

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발(8차년도)

조경희 · 이상현 · 박성배 · 박세희 · 최준호 · 박주현 · 문한빛 · 오시내 · 전성연
이재광 · 최정규 · 조현순 · 유인선 · 이재우 · 배철영 · 김동욱 · 박은지 · 김유미

국민건강보험

National Health
Insurance Service

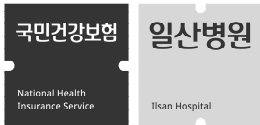
일산병원

Ilsan Hospital

연구보고서
2021-20-014

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발(8차년도)

조경희 · 이상현 · 박성배 · 박세희 · 최준호 · 박주현 · 문한빛
오시내 · 전성연 · 이재광 · 최정규 · 조현순 · 유인선 · 이재우
배철영 · 김동욱 · 박은지 · 김유미



국민건강보험 일산병원 연구소

[저 자]

책임 연구자:	국민건강보험 일산병원 가정의학과	조경희
공동 연구원:	국민건강보험 일산병원 가정의학과	이상현
	국민건강보험 일산병원 가정의학과	박성배
	국민건강보험 일산병원 내분비내과	박세희
	국민건강보험 일산병원 가정의학과	최준호
	국민건강보험 일산병원 가정의학과	박주현
	국민건강보험 일산병원 가정의학과	문한빛
	국민건강보험 일산병원 가정의학과	오시내
	국민건강보험 일산병원 연구소 연구분석부	전성연
	국민건강보험 일산병원 연구소 연구분석부	이재광
	국민건강보험 일산병원 연구소 연구분석부	최정규
	국립암센터	조현순
	이화여대 서울병원 가정의학과	유인선
	충북대학교병원 가정의학과	이재우
	메디에이지	배철영
	경상대학교 정보통계학과	김동욱
	국민건강보험공단 빅데이터전략본부	박은지
연구보조원:	국민건강보험 일산병원	김유미
도움주신 분	한국사회보장정보원 데이터연구소	박영규

연구관리번호		IRB 번호
NHIS-2021-1-459	NHIS-2021-2-155	NHIMC 2021-01-003
NHIS-2021-2-156	NHIS-2021-2-201	

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

머리말

건강IN은 인터넷상 검증되지 않고 무분별하게 서비스 되고 있는 건강정보로부터 국민건강을 보호하기 위하여 인터넷상 검증되지 않고 무분별하게 서비스 되고 있는 건강정보로부터 국민건강을 보호하기 위한 건강정보 전문사이트이다.

본 연구는 올해로 8년째 검증한 신뢰성 있는 건강정보를 제공하도록 건강인 사이트 관련 연구를 지원해 오고 있다. 다빈도 질병/처치/수술/검사/흡연/영양 등 다양한 건강정보를 개발해 왔으며, 공단이 보유하고 있는 국가건강검진자료를 이용한 개인별 맞춤형 건강정보 연구, 비만·흡연자를 위한 맞춤형 생활습관 개선 프로그램 연구, 개인의 위험도 평가 후 건강개선자료를 제공하는 건강나이 알아보기 연구, 대사성질환자를 위한 맞춤형 정보 연구 등을 수행해 왔다. 최근 수년간은 각종 암에 관한 질병 예측 연구와 함께 유전체연구, Digital Therapeutics(DTx) 관련 연구 등 4차 산업 혁명시대의 맞춤형 정보 관련 연구를 수행하였다.

2021년에는 노화관련질병(ARD)의 현황과 질병예측모델의 개발, 지불제도와 관련하여 계층적 질환군(NHIS-HCC) 위험조정 모델 정보화 현황, 건강-요양-복지 연계를 위한 정보 현황을 살펴보았다.

모두 14개의 노화관련질병(ARD)이 분석되었으며, 그 중에서 골다공증의 예측 모델과 환자 교육 자료를 개발하였다. 계층적 질환군(NHIS-HCC) 위험조정 모델은 추후 정책적 필요에 의한 새로운 준비가 요망되며, 건강정보와 요양정보의 연결로 맞춤형 보건의료 서비스 제공을 위한 Data set을 구축하였으며 추후 복지 정보(사회보장원의 PSR)의 연계는 개인별 건강-요양-복지 정보의 Dat-set이 구축될 것이다.

국가 PHR 사업에 맞추어 건강인도 신뢰 있는 건강 플랫폼을 위한 지속적인 고도화와 함께 앱을 통한 맞춤형 건강 정보 제공을 위한 지속적인 노력이 필요하다.

2022년 2월

국민건강보험 일산병원장

일산병원 연구소장

김성우
이천균

목차

요약	1
제1장 표본 코호트에서의 ARD 분석	9
제1절 서론	11
제2절 이론적 고찰	12
제3절 연구 내용 및 방법	21
제4절 분석 결과	22
제5절 결론 및 제언	28
제2장 검진 코호트에서의 ARD 분석	31
제1절 서론	33
제2절 이론적 고찰	34
제3절 연구 내용 및 방법	35
제4절 분석 결과	38
제5절 결론 및 제언	48
제3장 골다공증 예측 모델의 개발	51
제1절 서론	53
제2절 이론적 고찰	53
제3절 연구 내용 및 방법	57
제4절 분석 결과	61
제5절 결론 및 제언	88
제4장 골다공증 환자 교육 처방 메시지	91
제1절 서론	93
제2절 이론적 고찰	93
제3절 연구 내용 및 방법	95
제4절 분석 결과	98
제5절 결론 및 제언	101

제5장 계층적 질환군(NHIS-HCC) 위험조정 모델을 통한 개인별 건강위험 평가	103
제1절 서론	105
제2절 이론적 고찰	106
제3절 연구 내용 및 방법	109
제4절 분석 결과	112
제5절 결론 및 제언	115
제6장 Data 결합을 통한 공공서비스 제공에 관한 연구	117
제1절 서론	119
제2절 이론적 고찰	120
제3절 연구 내용 및 방법	120
제4절 분석 결과	121
제5절 결론 및 제언	127
참고문헌	129

표목차

<표 1-1> Top 5 leading specific causes of DALYs by age group in Korea, 2015	16
<표 1-2> 노인의 성별 만성질병 종류별 유병률 및 현 치료율	20
<표 1-3> 노화관련질병 및 주상병 코드	22
<표 1-4> 연도별 발생률	24
<표 1-5> 연령별 발생률	26
<표 2-1> ARD 별 최종 연구대상자	36
<표 2-2> 건강노화나이 공식	37
<표 2-3> 건강노화나이 그룹 기준	37
<표 2-4> 연구대상자 특성	38
<표 2-5> ARD 발생률(전체)	39
<표 2-6> ARD 발생률(남)	39
<표 2-7> ARD 발생률(여)	39
<표 2-8> 연령대별 ARD 발생률(전체)	40
<표 2-9> 연령대별 ARD 발생률(남)	41
<표 2-10> 연령대별 ARD 발생률(여)	41
<표 2-11> 건강노화나이 그룹별 ARD 발생률(전체)	42
<표 2-12> 건강노화나이 그룹별 ARD 발생률(남)	43
<표 2-13> 건강노화나이 그룹별 ARD 발생률(여)	44
<표 2-14> 성별 건강노화나이에 따른 ARD 발병 위험도	45
<표 2-15> 연령대별 건강노화나이에 따른 ARD 발병 위험도	45

<표 2-16> 성별 의료 관련 요인 특성	46
<표 2-17> 연령대별 의료 관련 요인 특성(전체)	46
<표 2-18> 연령대별 의료 관련 요인 특성(남, 여)	47
<표 2-19> 건강노화나이와 연령에 따른 의료 관련 요인 분석	48
<표 3-1> 일반검진 대상자 특성	62
<표 3-2> 일반검진 대상자 생존분석	64
<표 3-3> 65세 미만, 65세 이상 생존분석	66
<표 3-4> Random forest의 model accuracy with test set	70
<표 3-5> eXtreme Gradient Boosting의 model accuracy with test set ·	72
<표 3-6> K-nearest neighbor의 model accuracy with test set	73
<표 3-7> Lasso의 model accuracy with test set	76
<표 3-8> Logistic의 model accuracy with test set	77
<표 3-9> 여성의 골다공증 예측모델 비교	78
<표 3-10> Random forest의 model accuracy with test set	80
<표 3-11> eXtreme Gradient Boosting의 model accuracy with test set ·	82
<표 3-12> K-nearest neighbor의 model accuracy with test set	84
<표 3-13> Lasso의 model accuracy with test set	86
<표 3-14> Logistic의 model accuracy with test set	87
<표 3-15> 남성의 골다공증 예측모델 비교	88
<표 4-1> 건강메시지 제공 공공기관 웹페이지	94
<표 4-2> 10년 이내 골다공증 발생 위험 예측도 분류	97
<표 4-3> 골다공증 발생 요인별 위험메시지 시, 고려한 요인	98
<표 4-4> 골다공증 위험 공통 요인 메시지 시, 고려한 요인	98
<표 4-5> 골다공증 발생 위험 메시지	99
<표 4-6> 골다공증 위험 교정 가능한 요인의 건강신호등	99

<표 4-7> 골다공증 위험 교정 가능한 요인별 메시지	100
<표 4-8> 골다공증 위험 공통 요인별 메시지	101
<표 5-1> NHIS-HCC 위험조정 모델에서 질환군 계층화 분류 기준	111
<표 5-2> 계층적 질환군에 따른 HCC 위험지수	112
<표 6-1> 고양시 65세 이상 인구 대상 장기요양등급별 특성	121
<표 6-2> 고양시 65세 이상 인구 대상 재택의료 서비스 현황	122
<표 6-3> 고양시 65세 이상 인구 대상 장기요양등급별 특성 (Y 기각, Z 탈락)	122
<표 6-4> 고양시 65세 이상 인구 중 심한 장애인에서 장기요양등급별 특성(Y 기각, Z 탈락)	123
<표 6-5> 고양시 65세 이상 인구 중 심하지 않은 장애인에서 장기요양등급별 특성	124
<표 6-6> 고양시 신규 등급판정자 대상 거주지 분류	125
<표 6-7> 고양시 신규 등급판정자 대상 항목별 장기요양등급판정표 점수 ·	125

그림목차

[그림 1-1] Schematic representation of the proposed sequence of events moving from age-related alterations to multimorbidity	13
[그림 1-2] Main aging-related disease with age and incidence	14
[그림 1-3] Leading causes of global DALYs and percentage of total DALYs (1990 and 2019), and percentage change in number of DALYs and age-standardised DALY rates from 1990 to 2019	15
[그림 1-4] Diseases caused by threshold-crossing of a parameter affected by senescent cells are predicted to have an exponential incidence curve with a decline at old ages.	18
[그림 1-5] 92 age-related disease	19
[그림 1-6] 표본코호트 노화관련질병 연구 대상자 선정	21
[그림 1-7] 연도별 발생률	23
[그림 1-8] 연령별 발생률	26
[그림 2-1] 건강노화나이에 따른 ARD 발병률과 관련 요인 분석 대상자 선정	36
[그림 3-1] 고관절 골절의 연령별 발생률	54
[그림 3-2] 골다공증 환자의 연령-성별에 따른 의료이용률	55
[그림 3-3] 연구대상자 flow chart	58
[그림 3-4] Data split과 5-fold cross validation	61
[그림 3-5] Random forest의 modeling in training set	68
[그림 3-6] Random forest의 variable importance	69
[그림 3-7] Random forest의 cutoff point ROC curve	69

[그림 3-8] eXtreme Gradient Boosting의 modeling in training set	70
[그림 3-9] eXtreme Gradient Boosting의 variable importance	71
[그림 3-10] eXtreme Gradient Boosting의 cutoff point ROC curve	71
[그림 3-11] K-nearest neighbor의 variable importance	72
[그림 3-12] K-nearest neighbor의 cutoff point ROC curve	73
[그림 3-13] Lasso의 modeling in training set	74
[그림 3-14] Lasso의 variable importance	75
[그림 3-15] Lasso의 cutoff point ROC curve	75
[그림 3-16] Logistic의 variable importance	76
[그림 3-17] Logistic의 cutoff point ROC curve	77
[그림 3-18] Random forest의 modeling in training set	78
[그림 3-19] Random forest의 variable importance	79
[그림 3-20] Random forest의 cutoff point ROC curve	79
[그림 3-21] eXtreme gradient Boosting의 modeling in training set	80
[그림 3-22] eXtreme Gradient Boosting의 variable importance	81
[그림 3-23] eXtreme Gradient Boosting의 cutoff point ROC curve	81
[그림 3-24] K-nearest neighbor의 modeling in training set	82
[그림 3-25] K-nearest neighbor의 variable importance	83
[그림 3-26] K-nearest neighbor의 cutoff point ROC curve	83
[그림 3-27] Lasso의 modeling in training set	84
[그림 3-28] Lasso의 variable importance	85
[그림 3-29] Lasso의 cutoff point ROC curve	85
[그림 3-30] Logistic의 modeling in training set	86
[그림 3-31] Logistic의 cutoff point ROC curve	87
[그림 4-1] 골다공증 환자 교육 처방메시지 제공 단계	95
[그림 4-2] 처방 메시지 고려 위험 요인	96
[그림 5-1] 허혈성 심질환군의 계층화 단계 사례	108

요약



요약

1. 연구 배경 및 목적

건강in은 인터넷상 검증되지 않고 무분별하게 서비스 되고 있는 건강정보로부터 국민 건강을 보호하기 위하여 인터넷상 검증되지 않고 무분별하게 서비스 되고 있는 건강정보로부터 국민건강을 보호하기 위한 건강정보 전문사이트로써, 본 연구는 올해까지 8년에 걸쳐서 검증한 신뢰성 있는 건강정보를 제공하도록 건강인 사이트 관련 연구를 지원해 오고 있다. 다빈도 질병/처치/수술/검사/흡연/영양 등 다양한 건강정보를 개발해 왔으며, 공단이 보유하고 있는 국가건강검진자료를 이용한 개인별 맞춤형 건강정보 연구, 비만·흡연자를 위한 맞춤형 생활습관 개선 프로그램 연구, 개인의 위험도 평가 후 건강개선자료를 제공하는 건강나이 알아보기 연구, 대사성질환자를 위한 맞춤형 정보 연구 및 최근 수년간은 각종 암에 관한 질병 예측 연구와 함께 4차 산업 혁명시대의 맞춤형 정보 관련 연구를 수행하였다.

2021년 8년차 연구에서는 노화관련질병(ARD)의 분석과 함께, 그 중에서 골다공증 예측모델을 분석하였다(제1-4장). 시대적으로 준비가 필요한 계층적 질환군(NHIS-HCC) 위험조정 모델에 대한 분석(제5장), 그리고 Data Dam 사업의 기초로 결합자료의 현황을 살펴보고자 하였다(제6장).

구체적으로는 고령화 사회에서 가장 대표적으로 흔하며 의료이용과 의료비 증가의 원인이 되는 노화관련질병의 개념과 정의 그리고 연도별 발병률 및 연령별 발병률을 표본코호트로 살펴보았다(제1장). 건강검진 코호트를 이용하여 노화 관련 질환 분석 및 의료이용에 관한 현황을 알아보았다(제2장). 건강in의 탑재모델로 골다공증의 예측모델을 개발하고(제3장), 관련 교육자료 현황 분석과 자료를 개발하고자 하였다(제4장).

계층적 질환군(NHIS-HCC) 위험조정 모델을 통한 개인별 건강위험도를 제시하고 이를 추후 기준 연령군별 비교하여 상대 건강위험도를 제시한다면, 의료비 예측과 함께 개인의 건강관리 방향 및 의료 이용을 파악하는 데 도움이 될 것이다. 본 연구는 최근에 구축된 2019년도 표본코호트 자료를 이용하여 질환력을 통한 각 개인별 진료비 분석과

함께 개인별 의료이용을 조사하고자 하였으나, 필요한 통계 자료 수집의 한계로 방향 제시의 보고서로 구성하였다(제5장).

한국형 뉴딜 사업의 핵심인 데이터 댐에서 보건의료복지정보의 활용을 위해서는 데이터 셋 구축이 필요하다. 통합형 보건복지 서비스의 돌봄사업에서의 서비스 제공은 단순히 보건의료, 홈케어만이 아닌 지역사회 내 여러 자원(예를 들면, 주민센터, 보건소 등)을 활용하여 서로 연계된 서비스가 제공된다. 보건의료복지 데이터 셋의 원천 자료으로는 의료 정보, 요양정보 그리고 복지정보가 필요하다. 데이터의 축적과 분석, 그리고 서비스 활용 등을 위하여 데이터 연계의 현황과 방향을 모색해 보고자 하였다 (제6장).

2. 연구 결과

표본 코호트에서의 ARD 분석에서는 고령사회에 진입한 한국에서 이러한 형태를 가진 노화관련질환을 재정의하여 연령에 따라 질병부담이 증가하고 발생률이 증가되는 질환들 중 한국인에서 유병률이 높은 질환 14개를 선정하여 (암 제외) 연도별, 연령별 발생률을 확인 하였다. 노화관련질환중 2004년부터 2019년까지 연도별로 발생률이 감소한 질환은 만성폐쇄성폐질환, 울혈성심부전, 허혈성심질환 이었으며, 지속적으로 발생률이 증가한 질환은 고지혈증, 만성신질환, 백내장, 난청, 파킨슨병 이었다. 그리고 특징적으로 고혈압, 당뇨, 뇌혈관계질환, 골다공증, 골관절염, 연령관련 황반변성의 경우, 점차 감소되던 발병률이 2015년 이후 증가하는 경향을 보였다. 하지만, 각 질환들의 10세 단위의 연령별 발생률에서는 연도별 발생률에서 보였던 차이와 상관없이 거의 모든 질환들은 정도의 차이는 있으나 연령이 증가함에 따라 지수함수적으로 증가하다가 매우 높은 연령에서는 발생률이 감소하는 형태를 보였다. 울혈성 심부전의 경우만 연령증가에 따라 지속적으로 발생률이 증가하는 형태를 보였다. 각 질환들에서 가장 높은 발병률을 보이는 연령에는 약간의 차이가 있었으나 대부분 70~80세까지 발생률이 증가하다가 감소하는 것으로 보여진다.

건강노화나이에 따른 ARD 발병률과 관련 요인 분석 연구에서는 총 276,723명의 10년 코호트 자료를 기준으로 건강노화나이에 따른 ARD 발병률과 관련 요인을 분석하였다. 연령대가 증가할수록 모든 ARD 발생률이 높아졌고 건강노화나이가 나쁜 그룹일수록 ARD 발생률이 높아지는 결과를 보였다. 또한 건강노화나이가 증가할수록 ARD 발병 위험도가 모두 유의하게 높아지는 결과를 보였다. (HR: 고지혈증 1.259, 골관절염 1.062,

만성신질환 1.253, 만성폐쇄성폐질환 1.005, 백내장 1.031, 울혈성심부전 1.025) 그리고 의료비, 연간 의료비 증가, 외래 일수, 입원 일수 모두 남자보다 여자가 높았고, 연령대가 증가할수록 모두 증가하는 결과를 보였다. 또한 전체, 남, 여 모두 건강노화나이가 많을수록, 연령이 높을수록 의료 관련 요인이 유의하게 증가하는 결과를 보였다.

골다공증 예측모델 개발에서는 일반검진 수검자를 대상으로 위험인자 및 생활습관과 골다공증과의 관련성을 분석한 결과 여성의 경우 나이에 따라 가장 큰 격차를 보였으며, 당뇨병인 경우(RH 1.433), 저체중인 경우(HR 1.088), 폐경인 경우(HR 1.124) 위험도가 높게 나타났다. 남성의 경우에도 나이에 따라 가장 큰 격차를 보였으며, 당뇨병인 경우(HR 1.55), HGB 위험군(HR1.329), 저체중인 경우(HR 1.225) 위험도가 높게 나타났다. 머신러닝(Machine learning)과 logistic regression을 이용한 예측모델을 분석하였으며, 여성의 경우 다섯 개의 예측모델 중 AUC가 가장 높은 모델은 KNN(0.621)이었고 accuracy가 가장 높은 모델은 Lasso(0.5973) 이었으며, positive predictive value가 가장 높은 모델은 Lasso(0.4745)였다. 남성의 경우 다섯 개의 예측모델 중 AUC (0.615-0.679)가 가장 높은 모델은 Lasso(0.679)이었고 Accuracy (0.3804-0.47078)가 가장 높은 모델은 Random forest(0.4078) 이었으며, positive predictive value (0.0382-0.0487)가 가장 높은 모델은 KNN(0.0483)로 매우 낮게 나타났다.

골다공증 환자 교육 처방 메시지에서는 개발된 10년 골다공증 예측모델에 근거하여, 예측모델의 추정된 모형식에 포함된 변수를 참조하여 처방메시지를 작성하였다. 환자 교육 처방메시지를 작성 시, 예측모델 모형식에 고려된 골다공증 위험 변수들을 고려하였다. 골다공증 위험 처방메시지는 단계별 처방 메시지를 제공하는 것으로 작성하였다. 환자(또는 이용자)의 국가 검진 결과 정보를 불러오고, 이 정보를 토대로 골다공증 발생 위험 메시지를 제공하고, 이후 위험요인 중에 교정 가능한 요인들은 해당 요인별 메시지를 제공하고, 이후에 추가적인 골다공증 공통적인 위험 메시지를 제공하는 것으로 작성하였다. 개별 환자(또는 이용자)별로 국가건강검진 자료를 기반으로 계산한 10년 골다공증 발생위험도를 제공하고, 이후 동일연령대의 골다공증 위험도와 비교하여 어떠한 수준인지 나타내는 메시지로 개발하였다. 흡연, 운동, 음주, BMI (허리둘레)등의 교정가능한 요인들은 3분류로 분류하고, 위험에 해당하는 경우 요인별 처방메시지를 작성하였다. 바꿀 수 없는 (교정 불가능한) 요인 및 일반적인 골다공증 발생 위험으로 알려진 요인에 대한 메시지도 추가 작성 및 개발하였다. 고려한 공통 위험 요인으로는 나이, 골절 과거력, 골절 가족력, 폐경여부, 스테로이드 치료여부 등에 대하여 메시지를 작성하였다.

NHIS-HCC 위험조정 모델의 의료비 예측 및 분석은 자료 획득 및 시한의 문제로 2019년 표본코호트 자료의 분석은 못하였고, 방법론 및 기존 연구 결과(참고문헌7)로 대체하였다.

Data 결합을 통한 공공서비스 제공에 관한 연구에서는 2020년 1월 1일부터 2020년 12월 31일까지 고양시에 거주하는 65세 이상 노인인구 156,861명을 대상으로 하였다. 노인 1인당 연간 총 진료비는 3,125,160원으로 나타났다. 장기요양 등급자, 심한장애인, 심하지 않은 장애인 모두에서 일반 노인 평균에 비해 총 진료비가 높았다. 장기요양 등급별로는 등급이 높을수록 진료비용 및 입원 일수가 증가함을 관찰하였고 심한 장애인과 심하지 않은 장애인에서도 장기요양 등급이 높아질수록 총 진료비 및 입원일수가 증가함이 관찰되었다.

3. 결론 및 제언

표본코호트에서의 ARD 분석은 한국인을 대표하는 코호트를 기반으로 하여 연령질환 질병의 발생률을 보여준다는 점에서 의의가 있겠다. 이러한 노화관련질병에 대한 이해를 높이고 관련 요인들을 파악하여 질병의 예측 모델을 개발하여 예방함으로써 질환이 발생하지 않도록 교육하고 적절한 치료적 개입을 해주는 것은 매우 중요한 건강관리 시스템이 될 것으로 생각된다. 그러므로 현재 노화관련질병의 발생률을 파악한 첫 단계에서 더 나아가 연령 관련 질병 발생 메카니즘에 관한 분자생물학적 연구 및 사회 역학적 연구 등의 다각적인 연구가 필요할 것으로 생각된다.

건강노화나이에 따른 ARD 발병률과 관련 요인 분석 연구에서는 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발 3차년도에 이어 건강관리 분야에서의 건강노화나이 모델의 타당성과 유용성에 대한 후속 연구임에 의의가 있다. 분석 대상자 수와 질병 발생률의 제한으로 대표적인 6개의 ARD에 대해 분석하여 상당한 결과를 하였지만, 향후 성별, 연령별 ARD 발병률을 더 정확하게 예측하려면 연구 대상자 수를 늘려 다른 ARD 발병률을 추가적으로 분석하는 연구가 뒤따라야 할 것이다. 아울러 의료비 등을 비롯한 관련 요인 분석에서는, 향후 급여 의료비뿐만 아니라 비급여 의료비까지 포함시키고 질병 유무별 의료비를 세부적으로 예측할 수 있는 추가적인 연구가 필요할 것이다.

골다공증 예측모델에서는 일반검진 수검자에서 남녀 공통적으로 연령에 따라, 당뇨병이 있는 경우, 저체중인 경우, 또한 여성의 경우 폐경인 경우 발생률과 위험도가 높아

이전 연구들과 비슷한 결과를 보였다. 이러한 위험군에 대하여 골다공증 발생 예방을 위해 생활습관 관리 및 조기 진단과 치료가 필요할 것으로 판단된다. 골다공증 예측모델에서 machine learning과 logistic regression model은 비슷한 결과를 보였다. 여성의 예측 모델에서 예측도가 낮게 분석된 것은 골다공증 위험도와 관련된 항목이 아닌 검진 데이터 항목으로 모델을 구축하였기 때문인 것으로 판단된다. 남성에서도 모든 예측 모델에서 AUC와 Accuracy 가 높지는 않았으며, positive predictive value 가 매우 낮게 분석되었다. 추후 국가건강검진 데이터를 이용한 골다공증 예측모델 개발에서는, 건강검진 항목에 관련 factor가 많이 포함되도록 설문 추가 등의 단계적 예측 모델을 구축해야 할 것으로 판단된다.

골다공증 환자 교육 처방 메시지에서는 골다공증 예측 모델을 이용하여, 환자(또는 이용자)들의 건강행동양식에 변화를 주고자 하는 목적으로, 골다공증 환자 교육 처방 메시지를 개발하였다. 개발된 환자 교육 처방메시지는 단계별 메시지 형식으로 작성되었고, 건강 iN 웹페이지에 탑재 가능성을 고려하여, 환자(또는 이용자)의 국가 검진 결과 정보를 불러오고, 이 정보를 토대로 골다공증 발생 위험 메시지를 제공하고, 동시에 동일연령 위험메시지를 작성하였다. 실제 건강iN에 탑재하기 위해서는 구체적인 탑재방안이 있어야 할 것으로 사료된다. 또한 교육용 처방 메시지가 수검자들에게 효과적 전달되기 위해서 웹페이지 탑재 방식에 대한 충분한 논의가 필요할 수 있다. 이러한 골다공증 환자 교육용 처방 메시지를 통하여 환자 또는 이용자들이 본인들의 골다공증 위험도를 파악하고, 궁극적으로 골다공증 발생 위험을 낮출 수 있게 건강생활습관을 유지하는데 도움이 될 수 있을 것으로 사료된다.

계층적 질환군 위험조정모델은 실제로 의료비 예측 외에 기타 여러 의료 지표를 예측하는데 활용되기도 한다. 2019년 표본코호트에 대한 올바른 분석과 함께 추후 전수 분석의 연구가 필요하다. 추후 연구를 통해 구축한 고도화 된 NHIS-HCC 모델의 의료이용과 비용 분석은 추후 국가적 제도 혁신의 중요한 근거 자료가 될 것이다.

Data 결합을 통한 공공서비스 제공에 관한 연구를 통해 건강보험과 장기요양 데이터를 결합하여 의료 이용 및 장기요양 이용행태, 제공된 서비스 등을 확인하였다. 현재 한국사회보장정보원 복지 데이터 사업이 단계적으로 완성 중으로, 추후 의료와 요양 데이터에 복지 데이터(사회보장정보원의 PSR)를 연계하면 개인의 보건의료 이외에도 복지에 관한 여러 서비스의 적절성 및 방향에 대해 분석이 가능할 것이다.

제 1 장

표본 코호트에서의 ARD 분석

제1절 서론	11
제2절 이론적 고찰	12
제3절 연구 내용 및 방법	21
제4절 분석 결과	22
제5절 결론 및 제언	28

제 1 장

표본 코호트에서의 ARD 분석

제1절 서론

1. 연구 배경 및 필요성

한국은 OECD 주요 국가 중 가장 빠른 고령화 속도를 보여주고 있다.¹ 통계청 보고에 따르면 1970년부터 2018년까지 우리나라의 고령화 비율은 연평균 3.3%가 증가했고, 2026년에는 65세 이상 인구가 전체 인구에 20%를 차지하는 초고령 사회에 들어설 것으로 예상된다.² 이러한 초고령 사회로의 진입으로 인한 노인인구의 증가는 의료비용 및 건강보험의 재정 부담을 증가시키는 주요 원인이 된다.³ 일례로, 2019년 65세 이상 노인의 진료비는 전체 진료비 중 41.6%로 10년간 9.3% 증가하였으며, 1인당 연간 진료비는 491만원으로 전체 1인당 연간 진료비 168만 1천원의 3배 수준에 달하는 것으로 나타났다.⁴ 그러므로 연령이 증가함에 따라 증가하는 노화관련질병(Age related disease, ARD)을 미리 파악하여 적절히 관리하고 예방하는 일은 보건 의료 서비스 측면에서 매우 중요하다고 할 수 있겠다.

그러나 일관된 합의에 의해 정의된 ARD의 disease category가 정해져 있는 것은 없다. 연령 증가에 따라 자연스럽게 나타나는 정상적 노화 현상과 병적인 노화 현상의 구분이 명확하지 않으며, 노화에 따른 기능의 저하와 질병이 복합되어 나타나기 때문에 문헌 별로 조금씩 상이한 질환의 범주를 가진다.⁵⁻¹¹ 그러므로 합의 된 ARD의 정의에 따른 질환들을 정리하는 것은 본 과제가 목적으로 하는 개인 맞춤형 서비스 고도화를 위하여 노화 관련 질병을 예측하고 관련 요인들을 파악하는 일의 가장 중요한 첫걸음이라고 하겠다. 이를 위하여 문헌에 나타난 다양한 ARD를 정의하고 본 과제에서 새롭게 정의한 ARD에 속한 질환들의 발병률을 파악하고자 한다.

2. 연구의 목적

본 연구는, 노화 관련 질병을 예측하고 관련 요인들을 파악하여 질환의 예측 모델을 개발하고자 함에 있으며 이를 위해 가장 먼저 ARD를 재 정의하고, 정의한 ARD 질환의 발생률을 확인하여 향후 개인의 장기적인 질환 예방 및 관리에 도움이 되고자 한다.

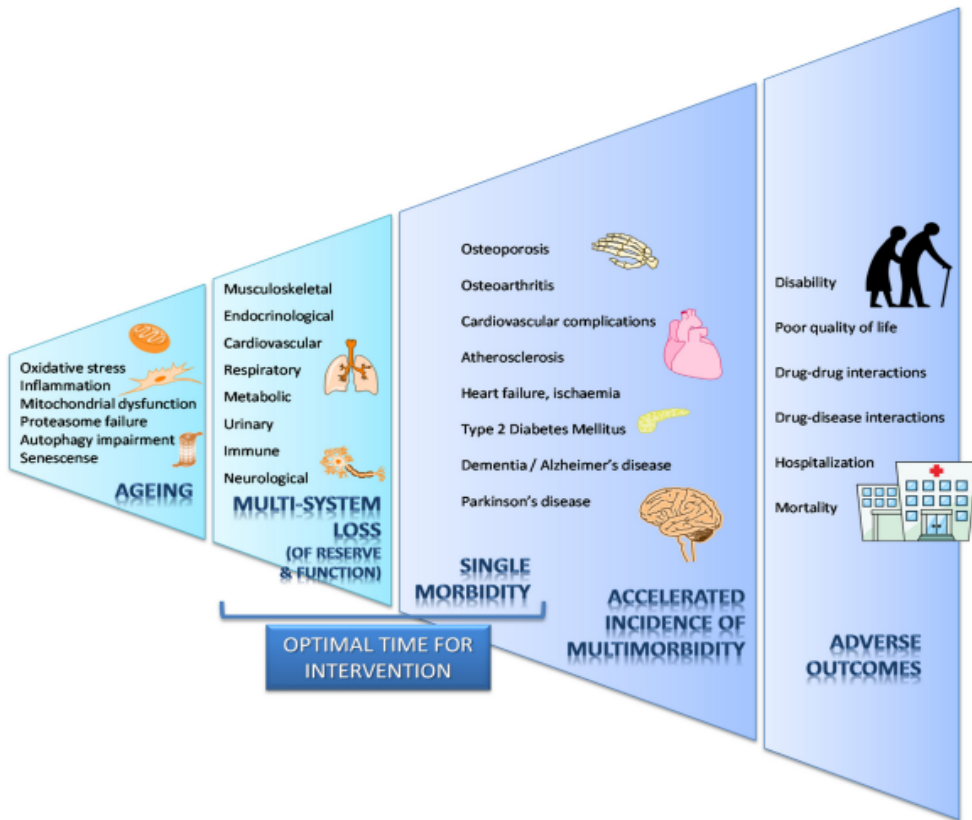
제2절 이론적 고찰

1. 노화 관련 질병(Age related disease)의 개념

1) 문헌적 고찰

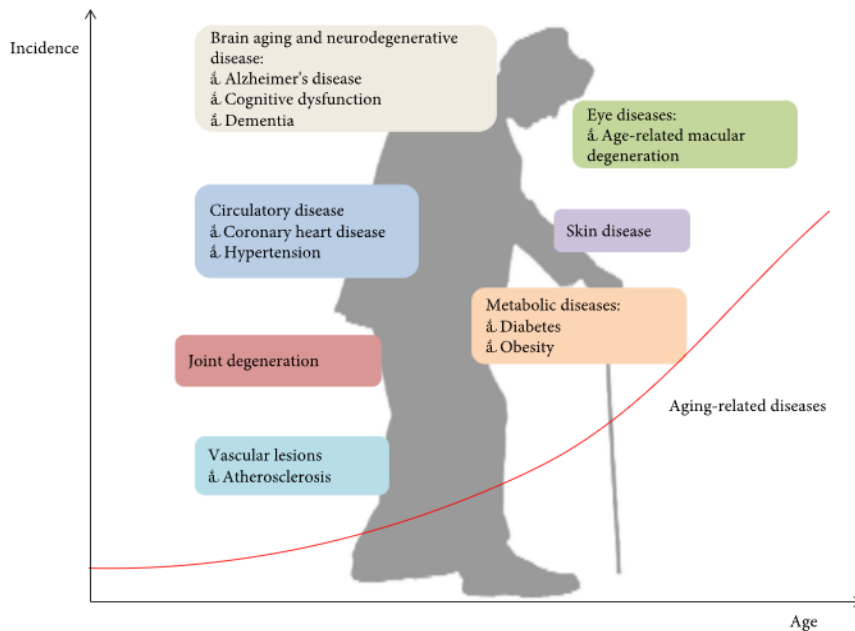
노인성 질환이란 노화에 따른 기능의 저하와 질병이 복합되어 발현되는 것으로 되어 있다.⁵ 노인성 질환은 노인증후군과 노인에서 흔한 질환으로 구분할 수 있는데, 노인증후군은 노인에서 나타나는 증상들로 치매, 섬망, 요실금, 골다공증 등이 이에 속하고, 노인에서 흔한 질환은 노인이 되기 전부터 발생할 수 있지만, 나이가 많아짐에 따라 발생률이 증가되는 질환으로 고혈압, 당뇨, 뇌졸중, 파킨슨 병 등이 포함된다.^{6,7}

아래의 그래프는, age-related alterations 에서 multi-morbidity로 가는 sequence 를 보여주고 있는데, 이에 의하면, Oxidative stress나 inflammation과 같은 노화 과정에서 나타나는 현상은 Musculoskeletal이나 cardiovascular 등 organ에 reserve 나 function의 loss를 유발할 수 있고, 이러한 multi-system loss로 하나의 질병이 발생하고 각 하나의 질병들이 진행되어 multi-morbidity를 가속화 시켜 adverse event가 발생 한다는 것을 보여주고 있다. 여기에서 말하는 age related disease는 cardiovascular complication, Atherosclerosis, Ischemia, Osteoporosis, Osteoarthritis, Type-2 diabetes mellitus, Alzheimer's disease, Parkinson disease 등이며, multi-system loss나 하나의 질병이 발생한 시점에서 적절한 intervention을 가하는 것이 향후의 다 질환 발생이나 adverse event를 발생하지 않도록 하는데 도움이 된다는 것 또한 표현하고 있다.⁸



[그림 1-1] Schematic representation of the proposed sequence of events moving from age-related alterations to multimorbidity

한 Review article에서는 나이가 증가함에 따라 incidence가 증가하는 질환들을 organ specific 하게 정리하여 ARD를 소개하였는데 이에 의하면 Brain aging and neuro- degenerative disease 카테고리에는 Alzheimer's disease나 dementia가 포함 되고, Circulatory disease에는 CHD, HTN이 Eye disease에는 Age-related macular degeneration가 그리고 joint degeneration, metabolic disease 등이 포함 되어 있다.⁹



[그림 1-2] Main aging-related disease with age and incidence

또한, Efraim Jaul and Jeremy Barron은,¹⁰ Somatic Disease and Multiple Chronic Conditions (Cardiovascular disease, Hypertension, Cancer, Osteoarthritis, Type-2 diabetes mellitus, Osteoporosis, Multiple Chronic Conditions)과 Psychological and cognitive (Cognitive aging, dementia, depression) 으로 구분하여 Age related disease를 정의 하였다.

