

건강위험요인에 따른 의료비지출  
크기 비교 연구

연세대학교 보건대학원

보건정보관리학과

김 혜 영

건강위험요인에 따른 의료비지출  
크기 비교 연구

지도 채 영 문 교수

이 논문을 보건학석사 학위논문으로 제출함

2004년 12월 일

연세대학교 보건대학원

보건정보관리학과

김 혜 영

김혜영의 보건학석사 학위논문을 인준함

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

연세대학교 보건대학원

2004년 12월 일

## 감사의 말씀

2년 반 동안의 대학원 생활을 마무리짓게 되었습니다. 부족한 점이 많은 저에게 대학원 생활동안 많은 경험과 좋은 만남을 이룰 수 있도록 도와주시고 격려해주신 모든 분들께 감사의 말씀을 올립니다.

부족한 저에게 논문의 시작부터 완성에 이르기까지 세심한 지도말씀과 끝까지 이끌어주신 채 영문 교수님께 머리 숙여 감사의 뜻을 올립니다.

바쁘신 중에도 귀중한 시간을 내어 부심을 맡아 지도해 주신 박태수 본부장님, 이동하 박사님께도 깊은 감사를 드립니다.

무엇보다 진학에 대한 바램으로 저에게 학문의 길을 열어주신 용왕식 부장님, 문광자 차장님께 진심으로 감사를 드립니다. 논문을 쓰도록 배려해 주신 홍갑표 지사장님, 정희자 부장님, 강형일 대리, 이위환 주임에게도 깊은 감사를 드립니다.

함께 논문을 쓰면서 고생한 동기들과 대학원 생활동안 많은 힘이 되어준 박정진 선생님, 미정 언니, 화영 언니, 진이, 이영곤 선생님께도 고마움을 전합니다.

오늘의 제가 있기까지 한없는 사랑으로 키워주시고 보살펴주신 부모님과 늘 옆에서 격려해주고 응원해 준 나의 사랑스런 남편에게 감사의 말씀을 올립니다. 마지막으로 부족한 제가 지금의 자리에 있을 수 있도록 많은 도움주신 모든 분들께 진심으로 감사 드립니다.

2004. 12.

김혜영 올림

# 차 례

국문요약 .....	iv
<b>I. 서 론</b> .....	1
1. 연구의 배경 및 필요성 .....	1
2. 연구목적 .....	4
<b>II. 문헌고찰</b> .....	5
<b>III. 이론적 배경</b> .....	7
1. 건강위험요인 .....	7
2. 데이터마이닝 .....	15
<b>IV. 연구방법</b> .....	26
1. 연구의 틀 .....	26
2. 연구자료 .....	27
3. 변수의 선정 .....	28
4. 분석방법 .....	30

<b>V. 연구결과</b>	33
1. 연구대상자의 일반적 특성	33
2. 건강위험요인 기간에 따른 의료비 평균 비교	39
3. 회귀분석 결과	42
4. 순차패턴 분석 결과	50
5. 활용방안	54
<b>VI. 고찰</b>	56
1. 연구자료 및 방법에 대한 고찰	56
2. 연구결과에 대한 고찰	59
<b>VII. 결론</b>	62
참고문헌	64
부록	69
ABSTRACT	73

## 표 차 례

표 1. 연구자료 .....	27
표 2. 의료수가 인상 지수 .....	29
표 3. 분석항목 .....	31
표 4. 연령군별 연구 대상자 .....	33
표 5. 연령별에 따른 의료비 현황 .....	34
표 6. 연구대상자의 건강위험요인 특성 비교 .....	36
표 7. 건강위험요인과 의료비의 연간 추이 분석 .....	38
표 8. 건강위험요인 기간에 따른 의료비 평균 지출 분석(성별) .....	40
표 9. 건강위험요인 기간과 의료비간의 상관관계 분석(성별) .....	43
표 10. 건강위험요인 기간과 의료비간의 상관관계 분석(연령별) .....	46
표 11. 건강위험요인 기간과 의료비간의 회귀분석(성별) .....	47
표 12. 건강위험요인 기간과 의료비간의 회귀분석(연령별) .....	49
표 13. 건강위험요인 기간과 의료비간의 순차패턴 분석(성별) .....	51
표 14. 건강위험요인 기간과 의료비간의 순차패턴 분석(연령별) .....	53

## 그 립 차 례

그림 1. 연구의 틀 .....	26
그림 2. 분석모형 .....	30

## 국문요약

만성퇴행성질환의 발생에 영향을 미치는 주요인자인 건강생활습관은 의료비 지출과 밀접한 관계를 가지고 있으며, 건강생활습관의 주요 위험요인은 흡연, 음주, 식이, 운동 등 건강과 관련된 생활양식이다. 본 연구에서는 개인의 생활습관요인과 의료비 지출 크기에 관한 분석을 하였다.

국민건강보험공단의 1998년, 2000년, 2002년 3년 동안 계속적으로 건강검진을 받은 대상자의 건강검진결과의 문진내역과 3개 질병군으로 진료한 내용의 의료비를 이용하여 분석하였다.

건강위험요인 기간에 따른 의료비 평균을 비교한 결과, 남자의 경우 음주기간이 1년과 2년의 집단차가 14,593.0원, 여자의 경우 식습관(불규칙)기간이 1년과 2년의 집단차가 77,102.1원, 연령별로는 60대의 경우 흡연기간이 1년과 2년의 집단차가 90,550.0원으로 높은 평균차가 있었다. 의료비 증가에 영향력을 미치는 주된 요인은 남자의 경우 음주기간이 증가할수록, 여자의 경우 흡연기간, 음주기간이 증가할수록, 연령별로는 40대가 음주기간이 증가할수록 의료비가 증가하였다. 건강위험요인의 기간에 따라서 의료비 증가의 영향력을 미치는 요인을 분석한 결과, 음주를 3년동안 지속적으로 하고 있는 경우 의료비 증가율이 남자는 6%, 여자는 38% 증가하였다.

결론적으로 건강위험요인 유무에 따른 의료비 지출 크기는 요인의 기간이 증가함에 따라서 의료비 지출이 증가하였다. 이에 본 연구는 우리나라의 직장근로자, 지역주민들에게 확대적용 할 수 있을 것으로 판단되고, 나아가 많은 보건분야에서 건강증진사업의 일환으로 활용할 수 있을 것으로 판단된다.



# I. 서 론

## 1. 연구의 배경 및 필요성

우리나라는 경제성장과 국민생활수준의 향상과 함께 감염성 질환은 어느 정도 퇴치되었으나, 암이나 뇌졸중, 심장병, 고혈압, 당뇨병 및 간질환 등의 만성퇴행성질환과 사고로 인한 부상이 전체 사망의 70%를 상회하는 것으로 나타나고 있다(보건복지부, 1995). 우리나라 사망원인 통계연보에 따르면 2003년 암으로 인해 6만여명이 사망하고, 뇌혈관질환으로 3만 6천여명, 심장질환으로 1만 7천여명, 당뇨병으로 1만 1천여명이 사망하여 암, 뇌혈관질환, 심장질환, 당뇨병으로 인한 사망자가 전체 사망자의 57.1%로 전체 사망자의 절반이상을 차지하고 있다(통계청, 2003). 이러한 만성질환의 발생에 영향을 미치는 주요인자는 건강생활습관과 밀접한 관계를 가지고 있다. Lalonde(1976)는 국민건강에 영향을 미치는 상대적 효과를 유전적 요인 16.7%, 환경적 요인 33.3%, 개인의 생활습관 33.3%, 의료서비스를 16.7%로 제시하고 있다. 한편 최근 우리나라의 한 연구에서도 건강에 가장 크게 영향을 미치는 요인은 개인의 생활습관이 52%를 차지하며, 유전적 요인 20%, 환경적 요인 29%이며, 의료서비스의 영향력은 8%에 불과한 것으로 나타나 우리나라가 상대적으로 개인의 생활습관이 매우 높은 것으로 나타났다(O'Donnell, 1999). 즉 만성퇴행성질환을 유발하는 주요 위험요인은 흡연, 음주, 식이, 운동 등 건강과 관련된 생활양식으로 밝혀지고 있다(US

DHHS, 1988).

1990년 미국인의 사망요인을 추정한 보고에서는 흡연으로 인한 요인이 가장 높고, 식이와 신체활동, 알코올 등이 주요 사망원인인 것으로 밝히고 있다(McGinnis, 1993). 이와 같이 각종 사망의 50%를 예방할 수 있는 요인 중 주요부분이 일상적인 개인의 건강실천 행위라고 볼 수 있다.

2001년 국민건강영양조사에 따르면 만성퇴행성질환으로 병원을 찾는 환자는 1999년 163만명에서 2002년 210만명으로 29.1%나 증가했다(보건복지부, 2004). 1999년부터 2002년까지 국민건강보험공단에 청구된 상세불명의 만성기관지염, 폐기종, 기타 만성폐질환 비용을 보면 1999년 745억원에서 2002년 815억원으로 9%나 증가했다(국민건강보험공단, 2003). 이처럼 만성퇴행성 질환은 발병률이 높아지므로 경제적으로 생산력이 있는 연령층이 노동력이 상실되므로 사회경제적으로 손실이 크다.

Wolfe는 생활방식이 건강에 영향을 미치는 것이 일정할 때 의료비 지출이 건강상태 증진에 도움을 줄 것이라는 결과를 일반적인 평균수명을 사용하는 것보다 건강상태에 영향을 미치는 흡연, 음주, 직업상의 위험 등에서 평균수명을 사용하여 타당할 결과를 도출하였다(Wolfe, 1986).

최근 의료 시스템의 가장 큰 문제점 중의 하나는 의료기술과 지식의 발전과 함께 급증하는 의료수요로 인한 기하학적인 의료비용의 증가이다. 따라서 의료비용을 감소시키기 위한 의료정책, 의료행위 및 의료내용 등에 대한 다각적인 연구가 활발히 진행되고 있다.

일차예방 및 건강증진의 의학적인 효과 즉, 질병예방, 질병발생 지연 감소, 치명률 감소, 생명 연장 등에 대한 다양한 연구 결과들이 발표되어져 왔다. 그러나 이들이 의료비용에 미치는 영향에 대한 연구는 그리 많지 않

으며 질병이 의료비용에 미치는 영향에 비하여 건강생활습관이 의료비용에 미치는 영향에 대한 연구 또한 많지 않다.

이에 본 연구에서는 최근의 변화된 추세에 맞추어 개인의 생활습관의 변화 또는 건강위험요인의 감소를 목표로 하고 있는 건강검진의 효과를 극대화하기 위해서 생활습관요인이 의료비 지출 크기에 영향을 미치는지를 분석해 보고자 한다.

## 2. 연구목적

이 연구는 데이터마이닝의 기법을 활용하여 개개인의 생활습관에 따른 건강위험요인과 의료비 지출 크기의 관한 분석으로써 생활습관과 의료비간의 상호관련성과 패턴에 대한 탐색을 통해 유용한 지식을 이끌어 내고 나아가 국민건강보험공단의 건강증진사업 정책 안정화 및 활성화를 위한 기초자료로 제시하고자 한다.

구체적인 목적을 살펴보면 다음과 같다.

첫째, 건강위험요인의 기간에 따른 의료비의 차이를 분석한다.

둘째, 의료비 증가에 영향력을 미치는 주된 건강위험요인을 분석한다.

셋째, 건강위험요인의 기간에 따라서 의료비 증가에 영향력을 미치는 요인을 분석한다.

넷째, 건강위험요인과 의료비용간의 상호관련성을 통하여서 보건분야의 확대적용 방안을 제시한다.

## II. 문헌고찰

건강위험요인과 질병발생의 관련성을 보고한 연구결과들은 많지만, 건강위험요인과 의료비 지출에 관련된 연구는 흔하지 않다.

Bertera는 45,976명의 근로자들을 대상으로 하여 1984년에서 1988년까지 건강위험요인을 측정하고 건강위험요인의 유무에 따른 의료비 지출의 차이를 살펴보았다. 그 결과 흡연, 비만, 과도한 음주, 고혈압 등 건강위험요인을 하나라도 가지고 있는 근로자가 그렇지 않은 근로자들에 비해서 통계학적으로 유의한 결근율(10~32%)을 나타냈다(Bertera, 1991).

Yen 등은 건강위험 평가도구(Health Risk Appraisal)에서 뽑은 12가지의 건강위험요인을 독립변수로 하고 의료비에 근로자들의 결근에 의한 손해비용을 종속변수로 하여 다중회귀분석을 실시한 결과, 34세 이하의 남자, 35세 이상의 남자 그리고 여자의 세 집단에서 통계적으로 유의( $p < 0.001$ )한 것으로 나타났다(Yen, 1992).

미국에서의 조사에 의하면 1985년에 흡연에 의한 의료비 및 노동력상실에 의한 연간 경제적 손실이 650억달러, 1998년 기준으로 음주를 함으로써 생기는 연간 경제 사회적 비용이 858억달러(O'Connor, 2004), 운동을 하지 않음으로써 생기는 직접적인 의료비 손실이 매년 760억달러(Benowitz, 2004)에 달하고 있다.

1995년 미국 질병관리예방센터(US Center for Disease Control and Prevention)에 의하면, 미국에서 흡연에 의하여 매년 500억달러 이상의 직접의료비용이 지출될 뿐만 아니라, 그 외에 부가적인 경제생산 손실액이

500억달러에 이른다고 한다. 미국의 통계를 보면 성인의 64%가 음주를 하고, 미국 사회가 매년 1천억 달러를 알콜 사용으로 인한 질병의 치료와 알콜 관련 경제적 손실로 예측하고 있다(MMWR, 2001).

