.2회, 평균 재원일수는 18.0일로 각각 OECD 평균의 3.7배, 2.25배 이상이고 연간 1인당 의약품 판매액도 690.9 US\$(PPP)로 OECD 평균을 상회하고 있음

〈표 1〉 2019-2020 OECD대비 한국의 주요 건강지표 현황

2019년 주요 지표	한국	OECD 평균
기대수명	83.3세	81.0세
회피가능사망률 (인구 10만명 당)	144명	199.7명
자살률 (인구 10만명 당)	24.6명	11.3명
연간 1인당 외래진료 횟수	17.2회	4.6회
연간 1인당 평균 재원일수	18.0일	8.0일
연간 1인당 의약품 판매액 (US\$ PPP)	690.9	532.9

* 자료: OECD. (2021); 통계청. (2021)

- 2020년 1인 가구는 전체의 31.7%로 가구소득은 전체가구의 가구소득 대비 1.2이하로 적으나 연간 의료비 지출은 전체가구보다 1.3배 이상 높음
- 가구유형에 따른 의료이용의 차이는 동일한 연령집단 내에서도 돌봄 필요나 부양부담에 영향을 미치는 인구학적 특성에 따라 건강격차가 발생하고 있음을 시사함

〈표 2〉 2019-2020 가구유형별 의료이용 현황

2020년 가구유형별 의료이용 현황	1인 가구	전체 가구
	6,643가구 (31.7%)	20,927 천 가구
가구소득	2,162만 원	5,924만 원
연간 의료비	95만 5천 원	68만5천 원
정기 건강검진 실천율	76.4%	80.7%

* 자료: 통계청. (2021)

2) 한국 보건의료정책의 생애주기별 접근

- 질병예방 및 건강증진을 강조하고 인구집단 내 건강격차를 완화하기 위해 국내외에선 특정 질병을 퇴치하거나 치료하는 수직적 접근 대신 생애주기 접근에 기반한 보건의료정책이나 개입전략을 추진하고 있음(Chandra, & Vogl, 2010)
- 한국 정부가 발표한 제1차 국민건강보험종합계획(2019~2023)은 계층과 세대에 구분 없는 건강보장의 형평성 확보 및 출생부터 노년까지 전 생애에 걸친 건강보장 구현을 비전으로 제시함
- 제5차 국민건강증진종합계획(Health Plan 2030, 2021~2030, 이하 HP2030)에서도 성, 계층, 지역 간 건강 형평성을 확보하고 출생부터 노년까지 전 생애주기에 걸친 건강권을 보장하는 것을 추구하고 있음

- 인구집단의 특징과 의료요구를 파악하고 중재방안을 수립하는 과정에선 이미 국가승인통계, 행정자료, 설문조사 등을 활용한 정책연구가 활발히 진행됨
- 정책의 개발과 평가를 위해 중재연구(implementation science)에 준하는 무작위대조실험이나 추적조사를 매년 수행하기 어렵기 때문에 기존 자료원의 활용이 필요함
- 그러나 기존 자료원을 활용한 연구는 개별 자료원이 갖는 표본추출, 변수구성, 정보관리 방법이 정책목적에 부합하지 않는 경우 분석은 다소 제한적일 수 있음
- 또한 기존 자료만으로 분석을 시도할 수 없을 수 있고 이론적 기반과 국외사례 고찰에 근거한 분석을 하더라도 분석결과의 엄밀성과 실효성을 입증하기 어렵다면 정책적 대안을 마련하고 우선순위를 설정하는 근거기반 정책형성을 실현하기 어려움(유중성, 전병유, 신광영, 이도훈, & 최성수, 2020)
- 건강보험심사평가원과 국민건강보험공단의 행정자료는 건강보험가입자와 의료급여수급권자를 대상으로 하는 전수자료로 의료서비스에 관련된 정보는 방대하지만 임상 의무기록과 사회경제적 특성에 대한 정보가 배제되어 건강격차, 건강불평등에 대한 포괄적 분석이 어려움(정현수, 김수영, & 김현수, 2019; Ryu, D. R, 2017)

3) 코로나19와 가명정보 결합

- 2018년 처음 발의됐던 데이터 3법¹⁾이 2020년 8월 코로나19 감염병 위기에 대응하는 과정에서 시행되면서 가명정보라는 개념이 도입됨
- 가명정보란 추가정보가 없으면 특정 개인을 알아볼 수 없게 보안조치한 정보로서 상업적 목적을 포함한 통계작성과 연구, 공익적 기록보전을 목적으로 하는 경우 개인정보를 제공했거나 수집에 동의한 당사자에게 추가적인 동의를 구하지 않더라도 활용이 가능함

1) 데이터 3법은 「개인정보 보호법」, 「정보통신망 이용촉진 및 정보보호 등에 관한 법률(약칭 : 정보통신망법)」, 「신용정보의 이용 및 보호에 관한 법률(약칭 : 신용정보법)」을 의미함

- 이후 개인정보에서 가명정보로의 변환과 결합을 지원하고 가명정보 처리한 자료의 활용을 증진하기 위한 법적 근거와 지침이 정비되면서 보건복지부 산하 가명정보결합기관(국민건강보험공단, 건강보험심사평가원, 국립암센터)을 운영하고 있음
- 보건의료 분야 자료를 가명정보로 결합해 정책연구에 활용한다면 코로나19를 비롯한 감염병 대비 및 대응에서 더하여 서비스제공의 효율성을 향상하고 건강형평성을 강화하기 위한 보건의료체계의 정책적 의사결정에 기여할 것으로 기대함

나. 연구 목적

- 본 연구는 2022년 제2기 가명정보 결합 건강(의료형평성) 분야의 선도사례²⁾로 선정된 과제이며, 연구목적은 가명정보결합 기법을 활용해 공익적 목적으로 수집 및 관리되던 세 가지 자료원을 결합하고 생애주기별 인구집단의 특성과 경험에 따라 의료이용이 달라지는지 파악하는 것임
- 가명정보결합자료를 분석한 결과는 각 자료원의 품질향상을 위한 기초자료로 활용되어 공익적 목적의 정책연구를 수행하는 공공기관 간 협업과 연구의 질적 향상을 도모하는데 활용될 수 있을 것으로 기대함
- 한국노동패널조사 자료와 건강보험 자격 및 청구자료, 건강검진사업 및 노인 장기 요양보험제도 자료를 결합하고 분석한 과정에서 정책적 함의를 도출하여 생애주기 접근에 기반한 보건의료 빅데이터 활용이 정책적 의사결정에 기여하기 위한 고려사항과 후속과제를 검토함

2) 개인정보위원회는 2021년 5대 분야 31개 유형의 자료에 대한 가명정보결합 1기 시범사례를 발굴하였고, 2022년에는 사회(소득불평등), 복지(장애인복지), 환경(저탄소경제), 건강(의료형평성) 분야의 2기 선도사례를 선정하여 가명정보결합을 지원함(대한민국 정책브리핑, 2022)

2. 연구 내용과 방법

가. 연구 내용

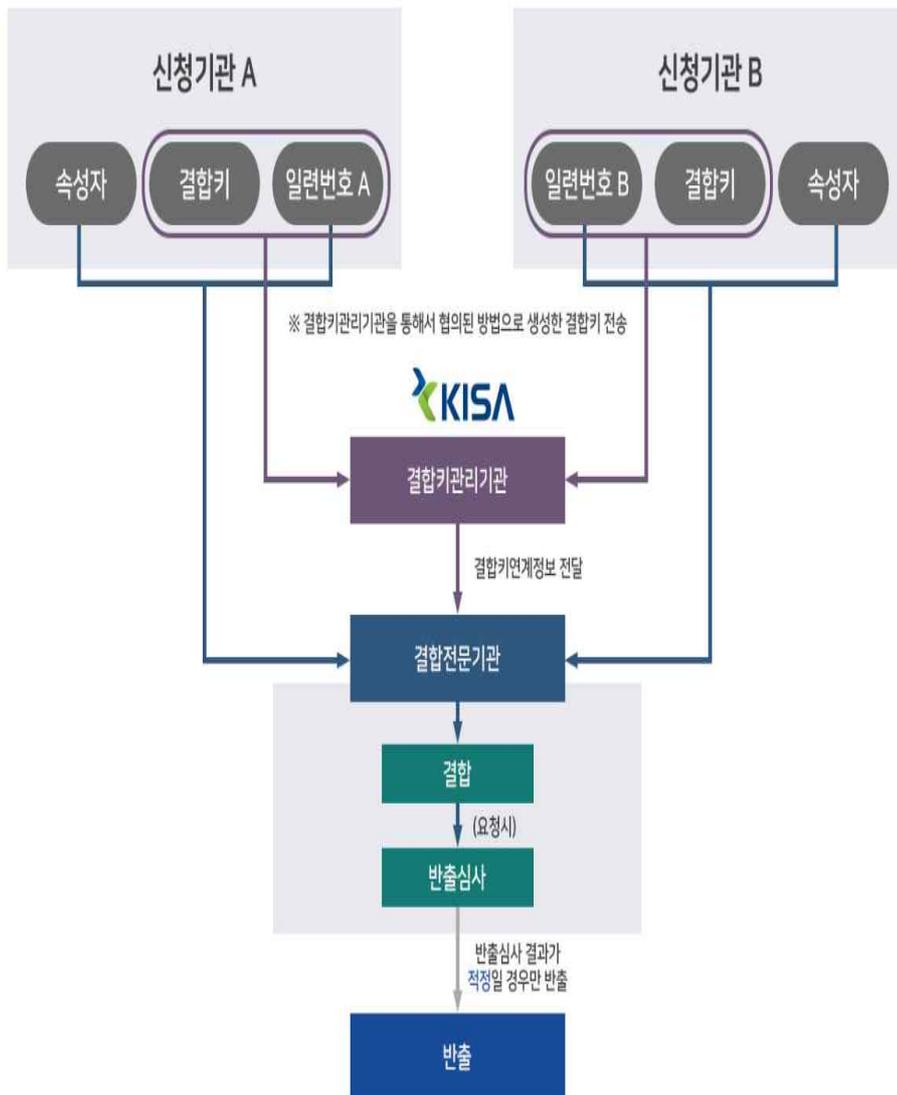
1) 생애주기 접근의 이론과 활용 검토

- 생애주기 접근은 삶의 연속성과 선후관계에 따른 인과성을 고려할 수 있다는 장점이 있지만 실제 생애주기와 생애사건에 대한 정의 및 그 분류기준은 활용 목적과 방법에 따라 편차가 존재하므로, 보건의로 영역에서 주요하게 다루지는 법제도에 따른 생애주기별 인구집단을 정의하고 해당 인구집단의 의료이용 현황을 검토함
 - 사회·복지·인구정책과 관련 법령 및 국가통계조사에서 정의하는 생애주기별 인구집단은 대부분 연령에 따라 집단을 구분하지만 구체적인 기준점이 다르기 때문에, 생애주기 접근을 강조한 『국민건강증진종합계획(Health Plan 2030, 이하 HP 2030)』에서는 건강증진사업의 대상자를 연령과 취약성에 따라 아동·청소년, 여성, 장애인, 노인으로 분류함
 - 「청년기본법」과 「노인복지법」을 기본으로 분류한 생애주기별 인구집단은 청년기(20~34세), 중장년기(35~64세), 노년기(65세 이상)이었으며, 국민들이 인식하는 연령기준은 아동이 12세 이하, 청년은 35세 이하, 노인은 70세 이상으로 법적 기준과 상이함(이유경, 2021)
- 가명정보결합 자료원을 활용해 생애주기별 다빈도 질환과 의료이용 양상을 분석한 기존 문헌을 고찰하여 분석대상 변수를 선정하고 분석모형 설계의 근거를 확보함
 - HP2030뿐만 아니라 『제1차 국민건강보험종합계획(2019~2023)』에서도 출생부터 노년까지 전 생애에 걸친 건강보장 구현을 강조하며 생애주기에 따른 서비스제공을 제시하고 있으나 사망원인통계나 건강보험통계연보 등 정책의 개발과 평가에 활용되는 국가통계에서는 여전히 연도별, 연령별, 성별에 따른 횡단면 분석결과를 제공함

2) 가명정보결합 추진

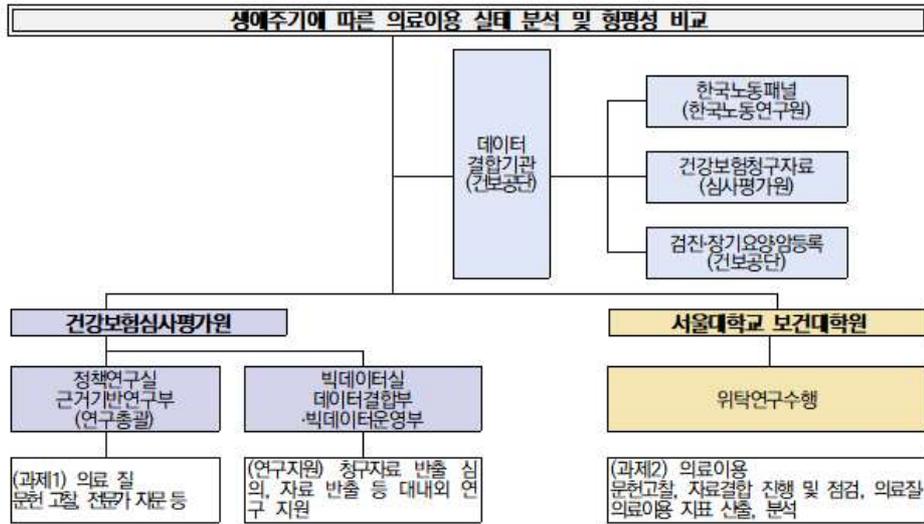
- 법률에 따라 가명정보 결합을 추진하기 위한 절차는 [그림 1]과 같고, 연구진을 포함한 참여기관과 그 역할은 [그림 2]와 같음
 - 연구진은 자료 간 결합률과 결합자료의 품질을 향상하기 위해 자료결합 진행 단계별 절차를 담당 전문 인력들과 고려사항 및 진행 현황을 점검하고 행정절차 및 기술적 지원을 제공함
 - 건강보험심사평가원, 국민건강보험공단, 한국노동연구원은 각 기관에서 관리하는 원 자료에서 결합키의 생성에 필요한 개인식별번호를 가명처리³⁾하여 일련번호와 함께 결합관리기관인 한국인터넷진흥원(Korea Internet & Security Agency, KISA)에 제공하고, 자료결합에 필요한 행정적, 실무적 협의를 지원함
 - KISA는 원 자료 수집에 참여한 대상자를 식별할 수 없는 형태로 가공한 가명정보로 결합키를 생성해 결합률을 확인하고 결합률에 따라 보완전략을 제안함
 - 결합전문기관으로 선정되어 연구에 참여하는 국민건강보험공단은 KISA에서 생성한 결합키를 전달받아 원 자료를 결합하고 반출심사를 거쳐 적정 판정을 받은 결합자료를 연구진이 이용할 수 있도록 분석환경을 제공하며 성과물을 반출함

3) 개인식별번호는 암호화 처리(영문과 숫자의 조합)하여 제공되며 이를 Key값이라고 함



[그림 1] 가명정보 결합 절차

* 자료: 가명정보보호위원회. 개인정보 가명처리-가명정보 결합. Retrieved from <https://www.pipc.go.kr/np/default/page.do?mCode=D040010000> (Access on 2022.04.12.)



[그림 2] 가명정보결합 참여 기관과 역할

3) 결합대상 자료 검토

- 가명정보결합 전 결합대상 자료원별 자료수집 방법에 따른 표본의 특징, 변수의 정의와 유형화 방법, 유지되거나 탈락하는 조건 및 그에 따른 조사 참여자의 일반적 특성을 검토함
 - 검토결과는 가명정보결합자료의 품질검토와 분석모형 설계에 활용함
 - 한국노동패널조사 자료 분석결과를 통해 성과 연령을 포함한 인구학적 특성과 결혼, 노동, 은퇴 등 생애사건의 경험을 중심으로 교육수준, 종사상지위, 가구의 소득과 지출, 노동 강도를 다양한 방법으로 변수화하고 의료이용의 인구집단별 차이를 파악하기 위한 독립변인을 구성함
 - 공개된 국가통계자료와 건강보험 청구 자료를 활용하여 한국인의 다빈도 질환과 주요 사망원인에서 성별, 연령별, 연도별 특징을 파악하고 특정 조건이 충족되었을 때 발현할 수 있는 질병발생 및 의료이용의 가능성을 탐색함

4) 가명정보결합자료 분석

- 한국노동패널조사 자료에서 얻은 독립변인에 따른 인구집단의 연도별, 연령별, 성별 평균 의료이용 빈도 및 비용의 차이를 분석하고 독립변인들이 의료이용에 미치는 효과크기를 추정함
- 생애주기별 인구집단의 인구학적 특성, 사회경제적 지위, 결혼, 노동, 은퇴 등 생애사건의 경험여부를 중심으로 주요 상병에 의한 외래 및 입원 의료비 지출규모를 분석을 위해서는 자료원의 한계점을 극복하고 자료생산단계의 품질 향상과 자료원간 부족한 부분을 조작적으로 채워 넣는 분석기법을 활용해 측정하고 산출한 값의 대표성을 확보해야 함(김남순 외, 2003; 김남순 외, 2018; 박은자 외, 2019; 오미애 외 2014; 오미애 외, 2015; 이해정, 2020)



[그림 3] 자료분석 절차

나. 연구 방법

1) 선행문헌 고찰

- 한국노동패널자료와 건강보험 청구자료를 활용한 보건학 분야 국내 연구를 고찰하여 연구 설계와 정책적 활용의 현황, 시사점, 제한점을 검토하고 가명정보결합 추진과정에서 고려해야 할 자료원의 특성과 예상 분석결과를 정리함