또 다른 Review article에서는, Immune-senescence의 개념으로 age가 증가함에 따라 Cardiovascular disease (Atherosclerosis, Hypertension, Cardiac failure), Articular damage (Osteoporosis, Osteoarthritis), Metabolic disease (Type-2 diabetes mellitus, Metabolic syndrome, Obesity), Cancers (Carcinoma, Melanoma, Breast cancer), Neurologic disease (Alzheimer's disease, Parkinson disease, Depression)의 발생이 증가한다고 설명 하였다.¹¹

정리해보면, 약간의 차이는 있지만 여러 논문들에서 공통적으로 정의하는 ARD에는 대부분의 만성질환에 해당하는 CVD, cerebrovascular disease, HTN, Cancer, Osteoarthritis, T2DM, Dementia, Parkinson's disease, Osteoporosis, Cataract, AMD, Hearing loss 등이 포함되는 것을 확인할 수 있었다.

2) 연령별 질병 부담률

노인의 인구 증가에 따라 노인 관련 질환으로 인해 발생하는 질병 부담은 지속적으로 증가하는 추세이다. 세계질병부담 연구 (Global Burden of Disease, GBD 2019)에 의하면, 장애보정생존년(disability adjusted life year, DALY)을 기준으로 하여 개별 상병의 질병부담(GBD)크기를 확인하였을 때, 과거 30년간 질병부담이 증가한 10대 질병에 허혈성 심장질환(1990-2019 DALYs 증가 %: + 50%), 당뇨병(+ 148%), 뇌졸중(+ 32%), 만성신질환(+ 93%), 폐암(+ 69%), 노인성 난청(+ 83%) 등 연령이 증가함에 따라 증가하는 질환에 해당하는 질병들이 6개나 해당 되는 것을 확인 하였다.¹²

A All ages

Leading causes 1990	Percentage of DALYs 1990	Leading causes 2019	Percentage of DALYs 2019	Percentage change in number of DALYs, 1990-2019	Percentage change in age-standardised DALY rate, 1990-2019
1 Neonatal disorders	10.6 (9.9 to 11.4)	1 Neonatal disorders	7.3 (6.4 to 8.4)	-32.3 (-41.7 to -20.8)	-32.6 (-42.1 to -21.2)
2 Lower respiratory infections	8.7 (7.6 to 10.0)	2 Ischaemic heart disease	7.2 (6.5 to 7.9)	50.4 (39.9 to 60.2)	-28.6 (-32.3 to -24.2)
3 Diarrhoeal diseases	7.3 (5.9 to 8.8)	3 Stroke	5.7 (5.1 to 6.2)	32.4 (22.0 to 42.2)	-35.2 (-40.5 to -30.5)
4 Ischaemic heart disease	4.7 (4.4 to 5.0)	4 Lower respiratory infections	3.8 (3.3 to 4.3)	-56.7 (-64.2 to -47.5)	-62.5 (-69.0 to -54.9)
5 Stroke	4.2 (3.9 to 4.5)	5 Diarrhoeal diseases	3.2 (2.6 to 4.0)	-57.5 (-66.2 to -44.7)	-64.6 (-71.7 to -54.2)
6 Congenital birth defects	3.2 (2.3 to 4.8)	6 COPD	2.9 (2.6 to 3.2)	25.6 (15.1 to 46.0)	-39.8 (-44.9 to -30.2)
7 Tuberculosis	3.1 (2.8 to 3.4)	7 Road injuries	2.9 (2.6 to 3.0)	2.4 (-6.9 to 10.8)	-31.0 (-37.1 to -25.4)
8 Road injuries	2.7 (2.6 to 3.0)	8 Diabetes	2.8 (2.5 to 3.1)	147.9 (135.9 to 158.9)	24.4 (18.5 to 29.7)
9 Measles	2.7 (0.9 to 5.6)	9 Low back pain	2.5 (1.9 to 3.1)	46.9 (43.3 to 50.5)	-16.3 (-17.1 to -15.5)
10 Malaria	2.5 (1.4 to 4.1)	10 Congenital birth defects	2.1 (1.7 to 2.6)	-37.3 (-50.6 to -12.8)	-40.0 (-52.7 to -17.1)
11 COPD	2.3 (1.9 to 2.5)	11 HIV/AIDS	1.9 (1.6 to 2.2)	127.7 (97.3 to 171.7)	58.5 (37.1 to 89.2)
12 Protein-energy malnutrition	2.0 (1.6 to 2.7)	12 Tuberculosis	1.9 (1.7 to 2.0)	-41.0 (-47.2 to -33.5)	-62.8 (-66.6 to -58.0)
13 Low back pain	1.7 (1.2 to 2.1)	13 Depressive disorders	1.8 (1.4 to 2.4)	61.1 (56.9 to 65.0)	-1.8 (-2.9 to -0.8)
14 Self-harm	1.4 (1.2 to 1.5)	14 Malaria	1.8 (0.9 to 3.1)	-29.4 (-56.9 to 6.6)	-37.8 (-61.9 to -6.2)
15 Cirrhosis	1.3 (1.2 to 1.5)	15 Headache disorders	1.8 (0.4 to 3.8)	56.7 (52.4 to 62.1)	1.1 (-4.2 to 2.9)
16 Meningitis	1.3 (1.1 to 1.5)	16 Cirrhosis	1.8 (1.6 to 2.0)	33.0 (22.4 to 48.2)	-26.8 (-32.5 to -19.0)
17 Drowning	1.3 (1.1 to 1.4)	17 Lung cancer	1.8 (1.6 to 2.0)	69.1 (53.1 to 85.4)	-16.2 (-24.0 to -8.2)
18 Headache disorders	1.1 (0.2 to 2.4)	18 Chronic kidney disease	1.6 (1.5 to 1.8)	93.2 (81.6 to 105.0)	6.3 (0.2 to 12.4)
19 Depressive disorders	1.1 (0.8 to 1.5)	19 Other musculoskeletal	1.6 (1.2 to 2.1)	128.9 (122.0 to 136.3)	30.7 (27.6 to 34.3)
20 Diabetes	1.1 (1.0 to 1.2)	20 Age-related hearing loss	1.6 (1.2 to 2.1)	82.8 (75.2 to 88.9)	-1.8 (-3.7 to -0.1)
21 Lung cancer	1.0 (1.0 to 1.1)	21 Falls	1.5 (1.4 to 1.7)	47.1 (31.5 to 61.0)	-14.5 (-22.5 to -7.4)
22 Falls	1.0 (0.9 to 1.2)	22 Self-harm	1.3 (1.2 to 1.5)	-5.6 (-14.2 to 3.7)	-38.9 (-44.3 to -33.0)
23 Dietary iron deficiency	1.0 (0.7 to 1.3)	23 Gynaecological diseases	1.2 (0.9 to 1.5)	48.7 (45.8 to 51.8)	-6.8 (-8.7 to -4.9)
24 Interpersonal violence	0.9 (0.9 to 1.0)	24 Anxiety disorders	1.1 (0.8 to 1.5)	53.7 (48.8 to 59.1)	-0.1 (-1.0 to 0.7)
25 Whooping cough	0.9 (0.4 to 1.7)	25 Dietary iron deficiency	1.1 (0.8 to 1.5)	13.8 (10.5 to 17.2)	-16.4 (-18.7 to -14.0)
27 Age-related hearing loss	0.8 (0.6 to 1.1)	26 Interpersonal violence	1.1 (1.0 to 1.2)	10.2 (3.2 to 19.2)	-23.8 (-28.6 to -17.8)
29 Chronic kidney disease	0.8 (0.8 to 0.9)	40 Meningitis	0.6 (0.5 to 0.8)	-51.3 (-59.4 to -42.0)	-57.2 (-64.4 to -48.6)
30 HIV/AIDS	0.8 (0.6 to 1.0)	41 Protein-energy malnutrition	0.6 (0.5 to 0.7)	-71.1 (-79.6 to -59.7)	-74.5 (-82.0 to -64.5)
32 Gynaecological diseases	0.8 (0.6 to 1.0)	46 Drowning	0.5 (0.5 to 0.6)	-60.6 (-65.2 to -53.6)	-68.2 (-71.9 to -62.8)
34 Anxiety disorders	0.7 (0.5 to 1.0)	55 Whooping cough	0.4 (0.2 to 0.7)	-54.5 (-74.6 to -16.9)	-56.3 (-75.6 to -20.3)
35 Other musculoskeletal	0.7 (0.5 to 1.0)	71 Measles	0.3 (0.1 to 0.6)	-89.8 (-92.3 to -86.8)	-90.4 (-92.8 to -87.5)

[그림 1-3] Leading causes of global DALYs and percentage of total DALYs (1990 and 2019), and percentage change in number of DALYs and age-standardised DALY rates from 1990 to 2019

또한, 2015년 한국인의 질병 부담을 살펴보면, 한국인의 질병 부담은 10만명당 29,476 DALYs로 산출되었고, 질환별로는 Low back pain, Diabetes mellitus, chronic obstructive pulmonary disease, Ischemic heart disease, Osteoarthritis,

Ischemic stroke의 순으로 비감염성질환(Non-communicable disease, NCD)에서의 질병 부담이 높게 나타난 것을 알 수 있다. 또한, 연령 별 질병 부담의 순위가 높은 질환들을 보면 연령이 증가할수록 Diabetes mellitus, Dementia, Ischemic heart disease, Stroke들이 나타나는 것을 확인 할 수 있다.¹³

<표 1-1> Top 5 leading specific causes of DALYs by age group in Korea, 2015

Age group, yr	Rank	Specific cause	% of age group total DALYs
0-9	1	Refraction and accommodation disorders	13.6
	2	Asthma	13.1
	3	Dental caries	12.3
	4	Viral skin diseases	8.3
	5	COPD	7.5
10-19	1	Low back pain	7.5
	2	Falls	5.4
	3	COPD	5.3
	4	Epilepsy	4.0
	5	Major depressive disorders	3.9
20-29	1	Low back pain	10.0
	2	Motorized vehicle with three or more wheels	5.7
	3	Schizophrenia	4.1
	4	Falls	3.7
	5	Self-harm	3.7
30-39	1	Low back pain	12.1
	2	Diabetes mellitus	8.7
	3	Abortion	4.3
	4	COPD	4.3
	5	Infertility	3.6
40-49	1	Low back pain	13.4
	2	Diabetes mellitus	12.1
	3	Cirrhosis of the liver	4.5
	4	COPD	3.4
	5	Osteoarthritis	3.3
50-59	1	Diabetes mellitus	11.1
	2	Low back pain	10.3
	3	Osteoarthritis	7.2
	4	Ischemic heart disease	5.1
	5	COPD	4.6
60-69	1	Diabetes mellitus	8.3
	2	Low back pain	7.5
	3	Osteoarthritis	6.8
	4	Ischemic heart disease	6.8
	5	Ischemic stroke	6.1
70-79	1	Alzheimer's disease and other dementias	9.4
	2	Ischemic stroke	8.2
	3	Ischemic heart disease	7.1
	4	Diabetes mellitus	5.3
	5	COPD	5.2
≥ 80	1	Alzheimer's disease and other dementias	19.0
	2	Ischemic stroke	9.8
	3	Ischemic heart disease	8.1
	4	COPD	5.5
	5	Falls	3.9

DALY = disability-adjusted life year, COPD = chronic obstructive pulmonary disease.

따라서, ARD를 정의함에 있어서 이러한 질병 부담의 관점에서 접근을 하여, 연령이 증가함에 따라 질병 부담이 증가하고 발생률이 증가하는 질환에 초점을 맞춰 재정의 하는 것이 필요할 것으로 고려된다.

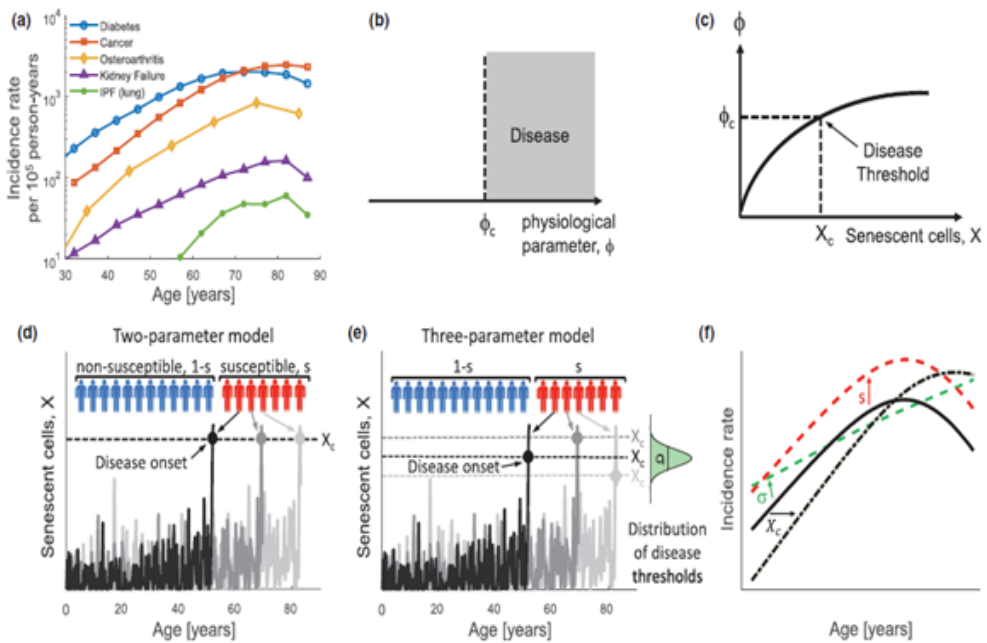
2. 노화 관련 질병(Age related disease, ARD)의 특징

노화관련질병은 일반적으로 연령이 증가함에 따라 증가하는 질병들을 말한다.¹⁴ 앞에서 언급하였듯이 조금씩 다른 질병들이 포함되고, 각 질병은 mutations, dysregulated homeostasis, fibrosis, degenerative processes등 각기 다른 원인으로 다양한 장기에 영향을 끼쳐 발생한다.⁸ 그러나 이러한 병리학의 차이점에도 불구하고 노화관련질병은 발병률 측면에서 보편적인 특징이 있다. 질병의 발병률을 연간 새롭게 발생한 질병의 수를 인구의 수로 나눈 값으로 정의 하였을 때, 연령 관련 질병의 발병률은 연령이 증가함에 따라 거의 지수함수의 형태로 증가한다. 또한, 이러한 발병률 곡선의 상승 부분의 기울기는 많은 연령 관련 질병에서 연간 6-8% 범위에서 유사하게 나타난다. 그러나 많은 질병의 경우 very old age에서는 발병률이 오히려 떨어지는 것이 관찰된다.^{15,16} 이러한 유사성은 다양한 질병의 발병을 지배하는 일반적인 생물학적 노화 과정이 있음을 암시한다고 볼 수 있다.¹⁴

이러한 노인관련질병의 지수함수적 발병률 증가 및 늦은 감소 [그림 1-4, (a)]는 질병의 발병을 유발하는 노화세포(Senescent cell)의 축적과 집단에서 개인들의 질병에 대한 감수성 차이로 설명될 수 있다.

노화세포는 다양한 스트레스에 대한 반응으로 세포 분열이 중단된 세포를 말하며 노화 관련 분비물들(Senescence Associated Secretion Profile, SASP)을 분비하여 염증을 유발하거나 정상 세포의 재생산을 억제한다.¹⁷ 노화 세포는 젊을 때에는 만들어지고 제거되는 turnover가 빠르게 나타나지만, 연령이 증가함에 따라 다양한 스트레스에 의한 생성은 증가하고, 제거되는 속도는 스스로 늦추기 때문에 연령이 증가할수록 기하함수적으로 증가하게 된다(Saturated-Removal model, SR model).¹⁴ 만일, 어떠한 질병이 physiological parameter가 임계점(Threshold) 이상을 넘으면 발생한다고 가정을 하면 [그림 1-4, (b)], 노화 세포는 SASP등을 분비하여 physiologic parameter에 영향을 미쳐 parameter들이 질환을 일으키는 임계점(Threshold)을 넘게 되어 질환이 발병된다[그림 1-4, (c)]. 따라서 연령 증가에 따라 노화세포가 지수함수로 증가하는 것은 연령 증가에 따라 질환이 지수함수로 증가되는 것과 관련이 있다. 그리고 매우 높은 연령에서 질환의 발병률이 감소하는 것은 population heterogeneity로 설명 할 수 있다. 각 질병은

genetic 또는 environment factor의 차이로 개인 마다 질병 발생의 감수성이 달라 질병 위험이 다르게 나타날 수 있다. 그러므로 질병의 발생, 질환의 발생 임계점이 낮은 사람들 [그림 1-4, (d), (e) susceptible fraction, s]에서부터 나타날 것이고, 질환의 발생 임계점이 높은 사람들 [그림 1-4, (d), (e) non-susceptible fraction, $1-s$]은 정상 노화에서는 질병이 발생하지 않을 것이다. 그러므로 매우 높은 연령에 있는 사람들의 경우 거의 이미 질병에 이환 되어 있는 상태이고, 남아있는 사람들은 질환의 발생 임계점이 높은 사람들이 대부분일 것이므로 새롭게 질병이 발생할 확률은 오히려 감소하게 되는 것이다[그림 1-4, (f)].¹⁴



[그림 1-4] Diseases caused by threshold-crossing of a parameter affected by senescent cells are predicted to have an exponential incidence curve with a decline at old ages.

3. 노화 관련 질병(Age related disease, ARD) 정의

본 연구진은, 노화 관련 질병이란 ‘연령이 증가함에 따라 질병부담이 증가하고, 발생률이 지수함수적으로 증가하는 질환이다’라고 정의하였다. 이에 따라, Lancet Public Health에서 보고한 92개 <표 1-2>의 Age related disease와 (세계 질병부담(GBD)에

있는 293가지 질병 중에서 연령에 따라 발생률 혹은 유병률이 증가 하는 질환이 145개를 제외하고 남은 질병들 중 연령 증가에 따라 발생률 및 유병률이 지수함수적으로 증가하지 않는 경우를 제외하여 추출함),¹⁸ 2020년 보건복지부 노인실태조사를 확인하여 한국 노인인구에서 유병률이 높은 질환들을 확인하여¹⁹ 하기와 같이 14개의 질환을 선정하였으며, 이번 연구에서는 만성질환에 초점을 맞춰 노화관련질환의 대표적인 질환인 ‘암’의 경우는 선정에서 제외하였다.

선정 된 ARD: 1. 고혈압, 2. 당뇨병, 3. 고지혈증, 4. 뇌혈관계질환, 5. 허혈성심질환, 6. 골다공증, 7. 골관절염, 8. 만성폐쇄성폐질환, 9. 울혈성심부전, 10. 만성신질환, 11. 백내장, 12. 황반변성, 13. 난청, 14. 파킨슨병

<p>Cardiovascular diseases Atrial fibrillation and flutter; endocarditis; hypertensive heart disease; intracerebral haemorrhage; ischaemic heart disease; ischaemic stroke; myocarditis; non-rheumatic calcific aortic valve disease; non-rheumatic degenerative mitral valve disease; other cardiomyopathy; other cardiovascular and circulatory diseases; other non-rheumatic valvular diseases; peripheral artery disease</p> <p>Chronic respiratory diseases Asbestosis; chronic obstructive pulmonary disease; coal worker pneumoconiosis; interstitial lung disease and pulmonary sarcoidosis; other pneumoconiosis; silicosis</p> <p>Communicable, maternal, neonatal, and nutritional diseases Diarrhoeal diseases; encephalitis; lower respiratory infections; pneumococcal meningitis; trachoma</p> <p>Diabetes and kidney diseases Chronic kidney disease due to type 2 diabetes mellitus; chronic kidney disease due to glomerulonephritis; chronic kidney disease due to other and unspecified causes</p> <p>Digestive diseases Cirrhosis due to NASH; pancreatitis; paralytic ileus and intestinal obstruction; peptic ulcer disease; vascular intestinal disorders</p> <p>Injuries Drowning; environmental heat and cold exposure; falls; foreign body in other body part; other transport injuries; other unintentional injuries</p>	<p>Neoplasms Acute lymphoid leukaemia; acute myeloid leukaemia; benign and in-situ intestinal neoplasms; bladder cancer; brain and nervous system cancer; breast cancer; chronic lymphoid leukaemia; chronic myeloid leukaemia; colon and rectum cancer; gallbladder and biliary tract cancer; Hodgkin lymphoma; kidney cancer; larynx cancer; lip and oral cavity cancer; liver cancer due to NASH; liver cancer due to alcohol use; liver cancer due to hepatitis C; malignant skin melanoma; mesothelioma; multiple myeloma; myelodysplastic, myeloproliferative, and other hematopoietic neoplasms; non-Hodgkin lymphoma; non-melanoma skin cancer (basal-cell carcinoma); non-melanoma skin cancer (squamous-cell carcinoma); oesophageal cancer; other benign and in-situ neoplasms; other leukaemia; other malignant neoplasms; ovarian cancer; pancreatic cancer; prostate cancer; stomach cancer; thyroid cancer; tracheal, bronchus, and lung cancer; uterine cancer</p> <p>Neurological disorders Alzheimer's disease and other dementias; motor neuron disease; Parkinson's disease</p> <p>Other non-communicable diseases Congenital musculoskeletal and limb anomalies; digestive congenital anomalies; endocrine, metabolic, blood, and immune disorders; other haemoglobinopathies and haemolytic anaemias</p> <p>Sense organ diseases Age-related and other hearing loss; age-related macular degeneration; cataract; glaucoma; other sense organ diseases; other vision loss; refraction disorders</p> <p>Skin and subcutaneous diseases Cellulitis; decubitus ulcer; fungal skin diseases; other skin and subcutaneous diseases; pyoderma</p>
--	---

[그림 1-5] 92 age-related disease

<표 1-2> 노인의 성별 만성질환 종류별 유병률 및 현 치료율

(단위: %)

특성	전체		남자		여자		2017년도 유병률 ²⁾
	유병률	치료율	유병률	치료율	유병률	치료율	
고혈압	56.8	98.7	57.0	99.0	56.6	98.5	59.0
뇌졸중(중풍, 뇌경색)	4.3	96.5	4.8	96.8	3.8	96.2	7.1
고지혈증(이상지질혈증)	17.1	96.9	13.5	96.1	19.9	97.3	29.5
협심증, 심근경색증	4.4	96.9	4.8	97.3	4.2	96.6	7.0
기타 심장질환	4.5	98.2	4.2	99.0	4.8	97.7	6.6
당뇨병	24.2	99.0	23.9	99.4	24.4	98.7	23.2
갑상선 질환	3.3	97.3	1.3	97.2	4.9	97.3	3.3
골관절염 또는 류머티즘 관절염	16.5	91.1	7.4	87.3	23.3	92.1	33.1
골다공증	8.5	90.1	1.7	87.4	13.7	90.3	13.0
요통, 좌골신경통	10.0	82.1	4.3	82.5	14.3	82.0	24.1
만성기관지염, 폐기종(CO모노트D)	1.5	91.8	2.2	94.2	0.9	87.7	1.5
천식	2.0	94.3	2.4	93.3	1.7	95.4	3.1
폐결핵, 결핵	0.1	82.9	0.2	74.9	0.1	100.0	0.3
백내장	5.3	74.1	6.0	76.8	4.8	71.5	7.1
녹내장	1.1	81.1	1.2	86.7	1.0	76.2	2.4
만성중이염	0.7	98.1	1.0	100.0	0.5	95.3	0.4
암(악성신생물)	1.8	88.5	2.4	90.2	1.4	86.3	3.8
위·십이지장궤양	5.0	93.6	4.9	95.8	5.0	92.0	9.1
간염	0.4	84.1	0.4	75.7	0.3	94.1	0.8
간경변	0.3	90.1	0.6	87.1	0.1	100.0	0.5
만성신장질환	1.1	95.1	1.4	91.8	0.9	98.8	1.9
전립선비대증	4.1	96.2	9.3	96.2	0.1	100.0	8.9
요실금	3.9	43.7	0.5	43.7	6.5	43.7	2.5
빈혈	1.8	75.9	0.7	77.0	2.5	75.6	2.1
피부병	0.9	84.3	1.1	91.0	0.7	76.3	1.9
우울증	1.5	81.3	0.8	90.7	2.1	78.4	3.0
치매	2.1	93.8	1.9	95.2	2.2	93.0	1.5
골절, 탈골 및 사고 후유증	1.5	87.0	1.1	86.7	1.8	87.1	1.7
불면증	2.0	78.9	1.1	78.3	2.6	79.1	3.5
파킨슨	0.6	100.0	0.7	100.0	0.4	100.0	1.0
노인성 난청	2.5	53.9	2.6	59.6	2.4	49.2	-
기타	1.6	94.4	1.8	93.8	1.4	95.1	-

주: 1) 전체응답자를 대상으로 한 분석결과임. / 2) 본인응답자를 대상으로 한 분석결과임.

제3절 연구 내용 및 방법

1. 연구 내용

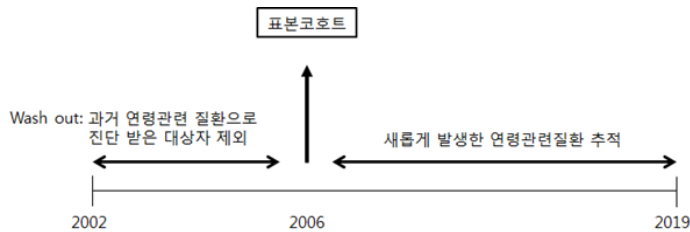
앞에서 선정한 14가지의 노화관련질병들이 한국인에서 어떠한 형태로 발병하는지 특징을 확인해보기 위한 연구를 진행하였다. 전 국민을 대표할 수 있는 국민건강보험공단의 표본코호트DB를 활용하여 노화관련질병의 연도별 발생률과 연령별 발생률을 확인하였다.

2. 연구 방법

1) 데이터셋 구축 및 분석방법

표본코호트DB는 2006년 1년 간 건강보험 및 의료급여 수급권자 자격을 유지한 전국민 중 100만 명을 표본으로 추출한 자료이며, 추가 표본으로 보험료 분위 결측군과 2006년부터 매년 신생아 표본이 추가되고 있다. 노화관련질병의 발생 여부를 확인하기 위해 명세서테이블(T20)의 주상병, 부상병 총 2개 항목에 포함된 상병분류코드를 사용 하였다. 상병코드는 KCD-8 진단코드 범주를 기준으로 의무기록 전문가, 임상과의 회의를 통해 선정하였다.

연도별 발생률은 각 노화관련질병별로 2002년부터 2005년까지 각 노화관련질병을 진단받은 이력이 있는 대상자를 제외한 후 2006년부터 2019년 사이 총 14년 동안 각 연도에 각 노화관련질병이 새롭게 발생한 대상자 수를 분자로 하고, 이전 연도에 각 노화관련질병이 발생하지 않은 대상자 수를 분모로 하여 발생률을 산출하였고, 연령별 발생률은 2019년 기준 연령별 10세 단위로 나누어 각 연령 그룹별 노화관련질병 발생한 대상자 수를 분자로 하고, 2019년 각 연령별 대상자의 수를 분모로 하여 연령 그룹별 각 노화관련질병 발생률을 산출하였다. 데이터 전처리와 발생률 산출을 위해 SAS 9.4를 이용하였다.



[그림 1-6] 표본코호트 노화관련질병 연구 대상자 선정

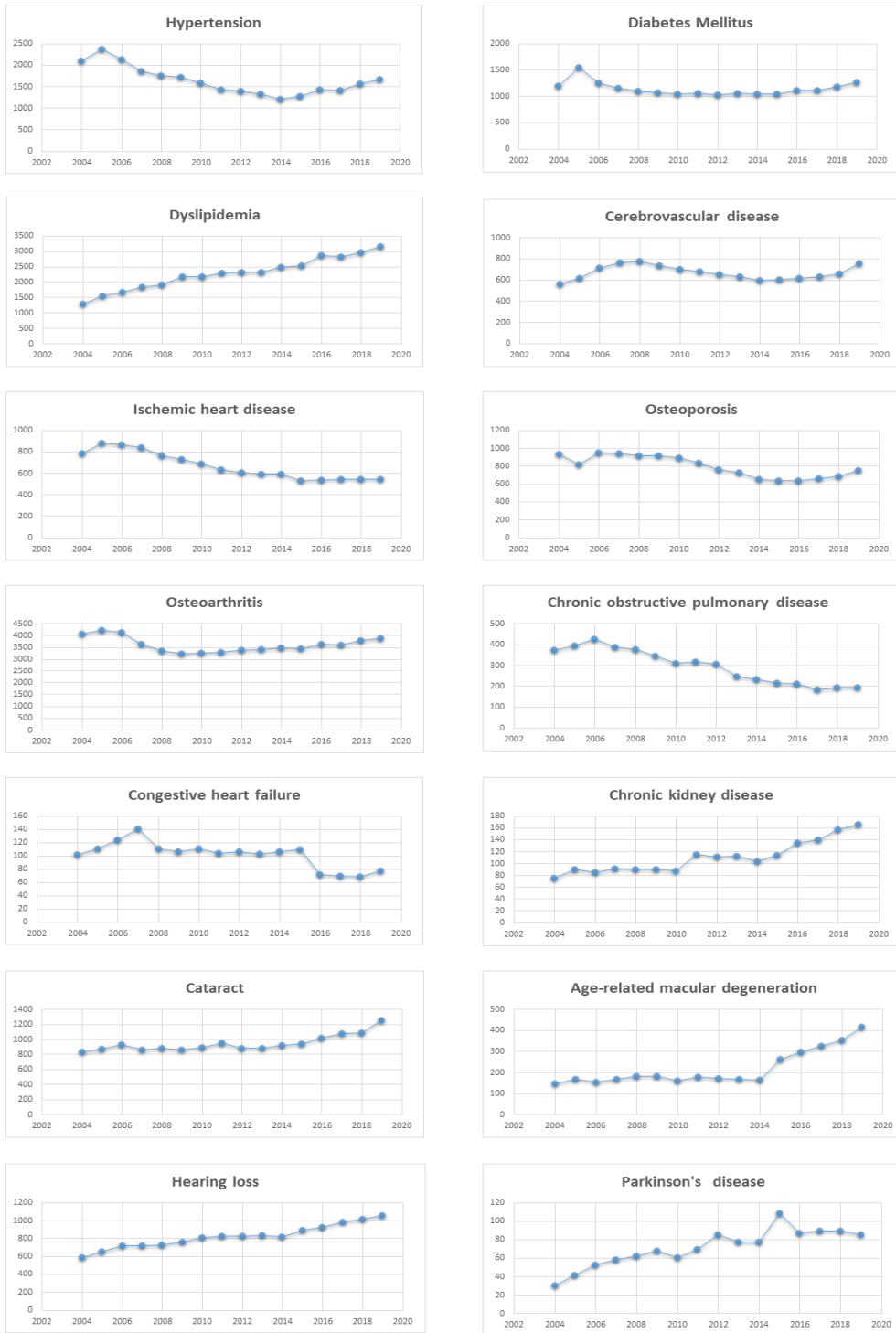
<표 1-3> 노화관련질병 및 주상병 코드

질환명	주상병코드
고혈압	I10, I11, I12, I13
당뇨	E10, E11, E12, E13, E14
고지혈증	E780, E781, E782, E783, E784, E785
뇌혈관계질환	I60, I61, I62, I63, I64, I65, I66, I67, I68, I69
허혈성 심질환	I20, I21, I22, I23, I24, I25
골다공증	M80, M81
골관절염	M00, M01, M02, M03, M04, M05, M06, M07, M08, M09, M10, M11, M12, M13, M14, M15, M16, M17, M18, M19
만성폐쇄성폐질환	J44
울혈성심부전	I500
만성신질환	N18, I12
백내장	H25, H26
황반변성	H353, H353
난청	H90, H91
파킨슨병	G20, G21

제4절 분석 결과

1. 연도별 발생률

10만명당 연도별 발생률을 질병에 따라 살펴보면, 고혈압은 2006년부터 점차 감소되다가 2014년에 최저 발생률에서 반등하여 점차 발생률이 증가하는 경향이 나타났다. 당뇨병의 경우 발생률은 점차 감소하다가 2018년부터 약간 증가 하였으나 증감폭이 크게 나타나지는 않았고, 고지혈증의 경우는 지속적으로 발생률이 상승하였다. 뇌혈관계 질환은 2008년도 까지는 증가하다가 이후 감소하는 추세에서 2016년부터는 점차 증가하는 경향을 보이며, 허혈성심장질환은 점차 감소하는 추세이다. 골다공증은 역시 점차 감소되다가 2016년 이후 증가하는 경향을 보였다. 골관절염은 2009년까지 감소하다가 이후 완만한 증가추세를 보였고, 만성폐쇄성폐질환의 발생률은 점차 감소하였다. 울혈성 심부전의 경우 2015년 이후 발생률이 감소된 후 비슷하게 유지되며, 만성신질환은 완만한 발생률을 보이다가 2015년 이후 발생률이 증가하였다. 백내장의 경우 지속적으로 완만하게 발생률이 증가하였으며, 연령 관련 황반변성은 비슷하게 유지되던 발병률이 2014년 이후 급속하게 증가되었고, 난청은 연도별로 완만하게 발생률이 증가되어 나타났으며, 파킨슨병의 경우 발생률 자체는 높지 않았으며 연도별로 상이하지만 전체적으로는 증가하는 추세를 보였다.