캐나다의 경우에도 흡연자가 내는 세수입은 78억달러인데 반해 이들로 인한 사회적 비용은 150억달러 정도이며 흡연으로 인한 보건의료비는 25억 달러(캐나다 달러)로 추정하고 있다. 또한 1995년 결근, 생산성 감소 등 흡연으로 인한 근로자 1인당 연간 비용손실액을 약 3,100달러 정도로 추계하고 있다(보건복지부, 건강길라잡이).

우리나라에서도 2회에 걸쳐 연구가 진행되었는데 1991년에 발표된 논문에 의하면 담배로 인한 추가 의료비만 1년에 1조 2천억원이며 사망, 결근, 화재 등으로 인한 재정적인 손실이 1년에 1조원이 된다고 하였고, 그 이외 환경오염, 간접흡연으로 인한 손상 등을 합하면 가히 천문학적인 숫자가 될 것으로 보고 있다(이규식, 1991). 또한 1995년에 발표된 자료에서는 담배로 인한 추가 의료비가 연간 2조 2천 6백억원에 직간접 경제손실액이 3조 5천억원으로 추정하고 있다(이규식, 1995). 비만으로 인한 질병부담을 직접과 간접 사회경제적 비용을 추계한 연구를 보면 4,600억원이었으며, 공식적인 보건의료체계 내·외부에서 지출된 총 비만의 사회경제적 비용은 약 1조 17억원에 달한다(문옥륜, 2000).

Dalton 등은 “Health, Safety and Environment Program“에서 건강위험요인, 의료서비스 이용, 건강 관련비용 지출 등에서 개선을 이룰 수 있다고 밝혔으며, 이러한 건강증진 프로그램은 건강위험요인을 줄이려는 노력을 실시함으로써 장기간에 걸친 의료비 지출의 감소를 가져올 수 있다는 연구 결과이다(Daton, 1991).

### Ⅲ. 이론적 배경

#### 1. 건강위험요인

##### 가. 건강위험요인의 정의 및 목적

###### 1) 건강위험요인의 정의

건강위험요인은 건강에 영향을 미친다고 여겨지는 개인의 행태를 말하는데, 이러한 건강위험요인에 대한 자료는 어떤 특수한 집단의 건강과 관련된 독특한 정보를 제공하게 된다. 한 집단의 사망률, 이환율과 함께 자료를 분석하면 그 집단의 건강증진 또는 질병예방을 위하여 어떤 부문에 있어서 어떤 점을 강조할 것인지를 결정할 수 있다(조비룡 등, 1997).

그런데, 건강위험요인을 직접으로 측정하는 방법은 설문조사(Survey Research)밖에 없다. 물론 술이나 담배의 소비량 등으로 간접적으로 건강위험요인을 예측할 수는 있으나 특정 집단의 건강위험요인에 대한 정보는 제공하지 못한다. 특히 이러한 간접적인 측정법은 조사할 수 있는 항목 또한 한정되어 있다. 예를 들어 안전벨트, 신체활동 등은 간접적인 방법으로 추정하기 어렵다(Somer, 1986).

미국에서는 1981년 미국의 3개 주에서 처음으로 주민들의 건강위험요인에 대한 조사를 시작하였는데, 1984년부터는 Center for Disease Control(CDC)를 주축으로 하여 전화 설문조사법으로 매달 시행하고 있으며(Gentry, 1985),

우리나라에서는 1989년부터 한국보건사회연구원에서 개별 면담 설문조사법으로 매 3년마다 시행하고 있다.

## 2) 건강위험요인의 목적

건강위험요인의 평가의 목적은

첫째 개인의 건강습관과 위험요인을 파악하고, 둘째 개인의 사망위험 및 여러 가지 원인에 의한 건강결과를 계산하여 주며, 셋째 개인의 위험요인을 변화시키기 위한 상담과 교육 자료를 제공한다.

이것은 행동과학의 영역에서 유래되었다기보다 공중보건학적 측면에서 유래되었다. 건강위험요인 평가는 환자 개개인의 생활습관을 변화시키는데 도움을 주고자 개발되었지만, 개인이 직면한 위험요인을 비교적 정확하게 계산하여 제공하므로 다른 영역으로의 잠재적인 발전 가능성이 있다(백유진, 2002). 건강위험요인 평가는 건강증진 프로그램을 평가하는데 있어 기초 혹은 결과 측정으로 혹은 연구목적으로 사용하기도 한다(백유진, 2002). 사업장과 같이 특정한 건강문제를 지닌 인구집단에게 건강위험요인 평가를 적용하는 시도들이 관심을 끌었는데 건강위험요인의 감소, 건강한 생활습관으로의 개선, 직무태도의 향상, 보건관리 비용의 절감 등 여러 가지 효과가 있다고 한다(Gibbs, 1985).



## 나. 건강위험요인의 선정

건강위험요인의 선정에는 미국의 ‘질병관리센터 건강위험평가’ (Center for Disease Control/Health Risk Appraisal)(Julie, 1991), 성낙진의 건강행동평가(성낙진, 1992)가 있다. 현재 미국의 질병관리센터의 건강위험평가의 경우는 37가지의 항목을 사용하고 있으며, 카터 센터의 건강위험평가의 경우는 45가지 항목을 사용하고 있다.

## 다. 건강위험요인의 중요성

인구집단의 관점에서 볼 때 건강위험요인이 질병이나 사망에 기여하는 부분, 즉, 인구 기여위험도(Population Attributable Risk, PAR)의 크기가 건강위험요인의 중요성을 나타내는 최고의 수단이다(김성원, 2002).

PAR은 특정 질병에 의한 모든 사망 중에서 한 위험요인이 기여한 비율로 정의되며, 상대위험도와 위험요인의 유병률에 의해 계산된다. 한 연구에서는 특정 질병에 의한 총 사망 중 0.5%이상을 차지한 모든 건강위험요인을 포함시킬 것을 제안하였다( $PAR \geq 0.5\%$ ). PAR 접근법은 상대위험도가 매우 높지 않는 한 일부 직업적 노출처럼 강력하지만 드문 위험요인으로 라디오 수리공, 전기기술자 등이 있었지만 이러한 기준에 근거하여 건강위험요인에서 제외시켰다. 하지만 그러한 건강위험요인에 노출된 개인의 관점에서 본다면 오로지 상대위험도가 중요할 뿐이고, 아무리 드물다 할지라도 강력한 건강위험요인은 포함되어야 한다(김성원, 2002).

## 라. 건강위험요인과 질병의 연관성에 대한 증거의 질

Canadian Task Force on the Periodic Health Examination이 사용한 것과 유사한 기준이 사용될 수 있다. Williams 등은 건강위험요인으로 포함될 3가지 범주를 마련하였는데, 첫째는 2개 이상의 대규모 전향적 연구, 둘째는 하나의 전향적 연구 및 여러 개의 후향적 연구, 셋째는 여러 개의 좋은 후향적 연구 등이다. 국내에는 질 좋은 전향적, 후향적 코호트 연구가 매우 드물다. 따라서 일부 질환의 건강위험요인에서는 국내의 환자-대조군 연구의 자료를 이용하거나 아예 질 좋은 외국의 연구 자료를 이용하고 있다(김성원, 2002).

## 마. 건강위험요인 평가 도구의 개발과 국내 현황

기능적 건강 상태의 평가를 일상 진료에서 이용하려면 ‘일정 기간 동안의 기능적 건강 상태’가 중요한 문제가 된다. 그리고 건강위험요인 평가를 위한 독립적인 지표가 필요한데 신체적 기능, 정신적 기능, 의존성, 사회적 기능, 그리고 의료 서비스의 이용에 이르기까지 여러 가지가 있다. 예를 들어 환자들의 증상이나 불편함(통증, 어지럼증 등)이 환자에게 얼마나 괴로움을 주는가를 평가하는 것도 한가지 지표가 될 수 있다(신호철, 2003).

건강위험요인을 평가하는 도구들은 타당도 조사가 안되어 있고, 사용하기가 복잡하고 길어서 실용성에 문제가 있는 경우가 많다. 1980년대에 들어서 건강증진의 중요성이 부각되면서 긍정적인 건강의 지표를 찾으려는 노력들이 많이 있었고, 세계보건기구의 건강에 대한 개념이 널리 받아들여지면서 전반적인 안녕 상태, 정신적인 상태 등을 측정하려는 시도가 이루어져 왔다(신호철, 2003).

최근 일차의료영역에서 쉽게 사용할 수 있는 전반적인 건강위험요인의 평가 도구가 개발되고 있는데 대표적으로 Dartmouth COOP 차트(Nelson 등, 1987)를 들 수 있다. 이 도구는 세계가정의학회(WONCA)에서 채택하여 COOP/WONCA 차트로 수정되어 전 세계의 가정의들의 일상 진료에서 사용하도록 권장되고 있기도 하다. 특히, 우리나라의 경우에는 주로 외국에서 개발된 기능적 건강위험요인 평가 도구들을 번역해서 사용해 왔는데 이들 번역된 도구들이 타당도 조사가 제대로 안되어 있는 경우가 많고 그 내용도 우리나라 사람들의 건강 행동 변화를 나타내지 못하는 경우가 많아서 정확도 자체에 문제가 있는 실정이다(신호철, 2003).

## 바. 건강위험요인의 주요 활용분야

### 1) 특정 직업군

현재 여러 작업장에는 자체적인 건강증진 프로그램이 있지만, 건강위험요인에 근거한 평가는 아주 매력적인 도구가 될 수 있다. 그러나 일반적인 형태의 건강위험요인의 평가는 그러한 세팅에는 적합하지 않을 수 있다.

사업장 근로자를 대상으로 하는 건강위험요인의 평가의 경우 작업환경, 작업유형, 근무시간 및 과거 고용력 등과 같은 직업 관련 위험요인을 배제할 수 없으므로 일반적 형태의 건강위험요인의 평가도구를 확장한 판본을 개발하기보다는 근로자를 대상으로 한 건강위험요인 평가도구를 따로 개발하는 것이 좋을 것이다. 특정 직업군에서 특정 질환의 발병 위험성이 높다는 것은 잘 알려져 있다. 이러한 특정 위험인자는 건강위험요인 평가가 취급하는 일반적 위험을 훨씬 상회할 수 있으므로 반드시 고려해야 한다. 최근에는 특정 직업군의 사망률을 과거보다 쉽게 얻을 수 있지만, 아직까지도 불안정하고 얻을 수 없는 경우도 많다(백유진, 2002).

### 2) 만성 질환자 군

건강위험요인은 주로 건강하다고 간주되는 인구집단을 대상으로 하므로 만성질환자들의 경우에도 특정 직업군과 같은 문제가 제기된다. 그러나 만성질환자들은 기존 질환 외의 위험요인들을 감소시키기 위해 건강위험 평가의 도움을 받을 수 있고, 건강행위를 수정하도록 동기 부여를 할 수 있다. 이러한 경우에는 환자가 기존에 가지고 있는 질환과 관련된 원인에 대한 위험도 평가는 타당하지 않을 수 있다고 주의를 환기하는 것이 바람직

하다. 특정 질환에 대한 예후를 산정하는 것은 윤리적이고 방법론적인 문제를 야기하므로, 일반적 행태의 건강위험 평가에서 추구되어서는 곤란하다. 물론 특정 직업군의 경우와 마찬가지로 특정 질환(당뇨병 등)에 대한 특수한 건강위험요인 평가도구를 개발할 수도 있겠지만, 이것은 질병을 중시하는 전문가 집단에서 해야 할 일이다. 그와 같은 경우에도 보다 건강한 생활습관으로의 변화가 중요함을 강조하여야 한다(백유진, 2002).

### 3) 연령별 적합성

상대위험도와 회귀 상수는 중간층 연령에서 도출되는 경우가 많으므로 나이가 어리거나 너무 노령인 경우에는 적용하기 어려운 경우가 많다. 더욱이 건강증진의 측면에서 보자면, 여러 조건이나 위험인자의 선택, 위험도 평가가 예측하는 향후의 연한 간격대(time span), 그리고 사망률을 주로 다루는 도구의 특성과 관련하여 여러 가지 문제가 노정된다. 특정연령군에 대한 다양한 결과의 출력 형태와 연한 간격대를 통합한 형태의 건강위험요인 평가도구가 필요하다(백유진, 2002).

### 4) 기타 활용분야

- 일차 진료의의 정기적 건강평가의 통합적 구성요소
- 지역사회의학 예방활동 혹은 건강증진 프로그램의 평가도구의 구성요소
- 고용 전 혹은 고용초기 건강평가
- 관리의학(HMO 등) 체계에서 첫 등록 평가
- 보험 계약 전후 평가
- 개인이나 단체에 대한 건강위험요인 평가 무료 제공

- 비영리 혹은 영리 단체의 건강증진프로그램의 핵심요소
- 건강 교육이나 건강서비스 프로그램의 효능과 효과 판정
- 임상 연구의 평가 도구

## 2. 데이터 마이닝

### 가. 데이터마이닝의 출현배경

경영층의 의사결정에 도움을 주는 고급정보를 가공하고 축적하는데 관심을 가졌던 의사결정지원시스템(DSS : Decision Support System)은 정보화의 개념을 조직의 하부계층의 반복업무를 지원하는 자동화 업무에서 전사적인 개념으로 확장시키는 역할을 하였다. 그런데 이를 구축하기 위하여 선결과제가 발견되었는데 바로 전사인 시스템을 통합 관리하는 통합 데이터베이스의 구축이었다. 각 부서별로 독립적으로 운영되는 시스템으로는 경영층이 의사결정을 내리는데 별로 도움이 되지 않았던 것이다. 통합 데이터베이스 구축이 어느 정도 이루어졌을 때 발생한 또 다른 문제점은 방대한 데이터와 정보들 가운데서 찾고자 하는 정보를 정확하고 빠르게 찾는다는 것이 아주 힘들다는 점이다. 이러한 문제는 실제 비즈니스에서 더욱 개선의 필요성이 드러났고, 결국 이를 해결하고자 하는 노력으로 등장한 개념 중에 하나가 바로 “데이터마이닝(Data Mining)”이다.

