

Modeling of Metabolic Syndrome Using Bayesian Network

Mi-Hyun Jin^a · Hyun-Ji Kim^b · Jea-Young Lee^{b,1}

^aSamsung Changwon Hospital, Sungkyunkwan University

^bDepartment of Statistics, Yeungnam University

(Received July 3, 2014; Revised September 18, 2014; Accepted October 6, 2014)

Abstract

Metabolic syndrome is a major factor for cardiovascular disease that can develop into a variety of complications such as stroke disease. This study utilizes a Bayesian network to model metabolic syndrome. In addition, we tried to find the best risk combinations to diagnose metabolic syndrome. We confirmed that the combinations are difference according to individual characteristics. The paper used data from 4,489 adults who responded to all health interview questions from the the 5th Korea National Health and Nutrition Examination Survey conducted in 2010.

Keywords: Bayesian network, metabolic syndrome, posterior probability.

1. 서론

대사증후군(metabolic syndrome)은 만성적인 대사 장애로 복부비만, 고중성지방혈증, 고혈당 등 여러 가지 만성질환이 한 개인에게서 한꺼번에 나타나는 것을 말한다. 이러한 대사증후군은 임상적으로 심혈관질환의 발생을 증가시킬 수 있다고 발표된 이후 널리 알려지게 되었으며, 다양한 합병증으로 발전될 수 있어 그 심각성이 커지고 있다. 그러므로 대사증후군에 대한 연구의 필요성이 매우 강조되고 있다 (Kim과 Park, 2012; Yoo 등, 2009). 대사증후군은 미국에서 20세 이상 성인의 23.7%, 50세 이상에서는 약 44%의 높은 유병률을 보이고 있으며 (Ford 등, 2002), 우리나라에서도 20대 이상 성인의 유병률은 1998년에 여자 14.2%, 남자 17.7%였으나 (Park 등, 2004) 2005년에는 여자 25.5%, 남자 27.1%로 증가하였다. 이는 2005 ~ 2008년 3년 동안의 국민건강영양조사에서 특히 여자의 당뇨병, 고혈압, 복부비만의 유병률이 소폭 상승한 결과로 보여진다 (Korea Centers for Disease Control and Prevention, 2008). 더불어 대사증후군은 인종, 성별, 연령, 가족력 등에 그 특성이 각각 다르게 나타난다고 보고되고 있다 (Jung 등, 2002; Cho 등, 2012; Ban 등, 2012; Han 등, 2012).

본 연구에서는 베이지안 네트워크 기법을 이용하여 대사증후군과 그 진단 기준이 되는 만성질환 5가지(복부비만, 고중성지방혈증, 고혈압, 저 HDL 콜레스테롤혈증, 고혈당)의 관계를 규명하고, 네트워크 모델을 구축한다. 더불어 대사증후군의 정의가 기관에 따라 다양하여 여러가지 정의 중에서 어느 기관의 정의가 가장 타당한지에 대해서는 명확하지 않은데, 모델링의 결과를 통해서 그 중 대표적인 세계보

¹Corresponding author: Department of Statistics, Yeungnam University, Kyungsan 712-749, Korea.

E-mail: jlee@yu.ac.kr

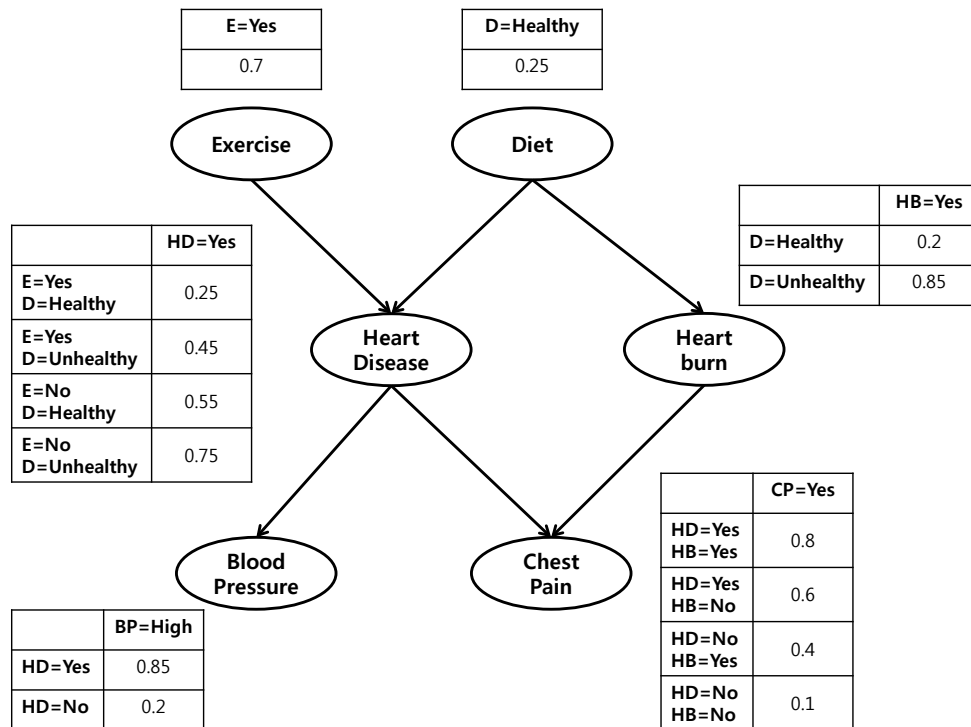


Figure 2.1. Bayesian network model of heart disease (Tan *et al.*, 2006).