[그림 1-7] 연도별 발생률

<표 1-4> 연도별 발생률

고혈압	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	2,127	1,856	1,762	1,724	1,593	1,429	1,402	1,323	1,207	1,279	1,437	1,423	1,563	1,673
전체 대상자 수	920,747	908,912	897,268	887,432	878,180	869,934	865,101	859,197	854,180	852,055	849,232	842,585	836,784	829,077
발생자 수	19,581	16,873	15,810	15,296	13,987	12,435	12,133	11,363	10,310	10,896	12,205	11,987	13,082	13,867
발생률	2.13	1.86	1.76	1.72	1.59	1.43	1.40	1.32	1.21	1.28	1.44	1.42	1.56	1.67

당뇨	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	1,251	1,151	1,097	1,071	1,046	1,061	1,034	1,056	1,041	1,050	1,112	1,111	1,188	1,275
전체 대상자 수	967,764	965,139	958,460	952,784	948,377	944,360	940,484	936,565	932,507	929,035	925,146	919,108	913,200	906,226
발생자 수	12,105	11,111	10,513	10,202	9,920	10,018	9,729	9,894	9,708	9,757	10,291	10,213	10,848	11,554
발생률	1.25	1.15	1.10	1.07	1.05	1.06	1.03	1.06	1.04	1.05	1.11	1.11	1.19	1.27

고지혈증	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	1,681	1,846	1,915	2,180	2,169	2,306	2,312	2,323	2,491	2,533	2,867	2,832	2,965	3,169
전체 대상자 수	983,471	978,946	965,438	953,561	937,722	923,481	907,638	891,399	877,282	861,210	846,926	825,942	806,693	786,955
발생자 수	16,536	18,073	18,490	20,783	20,335	21,296	20,986	20,710	21,854	24,279	23,388	23,919	24,939	24,939
발생률	1.68	1.85	1.92	2.18	2.17	2.31	2.31	2.32	2.49	2.53	2.87	2.83	2.97	3.17

뇌혈관계질환	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	713	764	780	738	699	679	654	632	597	601	619	632	657	758
전체 대상자 수	1,005,132	1,008,944	1,006,256	1,003,050	1,001,241	999,757	999,073	997,739	996,921	996,995	996,324	994,413	991,636	988,869
발생자 수	7,171	7,707	7,850	7,402	6,994	6,789	6,536	6,308	5,951	5,995	6,169	6,285	6,515	7,491
발생률	0.71	0.76	0.78	0.74	0.70	0.68	0.65	0.63	0.60	0.60	0.62	0.63	0.66	0.76

허혈성심질환	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	871	843	765	729	691	632	608	596	593	529	542	542	543	544
전체 대상자 수	993,762	995,291	990,594	987,171	985,102	982,905	982,193	980,907	980,269	979,286	978,704	976,928	974,277	971,133
발생자 수	8,653	8,393	7,577	7,200	6,805	6,214	5,974	5,842	5,809	5,184	5,305	5,297	5,286	5,285
발생률	0.87	0.84	0.76	0.73	0.69	0.63	0.61	0.60	0.59	0.53	0.54	0.54	0.54	0.54

골다공증	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	954	943	914	918	897	832	760	730	652	636	640	660	691	752
전체 대상자 수	991,039	991,929	986,442	981,770	977,914	973,564	970,424	967,579	965,015	964,098	962,601	960,280	957,155	953,415
발생자 수	9,456	9,357	9,017	9,012	8,771	8,102	7,372	7,060	6,292	6,135	6,164	6,337	6,617	7,173
발생률	0.95	0.94	0.91	0.92	0.90	0.83	0.76	0.73	0.65	0.64	0.64	0.66	0.69	0.75

골관절염	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	4,134	3,639	3,368	3,240	3,276	3,310	3,389	3,419	3,483	3,470	3,651	3,605	3,791	3,895
전체 대상자 수	874,646	843,531	815,805	792,941	774,316	755,913	738,821	720,612	703,395	685,954	670,024	650,631	633,224	614,019
발생자 수	36,161	30,698	27,480	25,690	25,370	25,017	25,039	24,641	24,501	23,805	24,460	23,458	24,003	23,917
발생률	4.13	3.64	3.37	3.24	3.28	3.31	3.39	3.42	3.48	3.47	3.65	3.61	3.79	3.90

만성폐쇄성폐질환	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	428	387	377	347	311	319	307	247	231	215	209	184	193	194
전체 대상자 수	1,008,975	1,014,715	1,014,958	1,015,170	1,016,477	1,018,279	1,020,378	1,021,068	1,023,158	1,025,554	1,027,235	1,027,772	1,028,068	1,027,514
발생자 수	4,317	3,930	3,825	3,527	3,164	3,244	3,133	2,523	2,360	2,202	2,150	1,887	1,987	1,990
발생률	0.43	0.39	0.38	0.35	0.31	0.32	0.31	0.25	0.23	0.21	0.21	0.18	0.19	0.19

울혈성심부전	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	124	141	111	106	111	104	106	103	106	110	72	70	69	77
전체 대상자 수	1,017,969	1,027,330	1,029,461	1,032,154	1,035,853	1,038,972	1,042,775	1,045,416	1,048,484	1,051,746	1,053,372	1,054,804	1,055,281	1,055,277
발생자 수	1,267	1,448	1,142	1,092	1,152	1,084	1,108	1,072	1,109	1,155	756	735	725	815
발생률	0.12	0.14	0.11	0.11	0.11	0.10	0.11	0.10	0.11	0.11	0.07	0.07	0.07	0.08

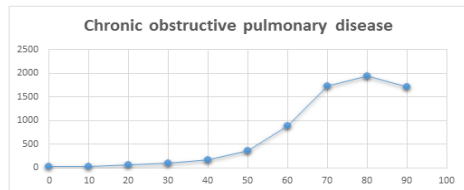
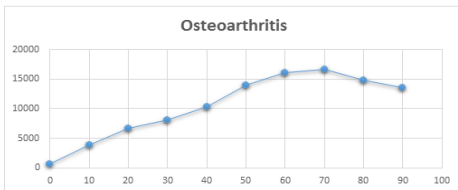
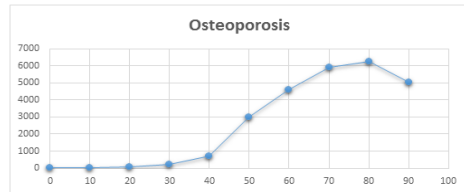
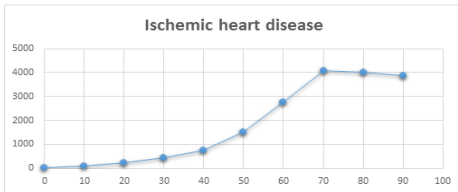
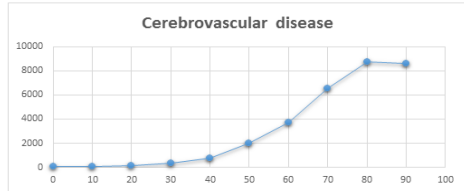
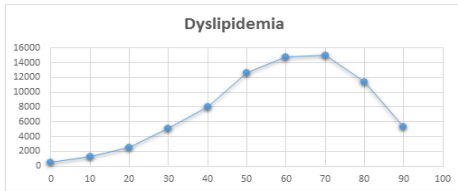
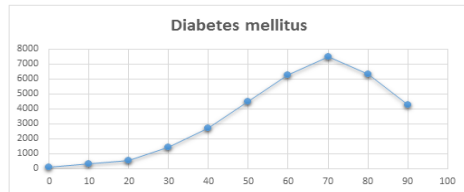
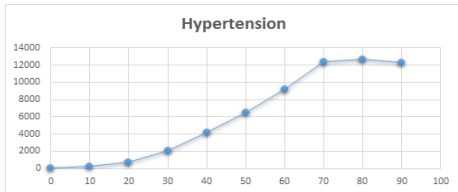
만성신질환	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	85	91	90	90	88	115	111	113	104	114	135	140	157	167
전체 대상자 수	1,018,323	1,027,969	1,030,743	1,033,510	1,037,106	1,040,588	1,044,006	1,046,462	1,049,144	1,052,300	1,054,361	1,055,095	1,054,976	1,054,008
발생자 수	866	935	926	934	914	1,194	1,163	1,179	1,088	1,203	1,424	1,476	1,655	1,755
발생률	0.09	0.09	0.09	0.09	0.09	0.11	0.11	0.11	0.10	0.11	0.14	0.14	0.16	0.17

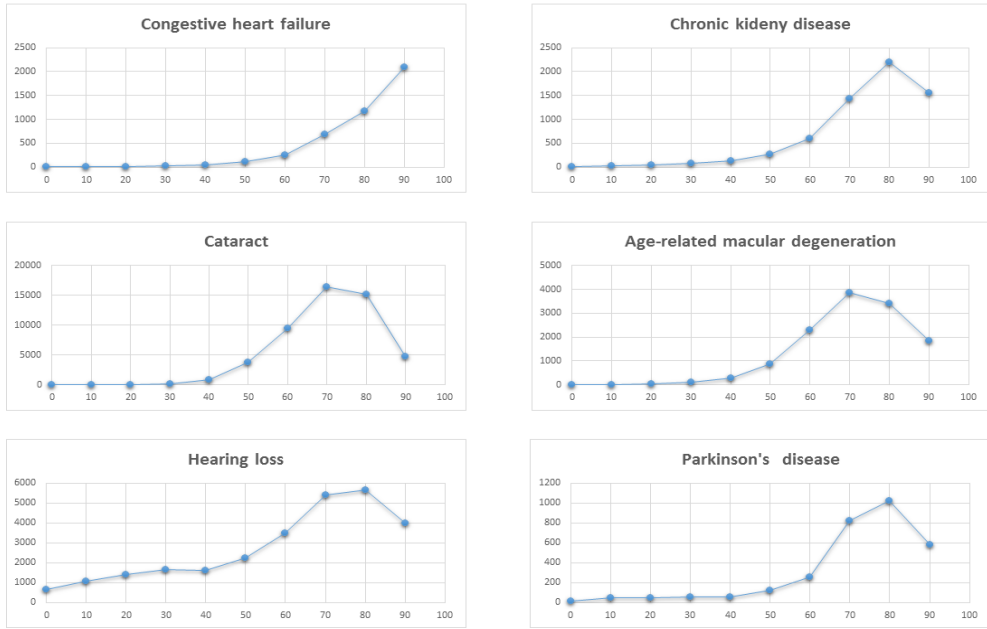
백내장	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
발생률 (10만명당)	928	867	885	868	893	955	884	882	927	942	1,024	1,077	1,088	1,260
전체 대상자 수	993,521	994,138	990,302	986,076	983,524	980,933	977,087	973,825	971,544	968,796	965,773	960,740	954,191	948,574
발생자 수	9,224	8,619	8,767	8,555	8,783	9,369	8,640	8,586	9,002	9,128	9,886	10,348	10,384	11,952
발생률	0.93	0.87	0.89	0.87	0.89	0.96	0.88	0.88	0.93	0.94	1.02	1.08	1.09	1.26

항목	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
황반변성														
발생률 (10만명당)	154	167	184	181	163	179	174	167	164	263	299	327	354	416
전체 대상자 수	1,017,115	1,026,146	1,028,181	1,029,794	1,032,231	1,034,751	1,037,393	1,039,063	1,041,194	1,044,611	1,045,297	1,044,579	1,042,582	1,040,138
발생자 수	1,566	1,716	1,893	1,863	1,681	1,857	1,802	1,739	1,707	2,752	3,125	3,418	3,688	4,322
발생률	0.15	0.17	0.18	0.18	0.16	0.18	0.17	0.17	0.16	0.26	0.30	0.33	0.35	0.42
난청														
발생률 (10만명당)	718	720	732	765	811	830	831	832	816	890	925	984	1,016	1,058
전체 대상자 수	1,005,248	1,008,436	1,005,137	1,001,983	999,476	995,900	992,691	988,475	984,485	981,825	977,258	971,463	964,014	955,725
발생자 수	7,222	7,256	7,354	7,667	8,108	8,262	8,251	8,226	8,035	8,742	9,039	9,563	9,792	10,110
발생률	0.72	0.72	0.73	0.77	0.81	0.83	0.83	0.83	0.82	0.89	0.92	0.98	1.02	1.06
파킨슨병														
발생률 (10만명당)	53	58	62	68	61	69	86	77	77	108	87	90	90	85
전체 대상자 수	1,020,231	1,030,203	1,033,239	1,036,197	1,039,875	1,043,351	1,047,355	1,049,913	1,052,973	1,056,531	1,058,105	1,059,204	1,059,267	1,058,699
발생자 수	536	598	640	700	633	722	899	812	813	1,146	918	951	949	902
발생률	0.05	0.06	0.06	0.07	0.06	0.07	0.09	0.08	0.08	0.11	0.09	0.09	0.09	0.09

2. 연령별 발생률

각 질병의 연령별 발생률은 peak 연령에는 조금씩 차이가 있었으나, 거의 모두 연령이 증가함에 따라 지수함수적으로 증가하다가 매우 높은 연령에서는 발생률이 감소하는 노화관련질병에 부합하는 특징을 보였다.





[그림 1-8] 연령별 발생률

<표 1-5> 연령별 발생률

고혈압	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	33	281	745	2,038	4,174	6,503	9,201	12,416	12,670	12,300	3,344
전체 대상자 수	86,786	102,557	138,776	137,285	148,289	123,859	63,756	21,005	5,951	813	829,077
발생자 수	29	288	1,034	2,798	6,190	8,054	5,866	2,608	754	100	27,721
발생률	0.03	0.28	0.75	2.04	4.17	6.50	9.20	12.42	12.67	12.30	3.34

당뇨	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	68	337	553	1,440	2,720	4,465	6,256	7,497	6,335	4,255	2,549
전체 대상자 수	86,824	102,140	138,534	138,725	154,899	142,051	85,267	38,123	16,702	2,961	906,226
발생자 수	59	344	766	1,998	4,214	6,342	5,334	2,858	1,058	126	23,099
발생률	0.07	0.34	0.55	1.44	2.72	4.46	6.26	7.50	6.33	4.26	2.55

고지혈증	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	484	1,331	2,488	5,124	8,060	12,663	14,810	15,018	11,451	5,296	6,333
전체 대상자 수	86,722	100,550	133,628	126,999	132,861	108,035	55,491	26,036	13,763	2,870	786,955
발생자 수	420	1,338	3,324	6,508	10,708	13,680	8,218	3,910	1,576	152	49,834
발생률	0.48	1.33	2.49	5.12	8.06	12.66	14.81	15.02	11.45	5.30	6.33

뇌혈관계질환	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	39	72	152	358	772	2,012	3,698	6,567	8,773	8,577	1,514
전체 대상자 수	86,789	102,686	140,376	143,073	166,393	164,036	109,348	52,871	20,359	2,938	988,869
발생자 수	34	74	214	512	1,284	3,300	4,044	3,472	1,786	252	14,972
발생률	0.04	0.07	0.15	0.36	0.77	2.01	3.70	6.57	8.77	8.58	1.51

허혈성심질환	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	7	86	221	422	758	1,522	2,757	4,088	4,023	3,886	1,088
전체 대상자 수	86,849	102,752	140,005	141,605	163,133	157,831	102,704	50,832	21,922	3,500	971,133
발생자 수	6	88	310	598	1,236	2,402	2,832	2,078	882	136	10,568
발생률	0.01	0.09	0.22	0.42	0.76	1.52	2.76	4.09	4.02	3.89	1.09

골다공증	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	25	35	94	240	694	3,019	4,588	5,939	6,236	5,025	1,504
전체 대상자 수	86,886	102,914	140,899	143,432	166,771	158,853	95,768	40,950	14,753	2,189	953,415
발생자 수	22	36	132	344	1,158	4,796	4,394	2,432	920	110	14,344
발생률	0.03	0.03	0.09	0.24	0.69	3.02	4.59	5.94	6.24	5.03	1.50

골관절염	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	687	3,926	6,641	8,081	10,313	13,932	16,146	16,634	14,796	13,550	7,779
전체 대상자 수	85,991	95,613	113,152	100,460	101,660	72,809	31,500	9,775	2,690	369	614,019
발생자 수	591	3,754	7,514	8,118	10,484	10,144	5,086	1,626	398	50	47,765
발생률	0.69	3.93	6.64	8.08	10.31	13.93	16.15	16.63	14.80	13.55	7.78

만성폐쇄성폐질환	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	33	33	68	89	164	368	890	1,728	1,945	1,711	387
전체 대상자 수	86,729	102,496	140,333	143,366	168,024	170,598	120,195	64,017	27,664	4,092	1,027,514
발생자 수	29	34	96	128	276	628	1,070	1,106	538	70	3,975
발생률	0.03	0.03	0.07	0.09	0.16	0.37	0.89	1.73	1.94	1.71	0.39

울혈성심부전	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	6	4	14	24	49	106	255	684	1,177	2,104	154
전체 대상자 수	86,867	102,927	141,141	144,629	170,243	174,407	126,450	71,297	32,467	4,849	1,055,277
발생자 수	5	4	20	34	84	184	322	488	382	102	1,625
발생률	0.01	0.00	0.01	0.02	0.05	0.11	0.25	0.68	1.18	2.10	0.15

만성신질환	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	16	31	41	83	136	265	606	1,427	2,205	1,557	333
전체 대상자 수	86,888	102,943	141,048	144,390	169,709	173,616	125,802	71,189	33,285	5,138	1,054,008
발생자 수	14	32	58	120	230	460	762	1,016	734	80	3,506
발생률	0.02	0.03	0.04	0.08	0.14	0.26	0.61	1.43	2.21	1.56	0.33

백내장	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	16	27	53	140	824	3,710	9,516	16,357	15,167	4,731	2,520
전체 대상자 수	86,871	102,710	140,597	143,975	168,059	164,570	99,938	31,901	8,558	1,395	948,574
발생자 수	14	28	74	202	1,384	6,106	9,510	5,218	1,298	66	23,900
발생률	0.02	0.03	0.05	0.14	0.82	3.71	9.52	16.36	15.17	4.73	2.52

황반변성	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	5	12	24	98	275	865	2,301	3,862	3,411	1,830	831
전체 대상자 수	86,890	102,866	140,678	144,190	169,418	172,318	122,836	66,449	29,902	4,591	1,040,138
발생자 수	4	12	34	142	466	1,490	2,826	2,566	1,020	84	8,644
발생률	0.00	0.01	0.02	0.10	0.28	0.86	2.30	3.86	3.41	1.83	0.83

난청	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	638	1,062	1,398	1,648	1,603	2,254	3,497	5,416	5,665	4,004	2,099
전체 대상자 수	85,293	97,922	131,429	133,982	156,811	157,466	108,724	56,090	24,362	3,646	955,725
발생자 수	544	1,040	1,838	2,208	2,514	3,550	3,802	3,038	1,380	146	20,060
발생률	0.64	1.06	1.40	1.65	1.60	2.25	3.50	5.42	5.66	4.00	2.10

파킨슨병	0	10	20	30	40	50	60	70	80	90	전체
발생률 (10만명당)	12	43	47	54	56	121	259	819	1,026	577	170
전체 대상자 수	86,898	102,913	140,967	144,459	170,133	174,643	127,450	72,516	33,525	5,195	1,058,699
발생자 수	10	44	66	78	96	212	330	594	344	30	1,804
발생률	0.01	0.04	0.05	0.05	0.06	0.12	0.26	0.82	1.03	0.58	0.17

제5절 결론 및 제언

1. 결론

노화관련질병에 속하는 질병들은 일관되게 정의되어 있지 않다. 그러나 여러 문헌에 의하면 공통적으로 포함하고 있는 질환들이 나타나는데 이들은 대개 만성질환에 속하며, 의료비 부담 증가 등 질병부담이 높은 질환들에 해당하였다. 또한, 노화관련질병은 다양한 병리학적 기전과 다양한 기관에서 발생한다고 하더라도 공통적인 특징을 보여주었는데, 질환에 상관없이 거의 모두에서 연령이 증가함에 따라 지수함수적으로 발생률이 증가하다가 매우 높은 연령이 되면 오히려 발생률이 감소하는 형태를 보인 것이다. 이러한 특징은 노화세포의 축적과 질환에 대한 감수성 차이로 설명되고 있다.

본 연구진은, 고령사회에 진입한 한국에서 이러한 형태를 가진 노화관련질병을 재정의 하여 연령에 따라 질병부담이 증가하고 발생률이 증가되는 질환들 중 한국인에서 유병률이 높은 질환 14개를 선정하여 (암 제외) 연도별, 연령별 발생률을 확인 하였다.

그 결과, 노화관련질병 중 2004년부터 2019년까지 연도별로 발생률이 감소한 질병은 만성폐쇄성폐질환, 울혈성심부전, 허혈성심질환 이었으며, 지속적으로 발생률이 증가한 질병은 고지혈증, 만성신질환, 백내장, 난청, 파킨슨병 이었다. 그리고 특징적으로 고혈압, 당뇨, 뇌혈관계질환, 골다공증, 골관절염, 연령관련 황반변성의 경우, 점차 감소되던 발병률이 2015년 이후 증가하는 경향을 보였다. 하지만, 각 질병들을 10세 단위의 연령별로 발생률을 확인하였더니 연도별 발생률에서 보였던 차이와 상관없이 거의 모든 질환들은 정도의 차이는 있으나 연령이 증가함에 따라 지수함수적으로 증가하다가 매우 높은 연령에서는 발생률이 감소하는 형태를 보였다. 울혈성 심부전의 경우만 연령증가에 따라 지속적으로 발생률이 증가하는 형태를 보였다. 각 질병들에서 가장 높은 발병률을 보이

는 연령에는 약간의 차이가 있었으나 대부분 70~80세까지 발생률이 증가하다가 감소하는 것으로 보여진다.

2. 제언

이 연구는, 한국인을 대표하는 코호트를 기반으로 하여 노화관련질병의 발생률을 보여 준다는 점에서 의의가 있겠다. 각 질병은 연도별 발생률에는 차이가 있었으나, 거의 모두 연령이 증가함에 따라 지수함수적으로 증가하고 매우 높은 연령에서는 발생률이 감소하는 공통적인 특징을 보여주었다.

이러한 노화관련질병들에 대한 이해를 높이고 관련 요인들을 파악하여 질병의 예측 모델을 개발하여 예방함으로써 질환이 발생하지 않도록 교육하고 적절한 치료적 개입을 해주는 것은 매우 중요한 건강관리 시스템이 될 것으로 생각된다.

그러므로 현재 노화관련질병의 발생률을 파악한 첫 단계에서 더 나아가 노화 관련 질병 발생 메커니즘에 관한 분자생물학적 연구 및 사회 역학적 연구 등의 다각적인 연구가 필요할 것으로 생각된다.

제2장

검진 코호트에서의 ARD 분석

제1절 서론	33
제2절 이론적 고찰	34
제3절 연구 내용 및 방법	35
제4절 분석 결과	38
제5절 결론 및 제언	48

제2장

검진 코호트에서의 ARD 분석

제1절 서론

1. 연구배경 및 필요성

2021년 건강in 고도화 사업은 개인 맞춤형 서비스 제공 고도화를 위하여 노화관련질환의 분석과 관련 질병의 지식기반 질환발생 예측모델의 개발과 함께, 질환 발생 감소를 위한 개인 맞춤형 건강관리 방법을 제공하고자 하고 있다.

연령이 증가함에 따라 증가하는 노화관련질환(Age related disease, ARD)을 미리 파악하여 적절히 관리하고 예방하는 일은 보건 의료 서비스 측면에서 매우 중요하지만, 일관된 합의에 의해 정의된 ARD의 disease category가 정해져 있는 것은 없다. 연령 증가에 따라 자연스럽게 나타나는 정상적 노화 현상과 병적인 노화 현상의 구분이 명확하지 않고, 노화에 따른 기능의 저하와 질병이 복합되어 나타나기 때문에 개개인이 조금씩 상이한 질환의 범주를 가진다.

이에 맞추어 제1장에서는 노화 관련 질병의 발병 정도와 관련 요인 분석을 해 보았고, 이번 연구에서는 개개인의 건강과 노화 그리고 질병 발현을 알아보고자 건강노화나이의 산출과 함께 이에 따른 노화관련질환의 발병 및 관련 여러 요인을 파악하고자 한다.

2. 연구의 목적

본 연구는 『2014년 개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발』 연구에 이은 후속 8차년도 연구로써, 건강노화나이에 따른 ARD 발병률과 관련 요인 분석하여, 이용자의 건강, 질병정보 제공을 통하여 질병 발생의 예방과 질병 관리에 도움이 되고자 한다.

제2절 이론적 고찰

1. 건강노화나이에 따른 ARD 발병률과 관련 요인 분석

평균수명의 증가와 출산율의 감소로 고령 인구의 증가는 세계적인 추세이다. UN, OECD 등 국제기관에서는 65세 이상 인구가 전체 인구에서 차지하는 비율이 7% 이상이면 고령화 사회(Aging Society), 14% 이상이면 고령 사회(Aged Society), 20% 이상이면 초고령 사회(Super-Aged Society)로 구분하고 있다. 미국은 1942년, 일본은 1970년에 각기 이미 고령화 사회가 되었으며, 한국은 2000년 고령화 사회에 진입했다. 2021년 9월말 기준으로 한국의 65세 이상 주민등록 인구는 853만 7천 명으로 전체 인구(5,182만 2천명)에서 차지하는 비율이 16.4%를 넘어섰다.¹ 고령 인구 비중은 계속 증가하여 2025년에는 20.3%로 우리나라가 초고령 사회에 진입할 것으로 전망된다.¹

2020년 기준 65세 이상 고령자 10만 명당 사망원인은 암(733.3명), 심장질환(331.9명), 폐렴(257.5명), 뇌혈관질환(225.3명), 알츠하이머병(91.4명) 순으로 나타났다.¹ 노화 관련 질환(Age Related Disease, ARD)은 노화가 증가함에 따라 빈도가 증가하면서 가장 흔히 나타나는 질환이다. 노화 관련 질환의 예로는 심혈관 질환, 관절염, 백내장, 골다공증, 당뇨병, 고혈압 및 알츠하이머병이 있다. 이러한 모든 질병의 발병률은 나이가 들수록 기하급수적으로 증가한다.²

국내 노인 1인당 연간 의료비 지출은 빠르게 증가하고 있으며, 이는 다른 연령대의 증가율보다 더 높은 것으로 알려져 있다. 또한, 우리나라의 의료비 지출 증가율은 다른 나라와 비교했을 때 높은 편에 속하며 앞으로도 인구 고령화로 의료비 지출 증가 속도는 더 빨라질 것으로 보인다.³ 2017년 중국에서 진행된 연령증가가 의료비에 미치는 영향에 대한 연구에서도 고령인구에서 사용되는 의료비 지출과 자원을 절약하기 위해 노인 인구에 대처할 것을 제안하였다.⁴ 현재 인구 고령화가 의료비 지출에 미친 영향을 심층적으로 분석한 연구는 많지 않은 편이다. 그동안 의료비 결정요인에 대한 선행연구들을 살펴보면 대부분이 국가별 65세 이상 노인인구의 비율이 국민의료비 지출에 어떤 영향을 주고 있는가 하는 수준에 머물러 있었다.⁵

이에 맞추어 본 연구는 고령화 사회에서 대표적으로 흔하고 의료 이용과 의료비 증가의 원인이 되는 노화 관련 질환을 선정하여 건강노화나이에 따른 노화 관련 질환 발병률과 관련 요인 분석을 수행하였다.

제3절 연구 내용 및 방법

1. 연구 내용

본 연구는 건강검진 코호트DB를 사용하여 검진 수검자를 대상으로 건강노화나이에 따른 ARD 발병률과 관련 요인을 분석하였다. 건강노화나이는 이전 3차년도 연구에서 만들었던 모델을 사용하였고, ARD는 고지혈증, 골관절염, 만성신질환, 만성폐쇄성폐질환, 백내장, 울혈성심부전을 Cox 비례 위험 모형을 통해 분석하였다. 의료 관련 요인으로는 의료비, 연간 의료비 증가, 외래 일수, 입원 일수를 회귀분석을 이용하여 진행하였다.

2. 연구 방법

1) 자료 개요

① 표본 코호트 DB

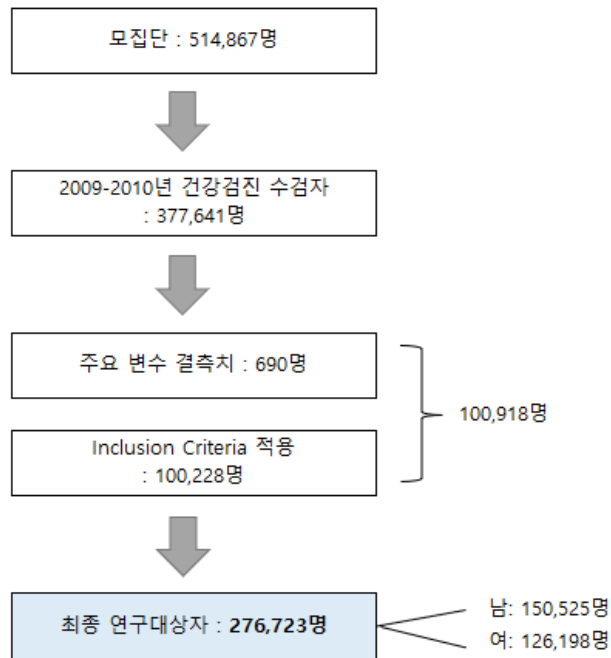
2006년 1년간 건강보험가입자 및 의료급여수급권자 자격을 유지한 전 국민의 약 2% 성·연령·가입자구분·보험료 분위·지역별 층화 추출한 100만 명에 대한 2002~2018(17년간)의 사회·경제적 현황(자격 및 보험료, 장애 및 사망), 의료이용 현황(진료 및 건강검진), 요양기관 정보를 코호트 형식으로 구축한 개인 식별이 불가능한 연구용 DB 이다.

② 건강검진 코호트 DB

건강검진 수검자 중심의 의료이용 및 건강결과 분석을 위해 2002~2003년 일반 건강검진 수검자 중 2002.12월말 기준 40세 ~ 79세의 건강보험 자격유지자 515만 명의 10%인 약 51만 명에 대한 2002~2018(17년간)의 자격 및 소득정보(사회경제적 변수), 병·의원 이용 내역 및 건강검진결과, 요양기관 정보를 코호트 형식으로 구축한 개인 식별이 불가능한 연구용 DB 이다.

2) 연구 자료 및 대상자 선정

연구 자료는 2002.12월 기준 건강보험 자격유지자 515만명의 10%를 추출한 514,866명 대상으로 2002-2019년 관찰한 건강검진 코호트DB를 사용하였다. 이 중 2009-2010년 건강검진 수검자 대상으로 건강노화나이 모델을 구성하는 주요 변수 결측치가 있는 사람을 제외하고, 이전 3차년도 연구에서 사용하였던 Inclusion Criteria를 적용하여 총 276,723명을 대상으로 2009-2019년 최대 10년을 관찰한 자료를 사용하였다.



[그림 2-1] 건강노화나이에 따른 ARD 발병률과 관련 요인 분석 대상자 선정

ARD 별 데이터 셋은 276,723명 대상자 중 2002-2008년 기준으로 ARD 진단받은 사람을 제외하였다. ARD 별 최종 연구대상자 수는 <표 2-1>과 같다.

<표 2-1> ARD 별 최종 연구대상자

ARD	All	Male	Female
고지혈증	217,922	120,759	97,163
골관절염	172,860	111,706	61,154
만성신질환	276,169	150,189	125,980
만성폐쇄성폐질환	200,950	113,785	87,165
백내장	238,169	133,023	105,146
울혈성심부전	273,144	149,111	124,033

3) DB별 사용 변수

자격 DB에서 성, 연령, 사망일자 변수를 사용하였고, 진료 DB에서는 영양개시일자, 주상병, 부상병, 요양일수, 입내원일수, 심결요양급여비용총액을 사용하였다. 건강노화나이 모델을 위해 건강검진 DB에서 키, 허리둘레, 수축기혈압, 공복혈당, 총콜레스테롤, 중성지방, 고밀도 콜레스테롤, 혈색소, 크레아티닌, AST, 감마지티피,

신사구체여과율 등을 사용하였다.

4) ARD(Age Related Disease)

ARD는 고지혈증, 골관절염, 만성신질환, 만성폐쇄성폐질환, 백내장, 울혈성심부전을 선정하였다. 각 변수의 koicd 코드는 다음과 같이 설정하였다.

고지혈증 : E7808, E781-E785

골관절염 : M150, M158, M159, M160, M161, M169, M170, M171, M179, M180, M181, M189, M190, M198, M199

만성신질환 : N18

만성폐쇄성폐질환 : J42-J45

백내장 : H25, H268, H269

울혈성심부전 : I500

5) 건강노화나이 모델과 건강노화나이 그룹

건강노화나이는 이전 3차년도 연구에서 만들었던 모델을 사용하였고 공식은 <표 2-2>와 같다. 건강노화나이 그룹은 건강노화나이와 주민등록나이의 차이의 표준편차를 기준으로 설정하였고 기준은 <표 2-3>과 같다.

<표 2-2> 건강노화나이 공식

구분	건강노화나이 공식
남	-127.6 -0.19*키 +0.88*허리둘레 +0.46*수축기혈압 +0.40*공복혈당 +0.57*혈색소 -0.17*eGFR +0.62*혈청지오티 +0.65*연령
여	-0.01 -0.45*키 +0.49*허리둘레 +0.24*수축기혈압 +0.22*공복혈당 +0.07*총콜레스테롤 +0.07*중성지방 -0.18*고밀도콜레스테롤 -0.07*eGFR +0.44*혈청지오티 +0.20*감마지티피 +0.36*연령

<표 2-3> 건강노화나이 그룹 기준

건강노화나이 그룹	기준
Group1	SD(건강노화나이 - 주민등록나이) ≤ -1
Group2	-1 < SD(건강노화나이 - 주민등록나이) ≤ 0
Group3	0 < SD(건강노화나이 - 주민등록나이) < 1
Group4	1 ≤ SD(건강노화나이 - 주민등록나이)

6) 통계적 분석 방법

건강노화나이 모델은 이전 3차년도 연구에서 만들었던 모델을 사용하였고, 건강노화나이에 따른 ARD 발병률 분석은 Cox 비례 위험 모형, Trend Test를 사용하였다. 그리고 건강노화나이에 따른 의료 관련 요인 분석은 선형 회귀분석을 이용하였다. 이상의 분석은 R Studio 3.3.3 ver. 프로그램으로 수행하였으며, p값이 0.05 미만이면 통계적으로 의미 있는 결과로 판단하였다.

제4절 분석 결과

1. 연구 대상자 특성

연구대상자의 일반적 특성은 아래의 <표 2-4>와 같다.

<표 2-4> 연구대상자 특성

Parameter	All(n=276,723)	Male(n=150,525)	Female(n=126,198)
AGE	58.91 ± 8.8	58.36 ± 8.7	59.57 ± 8.9
BMI	23.76 ± 2.6	23.85 ± 2.5	23.67 ± 2.6
WAIST	81.44 ± 7.7	83.98 ± 6.9	78.42 ± 7.4
BP_HIGH	123.36 ± 13.2	124.28 ± 12.6	122.26 ± 13.7
BP_LWST	76.41 ± 8.7	77.4 ± 8.4	75.24 ± 8.8
BLDS	96.22 ± 13.3	97.54 ± 13.8	94.65 ± 12.6
TOT_CHOLE	195.08 ± 30.9	191.51 ± 30.9	199.34 ± 30.3
TRIGLYCERIDE	130.92 ± 64.5	137.01 ± 67.7	123.66 ± 59.8
HDL_CHOLE	52.89 ± 12.0	51.03 ± 11.7	55.11 ± 12.1
LDL_CHOLE	115.8 ± 29.1	112.95 ± 29.2	119.21 ± 28.7
HMG	13.83 ± 1.4	14.63 ± 1.2	12.87 ± 1.0
CREATININE	0.91 ± 0.2	1.01 ± 0.2	0.79 ± 0.2
eGFR	83.79 ± 18.8	84.44 ± 18.7	83.03 ± 18.1
SGOT_AST	22.75 ± 7.2	25.3 ± 7.4	23.65 ± 6.8
SGPT_ALT	22.75 ± 10.7	24.71 ± 11.3	20.41 ± 9.4
GAMMA_GTP	30.38 ± 22.4	37.84 ± 25.0	21.48 ± 14.5

2. 건강노화나이에 따른 ARD 발병률 분석

1) 성별 ARD 발생률

전체 ARD 발생률은 <표 2-5>와 같다. 전체 ARD 발생률은 고지혈증, 골관절염,

백내장, 만성폐쇄성폐질환, 만성신질환, 울혈성심부전 순으로 높았다.

성별 ARD 발생률은 <표 2-6>, <표2-7>과 같다. 만성신질환을 제외한 5개의 질환에서 여자의 발생률이 남자보다 높았다.

<표 2-5> ARD 발생률(전체)

ARD	Event				Total	
	Yes		No		N수	비율(%)
	N수	비율(%)	N수	비율(%)		
고지혈증	95,928	44.0	121,994	56.0	217,922	100
골관절염	71,211	41.2	101,649	58.8	172,860	100
만성신질환	5,781	2.1	270,388	97.9	276,169	100
만성폐쇄성폐질환	61,300	30.5	139,650	69.5	200,950	100
백내장	74,221	31.2	163,948	68.8	238,169	100
울혈성심부전	3,186	1.2	269,958	98.8	273,144	100

<표 2-6> ARD 발생률(남)

ARD	Event				Total	
	Yes		No		N수	비율(%)
	N수	비율(%)	N수	비율(%)		
고지혈증	47,799	39.6	72,960	60.4	120,759	100
골관절염	39,120	35.0	72,580	65.0	111,700	100
만성신질환	3,966	2.6	146,223	97.4	150,189	100
만성폐쇄성폐질환	32,614	28.7	81,171	71.3	113,785	100
백내장	36,465	27.4	96,558	72.6	133,023	100
울혈성심부전	1,598	1.1	147,513	98.9	149,111	100

<표 2-7> ARD 발생률(여)

ARD	Event				Total	
	Yes		No		N수	비율(%)
	N수	비율(%)	N수	비율(%)		
고지혈증	48,129	49.5	49,034	50.5	97,163	100
골관절염	32,091	52.5	29,063	47.5	61,154	100
만성신질환	1,815	1.4	124,165	98.6	125,980	100
만성폐쇄성폐질환	28,686	32.9	58,479	67.1	87,165	100
백내장	37,756	35.9	67,390	64.1	105,146	100
울혈성심부전	1,588	1.3	122,445	98.7	124,033	100

2) 연령대별 ARD 발생률

전체 연령대별 ARD 발생률은 <표 2-8>과 같다. 골관절염, 만성신질환, 만성폐쇄성폐질환, 백내장, 울혈성심부전은 연령대가 증가할수록 ARD 발생률이 높아지는 결과를 보였다.

연령대별 ARD 발생률은 <표 2-9>, <표 2-10>과 같다. 골관절염, 만성신질환, 만성폐쇄성폐질환, 백내장, 울혈성심부전에서 남, 여 모두 연령대가 증가할수록 ARD 발생률이 높아지는 추세를 보였다.

<표 2-8> 연령대별 ARD 발생률(전체)

ARD	연령대	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
고지혈증	40≤AGE<50	12,957	39.3	19,977	60.7	32,934	100
	50≤AGE<65	60,752	46.1	71,035	53.9	131,787	100
	65≤AGE	22,219	41.8	30,982	58.2	53,201	100
골관절염	40≤AGE<50	10,136	31.9	21,647	68.1	31,783	100
	50≤AGE<65	46,649	41.6	65,591	58.4	112,240	100
	65≤AGE	14,426	50	14,411	50	28,837	100
만성신질환	40≤AGE<50	254	0.7	37,604	99.3	37,858	100
	50≤AGE<65	2,436	1.5	163,771	98.5	166,207	100
	65≤AGE	3,091	4.3	69,013	95.7	72,104	100
만성폐쇄성폐질환	40≤AGE<50	7,055	22.9	23,736	77.1	30,791	100
	50≤AGE<65	37,552	29.8	88,417	70.2	125,969	100
	65≤AGE	16,693	37.8	27,497	62.2	44,190	100
백내장	40≤AGE<50	3,962	10.6	33,405	89.4	37,367	100
	50≤AGE<65	45,949	29.5	109,651	70.5	155,600	100
	65≤AGE	24,310	53.8	20,832	46.2	45,202	100
울혈성심부전	40≤AGE<50	102	0.3	37,674	99.7	37,776	100
	50≤AGE<65	1,184	0.7	164,140	99.3	165,324	100
	65≤AGE	1,900	2.7	68,144	97.3	70,044	100

※ p-value<0.05(Significant)

<표 2-9> 연령대별 ARD 발생률(남)

ARD	연령대	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
고지혈증	40≤AGE<50	6,978	34.7	13,110	65.3	20,088	100
	50≤AGE<65	30,136	41.4	42,662	58.6	72,798	100
	65≤AGE	10,685	38.3	17,188	61.7	27,873	100
골관절염	40≤AGE<50	5,286	25.4	15,544	74.6	20,830	100
	50≤AGE<65	24,825	34.9	46,295	65.1	71,120	100
	65≤AGE	9,009	45.6	10,747	54.4	19,756	100
만성 신질환	40≤AGE<50	184	0.8	23,399	99.2	23,583	100
	50≤AGE<65	1,714	1.9	88,857	98.1	90,571	100
	65≤AGE	2,068	5.7	33,967	94.3	36,035	100
만성폐쇄성 폐질환	40≤AGE<50	4,061	20.6	15,694	79.4	19,755	100
	50≤AGE<65	19,848	27.8	51,641	72.2	71,489	100
	65≤AGE	8,705	38.6	13,836	61.4	22,541	100
백내장	40≤AGE<50	2,197	9.5	21,007	90.5	23,204	100
	50≤AGE<65	21,918	25.7	63,278	74.3	85,196	100
	65≤AGE	12,350	50.2	12,273	49.8	24,623	100
울혈성 심부전	40≤AGE<50	64	0.3	23,489	99.7	23,553	100
	50≤AGE<65	664	0.7	89,543	99.3	90,207	100
	65≤AGE	870	2.5	34,481	97.5	35,351	100

※ p-value<0.05(Significant)

<표 2-10> 연령대별 ARD 발생률(여)

ARD	연령대	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
고지혈증	40≤AGE<50	5,979	46.5	6,867	53.5	12,846	100
	50≤AGE<65	30,616	51.9	28,373	48.1	58,989	100
	65≤AGE	11,534	45.5	13,794	54.5	25,328	100
골관절염	40≤AGE<50	4,850	44.3	6,103	55.7	10,953	100
	50≤AGE<65	21,824	53.1	19,296	46.9	41,120	100
	65≤AGE	5,417	59.7	3,664	40.3	9,081	100
만성 신질환	40≤AGE<50	70	0.5	14,205	99.5	14,275	100
	50≤AGE<65	722	1.0	74,914	99.0	75,636	100
	65≤AGE	1,023	2.8	35,046	97.2	36,069	100

ARD	연령대	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
만성폐쇄성 폐질환	40≤AGE<50	2,994	27.1	8,042	72.9	11,036	100
	50≤AGE<65	17,704	32.5	36,776	67.5	54,480	100
	65≤AGE	7,988	36.9	13,661	63.1	21,649	100
백내장	40≤AGE<50	1,765	12.5	12,398	87.5	14,163	100
	50≤AGE<65	24,031	34.1	46,373	65.9	70,404	100
	65≤AGE	11,960	58.1	8,619	41.9	20,579	100
울혈성 심부전	40≤AGE<50	38	0.3	14,185	99.7	14,223	100
	50≤AGE<65	520	0.7	74,597	99.3	75,117	100
	65≤AGE	1,030	3.0	33,663	97.0	34,693	100

※ p-value<0.05(Significant)

3) 건강노화나이 그룹별 ARD 발생률

전체 건강노화나이 그룹별 ARD 발생률은 <표 2-11>과 같고 남, 여 건강노화나이 그룹별 ARD 발생률은 <표 2-12>, <표 2-13>과 같다. 고지혈증, 골관절염, 만성신질환, 백내장에서 건강노화나이가 나쁜 그룹일수록 ARD 발생률이 높아지는 결과를 보였다.