## 나. 데이터마이닝의 정의

Mine이란 의미는 채광하다. 즉, 거대한 더미 속에서 가치 있는 무언가를 캐낸다는 것이다. 데이터마이닝에서는 방대한 양의 데이터 속에서 쉽게 드러나지 않는 유용한 정보를 찾아내는 과정이다.

Meta Group의 Aaron Zornes의 의하면, 데이터마이닝은 대용량 데이터 베이스에서 알려져 있지 않은 실행 가능한 정보를 추출하는 지식발견 과정이라고 할 수 있다(강현철 등, 2002).

“Data Mining is [a] knowledge discovery process of extracting previously unknown, actionable information from very large databases.”

- Aaron Zornes, The META Group

Gartner Group의 Erick Brethenoux는 통계 및 수학적 기법뿐만 아니라, 패턴 인식 기술을 이용하여 많은 양의 데이터에서 의미 있는 상관관계, 패턴 및 트렌드를 발견하는 과정이라고 보다 구체적으로 정의한다(강현철 등, 2002).

“Data Mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories. using pattern recognition technologies as well as statistical and mathematical techniques.”

- Robert Small. Two Crows



이를 종합하여 볼 때 데이터마이닝은 많은 양의 데이터 속에서 의미 있는 정보를 캐는 과정으로 이해하면 된다(강현철 등, 2002). Data Mining은 흔히 Knowledge Discovery in Database(정보발견)라고도 불리우며, 그 외에 Knowledge Extraction(지식추출), Information Harvesting(정보추수), Data Archeology(정보고고학), Data Pattern Processing(자료패턴처리) 등으로도 불리운다.

#### **다. KDD와 데이터마이닝**

KDD(knowledge discovery in database)는 데이터로부터 지식을 추출하는 전 과정을 설명하는 의미로 데이터마이닝은 KDD 과정에서 탐사 단계만을 뜻하는 의미로 사용한다고 제안되었다. KDD에 대한 공식적인 정의는 “데이터로부터 암시적이며 이전에 알려지지 않은 잠재적으로 유용한 지식을 추출하는 것”이며, 새로운 기술이기보다는 기계-학습, 통계학, 데이터베이스 기술, 전문가 시스템 그리고 데이터 가시화 등의 분야가 모두 기여를 한 학제적인 연구 분야에 속한다(용환승, 1998).

#### **라. 데이터마이닝의 분석기법**

데이터마이닝 기법에는 일반적으로 통계학에서 말하는 여러 가지 분석 기법들을 포함하여 연관성규칙(Association Rules), 군집분석(Cluster analysis), 의사결정나무분석(Decision tree analysis), 신경망분석(Neural network analysis)와 같은 분석기법이 있다.

### 1) 연관성 규칙(Association Rules)

어떤 특정 문제에 대해 아직은 일어나지 않은 답을 얻고자 하는 예측(Prediction)의 문제나 고객들을 특정목적에 따라 분류(Segmentation)하는 문제가 아니라, 상품 혹은 서비스(이하 상품)의 거래기록(Historical) 데이터로부터 상품간의 연관성 정도를 측정하여 관성이 많은 상품들을 그룹화하는 클러스터링(clustering)의 일종이다(강현철 등, 2002). 연관성 규칙은 하나의 거래나 사건에 포함되어 있는 둘 이상의 품목들의 상호관련성을 발견하는 것이다. 사전에 분석 방향이나 목적이 특별히 없을 경우, 대용량 데이터에 대한 비목적성 데이터마이닝(undirected data mining) 분석기법으로서 연관성 분석은 매우 유용하다(최종후 등, 2000). 일반적인 연관성 분석을 위한 데이터는, 동일한 거래임을 나타내는 변수와 거래에 포함되는 품목을 다른 하나의 변수로 생각하는 일반적인 변수·관찰치 형태의 데이터와는 다른 형태를 갖는다. 따라서 거래내용에 대한 데이터를 변환 없이 그 자체로 이용할 수 있는 간단한 자료구조를 갖는 분석방법이다(강현철 등, 2002).

### 2) 군집분석(Cluster analysis)

군집분석(Cluster analysis)은 주어진 관찰치 중에서 유사한 것들을 몇몇의 집단으로 그룹화하여, 각 집단의 성격을 파악함으로써 데이터 전체의 구조에 대한 이해를 돕고자 하는 분석방법이다(김기영 등, 1998). 대용량 데이터에서 개개의 관찰치를 요약하는 것보다는 전체를 유사한 관찰치들의 군집(cluster)으로 구분하여 복잡한 전체보다는 그를 잘 대표하는 군집들을 관찰함으로써 전체 데이터에 대한 의미 있는 정보를 얻어낼 수 있게 된다. 따라서 동일한 군집의 개체들은 유사한 성격을 갖도록 서로 다른 군집에

속한 개체들 사이에는 그와 반대로 상대적으로 서로 다른 성격을 갖도록 군집이 형성되어야 할 것이다(김기영 등, 1998).

여러 군집방법 중에 대용량 데이터를 빠르게 처리할 수 있으며, 그 알고리즘도 비교적 간단한 방법으로 k-평균 군집방법이 있다(강현철 등, 2002). 이는 관찰치들 사이의 거리를 이용해 주어진 기준을 최적화 하도록 구현되므로 최적분리 군집방법이라고 하며 이 방법은 계보적 군집방법의 단점을 극복할 수 있고, 관찰치의 수가 많을 때 주로 이용하므로 데이터마이닝에 유용한 방법이라고 말할 수 있다(최종후 등, 2000).

### 3) 의사결정나무분석(Decision tree analysis)

의사결정나무분석(Decision tree analysis)은 의사결정규칙(decision rule)을 나무구조로 도표화하여 분류(classification)와 예측(prediction)을 수행하는 분석방법이다. 이 방법은 분류 또는 예측의 과정이 나무구조에 의한 추론규칙(inductio rule)에 의해서 표현되기 때문에 다른 방법들(예를 들면 신경망, 판별분석, 회귀분석 등)에 비해서 연구자가 그 과정을 쉽게 이해하고 설명할 수 있는 장점이 있다(허명희, 1998).

데이터마이닝(data mining)에서의 의사결정나무는 탐색(exploration)과 모형화(modeling)라는 두 가지 특성을 모두 가지고 있다고 할 수 있다. 즉, 의사결정나무는 판별분석(discrimination analysis) 또는 회귀분석(regression analysis) 등과 같은 모수적(parametric) 모형을 분석하기 위해서 사전에 이상치(outlier)를 검색하거나 분석에 필요한 변수 또는 모형에 포함되어야 할 교호효과를 찾아내기 위해서 사용될 수 있고, 그 자체가 분류 또는 예측 모형으로 사용될 수도 있다(강현철 등, 2002)

#### 4) 신경망분석(Neural network analysis)

신경망 또는 인공신경망(artificial neural networks)에 관한 연구는 뇌 신경생리학(neurophysiology)으로부터 영감을 얻어 시작되었다. 신경망분석은 자료분석 분야에서 복잡한 구조를 가진 자료에서의 예측(prediction) 문제를 해결하기 위해서 사용되는 유연한 비선형모형(nonlinear models)의 하나로 분류될 수 있으며, 신경망은 은닉마디(hidden units)라고 불리는 독특한 구성요소에 의해서 일반적인 통계모형과 구별되어진다(김대수, 1992). 은닉마디는 인간의 신경세포를 모형화한 것으로써 각 은닉마디는 입력변수들의 결합(combination)을 수신하여 목표변수에 전달하며 결합에 사용되는 계수(coefficient)들을 연결강도(synaptic weights)라고 부르며, 활성화함수는 입력값을 변환하고 이를 입력으로 사용하는 다른 마디로 출력하게 된다(강현철 등, 2002).

신경망에는 여러 가지 다양한 모형이 있으나, 그 중에서도 자료분석을 위해 가장 널리 사용되는 모형은 MLP(Multilayer Perceptron, 다층인식자) 신경망이고, MLP는 종종 범용근사자(universal approximator)라고 불리기도 하는데 이는 이론적으로 적절한 활성화함수와 적당한 수의 은닉마디를 가지는 MLP는 모든 비선형 곡선을 매우 정확하게 근사시킬 수 있다는 것을 의미한다. MLP는 입력층(input layer) 은닉마디로 구성된 은닉층(Hidden layer), 출력층(output layer)으로 구성된 전방향(feed-forward) 신경망이다(강현철 등, 2002).

## 라. 순차패턴

순차패턴은 연관규칙에 시간의 변이를 추가한 지식의 한 형태로서 1995년 정의된 이후 학계와 산업계에 응용범위를 넓히고 있으며 최근에는 사용자가 웹 사이트에 접속해서 방문하는 경로(user web navigation path)를 발견하여 사이트 구성과 광고 배치를 결정하는 등 적용범위가 다양화되어 가고 있다.

순차패턴을 찾는 문제는 사용자의 구매 정보 데이터베이스가 주어졌을 때 사용자가 정의한 최소 지지도(minimal support)를 만족하는 모든 시퀀스(sequence)를 찾아냄으로써, 시간의 경과에 따라 발생하는 사건이나 트랜잭션의 빈번한 패턴을 찾아내는 것이다.

순차패턴 추출 알고리즘에서 사용하는 기본적인 용어와 전제 조건 및 표기 방법에 대해서 정리하면 다음과 같다.  $I = \{i_1, i_2, \dots, i_m\}$ 를 모든 항목(item)들을 모아 놓은 집합이라고 하면, 시퀀스는  $S = s_1 \rightarrow s_2 \rightarrow \dots \rightarrow s_k$ ,  $s_j \subseteq I$  이고,  $1 \leq j \leq l$  이다. 시퀀스의 원소(element)라 불리는  $s_j$  는 항목 집합을 나타내며, 편의상  $s_j = (x_1 x_2 \dots x_m)$ ,  $x_k \in I$ ,  $1 \leq k \leq m$ 로 정의된다. 한 항목은 시퀀스의 한 원소에 최대 한번 발생할 수 있으며, 시퀀스의 여러 원소에 중복해서 발생할 수 있다. 그리고, 시퀀스의 길이는 그 시퀀스에 발생한 항목들의 수를 말하며  $l$ - 시퀀스로 표현된다.

주어진 시퀀스에서 마지막 원소를 제거해서 얻은 시퀀스를 해당 시퀀스의 Head라고 하고, 마지막 원소만으로 이루어진 시퀀스를 Tail이라고 한다. 지지도는 주어진 시퀀스 집합에 주어진 하나의 시퀀스가 포함되어 있는 원소 수를 뜻하고 Lift는 Head가 Tail에게 영향을 주는 수치를 계산한

것으로 다음과 같이 표현 된다.

$$\text{Lift}(S) = \text{Support}(S) / ( \text{Support}(\text{Head}(S)) * \text{Support}(\text{Tail}(S)) )$$

Lift는 1 이상인 경우 양의 상관관계가 존재하는 것을 의미한다.

예를 들어, 주어진 시퀀스 집합이 다음과 같이 주어졌다고 하면,

$$s1 = (1, 3, 6) \rightarrow (2, 4) \rightarrow (4)$$

$$s2 = (1, 6) \rightarrow (2, 4) \rightarrow (4)$$

$$s3 = (1, 2, 6) \rightarrow (4, 6)$$

이 경우 주어진 시퀀스  $(1, 6) \rightarrow (4) \rightarrow (4)$  는 두개의 시퀀스에 나타나므로  $2/3 = 67\%$ 의 지지도를 갖는다. 이 경우 Lift는  $(1, 6) \rightarrow (4)$ 가 세개의 시퀀스에 나타나므로, Lift는 100% 이다. 또 S1의 Head는  $(1, 3, 6) \rightarrow (2, 4)$ 이고 Tail은  $(4)$ 이다. 순차패턴은 한 대상자에 해당하는 시퀀스를 입력받아 준비된 시퀀스 집합에서 일치하는 패턴을 추출하고, 대상자의 다음 시퀀스를 예측하는 시스템이다. 시퀀스 집합내의 시퀀스는 전체 데이터에 대한 정보를 가지고 있으므로 적절한 예측을 할 수 있다. 실제 처리 방법은 준비된 시퀀스 집합내의 일치하는 시퀀스 중 가장 큰 지지도가 해당 질환이 나타날 가능성이 있고, 이와 독립적으로 일치된 패턴 중 Lift가 가장 높은 시퀀스가 주요 요인이 된다.

순차패턴의 주요 연구로는 Sequence pattern(Agrawal, 1995), Episode(Mannila, 1996), GSP(Srikang, 1996), SPIRIT(Garofalakis, 1999), Cyclic Association(Ozden, 1998), Temporal relation rule(이용준, 2001) 등 많은 연구들이 있으며, 고객 순차 구매 패턴, 의약 처방 관리, 자연재해(지진), 과학 및 실험데이터 분석, 웹 접근패턴 분석, DNA 시퀀스 분석 등 여러 분야에서 다양하게 응용하여 유용한 지식을 추출할 수 있다.

