



Osong Public Health and Research Perspectives

Journal homepage: <http://www.kcdcphrp.org>



Original Article

Developing the High-Risk Drinking Scorecard Model in Korea

Jun-Tae Han ^a, Il-Su Park ^{b,*}, Suk-Bok Kang ^c, Byeong-Gyu Seo ^c

^a Department of Student Aid Policy Research, Korea Student Aid Foundation, Daegu, Korea

^b Department of Health Management, Uiduk University, Gyeongju, Korea

^c Department of Statistics, Yeungnam University, Gyeongsan, Korea

ABSTRACT

Article history:

Received: August 9, 2018

Revised: August 20, 2018

Accepted: August 23, 2018

Keywords:

high-risk drinking, data mining, weighted multiple logistic regression, scorecard, Korea Community Health Survey

Objectives: This study aimed to develop a high-risk drinking scorecard using cross-sectional data from the 2014 Korea Community Health Survey.

Methods: Data were collected from records for 149,592 subjects who had participated in the Korea Community Health Survey conducted from 2014. The scorecard model was developed using data mining, a scorecard and points to double the odds approach for weighted multiple logistic regression.

Results: This study found that there were many major influencing factors for high-risk drinkers which included gender, age, educational level, occupation, whether they received health check-ups, depressive symptoms, over-moderate physical activity, mental stress, smoking status, obese status, and regular breakfast. Men in their thirties to fifties had a high risk of being a drinker and the risks in office workers and sales workers were high. Those individuals who were current smokers had a higher risk of drinking. In the scorecard results, the highest score range was observed for gender, age, educational level, and smoking status, suggesting that these were the most important risk factors.

Conclusion: A credit risk scorecard system can be applied to quantify the scoring method, not only to help the medical service provider to understand the meaning, but also to help the general public to understand the danger of high-risk drinking more easily.

<https://doi.org/10.24171/j.phrp.2018.9.5.04>
pISSN 2210-9099 eISSN 2233-6052

©2018 Korea Centers for Disease Control and Prevention. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Moderate alcohol consumption is generally known to reduce the risk of ischemic heart disease [1,2]. However, alcohol consumption has been recognized as one of the major risk factors of preventable mortality and morbidity. Binge drinking and heavy drinking have been associated with violence, poor management of diabetes, neurological damage, hypertension, hepatitis, gastrointestinal and heart disease, liver cirrhosis, cancers such as oral, rectal, and liver cancer, stroke, and alcohol dependence [3-5]. The 2016 Korea National Health and Nutrition Examination Survey [conducted by the Korea Centers for Disease Control and Prevention (KCDC)] revealed rates of monthly alcohol consumption and high-risk drinking

of 61.9% and 13.8%, respectively. This rate of high-risk drinking is very high compared to the rates reported by the World Health Organization for Africa (5.7%), the Americas (13.7%), the Eastern Mediterranean (0.1%), Europe (16.5%), South-East Asia (1.6%), and the Western Pacific Region (7.7%) [6]. In addition, many longitudinal studies in Korea have focused on the health effects of alcohol drinking [7-9].

Many recent studies have found that age, income level, employment status, smoking status, obesity, subjective assessment of health, and presence of spouse are all related to high-risk drinking [10,11]. Previous studies ranged from small scale sample surveys which were designed and surveyed individually, to large scale sample surveys which were collected the nationally. Some studies involved individuals

*Corresponding authors: Il-Su Park
Department of Health Management, Uiduk University, Gyeongju, Korea
E-mail: ispark@uu.ac.kr

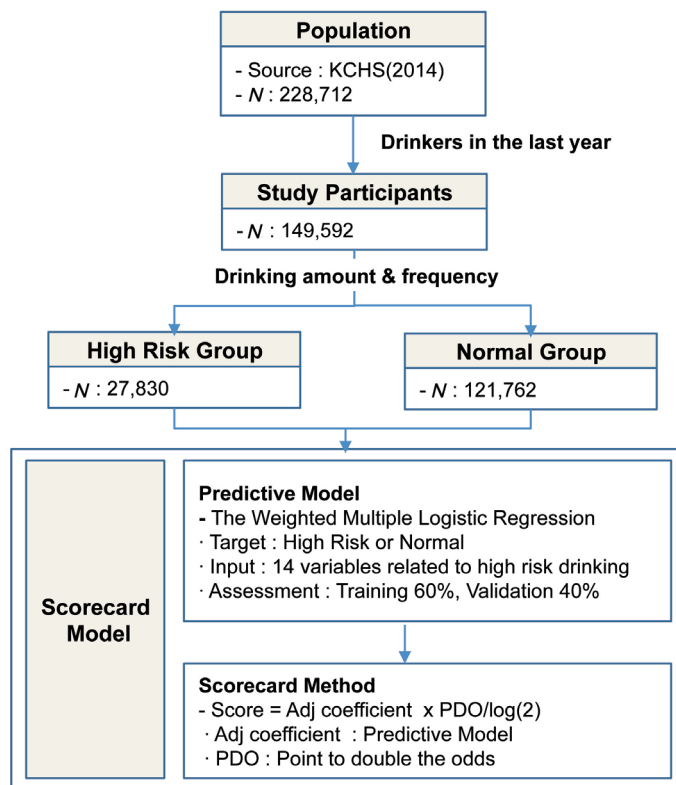


Figure 1. Framework of the study.