<표 2-11> 건강노화나이 그룹별 ARD 발생률(전체)

ARD	건강노화나이 그룹	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
고지혈증	Group1	10,294	32.5	21,373	67.5	31,667	100
	Group2	32,396	40.7	47,218	59.3	79,614	100
	Group3	34,818	47.9	37,849	52.1	72,667	100
	Group4	18,420	54.2	15,554	45.8	33,974	100
골관절염	Group1	9,102	38.1	14,781	61.9	23,883	100
	Group2	24,962	40.5	36,659	59.5	61,621	100
	Group3	24,543	41.8	34,133	58.2	58,676	100
	Group4	12,604	43.9	16,076	56.1	28,680	100
만성 신질환	Group1	506	1.4	36,208	98.6	36,714	100
	Group2	1,736	1.8	95,490	98.2	97,226	100
	Group3	2,183	2.3	92,345	97.7	94,528	100
	Group4	1,356	2.8	46,345	97.2	47,701	100

ARD	건강노화나이 그룹	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
만성폐쇄성 폐질환	Group1	8,114	30.7	18,348	69.3	26,462	100
	Group2	21,603	30.4	49,460	69.6	71,063	100
	Group3	20,946	30.3	48,103	69.7	69,049	100
	Group4	10,637	30.9	23,739	69.1	34,376	100
백내장	Group1	9,328	30.2	21,605	69.8	30,933	100
	Group2	25,740	30.7	58,047	69.3	83,787	100
	Group3	25,655	31.3	56,379	68.7	82,034	100
	Group4	13,498	32.6	27,917	67.4	41,415	100
울혈성 심부전*	Group1	414	1.1	35,895	98.9	36,309	100
	Group2	1,124	1.2	95,100	98.8	96,224	100
	Group3	1,059	1.1	92,463	98.9	93,522	100
	Group4	589	1.3	46,500	98.7	47,089	100

※ p-value<0.05(Significant), *p-value>0.05(NS)

<표 2-12> 건강노화나이 그룹별 ARD 발생률(남)

ARD	건강노화나이 그룹	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
고지혈증	Group1	6,521	30.8	14,685	69.2	21,206	100
	Group2	16,112	37.3	27,065	62.7	43,177	100
	Group3	17,001	43.1	22,462	56.9	39,463	100
	Group4	8,165	48.3	8,748	51.7	16,913	100
골관절염	Group1	6,031	34.1	11,662	65.9	17,693	100
	Group2	13,449	34.4	25,651	65.6	39,100	100
	Group3	13,328	35.4	24,325	64.6	37,653	100
	Group4	6,312	36.6	10,948	63.4	17,260	100
만성 신질환	Group1	363	1.5	23,959	98.5	24,322	100
	Group2	1,140	2.2	51,230	97.8	52,370	100
	Group3	1,555	3.1	48,965	96.9	50,520	100
	Group4	908	4.0	22,069	96.0	22,977	100
만성폐쇄성 폐질환*	Group1	5,411	30.4	12,380	69.6	17,791	100
	Group2	11,437	28.9	28,177	71.1	39,614	100
	Group3	10,844	28.0	27,845	72.0	38,689	100
	Group4	4,922	27.8	12,769	72.2	17,691	100

ARD	건강노화나이 그룹	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
백내장	Group1	5,782	27.3	15,372	72.7	21,154	100
	Group2	12,679	27.3	33,740	72.7	46,419	100
	Group3	12,251	27.2	32,752	72.8	45,003	100
	Group4	5,753	28.1	14,694	71.9	20,447	100
울혈성 심부전*	Group1	256	1.1	23,857	98.9	24,113	100
	Group2	549	1.1	51,456	98.9	52,005	100
	Group3	531	1.1	49,657	98.9	50,188	100
	Group4	262	1.1	22,543	98.9	22,805	100

※ p-value<0.05 (Significant), *p-value>0.05(NS)

<표 2-13> 건강노화나이 그룹별 ARD 발생률(여)

ARD	건강노화나이 그룹	Event				Total	
		Yes		No		N수	비율 (%)
		N수	비율(%)	N수	비율(%)		
고지혈증	Group1	3,773	36.1	6,688	63.9	10,461	100
	Group2	16,284	44.7	20,153	55.3	36,437	100
	Group3	17,817	53.7	15,387	46.3	33,204	100
	Group4	10,255	60.1	6,806	39.9	17,061	100
골관절염	Group1	3,071	49.6	3,119	50.4	6,190	100
	Group2	11,513	51.1	11,008	48.9	22,521	100
	Group3	11,215	53.3	9,808	46.7	21,023	100
	Group4	6,292	55.1	5,128	44.9	11,420	100
만성 신질환	Group1	143	1.2	12,249	98.8	12,392	100
	Group2	596	1.3	44,260	98.7	44,856	100
	Group3	628	1.4	43,380	98.6	44,008	100
	Group4	448	1.8	24,276	98.2	24,724	100
만성폐쇄성 폐질환*	Group1	2,703	31.2	5,968	68.8	8,671	100
	Group2	10,166	32.3	21,283	67.7	31,449	100
	Group3	10,102	33.3	20,258	66.7	30,360	100
	Group4	5,715	34.3	10,970	65.7	16,685	100
백내장	Group1	3,546	36.3	6,233	63.7	9,779	100
	Group2	13,061	35.0	24,307	65.0	37,368	100
	Group3	13,404	36.2	23,627	63.8	37,031	100
	Group4	7,745	36.9	13,223	63.1	20,968	100
울혈성 심부전*	Group1	158	1.3	12,038	98.7	12,196	100
	Group2	575	1.3	43,644	98.7	44,219	100
	Group3	528	1.2	42,806	98.8	43,334	100
	Group4	327	1.3	23,957	98.7	24,284	100

※ p-value<0.05 (Significant), *p-value>0.05(NS)

4) 성별 건강노화나이에 따른 ARD 발병 위험도

성별 건강노화나이에 따른 ARD 발병 위험도는 <표 2-14>와 같다. 고지혈증, 골관절염, 만성폐쇄성폐질환, 백내장의 경우 건강노화나이 1SD 증가에 따른 발병 위험도가 남자보다 여자가 높았고, 만성신질환, 울혈성심부전의 경우 건강노화나이에 따른 발병 위험도가 여자보다 남자가 높았다.

<표 2-14> 성별 건강노화나이에 따른 ARD 발병 위험도

ARD	All		Male		Female	
	HR	95% CI	HR	95% CI	HR	95% CI
고지혈증	1.259	1.25-1.26	1.219	1.21-1.23	1.269	1.26-1.28
골관절염	1.062	1.05-1.07	1.029	1.02-1.04	1.052	1.04-1.06
만성신질환	1.253	1.22-1.28	1.379	1.34-1.42	1.151	1.10-1.20
만성폐쇄성폐질환	1.005	1.00-1.01	0.964	0.95-0.97	1.036	1.02-1.05
백내장	1.031	1.02-1.04	1.009	1.00-1.02	1.015	1.01-1.03
울혈성심부전	1.025	0.99-1.06	1.018	0.97-1.07	1.013	0.96-1.06

5) 연령대별 건강노화나이에 따른 ARD 발병 위험도

연령대별 건강노화나이에 따른 ARD 발병 위험도는 <표 2-15>와 같다. 고지혈증, 백내장, 울혈성심부전에서 연령대가 올라갈수록 건강노화나이에 따른 ARD 발병 위험도는 작아지는 결과를 보였다.

<표 2-15> 연령대별 건강노화나이에 따른 ARD 발병 위험도

ARD	연령대	All		Male		Female	
		HR	95% CI	HR	95% CI	HR	95% CI
고지혈증	40≤AGE<50	1.32	1.30-1.35	1.28	1.25-1.31	1.36	1.33-1.40
	50≤AGE<65	1.26	1.25-1.27	1.22	1.20-1.23	1.26	1.24-1.27
	65≤AGE	1.22	1.21-1.24	1.19	1.17-1.22	1.23	1.21-1.25
골관절염	40≤AGE<50	1.06	1.04-1.08	1.04	1.01-1.07	1.07	1.04-1.10
	50≤AGE<65	1.08	1.07-1.09	1.05	1.04-1.06	1.06	1.04-1.07
	65≤AGE	1.04	1.02-1.05	1.02	1.00-1.04	1.02	0.99-1.04
만성신질환	40≤AGE<50	1.34	1.19-1.52	1.57	1.36-1.82	0.97	0.76-1.24
	50≤AGE<65	1.29	1.24-1.34	1.45	1.38-1.52	1.17	1.09-1.25
	65≤AGE	1.31	1.26-1.35	1.42	1.36-1.48	1.25	1.17-1.32
만성폐쇄성폐질환	40≤AGE<50	1.01	0.99-1.04	0.98	0.94-1.01	1.05	1.01-1.09
	50≤AGE<65	1.02	1.01-1.03	0.98	0.97-0.99	1.04	1.03-1.06
	65≤AGE	1.00	0.98-1.01	0.97	0.95-0.99	1.04	1.01-1.06

ARD	연령대	All		Male		Female	
		HR	95% CI	HR	95% CI	HR	95% CI
백내장	40≤AGE<50	1.06	1.03-1.10	1.05	1.01-1.10	1.06	1.01-1.11
	50≤AGE<65	1.05	1.04-1.06	1.04	1.02-1.05	1.03	1.02-1.04
	65≤AGE	1.05	1.04-1.31	1.06	1.04-1.08	1.02	1.00-1.04
울혈성 심부전	40≤AGE<50	1.20	0.99-1.46	1.17	0.91-1.51	1.25	0.93-1.69
	50≤AGE<65	1.14	1.08-1.21	1.16	1.08-1.25	1.13	1.04-1.23
	65≤AGE	1.04	0.99-1.09	1.01	0.95-1.08	1.04	0.98-1.11

6) 의료 관련 요인 특성

성별 의료 관련 요인 특성은 <표 2-16>과 같다. 의료비와 연간 의료비 증가, 외래일수, 입원일수 모두 남자보다 여자가 높았다.

연령대별 의료 관련 요인 특성은 <표 2-17>, <표 2-18>과 같다. 전체, 남, 여 모두 연령대가 증가할수록 의료비, 연간 의료비 증가, 외래일수, 입원일수는 증가하였다.

<표 2-16> 성별 의료 관련 요인 특성

요인	All (n=265,283)		Male (n=141,364)		Female (n=123,919)	
	Mean	SD	Mean	SD	Mean	SD
의료비	787,014	1,831,489	732,839	1,935,130	848,816	1,703,491
연간 의료비 증가	166,574	636,209	165,286	664,129	167,968	604,554
외래일수	39.3	77.8	35.8	81.6	43.3	73.0
입원일수	24.2	26.6	20.8	25.2	28.0	27.7

<표 2-17> 연령대별 의료 관련 요인 특성(전체)

요인	연령대	All (n=265,283)	
		Mean	SD
의료비	40≤AGE<50	444,040	1,082,543
	50≤AGE<60	590,302	1,442,980
	60≤AGE<70	919,507	1,994,039
	70≤AGE	1,375,093	2,637,757
연간 의료비 증가	40≤AGE<50	78,673	36,947
	50≤AGE<60	121,316	547,373
	60≤AGE<70	205,004	687,365
	70≤AGE	329,843	922,114

요인	연령대	All (n=265,283)	
		Mean	SD
외래일수	40≤AGE<50	19.8	41.7
	50≤AGE<60	26.5	51.3
	60≤AGE<70	45.8	86.9
	70≤AGE	79.1	118.0
입원일수	40≤AGE<50	14.0	15.9
	50≤AGE<60	18.4	19.4
	60≤AGE<70	28.2	27.7
	70≤AGE	41.3	37.4

<표 2-18> 연령대별 의료 관련 요인 특성(남, 여)

요인	연령대	Male (n=141,364)		Female (n=123,919)	
		Mean	SD	Mean	SD
의료비	40≤AGE<50	402,939	1,020,294	507,069	1,168,831
	50≤AGE<60	541,020	150,404	648,382	1,364,656
	60≤AGE<70	865,734	2,133,483	975,371	1,836,367
	70≤AGE	1,402,734	2,961,328	1,348,348	2,281,137
연간 의료비 증가	40≤AGE<50	76,500	393,543	81,894	331,341
	50≤AGE<60	124,438	597,490	117,756	483,927
	60≤AGE<70	216,970	740,820	193,270	630,349
	70≤AGE	331,672	932,800	328,362	913,387
외래 일수	40≤AGE<50	18.7	44.1	21.5	37.7
	50≤AGE<60	24.1	54.1	29.3	47.6
	60≤AGE<70	43.2	98.3	48.6	73.2
	70≤AGE	75.0	122.1	83.1	113.7
입원 일수	40≤AGE<50	12.2	14.9	16.7	16.9
	50≤AGE<60	15.6	17.5	21.7	21.0
	60≤AGE<70	24.7	26.4	31.9	28.5
	70≤AGE	38.6	37.8	43.9	36.7

7) 건강노화나이와 연령에 따른 의료 관련 요인 분석

건강노화나이와 연령에 따라 의료 관련 요인을 예측하는 회귀 식은 <표 2-19>와 같다. 전체, 남, 여 모두 건강노화나이가 많을수록, 연령이 높을수록 의료비, 연간 의료비 증가, 외래일수, 입원일수가 유의하게 증가하는 결과를 보였다.

<표 2-19> 건강노화나이와 연령에 따른 의료 관련 요인 분석

요인	구분	회귀 식
의료비	All	의료비 = -1326669.1+36072.4*건강노화나이+35694.6*AGE
	Male	의료비 = -1463024.8+3916.8*건강노화나이+37437*AGE
	Female	의료비 = -1140626.4+64494.5*건강노화나이+33126.3*AGE
연간 의료비 증가	All	연간 의료비 증가 = -407975.2+9055*건강노화나이+9805*AGE
	Male	연간 의료비 증가 = -422340.2+12295.2*건강노화나이+10151.5*AGE
	Female	연간 의료비 증가 = -400620.1+6934.3*건강노화나이+9583.7*AGE
외래 일수	All	외래일수 = -97.02137 + 2.76352*건강노화나이+2.30194*AGE
	Male	외래일수 = -91.69193+1.38003*건강노화나이+2.17421*AGE
	Female	외래일수 = -100.97251+3.77384*건강노화나이+2.40582*AGE
입원 일수	All	입원일수 = -38.270706+1.603635*건강노화나이+1.0535*AGE
	Male	입원일수 = -38.276128+0.620542*건강노화나이+1.008023*AGE
	Female	입원일수 = -35.595572+2.041787*건강노화나이+1.058103*AGE

※ p-value<0.05(Significant)

제5절 결론 및 제언

1. 결론

총 276,723명의 10년 코호트 자료를 기준으로 건강노화나이에 따른 ARD 발병률과 관련 요인을 분석하였다. 고지혈증, 골관절염, 만성폐쇄성폐질환, 백내장, 울혈성심부전은 남자보다 여자에서 발생률이 높았고, 만성신질환은 남자에서 발생률이 높았다. 대체적으로 연령대가 증가할수록 발생률이 높아지는 결과를 보였고, 건강노화나이가 나쁜 그룹일수록 발생률이 높아지는 결과를 보였다.

건강노화나이 증가에 따른 ARD 발병률 분석에서는 건강노화나이가 증가할수록 ARD 발병 위험도가 모두 유의하게 높아지는 결과를 보였다. (HR: 고지혈증 1.259, 골관절염 1.062, 만성신질환 1.253, 만성폐쇄성폐질환 1.005, 백내장 1.031, 울혈성심부전 1.025) 그리고 대체적으로 연령대가 올라갈수록 건강노화나이 증가에 따른 ARD 발병 위험도는 낮아졌다.

의료비, 연간 의료비 증가, 외래일수, 입원일수는 모두 남자보다 여자가 높았고, 연령대가 증가할수록 모두 증가하는 결과를 보였다. 또한 전체, 남, 여 모두 건강노화나이가

많을수록, 연령이 높을수록 의료 관련 요인이 유의하게 증가하는 결과를 보였다.

2. 제언

이번 연구는 개인별 맞춤형 통합건강관리 프로그램 연구 및 개발 3차년도에 이어 건강관리 분야에서의 건강노화나이 모델의 타당성과 유용성에 대한 후속 연구임에 의의가 있다. 분석 대상자 수와 질병 발생률의 제한으로 대표적인 6개의 ARD에 대해 분석하여 의미 있는 결과를 보고하였지만, 향후 성별, 연령별 ARD 발병률을 더 정확하게 예측하려면 연구 대상자 수를 늘려 다른 ARD 발병률을 추가적으로 분석하는 연구가 뒤따라야 할 것이다. 아울러 의료비등을 비롯한 관련 요인 분석에서는, 향후 급여 의료비뿐만 아니라 비급여 의료비까지 포함시키고 질병 유무별 의료비를 세부적으로 예측할 수 있는 추가적인 연구가 필요할 것이다.

제3장

골다공증 예측 모델의 개발

제1절 서론	53
제2절 이론적 고찰	53
제3절 연구 내용 및 방법	57
제4절 분석 결과	61
제5절 결론 및 제언	88

제3장 골다공증 예측 모델의 개발

제1절 서론

1. 연구배경 및 필요성

2021년 건강in 고도화 사업은 개인 맞춤형 서비스 제공 고도화를 위하여 지식기반 질환발생 예측모델의 고도화와 함께, 질환 발생 감소를 위한 개인 맞춤형 건강관리 방법을 제공하고자 하고 있다. 이에 맞추어 고령화 사회에서 가장 대표적으로 흔하며 의료 이용과 의료비 증가의 원인이 되는 노화관련질병(ARD)의 분석을 통하여 14개의 질병을 도출하였다(제1장).

본 연구에서는 기 도출된 노화관련질병 중 하나인 골다공증의 예측모형을 개발하고자 하였다. 개인 맞춤형 건강관리 방법 제공을 위하여 대표적 web site 에서의 골다공증의 교육자료 현황 분석과 함께 맞춤형 교육 자료를 개발하여(제4장), 2022년도에 골다공증 탑재를 위한 모델을 개발하고자 하였다.

2. 연구의 목적

개인 맞춤형 서비스 고도화를 위한 노화관련질병의 예측 및 탑재 모델 개발을 위해 대표적인 노화관련질병인 골다공증 질환에 대해 탑재용 예측모델을 개발한다.

제2절 이론적 고찰

1. 골다공증의 정의

세계보건기구(WHO)는 골다공증을 “골량의 감소와 미세구조의 이상을 특징으로 하는 전신적인 골격계 질환으로서 결과적으로 뼈가 약해져서 부러지기 쉬운 상태가 되는 질환”으로 정의하였고, 미국 국립보건원(NIH)에서는 “골강도의 약화로 골질의 위험성이 증가하는 골격계질환”으로 규정하였다. 골강도는 골량(quantity)과 골질(quality)로 결정되며, 골량은 주로 골밀도(BMD)로 표현되고 골질은 구조, 골교체율, 무기질화, 미세손

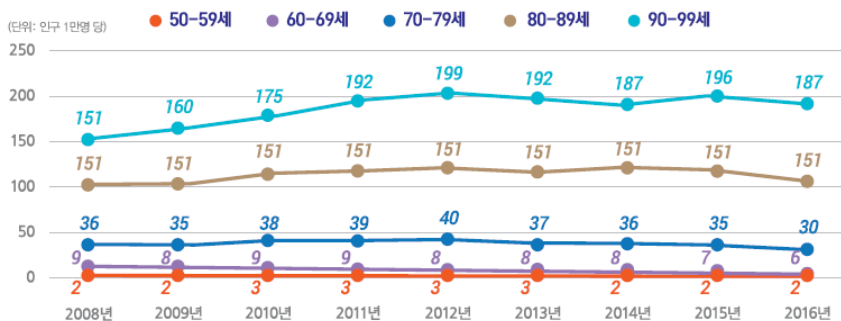
상 축적 등으로 구성된다.¹ 골다공증은 WHO의 가이드라인에 따라 골밀도 측정기로 골밀도를 측정해서 T 점수가 -2.5 이하를 보이는 경우를 골다공증, -1.0에서 -2.5를 보이는 경우 골감소증으로 진단한다.

골다공증은 원인, 연령, 임상적 특징에 따라 일차성 골다공증과 이차성 골다공증으로 분류하며, 일차성 골다공증은 성인에서 골다공증을 일으킬 만한 다른 질환이 동반되지 않은 상태에서 발생하는 골다공증을 말한다. 일차성 골다공증은 1형과 2형으로 나뉘며, 제1형 골다공증(폐경후 골다공증)은 여성에서 폐경 후 여성호르몬 박탈로 인하여 발생하는 골다공증이고 제2형 골다공증(노년골다공증)은 고령의 남녀에게 칼슘과 비타민D의 부족으로 인하여 발생하는 골다공증이다. 이차성 골다공증은 발병 연령에 상관없이 골다공증을 유발시키는 원인질환이 선행되어 발생하는 골다공증으로, 원인질환의 치료가 골다공증 치료의 가장 중요한 요소이다.²

2. 골다공증의 역학

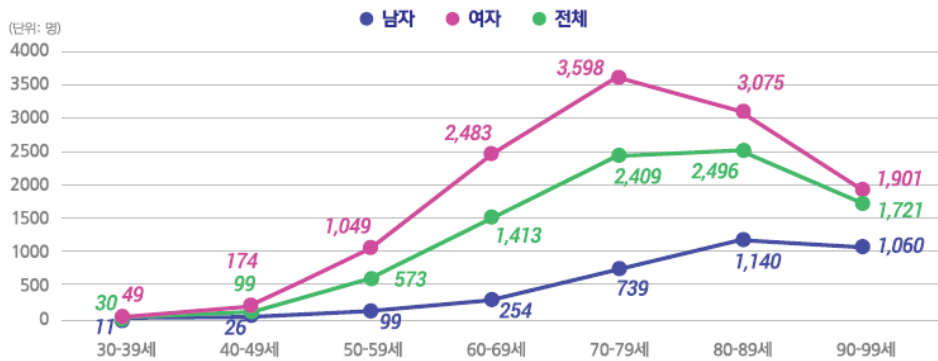
2008-2011 국민건강영양조사 자료에 따르면 50세 이상 골다공증 유병률은 22.4%, 골감소증의 유병률은 47.9%로 성인 5명 중 1명이 골다공증 환자이고 2명 중 1명은 골감소증이 있다. 70세 이상 여성은 68.5%가 골다공증 환자이며, 여성에서는 10세 단위로 연령이 증가할 때마다 골다공증이 2배씩 증가한다.³

골다공증성 골절은 여성에서 폐경 후 골밀도 감소와 연관이 있으며, 외상 이외 골절이 동반된 경우를 말한다. 주로 발생하는 부위는 척추, 요골, 대퇴골 및 상완골과 그 외 골반, 늑골, 발목 등이 있다.⁴ 2008-2016 건강보험공단 청구 자료에 의하면 50세 이상에서 골다공증 골절 발생 건수는 2008년 이후 점차 증가하는 양상이다. 50대에는 손목골절이 주로 발생하고 고령으로 갈수록 척추 및 고관절 골절의 발생률이 증가하였다.³



[그림 3-1] 고관절 골절의 연령별 발생률³
(2008-2016 건강보험공단 청구자료)

골절 후 1년 내 치명률이 고관절 골절은 남성이 여성의 1.5배, 척추골절은 2.2배이며, 골다공증 골절 발생 후 재골절은 4년까지 매년 비슷한 수가 발생하며 재골절도 척추골절이 가장 많이 발생하였다.³ 이와 같이 골다공증은 진단 및 치료가 중요한 질환이지만 골다공증 환자 10명 중 4명은 의료서비스 이용을 하지 않고 있고, 여성 골다공증 환자 10명 중 7명, 남성 골다공증 환자 10명 중 8명은 치료를 하지 않고 있다. 골절 후에도 42%에서만 약물 치료를 받았고 골절 부위 중 척추골절에서 약물 치료율이 53%로 가장 높았다. 골다공증 골절 발생 후 1년 내 약물 치료율은 여성이 남성의 2.3배 정도 높았으며, 연령이 높을수록 증가하다가 80세 이후 감소하였다. 골다공증 치료제 처방은 2016년까지 연평균 6%씩 꾸준히 증가하고 있으나, 지속 치료율은 6개월에 45.4%, 1년에 33.2%, 2년에 21.5%에 그치고 있다.³ 이처럼 골다공증은 유병율도 높고 골절 시 치명율도 높지만 관리가 잘 되지 않고 있어 앞으로 잘 예측하고 관리해야 할 질환으로 생각된다.



의료 이용률 = [(의료서비스를 이용한 골다공증 환자수)/(전체 골다공증 환자수)] x 100 (%)
 전체 골다공증 환자수 = [골밀도에 근거한 골다공증 유병율 (KNHANES 2008-2010)] x [해당연도 추정 인구] (n)
 자료는 2008-2012 건강보험공단 청구자료에서 추출함

[그림 3-2] 골다공증 환자의 연령-성별에 따른 의료이용률³
 (2008-2012 건강보험공단 청구자료)

한국 성인 골다공증 부담에 대한 국민건강보험 데이터베이스 연구에서는 2005-2008년까지 한국 성인의 골다공증 관련 의료 청구가 크게 증가하였고, 특히 폐경기 여성의 골다공증 부담이 현저한 것으로 나타났다.⁵ 한국 여성의 골다공증 유병률, 인식, 치료에 관한 제4차 국민건강영양조사를 이용한 연구에서는 골다공증 유병률은 높았으나 진단과 치료가 부족하였고, 정기적인 건강검진은 골다공증에 대한 인식과 치료율을 높이는 데 효과적인 전략이 될 수 있다는 결과를 보였다.⁶

3. 골다공증 위험인자

골다공증의 위험인자는 조절 가능한 위험인자와 조절 불가능한 인자로 구분할 수 있다. 골다공증의 발생은 연령이 증가하면서 같이 증가되고 골다공증과 골절이 여성들에게서 흔하지만 남성들에서도 연령이 증가하면서 골다공증의 발생 위험이 증가한다. 골다공증 및 골절의 가족력, 유전 및 환경적 요소들이 골다공증의 발생에 중요한 역할을 한다. 수술적 혹은 자연적 폐경은 빠른 골소실을 초래하며 폐경 후 5년 이상 동안 1년에 2-3% 정도 소실되는 것으로 알려져 있다. 에스트로젠뿐만 아니라 테스토스테론 결핍도 골다공증의 위험인자이지만 에스트로겐에 비하면 상대적으로 역할이 적다. 조절 가능한 위험인자로는 칼슘섭취, 비타민 D 섭취, 성호르몬부족, 만성적 글루코코르티코이드 등의 사용, 체중, 육체적 활동, 흡연이 있다.²

한국인을 대상으로 한 연구에서는 나이, BMI, 골절 과거력, 골다공증 가족력, 흡연, 음주, 운동, 소득, 학력, 칼슘 섭취, 초경 연령, 마지막 출산 연령, 자녀 수, 폐경, 폐경 후 기간, 호르몬 대체요법 등이 골다공증의 위험인자였다.⁷

또한 한국인 여성 골다공증 및 골감소증의 위험요인에 대한 지역사회 단면연구(2008)에서는 나이, 폐경상태, 체중, BMI, 학력, 골다공증 가족력, 흡연, 음주, 초경연령, 폐경연령, 자녀 수, 피임약 사용, 폐경 후 호르몬 치료 등을 위험인자로 선정하였다.⁸

4. 국가검진에서 골다공증 검사 현황

일반건강검진에서는 만 54세와 66세 여성을 대상으로 골다공증 검사를 시행하고 있으며, 의료급여생애전환기 검진에서도 만66세 여성을 대상으로 골다공증 검사를 시행하고 있다. 만66세 골다공증 검사는 2007년부터 시행되었고, 만54세 골다공증 검사는 2018년부터 시행되었다.

2008-2009년 국민건강영양조사를 이용한 50세 이상 성인의 골다공증 유병률과 국가검진 수검자 여성의 진단율 상승에 대한 연구에서는 골다공증에 대한 전국적인 검사 프로그램은 한국 노인 여성의 진단율을 높이는데 기여하였으나, 골다공증 진단 후 치료는 불충분한 것으로 나타났다.⁹

제3절 연구 내용 및 방법

1. 연구 내용

일반검진 수검자를 대상으로 골다공증과 나이, 성별, 비만도, 허리둘레, 혈압, 당뇨, 콜레스테롤, 생활습관(흡연, 음주, 운동)과의 관련성을 분석하였다. 또한 일반검진 대상자의 골다공증 발생 예측모형을 구축하고 적합도를 평가하였다.

2. 연구 방법

1) 연구 대상자 및 정의

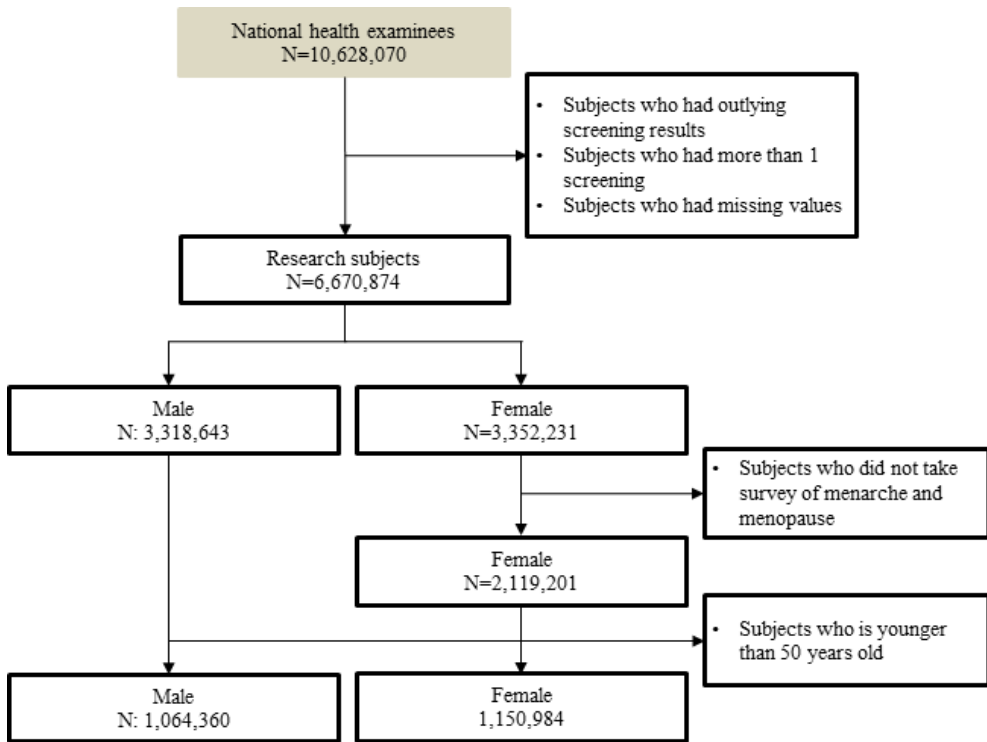
① 연구대상자

이 연구는 국민건강보험공단 맞춤형 자료(2009년)를 이용하였다. 국민건강보험공단에서 제공하는 맞춤형 자료는 우리나라 건강보험 가입자 전체의 자료 중 연구자가 원하는 대상자에 대하여 맞춤형으로 추출하여 제공하는 자료이며, 대상자의 2002년부터 2019년까지의 자격 정보, 진료내역, 처방내역, 요양기관 정보, 건강검진결과, 출생 및 사망 정보 등을 포함하고 있다. 자료 제공 시 해당 연구 대상자의 자료를 익명화 처리 후 연구자에게 제공하고 있다. 이 연구에서는 국민건강보험 가입자 중 2009년 건강검진을 실시한 10,628,070 명을 연구대상자로 선정하였다.

② 제외 기준

2009년 이전에 골다공증이 발생한 환자는 제외하였으며, 검진 결과를 토대로 예측모형이 구축되었기 때문에 이상치의 검진 결과를 가진 대상자를 제외하였다. 이상치의 기준은 하위값 $Q1-1.5*(Q3-Q1)$, 상위값 $Q3+1.5*(Q3-Q1)$ 로 선정하였다. 또한, 여성 골다공증 예측에 폐경 여부, 초경 나이 변수를 추가하기 위하여 여성 중 자궁경부암 검진을 받지 않은 대상자도 제외하였으며, 50세 미만 대상자 역시 제외하였다.

그 외 분석에 사용된 변수의 결측치를 가진 대상자를 제외하였으며, 최종 연구대상자는 총 2,215,344명이 선정되었다.



[그림 3-3] 연구대상자 flow chart

2) 분석내용 및 방법

① 골다공증 발생의 생존분석

2009년 국가건강검진 대상자에 대하여 검진 일자를 기준으로 10년간의 추적 관찰 기간 중에 골다공증 발생 여부에 대하여 생존 분석 (Cox proportional hazards ratio)을 시행하였다.

가. 종속 변수

종속변수는 골다공증 발생여부로 KCD 코드 (M80, M81, M82) 로 내원하였을 경우 골다공증 발생으로 정의하였다.

나. 위험인자 선정

이 연구에서는 자격 정보를 통한 성별, 나이와 검진 결과를 통한 BMI, 허리둘레, 혈압, 공복혈당수치, 청력이상여부, 시력, TG, LDL, HDL, GFR, HGB, GGT, SGPT, SGOT, Creatinine, 요단백, 흡연, 음주, 운동, 폐경여부, 초경나이 등이 위험요인으로 선정되었다.

나이는 50-60, 60-70, 70-80, 80세 이상으로 구분하였다, 체질량지수는 18.5 미만(저체중), 18.5~24.9(정상), 25이상(비만)으로 분류하였다. 당뇨병은 공복혈당 126 이상 또는 설문지에 의사에게 당뇨병 진단받은 경우로 정의하였으며, 공복 혈당이 100이 넘는 경우 공복혈당장애로 분류하였다. 고혈압은 수축기 140이상 또는 이완기 90이상 또는 설문지상 고혈압을 진단받은 경우로 정의하였으며, 수축기 120 이상 또는 이완기 80 이상의 경우 고혈압 전단계로 정의하였다. GFR (정상:90이상/중위험:60~90, 고위험: 60미만), 요단백 (정상/저위험:Trace (-/+), (+) / 고위험: (++) 이상) 등은 정상, 중위험, 고위험 3그룹으로 분류하였으며, 청력 (이상있음/이상없음), TG (200이상/200미만), HDL (40미만/40이상), LDL (160이상/160미만), HGB(12이상/12미만), SGOT (40이상/40미만), SGPT(40이상/40미만), 시력(1.0미만/1.0이상)은 위험, 정상 2그룹으로 분류하였다.

생활습관은 다음과 같이 분류하였다. 흡연의 경우 비흡연, 과거흡연, 현재흡연 세 가지로 분류한다. 음주의 경우 WHO 권고 기준에 따라 평균 하루 섭취 알코올 용량을 분류한다. 남자의 경우 0<저위험<40g, 40≤중위험<60g, 60g≤고위험으로 분류하였다. 여자의 경우 0<저위험<20g 20≤중위험<40g, 40g≤고위험으로 분류한다. 평균 하루 섭취 알코올 용량은 A x50ml(소주한잔 용량) x 0.22(소주 도수) x 0.8(알콜 함량비)x B/7으로 계산한다. (A: 하루 섭취하는 잔수, B: 1주일에 마시는 횟수). 운동의 경우 WHO에서 제공하는 IPAQ-K 기준에 의해 분류한다. 0≤저강도 활동<599 MET x min, 600≤중등도활동<2999MET x min, 3000MET x min< 고활동군으로 분류한다. 운동은 설문지에 근거하여 격렬한 신체활동을 한 시간이상 한 경우 주당 횟수 x 8.0MET x 60분으로 계산하고, 중등강도 신체활동을 한 시간 이상 한 경우 주당 횟수 x 4.0MET x 60분으로 계산, 저강도활동을 30분이상한 경우 걷기 주당 횟수 x 3.3MET x 30분으로 계산한다.

다. 통계적 분석

골다공증 발생에 따른 대상자의 기본적 특성 차이를 보기 위하여 Chi-square test 및 T-test 분석을 진행하였고, 위험인자들에 따른 골다공증 발생을 보기위하여 생존분석 (Cox proportional hazards ratio)을 시행하였다. 모든 분석은 SAS 9.4를 이용하였다.