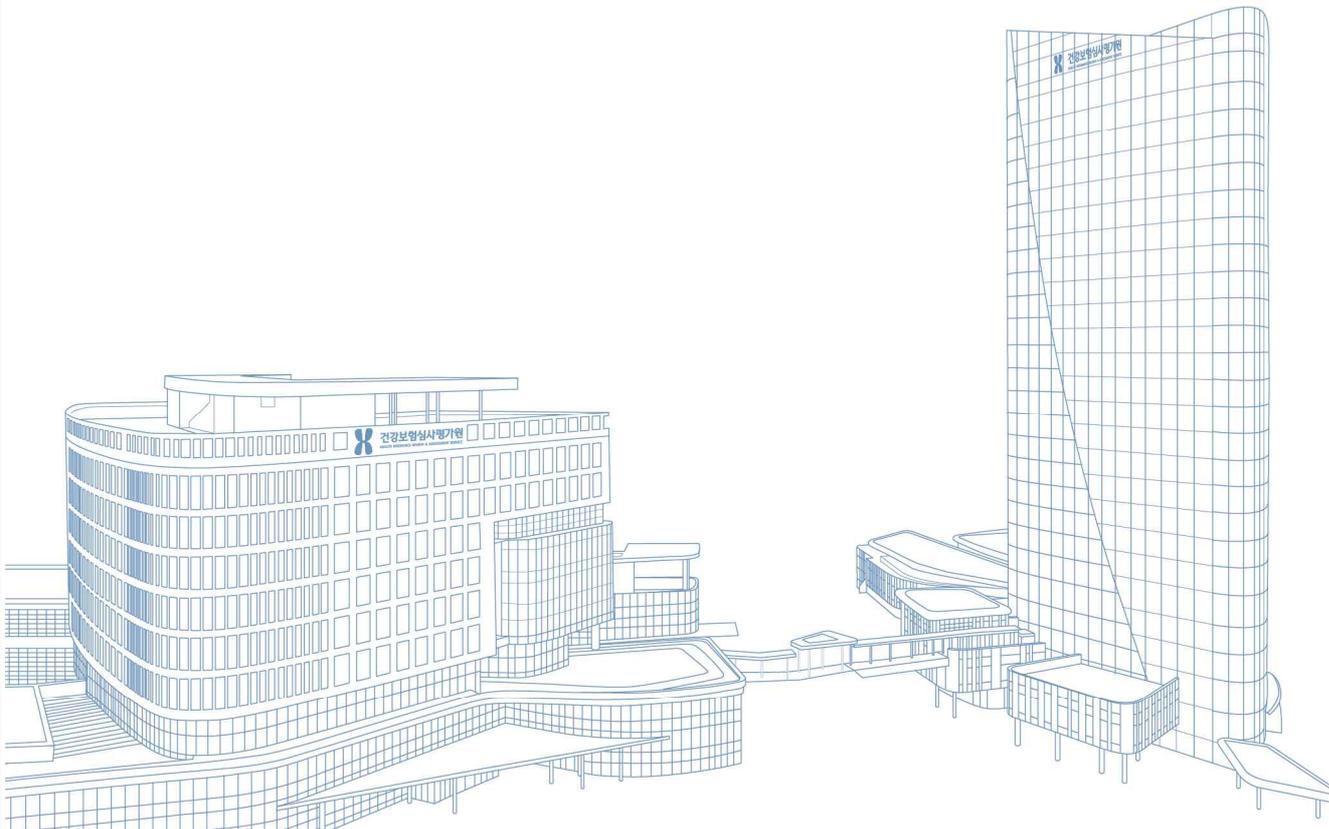
2) 가명정보결합자료 품질검토 및 의료이용 분석

- 가명정보결합자료의 결합률과 인구대표성을 포함한 품질을 검토함
 - 품질검토를 위해 자료원을 직접 생성 및 관리하거나 각 자료원에 대한 이해가

높고 통계학적 지식이 풍부한 전문가와 자문회의를 수행하여 자료분석을 위한 사전준비 및 분석전략의 적절성을 검증함

- 건강보험심사평가원과 건강보험공단의 자료는 전수조사 자료인 반면 한국노동패널자료는 비례추출로 구성된 종단적 조사자료로 표본의 탈락 및 결측을 고려해야 하므로 노동패널 자료를 기준으로 적어도 1회 이상 조사에 응답한 참여자를 결합대상으로 상정하여 결합기가 중복 결합되는 비율과 최종 결합률을 기준으로 가명정보 결합 적정성을 판단함(결합전문기관)

선행문헌 고찰



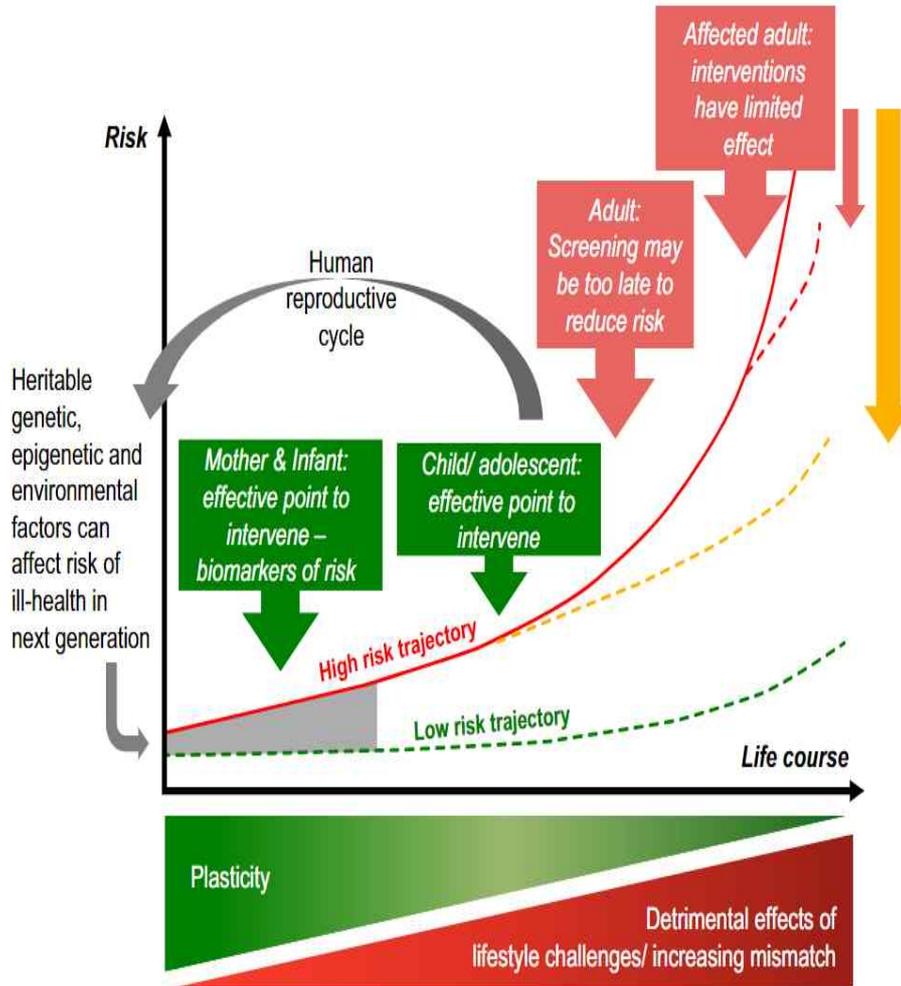
제2장 선행문헌 고찰

1. 이론적 배경

가. 생애주기 개념의 발전

- 인구집단 건강은 영양, 교육, 결혼, 직업, 거주환경 등 사회적 결정요인에 영향을 받으며, 생애주기 접근은 태아가 성장하는 주산기 발달부터 사망에 이르기까지 개인의 전 생애과정에서 경험하는 건강행태, 외상, 기능저하에 따라 건강수준이 달라진다는 관점임(Goodman, et al., 2016)
- 생애주기 접근은 발달적 관점과 구조적 관점으로 나눌 수 있음(Jones, et al., 2019)
 - 발달적 관점: 생애초기 발생부터 출생까지를 핵심적 생애기간으로 특정하고 임신부의 사회환경적 노출에 의한 생물학적 가변성이 후기건강에 영향을 미치는 점을 강조
 - 구조적 관점: 불균등한 사회구조 속에서 인종이나 소득에 따른 계층을 따라 건강위해 요인의 차등적 노출이 건강불평등으로 이어짐을 강조
- 생애주기 구분은 대체로 연령기준이지만, 기준점(cut-off)과 유형(category)은 시대적, 사회적 특성에 따라 변화해왔으며, 해당 시기에 기대되는 역할 및 규범에 대한 인식을 반영함(이윤경, 2021)
- 아동·청소년의 생애주기는 활용하는 주체나 목적에 따라 발달단계를 기준으로 영아기, 유아기, 아동기, 청소년기로 구분된 유형을 통합하거나 세분화할 수도 있고, 교육이라는 생애과업에 주목해 교육체계에 따라 학령전기와 학령기로 분류하기도 함
- 특정 생애주기 인간의 의사결정이나 과업달성 여부에 따른 행동변화를 설명하는 연구가 활발해지면서 노인인구의 노화를 질병이 아닌 발달과정의 일부로 취급하는 사회적 인식이 보편화되고, 사회규범과 생애경험의 상호작용이 개인의 신체적, 정신적, 사회적 건강에 미친 영향에 대한 이론이 발전함(Zacher, H., & Froidevaux, A., 2021)
- 만성질환이나 비감염성 질환 역시 영향요인과 노출시점을 특정할 수 없는 탓에 노인

성질환(geriatric illness)으로 취급되었지만 최근 생활양식의 변화와 진단의 기준 및 검사법의 발달로 발병연령이 점차 앞당겨지고 개인의 자가관리 및 건강행태에 따라 증증도와 예후를 개선하거나 악화할 수 있음에 대한 연구가 누적되면서 생애주기의 경험과 행태를 고려한 종단적 분석의 필요성이 증가함(Hanson, M. A., & Gluckman, P. D., 2014)



[그림 4] Lifecourse view of noncommunicable disease (NCD) risk

* 자료: Hanson, M. A., & Gluckman, P. D. (2014). Early developmental conditioning of later health and disease: physiology or pathophysiology?. *Physiological reviews*.

나. 생애주기 개념의 활용

1) 국제사회의 생애주기 의제화

- 국제사회에서는 생애주기 접근은 건강의 사회적 결정요인에 따른 건강불평등의 규명과 완화를 위해 다양한 연구, 정책, 사업을 설계 및 평가하는 방향으로 확장되고 있음
 - 건강불평등이란 개인이나 집단 간 건강 성취에서 나타나는 차이, 변이, 격차를 의미하고, 사회적, 경제적, 인구학적, 지리적으로 조직된 인구집단의 건강격차를 드러내고 이에 대한 정책적 개입을 촉구하기 위해 발전한 개념임(Kawachi, I., Subramanian, S. V., & Almeida-Filho, N., 2002; 정최경희, & 김유미., 2013)
 - 세계보건기구가 2008년 건강의 사회적 결정요인 위원회를 통해 건강의 사회적 결정요인이란 개념을 소개하는 보고서를 공개하면서 생태병리학적 요인과 임상적 처치의 교육, 소득, 직업을 일컫는 사회경제적 지위(Socio-Economic Status, SES)와 의료접근성, 사회적 포용성(Social inclusion), 물리적 환경 등 비임상적 건강의 영향요인이 강조됨(WHO, 2008)
 - 초기 세계보건기구(WHO, 2010)는 건강의 사회적 결정요인을 건강불평등의 구조적 결정요인과 건강으로 이르는 매개적 결정요인으로 구분해 설명했지만, 현재 사회적 결정요인은 온존(well-being), 지속가능성(sustainability), 형평성(equity)에 관한 윤리적, 철학적 차원에서 논의되기도 하고 사회체계이론, 생태학적 접근, 생애주기 접근 등 관점에 따라 다양한 이론적 틀을 마련하며 확장함
 - 질병예방과 건강증진을 추구함에 있어 국제사회는 인구집단의 건강수준 격차와 건강 불평등을 완화하고 보편적 건강보장을 달성하기 위해 예방접종 및 건강증진 등 지역 사회 단위 개입전략과 건강고려정책(Health in All policies)을 촉구하게 됨(WHO, 2014)

2) 국내 보건의료정책적 활용

- 국내에서 생애주기 접근을 적용한 대표적인 보건의료정책은 HP 2030으로, 인구집

단별 건강형평성 제고를 목표로 생애주기 인구집단별 주요 위험요인을 조기에 진단하고 관리하여 질병을 예방하는 보건의료사업을 유형화하고 목표치를 설정해 연도별, 성별 성과지표를 평가하고 있음

- 이전 종합계획들은 목표에 비해 추진체계와 전략이 모호하고 평가지표의 타당성과 실효성이 부족하다는 비판이 반복되면서 근거기반 정책 추진의 한계를 보완하기 위해 자료원과 사업, 중재, 제도 간의 유기적인 연계의 필요성이 대두됨(최용준, 윤태호, & 신동수, 2012; 장숙량, 이주현, 김창오, 허현희, 황종남, & 김태연, 2017)
- 2021년 발표된 HP2030은 성별, 지역별, 소득별 건강격차 완화를 강조한다는 점에서 형평성을 강조하고 있지만 사업의 성과를 평가하고 모니터링하기 위한 지표개발과 정책평가 연구에서 자료원 및 조사방법의 제한점이 지적되고 있음(김은아, 오유미, & 조성일, 2020)

3) 건강 및 의료의 생애주기 영향에 관한 실증연구

- 학계에서는 특정한 생애과업이나 생애사건의 경험이 인구집단의 건강에 미치는 영향을 분석하거나 특정시기 인구집단 내 건강격차를 설명하고자 하는 연구가 활발하게 이뤄졌으며, 건강격차에 대한 탐색적 연구뿐만 아니라 의료이용과 의료의 질에 관한 정책적 함의를 도출하고 있었음
- 북미와 유럽 지역에서 발표된 생애초기 아동의 건강악화가 후기건강에 미친 영향에 관한 연구를 고찰한 결과, 아동의 생애초기 건강악화가 10% 감소하면 연간 3백만 DALYs 혹은 1050억 달러 규모의 의료비 지출을 절감할 수 있음(Bellis, M. A., Hughes, K., Ford, K., Rodriguez, G. R., Sethi, D., & Passmore, J., 2019)
- 한국노동패널조사를 활용한 전향적 코호트 연구에서는 아동기 사회경제적 위치와 성인기 사망률 사이의 관계를 탐색하여 유년기 사회경제적 위치가 낮을수록 50세 이상 남성의 사망위험이 증가함을 밝힘(Khang YH, 2006)
- 국민건강영양조사와 통계청 사망률 자료를 연계한 연구에서는 사회경제적 지위를 대리하도록 개인의 교육수준과 직업을 재분류하여 사회경제적 지위가 낮을수록 모든 원인에 의한 사망위험(all-cause mortality)이 높아지고, 이러한 경향은 성, 연령, 특정 지표에 따라 교차함을 드러내 생애주기의 특성과 동일 생애주기 인구집단이

경험하는 건강격차를 제시함(Khang, Y. H., & Kim, H. R., 2016)

- 건강보험 청구자료를 이용해 특정질환의 사망률이 소득수준에 따라 달라짐을 규명한 연구결과, 치사율의 소득격차가 나타난 질환은 성별 차이가 있었고 그 영향요인은 해외연구와도 달라 사회경제적 불평등이 건강에 미치는 영향은 평가하는 기준과 대상에 따라 상대적으로 나타남을 시사함(Jung-Choi, K., Khang, Y. H., & Cho, H. J., 2011)
- 그 외 건강보험 청구자료나 의료패널을 활용하여 생애주기별 인구집단의 건강을 분석한 연구는 다수 진행되었으나, 사회경제적 요인이나 생애경험에 관련한 변수의 부족으로 인해 특정시기 다빈도질환의 유병률 및 사망률, 그리고 의료비 지출에 관한 분석에 집중되어 있음
- 교육, 노동, 결혼 등 광범위한 사회경제적 요소들을 추적조사하고 있는 한국노동패널 조사의 경우 건강이나 보건의료에 관련된 변수가 제한적이기 때문에 생애주기별 연구들 속에서도 보건학적 연구는 제한적이었음(조진호, 김유빈, & 이지은., 2018; 노가은, & 이지은., 2019; 이석용, & 이지은., 2020)

2. 인구집단의 건강과 의료이용에 관한 국내연구

가. 청구자료 및 임상자료를 활용한 연구

- 조혜민, 최혜미(2019)에 따르면, 1990년, 2000년, 2010년, 3개년도 의학지 논문에서 다중비교, 로지스틱 회귀분석, COX와 Kaplan-Meier 생존분석이 활용되나, Choi와 Kang(2009)은 통계기법 활용에서 여러 오류가 발생해 연구의 정확성과 전문성을 향상해야 할 필요성이 제기했음
- 2002년부터 2018년 사이 짝수해에 New England Journal of Medicine (NEJM; IF=70.670), Journal of the American Medical Association (JAMA; IF=51.273), Journal of Dental Research (JDR; IF=5.125)에서 발행된 총 2,645 편의 논문의 경향을 분석함

- 기술통계량, 메타분석, 요인분석을 분석대상에서 제외하고 F-검정을 분산분석으로 분류하였으며 일반화선형 모형과 로지스틱 회귀분석을 분리하여 정리한 분석결과, 분석유형별 학회지별 분석기법의 비율은 아래와 같음
- 상위 6개 통계기법은 생존분석, 로지스틱 회귀분석, 역학적 통계량, 교차표분석, t-검정과 분산분석이었고, 그 중 교차표분석과 분산분석은 유의하게 감소하는데 반해 일반화선형모형과 베이지안 방법의 사용이 유의하게 증가함
- 30% 이상의 논문에서 SAS를 사용하고 있는 가운데, R을 활용한 비율이 급격히 증가했고, 빅데이터와 머신러닝을 활용한 분석은 전체의 0.5% 내외였음

○ 많은 보건의료분야 연구가 여전히 특정 질환의 완치나 5년 내 생존 여부에 따라 건강개선을 평가하는 가운데, 김수진 외(2021)는 특정 질병의 치료에 집중하기보다 건강결과를 개선하기 위한 의료의 질적 향상을 강조하며 아래와 같은 의료의 질 지표별 현황을 분석함

- (의료의 질 지표) 2011년부터 2020년까지의 건강보험 맞춤형DB를 활용한 만성기 케어의 지표는 ‘외래 민감성 질환으로 인한 계획되지 않은 입원’, ‘천식, 당뇨, 뇌전증으로 인한 계획되지 않은 입원(19세 미만)’, ‘고혈압 혹은 당뇨병 진단 후 2년간의 복약순응도’가 있고, 급성기 케어 지표로는 ‘병원 입원이 필요하지 않은 급성질환으로 인한 응급입원’, ‘퇴원 후 30일 내 응급 재입원’, ‘울혈성 심부전 응급입원 1년 내 사망 및 입원’, ‘허혈성 뇌졸중 응급입원 1년 내 사망 및 입원’이 있음
- 그 외 만성기 케어 지표에는 2010년부터 2019년까지 국민건강영양조사 자료를 활용한 ‘만성질환을 가진 사람들의 삶의 질’ 지표와 2021년 2달간 진행한 전화설문조사로 얻은 ‘우리나라 의료시스템의 질환관리에 대한 역할수행’ 지표가 있음
- (외래 민감성 질환으로 인한 계획되지 않은 입원) 2016년부터 입원 10만 건당 약 285건에서 2020년 205건으로 연평균 7.9%, 코로나 전후 2년간(2019-2020)에는 21.7% 감소했고, 여성에 비해 남성의 입원이 전체 기간 조금씩 높은 가운데 소득수준별 차이는 1분위부터 5분위까지는 적은 반면 의료급여 수급권자에서 크게 증가하는 것으로 나타남
- (천식, 당뇨, 뇌전증으로 인한 계획되지 않은 입원(19세 미만)) 외래 민감성 질환으로

- 인한 계획되지 않은 입원과 유사한 양상을 띄고 있지만 그 감소세가 더욱 두드러졌고, 소득수준에 따라 매년 10만 건 당 건수와 순위가 달라짐
- (고혈압 혹은 당뇨병 진단 후 2년간의 복약순응도) 소득수준에 따라 고혈압약 복약순응도는 의료급여 수급권자가 가장 낮고 소득 5분위에서 가장 높았으며, 당뇨약은 2013년 이후 의료급여 수급권자의 복약순응도가 개선되고 있음
 - (병원 입원이 필요하지 않은 급성질환으로 인한 응급입원) 2016년 인구 10만 명당 약 1,023건에서 2020년 583건으로 연평균 13.1%, 코로나 전후 2년간 37.7% 감소했으며 의료급여 수급권자에서 가장 높은 수치를 보이고 있음
 - (퇴원 후 30일 내 응급 재입원) 2016년부터 조금씩 감소했지만 2019년 대비 2020년 5.8%가 감소했고 의료급여 수급권자의 재입원 비율이 제일 높았음
 - (울혈성 심부전 응급입원 1년 내 사망 및 입원과 허혈성 뇌졸중 응급입원 1년 내 사망 및 입원) 허혈성 뇌졸중은 해당 질환으로 인한 입원은 감소하고 다른 질환에 따른 입원이 증가한 반면, 울혈성 심부전은 해당 질환과 다른 질환으로 인한 입원이 모두 증가해서 뇌졸중 집중치료실의 확대와 상대적으로 재활치료 비율이 높은 것이 영향을 미쳤을 수도 있음