## 마. 데이터마이닝의 활용분야

데이터마이닝은 정의의 다양성만큼 활용분야도 매우 다양하다고 할 수 있다. 특히 기업의 의사결정 문제에서 많이 활용되고 있는데, 주요 활용분야를 나열하면 다음과 같다.

(Internet site [[http://hodori.inje.ac.kr/%7Ehcphm/index4\\_1.htm](http://hodori.inje.ac.kr/%7Ehcphm/index4_1.htm)] )

### 1) 환자 이탈방지 시스템 구축

병원의 경우는 일반 소매업과 달리 고객의 특정병원에 대한 선호도가 병원선택에 있어 보다 중요한 요소가 될 것이며, 제공하는 서비스는 소매업에 비해 매우 이질적인 서비스를 제공하는 공급원임과 동시에 인간의 생명과 직접적으로 연관되어 있는 서비스이므로 어느 특정 병원에 대한 신뢰가 높게 이루어지면 다른 병원으로 이탈하는 유인은 상당히 줄어들 수 있다. 따라서 고객의 병원이용률에 영향을 미치는 요인을 찾아내어 이를 마케팅 전략에 적극 활용한다면 환자의 이탈방지 시스템을 구축하는 근원이 될 것이다. 환자 이탈방지 시스템의 분석과정은 우선 고객의 이탈 현황을 파악한 수와 이탈환자의 특성을 데이터마이닝 과정을 통해 파악하여 이탈할 가능성이 높은 환자군을 구분시켜 예비 이탈환자에 대한 적극적인 마케팅 전략을 수립하여 체계적인 대응방안을 모색하는 방법으로 요약될 수 있다. 이와 같이 환자 이탈의 현상을 이해하고 이를 막기 위한 적절한 활동을 전개함으로써 기존 고객으로부터 발생하는 수익을 극대화시키는 것이 이탈방지 시스템의 주 목적인 것이다.

## 2) 특정환자 타겟 마케팅(Target Marketing)

특정 고객에 대한 타겟 마케팅 전략은 고객의 패턴을 모형화하여 실제 고객 마케팅업무에 반영하고 고객이 선호하는 서비스를 면밀히 조사하여 표적 고객집단을 선정하고 고객의 계층별 선호도를 이용하여 좀더 효과적이고 전략적인 시장세분화 전략을 수립하고 주제별 데이터마트(Data Mart)를 구성하여 이를 이용한 OLAP시스템을 구축함으로써 지역별, 진료과목별 현황이나 심지어 담당의사나 환자의 정보등 다양한 정보를 자유롭게 조회할 수 있는 시스템의 구축으로 이루어진다. 이와 같은 과정을 통해 얻어진 데이터를 이용하여 수량화, 최적척도변환 등을 이용한 대응분석 등을 통해 고객의 선호를 파악하고 우선 순위를 설정하여 우선 순위가 높은 고객에 대해서는 각 개인이 선호하는 서비스(특정 질환 예방검사, 예방 접종 등)를 적극 홍보함으로써 고객의 병원이용률을 유지시키고 우선 순위가 낮은 고객의 유지 및 신수요의 창출을 위해서는 좀더 차별화된 서비스를 개발하기 위한 투자 등 보다 다양한 고객 서비스 전략을 구상해야 한다. 또한 자병원에 국한된 정보뿐만 아니라 다른 병원의 정보가 서로 연계되는 경우 이러한 전략은 보다 큰 효과를 기대할 수 있을 것이다. 즉, 이를 통하여 자병원의 특성에 맞는 환자를 유치할 경우 환자 만족 증대는 물론 더 많은 수익을 가져다 줄 수 있기 때문에 고객과 잠재고객의 파악은 각 병원으로 하여금 보다 적극적인 마케팅 전략을 수립 및 실행시킬 수 있도록 할 것이다.



### 3) 고객의 신용관리

신용카드나 전자화폐 및 직불카드 등의 지불수단이 일반화되고 거의 모든 의료관련 지불체계의 통합전산화가 이루어지게 되면 각 병원은 고객의 신용관리를 직접 담당할 수 있게 되며 이로 인해 불량 신용거래를 판단하거나 예측할 수 있는 시스템이 구축되어야 할 것이다. 또한 건강보험을 포함한 사회보험 및 자동차보험과 같은 사(射)보험의 정보 또한 공유되면 지금까지 보험자들이 담당해 왔던 징수에 관한 사후관리업무 및 보험사고 방지 업무를 일선 병원에서 담당할 수 있게 될 것인데 이 경우 고객에 대한 신용관리는 필수 불가결한 업무가 될 것이다. 그 경우 데이터 마이닝에 의한 고객의 신용관리는 큰 역할을 담당할 것이다.

### 4) 질병의 진단 및 예방의료체계 수립

모든 환자의 인구통계학적, 환경적, 계층별, 지역별 등의 일반적 자료와 각 환자의 과거 진료실적이나 유전적 형질의 특성과 같은 개인적 자료가 공유되는 경우 데이터마이닝이 적용된다면 질병의 자동진단체계 및 예방의료체계는 보다 쉽게 수립될 것이며, 이는 바로 잠재적인 환자의 확보에 가장 중요한 전략이 될 것이다. 또한 병원은 이를 통해 각 병원의 특성에 적합한 의료서비스 상품을 개발하거나 연계상품을 발굴할 수 있을 것이다.

## IV. 연구방법

### 1. 연구의 틀

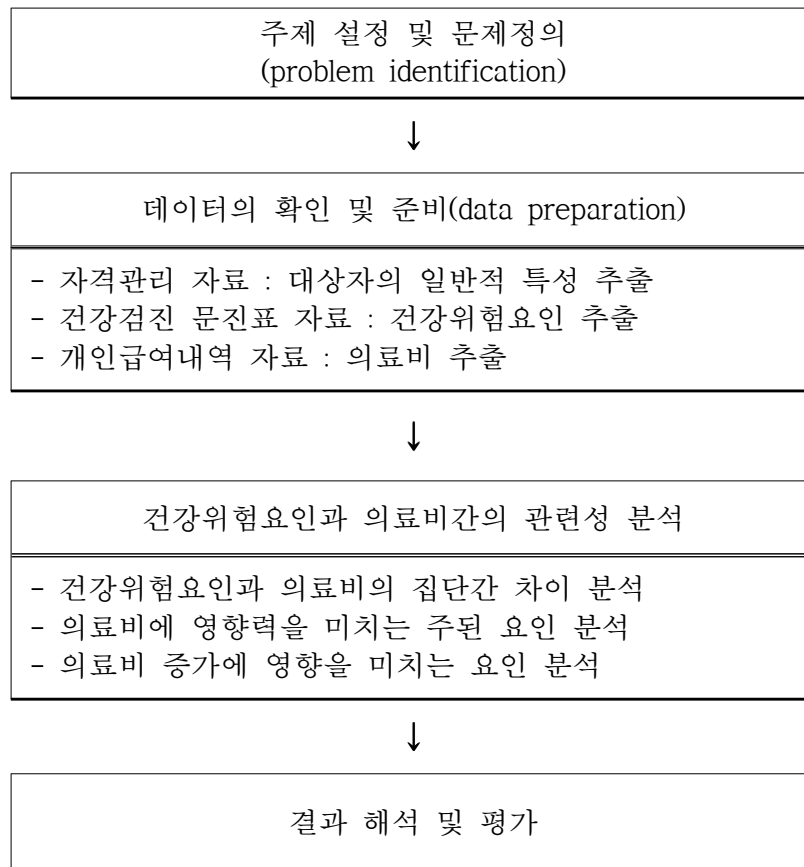


그림 1. 연구의 틀

## 2. 연구자료

본 연구는 국민건강보험공단의 1998년, 2000년, 2002년 3년 동안 계속적으로 건강검진을 받은 대상자 657,295건 중에서 연구에 필요한 변수를 추출한 것으로 대상자의 일반적 특성, 건강위험요인, 의료비 요인의 변수를 포함하고 있다.

대상자의 일반적 특성은 대상자의 성별, 연령을 알 수 있는 변수이고, 건강생활습관은 건강검진결과 문진표 내역 중에서 일부를 추출한 것으로 흡연습관, 음주습관, 운동습관, 식습관을 알 수 있는 변수이고, 의료비는 건강검진을 받은 대상자 중에서 고혈압(I10-15), 당뇨병(E10-14), 고지혈증(E78)으로 3년간 지출한 의료비용으로 요양기관에서 건강보험환자 진료에 소요된 공단부담금과 환자본인부담금을 합한 금액이며, 요양기관에서 청구한 진료비 중 심사·결정된 진료비를 말한다.

표 1. 연구자료

변수	세부내용	측정수준
대상자 일반적 특성	성별	1)남 2)여
	연령	1)20대 2)30대 3)40대 4)50대 5)60대 6)70대
건강위험요인	흡연습관	1)한다 2)안한다
	음주습관	1)한다 2)안한다
	운동습관	1)한다 2)안한다
	식습관	1)규칙적이다 2)불규칙적이다
의료비 증가	의료비	1)1998년 2)2000년 3)2002년

### 3. 변수의 선정

#### 가. 건강위험요인

건강위험요인은 연구 대상자 각각에 대해 1998년, 2000년, 2002년 건강진단 문진표 자료를 이용하여 다음과 같은 변수별 내용을 조사하였다.

흡연요인은 “비흡연자”와 “흡연자”로 구분하고 흡연자는 1일 흡연량을 조사하였다.

음주요인은 “비음주”와 “음주자”로 구분하고 음주자는 주간 음주횟수를 기준으로 조사하였다.

규칙적인 운동 요인은 “한다”와 “안한다”로 구분하고 안한다는 주간 운동횟수를 기준으로 조사하였다.

식습관 요인은 “불규칙군”과 “규칙군”으로 구분하고 불규칙군은 1일 3회 먹기는 하나 식습관이 불규칙적일 경우 불규칙군으로 구분하였다.

## 나. 의료비 요인

의료비는 건강검진결과에 따른 고혈압(I10-15), 당뇨병환(E10-14), 고지혈증(E78)으로 지출한 의료비로 공단부담액과 본인부담액을 포함한 금액을 적용하였다. 국민건강보험공단 외적 의료비의 건강보험 진료비에 포함되지 않는 식대, 지정 진료비, 병실료 차액, MRI 등 보험급여가 되지 않는 검사료 등은 제외하였다.

의료비는 4월 1일을 기준으로 하여 1998년부터 2002년 사이의 의료수가 인상률을 감안하여 할인하여 적용하였으며, 의료수가 인상지수는 표 2와 같다.

표 2. 의료수가 인상 지수 (단위 : %)

인상시점	인상률	인상지수
1998		100
1999. 11. 15	9.0	109
2000. 4. 1	6.0	115.54
7. 1	9.2	126.17
9. 15	6.5	134.37
2001. 1. 1	7.08	143.88
2002. 4. 1	-2.9	139.71

## 4. 분석방법

### 가. 분석모형 및 분석항목

본 연구에서의 분석모형은 첫째, 4가지 건강위험요인과 의료비의 차이를 각 기간별로 구분하여 연도별 총 의료비의 평균을 비교하였다.

둘째, 다변량 분석으로 연도별 총 의료비를 종속변수로 4가지 건강위험요인의 기간을 독립변수로 하는 다중회귀 분석을 실시하여 건강위험요인의 기간에 따라서 의료비 증가에 영향을 미치는 주된 요인을 분석하였다.

셋째, 데이터마이닝 중 순차패턴을 이용하여 4가지 건강위험요인과 기간에 따라서 의료비 증가에 영향력을 미치는 요인을 분석하였다.



그림 2. 분석모형

분석항목은 표 3과 같이 설명변수와 목표변수로 나누었다. 설명변수의 성별은 남, 여로 구분하고, 연령은 30대에서 60대까지 4개의 구간으로 구분하였다. 목표변수는 흡연, 음주, 식습관, 운동의 기간에 따라서 1년, 2년, 3년으로 구분하였다.

목표변수인 의료비는 1998년, 2000년, 2002년 3개년 동안의 건강검진결과 데이터가 모두 존재하는 집단중에서 고혈압(I10-15), 당뇨병(E10-14), 고지혈증(E78)으로 지출된 비용을 합한 금액이다.

표 3. 분석항목

변수	세부내용	측정수준
설명변수	성별	1)남 2)여
	연령	1)30대 2)40대 3)50대 4)60대
	흡연	1)1년간흡연 2)2년간흡연 3)3년간흡연
	음주	1)1년간음주 2)2년간음주 3)3년간음주
	식습관	1)1년간불규칙 2)2년간불규칙 3)3년간불규칙
	운동	1)1년간비운동 2)2년간비운동 3)3년간비운동
목표변수	의료비	1998, 2000, 2002년 모두 데이터가 존재하는 집단중에서 고혈압, 고지혈증, 당뇨병에 지출된 비용

## 나. 분석방법

분석방법의 도구는 SPSS 10.0과 SAS Enterprise Miner 4.0을 이용하여 건강위험요인에 따른 의료비 지출의 크기를 분석하기 위해서 분산분석(analysis of variance : ANOVA), 회귀분석(regression analysis), 연관성규칙(Association Rules) 중 순차패턴(Sequential Patterns)의 방법을 사용하였다.

분석방법은 첫째, 4가지 건강위험요인과 기간에 따라 의료비 차이를 분석기 위해 분산분석을 이용하였다. 4가지의 건강위험요인에 따라서 기간을 1년 지속군, 2년 지속군, 3년 지속군으로 나누어 집단별로 건강위험요인에 따른 의료비 차이를 비교하였다. 4가지의 건강위험요인을 독립변수로 하고, 의료비를 종속변수로 하여 실시하였으며, 일원 분산분석(one-way ANOVA)를 실시하였다.