② 골다공증 발생의 Machine learning 예측 모델

2009년 국가건강검진 대상자에 대해 검진 일자를 기준으로 10년 내에 골다공증

발생 여부를 통하여 예측모델을 구축하였다. Lasso model과 Logistic regression 모델은 연구대상자 전수로 예측모델을 구축하였으며, 그 외 나머지 machine learning model은 맞춤형 Room의 제한으로 인해 전체 대상자 중 20퍼센트를 무작위 표본 추출하여 예측모델을 구축하였다.

가. 변수 및 데이터 전처리

모델 구축 전 검진 결과 이상치를 제거하였으며, Correlation matrix를 통하여 높은 상관관계를 가지는 변수를 제거하였다 (Total cholesterol). 모든 변수는 0에서 1 사이의 값을 가지도록 rescaling 하였다.

남자에서 event 발생 (골다공증)이 너무 적어 sensitivity가 너무 낮게 예측모델이 구축되어, 이를 해결하기 위해 남성 예측모델을 구축할 때 골다공증 발생에 대한 over sampling, 그리고 미발생에 대한 under sampling 후 weight를 주었다.

나. 변수 및 모델 선정

변수 선정은 앞서 진행한 cox regression의 유의한 값과 stepwise selection을 통해 이루어졌다. 모델 구축 후 variable importance를 확인하였다.

모델은 총 5가지 모델로, Random forest model, eXtreme gradient boosting model, Lasso model, K-nearest Neighbor model, logistic regression model로 구축하였으며, 이 중 예측력이 가장 높은 모델을 최종 모델로 선정한다.

다. 모델 검증

예측모델의 검정력을 높이기 위해 5-fold cross validation을 시행하였으며, 5:5로 training set 과 test set을 나누어 training set에서 모델을 구축한 후, test set에서 검정하였다. 또한 tunelength를 통한 hyperparameter optimization 과정을 거쳐 모델을 선정하였다. Training set에서 구축된 모델의 검정을 위해, Test set에 적용 후 ROC 그래프, Model accuracy, Sensitivity, specificity, Positive predicted value, negative predicted value 등을 확인하였다.

③ 통계적 분석

위험인자들에 따른 골다공증 발생 및 stepwise selection을 보기 위하여 생존분석 (Cox proportional hazards ratio)을 시행하였다.

예측 모델 구축을 위해 Random forest model, eXtreme gradient boosting model, Lasso model, K-nearest Neighbor model, Multivariable logistic regression

을 사용하였으며, 모든 분석은 SAS 9.4와 R studio를 사용하였다.



[그림 3-4] Data split과 5-fold cross validation (출처: Siladittya, 2020)

제4절 분석 결과

1. 일반검진 대상자의 특성

최대 10년의 추적관찰 기간 중 여성은 전체 대상자 1,150,984명 중 38.19%인 439,616명에서 골다공증이 발생하였고 남성은 전체 대상자 1,064,360명 중 7.05%인 75,006명에서 골다공증이 발생하였다.

여성은 70대에서 발생률이 가장 높았고, 고혈압, 당뇨가 있는 경우 발생률이 높았다. GFR 고위험군 및 요단백 정상군, 시력 위험군에서 발생률이 높았으며, TG, HDL 위험군 및 LDL 정상군, HBG 위험군에서 발생률이 높았다. SGOT 위험군과 SGPT 정상군에서 발생률이 높았으며, BMI는 저체중인 경우 발생률이 높았다. 초경나이는 17세 이상인 경우 발생률이 높았으며 폐경인 경우 발생률이 높았다. 생활습관의 경우 비흡연군에서

발생률이 높았고, 저활동군과 저위험 음주군에서 골다공증 발생률이 높았다.

남성은 80세 이상에서 발생률이 가장 높았고, 고혈압과 당뇨병이 있는 경우 발생률이 높았다. GFR 고위험군 및 요단백 고위험군, 시력 위험군에서 발생률이 높았으며, TG, HDL, LDL 정상군, HBG 위험군에서 발생률이 높았다. SGOT 위험군과 SGPT 정상군에서 발생률이 높았으며, BMI는 저체중인 경우 발생률이 높았다. 생활습관의 경우 비흡연군에서 발생률이 높았고, 중등도 활동군과 저위험 음주군에서 발생률이 높았다.

<표 3-1> 일반검진 대상자 특성

	여성					남성				
	미발생		발생		p-value	미발생		발생		p-value
	빈도	%	빈도	%		빈도	%	빈도	%	
전체	711368	61.81	439616	38.19		989354	92.95	75006	7.05	
나이					<.0001					<.0001
50-59	443597	69.2	197465	30.8		559356	96.52	20149	3.48	
60-69	189513	52.33	172615	47.67		318924	90.16	34800	9.84	
70-79	71388	52.21	65334	47.79		103972	84.78	18662	15.22	
80이상	6870	62.05	4202	37.95		7102	83.58	1395	16.42	
흡연여부					<.0001					<.0001
비흡연	685315	61.71	425257	38.29		356338	92.18	30220	7.82	
과거흡연	8016	64.02	4505	35.98		321853	93.02	24165	6.98	
흡연	18037	64.67	9854	35.33		311163	93.78	20621	6.22	
운동					<.0001					<.0001
고활동군	65363	63.13	38175	36.87		124068	92.76	9690	7.24	
중등도 활동군	216366	62.78	128261	37.22		401632	93.64	27274	6.36	
저활동군	429639	61.13	273180	38.87		463654	92.42	38042	7.58	
음주					<.0001					<.0001
저위험 음주군	603255	60.89	387396	39.11		437642	91.92	38472	8.08	
중위험 음주군	105588	67.36	51153	32.64		523198	93.87	34176	6.13	
고위험 음주군	2525	70.3	1067	29.7		28514	92.36	2358	7.64	
고혈압					<.0001					<.0001
정상	235904	64.28	131089	35.72		252491	93.59	17280	6.41	
전기 고혈압	267011	62.93	157296	37.07		427371	93.68	28856	6.32	
고혈압	208453	57.95	151231	42.05		309492	91.47	28870	8.53	
당뇨병					<.0001					<.0001
정상	633497	62.54	379501	37.46		821682	93.29	59133	6.71	
공복혈당 장애	52166	60.2	34493	39.8		110644	92.84	8532	7.16	

	여성					남성				
	미발생		발생		p-value	미발생		발생		p-value
	빈도	%	빈도	%		빈도	%	빈도	%	
당뇨병	25705	50.08	25622	49.92		57028	88.6	7341	11.4	
GFR					<.0001					<.0001
정상	113664	67.38	55024	32.62		204242	95.04	10665	4.96	
중위험군	422634	63.03	247874	36.97		625280	93.41	44091	6.59	
고위험군	175070	56.15	136718	43.85		159832	88.76	20250	11.24	
요단백					<.0001					<.0001
정상	681040	61.78	421246	38.22		946855	92.98	71521	7.02	
중위험군	25785	62.18	15685	37.82		36426	92.61	2906	7.39	
고위험군	4543	62.85	2685	37.15		6073	91.3	579	8.7	
시력					<.0001					<.0001
정상	146922	67.59	70438	32.41		312170	95.15	15918	4.85	
위험군	564446	60.46	369178	39.54		677184	91.97	59088	8.03	
TG					<.0001					<.0001
정상	650379	61.98	399027	38.02		875773	92.85	67425	7.15	
위험군	60989	60.04	40589	39.96		113581	93.74	7581	6.26	
LDL					<.0001					<.0001
정상	613605	61.74	380250	38.26		910198	92.89	69623	7.11	
위험군	97763	62.22	59366	37.78		79156	93.63	5383	6.37	
HDL					<.0001					<.0001
정상	272312	62.6	162683	37.4		248383	92.59	19877	7.41	
위험군	439056	61.32	276933	38.68		740971	93.08	55129	6.92	
HGB					<.0001					<.0001
정상	605236	62.11	369260	37.89		973951	93.05	72705	6.95	
위험군	106132	60.14	70356	39.86		15403	87	2301	13	
SGOT					<.0001					<.0001
정상	706682	61.82	436406	38.18		980290	92.96	74198	7.04	
위험군	4686	59.35	3210	40.65		9064	91.82	808	8.18	
SGPT					<.0001					<.0001
정상	694670	61.8	429345	38.2		944388	92.92	71941	7.08	
위험군	16698	61.92	10271	38.08		44966	93.62	3065	6.38	
BMI					<.0001					<.0001
정상	463478	61.75	287114	38.25		654820	92.82	50657	7.18	
저체중	12066	56.19	9407	43.81		17951	89.76	2047	10.24	
비만	235824	62.24	143095	37.76		316583	93.42	22302	6.58	
초경나이					<.0001					
16세이하	405329	64.15	226545	35.85		-	-	-	-	
17세이상	306039	58.95	213071	41.05		-	-	-	-	
폐경					<.0001					
아니오	187495	69.6	81886	30.4		-	-	-	-	
예	523873	59.42	357730	40.58		-	-	-	-	

생존분석에서는 여성의 경우 나이에 따라 가장 큰 격차를 보였으며, 그 외 당뇨병, 체중, 폐경 여부에서 큰 차이를 보였다. 남성의 경우에도 나이에 따라 가장 큰 격차를 보였으며, 당뇨병, HGB, 체중에서 큰 차이를 보였다.

<표 3-2> 일반검진 대상자 생존분석

	여성			남성		
	HR	95% CI		HR	95% CI	
		Low	High		Low	High
전체						
나이	Ref	-	-	Ref	-	-
50-59						
60-69	1.476	1.465	1.487	2.337	2.294	2.381
70-79	1.392	1.378	1.407	3.104	3.032	3.178
80이상	1.054	1.021	1.087	3.119	2.946	3.303
흡연여부	Ref	-	-	Ref	-	-
비흡연						
과거흡연	0.985	0.956	1.014	1.037	1.019	1.055
흡연	1	0.98	1.021	1.003	0.985	1.022
운동	Ref	-	-	Ref	-	-
고활동군						
중등도활동군	1.015	1.003	1.027	1.02	0.997	1.044
저활동군	1.034	1.023	1.045	1.07	1.046	1.094
음주	Ref	-	-	Ref	-	-
저위험음주군						
중위험음주군	0.908	0.9	0.917	0.896	0.883	0.91
고위험음주군	0.85	0.8	0.902	0.946	0.907	0.986
고혈압	Ref	-	-	Ref	-	-
정상						
전기고혈압	1.002	0.994	1.009	0.951	0.933	0.969
고혈압	1.091	1.082	1.099	1.038	1.018	1.059
당뇨병	Ref	-	-	Ref	-	-
정상						
공복혈당장애	1.084	1.072	1.096	1.077	1.053	1.102
당뇨병	1.433	1.415	1.452	1.55	1.512	1.589
GFR	Ref	-	-	Ref	-	-
정상						
중위험군	1.044	1.034	1.054	1.023	1.001	1.045
고위험군	1.084	1.072	1.096	1.082	1.054	1.111
요단백	Ref	-	-	Ref	-	-

	여성			남성		
	HR	95% CI		HR	95% CI	
		Low	High		Low	High
정상						
중위험군	0.989	0.973	1.005	1	0.963	1.037
고위험군	0.94	0.905	0.977	1.059	0.975	1.149
시력						
정상	Ref	-	-	Ref	-	-
위험군	1.043	1.034	1.052	1.087	1.068	1.108
TG						
정상	Ref	-	-	Ref	-	-
위험군	1.013	1.002	1.023	0.971	0.948	0.995
LDL						
정상	Ref	-	-	Ref	-	-
위험군	1	0.991	1.009	0.973	0.946	1
HDL						
정상	Ref	-	-	Ref	-	-
위험군	0.969	0.963	0.975	0.928	0.912	0.943
HGB						
정상	Ref	-	-	Ref	-	-
위험군	1.05	1.042	1.059	1.329	1.274	1.386
SGOT						
정상	Ref	-	-	Ref	-	-
위험군	1.032	0.996	1.069	1.056	0.985	1.133
SGPT						
정상	Ref	-	-	Ref	-	-
위험군	1.054	1.033	1.075	1.035	0.998	1.074
BMI						
정상	Ref	-	-	Ref	-	-
저체중	1.088	1.066	1.111	1.225	1.171	1.281
비만	0.95	0.943	0.956	0.967	0.951	0.983
초경나이						
16세이하	Ref	-	-	-	-	-
17세이상	1.033	1.026	1.039	-	-	-
폐경						
아니오	Ref	-	-	-	-	-
예	1.124	1.115	1.133	-	-	-

65세 미만 생존 분석에서는 여성의 경우 나이에 따라 가장 큰 격차를 보였으며, 고혈압, 당뇨병, GFR, BMI, 폐경여부에서 큰 차이를 보였다. 남성의 경우에도 나이에 따라

가장 큰 격차를 보였으며, 고혈압, 당뇨병, GFR, HGB, BMI에서 큰 차이를 보였다.

65세 이상 생존분석에서는 여성의 경우 나이에 따라 가장 큰 격차를 보였으며, 당뇨병에서 큰 차이를 보였다. 남성의 경우에도 나이에 따라 가장 큰 격차를 보였으며, 당뇨병, HGB, BMI에서 큰 차이를 보였다.

<표 3-3> 65세 미만, 65세 이상 생존분석

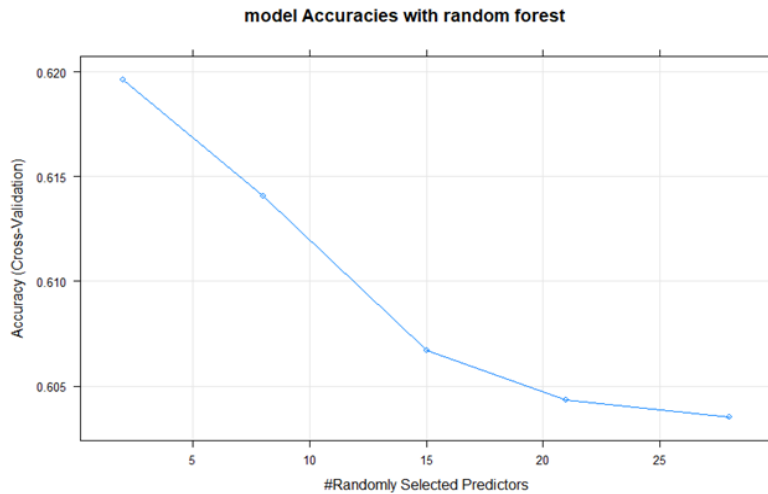
	65세 미만						65세 이상					
	여성			남성			여성			남성		
	HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI	
	Low	High		Low	High		Low	High		Low	High	
전체												
나이	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
50-59												
60-69	1.476	1.465	1.487	2.34	2.294	2.381	1.476	1.465	1.487	2.337	2.29	2.381
70-79	1.392	1.378	1.407	3.1	3.032	3.178	1.392	1.378	1.407	3.104	3.03	3.178
80이상	1.054	1.021	1.087	3.12	2.946	3.303	1.054	1.021	1.087	3.119	2.95	3.303
흡연여부												
비흡연	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
과거흡연	0.953	0.92	0.987	1.04	1.011	1.063	1.018	0.963	1.076	1.027	1	1.052
흡연	1.009	0.986	1.034	0.98	0.954	1.003	0.94	0.904	0.976	0.987	0.96	1.014
운동												
고활동군	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
중등도 활동군	1.008	0.995	1.022	0.97	0.935	0.999	1.006	0.984	1.029	1.046	1.01	1.08
저활동군	1.032	1.019	1.045	1.03	0.994	1.061	1.013	0.993	1.035	1.107	1.07	1.141
음주												
저위험 음주군	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
중위험 음주군	0.881	0.872	0.891	0.86	0.846	0.883	0.928	0.909	0.948	0.911	0.89	0.931
고위험 음주군	0.798	0.748	0.852	0.93	0.874	0.983	0.974	0.831	1.142	0.959	0.9	1.018
고혈압												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
전기 고혈압	1.028	1.02	1.037	0.98	0.95	0.999	0.978	0.963	0.993	0.945	0.92	0.973
고혈압	1.16	1.15	1.171	1.16	1.125	1.188	1.035	1.02	1.05	1.011	0.98	1.04
당뇨병												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
공복혈당 장애	1.097	1.082	1.113	1.08	1.046	1.115	1.065	1.046	1.085	1.079	1.05	1.115
당뇨병	1.445	1.42	1.471	1.6	1.542	1.661	1.42	1.394	1.447	1.522	1.47	1.573
GFR	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-

	65세 미만						65세 이상					
	여성			남성			여성			남성		
	HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI	
	Low	High		Low	High		Low	High		Low	High	
정상												
중위험군	1.075	1.064	1.086	1.12	1.091	1.148	0.995	0.968	1.022	1	0.96	1.045
고위험군	1.205	1.19	1.221	1.38	1.333	1.438	0.961	0.935	0.988	1.046	1	1.094
요단백												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
중위험군	0.985	0.966	1.004	1.05	0.995	1.105	0.987	0.96	1.014	0.946	0.9	0.997
고위험군	0.904	0.861	0.949	1.11	0.982	1.243	0.987	0.928	1.049	1.004	0.9	1.125
시력												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
위험군	1.094	1.084	1.104	1.2	1.169	1.221	0.975	0.95	1.001	1.061	1.03	1.097
TG												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
위험군	1.041	1.027	1.055	0.94	0.914	0.975	0.98	0.964	0.997	0.99	0.96	1.027
LDL												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
위험군	1.008	0.997	1.018	0.97	0.933	1.006	0.991	0.975	1.006	0.971	0.93	1.012
HDL												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
위험군	0.985	0.977	0.992	0.93	0.912	0.956	0.953	0.942	0.964	0.921	0.9	0.943
HGB												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
위험군	1.047	1.036	1.058	1.51	1.406	1.631	1.049	1.035	1.063	1.278	1.22	1.344
SGOT												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
위험군	1.053	1.008	1.1	1.03	0.928	1.143	1.014	0.956	1.076	1.105	1.01	1.214
SGPT												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
위험군	1.044	1.021	1.068	0.99	0.947	1.039	1.062	1.02	1.107	1.056	0.99	1.122
BMI												
정상	Ref	-	-	Ref	-	-	Ref	-	-	Ref	-	-
저체중	1.13	1.101	1.16	1.3	1.217	1.398	0.977	0.944	1.011	1.148	1.08	1.218
비만	0.952	0.944	0.96	1	0.974	1.02	0.984	0.972	0.995	0.954	0.93	0.977
초경나이												
16세이하	Ref	-	-	-	-	-	Ref	-	-	-	-	-
17세이상	1.071	1.063	1.079	-	-	-	1	0.989	1.011	-	-	-
폐경												
아니오	Ref	-	-	-	-	-	Ref	-	-	-	-	-
예	1.234	1.224	1.245	-	-	-	0.954	0.936	0.973	-	-	-

2. 골다공증 발생의 Machine learning 예측 모델 (여성)

1) Random forest model

Training set 에서 모델을 구축할 때, Cross validation을 통해 hyperparameter를 바꿔가면서 training set에서 반복적으로 모델을 구축하는 과정을 통해 가장 높은 정확도를 보이는 hyperparameter를 선정하게 된다. 다음 그래프와 표는 hyperparameter를 바꿔가면서 model accuracy를 측정한 결과이다. 여성 random forest model에서는 mtry가 2 일 때 training set에서 가장 높은 정확도를 보였으며 model accuracy는 0.6196이었다.

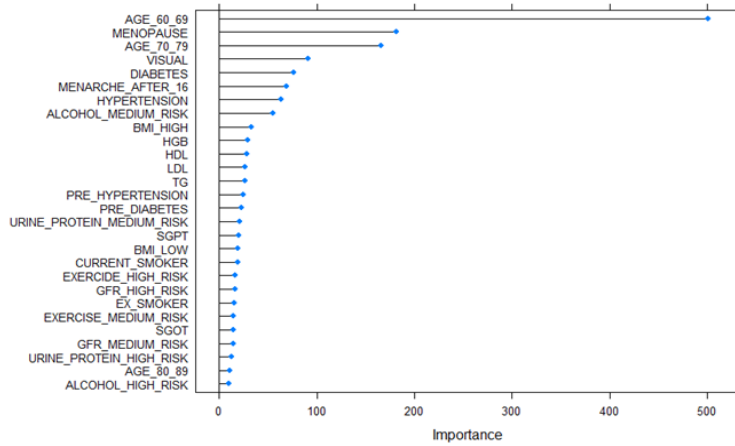


mtry	Accuracy	Kappa
2	0.619602	0.0004599903
8	0.614051	0.0602238425
15	0.606718	0.0664702036
21	0.604337	0.0646984581
28	0.603538	0.0630550295

[그림 3-5] Random forest의 modeling in training set

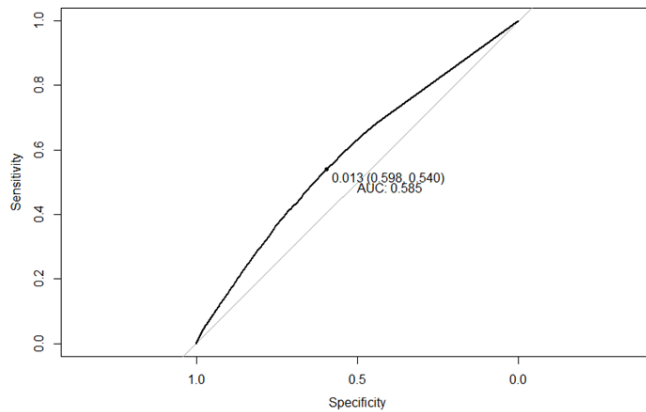
Variable importance는 모델을 구축할 때 어떤 변수에 중요도를 높게 봤는지를 나타낸다. 기존 통계모형처럼 해당 변수에 따른 값은 알 수 없지만, 모델을 구축할 때 어떤 변수를 중요하게 봤는지를 알 수 있다. Random forest의 variable importance는 나이 60-69세가 가장 높은 비중을 차지하였고, 그 다음이 폐경 여부,

그리고 나이 70-79세 순으로 높은 비중을 보였다.



[그림 3-6] Random forest의 variable importance

ROC curve는 cutoff를 계속 바꿔가면서 Y축의 sensitivity와 X축의 specificity로 그은 선을 말한다. 해당 ROC 선 아래의 면적을 AUC라고 하며, 모델의 설명력을 나타내는데 사용된다. 해당 모델의 AUC 값은 0.585이었으며 가장 적절한 cutoff는 0.013 으로 분석되었다.



[그림 3-7] Random forest의 cutoff point ROC curve

Training set에서 구축된 예측모델을 0.013의 확률값 cutoff를 적용하여 test set에서 모델 정확도를 평가하였을 때 model accuracy는 0.5757 이었으며, sensitivity는 0.5402, specificity는 0.5976, Positive predicted value는 0.4533, negative

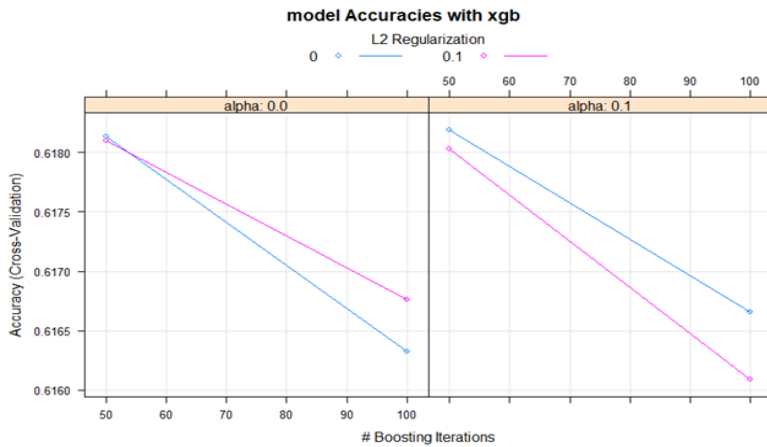
predicted value는 0,6779였다.

<표 3-4> Random forest의 model accuracy with test set

Random Forest	
AUC	0.585
Accuracy	0.5757
95% CI	(0.5729, 0.5786)
Sensitivity	0.5402
Specificity	0.5976
Pos Pred Value	0.4533
Neg Pred Value	0.6779
Balanced Accuracy	0.5689

2) eXtreme Gradient Boosting model

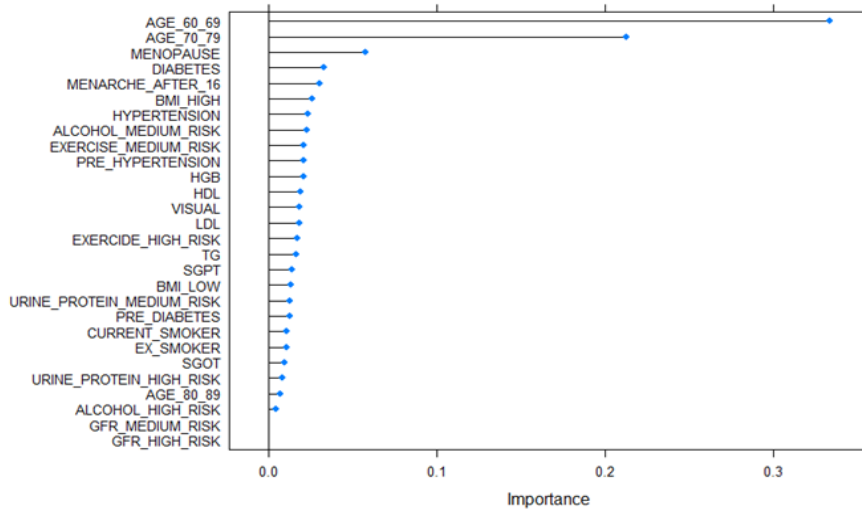
여성 extreme gradient boosting model에서는 lambda가 0, alpha가 0.1, nrounds가 50일 때 가장 높은 정확도를 보였으며 model accuracy는 0.6182이었다.



lambda	alpha	nrounds	Accuracy	Kappa
0	0	50	0.618134	0.056214
0	0	100	0.616327	0.060323
0	0.1	50	0.618186	0.056595
0	0.1	100	0.616657	0.060979
0.1	0	50	0.618099	0.055951
0.1	0	100	0.616761	0.06112
0.1	0.1	50	0.61803	0.056332
0.1	0.1	100	0.616092	0.05949

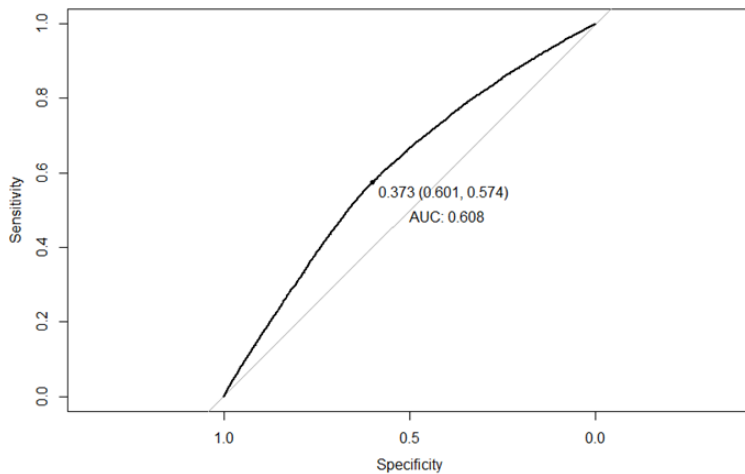
[그림 3-8] eXtreme Gradient Boosting의 modeling in training set

모델을 구축할 때 사용된 Variable importance에서는 나이 60-69세가 가장 높은 비중을 차지하였고, 그 다음이 나이 70-79세, 폐경 여부 순으로 높은 비중을 보였다.



[그림 3-9] eXtreme Gradient Boosting의 variable importance

모델의 AUC 값은 0.608이었으며 확률값의 cutoff는 0.373 으로 분석되었다.



[그림 3-10] eXtreme Gradient Boosting의 cutoff point ROC curve

0.373의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy 는 0.5906이었으며, sensitivity는 0.5739, specificity는 0.6009, Positive predicted

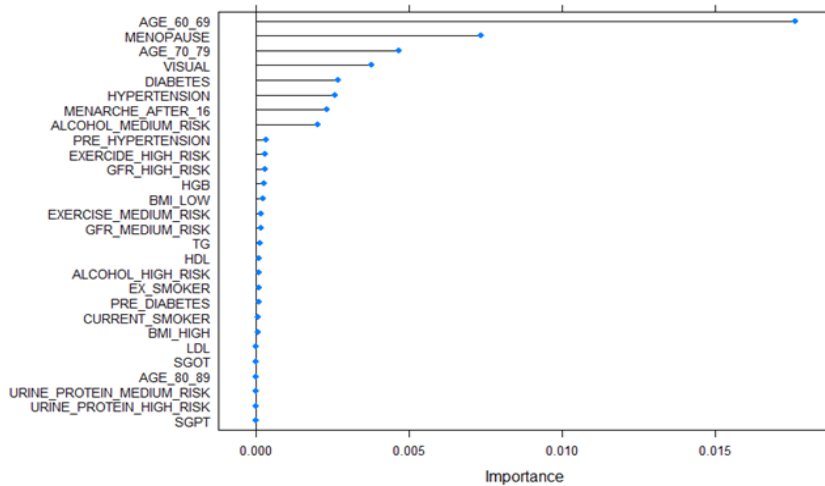
value는 0.4704, negative predicted value는 0.6954였다.

<표 3-5> eXtreme Gradient Boosting의 model accuracy with test set

XGB	
AUC	0.608
Accuracy	0.5906
95% CI	(0.5878, 0.5935)
Sensitivity	0.5739
Specificity	0.6009
Pos Pred Value	0.4704
Neg Pred Value	0.6954
Balanced Accuracy	0.5874

3) K-nearest neighbor model

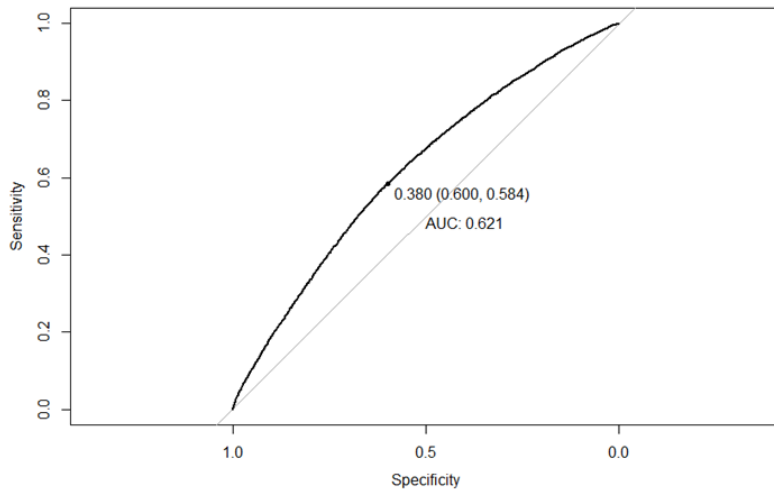
여성 K-nearest neighbor 모델의 Variable importance에서는 나이 60-69세가 가장 높은 비중을 차지하였고, 그 다음이 폐경 여부, 나이 70-79세 순으로 높은 비중을 보였다. K가 5일 때 Accuracy는 0.6110이었다.



K	Accuracy	Kappa
5	0.6110	0.0684

[그림 3-11] K-nearest neighbor의 variable importance

모델의 AUC 값은 0.621이었으며 확률값의 cutoff는 0.380 으로 분석되었다.



[그림 3-12] K-nearest neighbor의 cutoff point ROC curve

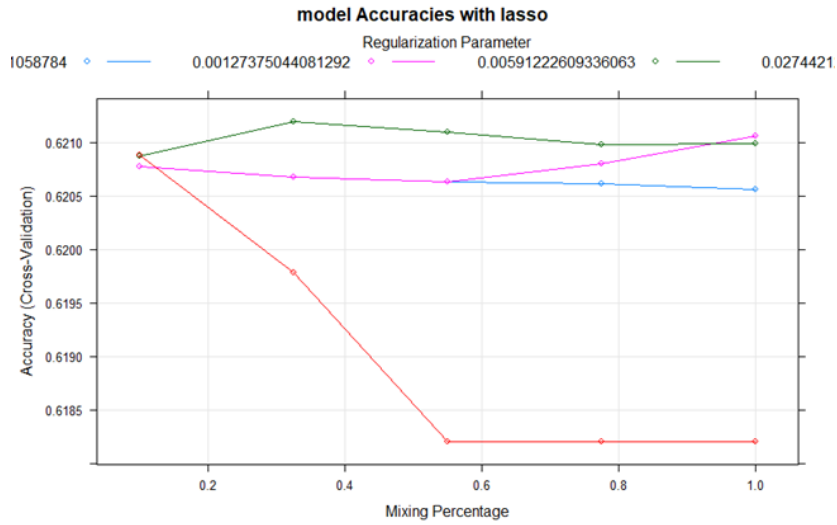
0.380의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy는 0.5933 이었으며, sensitivity는 0.5851, specificity는 0.5983, Positive predicted value는 0.4736, negative predicted value는 0.7001이었다.

<표 3-6> K-nearest neighbor의 model accuracy with test set

KNN	
AUC	0.621
Accuracy	0.5933
95% CI	(0.5904, 0.5961)
Sensitivity	0.5851
Specificity	0.5983
Pos Pred Value	0.4736
Neg Pred Value	0.7001
Balanced Accuracy	0.5917

4) Lasso model

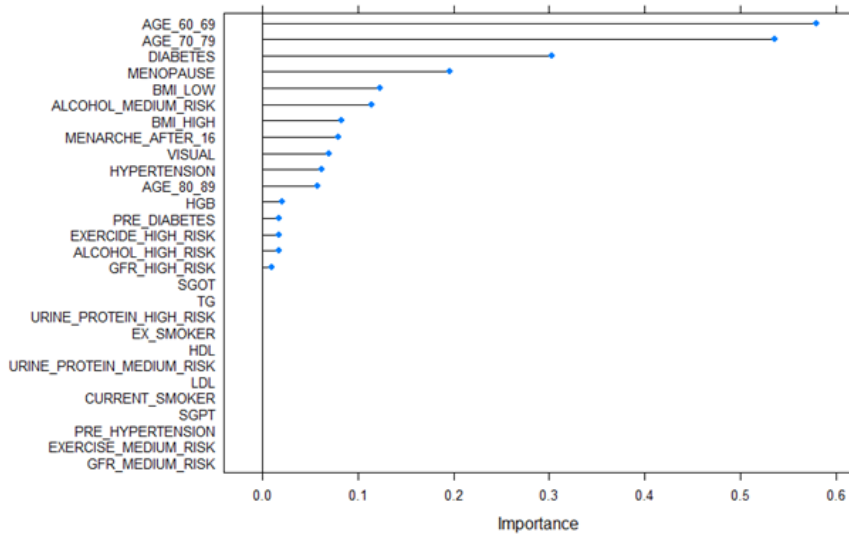
여성 Lasso model에서는 alpha가 0.325, lamda가 0.005912일 때 가장 높은 정확도를 보였으며 model accuracy는 0.6212였다.



alpha	lambda	Accuracy	Kappa
0.1	0.000274	0.620777	0.045657
0.1	0.001274	0.620777	0.045657
0.1	0.005912	0.620872	0.044704
0.1	0.027442	0.620877	0.024068
0.325	0.000274	0.620679	0.048009
0.325	0.001274	0.620679	0.048009
0.325	0.005912	0.621197	0.037066
0.325	0.027442	0.619783	0.012921
0.55	0.000274	0.620632	0.048269
0.55	0.001274	0.620632	0.048269
0.55	0.005912	0.621093	0.026567
0.55	0.027442	0.618202	0
0.775	0.000274	0.620617	0.048685
0.775	0.001274	0.620799	0.045519
0.775	0.005912	0.620975	0.023575
0.775	0.027442	0.618202	0
1	0.000274	0.620563	0.049044
1	0.001274	0.621062	0.042295
1	0.005912	0.620983	0.023166
1	0.027442	0.618202	0

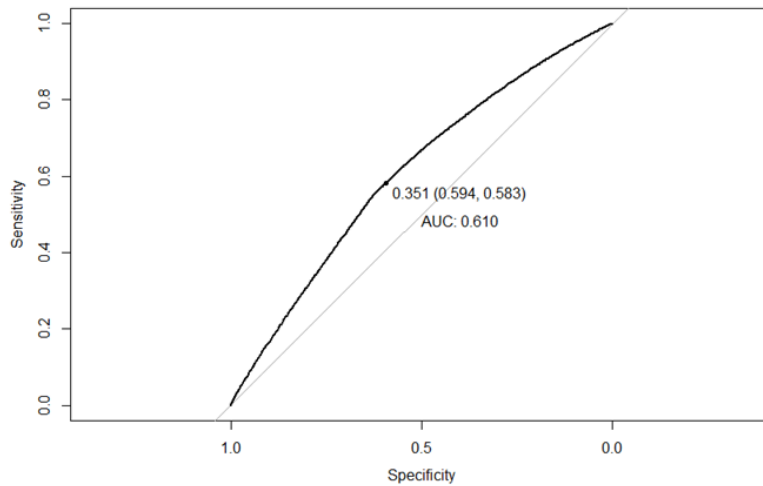
[그림 3-13] Lasso의 modeling in training set

모델을 구축할 때 사용된 Variable importance에서는 나이 60-69세가 가장 높은 비중을 차지하였고, 그 다음이 나이 70-79세, 당뇨병 순으로 높은 비중을 보였다.



[그림 3-14] Lasso의 variable importance

모델의 AUC 값은 0.610이었으며 확률값의 cutoff는 0.351 으로 분석되었다.