나. 한국노동패널을 활용한 연구

1) 한국노동패널조사 자료의 특성에 관한 연구

- 박민규, 김사라(2015)는 한국노동패널조사 자료의 자료생성과정에서 발생하는 표본마모와 모집단 대표성 저하의 문제를 해결하기 위한 가중치 활용을 제안하고 Duncan(1995)의 연구를 참고한 한국노동패널에서 제공하는 가중치 산출과정을 제시했으나 선행문헌에서 가중치 여부에 따른 분석결과를 비교해보면, 계수와 분산 추정량의 크기가 달라지긴 했으나 유의미한 차이가 있다고 볼 수 없다고 함
- 패널조사는 동일인을 T년 간 반복조사하기 때문에 전체 자료관측치의 수는 실제 자료를 제공하는 가구원의 수보다 크기 때문에 연도별 종단면 가중치를 혼합선형모형에 그대로 사용할 경우 실제 모집단의 T배가 되는 오류가 발생함

- 종단면 가중치를 활용하기 위해선 T년 간 누적된 자료에서 특정 t연도에 대응하는 모집단의 상대적 크기를 반영한 가중치로 표준화하고, 가중치 여부에 따른 모수추정에 미치는 영향이 달라지거나, 과대추정되는 경우를 고려해 가중치를 표본크기 기준으로 재표준화하는 두가지 단계를 거쳐야 함
 - 무엇보다 패널조사자료를 활용한 종단면 분석은 대부분 유한모집단의 동적 변화를 전제하므로 가중치를 활용하는 것이 바람직하고, 결과값이 유사하더라도 모형의 강건성을 보장해 연구의 질을 높여야 함
- 1차부터 11차까지 한국노동패널조사 자료의 표본마모를 분석한 이상협, 박찬용, 정성석, & 최혜미(2011)에 따르면, 1차(1998)부터 11차(2008)까지 원표본 유지율은 74.2%로 1차와 11차 표본의 일반적 특성 중에는 학력과 직업에 따른 표본분포의 차이가 나타났음
- 무응답이다가 다시 조사에 참여한 경우를 신규진입자로 취급하여 생존분석을 수행한 결과, 3차까지 초반 탈락자는 학력이 높을수록 많았으나 이후로는 곡선이 평행했으며, 7차까진 60세 이상의 탈락률이 가장 낮고 30세 미만의 탈락률이 가장 높았으나 탈락률 격차가 점차 줄어들었음
 - 조사가 진행될수록 특성별 격차나 탈락률 크기는 줄었으나, 미취업자의 탈락률은 지속해서 높았고, 시간이 지날수록 취업자와의 탈락률이 벌어졌음
 - 가구소득별 표본이탈과 그 영향요인을 분석한 홍만기, 최효미(2014)는 가구소득이 높을수록 무응답 탈락이, 낮을수록 가구소멸로 인한 탈락확률이 높으므로, 결측치 보정 없는 추정 시 불평등도를 과소평가할 위험을 강조했다음

2) 노동패널조사자료를 활용한 중고령자 대상 연구

- 노가은, 이지은(2019)에 따르면, 20년간 중고령자 대상 연구는 총 125건(학술지등재연구 70건, 학위논문 20건, 정책자료집 6건 등)으로 청년이나 여성 대상 연구에 비해 월등히 많고, 노동패널의 부가조사, 경제활동인구조사, 가계금융동향조사, 미국의 패널조사 등 여러 자료원을 결합한 연구가 시도되었음

- 종속변수를 기준으로 구분한 연구주제는 가계경제 및 세대에 관한 소득소비, 자산부채, 계층이동과 가구변화 등이 47건으로 가장 많았고, 은퇴와 재취업을 둘러싼 노동의 공급과 이동에 관한 연구가 46건, 삶의 질과 만족도가 21건으로 뒤를 이었으며 건강과 의료에 관한 연구는 6건이었음
- 패널의 구성상 의료비 지출과 의료비부담은 가구지출 변수로 분류되고, 건강보험과 민간의료보험의 가입여부와 그 종류는 연금과 보험으로 분류되던 가운데, 건강과 의료에 관한 변수는 다른 종속변수와 상관관계나 인과관계가 있는 독립변수로도 활발하게 활용되었으며, 고령화와 건강불평등을 중심으로 중고령자 세대의 경제적, 의료적 격차가 분석모형에 반영되었음
- 종속변수로는 주관적 건강, 건강기대여명, 활동장애, 만성질환 유병률, 의료사용여부 등이 변수로 활용되었고, 독립변수로는 건강증진행위를 통제하기 위해 정기적 건강진단 여부, 정기적 운동여부, 흡연과 음주 여부, 식사조절여부, 공적 의료기관 사용빈도가 활용되었고, 건강상태를 통제하기 위해 장애여부, 주관적 건강, 기능제약, 객관적 건강, BMI, 은퇴전 건강, 주요활동 유무가 통제되었음

3) 노동패널조사자료를 활용한 여성 대상 연구

- 이석용, 이지은(2020)에 따르면, 노동패널로 여성을 다룬 150건의 연구 중 학술지등재연구는 139건, 정책보고서가 5건이었고, 패널만 활용한 130건과 달리 20건의 연구는 경제활동인구조사, 생활시간조사, 해외 패널조사나 국가통계를 혼합해 분석을 수행하였으며, 부가조사를 활용한 연구는 13건으로 청년층, 중고령자층을 대상으로 하는 부가조사와 달리 여성을 특정한 부가조사는 아직 수행되지 않았기 때문일 수도 있음
- 종속변수를 기준으로 연구주제를 분류해보면, 근로조건이 70편, 노동공급과 이동이 33편, 결혼과 출산이 27편이었고 건강은 단 3편이었음
- 임신과 출산의 기간, 확률, 사건발생까지의 기간이 결혼과 출산으로 분류되면서 여성의 건강에는 만성질환 여부, 건강상태, 정기 건강검진 여부가 포함됐으며, 중고령자 대상 연구와 달리 의료비나 건강보험 가입 여부는 주요 분석대상이 되지 않았음

- 중고령자 대상 연구와 유사하게 여성 대상 연구에서도 건강 변수는 상관관계와 인과 관계 분석에 필요한 통제변수로 주로 활용되면서 여성의 취업을 다룬 노동공급 연구에선 연령대와 결혼상태, 손자녀에 따른 건강격차와 취업상태의 차이에 대한 분석이 이뤄졌고 여성가구주의 가계빈곤을 다룬 연구는 인적자본론적 관점에서 여성의 학력과 건강상태가 빈곤여부에 미치는 영향을 강조함
- 사회경제적 지위를 중심으로 건강에 미치는 영향요인을 살핀 연구에선 배우자의 동거여부와 자녀양육이 건강에 영향을 미치는 반면 학력에 따라 사회경제적 지위가 건강에 미치는 영향은 달라졌고, 취업여부는 신체적, 정신적 건강에 긍정적인 영향이 있음

3. 한국의 건강 및 의료이용

가. 건강보험 빅데이터

1) 건강보험심사평가원의 건강보험 청구자료

- 건강보험심사평가원은 의료기관이 제출한 환자의 상병명, 일반정보, 처치 및 수술료의 진료내역, 청구비용을 포함해 청구명세서, 약제와 의약품, 요양기관 등 모든 의료이용 건수별 정보를 수집 및 관리하고 있음
- 청구자료는 국내 건강보험가입자 및 의료급여 수급권자의 비급여 서비스를 제외한 모든 의료이용 정보를 포함하고 있어 대표성 높은 자료임
- 결합변수: 명세서식별번호, 환자식별번호, 연령, 성별, 진료개시년월일, 진료종료년월일, 내원일수, 주상병, 요양일수, 총 진료비, 총 처방일수 등

2) 국민건강보험공단의 건강보험 자격 및 건강검진과 장기요양 행정자료

- 건강보험 가입자의 자격과 보험료 징수를 관리하고 국민건강보험과 노인장기 요양보험의 요양급여비용을 지급하는 기관임
- 대상자를 식별할 수 없는 방식으로 가공한 정보(자격정보, 보험료분위, 건강검진 및 노인장기 요양보험 관련 이용 및 평가 정보)를 공익적 연구에 제공하고 있음(Ryu,

D. R, 2017)

- 자격 및 보험료: 개인일련번호, 가입자구분, 보험료, 보험료20분위, 실거주주소코드 (시군구), 사업장업종세분류, 종합장애등급코드, 장애유형, 사망일자
- 건강검진: 검진년도, 검진일자, 검사항목별 검사여부, 일반검진결과, 암검진결과, 구강검진결과, 생애주기구강검진결과 등
- 장기요양: 장기요양신청일자, 대상자구분, 경감구분, 주거형태, 수발자, 동거인, 환산 점수, 인정점수, 판정등급, 유효기간, 급여제공년월, 급여시작일자, 지급년월, 명세서 번호, 급여종별코드, 급여구분코드, 공단부담금, 본인부담금, 결정총액, 급여일수, 인력기관(요양기관기호, 등록일자, 폐업일자, 장기요양시설유무, 요양병원유무, 등)

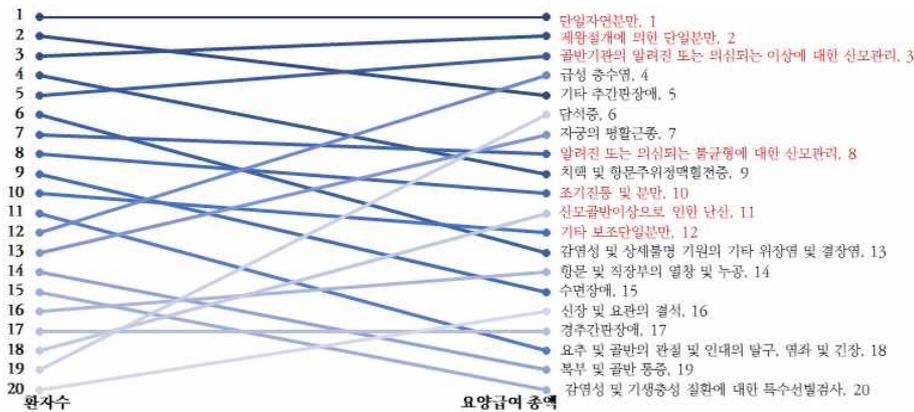


[그림 5] 결합 대상 건강보험 빅데이터

나. 연도별, 연령별, 성별 질병 및 사망 현황

1) 2020년 환자수별, 영양급여총액별 입원서비스를 이용한 다빈도질환 순위

- 건강보험심사평가원에서 공개하는 보건 의료 빅데이터를 활용해 상병코드 기준 2020년 입원서비스를 이용한 다빈도 질환을 살펴보면, 다빈도 질환은 연령대별 건강 수준과 생애과업에 따라 변함(건강보험심사평가원, 2022)
- 같은 질병도 연령대에 따라 총 환자수, 내원일수, 영양급여총액의 규모와 순위가 달라져 의료비부담 완화 및 건강증진을 위해서는 생애주기별 의료이용의 양과 질에 관한 분석의 필요성이 제기됨
- 각종 외상이나 손상으로 인한 골절이나 염증과 감염으로 인한 급성기 질환이 대부분 인 10대 및 20대와 달리 30대는 전 연령대 중 유일하게 자연분만이 환자 수 및 영양급여총액의 1순위를 차지하며 임신과 출산 관련 산부인과계 질환으로 인한 의료 이용이 대폭 늘어나는 시기임
- 남성을 포함해 임신과 출산 관련 상병코드가 높은 순위를 차지한 30대는 비교적 건강하고 생산성이 높아 질병으로 인한 의료이용은 상대적으로 적지만 생활습관 개선이나 조기치료가 필요한 추간판 장애, 결석, 담석증, 수면장애 등 증상이 나타나기 시작하는 시기이기도 함
- 출산률 감소로 인해 2018년부터 2020년까지 환자수가 줄어들고 있지만 20대 다빈도 질환으로 인한 입원서비스 영양급여총액은 동일기간 지속적으로 증가하고 있음



[그림 6] 2020년 30대 환자수 및 영양급여총액에 따른 다빈도 질환 순위

* 자료: 건강보험심사평가원, 보건 의료 빅데이터 개방 시스템.

제2장 선행문헌 고찰 ●●

- 근골격계 질환으로 인해 의료서비스를 이용하는 환자수가 증가하는 40대와 50대를 지나 60대부터는 질환당 진료비가 증가하고 백내장, 당뇨병, 뇌경색 등 만성·퇴행성 질환이 늘어나면서 방사선요법, 화학요법, 수혈, 예비치료, 완화의료, 등 '기타 의학 적 관리를 위한 보건서비스(Z51)'도 다빈도 입원서비스에 포함됨
- 2020년 코로나19 유행으로 '감염성 및 기생충성 질환에 대한 특수선별검사(Z11)'를 이용한 환자수가 모든 연령대에서 크게 증가함. 만 60세 이상 인구는 특수선별검사를 많이 이용(7만 5천명 이상)한 것 이외에도 원래의 다빈도 질환으로 인한 의료이용을 횟수와 요양급여 총액이 증가함



[그림 7] 2020년 60대 환자수 및 요양급여총액에 따른 다빈도 질환 순위

출처: 건강보험심사평가원. 보건의료빅데이터개방시스템.

- 환자 수에 따른 다빈도 질환의 연령대별 양상을 살펴보기 위해 입원 및 외래 의료서 비스를 이용하게 된 주상병을 분류한 결과, 2020년에는 코로나19 범유행으로 전반 적인 의료이용 환자수와 지출규모가 소폭 감소했으나 다빈도 질환 중에는 근골격계 통증, 급성 염증성 질환, 피부 및 피하조직의 병변과 같이 증증도가 낮은 질환이 높은 순위를 차지했음
- 그 중 급성 염증성 질환은 호흡기계, 소화기계, 피부, 안구 등 신체부위별로 입원과 외래 이용 빈도가 달라졌으며, 호흡기계와 소화기계는 입원 이용이 많은 반면 피부와 안구 등은 외래 이용이 많은 양상을 보였음

- 근골격계 질환은 전 연령대에서 입원 및 외래 서비스 이용이 모두 많았지만, 연령대에 따라 20대부터 40대까지 노동활동에 종사하는 연령대에서는 염좌, 긴장, 탈구, 골절 등 손상 및 외인에 의한 질환이 많은 반면, 50대 이상의 연령대에서는 추간판 장애나 관절통 같이 퇴행성 질환이 많았음

순위	입원							외래							
	20대	30대	40대	50대	60대	70대	80대	20대	30대	40대	50대	60대	70대	80대	
1	A09	O80	M51	M51	H25	H25	F00	K05	K05	K05	K05	K05	I10	I10	
2	J10	M51	K64	H25	M51	M17	J18	J20	J20	J20	I10	I10	K05	K05	순환기계
3	K35	O82	D25	M75	M17	I63	H25	K02	J30	K21	J20	J20	M17	M54	당뇨
4	Z11	K64	M75	K64	M48	M48	I63	K01	K02	M54	M54	M54	M54	M17	백내장
5	J18	O34	A09	M50	M75	Z11	S72	J30	M54	I10	K21	K21	J20	J20	
6	S93	A09	M50	A09	I20	J18	Z11	L23	K29	K29	K29	M17	E11	E11	소화기계
7	R10	O33	S33	M48	A09	I20	S32	K29	K21	J30	K04	E11	K21	F00	치주질환
8	S83	O60	N20	M23	Z11	F00	A41	A09	J03	K02	L23	K29	M48	M48	위장
9	S82	G47	H26	M17	R42	S32	S22	M54	J06	L23	E11	K04	K29	K29	근골격계
10	S62	O83	C50	H26	I63	M51	A09	J03	L23	K04	E78	E78	K08	K21	손상
11	J35	S33	K80	R42	K64	S22	I50	K21	K04	J06	K02	L23	L23	Z46	
12	J34	K35	D24	I20	H81	A09	G81	J06	A09	H10	M75	K02	K04	L23	
13	R50	D25	G47	S33	M23	C34	J15	H10	H10	J03	J30	M75	N40	K08	
14	S52	R10	Z11	H81	H26	R42	N39	K04	J02	M79	H10	H10	H10	N40	임신
15	M51	Z11	K35	Z11	S52	E11	M48	S93	Z34	M75	M17	M48	H25	H10	
16	J20	K60	R10	N20	E11	H26	G20	N76	J04	B35	M79	J30	K02	M79	코로나
17	J03	M50	M54	M54	S22	S72	E11	J02	J01	J02	B35	K08	M75	H04	
18	S92	O65	R42	S82	K63	K80	Z03	H00	N76	S33	J06	M79	M79	J00	
19	S42	K80	M23	I83	S82	N18	K80	L02	J00	A09	K03	B35	Z46	K59	
20	S33	N20	H81	K63	Z51	M75	N18	J04	S33	E78	H04	H04	H04	M81	

[그림 8] 2020년 연령대별 입원 및 외래 환자수에 따른 다빈도 질환 순위와 유형

* 자료: 건강보험심사평가원. 보건 의료 빅데이터 개방 시스템.