건기구(1988년)의 정의와 세계당뇨연맹(2005년)의 정의의 타당성을 확인하고자 했다. 또한 각 진단 조합의 위험도를 비교하여 가장 위험한 대사증후군의 진단 조합을 선별하고, 개개인의 특성에 따라 유병률과 그 특징이 달라짐을 고려하여 인구학적인 특성과 건강행태에 따른 진단 조합을 제시한다.

2장에서는 베이저안 네트워크를 소개하고, 대사증후군의 정의와 본 연구에 사용된 제5기 국민건강영양조사 자료에 대해서 소개한다. 그리고 3장에서는 국민건강영양조사 자료 이용하여 베이저안 네트워크 기법을 통해 대사증후군의 모델을 구축하고, 이를 사후확률로 설명한다. 마지막으로 4장에서는 3장의 결과를 바탕으로 결론을 제시한다.

2. 연구방법

2장에서는 대사증후군의 진단 모델을 구축하는데 사용될 베이저안 네트워크 기법을 소개하고, 대사증후군의 정의와 본 연구에 사용된 제5기 국민건강영양조사 자료에 대해서 소개한다.

2.1. 베이저안 네트워크 기법 소개

베이저안 네트워크는 확률 변수 집합사이의 확률적 관계를 네트워크로 표현하는 방법이다 (Heckerman, 1997; Tan 등, 2006). Figure 2.1은 베이저안 네트워크의 예제로, 심장병(heart disease; HD)이나 속쓰림(heart burn; HB)을 갖고 있는 환자를 모델링하는 베이저안 네트워크를 나타낸 것이다. 베이저안 네트워크는 Figure 2.1과 같이 두 가지 요소로 표현된다. 첫째는 노드 사이의 종속성 관계를 방향성 비

순환 그래프(directed acyclic graph; DAG)로 표현하는 것이고 둘째는 각 노드 사이의 관계를 확률표로 표현하는 것이다. 그래프에서 노드는 변수, 화살표는 변수 쌍 사이의 종속관계를 보여준다. Figure 2.1의 운동(exercise; E)과 식이요법(diet; D)은 독립변수이고 각각은 HD에 직접적인 영향을 가지고 있다. HD는 고혈압(blood pressure; BP)과 가슴통증(chest pain; DP)의 부모노드이고 BP와 DP는 HD의 자식노드가 된다. 그리고 D는 BP와 DP의 조상노드가 된다. 베이저안 네트워크의 한 노드가 부모노드가 알려져 있다면 그 노드는 자손이 아닌 노드들과 조건부 독립적이다. 그러므로 노드 CP와 D가 노드 BP의 자손이 아니기 때문에 BP는 CP와 D에 조건부 독립적이라고 할 수 있다.

일반적으로 베이저안 네트워크는 다른 노드들의 확률 값들을 기초로 특정 노드가 가질 값에 대한 조건부 확률을 계산하는데 이용할 수 있다. 따라서 하나의 베이저안 네트워크는 한 개체의 다른 속성들의 값이 주어졌을 때 분류 클래스 노드의 사후 확률 분포를 구해줌으로써 개체들에 대한 하나의 자동 분류기로 이용될 수 있다. 즉, 베이저안 네트워크를 기초로 분류 클래스를 확률적으로 예측할 수 있다. 베이저안 네트워크는 조건부 확률 계산에 베이지 정리를 이용하여 다음의 수식을 얻을 수 있다.

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}.$$

결과적으로 체인룰을 사용하면, 다음의 식이 만들어진다.

$$P(X_1, X_2, \dots, X_n) = P(X_1) \prod_{i=2}^n P(X_i | X_1, \dots, X_{i-1}).$$

다시 Figure 2.1의 예제를 살펴보면 심장병을 가질 가능성은 다음과 같이 계산할 수 있다. 쉬운 표기를 위해 운동의 이진 값을 $\alpha \in \{\text{Yes, No}\}$, 식사습관의 이진법을 $\beta \in \{\text{Healthy, Unhealthy}\}$ 로 표현하기로 가정한다. 이때 심장병을 가지고 있을 확률은 다음과 같다.