둘째, 회귀분석을 이용하여 4가지 건강위험요인과 의료비의 증가에 따른 변화를 분석하였으며, 4가지 건강위험요인들 중 의료비에 가장 영향력을 많이 미치는 요인을 추출했다.

셋째, 4가지 건강위험요인과 기간에 따른 의료비 증가를 분석하기 위해 순차패턴(Sequence Pattern)을 이용하여 시간의 경과에 따라 다수의 위험요소가 나타나는 패턴을 찾아내었다. 4가지의 건강위험요인과 의료비의 ‘증가’와 ‘감소’를 예측하기 위해서 대상집단을 4가지의 건강위험요인에 따라서 구분한 후 각각의 대상 집단에 대해 Sequence Pattern을 추출하고 Support/Confidence/Lift에 따라 Sequence Pattern을 선택하였다.



## V. 연구결과

### 1. 연구대상자의 일반적 특성

#### 가. 연구대상자의 성별, 연령별 비교

연구 대상자의 성별, 연령별 비교는 표 4와 같다. 연구대상자는 233,542명으로 남자는 165,440(70.8%), 여자는 68,102(29.2%)의 분포를 나타냈다.

연령별 분석한 결과 1998년 당시 20대가 4.1%, 30대가 33.7%, 40대가 39.2%, 50대가 19.5%, 60대가 3.4%, 70대가 0.2%를 차지한다. 이는 전체 연령비와 비교할 때 30대와 40대가 비교적 많은 비율을 나타낸다.

표 4. 연령군별 연구 대상자 (단위: 명, %)

연령대	남자	여자	계
20~29세	5,939(2.5)	3,560(1.5)	9,499(4.1)
30~39세	48,314(20.7)	30,341(13.0)	78,655(33.7)
40~49세	66,827(28.6)	24,744(10.6)	91,571(39.2)
50~59세	37,650(16.1)	7,861(3.4)	45,511(19.5)
60~69세	6,486(2.8)	1,371(0.6)	7,857(3.4)
70세이상	224(0.1)	225(0.1)	449(0.2)
계	165,440(70.8)	68,102(29.2)	233,542(100.0)

### 나. 연구 대상자의 연령별 의료비 분포도

연구 대상자 233,542명이 1998년, 2000년, 2002년 동안의 고혈압(I10-15), 당뇨병(E10-14), 고지혈증(E78)으로 지출한 연령별 의료비 지출현황은 표 5와 같다.

연령별로 3년간의 의료비 평균을 비교한 결과, 20대는 347.3원, 30대는 1,580.7원, 40대는 8,486.4원, 50대는 28,780.5원, 60대는 54,936.0원, 70대는 84,618.2원이었다. 시간의 경과에 따라서 의료비 변화에 차이가 있었으며, 연령대별로 연령의 증가에 따라서 평균 의료비의 증가를 보이고 있다.

표 5. 연령별에 따른 의료비 현황 (단위: 원)

변수	1998년	2000년	2002년	계
20~29세	134.2	373.8	534.0	347.3
30~39세	890.4	1,374.4	2,477.4	1,580.7
40~49세	4,570.9	7,625.1	13,263.2	8,486.4
50~59세	17,469.0	26,830.6	42,041.9	28,780.5
60~69세	36,058.0	51,134.4	77,615.5	54,936.0
70세 이상	61,408.2	84,031.1	108,415.3	84,618.2

#### 다. 연구대상자의 건강위험요인 특성 비교

연구대상자의 건강위험요인 특성 비교는 표 6과 같다. 흡연의 경우 흡연을 안하는 경우 54.0%, 흡연을 하는 경우가 46.0%로 나타났으며, 흡연을 하는 경우에는 3년 동안 흡연을 하는 경우가 43.3%로 통계적으로 유의한 차이가 있었다( $p < 0.001$ ). 음주의 경우 음주를 안하는 경우 35.4%, 음주를 하는 경우가 64.6%로 나타났으며, 음주를 하는 경우에는 3년 동안 음주를 하는 경우가 55.4%로 통계적으로 유의한 차이가 있었다( $p < 0.001$ ). 운동의 경우 운동을 하는 경우 45.7%, 운동을 안하는 경우가 54.3%로 나타났으며, 운동을 안하는 경우에는 3년 동안 운동을 안하는 경우가 38.4%로 통계적으로 유의한 차이가 있었다( $p < 0.001$ ). 식습관의 경우 식습관이 규칙적일 경우 11.6%, 식습관이 불규칙적일 경우 88.4%로 나타났으며, 식습관이 불규칙적인 경우에는 1년 동안 불규칙적인 경우가 53.1%로 통계적으로 유의한 차이가 있었다( $p < 0.001$ ).

표 6. 연구대상자의 건강위험요인 특성 비교

(단위: 원, %)

변수	남자	여자	계	$\chi^2$ 값
<b>흡연유무</b>				
흡연안함	58,341(25.0)	67,795(29.0)	126,136(54.0)	***
흡연함				
1년	3,250(1.4)	168(0.1)	3,418(1.5)	
2년	2,746(1.2)	23(0.0)	2,769( 1.2)	
3년	101,103(43.3)	116(0.0)	101,219(43.3)	
<b>음주유무</b>				
음주안함	29,675(12.7)	53,068(22.7)	82,743( 35.4)	***
음주함				
1년	5,418(2.3)	4,278(1.8)	9,696(4.2)	
2년	8,406(3.6)	3,210(1.4)	11,616(5.0)	
3년	121,941(52.2)	7,546(3.2)	129,487(55.4)	
<b>운동유무</b>				
운동함	96,975(41.5)	9,694(4.2)	106,669(45.7)	***
운동안함				
1년	1,525(0.6)	3,042(1.3)	4,567(1.9)	
2년	13,100(5.6)	5,739(2.5)	18,839(8.1)	
3년	40,115(17.2)	49,627(21.2)	89,742(38.4)	
<b>식습관유무</b>				
규칙적	14,866(6.4)	12,257(5.2)	27,123(11.6)	***
불규칙적				
1년	88,125(37.7)	35,900(15.4)	124,025(53.1)	
2년	18,948(8.1)	6,817(2.9)	25,765(11.0)	
3년	43,501(18.6)	13,128(5.6)	56,629(24.2)	

\*p<0.05    \*\*p<0.01    \*\*\*p<0.001

#### 다. 건강위험요인과 의료비의 연간 추이 분석

건강위험요인과 의료비의 연간 추이 분석은 표 7과 같다. 흡연의 경우 흡연을 안할 경우 1998년도에 비해 2002년도에 9,194.6원(26.1%) 증가하였으며, 흡연을 하는 경우 1998년도에 비해 2002년도에 흡연기간이 1년일 경우 14,636원(29.9%), 2년일 경우 10,048.0원(30.8%), 3년일 경우 11,390.4원(34.4%)으로 증가하였다.

음주의 경우 음주를 안한 경우 1998년도에 비해 2002년도에 8,877.6원(25.7%) 증가하였으며, 음주를 하는 경우 1998년도에 비해 2002년도에 음주기간이 1년일 경우 8,137.9원(27.3%), 2년일 경우 6,871.2원(22.7%), 3년일 경우 11,563.0원(29.7%)으로 증가하였다.

운동의 경우 운동을 할 경우 1998년도에 비해 2002년도에 12,733.3원(29.7%) 증가하였으며, 운동을 안한 경우 1998년도에 비해 2002년도에 운동을 안한 기간이 1년일 경우 10,619.5원(28.4%), 2년일 경우 8,163.3원(27.8%), 3년일 경우 7,572.4원(30.5%)으로 증가하였다.

식습관의 경우 식습관이 규칙적일 경우 1998년도에 비해 2002년도에 11,286.5원(27.7%) 증가하였으며, 식습관이 불규칙한 경우 1998년도에 비해 2002년도에 식습관 불규칙 기간이 1년일 경우 10,783.3원(28.6%), 2년일 경우 8,928.2원(30.1%), 3년일 경우 9,129.5원(34.4%)으로 증가하였다.

표 7. 건강위험요인과 의료비의 연간 추이 분석

(단위: 원, %)

	1998년	2000년	2002년	계	F값/x <sup>2</sup> 값
<b>흡연유무</b>					
흡연안함	7,503.3(21.3)	11,038.4(31.3)	16,697.9(47.4)	11,746.5	16.83***
흡연함					
1년	9,365.3(19.1)	15,566.6(31.8)	24,001.3(49.0)	16,311.1	
2년	6,383.2(19.5)	9,853.6(30.2)	16,431.2(50.3)	10,889.3	
3년	5,924.4(17.9)	9,856.3(29.8)	17,314.8(52.3)	11,031.8	
<b>음주유무</b>					
음주안함	7,422.4(21.5)	10,817.2(31.3)	16,300.0(47.2)	11,513.2	5.88**
음주함					
1년	5,864.3(19.7)	9,972.6(33.4)	14,002.2(46.9)	9,946.4	
2년	7,348.5(24.2)	8,764.1(28.9)	14,219.7(46.9)	10,110.8	
3년	6,482.6(18.4)	10,633.7(30.2)	18,045.6(51.3)	11,720.6	
<b>운동유무</b>					
운동함	8,562.9(19.9)	13,121.0(30.5)	21,340.2(49.6)	14,341.4	73.75***
운동안함					
1년	7,547.9(20.2)	11,638.7(31.2)	18,167.4(48.6)	12,451.3	
2년	6,048.6(20.6)	9,153.4(31.1)	14,211.9(48.3)	9,804.6	
3년	4,795.7(19.3)	7,638.9(30.8)	12,368.1(49.9)	8,267.6	
<b>식습관유무</b>					
규칙적	8,491.7(20.8)	12,513.7(30.7)	19,778.2(48.5)	13,594.5	42.87***
불규칙적					
1년	7,604.0(20.2)	11,733.1(31.1)	18,387.3(48.7)	12,574.8	
2년	5,959.8(20.1)	8,832.5(29.8)	14,888.0(50.2)	9,893.4	
3년	4,747.1(17.9)	7,916.4(29.8)	13,876.6(52.3)	8,846.7	

\*p<0.05    \*\*p<0.01    \*\*\*p<0.001

## 2. 건강위험요인 기간에 따른 의료비 평균 비교

### 가. 건강위험요인 기간에 따른 의료비 평균 비교 - 성별

건강위험요인 기간에 따른 의료비 평균 비교는 표 8과 같다. 남자의 경우 4가지의 건강위험요인 중 음주기간이 1년과 2년 집단을 비교했을 때 14,593.0원으로 의료비 지출에 높은 평균차가 있었다. 흡연기간은 1년과 2년 집단을 비교했을 때 13,767.7원, 비운동기간은 1년과 2년 집단을 비교했을 때 12,961.6원, 식습관(불규칙)기간은 1년과 3년 집단을 비교했을 때 27,68.1원으로 의료비 지출이 높은 평균차가 있었다.

여자의 경우 4가지의 건강위험요인 중 식습관(불규칙)기간이 1년과 2년 집단을 비교했을 때 77,102.1원으로 의료비 지출에 높은 평균차가 있었다. 흡연기간은 1년과 2년 집단을 비교했을 때 -5,537.2원, 음주기간은 1년과 2년 집단을 비교했을 때 46,515.5원, 비운동기간은 1년과 3년 집단을 비교했을 때 -34,803.4원으로 의료비 지출이 높은 평균차가 있었다.

표 8. 건강위험요인 기간에 따른 의료비 평균 지출 분석 - 성별

(단위: 원)

변수 (I)	(J)	평균차(I-J)	표준오차	95% 신뢰구간		F값/t²값
<b>남자</b>						
흡연기간						
1년	2년	13,767.7	9,154.60	-8,092.09	35,627.44	1.46
	3년	9,169.0	5,881.00	-4,873.89	23,211.93	
	2년	4,598.7	7,243.29	-12,697.20	21,894.52	
음주기간						
1년	2년	14,593.0	7,810.18	-4,056.53	33,242.48	2.72
	3년	2,764.0	6,080.04	-11,754.23	17,282.18	
	2년	11,829.0	5,231.61	-663.28	24,321.28	
비운동기간						
1년	2년	12,961.6	3,768.15	3,963.78	21,959.32	6.18*
	3년	4,181.4	3,024.72	-3,041.16	11,403.95	
	2년	8,780.2	3,204.82	1,127.54	16,432.76	
식습관(불규칙)기간						
1년	2년	-1,205.2	3,720.05	-10,088.12	7,677.67	0.79
	3년	2,768.1	2,653.98	-3,569.19	9,105.38	
	2년	-3,973.3	3,801.72	-13,051.25	5,104.61	
<b>여자</b>						
흡연기간						
1년	2년	-33,233.2	94,865.91	-263,154.40	196,687.93	0.18
	3년	-27,696.1	50,574.34	-150,270.27	94,878.11	
	2년	-5,537.2	98,298.85	-243,778.56	232,704.24	
음주기간						
1년	2년	-17,372.0	71,367.39	-190,341.08	155,597.18	0.81
	3년	-63,887.5	51,687.52	-189,159.58	61,384.68	
	2년	46,515.5	68,243.13	-118,881.55	211,912.53	
비운동기간						
1년	2년	-127,005.9	115,187.60	-406,179.60	152,167.80	0.83
	3년	-34,803.4	92,867.21	-259,880.43	190,273.64	
	2년	-92,202.5	77,270.19	-279,477.92	95,072.91	
식습관(불규칙)기간						
1년	2년	23,515.0	78,292.65	-166,238.48	213,268.49	1.23
	3년	77,102.1	49,579.60	-43,061.22	197,265.38	
	2년	-53,587.1	77,834.07	-242,229.13	135,054.98	

\*p<0.05    \*\*p<0.01    \*\*\*p<0.001



#### 다. 건강위험요인 기간에 따른 의료비 평균 분석 - 연령별

집단별 의료비 지출 평균 분석은 부록의 표 1~4과 같으며, 40대에서 음주기간과 비운동기간이 통계적으로 유의한 차이가 있었다( $p < 0.001$ ).