- 통계청이 발표하고 있는 사망원인통계 결과에 따르면, 2010년부터 2020년까지 지난 10년 간 성·연령별 사망자수는 연평균 6%씩 증가하고 있으며, 70대 이하 인구의 연평균 사망자수는 감소하는 반면 80세 이상 인구의 사망자수가 연평균 20%씩 증가하고 있음
- 사망자수 성비는 50대와 60대에서 가장 높았다가 급격히 감소하여 80대 이상에서 성비역전 현상이 나타나 남녀의 평균수명의 차이를 시사함

- 연령대별 10가지 주요 사망원인을 꼽아보면, 생애주기별 생활양식과 건강수준을 반영하듯 뚜렷한 차이를 보였고, 한국인 남녀 전체의 사망원인은 악성 신생물, 심장 질환, 폐렴 순이었음
- 선천성 질환이나 소아암이 주요 사망원인인 10대와 달리 20대부터 30대까지 인구의 사망원인은 2순위인 악성 신생물을 제외하면 고의적 자해, 운수사고, 외상이나 사고 처럼 질병이 아닌 다른 요인이 원인인 경우가 많았음
- 악성 신생물과 고의적 자해가 1, 2위를 다투는 40대부터 50대까지 인구의 기타 사망원인은 심장질환, 간질환, 뇌혈관질환으로 중증도와 치료시기에 따라 예후가 달라지는 질환 순이었고, 이러한 경향은 60대 이후로도 이어지면서 고의적 자해의 순위가 내려가고 사망 전 치료나 요양 과정에서 발생하는 폐렴, 패혈증이 늘어남

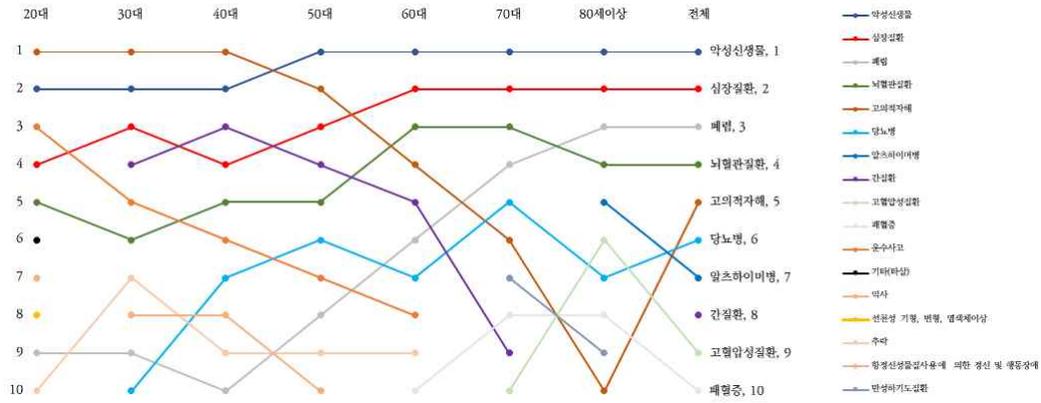
〈표 3〉 성·연령별 사망자 수 현황(2010년, 2015년, 2020년)

(단위: 명, %)

		연 령 (세)										
		전체*	0	1-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80 이상
남녀 전체	2010년	255,405	1,508	663	1,451	3,760	7,273	17,579	28,304	39,768	69,528	85,501
	2015년	275,895	1,190	494	905	2,634	5,418	14,161	28,753	36,596	73,065	112,638
	2020년	304,948	674	327	766	2,706	4,759	11,574	26,390	41,095	68,301	148,329
	연평균 증감률	(6)	(-24)	(-21)	(-19)	(-10)	(-13)	(-13)	(-2)	(1)	(-1)	(20)
남	2010년	142,358	888	388	959	2,384	4,697	12,697	21,130	27,812	40,404	30,955
	2015년	150,449	655	294	714	1,735	3,843	10,278	22,123	25,324	44,135	38,167
	2020년	165,163	387	186	486	1,660	2,968	7,737	19,171	29,774	43,649	59,136
	연평균 증감률	(5)	(-24)	(-22)	(-20)	(-11)	(-14)	(-15)	(-3)	(2)	(3)	(24)
여	2010년	113,047	620	275	492	1,376	2,576	4,882	7,174	11,956	29,124	54,546
	2015년	125,446	535	200	318	860	2,204	4,289	7,362	10,254	28,748	70,836
	2020년	139,785	287	141	280	1,046	1,791	3,837	7,219	11,321	24,652	89,193
	연평균 증감률	(7)	(-23)	(-20)	(-17)	(-9)	(-11)	(-8)	(-)	(-2)	(-5)	(18)
사망자 성 비 (남/여)	2010년	1.26	1.43	1.41	1.95	1.73	1.82	2.60	2.95	2.33	1.39	0.57
	2015년	1.20	1.22	1.47	1.85	2.06	1.68	2.30	2.91	2.57	1.54	0.59
	2020년	1.18	1.35	1.32	1.74	1.59	1.66	2.02	2.66	2.63	1.77	0.66

* 자료: 통계청. (2021). 2020년 사망원인통계 결과.

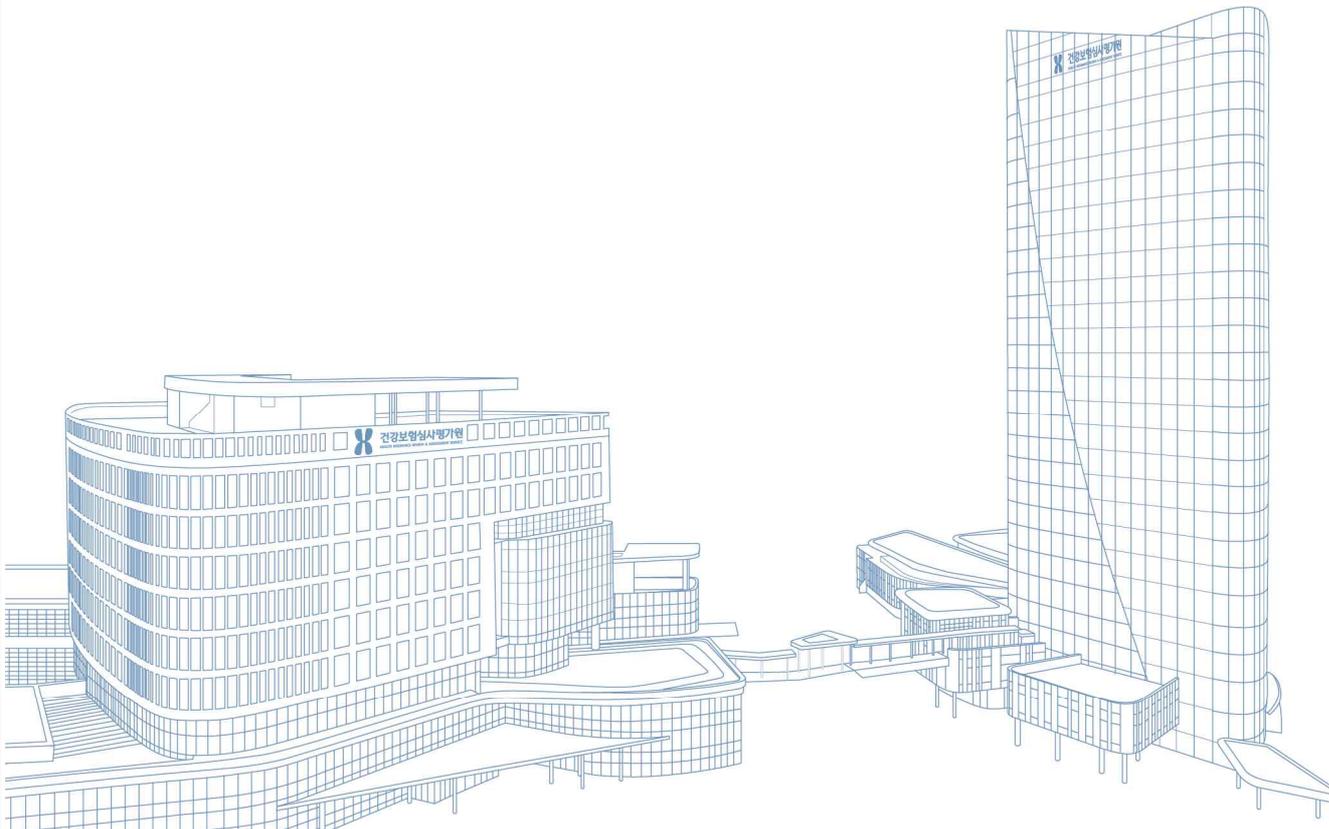
생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역



[그림 9] 2020년 연령대별 주요 사망원인 순위

* 자료: 통계청. (2021). 2020년 사망원인통계 결과.

가명정보결합자료 분석결과



제3장**가명정보결합과 분석결과****1. 가명정보결합 자료의 구조****가. 한국노동패널조사**

- 한국노동연구원에서 1998년부터 생성 및 관리하고 있는 한국노동패널조사는 고용 및 노동 특성뿐만 아니라 개인의 인구학적 특성, 교육수준과 혼인상태, 소비지출 행태를 매년 정기적으로 조사하고, 시기별, 주제별 부가조사를 비정기적으로 시행해 사회적 결정요인에 대한 포괄적 분석이 가능함
 - 원 표본은 제주도와 병역의무 이행자, 시설거주가구(원)을 제외하고 1998년 15세 이상 도시거주 가구주를 1995년 인구주택 총조사의 10% 표본조사구에서 5가구씩 임의추출하여 총 5,000가구를 매년 추적·조사하기 시작했음(장인성, 이지은, 신선옥, 정현상, 권익성, & 임태민, 2022)
 - 표본의 이탈과 마모를 고려하여 2005년 인구주택 총조사를 기준으로 2009년 1,415가구, 2016년 인구주택 총조사를 기준으로 2018년 5,044가구를 조사대상에 추가해왔으며, 가구주와 가구원을 조사대상으로 삼아 가구의 독립과 결혼, 동거, 입양을 통한 신규가구원을 포함한 결과, 제23차(2020년) 자료 기준 2009년도 통합표본 기준 6,711개 가구에 대한 조사에 성공해 13,577명이 조사되었음
- 본 연구에서는 2007년부터 2020년까지 적어도 1회 이상 한국노동패널조사에 참여한 자를 가명정보결합 대상으로 삼고, 가구테이블, 개인테이블, 직업력 테이블을 모두 결합한 하나의 데이터셋을 제공받았음
 - 민감정보 보호규정에 따라 소득변수 중 금융소득과 부동산소득을 합산한 값으로 제공받았으나 제공기관의 가명처리 과정에서 오류로 최종분석 변수로 활용하지 못함

나. 건강보험 빅데이터

- 건강보험심사평가원의 요양급여 청구자료는 2007년 1월부터 2021년 7월까지 의료기관에서 청구한 환자별 명세서 일반내역(T200), 진료내역(T300), 수진자 상병내역(T400), 원외처방전상세내역(T530)을 제공받음
- 건강보험공단의 자격·건강검진·장기요양 행정자료는 2011년부터 2020년까지 자료를 제공받음
 - 모든 빅데이터 데이터셋은 자료의 제공기관 및 결합기관 전문가들과의 협의를 통해 소수응답에 따른 개인식별 가능성을 차단하고 이상치 및 민감정보는 상하단코딩으로 범주화하는 마스킹 과정 거쳤음
 - ‘코로나19’는 민감상병으로 분류되어 마스킹됨

다. 가명정보결합 자료

1) 결합 결과

- 결합률과 결합자료 접근 전 결합대상을 정의해야 했으나 결합률을 최적화하기에 필요한 기초자료가 부재한 점을 고려하여 2011년부터 2020년까지 각 데이터셋의 모든 관찰개체를 가명정보 결합대상으로 정의함
- 한국인터넷진흥원(결합관리기관)에서 각 자료제공기관에서 제공한 결합정보를 활용해 Key를 생성하고 결합전문기관인 국민건강보험공단에서 개인정보 비식별화 자료의 심의, 연계, 반출을 수행함
- 자료제공기관별 결합을 위해 생성된 Key는 건강보험심사평가원이 56,870,039개, 건강보험공단이 37,268,095개, 노동연구원이 34,840개였음
 - 한국노동연구원 자료에서 결합된 Key는 총 21,162개(결합률 60.74%)로 건강보험심사평가원에서 19,881개, 건보공단에서 16,591개의 결합키가 생성되었고 세 개 자료에서 모두 결합한 Key는 15,310개였음
 - 자료간 결합 Key 외에 노동패널조사의 가구단위 결합이나 건강보험 청구자료의 명세서단위 사건을 구축하는 단계에서 활용되는 식별변수도 새로 변환하여 제공됨

〈표 4〉 자료원별 결합률과 개체수

(단위: 개, %)

자료제공기관	기관간 결합	생성된 Key	결합된 Key	결합률
건강보험심사평가원(A)		56,870,039	19,881	0.03
국민건강보험공단(B)		37,268,095	16,591	0.045
한국노동연구원(C)		34,840	21,162	60.741
	C - B 결합		16,591	
	C - A 결합		19,881	
	B - C 결합		15,310	
	3개 공통결합		15,310	

2) 주요 변수

- 결합자료에 포함된 주요 변수는 아래와 같으며, 장애정도, 공상구분 등은 세부유형별 응답치가 소수이므로 개인정보 비식별화를 위해 반출 전 재범주화 하였음

〈표 5〉 가명정보결합자료의 주요 변수

	노동패널	사망	
성	노사정 위원회 기준 비정규직	사망년월	
	정규직 근로자	건강보험 자격	
남성	비정규직 근로자	건강보험 가입유형	
		건강보험 보험료(10분위)	
연령 (20세 이상) 5세 단위	OECD기준 경제활동 상태	장애등급·유형	
	임금종사자	업종세분류	
학력수준	비임금종사자	건강보험 청구자료	
	실업구직자	서식구분	
중학교 졸업 이하	비경제활동인구	주상병	
	중사상지위	부상병	
고등학교 졸업	미취업	공상구분	
	상용직	요양급여총액	
전문대 졸업	임시일용직	본인부담금	
	자영업	보험자부담금	
대학 졸업 이상	무급가족노동	요양개시일	
		요양종료일	
결혼상태	근로규칙성	진료내역	
	서울	규칙적 근로	원외처방내역
미혼	경인	불규칙적 근로	
	기혼	계절성 근로	
이혼별거사별	광역시	노동강도	
	그 외		

가구경제	주간 총 근로시간		
가구 금융자산	주간 초과근로시간		
가구 금융부채	흡연		
가구 총 근로소득		안함	
보건의료 지출		가끔함	
연간 보건의료비 지출		흡연함	
연간 건강보험료 지출	음주		노인장기요양보험
현재건강상태		안함	기본정보
보통		가끔함	요양기관
안 좋음		음주함	지불비용
좋음			요양등급

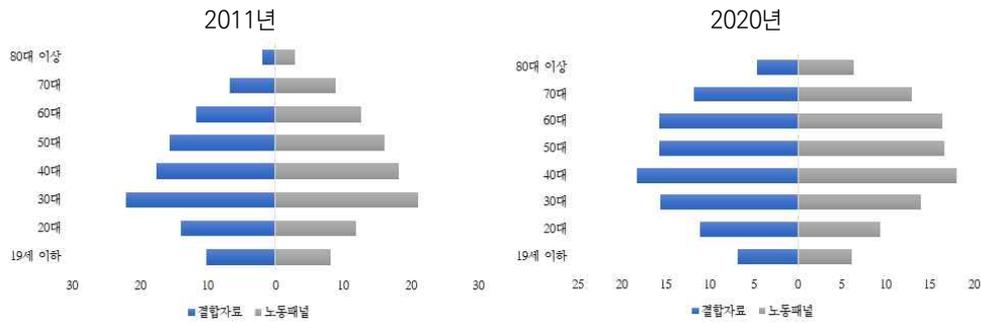
- 본 연구에서는 일반화 가능성을 점검하고 일반인구와의 의료이용 차이를 비교검토하기 위해 필요에 따라 성, 연령, 교육수준, 혼인상태 등 인구사회학적 특성과 종사상 지위, 건강보험료 등 노동·경제활동 특성, 건강보험 빅데이터에 포함된 주요 자격 및 급여 정보를 필요에 따라 선별·재가공하였음
- 연령변수는 자료의 특성과 생애주기별 의료이용의 차이의 검토에는 10세 단위 범주형을, 생애주기 특성에 따라 의료이용에 미치는 영향요인 분석에는 65세를 기준으로 이분형 범주를 활용함

2. 가명정보결합 자료의 특성

가. 결합대상자의 인구학적 특성

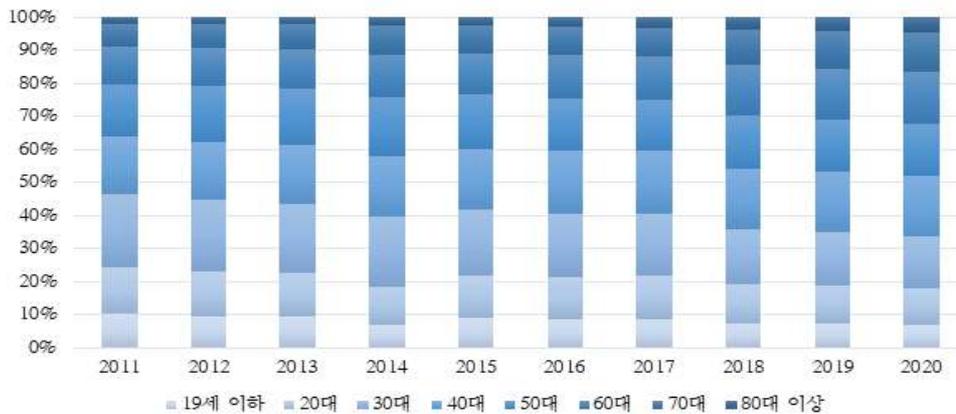
- 1998년부터 2020년까지 적어도 1회 한국노동패널조사에 참여했던 대상자 중 결합에 성공한 개체의 전체 관측치 212,586개에서 2011년 전이거나 직업력 변화로 인해 중복 생성된 경우, 건강보험 자격 정보가 결측인 경우, 사망일 이후 정보가 남아있는 경우를 제거함
- 원 자료인 한국노동패널조사에 비해 2011년과 2020년 결합자료의 10세 구간 연령 대별 분포는 비교적 유사했지만 80세 이상 인구의 결합율이 낮았고, 2011년에 비해 2020년에 중장년층의 비율이 늘어나는 고령화 현상이 관찰되었음

제3장 기명정보결합 분석결과 ●●



[그림 10] 결합자료와 한국노동패널조사의 2011년과 2020년 연령대 분포

- 동일기간 연령대별로 비율을 살펴보면, 20대 이하는 24.25%에서 17.98%로 감소했지만 60세 이상 고령자층은 20.37%에서 32.33%로 증가함
 - 20대 이하의 전체표본 대비 비율은 결합자료와 한국노동패널조사 모두 약 5%p씩 비율이 감소해 2011년에는 24.2%, 20.1%, 2020년에는 18.0%, 15.4%를 차지했고 한국노동패널보다 결합자료에서 20대 이하의 비율이 전체 표본에서 3-4%p 높음
 - 80세 이상 고령자층의 전체표본 대비 비율은 2011년 각각 1.9%, 2.9%로 1%의 차이를 보였지만 2020년에는 4.7%와 6.4%로 1.7%의 차이와 2-3%의 비율 증가함



[그림 11] 결합자료의 연령대별 분포

나. 연도별 일반적 특성

- 결합자료로 나타난 2011년부터 2020년까지 10년 간 인구사회학적 특성과 노동·경제활동 특성을 공개자료로 파악할 수 있는 결합 전 한국노동패널조사 원자료와 비교하면, 남성은 50.32%, 여성은 49.68%로 전체 조사참여자 남성(52.43%)과 여성(48.27%)의 성비와 유사했고 65세 이상 고령인구는 23.61%보다 적은 19.92%임
- 최종 졸업을 기준으로 교육수준을 구분해보면 고졸이 원 자료보다 1.84%p 적은 31.10%로 가장 많은 비율을 차지해 대졸이상, 중졸이하, 전문대 이상의 순서로 뒤를 이었으며 종사상 지위는 미취업자가 원자료보다 5.66%p 적은 38.14%로 가장 많았고 상용직, 자영업자, 임시·일용직, 무급가족종사자가 그 뒤를 이었음
- 한국노동패널조사와 건강보험 자격DB에 포함된 거주지 정보에는 연간 20명 내외의 응답 차이가 발견되었고, 자료원에 따라 서울, 경기도, 인천에 해당하는 수도권 거주자의 비율은 각각 51.47%, 51.62%를 차지하는 것으로 나타남
- 1인가구의 비율은 10년 평균 23.96%였지만 2011년 19.12%에서 2020년 28.17%로 급격한 증가했고, 등록장애인의 비율은 경증이 4.09%에서 5.38%, 중증이 1.8%에서 2.09%로 각각 4.62%, 1.96% 증가했음
- 대표적인 건강행태인 흡연과 음주를 살펴보면, 평균 과거 및 현재 흡연자는 9.08%, 19.46%로 원자료의 7.79%, 18.20%에 비해 적은 비율을 차지했고, 평균 과거 및 현재 음주자는 4.02%, 55.94%로 원자료의 3.76%, 51.96%보다 적었고, 흡연과 음주 모두 현재 상태에서 과거나 금연·금주로 응답하는 비율이 점차 늘어났음