$$\begin{aligned} P(\text{HD} = \text{Yes}) &= \sum_{\alpha} \sum_{\beta} P(\text{HD} = \text{Yes} | E = \alpha, D = \beta) P(E = \alpha, D = \beta) \\ &= \sum_{\alpha} \sum_{\beta} P(\text{HD} = \text{Yes} | E = \alpha, D = \beta) P(E = \alpha) P(D = \beta) \\ &= 0.25 \times 0.7 \times 0.25 + 0.45 \times 0.7 \times 0.75 + 0.55 \times 0.3 \times 0.25 + 0.75 \times 0.3 \times 0.75 \\ &= 0.49, \end{aligned}$$

그리고 심장병의 유무를 $\gamma \in \{\text{Yes, No}\}$ 로 나타내었을 때, 고혈압을 가지고 있을 확률은 다음과 같다.

$$\begin{aligned} P(\text{BP} = \text{High}) &= \sum_{\gamma} P(\text{BP} = \text{High} | \text{HD} = \gamma) P(\text{HD} = \gamma) \\ &= 0.85 \times 0.49 + 0.2 \times 0.51 \\ &= 0.5185, \end{aligned}$$

여기서 알고 있는 확률을 바탕으로 어떤 사람이 고혈압이 있을 때 심장병이 걸릴 확률을 구하면 다음과 같다.

$$\begin{aligned} P(\text{HD} = \text{Yes} | \text{BP} = \text{High}) &= \frac{P(\text{BP} = \text{High} | \text{HD} = \text{Yes}) p(\text{HD} = \text{Yes})}{P(\text{BP} = \text{High})} \\ &= \frac{0.85 \times 0.49}{0.5185} \\ &= 0.8033. \end{aligned}$$

그러므로 고혈압이 있을 때 심장병에 걸리지 않을 확률은 0.1967이다. 따라서 고혈압을 가진 사람은 심장병의 위험이 증가하는 것을 볼 수 있다.

2.2. 대사증후군의 정의

대사증후군은 1988년 ‘Syndrome X’라고 처음 명명 (Reaven, 1988)한 후, 1998년 세계보건기구(WHO; world health organization)에서 이를 대사증후군(metabolic syndrome)이라 명명하고 이의 진단 기준을 공식적으로 처음 제시하였다 (Consultation, 1999). 그 진단 기준은 복부비만, 고중성지방혈증, 고혈당, 저 HDL 콜레스테롤혈증, 고혈압 중 세 가지 이상의 요인을 가지는 경우이다.

이후, 2005년에 세계당뇨연맹(IDF; international diabetes federation)은 대사증후군의 국제적 정의를 도출하기 위해 새로운 진단 기준을 제시하였다. 복부비만이 인슐린 저항성을 가장 잘 반영하며, 인슐린 저항성을 일으키는데 가장 중요한 인자로 판단하기 때문에 복부비만을 포함하면서, 고중성지방혈증, 고혈당, 저 HDL 콜레스테롤혈증, 고혈압 중 두 가지 이상의 요인을 가지는 경우 대사증후군이라 진단하기로 하였다 (Brussels, 2006).

허리둘레(WC; waist circumference)

: 허리둘레가 남자 90cm 이상, 여자 80cm 이상인 자를 복부비만이라 함

중성지방(TG; triglyceride)

: 중성지방이 150mg/dl 이상 또는 고지혈증 약을 복용중인 자를 고중성지방혈증이라 함

혈압(BP; blood pressure)

: 수축기혈압 130mmHg 이상 또는 이완기 혈압이 85mmHg 이상이거나 혈압 조절 약을 복용중인 자를 고혈압이라 함

HDL 콜레스테롤(HDL-C; high density lipoprotein-cholesterol)

: HDL 콜레스테롤이 남자 40mg/dl 미만, 여자 50mg/dl 미만인 자 또는 약을 복용중인 자를 저 HDL 콜레스테롤혈증이라 함

공복시 혈당(FBS; fasting blood sugar)

: 공복 시 혈당이 100mg/dl 이상인 자 또는 인슐린 주사를 맞거나 당뇨병 약을 복용중인 자를 고혈당이라 함

2.3. 연구대상 - 국민건강영양조사

보건복지부의 질병관리본부가 주관하는 국민건강영양조사 자료는 보건 정책 수립 및 사업평가의 지표로 활용되며, 세계보건기구 (WHO)와 경제협력개발기구 (OECD)에 제공되는 자료이다. 이 조사 자료는 제1기(1998년), 제2기(2001년), 제3기(2005년), 제4기(2007 ~ 2009년), 제5기 (2010 ~ 2012년)으로 구분한다.

본 연구는 국민건강영양조사 제5기 자료 중 1차년도 자료(2010년)를 이용하였으며, 국민건강영양조사 중 건강검진과 건강의식행태조사 자료를 사용하였다. 조사 참여자 10,938명 중에서 우리나라 성인을 대상으로 한 것이므로 만 19세 미만이거나, 불충분한 응답으로 분석방법에 적용할 수 없는 응답자를 제외하여 4,489명을 최종 분석 대상자로 결정하였다.