30대의 경우 비운동기간이 1년과 2년 집단을 비교했을 때 1,851.4원, 40대의 경우 흡연기간이 1년 2년 집단을 비교했을 때 18,543.4원, 50대의 경우 음주기간이 1년과 2년 집단을 비교했을 때 34,722.7원, 60대의 경우 흡연기간이 1년과 2년 집단을 비교했을 때 90,950.0원으로 의료비 지출이 높은 평균차가 있었다.

### 3. 회귀분석 결과

#### 가. 상관관계 분석 결과 - 성별

의료비와 건강위험요인간의 Pearson 상관관계는 표 9와 같다. 남, 여 모두 의료비와 음주기간이 가장 큰 양의 상관관계( $r^2=0.010, 0.220$ )를 갖는데, 이러한 상관관계는 변환 이후에도 그대로 유지된다.

남자의 경우 종속변수와 독립변수들간의 상관성은 음주기간이 양의 상관관계를 식습관기간, 흡연기간, 운동기간이 음의 상관관계를 보였다. 여자의 경우 음주기간, 흡연기간이 양의 상관관계를 식습관기간과 운동기간이 음의 상관관계가 있었으며, 통계학적으로 유의하였다( $p < 0.001$ ).

독립변수들 사이의 상관관계는 남자의 경우 식습관기간은 음주기간, 흡연기간, 운동기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ). 음주기간은 식습관기간, 흡연기간, 운동기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ). 흡연기간은 식습관기간, 운동기간, 운동기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ). 운동기간은 식습관기간, 음주기간, 흡연기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ).

여자의 경우 식습관기간은 음주기간, 흡연기간, 운동기간이 양의 상관관계를, 음주기간은 식습관기간, 흡연기간, 운동기간이 양의 상관관계를 보였으며, 흡연기간은 식습관기간, 운동기간 양의 상관관계를, 운동기간이 음의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ). 운동기간은 식습관기간, 음주기간이 양의 상관관계를, 흡연기간이 음의 상관관계를 보였다.

표 9. 건강위험요인 기간과 의료비간의 상관관계 분석 - 성별

변수	상관계수	의료비	식습기간	음주기간	흡연기간	운동기간
남자	의료비		-.005	.003***	-.006***	-.004***
	식습기간	-.005		.063	.077***	.081*
	음주기간	.003	.063***		.246	.010***
	흡연기간	-.006	.077***	.246***		.090
	운동기간	-.004	.081***	.010*	.090***	
여자	의료비		-.145	.117***	.053***	-.021
	식습기간	-.145		.278	.321***	.092
	음주기간	.117	.278**		.557	.031
	흡연기간	.053	.321***	.557***		-.045
	운동기간	-.021	.092	.031	-.045	

\*p<0.05    \*\*p<0.01    \*\*\*p<0.001

## 나. 상관관계 분석 결과 - 연령별

연령별로는 30대의 경우 운동기간이 40대의 경우 식습관기간, 음주기간이 의료비의 가장 큰 양의 상관관계를 갖는데, 이러한 상관관계는 변환 이후에도 그대로 유지된다.

30대의 경우 종속변수와 독립변수들간의 상관성은 운동기간이 양의 상관관계를 식습관기간, 음주기간, 흡연기간이 음의 상관관계를 보였으며, 40대의 경우 식습관기간, 음주기간이 양의 상관관계를 흡연기간, 운동기간이 음의 상관관계를 보였으며, 50대의 경우 식습관기간, 음주기간, 흡연기간, 운동기간 모두 양의 상관관계를, 60대의 경우 음주기간이 양의 상관관계를 식습관기간, 흡연기간, 운동기간이 음의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ).

30대의 경우 식습관기간은 음주기간, 흡연기간, 운동기간이 양의 상관관계를, 음주기간은 식습관기간, 흡연기간이 양의 상관관계를, 운동기간은 양의 상관관계를, 흡연기간은 식습관기간, 운동기간, 흡연기간이 양의 상관관계를 보였으며, 운동기간은 식습관기간, 흡연기간이 양의 상관관계를, 음주기간이 음의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ).

40대의 경우 식습관기간은 음주기간, 흡연기간, 운동기간이 양의 상관관계를, 음주기간은 식습관기간, 흡연기간, 운동기간이 양의 상관관계를, 흡연기간은 식습관기간, 운동기간, 흡연기간이 양의 상관관계를, 운동기간은 식습관기간, 음주기간, 흡연기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ).

50대의 경우 식습관기간은 음주기간, 흡연기간, 운동기간이 양의 상관관

계를, 음주기간은 식습관기간, 흡연기간, 운동기간이 양의 상관관계, 흡연기간은 식습관기간, 운동기간, 흡연기간이 양의 상관관계를, 운동기간은 식습관기간, 음주기간, 흡연기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ).

60대의 경우 식습관기간은 음주기간, 흡연기간, 운동기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.05$ ). 음주기간은 식습관기간, 흡연기간, 운동기간이 양의 상관관계를 보였으며, 통계학적으로 유의하였다( $p < 0.001$ ). 흡연기간은 식습관기간, 운동기간, 흡연기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.001$ ). 운동기간은 식습관기간, 음주기간, 흡연기간이 양의 상관관계를 보였으며 통계학적으로 유의하였다( $p < 0.01$ ).

표 10. 건강위험요인 기간과 의료비간의 상관관계 분석 - 연령별

변수	상관계수	의료비	식습기간	음주기간	흡연기간	운동기간
30대	의료비		-.004	-.001	-.003***	.013***
	식습기간	-.004		.058***	.080***	.065
	음주기간	-.001	.058***		.231	-.003***
	흡연기간	-.003	.080***	.231***		.048
	운동기간	.013	.065***	-.003	.048***	
40대	의료비		.010	.014	-.006	-.019**
	식습기간	.010*		.065***	.078***	.110***
	음주기간	.014	.065***		.247***	.010
	흡연기간	-.006	.078***	.247***		.102***
	운동기간	-.019**	.110***	.010	.102***	
50대	의료비		.012	.002	.000	.007
	식습기간	.012		.064***	.068***	.079***
	음주기간	.002	.064***		.313***	.020
	흡연기간	.000	.068***	.313***		.108***
	운동기간	.007	.079***	.020*	.108***	
60대	의료비		-.026	.006	-.003	-.050
	식습기간	-.026		.060*	.032	.065*
	음주기간	.006	.060*		.260***	.013
	흡연기간	-.003	.032	.260***		.083***
	운동기간	-.050	.065*	.013	.083**	

\*p<0.05    \*\*p<0.01    \*\*\*p<0.001

## 나. 회귀 분석 결과 - 성별

1998년, 2000년, 2002년 3년간의 평균의료비를 종속변수로 하고, 4가지 생활습관 요인을 독립변수로 적합한 다중회귀의 표준화 회귀계수는 표 11과 같다. 남자의 경우는 음주기간이 증가할수록, 여자의 경우는 흡연기간, 음주기간이 증가할수록 의료비 지출이 그렇지 않은 경우보다 상대적으로 증가함을 보여주고 있다.

표준화계수의 절대값의 크기로부터 각 회귀변수들의 상대적 중요도를 짐작할 수 있는데 남자의 경우 음주기간, 운동기간, 식습관기간, 흡연기간 순이었고, 여자의 경우는 음주기간, 흡연기간, 운동기간, 식습관기간 순이었다.

표 11. 건강위험요인 기간과 의료비간의 회귀분석 - 성별

변수		표준화계수 (Beta)	95% 신뢰구간		t값
남자	흡연기간	-.007	-9,387.615	1,835.440	-1.319
	음주기간	.006	-2,566.742	8,311.906	1.035
	운동기간	-.003	-3,658.935	2,184.219	-.495
	식습관기간	-.005	-3,866.259	1,368.489	-.935
여자	흡연기간	.029	-53,165.423	68,710.793	.253
	음주기간	.156	-18,551.552	104,608.891	1.385
	운동기간	-.007	-84,874.925	79,063.996	-.070
	식습관기간	-.197	-104,743.106	-142.545	-1.987*

\*p<0.05    \*\*p<0.01    \*\*\*p<0.001

## 나. 회귀 분석 결과 - 연령별

30대의 경우는 식습관기간이 증가할수록, 40대의 경우 흡연기간, 음주기간이 증가할수록, 50대의 경우 흡연기간, 음주기간, 식습관기간이 증가할수록, 60대의 경우 음주기간, 운동기간이 증가할수록 의료비 지출이 그렇지 않은 경우보다 상대적으로 증가한다고 말할 수 있다.

표준화계수의 절대값의 크기로부터 각 회귀변수들의 상대적 중요도를 짐작할 수 있는데 30대의 경우 식습관기간, 음주기간, 운동기간, 흡연기간 순이었고, 40대의 경우 음주기간, 흡연기간, 운동기간, 식습관기간 순이었고, 50대의 경우 흡연기간, 식습관기간, 음주기간, 운동기간 순이었고, 60대의 경우 음주기간, 운동기간, 흡연기간, 식습관기간 순이었다.



표 12. 건강위험요인 기간과 의료비간의 회귀분석 - 연령별

변수	표준화계수 (Beta)	95% 신뢰구간		t값	
30대	흡연기간	-.004	-2,265.775	1,361.144	-.489
	음주기간	.000	-3,931.064	3,819.138	-.028
	운동기간	-.003	-4,612.904	3,306.532	-.323
	식습관기간	.013	-509.535	3,483.401	1.460
40대	흡연기간	.012	-926.058	6,479.269	1.470
	음주기간	.016	-180.663	14,910.811	1.913
	운동기간	-.009	-12,222.072	3,746.031	-1.040
	식습관기간	-.019	-9,368.190	-966.149	-2.411*
50대	흡연기간	.012	-4,266.005	13,565.815	1.022
	음주기간	.002	-16,439.790	18,933.923	.138
	운동기간	-.002	-19,159.537	15,542.522	-.204
	식습관기간	.006	-6,991.643	12,424.772	.549
60대	흡연기간	-.023	-43,444.858	19,904.940	-.729
	음주기간	.008	-50,299.118	64,935.589	.249
	운동기간	-.001	-56,471.772	55,017.958	-.026
	식습관기간	-.048	-56,863.828	7,352.099	-1.513

\*p<0.05    \*\*p<0.01    \*\*\*p<0.001

## 4. 순차패턴 분석 결과

### 가. 순차패턴 분석 결과 - 성별

지지도와 신뢰도가 각각에서 높은 값을 가지는 연관규칙을 표 13에 정리하였다. 남자의 경우 음주기간이 3년, 여자의 경우 음주기간이 3년, 흡연기간이 2년, 흡연기간이 3년이 향상도가 높은 연관규칙으로서, 지지도와 신뢰도에서도 비교적 높은 값을 나타내고 있다.

남자의 경우 “음주 →음주 →음주 & 의료비증가” 는 전체대상자 중 음주를 3년동안 지속적으로 하고 있는 경우에 의료비 증가율은 106%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 6% 상승시킴)이며, 음주를 3년동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 83.95%였다.

여자의 경우 “음주 →음주 →음주 & 의료비증가” 는 전체대상자 중 음주를 3년동안 지속적으로 하고 있는 경우에 의료비 증가율은 138%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 38% 상승시킴)이며, 음주를 3년동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 100.0%였다.

표 13. 건강위험요인 기간과 의료비간의 순차패턴 분석 - 성별

(단위: %)

변수	향상도	지지도	신뢰도	연 관 규 칙
남자	1.03	13.75	81.98	음주 → 음주 → 의료비증가
	1.06	13.23	83.95	음주 → 음주 → 음주 & 의료비증가
	1.04	12.74	83.63	흡연 → 흡연 → 의료비증가
	1.05	12.47	83.63	흡연 → 흡연 → 흡연 & 의료비증가
여자	1.15	45.55	83.33	음주 → 음주 → 의료비증가
	1.38	36.36	100.00	음주 → 음주 → 음주 & 의료비증가
	1.38	36.36	100.00	흡연 → 흡연 → 의료비증가
	1.38	36.36	100.00	흡연 → 흡연 → 흡연 & 의료비증가

## 나. 순차패턴 분석 결과 - 연령별

30대의 경우 “음주 →음주 → 의료비증가” 는 전체대상자 중 음주를 2년 동안 지속적으로 하고 있는 경우에 의료비 증가율은 104%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 4% 상승시킴)이며, 음주를 2년동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 78.08%였다.

40대의 경우 “음주 →음주 →음주 & 의료비증가” 는 전체대상자 중 음주를 3년동안 지속적으로 하고 있는 경우에 의료비 증가율은 104%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 4% 상승시킴)이며, 음주를 3년동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 82.81%였다.

50대의 경우 “음주 →음주 →음주 & 의료비증가” 는 전체대상자 중 음주를 3년동안 지속적으로 하고 있는 경우에 의료비 증가율은 103%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 3% 상승시킴)이며, 음주를 3년동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 83.38%였다.