〈표 6〉 2011 - 2016년 결합대상자의 일반적 특성

(단위: 명, %)

조사차수	14	15	16	17	18	19
조사년도	2011	2012	2013	2014	2015	2016
합계	7,117	7,136	7,089	6,770	7,134	7,218
성별						
남성	3,630 (51.0)	3,638 (51.0)	3,609 (50.9)	3,409 (50.4)	3,610 (50.6)	3,659 (50.7)
여성	3,487 (49.0)	3,498 (49.0)	3,480 (49.1)	3,361 (49.7)	3,524 (49.4)	3,559 (49.3)

제3장 기명정보결합 분석결과 ●●

연령							
64세 이하	6,232 (87.6)	6,162 (86.4)	6,039 (85.2)	5,630 (83.2)	5,905 (82.8)	5,898 (81.7)	
65세 이상	885 (12.4)	974 (13.7)	1,050 (14.8)	1,140 (16.8)	1,229 (17.2)	1,320 (18.3)	
학력수준							
초졸	938 (13.2)	931 (13.1)	926 (13.1)	910 (13.4)	909 (12.7)	905 (12.5)	
중졸	793 (11.1)	785 (11.0)	777 (11.0)	722 (10.7)	753 (10.6)	728 (10.1)	
고졸	2,355 (33.1)	2,329 (32.6)	2,276 (32.1)	2,180 (32.2)	2,218 (31.1)	2,218 (30.7)	
전문대졸	1,021 (14.4)	1,055 (14.8)	1,052 (14.8)	999 (14.8)	1,106 (15.5)	1,146 (15.9)	
대졸 이상	2,010 (28.2)	2,036 (28.5)	2,058 (29.0)	1,959 (28.9)	2,148 (30.1)	2,221 (30.8)	
결혼상태							
미혼	1,732 (24.3)	1,631 (22.9)	1,564 (22.1)	1,283 (19.0)	1,454 (20.4)	1,392 (19.3)	
기혼	4,668 (65.6)	4,763 (66.7)	4,767 (67.2)	4,719 (69.7)	4,869 (68.2)	4,987 (69.1)	
이혼·별거·사별	717 (10.1)	741 (10.4)	758 (10.7)	768 (11.3)	811 (11.4)	839 (11.6)	
노동패널조사 기준 거주지							
서울	1,609 (22.6)	1,567 (22.0)	1,528 (21.6)	1,388 (20.5)	1,420 (19.9)	1,468 (20.3)	
경기도·인천	2,133 (30.0)	2,154 (30.2)	2,143 (30.2)	2,065 (30.5)	2,169 (30.4)	2,201 (30.5)	
기타 광역시	1,371 (19.3)	1,380 (19.3)	1,392 (19.6)	1,368 (20.2)	1,421 (19.9)	1,404 (19.5)	
그 외	2,004 (28.2)	2,035 (28.5)	2,024 (28.6)	1,945 (28.7)	2,108 (29.6)	2,122 (29.4)	
결측	. (0.0)	. (0.0)	. (0.0)	4 (0.1)	16 (0.2)	23 (0.3)	
건보자격 기준 거주지							
서울	1,599 (22.5)	1,572 (22.0)	1,536 (21.7)	1,358 (20.1)	1,423 (19.9)	1,480 (20.5)	
경기도·인천	2,156 (30.3)	2,165 (30.3)	2,151 (30.3)	2,100 (31.0)	2,182 (30.6)	2,218 (30.7)	

생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역

기타 광역시	1,356 (19.1)	1,349 (18.9)	1,379 (19.5)	1,353 (20.0)	1,417 (19.9)	1,417 (19.6)
그 외	2,006 (28.2)	2,050 (28.7)	2,023 (28.5)	1,959 (28.9)	2,112 (29.6)	2,103 (29.1)
종사상지위						
미취업	2,831 (39.8)	2,790 (39.1)	2,751 (38.8)	2,563 (37.9)	2,729 (38.3)	2,661 (36.9)
상용직	2,343 (32.9)	2,401 (33.7)	2,406 (33.9)	2,321 (34.3)	2,500 (35.0)	2,585 (35.8)
임시·일용직	747 (10.5)	736 (10.3)	742 (10.5)	749 (11.1)	760 (10.7)	790 (10.9)
자영업자	994 (14.0)	999 (14.0)	981 (13.8)	943 (13.9)	955 (13.4)	1,002 (13.9)
무급가족종사자	202 (2.8)	210 (2.9)	209 (3.0)	194 (2.9)	190 (2.7)	180 (2.5)
건강보험 가입유형						
직장 가입자와 피부양자	4,918 (69.1)	5,048 (70.7)	5,087 (71.8)	4,938 (72.9)	5,271 (73.9)	5,400 (74.8)
지역 가입자와 세대원	2,003 (28.1)	1,887 (26.4)	1,812 (25.6)	1,651 (24.4)	1,676 (23.5)	1,632 (22.6)
의료급여 수급권자	196 (2.8)	201 (2.8)	190 (2.7)	181 (2.7)	187 (2.6)	186 (2.6)
1인가구 여부						
해당 없음	5,756 (80.9)	5,695 (79.8)	5,629 (79.4)	5,332 (78.8)	5,498 (77.1)	5,459 (75.6)
해당 있음	1,361 (19.1)	1,441 (20.2)	1,460 (20.6)	1,438 (21.2)	1,636 (22.9)	1,759 (24.4)
등록장애 정보						
해당 없음	6,698 (94.1)	6,699 (93.9)	6,658 (93.9)	6,343 (93.7)	6,692 (93.8)	6,780 (93.9)
중증 등록장애인	128 (1.8)	140 (2.0)	140 (2.0)	131 (1.9)	137 (1.9)	130 (1.8)
경증 등록장애인	291 (4.1)	297 (4.2)	291 (4.1)	296 (4.4)	305 (4.3)	308 (4.3)
흡연						
안 함	4,878 (68.5)	4,916 (68.9)	5,021 (70.8)	4,788 (70.7)	5,098 (71.5)	5,139 (71.2)
과거 흡연함	578 (8.1)	572 (8.0)	537 (7.6)	531 (7.8)	615 (8.6)	646 (9.0)

제3장 기명정보결합 분석결과 ●●

현재 흡연함	1,660 (23.3)	1,647 (23.1)	1,531 (21.6)	1,451 (21.4)	1,421 (19.9)	1,433 (19.9)
음주						
안 함	2,989 (42.0)	2,846 (39.9)	2,898 (40.9)	2,746 (40.6)	2,813 (39.4)	2,805 (38.9)
과거 음주함	194 (2.7)	211 (3.0)	201 (2.8)	228 (3.4)	257 (3.6)	278 (3.9)
현재 음주함	3,932 (55.3)	4,078 (57.2)	3,990 (56.3)	3,796 (56.1)	4,064 (57.0)	4,135 (57.3)

* 3명 이하 결측치가 발생한 경우 별도 표기하지 않음

〈표 7〉 2017 - 2020년 결합대상자의 일반적 특성

(단위: 명, %)

조사차수	20	21	22	23	결합자료
조사년도	2017	2018	2019	2020	합 계
합계	7,323	12,335	11,888	11,680	85,690
성별					
남성	3,743 (51.1)	6,136 (49.7)	5,898 (49.6)	5,790 (49.6)	43,122 (50.3)
여성	3,580 (48.9)	6,199 (50.3)	5,990 (50.4)	5,890 (50.4)	42,568 (49.7)
연령					
64세 이하	5,936 (81.1)	9,408 (76.3)	8,883 (74.7)	8,524 (73.0)	68,617 (80.1)
65세 이상	1,387 (18.9)	2,927 (23.7)	3,005 (25.3)	3,156 (27.0)	17,073 (19.9)
학력수준					
초졸	892 (12.2)	1,798 (14.6)	1,738 (14.6)	1,695 (14.5)	11,642 (13.6)
중졸	719 (9.8)	1,317 (10.7)	1,276 (10.7)	1,262 (10.8)	9,132 (10.7)
고졸	2,209 (30.2)	3,767 (30.5)	3,579 (30.1)	3,516 (30.1)	26,647 (31.1)
전문대졸	1,191 (16.3)	1,774 (14.4)	1,722 (14.5)	1,727 (14.8)	12,793 (14.9)
대졸 이상	2,312 (31.6)	3,679 (29.8)	3,573 (30.1)	3,480 (29.8)	25,476 (29.7)

생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역

결혼상태

미혼	1,389 (19.0)	1,982 (16.0)	1,803 (15.2)	1,694 (14.5)	15,924 (18.6)
기혼	5,076 (69.3)	8,750 (70.9)	8,493 (71.4)	8,371 (71.7)	59,463 (69.4)
이혼·별거·사별	858 (11.7)	1,603 (13.0)	1,592 (13.4)	1,615 (13.8)	10,302 (12.0)

노동패널조사 기준 거주지

서울	1,470 (20.1)	2,512 (20.4)	2,402 (20.2)	2,325 (19.9)	17,689 (20.6)
경기도·인천	2,289 (31.3)	3,847 (31.2)	3,729 (31.4)	3,682 (31.5)	26,412 (30.8)
기타 광역시	1,408 (19.2)	2,361 (19.1)	2,280 (19.2)	2,233 (19.1)	16,618 (19.4)
그 외	2,120 (29.0)	3,554 (28.8)	3,413 (28.7)	3,377 (28.9)	24,702 (28.8)
결측	36 (0.5)	61 (0.5)	64 (0.5)	63 (0.5)	269 (0.3)

건보자격 기준 거주지

서울	1,483 (20.3)	2,533 (20.5)	2,426 (20.4)	2,340 (20.0)	17,750 (20.7)
경기도·인천	2,288 (31.2)	3,823 (31.0)	3,727 (31.4)	3,671 (31.4)	26,481 (30.9)
기타 광역시	1,429 (19.5)	2,371 (19.2)	2,275 (19.1)	2,226 (19.1)	16,572 (19.3)
그 외	2,123 (29.0)	3,608 (29.3)	3,460 (29.1)	3,443 (29.5)	24,887 (29.0)

종사상지위

미취업	2,627 (35.9)	4,723 (38.3)	4,502 (37.9)	4,508 (38.6)	32,685 (38.1)
상용직	2,666 (36.4)	4,160 (33.7)	4,096 (34.5)	4,032 (34.5)	29,510 (34.4)
임시·일용직	839 (11.5)	1,428 (11.6)	1,387 (11.7)	1,296 (11.1)	9,474 (11.1)
자영업자	1,001 (13.7)	1,694 (13.7)	1,602 (13.5)	1,563 (13.4)	11,734 (13.7)
무급가족종사자	190 (2.6)	330 (2.7)	301 (2.5)	281 (2.4)	2,287 (2.7)

건강보험 가입유형

제3장 가명정보결합 분석결과 ●●

직장 가입자와 피부양자	5,541 (75.7)	9,071 (73.5)	8,730 (73.4)	8,633 (73.9)	62,637 (73.1)
지역 가입자와 세대원	1,595 (21.8)	2,869 (23.3)	2,786 (23.4)	2,675 (22.9)	20,586 (24.0)
의료급여 수급권자	187 (2.6)	395 (3.2)	372 (3.1)	372 (3.2)	2,467 (2.9)
1인가구 여부					
해당 없음	5,480 (74.8)	9,242 (74.9)	8,679 (73.0)	8,390 (71.8)	65,160 (76.0)
해당 있음	1,843 (25.2)	3,093 (25.1)	3,209 (27.0)	3,290 (28.2)	20,530 (24.0)
등록장애 정보					
해당 없음	6,871 (93.8)	11,467 (93.0)	11,033 (92.8)	10,808 (92.5)	80,049 (93.4)
중증 등록장애인	138 (1.9)	249 (2.0)	246 (2.1)	244 (2.1)	1,683 (2.0)
경증 등록장애인	314 (4.3)	619 (5.0)	609 (5.1)	628 (5.4)	3,958 (4.6)
흡연					
안 함	5,210 (71.2)	8,937 (72.5)	8,677 (73.0)	8,560 (73.3)	61,224 (71.4)
과거 흡연함	675 (9.2)	1,267 (10.3)	1,200 (10.1)	1,161 (9.9)	7,782 (9.1)
현재 흡연함	1,438 (19.6)	2,127 (17.2)	2,011 (16.9)	1,957 (16.8)	16,676 (19.5)
음주					
안 함	2,771 (37.8)	4,917 (39.9)	4,781 (40.2)	4,733 (40.5)	34,299 (40.0)
과거 음주함	312 (4.3)	583 (4.7)	600 (5.1)	584 (5.0)	3,448 (4.0)
현재 음주함	4,240 (57.9)	6,831 (55.4)	6,507 (54.7)	6,360 (54.5)	47,933 (55.9)

* 3명 이하 결측치가 발생한 경우 별도 표기하지 않음

- 경제활동 참여자의 주당 근로시간과 초과근로시간의 평균은 2011년부터 2020년까지 40.7시간에서 36.5시간(표준편차 15.2~18.9)으로, 9.2시간에서 6.9시간(표준편차 4.8~7.7)으로 꾸준히 감소했음

〈표 8〉 2011 - 2020 결합자료의 노동시간 특성

(단위: 시간)

	주당 평균 근로시간					주당 평균 초과근로시간				
	Obs	Mean	Std.	Min	Max	Obs	Mean	Std.	Min	Max
2011	3,090	40.7	18.9	1	168	775	9.2	6.8	0	56
2012	3,136	39.7	18.4	1	144	757	8.9	6.1	0	50
2013	3,147	39.6	18.4	1	168	762	9.3	7.4	0	60
2014	3,070	40.8	17.9	1	143	824	9.5	7.7	0	71
2015	3,258	39.6	17.5	1	144	807	9.0	6.8	1	50
2016	3,375	39.5	16.6	1	104	832	8.6	6.7	0	50
2017	3,505	39.1	16.8	1	140	769	8.7	7.1	0	68
2018	5,585	37.0	17.0	1	140	938	8.0	7.7	0	100
2019	5,482	36.9	15.5	1	100	715	7.5	6.0	0	60
2020	5,325	36.5	15.2	1	136	555	6.9	4.8	0	32

- 가구 총 근로소득의 평균, 표준편차, 최소값, 최대값, 10분위를 살펴보면, 2011년부터 2020년까지 연도별 가구단위 근로소득의 격차는 줄어들면서 전반적인 소득분위가 상승해 평균 근로소득도 증가했으며 최대값은 97% 상단코딩이 적용되어 매해 동일했음

〈표 9〉 2011 - 2020년 결합자료의 가구 근로소득 특성

(단위: 만 원)

	연간 가구 총 근로소득				
	Obs	Mean	Std.	Min	Max
2011	6,418	4,095.0	3,700.3	20	18,160
2012	6,412	4,284.1	2,778.0	6	18,160
2013	6,355	4,448.6	2,910.5	6	18,160
2014	6,017	4,566.4	2,956.4	6	18,160
2015	6,342	4,759.5	3,085.9	3	18,160
2016	6,367	4,839.5	3,071.7	10	18,160
2017	6,482	5,053.1	3,258.1	10	18,160
2018	10,522	4,846.3	3,189.6	50	18,160
2019	10,123	5,006.5	3,260.3	26	18,160
2020	9,953	5,143.7	3,330.6	10	18,160

〈표 10〉 2011 - 2020 결합자료의 연간 가구 총 근로소득 10분위 경계값

(단위: 만 원)

차수	14	15	16	17	18	19	20	21	22	23
년도	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
1분위	1,200	1,200	1,200	1,300	1,440	1,440	1,445	1,200	1,320	1,392
2분위	1,890	2,040	2,040	2,160	2,352	2,400	2,400	2,300	2,400	2,400
3분위	2,400	2,620	2,720	2,800	2,990	3,000	3,120	3,000	3,120	3,246
4분위	3,000	3,221	3,320	3,400	3,600	3,600	3,840	3,600	3,826	3,960
5분위	3,600	3,680	3,890	4,000	4,200	4,320	4,620	4,400	4,644	4,800
6분위	4,200	4,400	4,680	4,800	4,900	5,040	5,320	5,160	5,400	5,500
7분위	4,800	5,160	5,400	5,600	5,817	6,000	6,060	6,000	6,200	6,396
8분위	6,000	6,072	6,360	6,600	6,960	7,020	7,300	7,200	7,400	7,640
9분위	7,440	7,800	8,160	8,508	9,000	3,000	9,506	9,240	9,480	9,600
10분위	18,160	18,160	18,160	18,160	18,160	18,160	18,160	18,160	18,160	18,160
대상자	6,418	6,412	6,355	6,017	6,342	6,367	6,482	10,522	10,123	9,953

- 노동패널조사에서 자가응답으로 수집한 문항과 국민건강보험제도의 운영을 위해 수집·관리되고 있는 자격 및 진료에 대한 건강보험 빅데이터의 건강보험료와 보건의료비를 비교함
- 2011년부터 2020년까지 연간 건강보험료는 노동패널과 빅데이터 기준 평균 10.9만원, 10.7만원으로 연평균 증가율은 각각 3.4%, 4.5%이었고, 보건의료비(법정본인부담금) 지출금액은 평균 6.7만원, 32.9만원으로 연평균 증가율은 각각 2.2%, 8.4%를 기록해 건강보험에 비해 보건의료비에서 자료원에 따른 응답의 차이가 컸음

〈표 11〉 2011 - 2020년 결합자료의 건강보험료 특성

(단위: 만 원)

	노동패널 기준 연간 건강보험료					빅데이터 기준 연간 건강보험료				
	Obs	Mean	Std.	Min	Max	Obs	Mean	Std.	Min	Max
2011	7,117	9.09	8.18	0	110	6,907	8.66	6.63	0	59.5
2012	7,101	9.73	8.99	0	90	6,914	9.27	7.31	0	114.1
2013	7,043	9.89	8.74	0	80	6,871	9.69	8.14	0	219
2014	6,753	10.42	9.66	0	100	6,657	10.10	7.94	0	106.6

생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역

2015	7,097	10.88	9.72	0	100	6,997	10.50	8.63	0	225.8
2016	7,213	11.51	10.90	0	164	7,078	10.85	8.86	0	227.8
2017	7,323	12.09	10.62	0	120	7,188	11.27	9.53	0	239
2018	12,314	11.44	10.48	0	100	12,128	11.71	9.74	0	243.7
2019	11,888	11.86	11.46	0	150	11,677	12.09	10.44	0	184.6
2020	11,680	12.25	11.27	0	117	11,464	12.90	10.99	0	189.6

〈표 12〉 2011 - 2020년 결합자료의 보건의료비 지출 특성

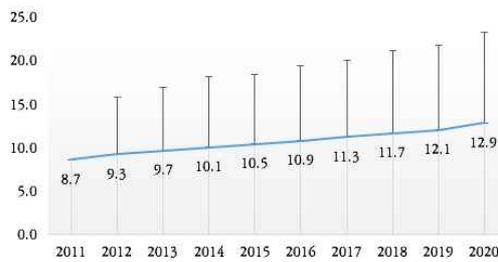
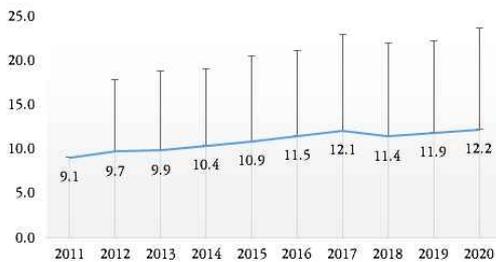
(단위: 만 원)