건강검진 자료는 신장, 체중, 혈압 등의 검진조사는 전문조사 수행 팀이 이동검진센터에서 직접 계측, 검진을 하였고, 흡연, 음주 등의 건강행태 설문조사는 자기기입방식과 면접조사를 병행하였다. 조사방법에 대해서는 국민건강영양조사 보고서에 상세히 기술되어 있으며 (Korea Centers for Disease Control and Prevention, 2010), 본 연구에서는 대사증후군 진단 요인 5가지와 인구사회학적 특성(성별, 연령, 교육수준, 결혼여부)과 건강행태(흡연여부, 음주여부, 주관적 체형 인식, 최근 1년간 체중변화여부,

Table 3.1. Frequency of components for metabolic syndrome.

	Overall	Male	Female
	(<i>N</i> = 4, 489)	(<i>N</i> = 1, 958)	(<i>N</i> = 2, 531)
Metabolic syndrome	1,078 (24.0%)	518 (26.5%)	560 (22.1%)
3 component	684 (63.5%)	331 (63.9%)	353 (63.0%)
4 component	315 (29.2%)	150 (29.0%)	165 (29.5%)
5 component	79 (7.3%)	37 (7.1%)	42 (7.5%)

Table 3.2. Demographic characteristics.

Characters		Nomal	Metabolic syndrome	<i>p</i> -value
		(<i>N</i> = 3, 411)	(<i>N</i> = 1, 078)	
Gender	Male	1,440 (42.2%)	518 (48.1%)	0.001
	Female	1,971 (57.8%)	560 (51.9%)	
Age		44.25 ± 14.79	54.95 ± 13.52	< 0.001
Height (cm)		163.28 ± 8.72	161.95 ± 9.83	< 0.001
Weight (kg)		60.72 ± 10.55	68.70 ± 12.38	< 0.001
BMI (kg/m ²)		22.69 ± 2.92	26.05 ± 3.03	< 0.001
WC (cm)	Male	81.59 ± 7.87	91.48 ± 7.53	< 0.001
	Female	74.74 ± 8.30	86.40 ± 7.48	< 0.001
TG (mg/dl)		105.36 ± 77.89	210.22 ± 147.90	< 0.001
BP (mmHg)	Systolic	113.11 ± 15.46	129.58 ± 16.60	< 0.001
	Diastolic	72.58 ± 9.61	80.71 ± 10.76	< 0.001
HDL-C (mg/dl)	Male	48.00 ± 10.23	39.97 ± 8.88	< 0.001
	Female	53.47 ± 10.42	43.24 ± 8.09	< 0.001
FBS (mg/dl)		93.54 ± 19.70	104.29 ± 20.74	< 0.001

주요 만성질환에 대한 가족력) 변수를 사용하였다. 그 중 가족력에 대해서는 “귀 가족 중에서 의사로부터 다음의 질병을 진단받으신 적이 있습니까?”로 질문하여 부모, 형제자매 각각 고혈압, 고중성지방혈증, 허혈성 심질환, 뇌졸중, 당뇨병을 진단받았는지 질문하고 각 질병에 대해 부모나 형제 중에 한 사람이라도 해당 질병을 앓고 있는 경우 가족력이 있는 것으로 정의하였다.

3. 적용 및 결과

3.1절에서는 실험 대상군의 특성에 대해서 살펴보고, 3.2절에서는 베이저안 네트워크 기법을 이용하여 대사증후군의 모델을 구축한다. 이를 이용하여 3.3절에서는 각 진단 조합별 위험도를 계산하여 비교하고 가장 위험한 진단 조합을 선별한다.

3.1. 대상군의 특성

총 4,489명의 대상자의 평균 연령은 46.81 ± 15.19세였으며, 그 중 남자가 43.6%, 여자가 56.4%였다. Table 3.1은 대사증후군의 유병률을 나타낸 표이다. 본 연구에서의 대사증후군 유병률은 24.0%이며, 성별에 따른 유병률은 남자가 26.5%, 여자가 22.1%로 나타났다. 대사증후군이라 진단 받은 사람들 중에서 3가지 위험 요인에 해당이 되는 사람은 63.5%로 가장 많았고, 4가지 위험 요인에 해당되는 사람과 5가지 요인에 해당되는 사람이 각각 29.2%, 7.3%로 나타났다.