[그림 3-15] Lasso의 cutoff point ROC curve

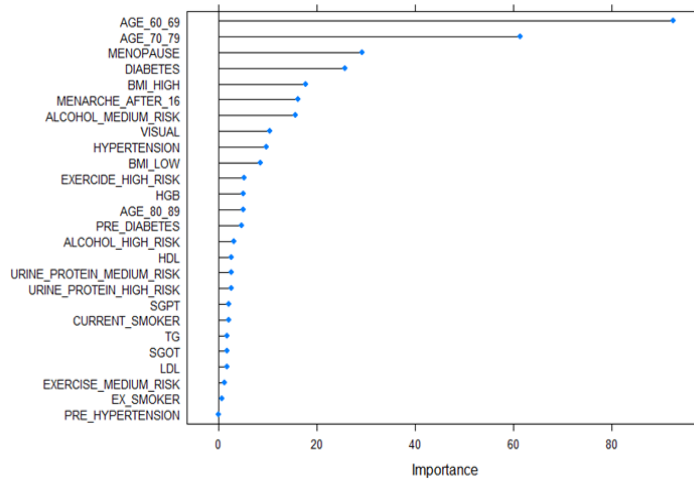
0.351의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy 는 0.5898이었으며, sensitivity는 0.5832, specificity는 0.5939, Positive predicted value 는 0.4703, negative predicted value는 0.6973이었다.

<표 3-7> Lasso의 model accuracy with test set

LASSO	
AUC	0.610
Accuracy	0.5898
95% CI	(0.5885, 0.5911)
Sensitivity	0.5832
Specificity	0.5939
Pos Pred Value	0.4703
Neg Pred Value	0.6973
Balanced Accuracy	0.5885

5) Logistic regression

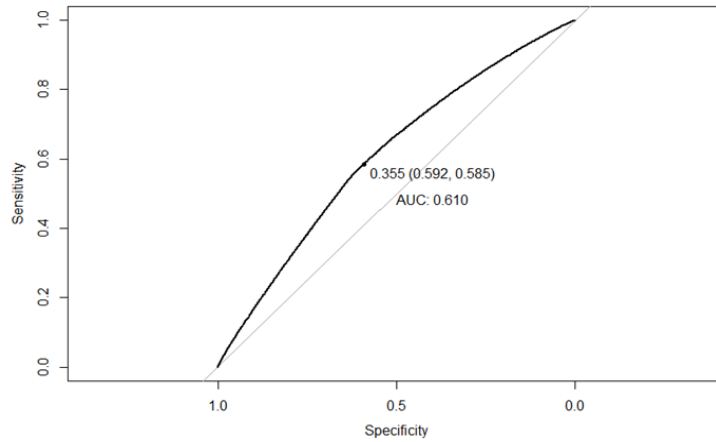
Logistic 모델을 구축할 때 사용된 Variable importance에서는 나이 60-69세가 가장 높은 비중을 차지하였고, 그 다음이 나이 70-79세, 폐경 여부 순으로 높은 비중을 보였다.



Accuracy	Kappa
0.6206	0.0545

[그림 3-16] Logistic의 variable importance

모델의 AUC 값은 0.610이었으며 확률값의 cutoff 는 0.355 으로 분석되었다.



[그림 3-17] Logistic의 cutoff point ROC curve

0.355의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy 는 0.5897이었으며, sensitivity는 0.5839, specificity는 0.5933, Positive predicted value는 0.4703, negative predicted value는 0.6975였다.

<표 3-8> Logistic의 model accuracy with test set

Logistic	
AUC	0.610
Accuracy	0.5897
95% CI	(0.5884, 0.591)
Sensitivity	0.5839
Specificity	0.5933
Pos Pred Value	0.4703
Neg Pred Value	0.6975
Balanced Accuracy	0.5886

6) 여성의 골다공증 예측모델 비교

다섯 개의 예측모델 중 AUC가 가장 높은 모델은 KNN model(0.621)이었고 accuracy가 가장 높은 모델은 Lasso model(0.5973)이었으며, positive predictive value가 가장 높은 모델은 Lasso model(0.4745)이었다.

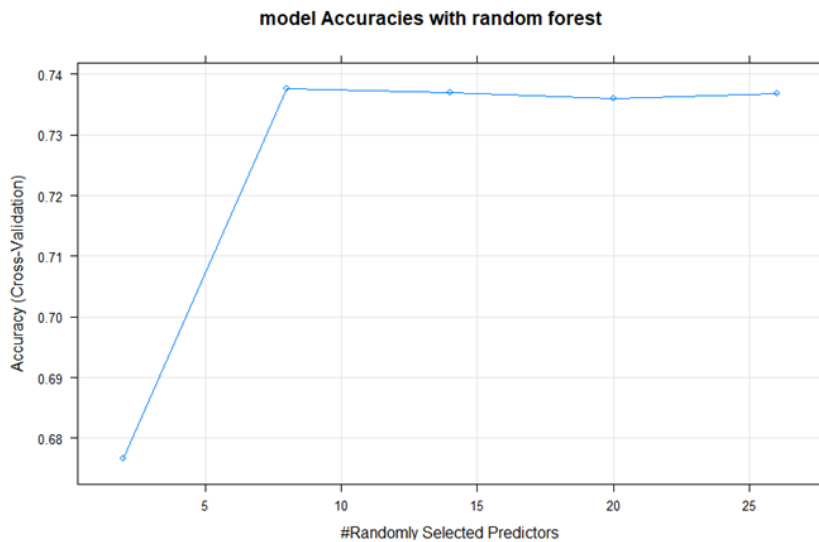
<표 3-9> 여성의 골다공증 예측모델 비교

	Random forest	XGB	KNN	Lasso	Logistic
AUC	0.585	0.608	0.621	0.610	0.610
Accuracy	0.5757	0.5906	0.5933	0.5898	0.5897
95% CI	(0.5729, 0.5786)	(0.5878, 0.5935)	(0.5904, 0.5961)	(0.5885, 0.5911)	(0.5884, 0.591)
Sensitivity	0.5402	0.5739	0.5851	0.5832	0.5839
Specificity	0.5976	0.6009	0.5983	0.5939	0.5933
Pos Pred Value	0.4533	0.4704	0.4736	0.4703	0.4703
Neg Pred Value	0.6779	0.6954	0.7001	0.6973	0.6975
Balanced Accuracy	0.5689	0.5874	0.5917	0.5885	0.5886

3. 골다공증 발생의 Machine learning 예측 모델(남성)

1) Random forest model

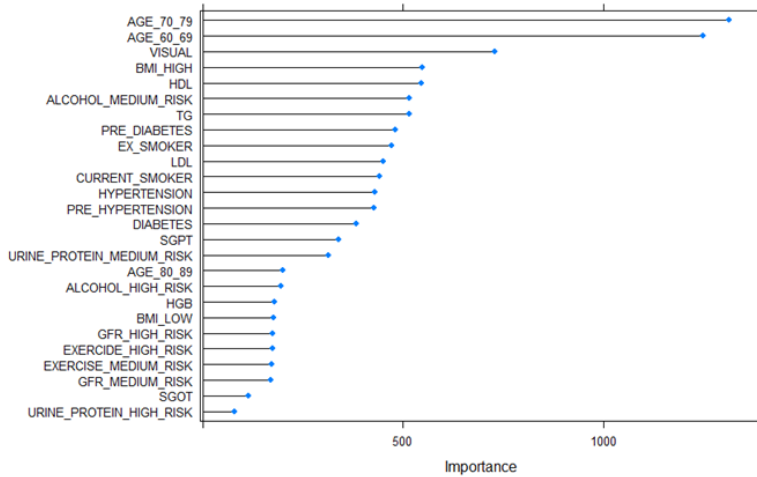
남성 random forest model에서는 mtry가 8일때 가장 높은 정확도를 보였으며 model accuracy는 0.7376이었다.



mtry	Accuracy	Kappa
2	0.6766818	0.3133850
8	0.7375661	0.4538934
14	0.7369237	0.4523373
20	0.7359788	0.4506265
26	0.7367725	0.4523625

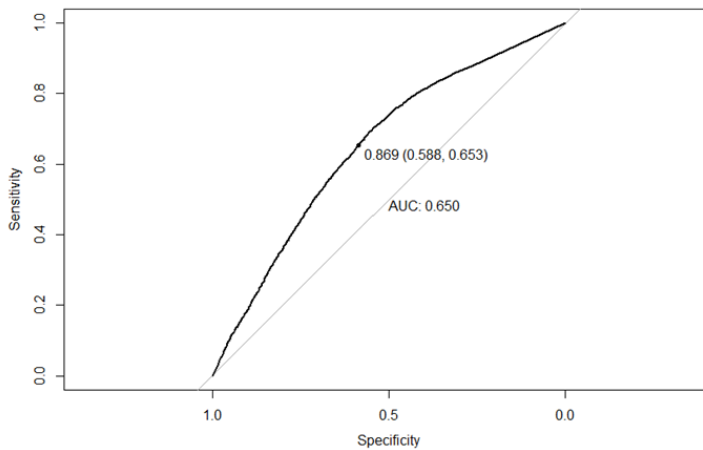
[그림 3-18] Random forest의 modeling in training set

모델을 구축할 때 사용된 Variable importance에서는 나이 70-79세가 가장 높은 비중을 차지하였고, 그 다음이 나이 60-69세, 시력 순으로 높은 비중을 보였다.



[그림 3-19] Random forest의 variable importance

모델의 AUC 값은 0.650이었으며 확률값의 cutoff는 0.869 으로 분석되었다.



[그림 3-20] Random forest의 cutoff point ROC curve

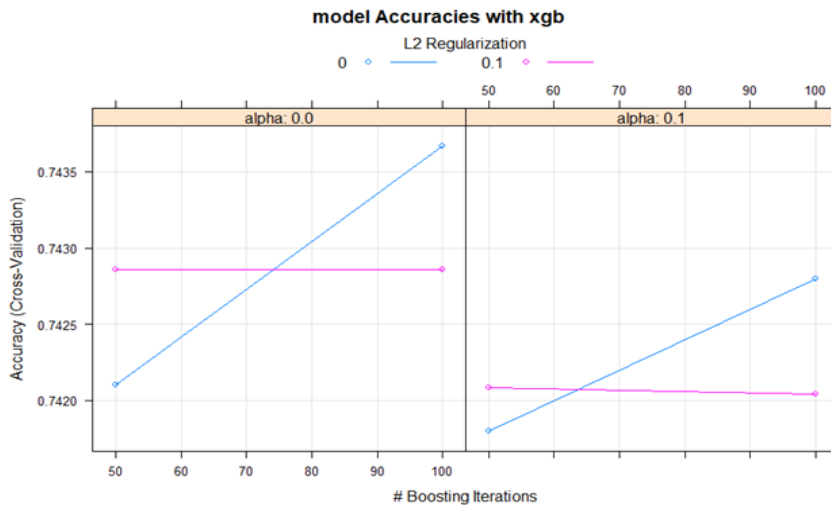
0.869의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy는 0.4078이었으며, sensitivity는 0.3466, specificity는 0.4124, Positive predicted value는 0.0424, negative predicted value는 0.3466이었다.

<표 3-10> Random forest의 model accuracy with test set

Random Forest	
AUC	0.650
Accuracy	0.4078
95% CI	(0.4049, 0.4108)
Sensitivity	0.3466
Specificity	0.4124
Pos Pred Value	0.0424
Neg Pred Value	0.3466
Balanced Accuracy	0.3795

2) eXtreme Gradient Boosting model

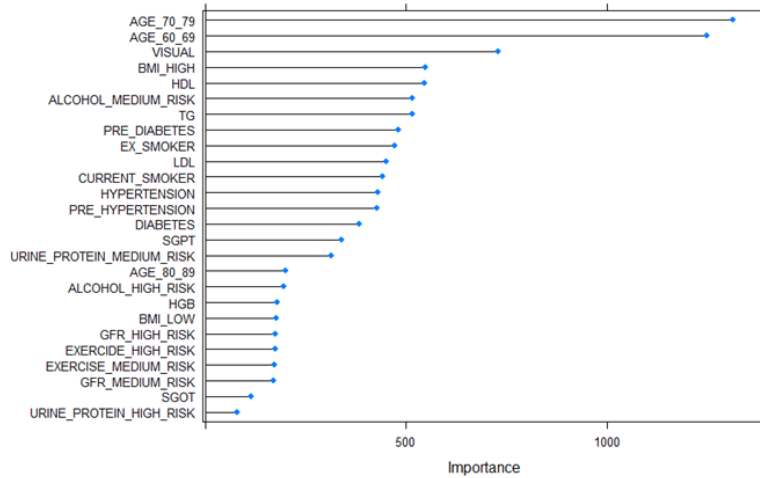
남성 extreme gradient boosting model에서는 lambda가 0, alpha가 0, nrounds가 100일 때 가장 높은 정확도를 보였으며 model accuracy는 0.7437이었다.



lambda	alpha	nrounds	Accuracy	Kappa
0	0	50	0.742101	0.459731
0	0	100	0.74367	0.464012
0	0.1	50	0.741799	0.458465
0	0.1	100	0.742801	0.461217
0.1	0	50	0.742857	0.46085
0.1	0	100	0.742857	0.461144
0.1	0.1	50	0.742082	0.459915
0.1	0.1	100	0.742045	0.460019

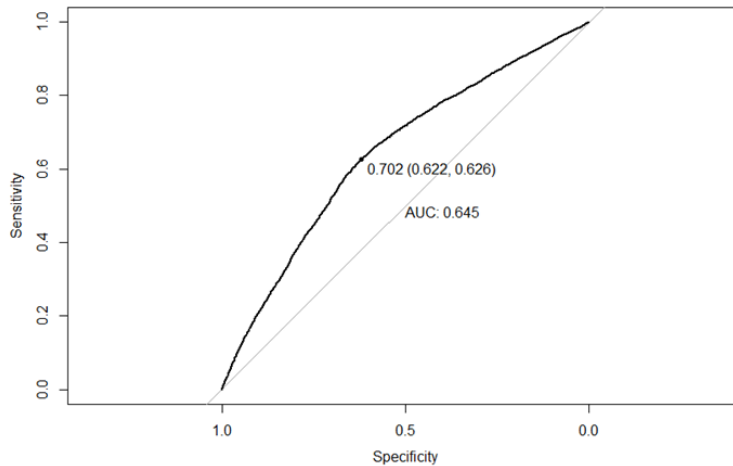
[그림 3-21] eXtreme gradient Boosting의 modeling in training set

모델을 구축할 때 사용된 Variable importance에서는 나이 60-69세가 가장 높은 비중을 차지하였고, 그 다음이 나이 70-79세, 시력 순으로 높은 비중을 보였다.



[그림 3-22] eXtreme Gradient Boosting의 variable importance

모델의 AUC 값은 0.645이었으며 확률값의 cutoff는 0.702 으로 분석되었다.



[그림 3-23] eXtreme Gradient Boosting의 cutoff point ROC curve

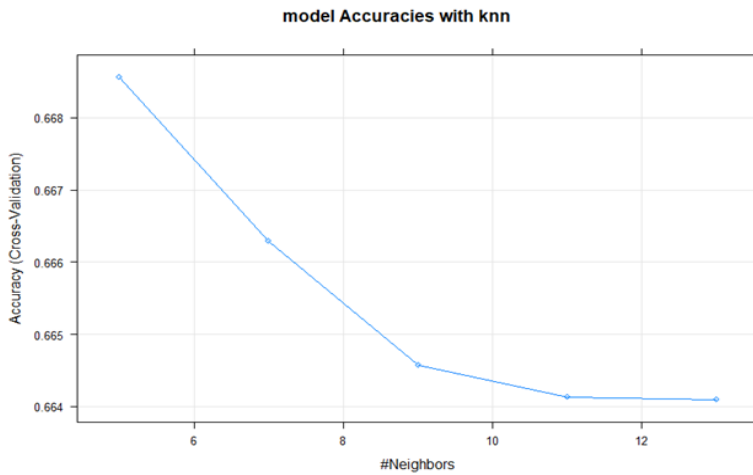
0.702의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy 는 0.3807이었으며, sensitivity는 0.3720, specificity는 0.3813, Positive predicted value는 0.0432, negative predicted value는 0.3719였다.

<표 3-11> eXtreme Gradient Boosting의 model accuracy with test set

XGB	
AUC	0.645
Accuracy	0.3807
95% CI	(0.3777, 0.3836)
Sensitivity	0.3720
Specificity	0.3813
Pos Pred Value	0.0432
Neg Pred Value	0.3719
Balanced Accuracy	0.3766

3) K-nearest neighbor model

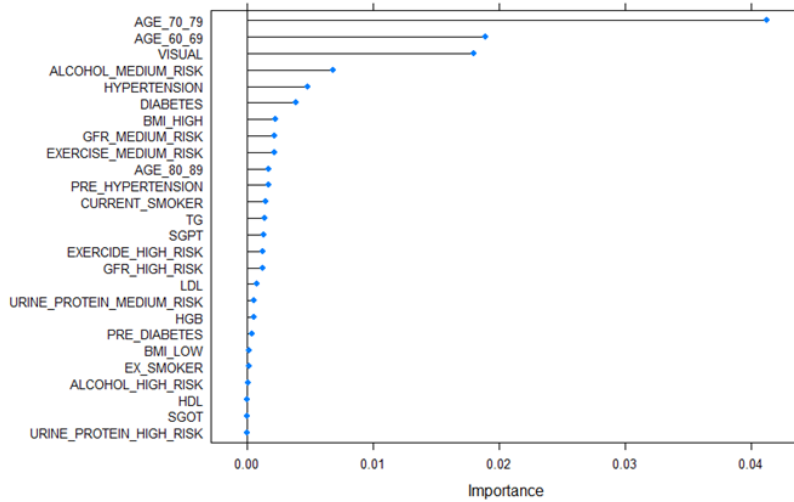
남성 K-nearest neighbor model에서는 K가 5일 때 가장 높은 정확도를 보였으며 model accuracy는 0.6686이었다.



k	Accuracy	Kappa
5	0.668556	0.306092
7	0.666289	0.300614
9	0.664569	0.296732
11	0.664135	0.294955
13	0.664097	0.294503

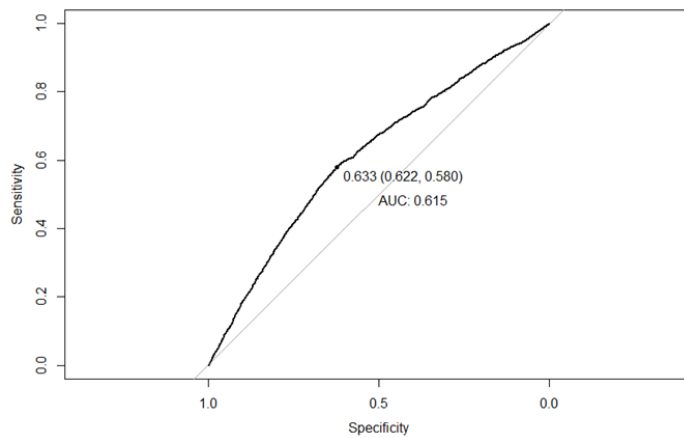
[그림 3-24] K-nearest neighbor의 modeling in training set

모델을 구축할 때 사용된 Variable importance에서는 나이 70-79세가 가장 높은 비중을 차지하였고, 그 다음이 나이 70-79세, 시력 순으로 높은 비중을 보였다.



[그림 3-25] K-nearest neighbor의 variable importance

모델의 AUC 값은 0.615이었으며 확률값의 cutoff는 0.633 으로 분석되었다.



[그림 3-26] K-nearest neighbor의 cutoff point ROC curve

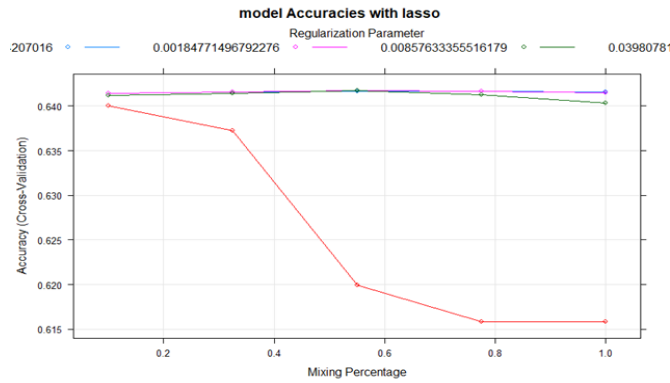
0.633의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy 는 0.3804이었으며, sensitivity는 0.4202, specificity는 0.3774, Positive predicted value는 0.0483, negative predicted value는 0.4202였다.

<표 3-12> K-nearest neighbor의 model accuracy with test set

KNN	
AUC	0.615
Accuracy	0.3804
95% CI	(0.3775, 0.3833)
Sensitivity	0.4202
Specificity	0.3774
Pos Pred Value	0.0483
Neg Pred Value	0.4202
Balanced Accuracy	0.3988

4) Lasso model

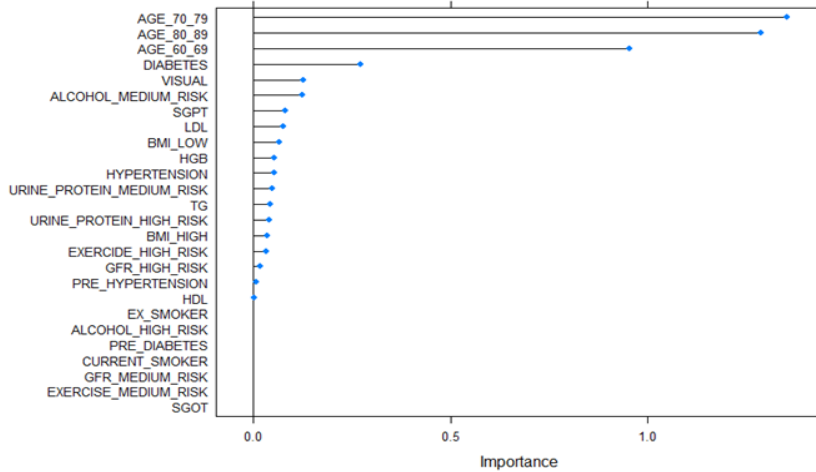
남성 Lasso model에서는 alpha가 0.55, lamda가 0.008576일 때 가장 높은 정확도를 보였으며 model accuracy는 0.64178이었다.



alpha	lambda	Accuracy	Kappa
0.1	0.000398	0.641448	0.2692452
0.1	0.001848	0.641448	0.2692452
0.1	0.008576	0.641184	0.2660260
0.1	0.039808	0.640016	0.2504044
0.325	0.000398	0.641563	0.2700428
0.325	0.001848	0.641613	0.2700373
0.325	0.008576	0.641391	0.2672390
0.325	0.039808	0.637233	0.2430391
0.55	0.000398	0.64169	0.2704315
0.55	0.001848	0.641758	0.2702603
0.55	0.008576	0.64177	0.2687425
0.55	0.039808	0.619955	0.1636291
0.775	0.000398	0.641655	0.2707774
0.775	0.001848	0.64167	0.2699532
0.775	0.008576	0.641288	0.2680907
0.775	0.039808	0.615862	0.1505227
1	0.000398	0.641594	0.2708568
1	0.001848	0.64149	0.2697869
1	0.008576	0.640338	0.2600643
1	0.039808	0.615801	0.1503707

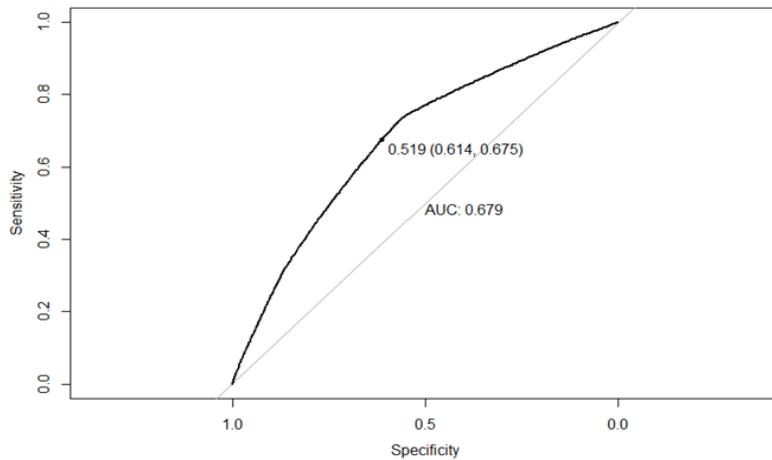
[그림 3-27] Lasso의 modeling in training set

모델을 구축할 때 사용된 Variable importance에서는 나이 70-79세가 가장 높은 비중을 차지하였고, 그 다음이 나이 80-89세, 나이 60-69세 순으로 높은 비중을 보였다.



[그림 3-28] Lasso의 variable importance

모델의 AUC 값은 0.679이었으며 확률값의 cutoff는 0.519 으로 분석되었다.



[그림 3-29] Lasso의 cutoff point ROC curve

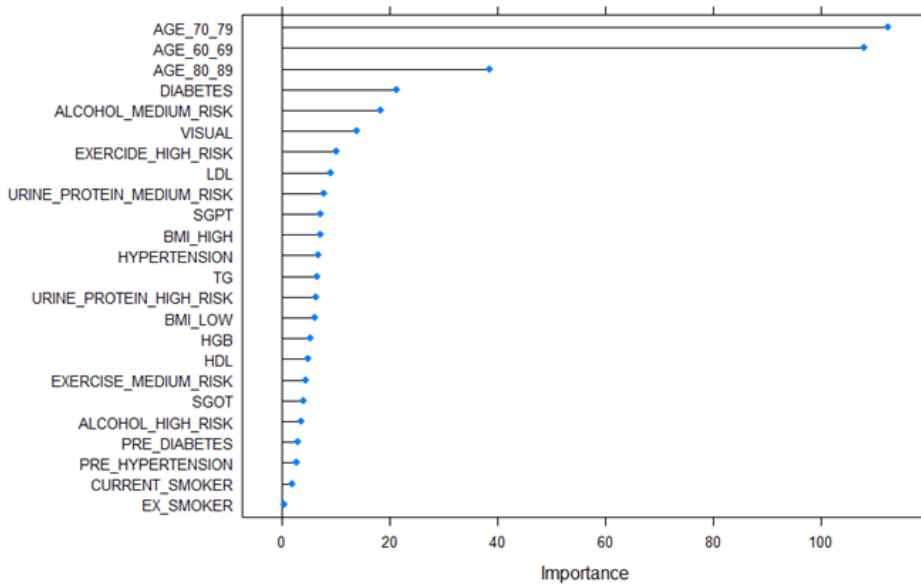
0.519의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy는 0.3819이었으며, sensitivity는 0.3251, specificity는 0.3863, Positive predicted value는 0.0388, negative predicted value는 0.8825였다.

<표 3-13> Lasso의 model accuracy with test set

LASSO	
AUC	0.679
Accuracy	0.3819
95% CI	(0.3806, 0.3832)
Sensitivity	0.3251
Specificity	0.3863
Pos Pred Value	0.0388
Neg Pred Value	0.8825
Balanced Accuracy	0.3557

5) Logistic regression

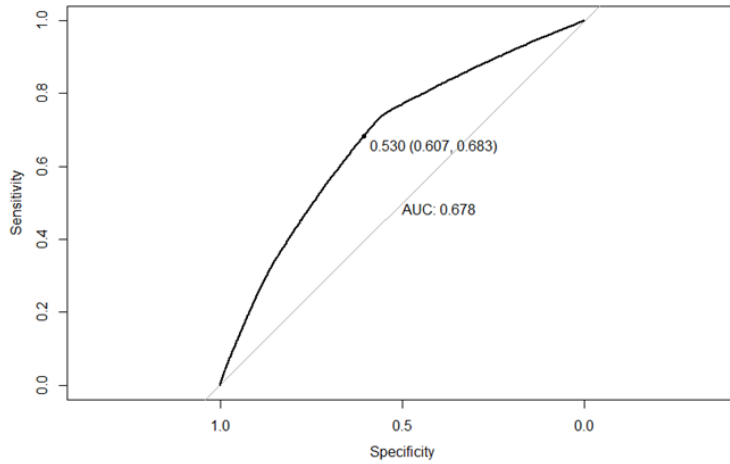
Logistic 모델을 구축할 때 사용된 Variable importance에서는 나이 70-79세가 가장 높은 비중을 차지하였고, 그 다음이 나이 80-89세, 나이 60-69세 순으로 높은 비중을 보였다.



Accuracy	Kappa
0.6414	0.2708

[그림 3-30] Logistic의 modeling in training set

모델의 AUC 값은 0.678이었으며 확률값의 cutoff는 0.530으로 분석되었다.



[그림 3-31] Logistic의 cutoff point ROC curve

0.530의 확률값 cutoff를 적용하여 모델 정확도를 평가하였을 때 model accuracy는 0.3885이었으며, sensitivity는 0.3159, specificity는 0.3940, Positive predicted value는 0.0382, negative predicted value는 0.8831이었다.

<표 3-14> Logistic의 model accuracy with test set

Logistic	
AUC	0.678
Accuracy	0.3885
95% CI	(0.3872, 0.3898)
Sensitivity	0.3159
Specificity	0.3940
Pos Pred Value	0.0382
Neg Pred Value	0.8831
Balanced Accuracy	0.3550

6) 남성의 골다공증 예측모델 비교

다섯 개의 예측모델 중 AUC가 가장 높은 모델은 Lasso(0.679)이었고 accuracy가 가장 높은 모델은 Random forest(0.4078)이었으며, positive predictive value가 가장 높은 모델은 KNN(0.0483)이었다.

<표 3-15> 남성의 골다공증 예측모델 비교

	Random forest	XGB	KNN	Lasso	Logistic
AUC	0.650	0.645	0.615	0.679	0.678
Accuracy	0.4078	0.3807	0.3804	0.3819	0.3885
95% CI	(0.4049, 0.4108)	(0.3777, 0.3836)	(0.3775, 0.3833)	(0.3806, 0.3832)	(0.3872, 0.3898)
Sensitivity	0.3466	0.3720	0.4202	0.3251	0.3159
Specificity	0.4124	0.3813	0.3774	0.3863	0.3940
Pos Pred Value	0.0424	0.0432	0.0483	0.0388	0.0382
Neg Pred Value	0.3466	0.3719	0.4202	0.8825	0.8831
Balanced Accuracy	0.3795	0.3766	0.3988	0.3557	0.3550

제5절 결론 및 제언

1. 결론

일반검진 수검자를 대상으로 위험인자 및 생활습관과 골다공증과의 관련성을 분석한 결과 여성의 경우 나이에 따라 가장 큰 격차를 보였으며, 당뇨병인 경우(RH 1.433), 저체중인 경우(HR 1.088), 폐경인 경우(HR 1.124) 위험도가 높게 나타났다. 남성의 경우에도 나이에 따라 가장 큰 격차를 보였으며, 당뇨병인 경우(HR 1.55), HGB 위험군(HR1.329), 저체중인 경우(HR 1.225) 위험도가 높게 나타났다.

65세 미만에서는 여성의 경우 나이에 따라 가장 큰 격차를 보였으며, 고혈압(HR 1.16), 당뇨병(HR 1.445), GFR 고위험군(HR 1.205), 저체중(HR 1.13), 폐경인 경우(HR 1.234) 위험도가 높게 나타났다. 남성의 경우에도 나이에 따라 가장 큰 격차를 보였으며, 고혈압(HR 1.16), 당뇨병(HR 1.6), GFR 위험군(HR 1.38), HGB 위험군(HR 1.51), 저체중인 경우(HR 1.3) 위험도가 높게 나타났다.

65세 이상에서는 여성의 경우 나이에 따라 가장 큰 격차를 보였으며, 당뇨병(HR 1.42)인 경우 위험도가 높게 나타났다. 남성의 경우에도 나이에 따라 가장 큰 격차를 보였으며, 당뇨병(HR 1.447), HGB 위험군(HR 1.278), 저체중(HR 1.148)에서 위험도가 높게 나타났다.

머신러닝(Machine learning)을 이용한 예측모델은 Random forest model, eXtreme Gradient Boosting model, Lasso model, K-nearest neighbor model을 이용하여 분석하였고, 기존 통계방법인 logistic regression을 이용한 예측모델과 비교 분석하였다.

여성의 경우 다섯 개의 예측모델 중 AUC가 가장 높은 모델은 KNN(0.621)이었고

accuracy가 가장 높은 모델은 Lasso(0.5973) 이었으며, positive predictive value가 가장 높은 모델은 Lasso(0.4745)였다.

남성의 경우 다섯 개의 예측모델 중 AUC (0.615-0.679)가 가장 높은 모델은 Lasso(0.679)이었고 Accuracy (0.3804-0.47078)가 가장 높은 모델은 Random forest (0.4078) 이었으며, positive predictive value (0.0382-0.0487)가 가장 높은 모델은 KNN(0.0483)로 매우 낮게 나타났다.

2. 제언

일반검진 수검자에서 남녀 공통적으로 연령에 따라, 당뇨병이 있는 경우, 저체중인 경우, 또한 여성의 경우 폐경인 경우 발생률과 위험도가 높아 이전 연구들과 비슷한 결과를 보였다. 이러한 위험군에 대하여 골다공증 발생 예방을 위해 생활습관 관리 및 조기 진단과 치료가 필요할 것으로 판단된다.

골다공증 machine learning 예측모델에서 여성의 예측 모델 중 AUC와 accuracy, positive predictive value가 높게 분석된 모델은 K-nearest neighbor model과 Lasso model로, 추후 보완하여 이들 모델을 이용하여 골다공증 예측에 활용할 수 있을 것으로 생각된다. 예측도가 낮게 분석된 것은 골다공증 위험도와 관련된 항목이 아닌 검진데이터 항목으로 모델을 구축하였기 때문인 것으로 판단된다. 남성에서도 모든 예측 모델에서 AUC (0.615-0.679)와 Accuracy (0.3804-0.47078), positive predictive value (0.0382-0.0487)가 매우 낮게 분석되었다. 다만 분석 환경의 한계로 인해 일부 모델은 전수 데이터로 구축되었고, 일부 모델은 표본추출에 의한 데이터로 구축되었기 때문에, 직접적인 비교에 주의가 필요하다. 또한, 대부분의 모델에서 accuracy의 95% CI 값들을 보게 되면 비슷한 구간이 많아, 실질적으로 어떤 모델이 더 뛰어나다고 정확히 말할 수 없다.

추후 국가건강검진 데이터를 이용한 골다공증 예측모델 개발에서는, 건강검진 항목에 관련 factor가 많이 포함되도록 설문 추가 등의 단계적 예측 모델을 구축해야 할 것으로 판단된다.

제4장

골다공증 환자 교육 처방 메시지

제1절 서론	93
제2절 이론적 고찰	93
제3절 연구 내용 및 방법	95
제4절 분석 결과	98
제5절 결론 및 제언	101

제4장

골다공증 환자 교육 처방 메시지

제1절 서론

1. 연구 배경 및 필요성

건강메세지란, 건강정보를 제공해 개인과 공동체가 건강을 관리할 수 있도록 하는 메시지이다. 이는 건강문제에 대한 인식을 높이고, 고정관념이나 편견을 줄이며, 태도 및 행동 변화를 통한 건강증진을 하는데 목적이 있다. 또한 금연, 절주, 균형 잡힌 영양 섭취, 적절한 운동 등의 행동변화를 촉구하는 커뮤니케이션을 통하여 질환을 예방하고, 건강을 지키는데 에 목적이 있다.

골다공증 예측 모델을 개발하고, 이 예측모델을 통하여, 수검자들의 건강행동양식에 변화를 주고자 함에는 교육용 처방 메시지가 어떻게 전달되는 지에 따라서 예측모델의 활용도는 높일 수 있을 것이다. 즉, 개발된 골다공증 예측 모델의 가치를 잘 활용하기 위해서는 이 예측 모델을 활용한 환자용 (또는, 이용자용) 건강 메시지의 개발이 필요할 것으로 사료된다.

2. 연구의 목적

개발된 골다공증 예측 모델을 활용하여, 건강IN 웹페이지에 탑재할 수 있는 골다공증 환자용 처방 메시지를 개발하는데 소연구 목적이 있다.

제2절 이론적 고찰

1. 골다공증 환자 교육 처방 메시지

건강메세지란, 건강정보를 제공해 개인과 공동체가 건강을 관리할 수 있도록 하는 메시지이다. 이는 건강문제에 대한 인식을 높이고, 고정관념이나 편견을 줄이며, 태도 및 행동 변화를 통한 건강증진을 하는데 목적이 있다. 또한 금연, 절주, 균형 잡힌 영양 섭취, 적절한 운동 등의 행동변화를 촉구하는 커뮤니케이션을 통하여 질환을 예방하고,

건강을 지키는데 에 목적이 있다.

또한 공공기관은 공적인 건강정보를 통하여 다양한 건강 메시지를 국민에게 제공해야 할 의무가 있다. 다양한 공적인 건강정보들은 암, 감염병, 희귀난치성 질환 등의 질병정보, 국가 건강검진의 결과에 관한 정보, 의약품에 관한 정보, 식생활 및 운동에 관한 생활습관 정보, 예방접종에 관한 정보, 해외여행에 관한 건강정보 등이 포함 될 것이다. 공공기관에서 웹페이지를 통하여 제공하고 있는 건강정보의 예시는 <표 4-1>과 같다.