	노동패널 기준 연간 보건의료비					빅데이터 기준 연간 보건의료비				
	Obs	Mean	Std.	Min	Max	Obs	Mean	Std.	Min	Max
2011	7,117	6.07	8.18	0	520	7,117	22.31	37.16	0	494
2012	7,123	6.15	8.99	0	300	7,136	23.62	40.27	0	859
2013	7,079	6.45	8.74	0	250	7,089	25.43	45.41	0	1,119
2014	6,770	6.46	9.66	0	250	6,770	27.43	46.35	0	1,383
2015	7,134	6.19	9.72	0	200	7,134	29.43	48.51	0	940
2016	7,218	6.66	10.90	0	350	7,218	33.47	59.24	0	1,523
2017	7,323	7.10	10.62	0	350	7,323	35.66	61.64	0	1,285
2018	12,333	7.19	10.48	0	650	12,335	40.12	64.40	0	1,621
2019	11,888	7.52	11.46	0	359	11,888	44.94	70.75	0	1,180
2020	11,680	7.36	11.27	0	1,000	11,680	46.22	79.72	0	1,796

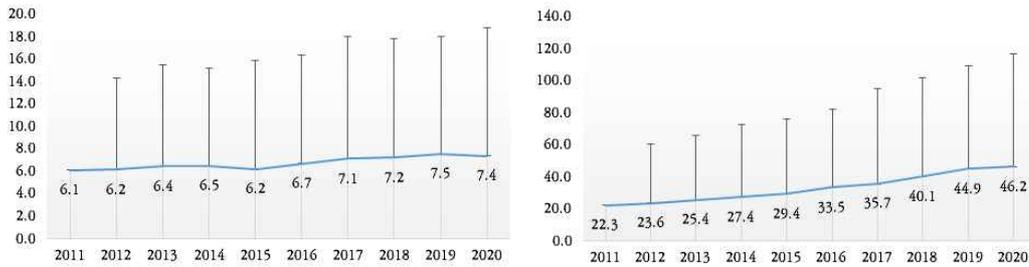
노동패널조사자료

건강보험 빅데이터

건강보험료



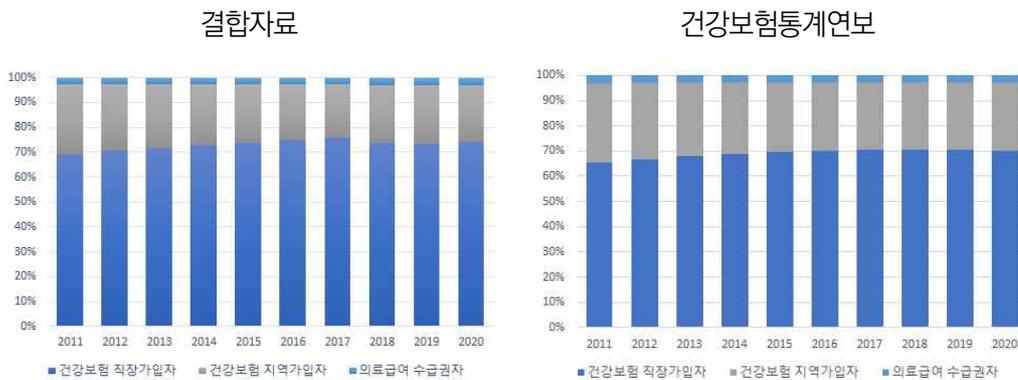
보건의료비



[그림 12] 자료원별·연도별 건강보험료 및 보건의료비 추이 (단위: 만원)

다. 건강 및 의료이용 관련 특성

- 건강보장 유형을 건강보험 직장가입자(피부양자 포함), 건강보험 지역가입자(세대주 및 세대원 포함), 의료급여 수급권자로 구분해 2011년부터 2020년 결합자료와 건강보험통계연보를 비교해보면, 결합자료의 건강보험 직장가입자의 비율은 2011년 69.1%, 2020년 73.9%로 건강보험통계연보에 비해 3.8%p, 3.7%p 많았고 지역가입자의 비율은 2011년 28.2%, 2020년 22.9%로 통계연보에 비해 3.4%p, 3.9%p 적었음
- 건강보험통계연보로 산출한 전인구 내 의료급여 수급권자 비율은 연도별 변동이 있었지만 결합자료 내 의료급여 수급권자 비율은 시간의 흐름에 따라 증가했음



[그림 13] 자료원별·연도별 건강보장 유형

- 노인장기 요양보험 인정조사 신청자는 전체의 약 0.93%였고 65세 이상 고령 신청자의 비율은 88.9%에서 98.1%로 증가했으며 해당 기간 연도별 결합자료 대비 등급인정자의 비율은 0.69% 수준이었지만 신청건수 대비 인정률은 74.7%로 10년간 연평균 인정조사 신청자와 장기요양등급 인정자의 증가율은 각각 21.43%, 25.88%였음
- 건강보험과 달리 노인장기 요양보험으로 장기요양서비스를 이용하고자 하는 자는 인정조사를 신청해 장기요양등급을 판정받아야 하고, 등급의 판정이나 인정조사를 철회하지 아니하고 시설급여 혹은 재가급여를 이용할 수 있는 1등급부터 6등급까지 등급을 받아야 하므로 이에 해당하는 장기요양등급 인정자를 별도로 산출함

〈표 13〉 2011 - 2020년 노인장기요양보험등급 신청 및 인정 현황

(단위: 명, %)

년도	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
전체 표본	7,117	7,136	7,089	6,770	7,134	7,218	7,117	7,136	7,089	6,770
인정조사 신청자	27	31	36	47	50	61	80	133	173	155
(고령 신청자 비율)	88.9	96.8	88.9	93.6	96.0	91.8	95.0	96.2	96.5	98.1
장기요양등급 인정자	15	21	25	38	35	47	60	102	130	119
(신청 대비 인정률)	55.6	67.7	69.4	80.9	70.0	77.0	75.0	76.7	75.1	76.8

- 결합대상자 중 2011년부터 2020년 사이 사망정보가 연계된 자는 총 429명, 해당기간 조사에 참여했던 자 중 사망정보가 연계된 자는 총 371명으로 한국노동연구원이 발표한 노동패널 참여자 중 사망자 수보다 적었음

〈표 14〉 2011 - 2020년 집계방식별 사망자 수 비교

(단위: 명)

년도	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	N
조사기관 집계	87	82	90	51	82	73	67	96	68	91	787
결합대상 전체	3	15	21	25	30	39	45	72	85	94	429
해당연도 참여자	2	13	17	22	23	32	37	63	79	83	371

제3장 기명정보결합 분석결과 ●●

- 10세 단위 연령대별 2011년과 2020년 외래와 입원 의료서비스를 이용한 자의 다빈도 상병을 5순위까지 산출하여 비교함
 - 다빈도 순위를 산출하기 위한 기준은 명세서 청구건별 주상병의 빈도로 외래와 입원 의료서비스를 이용할 당시 환자의 부상병(동반상병), 질환의 중증도, 의료이용 강도, 의료비 부담정도 등은 고려하지 아니하였으며, 질환을 특정할 수 없이 마스킹된 상병이나 최소 4번 이하 입원으로 개인식별 위험이 높은 상병은 순위에서 배제함
- 2011년 대비 2020년 상병 유형과 상병당 외래방문 횟수는 모두 증가한 가운데, 50대 이하 청장년층의 외래이용 주요 양상은 치과 관련 질환(매물치 및 매목치(K01), 치아우식(K02), 치수 및 근단주위조직의 질환(K04), 치은염 및 치주질환(K05)의 순위가 상승하였고 임신부의 산전관리를 위한 정상임신의 관리(Z34)가 20대 다빈도상병 순위에서 사라지고 30대에서도 3순위에서 5순위로 하락함
- 50대 이상 중고령층부터 상·하기도감염(J20, J30)에 의한 의료이용보다 등통증(M54), 요추탈구(S33), 척추병증(M48), 무릎관절증(M17)처럼 근골격계 질환이나 고혈압(I10), 당뇨(E11), 만성신장병(N18)처럼 만성질환에 의한 외래이용이 증가함

〈표 15〉 연령대별 2011년, 2020년 외래의료서비스 이용 시 다빈도 주상병 Top 5

		20대 미만	20대	30대	40대	50대	60대	70대	80대 이상
2011	1순위	K04	Z34	J20	M54	I10	I10	M54	M17
	2순위	J20	J20	M54	I10	M54	M17	I10	I10
	3순위	S93	K04	Z34	K05	K05	E11	M17	M54
	4순위	K02	M54	K04	J20	M75	N18	E11	E11
	5순위	M54	J03	K05	K04	J20	M75	M48	M48
2020	1순위	K05	K05	K05	K05	I10	I10	M54	I10
	2순위	J20	M54	M54	I10	K05	M54	I10	M54
	3순위	J30	K02	J20	M54	M54	E11	M17	E11
	4순위	K01	J20	S33	N18	E11	M17	E11	M48
	5순위	M54	K01	Z34	S33	N18	K05	M48	M17

- 외래와 마찬가지로 입원한 상병의 유형과 상병당 횟수가 크게 증가한 가운데, 2011년 대비 2020년 다빈도 질환으로 암질환 관련 신생물(구강암(C05), 위암(C16), 결장암(C18), 직장암(C20), 폐암(C34), 유방암(C50))이 여러 연령대에서 다양한 유형으로 나타남
 - 명세서 기준으로 산출한 것이므로 노동패널 표본 대비 암진단을 받은 환자의 수는 후속연구를 통해 밝혀져야 할 부분으로, 2011년 20대 5순위 상병이었으나 2020년에는 50대와 60대에서 각각 4위, 5위를 기록한 유방암처럼 환자의 치료 혹은 사망에 따라 연도별 암종류에 따른 입원빈도가 변동할 것을 고려해야 함
- 외래에서와 같이, 임신부의 출산을 위한 자연분만(O80)과 제왕절개(O82)는 2011년 20대와 30대에서 나란히 2순위, 3순위, 1순위, 2순위를 기록한 것과 대조적으로 2020년에는 자연분만만 각각 4순위와 1순위를 차지하였고, 2011년 유일한 20대 미만 다빈도 상병을 기록한 무월경(N91)은 2020년 허리(S33), 손(S62), 발(S82) 부위의 손상 혹은 외상으로 인한 입원이 증가함에 따라 순위에서 제외됨
- 2011년과 2020년 뇌경색증(I63)과 기타 뇌혈관질환(I67)을 제친 중고령층의 입원 치료 목적은 60대와 70대에선 백내장(H25)이었고 80세 이상은 치매(F00)였으며, 완화의료(Z51)를 위한 입원이 60대와 70대에서 새로 진입함

〈표 16〉 연령대별 2011년, 2020년 입원의료서비스 이용 시 다빈도 주상병 Top 5

		20대 미만	20대	30대	40대	50대	60대	70대	80대 이상
2011	1순위	N91	I69	O80	G81	C16	H25	H25	I63
	2순위	-	O80	O82	M51	I69	C20	C20	F00
	3순위	-	O82	G81	M54	K21	I20	S22	M48
	4순위	-	I84	I84	S83	M51	M51	M17	I10
	5순위	-	C50	M51	S32	I63	C16	J18	G30
2020	1순위	S33	K64	O80	M51	M51	H25	H25	F00
	2순위	K35	K21	M51	M75	M75	M48	I63	I63
	3순위	A09	M54	A09	M54	S82	Z51	Z51	S72
	4순위	S62	O80	G82	I67	C50	I63	C34	M48
	5순위	S82	A09	K60	C05	C18	C50	N18	G20

제3장 가명정보결합 분석결과 ●●

- 의료서비스 유형을 전체, 외래, 입원, 한방, 외래처방으로 나누어 해당 서비스의 이용 여부를 살펴보면, 연간 적어도 한 번 이상 의료서비스를 이용한 자는 전체의 86.53%에서 89.11%로 증가했고, 외래서비스 이용경험 비율도 이와 동일했으며 외래약제처방 경험율은 82.86%에서 86.0%로 증가함

〈표 17〉 연도별 의료서비스 유형별 이용 여부

(단위: 명, %)

		2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
전체	미이용	959	939	896	818	899	844	870	1,343	1,295	1,391
		13.47	13.16	12.64	12.08	12.6	11.69	11.88	10.89	10.89	11.91
	이용	6,158	6,197	6,193	5,952	6,235	6,374	6,453	10,992	10,593	10,289
		86.53	86.84	87.36	87.92	87.4	88.31	88.12	89.11	89.11	88.09
외래	미이용	959	939	896	818	899	844	870	1,343	1,295	1,391
		13.47	13.16	12.64	12.08	12.6	11.69	11.88	10.89	10.89	11.91
	이용	6,158	6,197	6,193	5,952	6,235	6,374	6,453	10,992	10,593	10,289
		86.53	86.84	87.36	87.92	87.4	88.31	88.12	89.11	89.11	88.09
입원	미이용	6,363	6,275	6,268	5,923	6,210	6,190	6,303	10,495	10,056	9,972
		89.41	87.93	88.42	87.49	87.05	85.76	86.07	85.08	84.59	85.38
	이용	754	861	821	847	924	1,028	1,020	1,840	1,832	1,708
		10.59	12.07	11.58	12.51	12.95	14.24	13.93	14.92	15.41	14.62
한방	미이용	5,175	5,155	4,981	4,726	5,093	5,209	5,269	8,729	8,366	8,502
		72.71	72.24	70.26	69.81	71.39	72.17	71.95	70.77	70.37	72.79
	이용	1,942	1,981	2,108	2,044	2,041	2,009	2,054	3,606	3,522	3,178
		27.29	27.76	29.74	30.19	28.61	27.83	28.05	29.23	29.63	27.21
외래처방	미이용	1,220	1,177	1,135	1,048	1,155	1,084	1,105	1,692	1,664	1,858
		17.14	16.49	16.01	15.48	16.19	15.02	15.09	13.72	14.00	15.91
	이용	5,897	5,959	5,954	5,722	5,979	6,134	6,218	10,643	10,224	9,822
		82.86	83.51	83.99	84.52	83.81	84.98	84.91	86.28	86.00	84.09
총합		7,117	7,136	7,089	6,770	7,134	7,218	7,323	12,335	11,888	11,680
		100	100	100	100	100	100	100	100	100	100

- 개인별 연간 외래와 입원 이용여부를 살펴보면, 외래 및 입원 의료서비스 평균 이용율은 87.9%, 13.3%로 연평균 이용율은 0.2%, 3.65% 증가했고, 종사상 지위에 따른

비율은 미취업자(33.6%, 6.4%), 상용직 근로자(30.0%, 3.3%), 자영업자(12.3%, 1.9%), 임시일용직 근로자(9.7%, 1.3%), 무급가족종사자(2.4%, 0.4%) 순이었음

〈표 18〉 연도별 종사상 지위에 따른 외래서비스 이용 비율 추이

(단위: %)

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
외래 이용										
미취업	34.5	34.0	33.9	33.3	33.6	32.7	31.8	33.9	33.8	34.1
상용직 근로자	28.3	28.8	29.3	29.8	30.2	31.2	31.7	29.8	30.4	30.0
임시·일용직 근로자	9.0	9.1	9.3	9.8	9.3	9.8	10.2	10.5	10.4	10.0
자영업자	12.3	12.4	12.3	12.5	12.0	12.5	12.2	12.5	12.2	11.9
무급가족종사자	2.5	2.6	2.6	2.6	2.3	2.2	2.3	2.4	2.3	2.1
외래 미이용자	13.5	13.2	12.6	12.0	12.6	11.7	11.9	10.9	10.9	11.9
총 합	100	100	100	100	100	100	100	100	100	100

〈표 19〉 연도별 종사상 지위에 따른 입원서비스 이용 비율 추이

(단위: %)

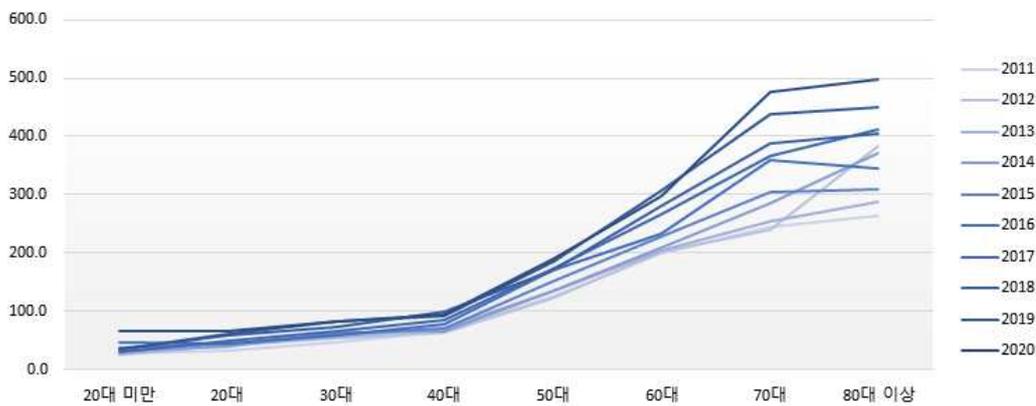
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
입원 이용										
미취업	5.2	5.7	5.5	5.9	6.2	6.6	6.4	7.2	7.7	7.2
상용직 근로자	2.4	3.1	3.3	3.1	3.0	3.8	3.9	3.5	3.3	3.4
임시·일용직 근로자	0.9	1.1	0.9	1.4	1.2	1.4	1.4	1.6	1.8	1.6
자영업자	1.8	1.8	1.6	1.7	2.1	2.1	1.9	2.1	2.1	2.1
무급가족종사자	0.4	0.3	0.4	0.4	0.4	0.4	0.3	0.5	0.5	0.3
입원 미이용자	89.4	87.9	88.4	87.5	87.0	85.8	86.1	85.1	84.6	85.4
총 합	100	100	100	100	100	100	100	100	100	100

라. 의료비 지출

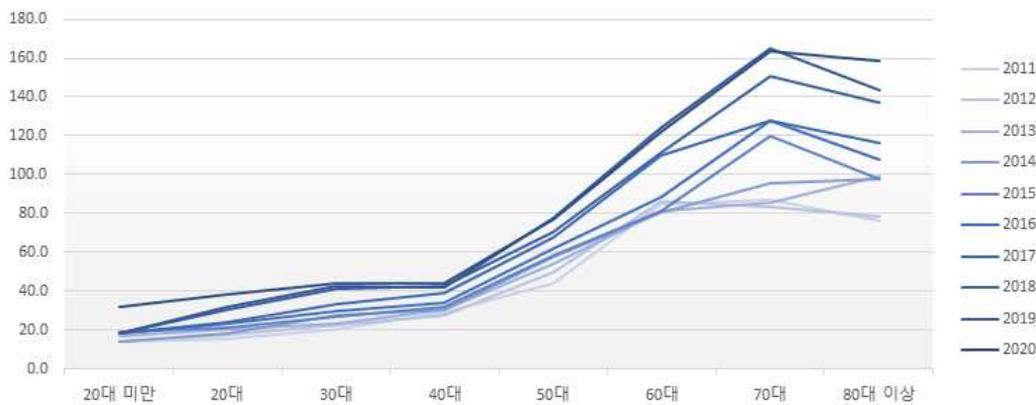
- 연간 단 한 번도 의료이용 이력이 없는 자를 포함해 전체 표본의 연도별·연령대별 만원 단위 평균 요양급여총액을 모든 유형의 의료서비스와 병·의원급 의료기관의 외래 및 입원 의료서비스로 나누어 각각 산출함
- 청구서식에 따라 한방과 약제를 포함해 외래에서 정액제로 분류하는 일부 정신 및 투석 관련 서비스는 외래 및 입원 비용에서 제외되어 연도별·연령대별 요양급여총액은 외래와 입원 총액보다 높음