Table 3.2는 일반적인 특성을 t검정이나 교차분석을 통해 대사증후군의 유무에 따라 비교한 결과를 나

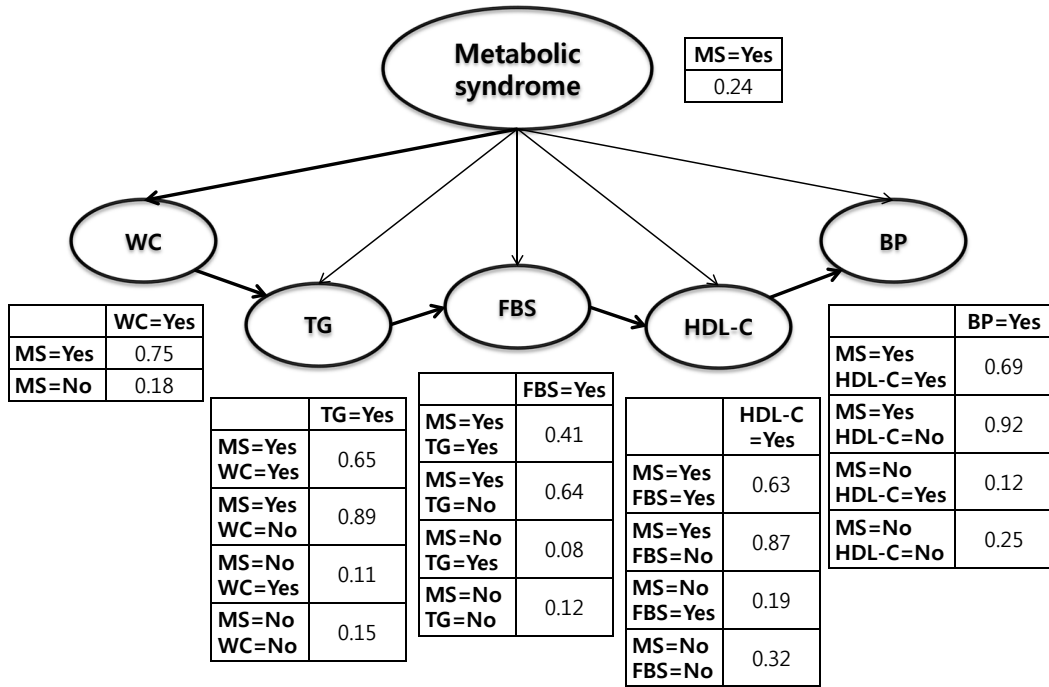


Figure 3.1. Bayesian network model of metabolic syndrome.

타낸 표이다. 대사증후군이라 진단 받은 군은 평균 연령이 54.95세로 정상군보다 10.70세가량 높았고, 키는 정상군보다 1.33cm가량 낮았고 체중은 7.98kg 높았다. 허리둘레, 중성지방, 혈압, 공복 시 혈당이 대사증후군을 진단받은 군에서 유의하게 높게 나타났고, HDL 콜레스테롤의 수치는 유의하게 낮았다.

3.2. 대사증후군 모델링

Figure 3.1은 대사증후군과 대사증후군을 진단하는 5가지 위험요인(복부비만, 고중성지방혈증, 고혈당, 저 HDL 콜레스테롤혈증, 고혈압)의 베이저안 네트워크를 IBM SPSS Modeler 버전 14.1을 통해 구축한 결과이다. 그림에서 각 변수는 이진 값(yes 또는 no)을 갖는 것으로 가정한다.

베이저안 네트워크를 구축한 결과, 대사증후군의 진단 여부에 대한 노드가 0.24의 사전확률을 가지고 있고, 복부비만, 고중성지방혈증, 고혈당, 저 HDL 콜레스테롤혈증, 고혈압 노드는 조건부 확률을 포함한다. 즉, 우리나라의 19세 이상 성인들의 대사증후군 유병률이 24%인 것으로 나타났고, 복부비만, 고중성지방혈증, 고혈당, 저 HDL 콜레스테롤혈증, 고혈압 노드와 각각 직접적인 관련이 있으며 특히 복부비만이 대사증후군에 가장 영향을 많이 미치고, 복부비만을 중심으로 복부비만은 고중성지방혈증에, 고중성지방혈증은 고혈당에, 고혈당은 저 HDL 콜레스테롤혈증에, 저 HDL콜레스테롤혈증은 고혈압에 영향을 미치는 등 5가지 요인들이 연쇄적인 네트워크를 갖는 것으로 분석된다. 여기서 단일 요인으로 봤을 경우, 복부비만이 가장 큰 영향력을 갖는다고 나타났다. 이는 복부비만을 갖고 있으면서, 다른 4가지 요인 중에 2가지에 해당될 경우 대사증후군이라는 진단을 내린다는 세계당뇨연맹의 대사증후군 정의가 타당하다는 것을 보여준다.