60대의 경우 “음주 →음주 → 의료비증가” 는 전체대상자 중 음주를 2년 동안 지속적으로 하고 있는 경우에 의료비 증가율은 105%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 5% 상승시킴)이며, 음주를 3년동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 81.54%였다.

표 14. 건강위험요인 기간과 의료비간의 순차패턴 분석 - 연령별

(단위: %)

변수	향상도	지지도	신뢰도	연 관 규 칙
30~39세	1.04	15.16	78.08	음주 → 음주 → 의료비증가
	1.01	18.62	76.09	음주 → 음주 → 음주 & 의료비증가
	1.03	18.35	77.53	흡연 → 흡연 → 의료비증가
	1.02	17.29	76.47	흡연 → 흡연 → 흡연 & 의료비증가
40~49세	1.03	18.51	81.88	음주 → 음주 → 의료비증가
	1.04	17.91	82.81	음주 → 음주 → 음주 & 의료비증가
	1.03	17.91	81.94	흡연 → 흡연 → 의료비증가
	1.04	17.45	82.73	흡연 → 흡연 → 흡연 & 의료비증가
50~59세	1.01	19.55	81.69	음주 → 음주 → 의료비증가
	1.03	18.61	83.38	음주 → 음주 → 음주 & 의료비증가
	1.01	18.00	81.40	흡연 → 흡연 → 의료비증가
	1.02	17.67	82.13	흡연 → 흡연 → 흡연 & 의료비증가
60~69세	1.05	19.41	81.54	음주 → 음주 → 의료비증가
	1.04	19.41	80.30	음주 → 음주 → 음주 & 의료비증가
	1.05	19.05	81.25	흡연 → 흡연 → 의료비증가
	1.05	19.05	81.25	흡연 → 흡연 → 흡연 & 의료비증가

## 5. 활용방안

본 연구는 건강위험요인과 의료비지출 크기를 비교한 연구로서 다음과 같은 활용방안을 제시하고자 한다.

첫째, 건강위험요인의 기간에 따라서 의료비 지출이 증가하는 대상자에 대하여 개인의 건강위험요인을 변화시키기 위한 상담과 교육자료를 제시할 수 있으며, 건강위험요인이 증가하는 대상자들에 대해서 건강증진목표를 설정하고 범국민적인 건강생활실천운동을 전개할 수 있다. 미국의 경우 'Healthy People 2000', 'Healthy People 2010', 영국의 경우 'The Health of the Nation(2000년)', 'Our Healthier Nation(2010년)', 일본의 경우 '활력있는 인생 80 건강계획(1988~1997년)', '건강일본 21(1998~2010년)' 으로 선진국에서 이미 1980년부터 실시해오고 있다.

둘째, 의료비 증가에 영향을 미치는 건강위험요인을 우선순위로 선정하여 유병률이 높아지는 요인, 사회경제적 부담을 야기하는 요인, 최근 10년간 의료비 증가추세에 영향을 미치는 요인, 예방사업의 비용·효과가 큰 요인 등으로 구분하여 의료비 증가에 가장 영향력을 미치는 요인을 중심으로 국민의 건강증진을 위한 종합적·체계적 관리를 한다면 의료비 지출도 감소하게 될 것이다. 예를 들면 영양, 운동, 정신건강 전문가 활용을 통한 건강증진 프로그램 확대, 운동상해 및 성인병 예방프로그램 개발·적용·보급 선도 등이다.

셋째, 건강위험요인이 증가하는 대상자에 대하여 다른 보건사업과의 연계사업을 실시 할 수 있다. 국민건강보험공단의 건강검진과 연결되어 건강위험요인 관리를 적극적으로 할 수 있을 것이다. 또한 정기적인 검진의 경우 건강위험요인을 같이 활용하고 추적 관찰에서도 이용 할 수 있다. 예방적 차원의 환자 관리에 대해 건강위험요인의 감소에 대해 동기를 유발할 수 있으며, 현재 계속적으로 증가되고 있는 만성질환자들에 대해서 건강위험요인은 중요한 역할을 할 것이다.

넷째, 본 연구의 결과를 Electronic Health Record(EHR)와의 연계를 통해 개인의 의료 정보에 대해 효율적인 개입과 응용이 가능한 상태가 될 것이다. EHR을 통하여서 건강위험요인과 개인급여자료들을 연관시켜 이들의 상관관계를 통하여서 건강과 다른 지표들과의 연관성을 이용하여 요인의 증가를 감소시킬 수 있을 것이다.

## VI. 고 찰

### 1. 연구자료 및 방법에 대한 고찰

연구 대상자는 국민건강보험공단이 매 2년마다 실시하는 건강검진 수검자 중 3년동안 계속적으로 검진을 받은 대상자를 분석자료로 이용하였으며, 이에 다음과 같은 제약점이 존재하였다.

첫째, 본 연구에서 이용한 4가지 건강위험요인은 공단에서 실시하는 건강진단에서 얻어진 건강검진자료의 문진표로 당시 자기기입식으로 작성한 문진표를 기준으로 각 변수들의 값을 취하였고, 4가지 건강위험요인의 응답 결과에 대한 타당도를 검증하지는 못하였다. 이 자료는 공단에서 1998년도 건강진단의 문진표에 의한 건강행위와 과거 질병 이환력 및 현재 건강상태를 계속적으로 조사해 오고 있었으므로 본 연구에 수검자들은 연령으로 볼 때 1998년부터 3차례 작성해 본 경험이 있는 사람들인 관계로 응답결과에 대한 신뢰도는 인정할 수 있을 것으로 본다.

둘째, 의료비 자료는 공단의 개인급여내역 자료로부터 수집하였고, 이 자료는 비교적 객관적인 기준에 따른 자료이다. 즉, 영양급여기준과 진료수가기준에 의하여 체계적으로 분류된 지침에 의하여 의료비 관련 변수들이 산정되었기 때문에 객관성과 신뢰성을 가질 수 있고, 또한 공단은 이 자료



를 계속적으로 데이터웨어하우스로 관리하므로 업무적으로 효율성 극대화는 물론 보건 분야의 연구에서도 많은 역할을 할 수 있는 중요한 의료비 관련 자료로 볼 수 있다.

셋째, 건강위험요인의 기간을 분석함에 있어 수집한 자료의 측정횟수의 제한성이 있었으나 이 자료는 건강위험요인들의 변화를 단 3회의 문진표로 측정하였기 때문에 측정한 사이의 기간이나 측정 이후의 기간 동안의 건강습관의 변화를 알 수 없는 단점이 있다.

넷째, 건강과 질병에 관련된 변수가 매우 다양한 반면, 본 연구에서는 공단의 자료를 이용했기 때문에 가용한 자료의 제한성이 있었다. 즉, 스트레스 등 정신적 요인, 만성질환상태, 복용하는 약물 등의 의학적 요소 및 정확한 소득수준, 교육정도 등의 사회경제학적인 요소들을 연구에 포함시킬 수 없었기 때문에 연구결과에 교란 변수들로 작용할 수 있는 이들 변수들의 영향을 충분히 고려하지 못한 단점이 있었으나 건강위험요인과 의료비 지출의 차이를 연구한 Dalton(1991), Bertera(1991), Yen(1992), MacKenzie(1994), 이규식(1991) 등도 기업체·공단의 신체검사 및 건강진단 자료를 활용한 연구 결과들이다. 미국과 캐나다 등 선진국에서는 건강위험요인과 의료비 지출의 크기 비교 연구를 위하여 의료보험이나 생명보험의 건강진단 자료를 수집하여 역학적 연구에 활용하고 있고, 우리나라에서도 최근에 계속적으로 공단의 건강검진자료를 가지고 건강위험평가도구 및 건강위험요인과 질병에 연관성에 대한 많은 연구들을 하고 있다. 이들 자료로부터 얻어진 연구결과의 신뢰도를 높이기 위해서는 공단의 건강검진에 사용되는 문진표

의 건강위험 평가도구의 표준화와 세분화할 필요가 있을 것으로 생각된다.

다섯째, 의료비 범위가 국민건강보험공단에서 추계하는 공식적인 의료비용 중에서 병·의원에 국한된 의료비를 이용하였기 때문에 약국이나 한의원 그리고 비공식적인 건강보조식품이나 일반 매약에 대한 의료비는 합산되지 못하였다. 현행 우리나라 건강보험제도는 비급여 대상을 정하고 있고 또한 정확한 진단을 위하여 사용되는 MRI 등 첨단 진단장비에 의한 진료비는 건강보험 진료비에 포함되지 않으므로 전액 환자가 부담해야 되는 등 보험외적 의료비가 제도화 내지는 관례화되어 있다.

여섯째, 의료비용 자료를 누적한 기간이 총 36개월로 비교적 긴 기간이지만, 건강위험인자들이 질병을 일으키거나 건강상의 문제를 야기하여 의료비용의 변화를 유발하는 시점이 비교적 장시간 이후일 것이라고 가정되기 때문에 보다 더 장기간의 추적관찰이 필요하다고 하겠다. 이 논문의 결과로 미루어 향후 기대되는 연구로서는 보다 장기간의 관찰을 통한 건강위험요인과 의료비용에 대한 영향 및 건강위험인자들의 적극적인 교정을 시도하는 경우 의료비용에 미칠 수 있는 영향, 사회경제적 요소들이 의료비용에 미치는 영향, 공식적인 의료비용 이외의 비공식적인 의료비용과 건강위험요인요인 및 만성질환과의 연관성을 들 수 있겠다.

## 2. 연구결과에 대한 고찰

본 연구의 주된 연구목표는 통계학적 추정절차에 따라 각각의 건강위험요인별로 의료비의 지출의 크기를 성별, 연령별로 구분하여 상대적 크기를 추정하는데 있다. 건강위험요인별로 의료비 지출 크기를 산출하기 위해서는 의료비의 산출이 필요하고, 의료비 산출을 하기 위해서는 건강위험요인에서 산출된 의료비를 비교함으로써 건강위험요인별 의료비의 상대적 크기를 산출할 수 있다.

건강위험요인과 의료비와의 관련성을 알아보기 위한 모형에서 4가지의 건강위험요인을 가지고 있는 경우 그렇지 않는 경우보다 의료비 지출이 많았다. 즉, 흡연이나 음주가 향후에 만성질환 발병에 직접적인 의료비 지출의 원인이라고는 말할 수 없지만, 본 연구모형에서 얻어진 결과는 국내·외의 문헌조사를 하면 건강위험요인이 의료비 지출에 직간접 경제적 손실이 있는 것으로 나타났다.

98년 건강위험요인들과 36개월간의 의료비를 누적하여 관련성을 살펴본 결과 연령이 증가할수록 유의하게 의료비용이 증가하였으며, 연령은 의료비용의 가장 큰 연관성을 나타냈다. 남자가 여자보다 유의하게 의료비가 높았고 시간의 경과에 따라서도 의료비가 높게 나타났다. 1992년 미국의 Surgeon General은 흡연자는 비흡연자에 비하여 평생 6,000달러의 의료비를 더 지출한다고 보고하여 흡연자가 평생 사용하는 의료비가 비흡연자보다 많다고 하였다(MacKenzie, 1994). 1997년 네덜란드에서 조사한 바에 의

하면 흡연자와 비흡연자의 의료비를 연령에 따라 조사한 결과, 1인당 의료비는 연령이 증가함에 따라 증가하며, 전 연령층에서 흡연자의 의료비가 비흡연자의 의료비가 높았다. 그러나 흡연자 전체의 의료비는 연령이 많아짐에 따라 증가하다가 70세 이후에는 급격히 감소하여 70세 이후에는 전체 비흡연자의 의료비가 전체 흡연자의 의료비보다 월등히 많음을 추산하였다 (Barendregt, 1997)

남자의 경우 흡연기간이 1년과 2년이 높은 평균차(14,593.0원)를 보였으며, 여자의 경우 식습관(불규칙)기간이 1년과 3년이 높은 평균차(23,515.0원)를 보였으며, 연령별로는 60대의 흡연기간이 1년과 3년이 높은 평균차(90,950.0원)를 보였다. 즉, 남자의 경우 흡연에서 발병전 위험요인을 갖지 않은 비흡연 환자에 비해 위험요인을 갖고 있는 흡연 환자의 의료비 지출이 많았다. 담배를 피우는 것이 질병이 발생되었을 때 의료비 지출의 직접적인 원인이라고 생각함은 무리다. 흡연과 일생동안의 의료이용에 있어서도 각 연령층에서 흡연자가 비흡연자보다 의료이용이 많고 기대수명도 단축되는 것으로 나타났으며(Hodgson, 1992), 7가지 생활습관 요인 중 흡연이 가장 높은 의료비 초과 지출 요인인 것으로 연구되었으며 현재 흡연자가 비흡연자보다 연간 의료비를 960달러 초과지출 되었다고 한다(Betera, 1991).

결론적으로 건강위험요인 유무에 따른 의료비 지출의 크기는 요인의 기간이 증가함에 따라서 의료비 지출이 많았다. 이러한 결과는 건강위험요인의 직접적인 원인이라고 해석하는 것은 무리가 있으나 본 연구결과는 선행 연구자들의 연구결과와 일치되는 연구결과를 얻었다. 따라서 본 연구는 우

리나라의 직장근로자, 지역주민들에게 확대 적용 할 수 있을 것으로 판단되고, 나아가 보건 분야에서도 건강증진사업의 일환으로 활용할 수 있을 것으로 판단된다.