제3장 가명정보결합 분석결과 ●●

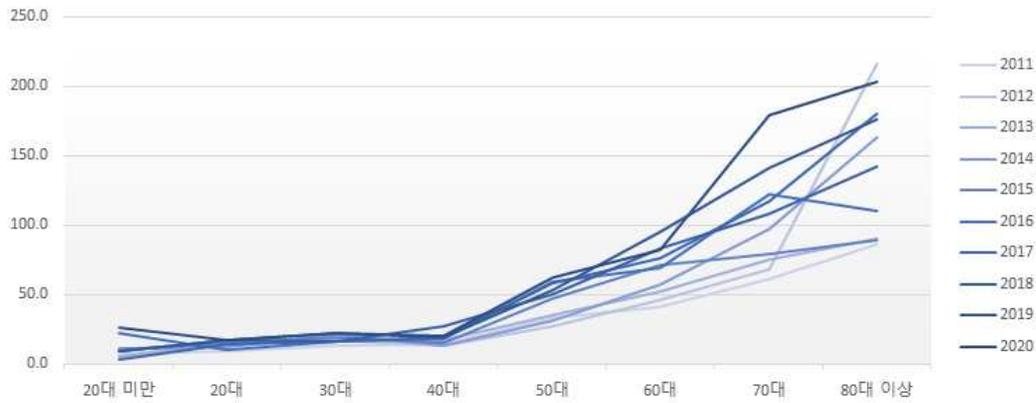
- 10년간 평균 요양급여총액은 모든 연령대에서 증가해 총액은 40대 이하보다 40대 이후에서 고령일수록 증가폭이 가파르게 커졌지만, 연평균 증가율은 20대가 8.0%로 가장 높았고 70대(7.7%)와 80대 이상(7.4%) 고령 집단 순서로 높았음
- 치과, 한방, 외래처방을 제외한 의원, 병원, 보건기관에서 청구된 외래와 입원 요양급여 총액은 50대 이상 인구집단에서 점차 증가했고 외래와 입원 이용률에 따라 외래이용은 70대에서 가장 높고 80대 이상부터 소폭 감소한 반면, 입원 이용률은 10-15%로 낮고 연도에 따라 80대 이상의 의료비 지출이 70대보다 높았음



[그림 14] 연도별·연령대별 전체 의료이용에 따른 요양급여총액 (단위: 만 원)

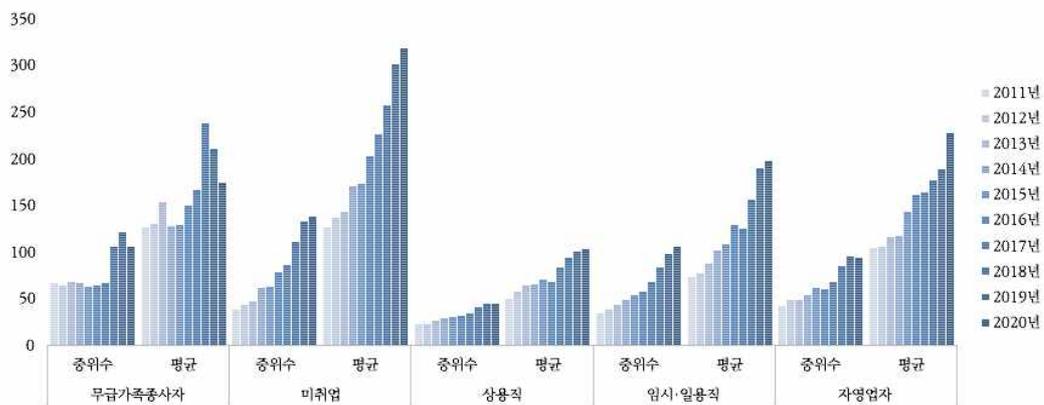


[그림 15] 연도별·연령대별 외래서비스 이용에 따른 요양급여총액 (단위: 만 원)



[그림 16] 연도별·연령대별 입원서비스 이용에 따른 요양급여총액 (단위: 만 원)

- 모든 유형을 포함한 전체 의료서비스의 연간 요양급여총액은 극단치에 강건한 중위수가 평균보다 적었고, 이러한 차이는 종사상 지위에 따라 상용직과 임시·일용직 근로자보다 자영업자, 무급가족종사자, 미취업자에게서 더 컸고 외래와 달리 입원에 따른 의료비 중위수는 항상 0원임



[그림 17] 연도별·종사상지위별 전체 요양급여총액 (단위: 만 원)

3. 생애주기별 인구집단 특성에 따른 의료이용과 의료비 지출

- 의료서비스를 이용하는 원인질환과 의료비 지출의 변화가 가장 두드러진 2018년부터 2020년까지 3개년도 자료로 의료서비스의 유형별 이용여부가 인구집단의 특성에 따른 차이가 존재하는지 검토하고자 로지스틱 고정효과 패널회귀분석을 수행함
- 모든 의료이용자는 적어도 1회 이상 외래 의료서비스를 이용하기 때문에 전체 의료서비스와 외래 의료이용서비스의 분석결과가 동일하게 산출됨
- 인구사회학적 특성, 종사상지위와 건강보험 자격, 건강행태에 따라 살펴보면, 외래이용은 교육, 결혼, 종사상 지위에 따른 차이가 있다고 볼 수 없지만 2018년 대비 2020년 외래이용 확률은 0.65배였음
- 2018년부터 3년간 증가한 연도별 표본에서 미취업자의 입원이용율이 상승했는데, 입원의료서비스를 이용할 확률은 미취업자에 비해 상근직 근로자가 0.72배, 임시일용직 근로자가 0.7배, 금주자에 비해 현재 음주자가 0.63배였음

〈표 20〉 2018-2020년 결합자료의 의료이용 유형별 이용여부 영향요인

		전체 의료서비스		외래 의료서비스		입원 의료서비스	
		OR	표준오차	OR	표준오차	OR	표준오차
연령	(ref. 64세 이하)						
	65세 이상	1.42	0.6186	1.42	0.6186	1.38	0.1796
교육수준	(ref. 고졸 이하)						
	전문학사 이상	0.00	647.1	0.00	647.1	3.11	0.8777
결혼상태	(ref. 미혼)						
	기혼	1.41	0.6385	1.41	0.6385	2.13	0.4095
	이혼·별거·사별	0.70	1.0533	0.70	1.0533	1.45	0.4861
종사상지위	(ref. 미취업자)						
	상근직	1.14	0.2097	1.14	0.2097	0.72***	0.1249
	임시·일용직	1.09	0.2625	1.09	0.2625	0.7***	0.1215
	자영업자	0.70	0.3495	0.70	0.3495	0.72*	0.1795
	무급가족종사자	0.00	523.20	0.00	523.20	0.83	0.3467
흡연	(ref. 비흡연)						
	과거 흡연자	0.48	0.6129	0.48	0.6129	1.00	0.2462
	현재 흡연자	0.56	0.4584	0.56	0.4584	0.81	0.2493
음주	(ref. 비음주)						
	과거 음주자	1.22	0.8095	1.22	0.8095	0.74	0.2227

생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역

	현재 음주자	0.58	0.386	0.58	0.386	0.63***	0.1777
건강보험	(ref. 직장가입자)						
	지역가입자	0.77	0.171	0.77	0.171	1.04	0.0939
	의료급여 수급권자	0.24	1.2201	0.24	1.2201	0.89	0.3475
시간	(ref. 2018년)						
	2019년	0.92	0.0841	0.92	0.0841	1.05	0.0411
	2020년	0.65***	0.0832	0.65***	0.0832	0.98	0.0429

p<0.1, **p<0.05, *p<0.01

- 여성에 한정해 외래와 입원 이용에 영향을 미치는 요인을 살펴보면, 전체 표본과 마찬가지로 2018년에 비해 2020년 외래이용 확률은 0.59배였고 입원이용은 미취업자보다 상근직이 0.71배, 임시일용직이 0.55배로 유의했음

〈표 21〉 2018-2020년 여성의 의료이용 유형별 이용여부 영향요인

		외래 이용		입원 이용	
		OR	표준오차	OR	표준오차
연령	(ref. 64세 이하)				
	65세 이상	0.00	1512.7	1.27	0.257
교육수준	(ref. 고졸 이하)				
	전문학사 이상	0.00	2202.9	1.72	0.959
결혼상태	(ref. 미혼)				
	기혼	7,799,988	1490.8	2.47*	0.516
	이혼·별거·사별	4.88	2172.1	1.74	0.606
종사상지위	(ref. 미취업자)				
	상근직	1.50	0.340	0.71**	0.161
	임시·일용직	0.94	0.410	0.55***	0.159
	자영업자	0.80	0.556	0.79	0.257
	무급가족종사자	0.00	1053.9	0.90	0.396
흡연	(ref. 해당없음)				
	과거 흡연자	.	.	0.77	0.889
	현재 흡연자	.	.	0.62	0.720
음주	(ref. 해당없음)				
	과거 음주자	1.69	1.327	1.11	0.316
	현재 음주자	0.82	0.591	0.71	0.231
건강보험	(ref. 직장가입자)				
	지역가입자	0.55	0.315	1.00	0.125
	의료급여 수급권자	0.00	2202.9	0.87	0.519
시간	(ref. 2018년)				
	2019년	1.03	0.152	1.02	0.057
	2020년	0.59***	0.144	0.92	0.059

p<0.1, **p<0.05, *p<0.01

제3장 가명정보결합 분석결과 ●●

- 남성의 외래와 입원 이용에 영향을 미치는 요인도 전체 표본과 같이 2018년에 비해 2020년 외래이용 확률은 0.69배였지만 입원이용에서 종사상지위에 따른 차이는 나타나지 않았고, 금주자보다 과거 음주자가 0.49배, 현재 음주자가 0.51배였음

〈표 22〉 2018-2020년 남성의 의료이용 유형별 이용여부 영향요인

		외래 의료서비스		입원 의료서비스	
		OR	표준오차	OR	표준오차
연령	(ref. 64세 이하)				
	65세 이상	2.23	0.729	1.49	0.252
교육수준	(ref. 고졸 이하)				
	전문학사 이상	0.00	444.7	97,971	197.4
결혼상태	(ref. 미혼)				
	기혼	0.79	0.710	1.61	0.694
	이혼·별거·사별	0.51	1.130	1.09	0.836
종사상지위	(ref. 미취업자)				
	상근직	0.98	0.268	0.75	0.200
	임시일용직	1.26	0.349	0.98	0.194
	자영업자	0.62	0.461	0.68	0.252
	무급가족종사자	0.00	545.3	0.58	0.713
흡연	(ref. 해당없음)				
	과거 흡연자	0.50	0.611	1.10	0.261
	현재 흡연자	0.59	0.459	0.88	0.269
음주	(ref. 해당없음)				
	과거 음주자	0.85	1.047	0.49***	0.330
	현재 음주자	0.47	0.537	0.51***	0.282
건강보험	(ref. 직장가입자)				
	지역가입자	0.88	0.206	1.09	0.143
	의료급여 수급권자	0.31	1.308	0.94	0.472
시간	(ref. 2018년)				
	2019년	0.88	0.102	1.10	0.060
	2020년	0.69***	0.102	1.04	0.063

* *p<0.1, **p<0.05, ***p<0.0

- 의료서비스 유형별 의료이용에 따른 요양급여총액에 미치는 영향요인을 살펴보기 위해 표본의 특성과 의료비 변수의 분포를 고려해 일반화선형모형 회귀분석을 수행함
- 의료이용 여부를 파악하기 위한 로지스틱 고정효과 패널분석과 다르게 분석모형에 활용된 요인들은 통계적으로 유의미한 영향을 미치고 있었고 그 중 성별에 따른

요양급여총액은 외래서비스의 경우 남성보다 여성이 0.9% 높았지만 입원서비스는 남성보다 여성이 1.1% 낮아 의료서비스의 유형에 따라 성별 의료비 부담에 차이가 있었음

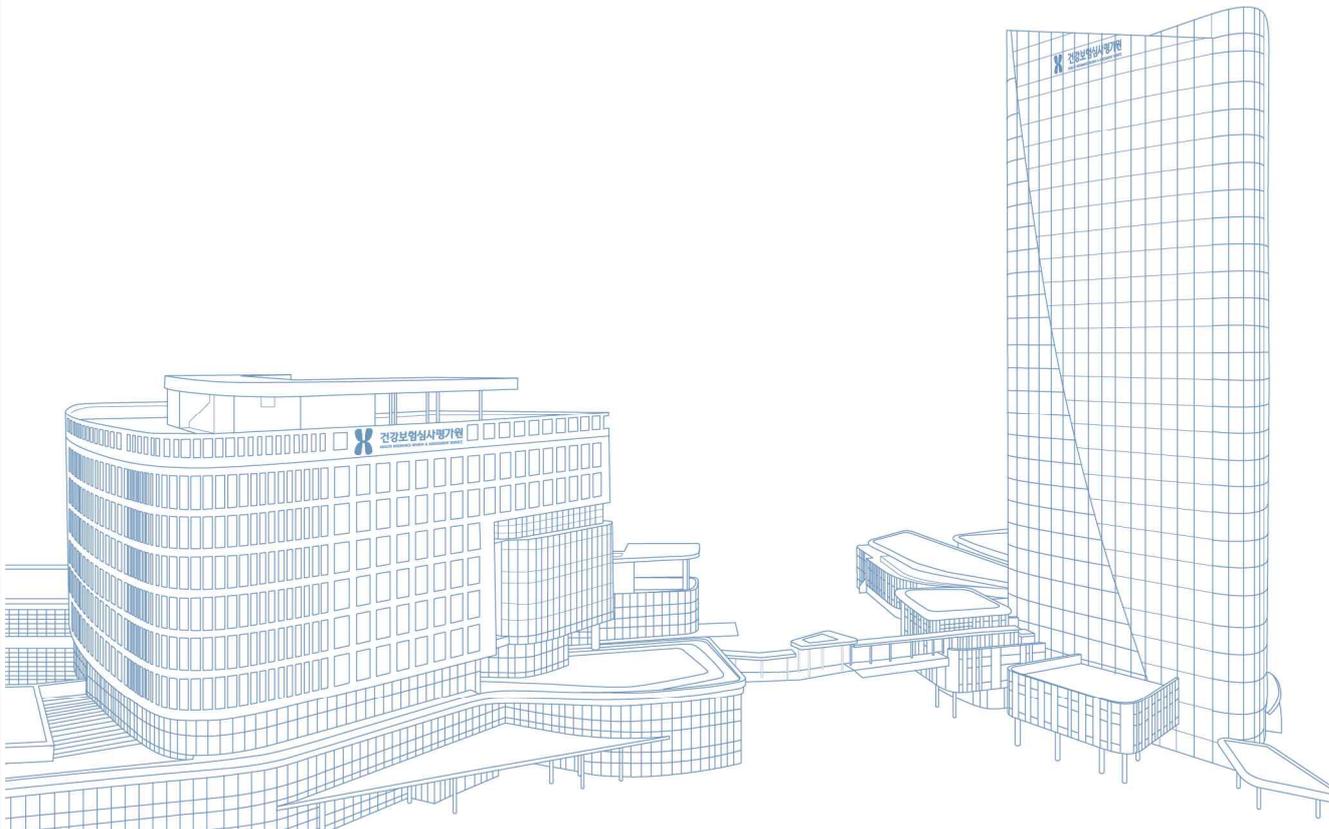
- 64세 이하 청장년층보다 65세 이상 고령인구의 요양급여총액은 전체서비스를 포괄할 경우 5.4% 높았지만, 외래서비스는 4.8%, 입원서비스는 2.7% 높았음($p < 0.0001$)
- 교육수준이 낮은 집단보다 높은 집단에서 의료비 지출은 1.0-2.0% 유의하게 낮았고 미혼보다 기혼, 이혼별거사별 등을 경험한 경우 의료비 지출은 서비스 유형에 따라 1.0%에서 4.2% 높았음
- 노인이나 여성이 많이 포함된 미취업자보다 상근직과 임시일용직, 자영업자가 전체 의료비 지출이 1.8%, 0.8% 0.7% 더 낮은 경향이 나타났고 외래에선 상근직(12%)과 임시일용직(0.4%), 입원에선 상근직(0.7%)과 자영업자(0.7%)의 의료비 지출이 미취업자보다 적었음
- 건강추구행태의 경우 흡연 이력에 따라 과거 흡연자의 의료비는 금연자보다 높았고 현재 흡연자는 금연자보다 낮은 의료비를 지출함에 따라 과거 흡연자의 총 의료비, 외래의료비가 각각 1.3%, 1.1% 높았고 현재 흡연자는 0.8%, 1.0% 낮았음
- 음주 이력에 따라서도 과거 음주자의 의료비는 금주자보다 높았지만, 현재 음주자의 의료비는 낮게 지출되어 과거 음주자의 총의료비, 외래의료비, 입원의료비는 금주자보다 각각 2.1%, 1.8%, 0.7%씩 높았지만 현재 음주자는 0.9%, 0.7%, 0.7%씩 낮았음
- 건강보험 가입유형에 따라 직장가입자에 비해 지역가입자의 외래지출이 0.1% 낮았고, 의료급여 수급권자의 총 의료비 지출이 4.2% 더 많았으며 이는 의료급여 수급권자에 한해 적용되는 투석과 정신질환 관련 정액제 서비스 이용에 따른 영향일 것으로 추측할 수 있음
- 2018년 대비 2019년과 2020년에는 매해 총 의료비 지출이 0.7%씩 증가했고 0.7%씩 증가한 외래보다 입원에서 0.8%와 1.4%씩 지출이 더 많이 증가했음

〈표 23〉 2018-2020년 결합자료의 의료이용 유형별 요양급여총액 영향요인

		전체	외래	입원
성	(ref. 남성)			
	여성	0.000	0.009***	-0.011***
연령	(ref. 64세 이하)			
	65세 이상	0.055***	0.048***	0.027***
교육수준	(ref. 고졸 이하)			
	전문학사 이상	-0.027***	-0.019***	-0.010***
결혼상태	(ref. 미혼)			
	기혼	0.033***	0.026***	0.010**
	이혼·별거·사별	0.042***	0.028***	0.022***
종사상지위	(ref. 미취업자)			
	상근직	-0.018***	-0.012***	-0.017***
	임시·일용직	-0.008***	-0.004**	-0.011***
	자영업자	-0.007***	-0.001	-0.017***
	무급가족종사자	-0.002	-0.003	-0.004
흡연	(ref. 해당없음)			
	과거 흡연자	0.013***	0.011***	-0.001
	현재 흡연자	-0.007***	-0.009***	0.003
음주	(ref. 해당없음)			
	과거 음주자	0.021***	0.018***	0.017***
	현재 음주자	-0.009***	-0.007***	-0.007**
건강보험	(ref. 직장가입자)			
	지역가입자	0.000	-0.001***	-0.002
	의료급여 수급권자	0.042***	0.039	0.006
시간	(ref. 2018년)			
	2019년	0.007***	0.007***	0.008***
	2020년	0.007***	0.007***	0.014***
상수항		2.658***	2.605***	2.716***