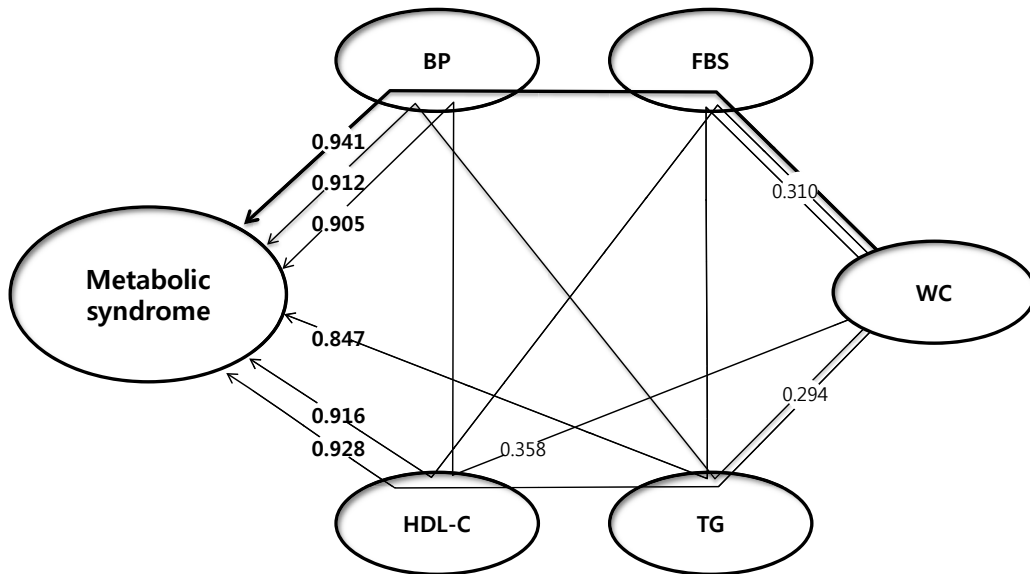


Figure 3.2. The risk combinations including waist circumference for metabolic syndrome.

Table 3.3. Posterior probability of risk combinations including waist circumference for metabolic syndrome.

Dependent variable	Risk factors	Posterior probability
Metabolic syndrome	FBS, BP	0.941
	TG, HDL-C	0.928
	FBS, HDL-C	0.916
	TG, BP	0.912
	HDL-C, BP	0.905
	FBS, TG	0.847

WC; wasit circumference, TG; triglyceride, BP; blood pressure, HDL-C; high density lipoprotein - cholesterol, FBS; fasting blood sugar

3.3. 진단 조합 선별

3.2절에서 구축된 대사증후군의 베이저안 네트워크 모델에서, 각 위험요인들의 확률적 관계를 이용하여 진단 조합의 위험도 분석을 통해 진단 조합을 선별했다. 3.2절에서 타당성이 확인 된 세계당뇨연맹의 대사증후군 정의(복부비만을 갖고 있으면서, 고중성지방혈증, 고혈당, 저 HDL 콜레스테롤혈증, 고혈압 중 2가지가 추가로 해당될 경우)를 사용했다. 이에 따른 진단 조합별 위험도의 계산 결과를 Table 3.3과 Figure 3.2에 나타내었다. 그 결과, 복부비만과 고혈당, 고혈압이 한 개인에게서 복합적으로 나타나는 경우의 위험도가 0.941로 가능한 6개의 조합 중에서 가장 높게 나타났다. 여기서 위험도는 $P(MS = Yes|WC = Yes, FBS = Yes, BP = Yes) = 0.941$ 로 사후확률을 의미한다. 그리고 복부비만과 고중성지방혈증, 저 HDL 콜레스테롤혈증의 조합은 0.928로 두 번째로 높은 위험도를 가지는 것으로 확인했다. 또한 복부비만과 고혈당, 저 HDL 콜레스테롤혈증의 조합의 위험도는 0.916이었으며, 복부비만, 고중성지방혈증, 고혈압의 조합과 복부비만, 저 HDL 콜레스테롤혈증, 고혈압의 조합이 각각 0.912와 0.905의 위험도를 가지는 것으로 나타났다. 마지막으로 복부비만, 고혈당, 고중성지방혈증의 위험요인이 한 개인에게서 나타나는 경우의 대사증후군의 위험도는 0.847로 6개의 조합 중에서 가장 위험도가 낮았다.

Table 3.4. The best risk combinations for metabolic syndrome by demographic characteristics.

Demographic characteristics		Risk factors	Posterior probability
Gender	Male	WC, HDL-C, BP	0.951
	Female	WC, HDL-C, TG	0.920
Age	< 50s	WC, FBS, BP	0.942
	≥ 50s	WC, TG, HDL-C	0.960
Education	≤ Middle school	WC, TG, HDL-C	0.958
	High school	WC, FBS, BP	0.922
	≥ University	WC, FBS, BP	0.943
Marital	Single	WC FBS, BP	0.943
	Married	WC FBS, BP	0.918

WC; wasit circumference, TG; triglyceride, BP; blood pressure, HDL-C; high density lipoprotein - cholesterol, FBS; fasting blood sugar

Table 3.5. The best risk combinations for metabolic syndrome by health behaviour.