<표 4-1> 건강메시지 제공 공공기관 웹페이지

기관	주소	내용
국가암정보 센터	https://www.cancer.go.kr/	암예방 정보, 암의 종류, 진단, 치료, 재발 및 전이, 암환자 생활, 암환자 증상관리, 암환자 식생활, 암 관련 통계등 제공
질병관리청	https://www.kdca.go.kr/	건강위해정보, 희귀질환정보, 건강관련 통계 제공
식품안전 나라	https://www.foodsafetykorea.go.kr/main.do	식품안전지식, 건강기능식품, 식품영양정보 등 안전하고 건강한 먹거리 이용을 위한 정보 제공
예방접종도우미	https://nip.kdca.go.kr/irgd/index.html	예방접종 길잡이, 예방접종 대상 감염병, 국가예방접종 사업 등 예방접종 관련 정보 제공
금연길라잡이	https://www.nosmokeguide.go.kr/index.do	금연 프로그램 및 금연관련 상담관련, 관련 통계 등 금연관련 건강정보 제공
국민건강정보포털	https://health.kdca.go.kr/	다양한 질환에 대한 건강정보, 심뇌혈관질환정보, 의료기관찾기 서비스, 만성질환 건강통계등 정보 제공

건강메시지가 효과적으로 이용자 또는 환자에게 전달되려면, 몇가지 고려사항이 있을 수 있다. 첫째는, 전달대상에 대한 고려이다. 건강메시지는 의료인 대상이라기보다 일반인 대상으로 하기 때문에, 일반인들이 이해하기 쉬운 내용이 되어야 한다. 이를 위해서는 작성된 메시지가 대부분의 사람들이 이해하기 쉬운 단어, 용어를 사용 하였는가? 의학 전문용어는 피하였는가? 대화체로 쉽고, 짧은 단문으로 되었는가? 주제가 명확한가? 에 대한 고려가 되어야 할 것이다. 두 번째는 전달내용에 대한 정확성에 대한 고려가 필요하다. 제공되는 정보가 의학적으로 신뢰할 수 있는 명확한 근거를 가진 자료인지 평가가 필요하고, 그 방법은 신뢰할 수 있는 정보 출처와 분야별 전문가의 의견 수렴이 필요할 수 있다. 마지막으로 전달 방식에 대한 고려가 필요하다. 자세한 정보 제공보다는 명료한 제시가 대상자에게 건강 행동변화를 일으키기 유리하며, 내용이 많은 정리본 보다는 리플릿이나 이미지 등을 활용하여 중요내용 위주로의 요약하는 것이 더 효과적일 수 있다. 또한 작성된 메시지가 문화적 가치에 기초하였는지, 삽화와 메시지가 문화를

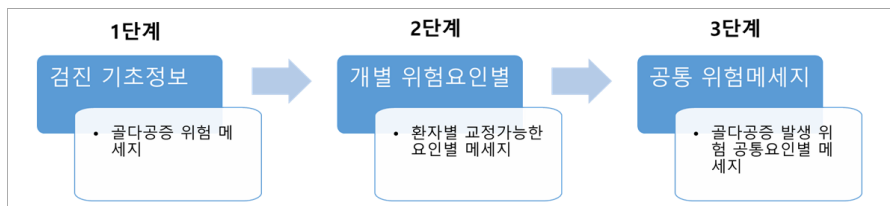
존중하면서 흥미를 유발시킬 수 있는지에 대한 고려도 필요하다.

이러한 이론적 배경으로, 골다공증 환자 교육 처방메시지 작성을 위해서는 학회 및 공공기관 등에서 기존 작성된 환자 교육용 자료를 참고 및 활용하는 것이 효율적으로 사료된다. 이에 관한 자료로는 국내에서는 국민건강정보포털, 질병관리청 및 보건소의 골다공증 환자 교육용 자료, 한국 골대사 학회에서 발간한 소책자 등이 있고, 해외자료는 세계보건기구, 세계 골다공증 재단 (International Osteoporosis Foundation, IOF), 미국 국립 골다공증 협회(National Osteoporosis Foundation)등이 있다.¹⁻⁷

제3절 연구 내용 및 방법

1. 연구 내용

개발된 10년 골다공증 예측모델에 근거하여, 예측모델의 추정된 모형식에 포함된 변수를 참조하여 처방메시지를 작성하였다. 머신러닝을 통한 예측모델 투입변수로는 성별, 나이, BMI, 허리둘레, 혈압, 공복혈당수치, 청력이상여부, 시력, Total cholesterol, TG, LDL, HDL, GFR, HGB, GGT, SGPT, SGOT, Creatinine, 요단백, 흡연, 음주, 운동, 폐경여부, 초경나이 등이 있었고, Cox proportional model을 통한 예측모델에는 나이, 흡연여부, 운동, 음주, 고혈압여부, 당뇨병여부, GFR, 요단백검사, 시력, TG, LDL, HGB, GOT, GPT, BMI, 허리둘레, 초경나이, 폐경여부가 포함되었다. 환자 교육 처방메시지를 작성 시, 예측모델 모형식에 고려된 골다공증 위험 변수들을 고려하였다. 골다공증 위험 처방메시지는 Flow model을 기반으로 단계별 처방 메시지를 제공하는 것으로 작성하였다. 건강 iN 웹페이지에 탑재 가능성을 고려하여, 환자(또는 이용자)의 국가 검진 결과 정보를 불러오고, 이 정보를 토대로 골다공증 발생 위험 메시지를 제공하고, 이후 위험요인 중에 교정 가능한 요인들은 해당 요인별 메시지를 제공하고, 이후에 추가적인 골다공증 공통적인 위험 메시지를 제공하는 것으로 계획하였다(그림 4-1).

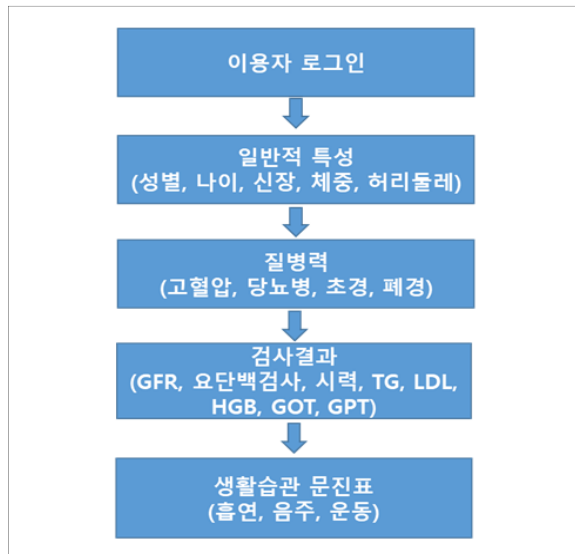


[그림 4-1] 골다공증 환자 교육 처방메시지 제공 단계

2. 연구 방법

1) 골다공증 발생 위험 메시지

먼저, 개발된 골다공증 예측모델에 포함된 변수로는 Cox proportional model의 경우, 나이, 흡연여부, 운동, 음주, 고혈압여부, 당뇨병여부, GFR, 요단백검사, 시력, TG, LDL, HGB, GOT, GPT, BMI, 허리둘레, 초경나이, 폐경여부를 고려되었고, 이는 환자(또는 이용자)의 국가 검진 결과 정보를 불러오는 것을 가정으로 골다공증 발생 위험 메시지를 제공하는 것으로 개발되었다. (그림 4-2) 이러한 스크래핑된 데이터를 기반으로, 골다공증 위험 메시지를 작성하였다. 골다공증 위험 메시지는 “귀하의 10년 이내 골다공증 발생 위험도는 (매우낮음), (낮음), (보통), (높음), (매우높음) 수준이며, 10년 골다공증 발생 위험률은 ___%입니다.” 라는 전달방식으로 환자별 발생 위험률이 얼마 정도인지 인지하기 쉽게 작성하였다. 골다공증 예측 모형식 결과를 참고할 때, 남자의 경우 전체 연구집단에서 10년 발생률은 7.05%였고, 여자의 경우 38.2%였고, 이러한 예측 모델 결과를 토대로, 연구자들의 논의 후, (표 4-2)처럼 위험도 구간을 설정하였다.



[그림 4-2] 처방 메시지 고려 위험 요인

<표 4-2> 10년 이내 골다공증 발생 위험 예측도 분류

위험도 분류	남성	여성
매우낮음	<1%	<5%
낮음	1-<2%	5-<10%
보통	2-<4%	10-<21%
높음	4-<8%	21-<40%
매우 높음	8%~	40%~

골다공증 위험 메시지 제공이후, 동일연령 대비 위험도 메시지를 제공하는 경우는 환자(또는 이용자)의 해당연령을 위험도 예측모형식에 넣고 나머지 위험요인이 평균으로 계산하면, 동일연령의 위험도를 산출할 수 있다. 산출가능하다는 전제하여, 동일연령 위험률을 제공하고, 환자(또는 이용자)의 골다공증 위험률이 동일 연령 위험률 대비 낮은지, 높은지 그 수준에 대하여 메시지를 작성하였다.

2) 골다공증 위험 교정 가능한 요인별 메시지

개발된 골다공증 예측 모형식에서, 위험 교정 가능한 요인은 흡연, 운동, 음주, BMI (허리둘레) 으로 논의 되었다. 교정 가능한 요인들은 환자(또는 이용자)의 생활 습관의 변화를 유도하는 메시지이므로, 검사실 결과들 (GFR, 요단백검사, 시력, TG, LDL, HGB, GOT, GPT, BMI) 및 교정 불가능한 요인 (나이, 질환력, 초경나이, 폐경 여부)는 교정 가능한 요인별 메시지를 고려하지 않았다. 개발된 골다공증 예측 Cox 모델 분석에서, 흡연은 비흡연, 과거흡연, 흡연으로 분류가 되었고, 운동은 저활동, 중등도활동, 고활동으로, 음주는 고위험, 중위험, 저위험으로, BMI는 낮음, 정상, 비만으로 분류가 되어 교정 가능한 요인들은 3분류로 구분되었다. 이러한 3단계 분류된 결과를 바탕으로, 교정 가능한 요인들은 (정상), (주위), (위험) 구간을 설정이 가능하였으며, Cox 모델 분석 결과상 골다공증 발생 위험은 흡연군에서, 저활동군에서, 저위험음주군에서, BMI가 낮은 군에서 그 위험이 높았다. 음주의 경우, 오히려 저위험음주군에서 골다공증 발생위험이 높았지만, 이는 국가검진 수검자의 음주 설문은 자가설문이므로, bias가 존재하기 때문에, 요인별 메시지 제공시 음주항목은 제외하는 것으로 논의되었다. 또한 허리둘레의 경우 BMI와 중복되는 요인이므로 요인별 메시지 제공시 제외하는 것으로 논의하였다. (표 4-3)

<표 4-3> 골다공증 발생 요인별 위험메시지 시, 고려한 요인

요인	분류구간	권장수치
BMI	3분류	>18.5
흡연	3분류	비흡연
운동	3분류	고활동

3) 골다공증 위험 공통 요인별 메시지

환자(또는 이용자)가 위험 교정 가능한 요인별 메시지 제공으로는 행동변화를 일으키지 않을 수 있으므로, 조금 더 행동변화를 유발할 수 있는 메시지 제공을 계획하였다. 바꿀 수 없는 (교정 불가능한) 요인 및 일반적인 골다공증 발생 위험으로 알려진 요인에 대하여 메시지를 작성하였다. 고려한 공통 위험 요인으로는 나이, 골절 과거력, 골절 가족력, 폐경여부, 스테로이드 치료여부 등을 고려하였다. 제공 방식은 일반인들이 이해하기 쉬운 내용으로, 대부분의 사람들이 이해하기 쉬운 단어, 용어를 사용 하였다. 또한 제공되는 정보가 의학적으로 신뢰할 수 있는 명확한 근거를 위하여 보건복지부, 대한 골대사학회 및 세계 골다공증 재단 (International Osteoporosis Foundation, IOF)에서 공개한 골다공증 관련 환자 교육용 자료들을 참고하였다.¹⁻⁷ 또한, 자세한 정보 제공보다는 명료한 제시가 대상자에게 건강 행동변화를 일으키기 유리하므로, 간결한 내용으로 작성하였다. (표 4-4)

<표 4-4> 골다공증 위험 공통 요인 메시지 시, 고려한 요인

요인
나이, 골절 과거력, 골절 가족력, 폐경여부, 스테로이드 치료여부

제4절 분석 결과

1. 골다공증 발생 위험 메시지

앞서, 환자 교육 처방메시지 연구방법에서 설명한바와 같이, 골다공증 발생 위험 메시지는 표 4-5와 같이 작성하였다. 개별 환자(또는 이용자)별로 국가건강검진 자료를 기반으로 계산한 10년 골다공증 발생위험도를 제공하고, 이후 동일연령대의 골다공증 위험도와 비교하여 어떠한 수준인지 나타내는 메시지로 계획하였다.

<표 4-5> 골다공증 발생 위험 메시지

내용
1. 귀하의 10년 골다공증 발생 위험도 (보통), (약간높은), (높은), (매우높은), (위험), (매우위험) 수준이며, 골다공증 발생 확률은 ___%입니다.
2. 동일연령대의 골다공증 발생위험도는 ___%입니다.
3. 귀하의 골다공증 발생 위험도는 동일연령 평균과 비교하여 (낮은), (비슷한), (높은) 수준입니다.

2. 골다공증 위험 교정 가능한 요인별 메시지

개발된 골다공증 예측 모형식에서, 위험 교정 가능한 요인은 흡연, 운동, 음주, BMI (허리둘레) 으로 논의 되었다. 3단계 분류된 교정 가능한 요인들은 건강신호등상 정상, 주의, 위험으로 구분이 가능하였다. (표 4-6) 요인별 메시지는 위험에 해당하는 요인별 분류 시 (BMI : 낮음, 흡연 : 현재흡연, 운동 : 중등도 및 저활동) 메시지를 제공하는 것으로 계획하였다. 요인별 메시지는 보건복지부, 대한 골대사학회 및 세계 골다공증 재단에서 공개한 골다공증 관련 환자 교육용 자료들을 참고하였다.¹⁻⁷ (표 4-7)

<표 4-6> 골다공증 위험 교정 가능한 요인의 건강신호등

요인	분류구간	건강신호등
BMI	낮음	위험
	정상	정상
	비만	주의
흡연	흡연	위험
	과거흡연	주의
	비흡연	정상
운동	고활동	정상
	중등도 활동	위험
	저활동	위험

<표 4-7> 골다공증 위험 교정 가능한 요인별 메시지

요인	분류구간	메시지
BMI	낮음	체질량지수가 낮으면 골다공증 위험이 증가하는 것으로 알려져 있습니다. 적정 체중을 유지하기 위하여 충분하고 골고른 영양소를 섭취하는 것을 권장드립니다.
흡연	흡연	흡연은 뼈 건강에 안 좋은 영향을 미치며, 장에서 칼슘 섭취를 감소시키고 칼슘대사에 영향을 주는 호르몬 농도를 변화시키는 것으로 골다공증 위험 발생이 증가 되는 것으로 알려져 있습니다. 흡연은 뼈 건강 뿐 아니라, 건강에 해로우므로 반드시 금연을 권장드립니다.
운동	중등도 활동, 저활동	신체활동 부족은 뼈와 근육 손실을 유발합니다. 규칙적으로 체중이 실리는 운동(달리기, 걷기, 줄넘기 등)과 근력운동은 뼈와 근육을 단련하고 강화하는 데 도움이 됩니다. 일정한 육체적 활동을 지속하면 노화에 의한 골소실이 감소됩니다. 노인에게도 운동을 통한 체력단련으로 근육의 무게를 증가시키면 골소실을 감소시킬 수 있습니다. 골량을 유지하기 위한 효과적인 스트레칭도 좋은 체중부하 운동으로 알려져 있습니다.

3. 골다공증 위험 공통 요인별 메시지

바꿀 수 없는 (교정 불가능한) 요인 및 일반적인 골다공증 발생 위험으로 알려진 요인에 대한 메시지를 작성하였다. 이를 골다공증 위험 공통 요인별 메시지로 안내하였고, 고려한 공통 위험 요인으로는 나이, 골절 과거력, 골절 가족력, 폐경여부, 스테로이드 치료여부 등에 대하여 메시지를 작성하였다. (표 4-8) 위험요인별 메시지와 같이, 공통 요인별 메시지도 보건복지부, 대한 골대사학회 및 세계 골다공증 재단에서 공개한 골다공증 관련 환자 교육용 자료들을 참고하였다.¹⁻⁷

<표 4-8> 골다공증 위험 공통 요인별 메시지

요인	메시지
나이	뼈는 일생동안 지속적으로 변화합니다. 일생동안 뼈의 양이 최대가 되는 시기는 20세~30세 경입니다. 이후 조금씩 감소하다 50세 이후 급격히 감소합니다. 따라서, 골다공증 위험은 나이에 따라 증가합니다. 특별한 임상적 위험 요인이 없더라도, 60세 이상은 다음 건강검진에서 뼈 건강에 대하여 상담해야 합니다.
골절 과거력	가벼운 낙상으로 뼈가 부러진 적이 있다면 골다공증의 원인될 수 있습니다. 한번 부러진 뼈는 새로운 골질의 주요한 위험요인이며, 골다공증 골절 이후 재골절 위험도는 2~10배로 증가합니다. 일단 척추 골절이 발생하게 되면 5명중 1명에서는 1년 이내에 또 다른 척추골절이 발생하며, 향후 대퇴골 골절의 발생은 3~4.5배 증가합니다. 따라서 골다공증 골절이 발생하면 적극적인 추적관찰이 필요합니다.
골절 가족력	부모가 대퇴골 골절을 겪은 경우 그 자녀는 골다공증 발병과 골절 위험이 클 수 있습니다. 골다공증은 여러 가지 인자가 복합적으로 작용해서 발생하나, 유전적 요인은 최대골량을 획득하는데 70%가량 영향을 주는 것으로 알려져 있습니다. 따라서 부모 중 등이 굽은 사람이 있거나 골다공증 가족력이 있다면, 일찍부터 뼈 건강에 관심을 가지는 것이 좋습니다.
폐경 여부	에스트로겐(여성호르몬)은 뼈를 보호하는 효과가 있습니다. 여성의 경우 폐경이후 급격한 뼈의 손실이 나타나며 일생 중 폐경후 첫 5년동안 뼈의 손실이 가장 많습니다. 폐경후 골다공증 환자의 50% 이상에서 일생동안 골절을 한번 이상 경험하며 뼈가 잘 부러지는 부위는 주로 척추입니다. 따라서 폐경이 된 여성은 골다공증 위험이 커지므로 골다공증 검사가 필요할 수 있습니다.
스테로이드 치료 여부	스테로이드는 천식, 류마티스관절염 및 일부 염증성 질환등의 질환에 흔히 처방됩니다. 3개월 연속 스테로이드를 사용하면 뼈 손실로 이어질 수 있고, 첫 3-6개월 치료기간동안 손실속도가 가장 빠릅니다. 따라서 스테로이드 약물을 투여중인 경우에는 뼈 건강에 대하여 상담해야 합니다.

제5절 결론 및 제언

1. 결론

골다공증 예측 모델을 이용하여, 환자(또는 이용자)들의 건강행동양식에 변화를 주고자 하는 목적으로, 골다공증 환자 교육 처방메시지를 개발하였다. 개발된 환자 교육 처방메시지는 단계별 메시지 형식으로 작성되었고, 건강 IN 웹페이지에 탑재 가능성을 고려하여, 환자(또는 이용자)의 국가 검진 결과 정보를 불러오고, 이 정보를 토대로 골다공증 발생 위험 메시지를 제공하고, 동시에 동일연령 위험메시지를 작성하였다. 다음

단계는 환자별 스크래핑된 건강검진정보 중 위험요인 중에 교정 가능한 요인(BMI, 흡연, 운동)들에 대한 해당 요인별 메시지를 작성하였다. 마지막 단계는 행동변화를 조금 더 유발할 수 있게, 바꿀 수 없는 (교정 불가능한) 요인 및 일반적인 골다공증 발생 위험으로 알려진 요인에 대하여 메시지를 작성하였다.

2. 제언

골다공증 환자 교육 처방 메시지를 개발 하였지만, 실제 건강iN에 탑재하기 위해서는 구체적인 탑재방안이 있어야 할 것으로 사료된다. 또한 교육용 처방 메시지가 수검자들에게 효과적 전달되기 위해서 웹페이지 탑재 방식에 대한 충분한 논의가 필요할 수 있다. 이를 통하여 궁극적으로는 환자 또는 이용자들이 본인들의 골다공증 위험도를 파악하고, 골다공증 발생 위험을 낮출 수 있게 건강생활습관을 유지하도록 활용할 수 있을 것으로 사료된다.

제5장

계층적 질환군(NHIS-HCC) 위험조정 모델을 통한 개인별 건강위험 평가

제1절 서론	105
제2절 이론적 고찰	106
제3절 연구 내용 및 방법	109
제4절 분석 결과	112
제5절 결론 및 제언	115

제5장

계층적 질환군(NHIS-HCC) 위험조정 모델을 통한 개인별 건강위험 평가

제1절 서론

1. 연구배경 및 필요성

노인 인구의 증가는 의료비 증가를 유발하고 있다. 또한 첨단 의료기기와 기술의 발달로 질환의 진단과 치료에 드는 의료비용은 점차 증가하고 있다. 이러한 의료비 증가는 건강보험 등 보건의료 보장체제를 위협하는 문제로 인식되고 있다.

한국의 노인인구의 비중이 20%인 초고령 사회로의 진입은 2026년일 것으로 전망되며, 2060년에는 노인인구 비율이 40%를 초과할 것으로 예측된다.¹ 주요 선진국 등에서 노인 인구 증가는 공통적으로 나타나는 현상이지만, 특히 우리나라의 경우 그 속도가 빠르다. 이유로는 급속히 진전되는 저출산·고령화 등이 있겠다.

우리나라의 경상의료비는 1980년 이후 지속적으로 증가하고 있다. 2017년 7.6%로 지속적인 증가를 보인다. 이와 같이 노인 인구가 급속히 증가하며, GDP 대비 경상의료비 비율이 증가할 경우 건강보험 재정이 점차 악화되는 문제에 직면할 수 있다. 의료비 증가는 현재 세계 각국에서 우려하고 있는 문제이다. 그래서 증가하는 의료비를 적절하게 관리하기 위한 보건의료 개혁은 여러 나라에서 시도되고 있다. 현재 시도되고 있는 방식은 의료보험 관리 방식이다. 의료보험 관리는 의료보험 가입자들 중에서 고의료비용 환자들이나 의료 이용도가 높은 과다의료 이용자들에게 대한 중재를 목표로 한다.

최근 각국에서는 의료비 관리를 위해 위험 조정 개념을 도입하고 있다. 위험조정은 의료보험 가입자의 건강상태를 반영하여 의료보험 지불계획, 공급자들에 대한 지불계획, 가입자 개인 또는 집단의 보험료를 조절하는 과정이며² 다양한 위험수준을 반영하는 보정된 지불제도이다.

과거 위험조정은 성별이나 연령에 따른 인구학적 요인만을 근거로 환자 군을 구분하여 적용하였다. 추가로 나온 모델들에서는 주 질환 군을 위험조정 모델에 추가하였다. 최근

의 미국의 CMS-HCC에서는 단일 질환 군에 의한 분류를 넘어서서, 계층화된 질환 군을 추가하였다. 이는 더 높은 단위의 모델이라 할 수 있겠다.

계층적 질환군에 대한 우리나라의 연구는 아직 부족한 상황이다. 한기명^{3,4} 전기홍⁵ 등의 연구에서 국민건강보험 표본 코호트 자료를 이용하여 계층적 질환 군의 부분적 분석이 이루어 졌으나, 김윤⁶ 등은 HCC 계수 산출을 위해 전체 인구 집단을 대상으로 한 연구가 필요함을 제시하기도 하였다. 또한 일산병원 이상현 등의 연구에서는 2016년과 2017년 국민건강보험을 이용한 연 2천3백만명 이상의 전수 자료와 2006년부터 2015년까지 10년도 국민건강보험 코호트 자료를 이용하여 계층적 질환 군 위험조정 모델을 기반으로 의료비 예측과 사망률 분석을 하기도 하였다.⁷

NHIS-HCC모형의 실제 활용에 대한 검증이 필요하여, 새로 구축된 2019년 표본코호트 자료로 바탕으로 의료비를 예측하고 이를 실제 의료비와 비교하여 모형을 재평가하고자 하였다.

2. 연구의 목적

국민건강보험 일산병원 이상현 연구팀에서는 2006년부터 2015년까지 10년도 국민건강보험 코호트 자료를 이용하여 구축된 NHIS-HCC모형을 구축하였다.² 기 연구를 기반으로 하여, 2019년 최근에 구축된 표본 코호트 자료를 이용하여, 의료비를 예측하고 실제 의료비와 비교하여 모형의 실용성을 검증하는 것이 이 연구의 목적이다. HCC 위험지표를 근거로 계층적 질환군별 2019년 의료비를 예측하고 이를 실제 사용된 의료비와 비교하고자 하였다.

제2절 이론적 고찰

1. 위험조정 모델

위험조정이라는 용어에서 위험은 의료비용과 연관된다. 의료비용이 더 부담이 될 것으로 예상되는 즉 건강상태가 좋지 않은 사람에게 지불을 더해야 하는 위험을 의미한다.

‘위험조정’은 의료비용에 대한 위험을 반영한 지불체도로 보험자에게 공정한 보상하는 것을 목적으로 한다. 위험조정은 지불 계획을 설정하고 의료보험 가입자 개인 또는 집단의 보험료를 조정하는 것이다.²

이상적인 위험조정은 의료공급자들에게 공평한 보상을 하고, 가입자들이 가입할 보험을 선택하는 등 기회를 줄 수 있게 된다.^{8,9}

위험조정에는 위험평가가 필요하다. 위험평가에는 우선 인구학적 요인이 포함된다. 인구학적 요인으로는 흔히 사용하는 성별 소득 등이 있다. 다음은 건강에 대한 내용이 들어가게 된다. 건강정보에는 질환명이 주요 요인으로 들어가며 투약 내역, 사망 등이 포함된다. 건강보험자료에는 입원이나 외래 환자 의료보험 청구자료, 약처방 청구자료가 이용된다.

위험 조정은 과거에는 인구 성별 요인만 포함하였다. 하지만 최근 건강정보를 첨가하여 예측력을 향상 시켰다.¹⁰ 건강과 관련된 요소는 의무기록의 주진단에 따른 질환별 요인이 들어가고 특히 질환을 계층화 시킨 계층적 질환군 모델이 포함된다.

위험평가를 통해 인구학적 요인과 건강상태 요인에 따른 위험지수가 산출되면, 나이와 성별 인자 그리고 질환 보유 정도 인자에 따라 각 개인이 특정 위험 점수를 가지게 된다. 이 위험점수를 통해 여러 집단을 비교 할 수 있고 의료비를 균등하게 분배할 수 있게 된다.⁷

2. 계층적 질환군 위험조정 모델의 발전

위험조정 모델은 여러 국가들에서 사용되고 있다. 미국에서는 의료보험 보장을 받지 못하는 국민들의 대한 문제가 지속적으로 제기 되었다. 그래서 CMS(the Center for Medicare and Medicaid Services)에서는 위험조정 모델을 개발하기 위해 노력해 왔다.

이러한 진단적 비용집단-계층적 질환군 위험조정모델(Diagnostic Cost Group Hierarchical Condition Category, DCG-HCC model)은 연령, 성별, 질환력을 근거로 HCC 위험지수를 산출하여 위험 평가를 한다. 2004년 CMS-HCC V12 모델에서 2019년 CMS-HCC V23 모델로 지속적인 변화를 이루어왔다. 계층적 질환군 위험조정 모델을 도입한 CMS-HCC모델은 2004년 R2가 0.098에서 2019년 대상군에 따라 0.098~0.131로 증가되어 약 10% 이상의 설명력을 보이고 있다.¹¹

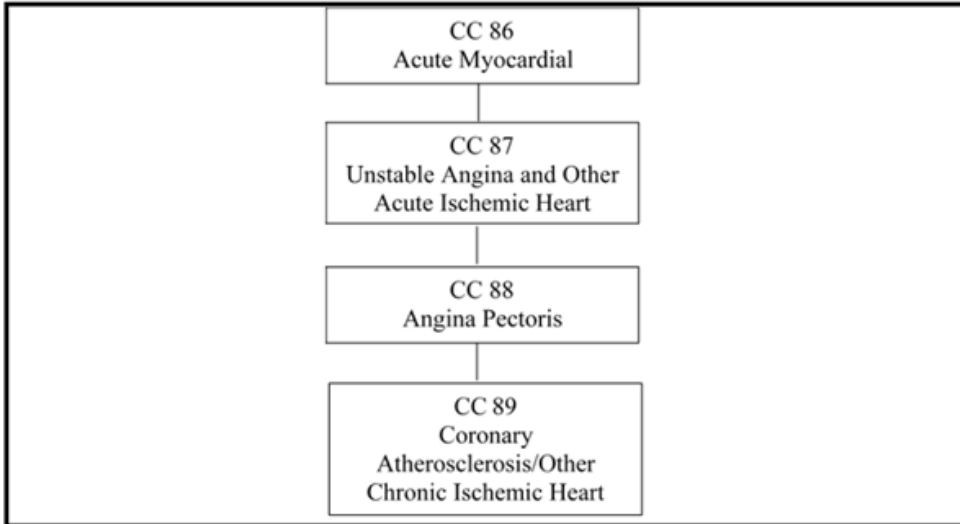
3. 질환군 분류 및 계층화

1) 질환군의 분류

CMS HCC

CMS-HCC 위험조정 모델은 일차적으로 204개 질환군 전체를 대상으로 질환군 계층화 분류 작업을 실시하고, 전체 204개 질환군 중 최종적으로 CMS-HCC V23 모델에 포함된 83개 질환군을 대상으로 상위 질환군과 하위 질환군을 나누어 계층화 분류 기준을 정하여 질환군을 최종 계층화한다. 의무 기록 등 의료 자료에 상위

질환군과 하위 질환군이 함께 존재하는 경우 질환 중증도가 높고 의료비용이 많이 드는 상위 질환군에 하위 질환군을 포함시켜 HCC 위험점수 산출시 중복 계산되지 않도록 한다. 질환군의 계층화 과정을 통해 더 정확한 의료비 추정이 가능하다.



출처: Report to Congress: Risk Adjustment in Medicare Advantage. 2018.12

[그림 5-1] 허혈성 심질환군의 계층화 단계 사례

계층적 질환군 모델은 질환의 중증도와 관련 의료비용 수준이 기록된 의료보험 가입자의 자료를 통해 계층화를 부여하는 것이다. 예를 들어 CMS-HCC에서 허혈성 심질환은 4개의 질환군(CC86~89)으로 분류된다(그림 1-1). 어떤 환자의 의무기록에 이 질환군이 함께 존재할 경우, 급성심근경색 질환군(CC86)이 질환 중증도가 높고 의료비용이 많이 발생하므로 상위 HCC에 해당하고, 하위 HCC로 불안정 협심증 질환군(CC87)과, 협심증 질환군(CC88), 만성 허혈성 심질환군(CC89)는 포함되어 예측 의료비 계산에서 중복되지 않도록 제외된다.

또한 두 가지 이상의 질병이 함께 있으면서 상호작용으로 총의료비가 각 질환군에 의해 예상되는 의료비보다 크다고 예상되는 경우 상호작용 요인을 추가하여 예측 의료비 계산에 포함한다.

4. 계층적 질환군 위험조정 모델의 적용

CMS-HCC 위험조정모델은 메디케어 가입자들과 특정 대상군의 의료비 예측이 위험조정的重要한 목표이기 때문에 몇 가지 집단으로 구분하여 적용하고 있다. ACO (Accountable Care Organization)에서도 계층적 질환군 위험조정 모델이 적용된다.¹² PACE 서비스를 위한 각 참여자의 HCC 위험점수는 CMS-HCC 모델을 이용하여 계산되고, PACE 서비스를 위한 지불금액은 계산된 개인별 HCC 위험점수에 노쇠 요인(frailty factor)이 더해져 최종 비용이 결정된다.¹³ 대상자의 HCC 위험점수는 CMS-HCC 위험조정 모델에 의해 결정된다.¹⁴ 또한 PCMH 인증 기준에서도 CMS-HCC가 환자를 위험도에 따라 분류하여 관리하는 분류 방법으로도 제시되고 있다.¹⁵ 미국 보건복지부(The U.S. Department of Health and Human Services)의 HHS-HCC 모델은 HCC 위험점수를 이용하여 각 보험의 예측 의료비용을 산출하고, 이를 각 주의 기준평균 보험료와 비교하여 위험점수가 낮은 보험사로부터 위험점수가 높은 보험사의 손실분을 조정한다.¹⁶

한기명⁴⁵ 전기홍⁶ 등의 연구에서 계층적 질환군으로 지역별 의료비 예측을 시도하였다. 김윤 등은 포괄적 일차의료 제공을 위한 제도적 지원 방안 연구에서⁷ 계층적 질환군 연구결과를 제시하며, 제한점으로 HCC 계수 산출하기에 표본 수가 적은 점을 명시하였다. 이상현 등 연구에서는 2016년과 2017년 국민건강보험을 이용한 연 2천3백만명 이상의 전수 자료와 2006년부터 2015년까지 10년도 국민건강보험 코호트 자료를 이용하여 계층적 질환군 위험조정 모델을 구축한 바 있다.⁷

제3절 연구 내용 및 방법

1. 연구 내용

2016년과 2017년 국민건강보험을 이용한 연 2천3백만명 이상의 전수 자료와 2006년부터 2015년까지 10년도 국민건강보험 코호트 자료를 이용하여 NHIS-HCC모델을 개발한 바 있다. 본 연구에서는 기존 연구에서 개발한 NHIS-HCC모형을 바탕으로 2019년 의료비를 예측하고 이를 실제 의료비와 비교하여 모형을 재평가하고자 하였다. 2019년 자료 획득 및 시한의 문제로 2019년 자료의 분석은 못하였고, 기존 2019년 보고서의 방법론 및 기존 연구 결과를 수록하였다.⁷

2. 방법

1) 질환군의 분류와 계층화

① 질환군의 분류

질환군 분류를 위해 약 71,000개의 ICD-10 진단 코드와 29,388개의 KCD-6의 진단 코드를 매칭한 후 4,673개의 진단군(Diagnostic Groups; DXGs)으로 정리하였다.

계층적 질환군 위험조정 모델은 입원 주질환이나 각 질환별로 각각 분석하지 않고, 비슷한 위험 정도를 가지고 유사한 질환 특성을 가지는 질환들을 질환군으로 통합한다. 이에 유사한 질환 특성을 가지는 질환들로 전체 204개의 질환군으로 분류한다.

우리나라 질환군 특성에 따른 계층적 질환군 모델을 구축하기 위하여 7명의 가정의학 전문의와 2명의 예방의학 전문의, 2명의 보건 통계 전문가 등이 질병군 선정작업을 하였다. 계층적 질환군 모델에 포함할 질환군은 CMS-HCC 선정 기준 10가지를 참조하였고, 국내 유병율, 만성질환 및 질환 중증도 등의 자료를 근거로 하여 선정 작업이 이루어졌다. 71개의 질환군을 선정하였다.

② 질환군의 계층화

NHIS-HCC 위험조정 모델에 적합한 질환군을 선정 후 관련 질환군의 위험 정도를 고려하여 상위 계층적 질환군과 하위 계층적 질환군으로 구분하여 계층화했다. 이는 하위 질환군이 상위 질환군과 함께 존재할 경우 상위 계층적 질환군에 포함하여 분석함을 의미한다.

심혈관 질환군에서도 건강보험 청구 자료에 HCC86 급성심근경색 질환군과 HCC87 협심증 질환군이 함께 존재하는 경우, 상위 계층적 질환군이 HCC86에 하위 질환군인 HCC87을 포함하여 분석하였다.

이런 과정을 거쳐 이번 NHIS-HCC 위험조정 모델 연구는 71개 질환군과 성별 연령대 10개군을 이용하여 분석하였다.

<표 5-1> NHIS-HCC 위험조정 모델에서 질환군 계층화 분류 기준

상위 HCC	상위 계층적 질환군	하위 HCC
8	Metastatic Cancer and Acute Leukemia	9, 10, 11, 12
9	Lung and Other Severe Cancers	10, 11, 12
10	Lymphoma and Other Cancers	11, 12
11	Colorectal, Bladder, and Other Cancers	12
17	Diabetes with Acute Complications	18, 19
18	Diabetes with Chronic Complications	19
27	End-Stage Liver Disease	28, 29, 80
28	Cirrhosis of Liver	29
46	Severe Hematological Disorders	48
54	Drug/Alcohol Psychosis	55, 56
55	Drug/Alcohol Dependence	56
57	Schizophrenia	58, 59, 60
58	Reactive and Unspecified Psychosis	59, 60
59	Major Depressive, Bipolar, and Paranoid Disorders	60
70	Quadriplegia	71, 72, 103, 104, 169
71	Paraplegia	72, 104, 169
72	Spinal Cord Disorders/Injuries	169
83	Respiratory Arrest	84
86	Acute Myocardial Infarction	87, 88
87	Unstable Angina and Other Acute Ischemic Heart Disease	88
99	Intracranial Hemorrhage	100
103	Hemiplegia/Hemiparesis	104
106	Atherosclerosis of the Extremities with Ulceration or Gangrene	107, 108, 161, 189
107	Vascular Disease with Complications	108
110	Cystic Fibrosis	111, 112
111	Chronic Obstructive Pulmonary Disease	112
114	Aspiration and Specified Bacterial Pneumonias	115
135	Acute Renal Failure	136, 137, 138
136	Chronic Kidney Disease, Stage 5	137, 138
137	Chronic Kidney Disease, Severe (Stage 4)	138
157	Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone	158, 161
158	Pressure Ulcer of Skin with Full Thickness Skin Loss	161

* 하위 HCC가 상위 HCC와 함께 발생시 상위 HCC에 포함하여 계층화 사례: HCC8 전이상 암 질환군과 HCC9 폐암 질환군이 함께 있을 경우 하위 질환군인 HCC9가 상위 질환군인 HCC8에 포함되어 HCC 점수 계산 및 분석에서 제외

3. NHIS-HCC 위험조정 모델의 의료비 예측

2019년 표본코호트 자료를 이용하여 계층적 질환군 위험조정 모델에 따른 회귀분석으로 계층적 질환군 각각의 HCC 위험지수를 산출하고자 하였다. HCC 위험지수를 근거로 계층적 질환군별 2019년 의료비를 예측하고 이를 실제 사용된 의료비와 비교하고자 하였다. 2019년 자료 획득 및 시한의 문제로 2019년 자료의 분석은 못하였다.