* *p<0.1, **p<0.05, ***p<0.01

결론



제4장 결론

1. 연구 결과

- 본 연구는 가명정보를 활용해 한국노동패널조사와 건강보험 빅데이터의 정보를 연계하여 결합자료를 구축하고 표본의 일반적 특성 및 의료이용을 분석하였음
 - 자료원에 따라 연계할 수 있는 기간이 달라 건강보험 빅데이터의 자료제공 기간을 기준으로 분석기간은 2011년부터 2020년까지로 정함
 - 분석에는 전체 표본의 생애주기와 생애사건을 살펴보기 위해 결측치가 적은 성, 연령, 교육수준, 혼인상태 등 인구사회학적 특성과 종사상 지위, 건강보험료 등 노동·경제활동 특성과 건강보험 빅데이터의 자격 및 급여 정보를 활용함
 - 건강보험 빅데이터와 결합하여 등록장애 중증도, 등록된 거주지, 건강보험 가입유형, 건강보험료, 보건의료비(본인 부담금), 노인장기요양보험 인정조사 신청 및 등급인정, 사망을 추가로 살펴보았음
- 원 자료 대비 결합자료의 연령대별 인구구조에선 80세 이상 인구의 결합율이 낮아 65세 이상 고령인구의 비율이 원자료보다 적었지만 시간이 흐를수록 30대 이하 청년층의 비율이 감소하여 2020년에는 40대 이상 중장년층의 비율이 증가함
 - 남녀 간 성비를 포함해 교육수준, 종사상지위, 거주지, 흡연 및 음주 행태, 노동시간, 등은 원 자료인 한국노동패널조사와 비교해 큰 차이는 없었음
 - 건강보험료는 가구구성원 간 의료이용 경험 및 의료비지출 금액에 대한 정보를 가구단위로 응답하여 보건의료비 지출금액이 개인단위 법정본인부담금에 비해 낮았고 상단코딩을 적용한 한국노동패널조사의 건강보험료는 빅데이터의 변수에 비해 연도별 최대값은 더 낮은 경향을 보였음
- 한국노동패널조사에서 결합된 자료에서 장애인, 노인장기 요양보험 등급인정자, 사망자를 파악했으나 해당하는 대상자의 수가 적어 개인을 식별할 위험이 높아 별도 분석을 변수로 포함하지는 않음

- 고령인구의 결합률이 낮았기 때문에 건강보험 빅데이터로 확인한 사망자 수는 조사기관의 집계에 비해 적게 나타났지만 2020년까지 수치상 차이가 점차 감소했고 이는 자료원 간 결합률이 결합자료 분석의 품질과 정확도에 영향을 미침을 드러냄
- 연도별 연령대별 다빈도상병 5순위를 산출한 결과, 건강보험심사평가원에서 제공하는 일반인구의 다빈도상병통계와 유사하게 전연령대에서 근골격계 질환에 의한 의료이용이 많이 나타났으나, 그 외 질환의 경우 2만 명이 안되는 표본자료의 특성상 소수 인원이 특정상병으로 서비스를 반복적으로 이용하는 양상이 연도별, 연령대 내 순위변동에 영향을 미친 것으로 나타남
- 20-30대에선 2011년 대비 2020년 임신과 출산에 의한 의료이용이 외래와 입원 모두 감소했으며, 2018년 한국노동패널조사에 인구대표성 확보를 위한 추가표본이 유입되었음을 감안하면 이는 자료원 내 표본의 연령상승에 의한 작용이라기 보다는 해당 연령대 임신과 출산을 경험하는 여성의 수가 줄어들면서 다른 상병으로 인한 의료서비스 이용이 증가함에 따른 복합적 결과로 볼 수 있음
- 고령층의 외래서비스 이용 시 다빈도상병은 고혈압 외에도 처방이나 투석 등 처치를 목적으로 자주 의료기관에 방문해야 하는 만성신부전과 당뇨병이 일반인구를 대상으로 한 외래 다빈도 상병통계와 비교해 높은 순위를 차지하고 있는 반면 입원은 암, 뇌졸중, 치매처럼 입원이 불가피한 질병에 따라 순위와 세부 상병코드가 달라졌음
- 다만 입원여부에 따라 실손보험 한도가 달라지는 백내장수술은 2011년과 2020년 모두 60-70대의 1순위 다빈도 상병이었고 2015년 수가가 인상된 완화의료는 2020년 60-70대 다빈도 상병 3위에 기록되는 등 정책적, 제도적 요인에 따른 영향도 표본 내 다빈도상병 순위에서는 관찰되는 것이 확인되었음
- 결합자료를 최대한 활용해 의료비 지출 추이와 영향요인을 분석하기 위해 연간 외래나 입원 의료서비스 이용률을 산출한 뒤 미이용도 분석대상에 포함하였음
- 연도별, 연령별 의료비 지출은 증가하는 양상이었으나 미이용군의 비율에 따라 평균과 중위수의 격차가 발생하였음
- 외래 서비스의 이용은 전체 표본, 남성, 여성 모두 2018년 대비 2020년에 코로나로 인해 감소하였으며 입원 서비스의 이용은 전체 표본과 여성의 미취업자 대비 취업자의 종사상지위와 전체 표본과 남성의 현재 음주여부에 따라 통계적으로 유의한 영향

을 받았음

- 인구학적 특성과 건강행태는 법정본인부담금과 보험자부담금을 합산한 요양급여총액에 영향을 미친 반면 건강보험 가입 유형에 따른 차이는 적거나 부분적으로 나타났고, 한방과 의료급여 대상 정액제 서비스를 포함한 전체 서비스 의료비지출은 직장가입자 대비 의료급여 수급권자가 더 높았지만 외래의료비는 직장가입자 대비 지역가입자에서 적었고 가입유형에 따른 입원의료비의 차이는 통계적으로 유의하지 않음
- 건강보험 빅데이터로는 분석범위 밖이었던 인구학적 특성에 따른 지출을 살펴보면, 교육수준은 높을수록 의료비 지출이 적었고 미혼일 경우에 비해 기혼, 이혼, 별거, 사별일 경우 의료비 지출이 늘어났고 종사상지위에 따른 정도의 차이는 있을지라도 미취업자에 비해 취업자의 의료비지출이 더 적은 것으로 나타났지만 여러 특성이 교차하면서 발생하는 효과는 분석되지 못함

2. 연구 의의

- 결합자료의 인구학적 특징과 의료이용 현황을 국가통계 및 공공데이터와 비교해 가명정보결합자료의 선도 사례로서 기초정보를 제공하고 가명정보결합 기법을 활용한 공공데이터 연계가 기존 자료원의 한계를 보완할 수 있다는 의의를 확인함
- 본 연구에선 결합과정과 결합된 자료의 특성을 기술하고 생애주기 인구집단의 특성에 따른 의료이용을 파악하는 것을 목표로 함
- 개인의 교육수준, 직업, 개인이 인지한 소득수준 등 다양한 건강의 사회적 결정요인을 포함한 의료이용 분석으로 청구자료로 파악할 수 없었던 건강격차의 현황을 파악하고 정책적 개입이 필요한 인구집단과 그 개입전략을 특정하기 위한 기초자료로 활용 가능함
- 단, 건강이 악화되었거나 질병 및 외상으로 인한 장기 입원 및 거주시설 입주로 인해 중도탈락 및 조사거부에 해당하는 대상은 분석대상에서 제외하였으므로 이에 관한 후속연구가 요구됨
- 자료연계의 핵심자료가 내재한 표본수, 결측치, 보고편향, 경쟁사건으로 인한 삐뚤림의 문제로 인해 특정 질병이나 소수 인구집단에 관한 분석은 여전히 제한적일 수

있음

- 개인과 시간의 특성을 통제한 고정효과모형을 사용한 분석결과를 통해 한국인의 전반적인 의료이용 접근성에 비해 의료비 부담이 증가하고 있음을 확인하였으며 영향요인과 그 효과크기는 전체 표본을 대상으로 한 분석과 성별 층화분석에서 다르게 나타날 수 있음을 파악하였음
- 본 연구에서 아직 활용하지 못한 건강검진과 장기요양 등을 활용해 의료이용과 건강행태의 인과관계 및 효과를 평가하거나 외래와 입원 의료서비스 이용 당시 상병의 특성 및 과거 의료이용 이력에 따른 예후, 건강행태 변화 등을 모니터링 및 평가하기 위해서는 더욱 세밀한 후속연구가 필요함

3. 정책적 제언

- 표본자료의 한계에도 불구하고 다수 인구집단에 해당하거나 단기간 치사율이 높은 급성중증질환은 아니라는 이유로 빅데이터 기반 연구에서는 주목받지 못했던 완화의료나 만성신부전, 손상 및 외상 관련 질환 등 의료이용의 원인들을 폭넓게 탐색할 수 있으므로 가명정보결합자료를 활용한 정책연구가 활발해져야 함
- 생애주기와 생애사건에 따른 단면적 영향을 살펴보았으나 배우자가 사별한 고령의 여성의 교육수준에 따른 의료비 지출 차이나 혼인상태 및 종사상지위의 변화로 인해 건강행태가 달라진 남성의 의료기관 방문 빈도의 변화 등 서로 다른 특성이 교차하면서 발생할 수 있는 인구집단의 건강 및 의료이용 격차에 대한 심층분석이 필요함
- 가명정보결합의 활용도를 높이기 위해선 결합전문기관으로 지정되지 않더라도 자료제공기관이 연구윤리를 준수해 자료를 가공 및 제공할 수 있도록 관련 절차와 업무에 관한 지침을 정비하고 교육, 사례, 심의가 원활하게 이뤄질 수 있도록 인적, 경제적 지원을 강화해야 함
- 학계와 공공기관은 기관의 자체적인 규정이나 담당자의 재량에 맡겨진 영역을 파악하고 관계기관 간 법률 조항을 해석하거나 자료원을 검증하는 기준의 차이에 대한 협의를 활성화하여 제도가 공익적 목적을 수행할 수 있도록 기여해야 함
- 가명정보결합 기법을 활용한 자료의 접근, 연계, 활용의 투명성 및 효율성이 개선한

다면 보건의료정책 시행과정에서 수집·관리하는 행정자료에서 관찰할 수 없었던 일반인구의 건강 관련 행태 및 의사결정, 기타 사회경제 문화적 특성들을 포함한 보건학적 연구에 활용도가 높을 것으로 기대됨

참고 문헌

- 건강보험심사평가원. 보건의료빅데이터개방시스템. Retrieved from <https://opendata.hira.or.kr/home.do> (Access on 2022.04.12.)
- 건강보험심사평가원. (2022). 의약품 및 의료기기 생애주기별 국내외 급여 관리제도 비교.
- 김동진, 채수미, 최지희, 김창엽, 김새롬, 박유경, ... & 박여리. (2016). 국민의 건강수준 제고를 위한 건강행태성 모니터링 및 사업 개발.
- 김남순, 송현종, 김명희, 강영아, 이희영, & 광민선. (2003). 한국여성의 건강통계.
- 김남순, 박은자, 전진아, 김명희, 김성이, 김유미, ... & 차미란. (2018). 한국 여성의 건강통계 산출 및 만성질환 치료·관리의 성별차이 분석.
- 김선제, 황지영, & 선정연. (2020). 보건의료환경이 일차의료 질 향상에 미치는 영향: 2008-2017년 OECD 보건통계 자료를 활용한 패널분석. 보건정보통계학회지, 45(3), 302-310. 김은아, 오유미, & 조성일. (2020). 국민건강증진종합계획 2030의 수립 개선 전략: 작성지침 준수도 평가를 통하여. 보건교육건강증진학회지, 37(1), 33-44.
- 김수진. (2022). 의료의 질과 격차 현황: 의료시스템의 건강 성과를 중심으로. 보건복지 Issue & Focus, 428, 1-10.
- 노가은, & 이지은. (2019). 한국노동패널조사(KLIPS) 활용 중·고령층 연구 현황: 1998~2018. 노동연구원.
- 대한민국 정책브리핑. (2022.02.18). 개인정보위, 제2기 가명정보 결합 선도사례 중점 추진. <https://www.korea.kr/news/policyNewsView.do?newsId=148899127>
- 박민규, & 김사라. (2015). 한국노동패널조사 자료의 분석을 위한 패널 가중치 산출 및 사용방안 사례 연구. 조사연구, 16(1), 49-71.
- 박은자, 정연, 서제희, 배정은, 이나경, 김은주, ... & 박현아. (2019). 제 2기 한국의료패널 구축·운영을 위한 연구.
- 오미애. (2015). 보건복지분야 데이터 연계 필요성 및 활용방안. 보건복지포럼, 2015(9), 17-28.
- 오미애, 최현수, 김용대, 이용희, & 진재현. (2014). 보건복지통계정보 생산 및 활용 촉진을 위한 마이크로데이터 통합 연계 방안. 한국보건사회연구원.

- 유종성, 전병유, 신광영, 이도훈, & 최성수. (2020). 증거기반 정책연구를 위한 행정자료의 활용. 한국사회정책, 27(1), 5-37.
- 이상협, 박찬용, 정성석, & 최혜미. (2011). 한국노동패널 탈락 분석. 한국데이터정보과학회지, 22(1), 1-8.
- 이석용, & 이지은. (2020). 한국노동패널조사(KLIPS) 활용 여성 연구 현황(1999-2019). 한국노동연구원.
- 이윤경. (2021). 생애주기의 연령기준과 연령구분에 대한 인식. 보건복지포럼, 2021(3), 39-54.
- 이해정. (2020). 패널 데이터 품질 개선을 위한 항목무응답 대체 방법. 보건복지포럼, 2020(8), 43-60.
- 장숙량, 이주현, 김창오, 허현희, 황종남, & 김태연. (2017). 국민건강증진종합계획 건강형평성 대표지표 개발 및 건강격차 해소 전략. 보건교육건강증진학회지, 34(4), 41-57.
- 장인성, 이지은, 신선옥, 정현상, 권익성, & 임태민. (2022). 한국노동패널조사, 1998-2020 [1-23 차]. 정혜민, 김현주, & 이진용. (2021). 국내 반복입원의 현황과 환자 특성: 외래진료 민감질환을 중심으로. 한국의료질향상학회지, 27(2), 45-56.
- 정현수, 김수영, & 김현수. (2019). 건강보험 빅데이터를 이용한 임상 연구: 실제와 전망. 대한내과학회지, 94(6), 463-470.
- 정최경희, & 김유미. (2013). 건강 수준에서의 불평등: 우리나라의 현황. Journal of the Korean Medical Association, 56(3), 167-174.
- 조진호, 김유빈, & 이지은. (2018). 한국노동패널조사(KLIPS) 활용 청년 연구 현황 1999~2017. 노동연구원.
- 조혜민, & 최혜미. (2019). 의학과 치의학 연구에서 통계 활용의 최근 추세. 한국데이터정보과학회지, 30(6), 1341-1349.
- 최용준, 윤태호, & 신동수. (2012). 건강 형평성의 관점에서 본 제 3 차 국민건강증진종합계획 평가. 비판사회정책, (37), 367-400.
- 통계청. (2021). 2020년 사망원인통계. 대전: 통계청.
- 통계청. (2021). 2021년 통계로 보는 1인가구.
- 홍민기, & 최효미. (2014). 노동패널 표본이탈과 불평등 지표. 조사연구, 15(4), 95-122.
- Bellis, M. A., Hughes, K., Ford, K., Rodriguez, G. R., Sethi, D., & Passmore, J. (2019). Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and

meta-analysis. *The Lancet Public Health*, 4(10), e517-e528.

- Chandra, A., & Vogl, T. S. (2010). Rising up with shoe leather? A comment on Fair Society, Healthy Lives (the Marmot Review). *Social science & medicine*, 71(7), 1227-1230.
- Delaney, L., & Smith, J. P. (2012). Childhood health: trends and consequences over the life-course. *The Future of Children/Center for the Future of Children, the David and Lucile Packard Foundation*, 22(1), 43.
- Goodman, J. M., Boone-Heinonen, J., Richardson, D. M., Andrea, S. B., & Messer, L. C. (2018). Analyzing Policies Through a DOHaD Lens: What Can We Learn?. *International journal of environmental research and public health*, 15(12), 2906.
- Hanson, M. A., & Gluckman, P. D. (2014). Early developmental conditioning of later health and disease: physiology or pathophysiology?. *Physiological reviews*.
- Hwang, J. H., & Rhee, S. J. (2002). A study on the identification of risk factors for unplanned readmissions in a university hospital. *Journal of Korean Public Health Nursing*, 16(1), 201-212.
- Jones, N. L., Gilman, S. E., Cheng, T. L., Drury, S. S., Hill, C. V., & Geronimus, A. T. (2019). Life course approaches to the causes of health disparities. *American journal of public health*, 109(S1), S48-S55.
- Jung-Choi, K., Khang, Y. H., & Cho, H. J. (2011). Socioeconomic differentials in cause-specific mortality among 1.4 million South Korean public servants and their dependents. *Journal of Epidemiology & Community Health*, 65(7), 632-638.
- Khang, Y. H. (2006). Relationship between childhood socio-economic position and mortality risk in adult males of the Korea Labour and Income Panel Study (KLIPS). *Public health*, 120(8), 724-731.
- Khang, Y. H., & Kim, H. R. (2016). Socioeconomic Inequality in mortality using 12-year follow-up data from nationally representative surveys in South Korea. *International journal for equity in health*, 15(1), 1-11.
- Kawachi, I., Subramanian, S. V., & Almeida-Filho, N. (2002). A glossary for health inequalities. *Journal of Epidemiology & Community Health*, 56(9), 647-652.
- OECD (2021), *Health at a Glance 2021: OECD Indicators*, OECD Publishing, Paris, <https://doi.org/10.1787/ae3016b9-en>.



- Sun, C., Velazquez, M. A., & Fleming, T. P. (2016). DOHaD and the periconceptional period, a critical window in time. In *The epigenome and developmental origins of health and disease* (pp. 33-47). Academic Press.
- Ryu, D. R. (2017). Introduction to the medical research using national health insurance claims database. *The Ewha Medical Journal*, 40(2), 66-70.
- WHO. (2008). *Closing the gap in a generation: health equity through action on the social determinants of health: Commission on Social Determinants of Health final report*. World Health Organization Commission on Social Determinants of Health, & World Health Organization.
- World Health Organization. (2010). *A conceptual framework for action on the social determinants of health*.
- World Health Organization. (2014). *Health in all policies: Helsinki statement. Framework for country action*.
- World Health Organization. (2022). *World health statistics 2022: monitoring health for the SDGs, sustainable development goals*. Geneva: World Health Organization.
- Zacher, H., & Froidevaux, A. (2021). Life stage, lifespan, and life course perspectives on vocational behavior and development: A theoretical framework, review, and research agenda. *Journal of Vocational Behavior*, 126, 103476.

생애주기에 따른 의료이용 실태분석 및 형평성 비교 연구용역

발행일 : 2022년 12월

발행인 : 김 선 민

편집인 : 이 진 용

발행처 : 건강보험심사평가원 심사평가연구소
강원도 원주시 혁신로 60(반곡동)

대표전화 : 1644-2000

홈페이지 : www.hira.or.kr

※ 이 보고서는 무단으로 복제나 인용을 할 수 없습니다.
(저작권법 제136조 등 관련법 적용)