Health behaviour		Risk factors	Posterior probability
Drinking	Non-drinker	WC, TG, HDL-C	0.928
	Drinker	WC, FBS, BP	0.949
Family history	No	WC, FBS, BP	0.947
	Yes	WC, TG, HDL-C	0.931
Weight change	No	WC, FBS, BP	0.953
	Yes	WC, FBS, BP	0.917
Subjective body type	Thin	WC FBS, BP	0.975
	Normal	WC FBS, BP	0.957
	Fat	WC FBS, BP	0.932

WC ; wasit circumference, TG ; triglyceride, BP ; blood pressure, HDL-C ; high density lipoprotein - cholesterol, FBS ; fasting blood sugar

Table 3.4는 대사증후군의 유무에 따라 대상자의 일반적인 특성에서 유의한 차이를 보인 성별, 연령, 교육 수준, 결혼 여부의 응답에 따라 대사증후군의 위험도를 나타낸 결과이다. 성별에 따라 보면 남성의 경우 복부비만, 저 HDL 콜레스테롤혈증, 고혈압 조합의 위험도가 0.951, 여성의 경우에는 복부비만, 저 HDL 콜레스테롤혈증, 고중성지방혈증 조합의 위험도가 0.920로 가장 높게 나타났다. 두 번째로 연령에 따라 보면 50대 이상에서 복부비만, 고중성지방혈증, 저 HDL 콜레스테롤혈증의 조합이 위험도가 0.960으로 가장 높게 나타났다. 세 번째로 교육수준에 따라 보면 중졸 이하의 경우에 복부비만, 고중성지방혈증, 저 HDL 콜레스테롤혈증의 조합이 높은 위험도를 가지는 것으로 나타났다. 결혼여부에 따라서는 미혼과 기혼에 상관없이 복부비만, 고혈당, 고혈압의 조합이 가장 위험한 조합인 것으로 나타났다.

Table 3.5은 건강행태 설문문항에 따라 대사증후군의 위험도를 정리한 결과이다. 월간 음주 여부에 따라서는 비음주자의 경우 복부비만, 고중성지방혈증, 저 HDL 콜레스테롤혈증의 조합이 0.928로 가장 위험한 조합이었다. 가족력이 있는 경우도 복부비만, 고중성지방혈증, 저 HDL 콜레스테롤혈증의 조합이 0.931로 가장 위험한 조합이었다. 그러나 그 외 최근 1년간 체중변화여부, 주관적인 체형인식의 경우에는 특성에 차이 없이 복부비만, 고혈당, 고혈압의 조합이 가장 위험한 조합으로 나타났다.

4. 결론

본 연구는 제5기 국민건강영양조사(2010 ~ 2012년)의 1차년도 대상자인 8,958명 중에서, 19세 이상 성

인이며 분석에 적합한 4,489명을 최종 대상으로 선정한 자료로 베이저안 네트워크 기법을 적용하여 대사증후군의 모델을 구축하였다. 그 결과, 복부비만이 가장 위험한 요인이며, 복부비만을 중심으로 고중성지방혈증, 고혈당, 저 HDL 콜레스테롤혈증, 고혈압의 순서로 연쇄적인 네트워크를 형성하였다. 즉, 복부비만을 가장 위험한 요인으로 보고, 복부비만을 포함한다는 세계당뇨연맹의 정의가 타당하다는 것을 확인하였다. 더불어 복부비만을 포함한 여러 가지 진단 조합 중에서 복부비만과 고혈당, 고혈압의 조합이 0.941로 가장 위험한 조합임을 밝혔다. 또한, 개개인의 특성(성별, 연령대, 가족력 등)에 따라 대사증후군의 유병률이 달라짐을 고려하여 분석한 결과, 대상자의 성별에 따라 분석한 결과에서 복부비만과 저 HDL 콜레스테롤혈증의 경우는 동일하나, 남성의 경우 고혈압을 가진 조합이 0.951로 가장 위험한 조합으로, 여성의 경우 고중성지방혈증을 가진 조합이 0.920로 가장 위험한 조합으로 나타나 성별에 차이가 있다는 것을 확인했다. 나이에 따라 분석한 결과 50대 이상의 경우 복부비만, 고중성지방혈증, 저 HDL 콜레스테롤혈증의 조합이 0.960으로 가장 위험한 조합으로 나타났다. 그리고 건강행태에 따라서 살펴보면 음주를 하지 않는 경우와 가족력이 있는 경우에는 복부비만, 고중성지방혈증, 저 HDL 콜레스테롤혈증의 위험도가 각각 0.951, 0.960으로 가장 위험한 조합으로 나타났고, 그 외의 경우에는 복부비만, 고혈당, 고혈압의 조합이 가장 위험한 조합으로 나타났다. 이렇게 개개인의 특성에 따라 나누어 본 경우, 대사증후군의 특징이 조금씩 다르게 나타난다는 것을 확인하였고, 더불어 진단 조합의 위험도가 상승하는 것을 확인하였다. 즉, 개개인의 특성에 따라 진단을 하면, 좀 더 정확히 예측 할 수 있다.