제4절 분석 결과

HCC 위험지수 (예시, 2019년 연구보고서⁷ 결과)

<표 5-2> 계층적 질환군에 따른 HCC 위험지수

HCC	설명	위험지수
		성-연령 위험지수
male_0_19	19세 이하 남성	318,768
male_20_39	20-39세 남성	81,467
male_40_64	40-64세 남성	510,224
male_65_74	65-74세 남성	923,736
male_75	75세 이상 남성	916,603
female_0_19	19세 이하 여성	190,029
female_20_39	20-39세 여성	33,755
female_40_64	40-64세 여성	339,129
female_65_74	65-74세 여성	956,105
female_75	75세 이상 여성	1,258,596
		HCC 위험지수
HCC1	HIV/AIDS	6,132,945
HCC2	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	10,092,576
HCC8	Metastatic Cancer and Acute Leukemia	9,730,649
HCC9	Lung and Other Severe Cancers	2,527,758
HCC10	Lymphoma and Other Cancers	3,376,659
HCC11	Colorectal, Bladder, and Other Cancers	2,131,208
HCC12	Breast, Prostate, and Other Cancers and Tumors	1,031,187
HCC17	Diabetes with Acute Complications	612,230
HCC18	Diabetes with Chronic Complications	983,893
HCC19	Diabetes without Complication	508,573

HCC	설명	위험지수
HCC21	Protein-Calorie Malnutrition	1,041,091
HCC23	Other Significant Endocrine and Metabolic Disorders	500,563
HCC27	End-Stage Liver Disease	271,081
HCC28	Cirrhosis of Liver	525,609
HCC29	Chronic Hepatitis	503,031
HCC33	Intestinal Obstruction/Perforation	2,133,743
HCC34	Chronic Pancreatitis	1,370,685
HCC35	Inflammatory Bowel Disease	1,527,214
HCC40	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	791,498
HCC46	Severe Hematological Disorders	1,842,035
HCC48	Coagulation Defects and Other Specified Hematological Disorders	1,552,164
HCC51	Dementia	2,222,949
HCC54	Substance Use with Psychotic Complications	3,045,701
HCC55	Substance Use Disorder, Moderate/Severe, or Substance Use with Complications	1,719,268
HCC56	Substance Use Disorder, Mild, Except Alcohol and Cannabis	2,605,296
HCC57	Schizophrenia	2,970,834
HCC58	Reactive and Unspecified Psychosis	1,594,413
HCC59	Major Depressive, Bipolar, and Paranoid Disorders	791,487
HCC60	Personality Disorders	165,138
HCC70	Quadriplegia	16,469,924
HCC71	Paraplegia	7,991,226
HCC72	Spinal Cord Disorders/Injuries	1,096,818
HCC73	Amyotrophic Lateral Sclerosis and Other Motor Neuron Disease	677,163
HCC74	Cerebral palsy	2,560,032
HCC78	Parkinson's disease	712,379
HCC79	Seizure Disorders and Convulsions	1,611,609
HCC80	Coma, Brain Compression/Anoxic Damage	2,559,456
HCC83	Respiratory Arrest	4,124,734
HCC84	Cardio-Respiratory Failure and Shock	3,908,069
HCC85	Congestive Heart Failure	1,111,437
HCC86	Acute Myocardial Infarction	1,521,612
HCC87	Unstable Angina and Other Acute Ischemic Heart Disease	914,636
HCC88	Angina Pectoris	311,137

HCC	설명	위험지수
HCC96	Specified Heart Arrhythmias	748,061
HCC99	Intracranial Hemorrhage	3,035,355
HCC100	Ischemic or Unspecified Stroke	477,713
HCC103	Hemiplegia/Hemiparesis	11,408,901
HCC104	Monoplegia, Other Paralytic Syndromes	2,506,475
HCC106	Atherosclerosis of the Extremities with Ulceration or Gangrene	3,498,171
HCC107	Vascular Disease with Complications	630,797
HCC108	Vascular Disease	310,759
HCC110	Cystic Fibrosis	3,699,945
HCC111	Chronic Obstructive Pulmonary Disease	444,341
HCC112	Fibrosis of Lung and Other Chronic Lung Disorders	1,640,894
HCC114	Aspiration and Specified Bacterial Pneumonias	1,112,748
HCC115	Pneumococcal Pneumonia, Empyema, Lung Abscess	1,227,195
HCC122	Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	235,823
HCC127	Cataract	409,242
HCC135	Acute Renal Failure	3,933,041
HCC136	Chronic Kidney Disease, Stage 5	15,418,633
HCC137	Chronic Kidney Disease, Severe (Stage 4)	787,999
HCC138	Chronic Kidney Disease, Moderate (Stage 3)	636,367
HCC157	Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone	10,047,586
HCC158	Pressure Ulcer of Skin with Full Thickness Skin Loss	8,743,433
HCC161	Chronic Ulcer of Skin, Except Pressure	1,912,884
HCC162	Severe Skin Burn or Condition	2,026,163
HCC167	Major Head Injury	533,696
HCC169	Vertebral Fractures without Spinal Cord Injury	1,071,408
HCC170	Hip Fracture/Dislocation	3,870,100
HCC173	Amputations and Complications	2,331,445
HCC176	Complications of Specified Implanted Device or Graft	5,619,124

성-연령 위험지수: 성별 연령대 회귀계수와 상수항을 각각 합산한 값
<이상현 외. 계층적 질환군(NHIS-HCC) 위험조정 모델의 의료비 예측과 사망률 분석. 국민건강보험 일산병원 보고서 2019 결과>

제5절 결론 및 제언

이번 연구에서는 2019년 표본 코호트 자료를 이용하여 계층적 질환군 위험조정 모델을 구현하여, 실제의료비와 예측의료비를 비교하여 추후 지불제도 변화 시 이용할 로직을 점검하고자 하였다. 연구 자료 획득상의 문제로 추후 방법론의 제시에 머물렀고, 기존 연구 결과물⁷⁾을 수록하였다.

추후 검증되는 NHIS-HCC를 바탕으로 의료비를 예측하여 제한된 정부자원을 효율적으로 배분하는 근거 자료 활용한다면, 올바른 의료 이용에 대한 정보와 함께 효과적인 재정 계획을 편성할 수 있을 것으로 판단된다.

제6장

Data 결합을 통한 공공서비스 제공에 관한 연구

제1절 서론	119
제2절 이론적 고찰	120
제3절 연구 내용 및 방법	120
제4절 분석 결과	121
제5절 결론 및 제언	127

제6장

Data 결합을 통한 공공서비스 제공에 관한 연구

제1절 서론

1. 연구의 배경 및 필요성

보건의료정보의 활용을 위해서는 우선 보건의료 데이터 셋 구축이 필요하다. 이러한 보건의료 데이터 셋의 원천 자원으로는 한국형 뉴딜 사업의 핵심인 데이터 댐으로 마련할 수 있다. 또한 인공지능 활용을 위해서는 고품질의 학습용 데이터가 필요하다. 데이터 댐을 통해 수집·축적된 데이터의 안전한 학습과 인공지능개발 및 활용을 지원할 수 있도록 해야 한다.

보건의료데이터가 주로 병원중심에서 범위를 확장하여 지역사회의 여러 보건의료정보를 포함해야한다는 수요가 높아지고 있다. 인구가 고령화되면서 점차 지역사회 중심으로 주거, 보건의료, 돌봄 서비스를 제공하는 통합형 보건복지 서비스가 필요하게 되었다. 돌봄사업에서 서비스 제공은 단순히 홈케어만이 아닌 지역사회 내 여러 자원(예를 들면, 주민센터, 보건소 등)을 활용하여 서로 연계된 서비스를 제공 하게 된다. ICT 기술의 발전을 통해 최첨단 장비 기술을 활용한 돌봄케어 서비스로 거동이 불편한 노인이나 장애인을 대상으로 모니터링 서비스를 제공하여 서비스가 가능하다. community health record 기반의 수집된 데이터를 추적하고, 분석하고 이를 서비스에 활용하고 지역사회를 혁신할지 등에 관한 보건의료복지 데이터 연계의 현황과 방향을 모색해 보고자 하였다.^{1,2}

본 연구에서는 현재 분리 운영 중인 건강보험 Data 와 장기요양 Data를 연계하여, 연계 Data set 의 유용성을 확인해 보고자 하였다.

2. 연구의 목적

보건의료 Data-set 구성을 위해 장기요양보험과 건강보험 자료의 적합성을 검증하며 두 자료를 결합하여 총 진료비 및 의료 이용 행태 등의 관련 요인들을 분석하는 것을 목표로 한다.

제2절 이론적 고찰

‘데이터 댐’은 정부가 2020년 7월 14일 확정·발표한 정책인 ‘한국판 뉴딜’의 10대 대표 과제 중의 하나이다. 10대 대표과제는 디지털 뉴딜(3개), 그린 뉴딜(3개), 융합과제(4개)로 구성되어 있는데, 데이터 댐은 디지털 뉴딜 분야에 속한다. ‘데이터 댐’ 프로젝트는 양질의 데이터 공급을 확대하고, 국가 전반의 데이터, 인공지능 활용을 확산하며, 데이터 인공지능의 핵심 인프라를 확충하는 사업으로 최종적으로 인공지능 학습용 데이터 및 빅데이터 플랫폼 구축을 목표로 한다.³

현재까지는 보건의료 관련 기관은 각자의 설립목적에 따라 보건의료 데이터를 수집하여 축적하고 있으나 기관 간 데이터에 대한 통합·연계 체계가 부재된 상태이다. 그래서 공공 정보시스템의 연계를 통한 보건복지 빅데이터 서비스 체계를 구축하여 분석의 폭을 넓히는 것이 필요한 시점이다.⁴⁵ 여러 분절화 된 보건의료정보의 통합적 활용을 위해서는 우선 관련 자료의 데이터 셋 구축이 필요하다.

단계적으로는 현재 2022년까지 구축중인 PHR (Personal Health Record) 사업 및 사회보장 정보의 PSR (Personal Social Affair Record) 사업의 완성과 함께 추후 연계가 요망된다.

제3절 연구 내용 및 방법

1. 연구 내용

1) 고양시 65세 노인 인구 대상 분석

경기도 고양시의 노인 인구 현황을 알기 위해 장기요양등급별 인원수, 재택의료 서비스(가정간호, 방문간호, 장애인주치의, 일차의료 만성질환, 방문진료)의 연간 이용자수와 진료건수를 분석하였다.

장기요양등급과 장애등급에 따라 정상노인군, 중증장애노인, 경증장애노인으로 분류하였으며 장기요양등급에 따라 총진료비, 본인부담금, 공단부담금, 입원일수, 외래일수, 처방일수를 분석하였다.

2) 6월 7월 등급판정위원회 신규등급자 분석

고양시 건강보험공단 지사에서 2021년 6월 7월 신규 장기요양 등급판정자를 대상으로 거주지와 장기요양등급판정조사표의 항목 점수를 분석하였다. 각 항목은 신체

기능영역, 사회생활 기능 영역, 인지기능 영역, 행동변화 영역, 간호처치 영역, 재활영역, 환경평가영역이다.

2. 연구 방법

2020년 1월 1일부터 2020.12.30일의 1년을 기준으로 고양시에 거주하는 1,085,591명 중 65세 이상 노인인구 156,861명을 대상으로 하였다. 건강보험데이터와 장기요양 데이터를 결합하여 대상자의 질병 및 의료기관 이용행태의 의료 정보와 장기 요양 서비스 자료 및 제공 서비스의 요양자료를 분석하였다. 최종적으로 서비스 대상자 제공서비스 현황 분석 및 질병, 요양등급, 사회 경제적 요인을 분석하였다.

제4절 분석 결과

1. 고양시 65세 이상의 노인 인구 현황 파악

2020년 1월 1일부터 2020년 12월 31일 기준으로 고양시 인구 1,085,591명 중 65세 이상의 인구는 156,861명이며 그 중 장기요양 등급 판정자(장기요양 1-5등급)는 16,092명, 등급외자는 2,408명이다. 장애등급별로 심한장애인은 5692명, 심하지 않은 장애인은 15,295명, 독거는 37,159명이다. 재택서비스 중 가정간호 이용자는 3079명, 방문간호는 159명, 일차의료 방문진료는 13명, 장애인 주치의 2명, 일차의료 만성질환 시범사업 이용자는 2711명이다. 노인 1인 당 연간 총 진료비는 3,125,160원이다.

<표 6-1> 고양시 65세 이상 인구 대상 장기요양등급별 특성

지역구		덕양구	일산동구	일산서구	합계
고양시 전체인구		479679	302339	303573	1,085,591
고양시 65세이상인구	소계	74894	40895	41072	156,861
	남	32351	17741	17495	67,587
	여	42543	23154	23577	89,274
차상위	소계	742	239	202	1,183
	남	254	76	65	395
	여	488	163	137	788
소득분위	1~10	27613	13136	12830	53,579
사망자		177	94	104	375
독거		19220	8983	8956	37,159

지역구		덕양구	일산동구	일산서구	합계	
장기요양 등급판정	소계	7244	4617	4231	16,092	
	장기요양 등급자	1등급	308	291	237	836
		2등급	684	466	369	1,519
		3등급	2094	1417	1247	4,758
		4등급	3146	1865	1785	6,796
		5등급	1012	578	593	2,183
	소계	1137	616	655	2,408	
	등급외자	인지지원등급	214	133	154	501
		등급외 A,B,C	923	483	501	1,907

<표 6-2> 고양시 65세 이상 인구 대상 재택의료 서비스 현황

	이용자수	진료건수
가정간호	3,079	20,702
방문간호	159	930
장애인주치의	2	5
일차의료만성질환	2,711	13,633
일차의료방문진료(왕진)	13	20

2. 장기요양 등급별 총 진료비 및 의료이용행태

1) 정상 노인군의 장기요양 등급별 비교

65세 이상에서 장기요양 등급 및 심한장애, 심하지 않은 장애로 나누어 의료비용, 의료이용 행태를 비교하였다. 노인 1인당 총 진료비에 비해 장기요양 등급자는 총 진료비가 높은 경향을 보인다. 정상 노인군에서는 장기요양 등급이 높을수록 총진료비 및 본인 부담금, 공단 부담금이 증가하였다. 등급이 높을수록 연간 입원일수 역시 많았고 외래 내원일수는 등급이 낮을수록 높은 경향을 보였다. 약제 처방 일수는 등급이 낮을수록 많았다.

<표 6-3> 고양시 65세 이상 인구 대상 장기요양등급별 특성(Y 기각, Z 탈락)

요양 등급	빈도	총진료비	본인 부담금	공단 부담금	입원 일수	내원 일수	처방 일수
미신청	122293	2,269,217	498,885	1,762,624	15.2	30.3	471.5
1	443	17,123,290	3,165,314	13,935,732	196.9	28.3	460.9
2	869	11,061,128	2,149,631	8,897,758	155.9	26.6	527.9

요양 등급	빈도	총진료비	본인 부담금	공단 부담금	입원 일수	내원 일수	처방 일수
3	2893	7,062,031	1,373,255	5,672,338	103.9	28.1	584.4
4	4590	5,436,739	997,512	4,426,364	65.5	32.0	623.5
5	1704	3,463,192	767,192	2,686,545	46.2	30.7	637.3
6	401	3,503,965	780,276	2,721,462	50.0	31.2	607.5
A	680	5,004,489	924,697	4,069,274	64.8	40.7	654.4
B	381	5,601,511	1,003,791	4,583,505	82.2	42.3	644.2
C	162	3,676,171	664,400	2,984,602	36.0	39.0	625.0
Y	300	8,684,989	1,701,920	6,950,150	83.2	41.1	643.3
Z	1158	4,795,741	854,179	3,929,115	62.4	34.9	614.6

2) 중증 장애인군 장기요양 등급별 비교

중증 장애인군에서는 등급이 높을수록 총진료비 및 본인 부담금, 공단 부담금이 증가하였다. 등급이 높을수록 연간 입원일수 역시 많았고 외래 내원일수는 등급이 낮을수록 높은 경향을 보였다. 약제 처방 일수는 등급이 낮을수록 많았다.

<표 6-4> 고양시 65세 이상 인구 중 심한 장애인에서 장기요양등급별 특성(Y 기간, Z 탈락)

요양 등급	빈도	총진료비	본인부담금	공단부담금	입원 일수	내원 일수	처방 일수
미신청	2747	8,041,256	1,071,423	6,919,791	73.6	43.9	516.8
1	282	23,726,966	4,280,351	19,306,877	208.4	33.1	452.2
2	399	15,967,290	2,645,784	13,297,784	195.8	30.4	511.9
3	903	14,342,340	2,191,457	12,118,813	141.8	43.1	547.6
4	797	11,539,236	1,385,020	10,118,672	70.1	53.5	582.4
5	111	7,830,562	1,028,759	6,786,752	63.4	51.2	673.7
6	20	6,022,370	699,138	5,322,192	124.3	49.4	514.2
A	164	11,486,155	1,298,208	10,123,414	53.9	60.4	593.4
B	64	6,027,542	766,173	5,176,450	153.9	43.0	611.6
C	25	11,182,582	1,119,524	10,063,058	19.7	57.2	724.1
Y	21	24,874,246	2,899,098	21,950,058	105.5	69.3	566.2
Z	159	9,882,213	1,457,919	8,387,057	109.7	43.2	565.0

3) 경증 장애인군 장기요양 등급별 비교

경증 장애인군에서는 등급이 높을수록 총진료비 및 본인 부담금, 공단 부담금이 증가하였다. 등급이 높을수록 연간 입원일수는 많았다. 내원 일수와 처방 일수는 등급이 낮을수록 많은 경향을 보인다.

<표 6-5> 고양시 65세 이상 인구 중 심하지 않은 장애인에서 장기요양등급별 특성

요양 등급	빈도	총 진료비	본인 부담금	공단 부담금	입원 일수	내원 일수	처방 일수
미신청	11297	3,244,529	640,586	2,578,558	20.6	37.5	569.6
1	111	21,541,291	4,068,006	17,446,134	195.1	35.0	551.4
2	251	10,593,313	1,908,432	8,635,641	144.5	26.2	530.2
3	962	7,472,941	1,451,785	5,991,305	90.2	30.8	646.0
4	1409	5,608,731	981,555	4,602,842	54.3	37.5	676.9
5	368	4,651,696	877,347	3,760,915	43.2	36.5	741.0
6	80	3,204,614	676,075	2,506,820	31.8	35.0	620.9
A	244	4,665,227	825,092	3,805,962	45.5	42.4	695.6
B	141	4,877,448	880,052	3,974,798	60.3	43.5	685.5
C	46	5,560,682	555,953	4,981,797	23.8	42.2	614.9
Y	60	10,352,140	1,765,191	8,414,746	71.4	46.2	668.8
Z	326	7,310,229	1,143,748	6,083,952	72.0	43.4	690.9

4) 그룹 간 비교

노인 1인당 총 진료비에 비해 장기요양 등급자, 중증장애인, 경증 장애인 모두 총 진료비는 높았다. 중증장애인이면서 요양등급 1, 2등급의 경우 정상노인, 경증장애인 1, 2등급에 비해 총진료비와 입원일수가 많은 경향을 보였다. 경증장애인군과 정상노인군의 2, 3, 4등급의 경우 총진료비, 본인부담금, 공단부담금은 경증장애인군에서 상대적으로 높으며 입원일수는 정상노인군이 다소 많았으며 외래내원일수는 경증장애인군에서 높았다.

3. 장기요양등급판정위원회 연계 데이터

1) 대상자 거주형태 분석

2021년 6, 7월 건강보험공단 고양시 지사에서 장기요양 등급 판정을 받은 판정자 1,811명을 대상으로 분석하였다. 1등급 판정자 중 자택에 있는 비율은 32%, 2등급은

38%, 3등급은 50%, 4등급은 77%, 5등급은 86%이다.

<표 6-6> 고양시 신규 등급판정자 대상 거주지 분류

	자택	노인요양 시설	노인요양 공동생활 가정	양로시설	요양병원	기타	합계
1등급	36	41	2	0	32	2	113
2등급	74	81	2	0	36	1	194
3등급	266	195	5	1	65	1	533
4등급	429	102	1	1	20	4	557
5등급	199	19	2	3	5	2	230
등급외A	69	0	0	0	1	0	70
등급외B	23	0	0	0	0	0	23
등급외C	13	0	0	0	0	0	13
인지지원	74	2	0	0	1	1	78
합계	1183	440	12	5	160	11	1811

2) 대상자 등급별 장기요양등급 점수 비교

1등급의 경우 신체기능영역, 사회생활기능영역, 인지영역, 간호처치영역, 재활영역에 걸쳐 모두 점수가 높았으며 5등급인 치매등급에서만 행동변화영역의 점수가 높았다.

<표 6-7> 고양시 신규 등급판정자 대상 항목별 장기요양등급판정표 점수

등급	명수	신체기능 영역	사회생활 기능영역	인지기능 영역	행동변화 영역	간호처치 영역	재활 영역	환경 평가
1등급	113	37.0	38.7	8.2	0.4	1.3	18.3	10.3
2등급	194	30.3	37.3	6.2	0.8	0.5	14.6	10.4
3등급	533	22.6	34.5	5.0	0.8	0.1	12.0	10.4
4등급	557	17.5	30.5	4.1	0.7	0.1	10.8	10.4
5등급	230	15.7	28.1	3.9	1.4	0.0	10.2	10.4
등급외A	70	15.9	25.2	2.5	0.2	0.1	10.4	10.5
등급외B	23	14.7	21.6	2.0	0.0	0.0	10.0	10.8
등급외C	13	13.0	13.2	1.2	0.1	0.2	10.0	10.6
인지지원 등급	78	13.7	23.3	2.8	0.9	0.0	10.0	10.5

4. 결합정보 사용의 실례

- 고양시 재택의료 통합서비스 시범사업

고양시 재택의료 통합서비스 시범사업은 거동이 불편하여 병원 내원이 어려운 환자에게 코디네이션하여 개인별 맞춤형 재택서비스를 제공하는 것을 목표로 한다. 건강보험공단 시범사업 형태로 건강보험과 장기요양 데이터 활용에 유리한 측면이 있다.

이러한 결합 데이터는 사업 대상자 선정 및 개인별 맞춤형 서비스 제공에 활용되었다. 초기 사업 대상으로 외상환자를 선정하기 위하여 장기요양 1,2,3등급, 자택에서 가정간호서비스 받는 환자, 장애 등급 소지자를 대상으로 하였다. 개인별 맞춤형 서비스 제공하기 위해 개인정보 활용 동의자를 대상으로 타병원의 병력, 투약력 등의 건강보험 정보, 등급자의 장기요양 서비스 이력 등의 장기요양 정보, 공단 장기요양등급 인정조사표의 일부 항목을 참고할 수 있었다.

재택의료 시범사업의 특성상 데이터 결합이 되지 않으면 환자나 보호자의 기억이나 의무기록에 의존한 병력, 투약력 수집이 되어 방문 진료의가 환자를 파악하기가 어렵다. 또한 환자 초기평가에 코디네이터의 업무가 집중되게 되어 환자의 통합적 관리에 어려움이 있다.

시범사업 측면에서 건강보험과 장기요양 데이터 결합의 장점으로 개인별 맞춤형 재택의료 통합 서비스를 제공할 수 있고, 점차 늘어나는 재택의료 수요에 미리 준비할 수 있으며, 주치의적 접근으로 환자의 지속적인 관리를 위해 필요하다.

5. 한국사회보장정보원 Data 현황

한국사회보장정보원은 각종 사회복지 급여 및 서비스 지원 대상자의 자격과 이력에 관한 정보를 통합 관리하고, 지자체의 복지업무 처리를 지원하는 기관이다. 기존 시·군·구별 서울행정시스템의 31개 업무 지원시스템 중 복지분야를 분리하여 개인별 가구별 DB로 중앙에 통합 구축한 정보시스템인 사회보장정보시스템(행복e음)을 개설하였고 현재 차세대 정보 시스템 구축 중이다. 기간은 2020년 12월 20일부터 2023년 12월까지이며 예산은 118,354억원으로 현재 단계적으로 개통 중이며 빅데이터센터 고도화가 진행 중이다.

제5절 결론 및 제언

1. 결론

노인 1인당 총 진료비에 비해 장기요양 등급자, 중증장애인, 경증 장애인 모두 총 진료비가 높은 경향을 보인다. 정상노인군, 심한 장애인, 심하지 않은 장애인 군에서 모두 장기요양등급이 높을수록 총진료비, 본인부담금, 공단부담금 및 입원일수가 많았으며 내원일수 및 처방일수는 대체적으로 적음을 알 수 있었다. 심한 장애인군의 장기요양 등급자(1-5등급, 인지지원 등급) 모두 경증 장애인, 정상 노인 장기요양 등급자에 비해 총진료비 및 입원 일수가 많았다.

2. 제언

장기요양과 건강보험 데이터를 결합하여 분석을 진행하였다. 이러한 데이터 셋을 이용하여 고양시 재택의료 통합서비스 시범사업에서 대상자 선정 및 개인별 맞춤형 서비스 제공에 활용하였다.

이처럼 건강보험 데이터와 장기요양 데이터를 결합하여 의료 이용 및 장기요양 이용 관련 분석 및 이를 통한 지역사회 공공서비스 제공과 관련된 연계 연구가 제안된다. 또한 장기요양 요양급여 비용 데이터를 활용하여 일인당 총 의료비용을 파악하는 것도 의미가 있다.

한국 사회보장정보원의 차세대 사업이 완성되면 추후 의료, 요양, 복지 서비스의 결합 Data set 구축과 제공서비스의 적절성 및 방향에 대한 추가 분석이 가능 할 수 있다. 또한 결합된 데이터는 국가 전체적으로 의료, 요양, 복지 서비스의 통합된 제공을 위해 이용될 수 있다.

의료, 요양, 복지 서비스 결합데이터가 구축되면 현재 전국 16개 지자체에서 이루어지고 있는 통합돌봄사업 등의 지역보건의료에 적용하여 대상자 선정 및 개인별 맞춤형 서비스 제공에 이용할 수 있다.

참고문헌

참고문헌

제1장 표본코호트에서의 ARD 분석

1. fki.or.kr, [Homepage on the Internet]. 한국경제연구원, 저출산 고령화 추이 국제비교와 시사점; [cited 2021,09] Available from: http://www.fki.or.kr/FkiAct/Promotion/Report/View.aspx?content_id=2d29dee0-dd04-46dc-9865-abdcfe33ec14&cPage=9&search_type=0&search_keyword=44qo44dithhpwr
2. kostat.go.kr, [Homepage on the Internet]. 대전: 통계청; [cited 2021, 09]. Available from: http://kostat.go.kr/portal/korea/kor_nw/1/1/index.board?bmode=read&aSeq=385322
3. 오인환, 윤석준, 김은정, 한국인의 질병부담, 대한의사협회지, 2011;54(6):646-652
4. hira.or.kr, [Homepage on the Internet] 건강보험심사평가원; [cited 2021,09] Available from: <https://www.hira.or.kr/bbsDummy.do?pgmid=HIRAA020045030000&brdScnBltno=4&brdBltNo=2413&pageIndex=1#none>
5. 김철호. 노인질환의 특징. 대한내과학회지 2006;71(2):844-847.
6. 유형준. 노년기질환의 특징-노인증후군, 대한내과학회지 2009; 77(4):1073-1076
7. Won CW, Yoo HJ, Yu SH, Kim CO, Dumlao LCI, Dewiasty E et al. Lists of geriatric syndromes in the Asian-Pacific geriatric societies. European Geriatric Medicine. 2013;4(5):335-338.
8. Inês Figueira, Adelaide Fernandes, Aleksandra Mladenovic Djordjevic, Andres Lopez-Contreras, Catarina M. Henriques, Colin Selman, et al. Interventions for age-related diseases: Shifting the paradigm, Mechanisms of Ageing and Development. 2016: 160; 69-92
9. Yanfei Liu, Weiliang Weng, Rui Gao, Yue Liu. New Insights for Cellular and Molecular Mechanisms of Aging and Aging-Related Diseases: Herbal Medicine as Potential Therapeutic Approach, Oxidative Medicine and Cellular Longevity 2019
10. Efraim Jaul, Jeremy Barron. Age-related diseases and clinical and public health implications for the 85 years old and over population. Frontiers in public health, 2017: 5; 335.

11. Florencia Barbé-Tuana, Giselle Funchal, Carine Raquel Richter Schmitz, Rafael Moura Maurmann, Moisés E. Bauer. The interplay between immunosenescence and age-related diseases. *Seminars in Immunopathology* 2020; 42; 545-557
12. GBD 2019 Diseases and Injuries Collaborators, Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 396(10258);1204-1222.
13. Young-Eun Kim, Hyesook Park, Min-Woo Jo, In-Hwan Oh, Dun-Sol Go, Jaehun Jung et al. Trends and patterns of burden of disease and injuries in Korea using disability-adjusted life years. *J Korean Med Sci.* 2019; 34(Suppl 1)
14. Itay Katzir, Miri Adler, Omer Karin, Netta Mendelsohn-Cohen, Avi Mayo, Uri Alon. Senescent cells and the incidence of age-related diseases. *Aging Cell.* 2021 Mar;20(3):e13314
15. Belikov, A. V.. Age-related diseases as vicious cycles. *Ageing Research Reviews*, 2019;49;11-26.
16. Zenin, A. , Tsepilov, Y. , Sharapov, S. , Getmantsev, E. , Menshikov, L. I., Fedichev, P. O. et al. Identification of 12 genetic loci associated with human healthspan. *Communications Biology.* 2019; 2;41
17. Judith Campisi, Ladislav Robertb. Cell Senescence: Role in Aging and Age-Related Diseases. *Aging: Facts and Theories.* 2014;39:45-61
18. Angela Y Chang, Vegard F Skirbekk, Stefanos Tyrovolas, Nicholas J Kassebaum, Joseph L Dieleman. Measuring population ageing: an analysis of the Global Burden of Disease Study 2017. *The Lancet Public Health.* 2019; 4(3): e159-e167

제2장 검진코호트에서의 ARD 분석

1. 통계청 [Internet]. 2021 고령자 통계 ; Availalbe from: https://kostat.go.kr/portal/korea/kor_nw/1/1/index.board?bmode=read&aSeq=403253
2. Belikov AV. (2019). Age-related diseases as vicious cycles. *Ageing research reviews*, 49, 11-26.

3. 정채림. “건강보험 표본코호트 자료를 이용한 의료비 증가 요인 분석.” 국내박사학위논문 서울대학교 보건대학원, 2015. 서울
4. Wang C, Li F, Wang L, et al. (2017). The impact of population aging on medical expenses: a big data study based on the life table. *Bioscience trends*, 2017; 11(6):619-631.
5. 국민건강보험공단(2017). 고령사회를 대비한 노인의료비 효율적 관리방안 연구용역 보고서
6. Kang YG, Cho KH, Bae CY, et al. (2018). Biological age as a health index for mortality and major age-related disease incidence in Koreans: National health Insurance service-health screening 11-year follow-up study. *Clinical interventions in aging*, 13, 429.
7. von Wyl V. (2019). Proximity to death and health care expenditure increase revisited: A 15-year panel analysis of elderly persons. *Health economics review*, 9(1), 1-16.
8. Gregersen FA. (2014). The impact of ageing on health care expenditures: a study of steepening. *The European Journal of Health Economics*, 15(9), 979-989.
9. Atella V, Piano Mortari A, Kopinska J, et al. (2019). Trends in age related disease burden and healthcare utilization. *Aging cell*, 18(1), e12861.
10. Franceschi C, Garagnani P, Morsiani C, et al. (2018). The continuum of aging and age-related diseases: common mechanisms but different rates. *Frontiers in medicine*, 5, 61.

제3장 골다공증 예측 모델의 개발

1. 골다공증 진료지침. 대한골대사학회. 2018. p.12-15.
2. 골다공증 제5판. 대한골대사학회. 2016. p.160-167
3. 골다공증 및 골다공증 골절 FACT SHEET 2019. [cited 2019 Nov 28] Available from:<https://www.ksbmr.org/bbs/index.html?code=fact&category=&gubun=&page=1&number=702&mode=view&keyfield=&key=>
4. 노성현 외. 대한민국 국민의 골다공증 치료의 경향 및 골다공증 환자들의 치료 여부에 따른 골다공증성 압박성 골절 발생 빈도 연구. 일산병원 연구소 연구보고서 2020-20-012. p.9-11.

5. Hyung Jin Choi et al, Burden of osteoporosis in adults in Korea: a national health insurance database study. J Bone Miner Metab(2012) 30:54-58.
6. Kyaе Hyung Kim et al, Prevalence, awareness, and treatment of osteoporosis among Korean women: The Forth Korea National Health and Nutrition Examination Survey. Bone 50(2012) 1039-1047.
7. Chan Soo Shin et al, Prevalence and risk factors of osteoporosis in Korea: A community-based cohort study with lumbar spine and hip bone mineral density. Bone 47 (2010) 378-387.
8. 최지엽 외. 한국인 여성 골다공증 및 골감소증의 유병률과 위험요인: 지역사회 단면연구. 대한폐경회지 제14권 제1호 2008.
9. Young Jun Choi et al, The prevalence of osteoporosis in Korean adults aged 50 years or older and the higher diagnosis rates in women who were beneficiaries of a national screening program: The Korea National Health and Nutritional Examination Survey 2008-2009. Journal of Bone and Mineral Research, Vol. 27, No.9, September 2012, pp1879-1886.

제4장 골다공증 환자 교육 처방 메시지

1. 골다공증 예방관리 자주하는 질문. 보건복지부. ISBN 978-89-6838-033-4, 2014.
2. 골다공증을 아세요?. 대한골대사학회. Available from: <https://www.ksbmr.org/info/index.php>
3. 골다공증의 진료지침. 대한골대사학회. 2020. Available from: <https://www.ksbmr.org/info/index.php>
4. LOVE YOUR BONES. International Osteoporosis Foundation. Available from: <https://www.osteoporosis.foundation/>
5. THAT'S OSTEOPOROSIS. International Osteoporosis Foundation. Available from: <https://www.osteoporosis.foundation/>
6. IOF Osteoporosis Risk Check. International Osteoporosis Foundation. Available from: <https://www.osteoporosis.foundation/patients#osteoporosis-riskcheck>

7. Know your risk for osteoporosis. International Osteoporosis Foundation. Available from: <https://www.osteoporosis.foundation/>

제5장 계층적 질환군(NHIS-HCC) 위험조정 모델을 통한 개인별 건강위험 평가

1. 통계청. 장래인구추계.; 2011.
2. Winkelman R, Mehmud S. A Comparative Analysis of Claims-Based Tools for Health Risk Assessment. Soc Actuar. 2007.
3. 한기명. 위계적 질환군 (HCC) 위험조정모델 기반 의료비용 예측. 아주대학교 대학원 박사학위논문. 2016.
4. 한기명, 유미경, 전기홍. 위계적 질환군 위험조정모델 기반 의료비용 예측. 보건행정학회지. 2017;27(2):149-156.
5. 이윤성, 김윤, 전기홍, 강희정. 미래보건의료체계 개선방향 마련을 위한 정책연구. 대한의학회 보고서. 2017.
6. 김윤, 광미영, 안아름, 최용준. 포괄적 일차의료 제공을 위한 제도적 지원 방안 연구. 서울대학교 산학협력단. 2016.
7. 이상현 외. 계층적 질환군(NHIS-HCC) 위험조정 모델의 의료비 예측과 사망률 분석. 국민건강보험 일산병원 보고서2019
8. Downs M. Risk Assessment and Risk Adjustment. Am Acad Actuar. 2010;5:1-8. www.actuary.org. Accessed October 24, 2019.
9. Hall MA, Fred JD, Turnage EL. Issue Brief Risk Adjustment Under the Affordable Care Act: A Guide for Federal and State Regulators. Commonw Fund pub. 2011;7:1-11.
10. Newhouse JP. Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection. J Econ Lit. 1996;34:1236-1263. <https://www.jstor.org/stable/2729501>. Accessed October 24, 2019.
11. 1Centers for Medicare and Medicaid Services. Report to Congress: Risk Adjustment in Medicare Advantage.; 2018. <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf>.

12. Yeatts JP, Sangvai DG. HCC Coding, Risk Adjustment, and Physician Income: What You Need to Know. *Fam Pract Manag.* 2016;24-27..
13. CMS. Programs of All-Inclusive Care for the Elderly (PACE) Manual.
14. CMS. 2018 Comprehensive Primary Care Plus Payment Methodologies; 2017.
15. NCQA. PCMH Standards and Guidelines, Edition 2, Version3.; 2017.
16. Health Exchange Risk Adjustment: HHS-HCC.
17. <http://www.millimanriskadjustment.com/Solutions/HHS-HCC/>. Accessed October 30.

제6장 Data 결합을 통한 공공서비스 제공에 관한 연구

1. 조경희, 강성홍, 박혜진, 이영성 등. 지역의료정보 체계의 과제와 발전방향. 2020. 보건복지부 정책 리포트
2. 남상요. 노인을 위한 커뮤니티케어 현황과 사례. 2020. 대한임상노인의학회 학술집
3. <https://www.knewdeal.go.kr/>
4. 송태민·이상영·이기호·박대순·진달래·류시원·장상현(2011), “u-Health 현황과 정책과제,” 한국보건사회연구원.
5. 건강보험심사평가원. 결합전문기관의 역할. 2020 병원 CIO 포럼 학술집
6. 이우식 등. 사회 보장 데이터 활용과 빅데이터 플랫폼. 2021 한국빅데이터학회 학술자료집

연구보고서 2021-20-014

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발(8차년도)

발행일	2022년 2월 28일
발행인	김성우
편집인	이천균
발행처	국민건강보험 일산병원 연구소
주소	경기도 고양시 일산동구 일산로 100(국민건강보험 일산병원)
전화	031) 900-6977, 6985
팩스	0303-3448-7105~7
인쇄처	지성프린팅 (02-2278-2490)



(우)10444 경기도 고양시 일산동구 일산로 100(백석1동 1232번지)
대표전화 1577-0013 / 팩스 031-900-0049
www.nhimc.or.kr

개인별 맞춤형 통합 건강관리 프로그램 연구 및 개발(8차년도)