## Ⅶ. 결 론

건강위험요인에 따른 의료비 크기를 비교하기 위해 각 건강위험요인 여부에 따른 의료비 지출의 크기를 분석하였다. 연구자료는 국민건강보험관리공단 피보험자 233,452명이 1998년 1월 1일부터 2002년 12월 31일사이 에 받은 건강검진 문진표와 3개 질병군으로 진료한 내용을 비교한 연구로서 주요결과는 다음과 같다.

1. 건강위험요인의 기간과 의료비의 연간 추이 분석을 보면 건강위험요인 의 기간이 증가할수록 의료비가 증가하였고 통계적으로 유의한 차이가 있었다.
2. 남자의 경우 음주기간이 1년과 2년 집단을 비교할 때 14,593.0원으로 높은 평균차가 있었고, 여자의 경우 식습관(불규칙)기간이 1년과 2년 집단을 비교할 때 77,102.1원으로 의료비 지출에 높은 평균차가 있었다. 연령별 로는 60대의 경우 흡연기간이 1년과 2년 집단을 비교할 때 90,950.0원으로 의료비 지출에 높은 평균차가 있었다.
3. 남자의 경우 음주기간이 증가할수록 여자의 경우 흡연기간, 음주기간이 증가할수록 의료비 지출이 증가하였고, 연령별로는 40대가 음주기간이 증가할수록 의료비가 증가하였다.

4. 남자의 경우 전체대상자 중 음주를 3년 동안 지속적으로 하고 있는 경우에 의료비가 증가될 비율은 106%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 6% 상승시킴)이며, 음주를 3년 동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 83.95%였다. 여자의 경우 전체대상자 중 음주를 3년동안 지속적으로 하고 있는 경우에 의료비가 증가될 비율은 138%(지속적인 음주를 하고 있는 사람이 그렇지 않은 사람보다 의료비 증가율을 38% 상승시킴)이며, 음주를 3년동안 지속적으로 하고 있는 사람들 중 의료비가 증가될 확률은 100.0%였다.

이 연구에서는 건강위험요인의 기간과 의료비 지출에 차이가 있는 것으로 나타났으며, 이상의 결과에서 건강위험요인이 의료비 지출에 영향을 준다는 것을 확인할 수 있었다.

## 참고문헌

국민건강보험공단. 2003년 건강보험공단 통계연보, 2003

강현철, 한상태, 최종후, 김은석, 김미경. SAS Enterprise Miner 4.0을 이용한 데이터마이닝-방법론 및 활용. 자유아카데미, 2002.

김기영, 전명식. SAS 군집분석. 자유아카데미, 1998

김대수. 신경망 이론과 응용. 하이테크정보, 1992

김성원. 한국형 건강위험평가 도구의 개발과정. 가정의학회지 2002; 23(11): 268-275

조비용, 허봉렬. 건강위험요인에 대한 전화설문조사의 한국에서의 방법론적 고찰. 가정의학회지 1997; 1055(10): 1-9

문옥륜. 한국인 비만의 역학적 특성에 따른 비만관리 전략개발연구. 건강증진 보고서. 보건복지부, 2000

백유진. 건강위험평가의 정의. 가정의학회지 2002; 23(11): 260-267

보건복지부. 1993년도 국민영양조사 보고서, 1995



보건복지부. 2003년 보건복지부 통계연보, 2004

보건복지부. homepage 건강길라잡이, <http://healthguide.kihasa.re.kr/>

신호철. 건강 상태 평가. 가정의학회지 2003; 24(4):339-348

성낙진. 종합검진환자의 건강행동평가. 가정의학회지 1992; 13(6):541-51

이규식, “The Consequences of Smoking in Korea“, Institute of Population & Health Service Research at Yonsei University, 1991

이규식, “The Consequences of Smoking in Korea“, Institute of Population & Health Service Research at Yonsei University, 1995

용환승. 데이터 마이닝. 그린, 1998

최종후, 강현철, 한상태, 김은석, 김미경. SAS Enterprise Miner 4.0을 이용한 데이터마이닝-기능과 사용법. 자유아카데미, 2000.

통계청. 2003년 사망원인통계결과, 2004

허명희. SAS 분산분석. 자유아카데미, 1988

Barendregt JJ, Bonneux L, van der Maas PJ. The health care costs of smoking. *New Eng J Med* 1997; 337:1052-7

Benowitz NL. Tobacco, In Lee Goldman, Dennis Ausiello editors, *Cecil Textbook of Medicine*, 22nd edition, Philadelphia; Saunders, 2004; 57-61

Bertera RB. The effects of behavioral risks on absenteeism and health-care costs in the workplace. *J Occup Med*. 1991 Nov; 33(11): 1919-24

Dalton BA, Harris JS. A comprehensive approach to corporate health management. *J Occup Med*. 1991 Mar; 33(3): 338-47

Gentry EM, Kalsbeek WD, Hogelin GC, Jones JT, Gaines KL, Forman MR et al. The behavioral risk factor surveys: II. Design, Methods, and estimates from combined state data. *Am J Prev Med*, 1985; Nov-Dec, 1(6): 9-14

Gibbs JO. Worksite health promotion. Five year trend in employee health care cost. *JOM* 1985; 28: 217-25

Hodgson TA. Cigarette smoking and lifetime medical expenditures. *The Milbank Quarterly*, 1992; Vol.70(1): 81-125

Julie AG, Besty F. Comparing the Predictive accuracy of Health risk appraisal: The Centers for disease control versus Carter center program. *Am J Pub Health* 1991; 81(10): 1296-301.

MacKenzie TD, Bartecchi CE, Schrier RW. The human costs of tobacco use—second of two parts. *New Eng J Med* 1994; 330:975-980.

McGinnis, J. M. Foege, W. H. Actual Causes of Death in the United States. *JAMA*, Vol.270, No.18, 1993

O'Donnell M. Health Promotion. An Emerging Strategy for Health Enhancement and Business Cost Savings in Korea(Unpulished). 1999

PG O'Connor. Alcohol Abuse and Dependence, In Lee Gold-man, Dennis Ausiello editors, *Cecil Textbook of Medicine*, 22nd edition Philadelphia: Saunders, 2004; 74-81

Somer AR, Weisfeld VD. Individual behavior and health. *Public Health and preventive medicine: Appleton-Century-Crofts*, Norwalk, CT, 1986; 983-997

US Center for Disease Control and Prevention: Increasing physical activity: A Report on recommendations of the Task Force on Community Preventive Services. *MMWR* 2001; 50(RR 18): 1-18

U.S. DHHS. The Surgeon General's Report on Nutrition and Health, U.S. Government Printing Office. Washington, D.C. 1988

Wolfe, B. L., Health Status and Medical Expenditures: Is There A Link, *Soc. Sci. Med*, 1986; pp.993--999

Yen LT, Dee WE, Pamela W. Prediction of prospective medical claims and absenteeism costs for 1284 hourly workers from a manufacturing company. J Occup Med; 1992, Apr; 34(4): 428~35

## 【부 록】

표 1. 건강위험요인 기간에 따른 의료비 평균 지출 분석 - 30대

(단위 : 원)

(I)	(J)	평균차(I-J)	표준오차	95% 신뢰구간		F값/x <sup>2</sup> 값
흡연기간						
1년	2년	-4,224.88	6,339.38	-19,363.64	10,913.87	0.36
	3년	-22.48	4,208.75	-10,073.19	10,028.24	
	2년	-4,202.40	4,899.91	-15,903.63	7,498.82	
음주기간						
1년	2년	-4,644.23	5,526.92	-17,842.79	8,554.34	0.46
	3년	-1,594.27	4,451.85	-12,225.50	9,036.96	
	2년	-3,049.95	3,508.30	-11,427.95	5,328.05	
비운동기간						
1년	2년	-938.88	2,474.72	-6,848.64	4,970.88	1.03
	3년	-2,790.25	2,067.79	-7,728.25	2,147.74	
	2년	1,851.38	2,124.04	-3,220.94	6,923.70	
식습관(불규칙)기간						
1년	2년	-4,935.17	2,542.02	-11,005.63	1,135.29	2.62
	3년	649.18	1,839.99	-3,744.80	5,043.16	
	2년	-5,584.35	2,486.79	-11,522.92	354.22	

\* p<0.05    \*\*p<0.01    \*\*\*p<0.001

표 2. 건강위험요인 기간에 따른 의료비 평균 지출 분석 - 40대

(단위 : 원)

(I)	(J)	평균차(I-J)	표준오차	95% 신뢰구간		F값/x <sup>2</sup> 값
흡연기간						
1년	2년	18,543.38	12,846.69	-12,905.05	49,991.82	1.09
	3년	9,508.94	8,389.33	-11,027.97	30,045.85	
	2년	9,034.44	10,054.29	-15,578.26	33,647.14	
음주기간						
1년	2년	13,352.41	10,964.66	-13,488.84	40,193.66	3.47*
	3년	-6,111.11	8,319.97	-26,478.23	14,256.01	
	2년	19,463.52	7,590.26	882.72	38,044.31	
비운동기간						
1년	2년	17,762.45	5,486.28	4,332.15	31,192.75	5.97**
	3년	12,546.63	4,356.64	1,881.66	23,211.61	
	2년	5,215.82	4,580.66	-5,997.55	16,429.19	
식습관(불규칙)기간						
1년	2년	-5,624.02	5,310.02	-18,622.85	7,374.81	1.01
	3년	-4,579.09	3,748.56	-13,755.50	4,597.32	
	2년	-1,044.93	5,501.98	-14,513.66	12,423.79	

\* p<0.05    \*\*p<0.01    \*\*\*p<0.001

표 3. 건강위험요인 기간에 따른 의료비 평균 지출 분석 - 50대

(단위 : 원)

(I)	(J)	평균차(I-J)	표준오차	95% 신뢰구간		F값/x <sup>2</sup> 값
흡연기간						
1년	2년	-6,789.98	30,727.12	-82,016.82	68,436.85	0.03
	3년	-486.47	17,398.25	-43,081.26	42,108.31	
	2년	-6,303.51	26,056.51	-70,095.64	57,488.63	
음주기간						
1년	2년	34,722.70	25,153.95	-26,859.77	96,305.17	1.17
	3년	9,957.18	19,004.86	-36,570.95	56,485.32	
	2년	24,765.52	17,605.81	-18,337.42	67,868.46	
비운동기간						
1년	2년	9,722.75	13,055.91	-22,241.03	41,686.53	0.74
	3년	-4,082.17	9,990.81	-28,541.91	20,377.57	
	2년	13,804.92	11,361.35	-14,010.19	41,620.04	
식습관(불규칙)기간						
1년	2년	-8,047.14	12,669.08	-39,063.86	22,969.59	0.61
	3년	-9,471.73	9,064.48	-31,663.59	12,720.13	
	2년	1,424.59	13,276.99	-31,080.44	33,929.63	

\* p<0.05    \*\*p<0.01    \*\*\*p<0.001

표 4. 건강위험요인 기간에 따른 의료비 평균 지출 분석 - 60대

(단위 : 원)

(I)	(J)	평균차(I-J)	표준오차	95% 신뢰구간		F값/x <sup>2</sup> 값
흡연기간						
1년	2년	90,950.02	92,806.68	-136,563.48	318,463.52	0.50
	3년	24,217.92	57,668.11	-117,154.15	165,589.99	
	2년	66,732.10	75,444.29	-118,217.89	251,682.09	
음주기간						
1년	2년	12,777.21	80,285.05	-184,039.79	209,594.21	0.05
	3년	-2,590.91	65,807.59	-163,916.73	158,734.91	
	2년	15,368.12	50,341.82	-108,043.72	138,779.96	
비운동기간						
1년	2년	63,470.59	42,478.56	-40,664.64	167,605.82	1.70
	3년	57,444.63	33,177.44	-23,889.11	138,778.38	
	2년	6,025.96	36,425.95	-83,271.45	95,323.36	
식습관(불규칙)기간						
1년	2년	25,295.36	39,549.57	-71,659.54	122,250.25	0.39
	3년	24,176.96	32,732.23	-56,065.37	104,419.30	
	2년	1,118.39	45,182.68	-109,645.93	111,882.72	

\* p<0.05    \*\*p<0.01    \*\*\*p<0.001



## ABSTRACT

### Comparative research on medical expenses according to health risk factors

Kim, Hye Young

Graduate School of Public Health

Yonsei University

(Directed by Professor Young Moon Chae, Ph. D.)

The health habits affecting the occurrence of chronic degenerative diseases are in high correlation with medical expenses. Major risk factors of health habits are embedded in the daily life style such as smoking, drinking, diet and physical exercises. This research report analyzes the correlation between medical expenses and individual living habits.

Data used for this research are 1)the medical examination results from the patients who took medical examination consecutively for the years 1998, 2000 and 2002 at the National Health Insurance Corporation and 2)the medical expenses spent for 3 categories of diseases.

The mean of the medical expenses calculated by the health risk factors showed that men who has been drinking for the second year spent ₩14,593.0 more than men drinking for one year. Women whose eating habits are irregular for the second year spent ₩77,102.1 more than women with the same habits for one year. The group in the age of the sixties, smokers for second year spent ₩90,550.0 more than the smokers for one year. The major factors that affect the increase of medical expenses are the drinking period for men, smoking and drinking period for women. For the people in their forties, medical expenses increased as the drinking period increased. The increase rates for the medical expenses are 6% and 38% for men and women, respectively, when drinking habit is continued for 3 consecutive years.

In conclusion, the amount of medical expenses depends on the period of time where the risk factors are present. The longer the period, the larger are the medical expenses. The results of this research may be applied for office workers and local residents and be used for the national programs developed for the health care of the country.

---

Key word : Medical expenses, Health risk factors