References

- Ban, S. M., Lee, K. J. and Yang, J. O. (2012). The effect of participation in a combined exercise program on the metabolic syndrome indices and physical fitness in the obese middle-aged women, *Journal of the Korean Data & Information Science Society*, **23**, 703–715.
- Brussels (2006). *The IDF Consensus Worldwide Definition of the Metabolic Syndrome*, IDF Communication, Belgium, 1–24.
- Cho, Y. C., Kwon, I. S., Park, J. Y. and Shin, M. W. (2012). Prevalence of metabolic syndrome and its associated factors among health checkup examines in a university hospital, *Journal of the Korea Academia-Industrial Cooperation Society*, **13**, 5317–5325.
- Consultation, W. (1999). *Definition, Diagnosis and Classification of Diabetes Mellitus and Its Complications*, World Health Organization Department of Noncommunicable Disease Surveillance Geneva.
- Ford, E. S., Giles, W. H. and Dietz, W. H. (2002). Prevalence of the metabolic syndrome among US adults, *The Journal of the American Medical Association*, **287**, 356–359.
- Han, J. M., Lee, K. J. and Yang, K. O. (2012). The effects of the 16-weeks combined exercise program on metabolic syndrome and autonomic nerve system of low-level physical strength group, *Journal of the Korean Data & Information Science Society*, **23**, 787–796.
- Heckerman, D. (1997). *Bayesian Networks for Data Mining*, Kluwer Academic Publishers, London.
- Jung, C. H., Park, J. S., Lee, W. Y. and Kim, S. W. (2002). Effects of smoking, alcohol, exercise, level of education, and family history on the metabolic syndrome in Korean adults, *Korean Journal of Medicine*, **63**, 649–659.
- Kim, M. K. and Park, J. H. (2012). Metabolic syndrome, *Journal of the Korean Medical Association*, **55**, 1005–1013.
- Korea Centers for Disease Control and Prevention (2008). *The third Korea National Health and Nutrition Examination Survey(KNHANES 3)*, Ministry of Health, Welfare and Family Affairs, Seoul.
- Korea Centers for Disease Control and Prevention (2010). *The fifth Korea National Health and Nutrition Examination Survey(KNHANES 5-1)*, Ministry of Health, Welfare and Family Affairs, Seoul.
- Park, H. S., Oh, S. W., Cho, S. I., Choi, W. H. and Kim, Y. S. (2004). The metabolic syndrome and associated lifestyle factors among South Korean adults, *International Journal of Epidemiology*, **22**, 328–336.
- Reaven, G. M. (1988). Diabetes, *Role of Insulin Resistance in Human Disease*, **37**, 1595–1607.
- Tan, P., Steinbach, M. and Kumar, V. (2006). *Introduction to Data Mining*, Addison Wesley Longman,

California.

Yoo, J. S., Jung, J. I., Park, C. G., Kang, S. W. and Ahn, J. A. (2009). Impact of life style characteristics on prevalence risk of metabolic syndrome, *Journal of Korean Academy of Nursing*, **39**, 594–601.

베이지안 네트워크를 이용한 대사증후군 모델링

진미현^a · 김현지^b · 이제영^{b,1}

^a성균관대학교 삼성창원병원, ^b영남대학교 통계학과

(2014년 7월 3일 접수, 2014년 9월 18일 수정, 2014년 10월 6일 채택)

요약

대사증후군은 뇌졸중이나 심혈관 질환 등 다양한 합병증으로 발전될 수 있어 그 심각성이 커지고 있으며, 우리나라의 유병률이 증가하는 추세를 보이고 있어 연구의 필요성이 강조되고 있다. 본 연구는 베이지안 네트워크를 활용하여 대사증후군과 그 진단기준이 되는 5가지 이상 징후인 복부비만, 고중성지방혈증, 고혈압, 저 HDL 콜레스테롤혈증, 고혈당의 관계에 대한 모델을 구축하고자 하였다. 추가적으로 대사증후군의 가장 위험한 진단조합을 선별하였고, 개인의 특성에 따라 대사증후군의 특징도 다르게 나타난다는 것을 확인하였다. 사용된 데이터는 제5기 국민건강영양조사 중 2010년 자료로써 건강설문조사의 모든 문항에 응답한 성인 4,489명의 데이터이다.

주요용어: 대사증후군, 베이지안 네트워크, 사후확률.

¹교신저자: (712-749) 경북 경산시 대동 214-1, 영남대학교 통계학과. E-mail: jlee@yu.ac.kr