KCHS = Korea Community Health Survey; PDO = point to double the odds.

exposed to high-risk drinking thus were a group that needs an improvement of drinking culture the most. Logistic regression analysis was the preferred method of most studies to detect risk factors for high-risk drinking. However, there was a difficulty in interpreting the results of high-risk drinking predictions, generally or utilizing it in medical service.

The purpose of this study was to develop a predictive model of high-risk drinking in Korea using data mining. Scorecards for high-risk drinking may be used by employing the developed prediction model.

Materials and Methods

1. Study design

This study was a secondary analysis of data that was collected in a nationally representative cross sectional and population-based survey conducted by the KCDC. The overall framework of this study is shown in Figure 1.

2. Subjects

This study was based on data acquired in the Korea

Community Health Survey (KCHS) from 2014. The KCHS is a national health survey conducted since 2008 to provide population-based estimates of health indicators to be used for the development and assessment of public health policies and programs. The 2014 KCHS used a multistage sampling design to obtain a representative sample of adults aged 19 years or older. Those aged 19 years or older were initially selected. Subjects who did not respond to the questionnaires about sociodemographic variables and health-related variables were excluded. After exclusion criteria were applied, 149,592 subjects were included in the final analysis.

3. Study variables

The target variable in this study was the alcohol consumption pattern. Any person who had drunk any kind of alcoholic beverage during the past year was classified as a current drinker and was asked more questions on the quantity consumed in a typical day and the drinking frequency. "High-risk" was defined as the male respondents who consumed more than 7 drinks twice a week or more, as well as female respondents who consumed more than 5 drinks twice a week or more. All others were defined as "normal."

For the comprehensive analysis of the various factors associated with high-risk alcohol consumption, health behaviors, sociodemographic variables, and self-rated health status, including mental health, were selected as independent variables. The sociodemographic variables were gender, age, marital status, monthly household income, education level, and occupation. The lifestyle and health-related variables were over-moderate physical activity (participated in moderate physical activity for 5 days or more per week, and for 30 minutes or more per activity, or in vigorous activity for 3 days or more per week, and for 20 minutes or more per activity), eating a breakfast regularly, current smoking status, health check-ups during the past 2 years (proxy variable indication for interest on self-health care [12-14]), experience of depression (yes, no), subjective health status (good, bad), subjective stress recognition (yes, no), and obesity status (Table 1).

4. Statistical analysis

Data analysis, predictive model and scorecard development were performed with SAS version 9.4. In order to calculate the total population that the sample would represent, the stratification variables and sampling weights designated by the KCDC were employed. All data were described as unweighted frequency, and weighted percentage. χ^2 -test for categorical variables were performed. The high-risk drinking predictive model was built on the training set and tested the validity of the models on the validation set. The data set was divided into the training data set (60%) and the validation data set (40%).

Table 1. Data description in the analysis.

Variable	Definition
Target	Male respondents who consumed more than 7 drinks twice a week or more, as well as female respondents who consumed more than 5 drinks twice a week or more, were defined as high-risk drinkers. (1: High-risk, 0: Normal)
Gender	Male, Female
Age (y)	19~29, 30~39, 40~49, 50~59, 60~69, ≥ 70
Marital status	Married Others: Never married, separated, divorced, widowed
Monthly household income (million KRW)	< 0.5, 0.5~1.0, 1.0~2.0, 2.0~3.0, 3.0~4.0, 4.0~5.0, 5.0~6.0, ≥ 6.0
Occupation	- Administrative officer: Administrative, management, or professional occupation - Clerical officer: Business and financial operations occupations - Service and sales worker: Sales and related occupations - Farmer and fisher: Farming, fishing, and forestry occupations - Elementary work: Installation, maintenance, and repair occupations/labors - Other
Input	
Educational level	Uneducated, elementary school, middle school, high school, university or higher
Health check-up	Yes, No
Experience of depression	Yes, No
Subjective health status	Good, Bad (Fair or poor)
Over-moderate physical activity	Yes, No
Subjective stress recognition	Yes, No
Current smoking	Yes, No
Obesity	Yes: BMI ≥ 25 kg/m ² , No: BMI < 25 kg/m ²
Eating a breakfast regularly	Yes, No

The training set contained 90,015 cases (60%) represented by 73,250 normal cases and 16,765 high-risk cases. The validation set comprised 59,577 cases (40%), divided into 48,512 normal cases and 11,065 high-risk cases.

In this study, the weighted multiple logistic regression model was employed to develop the high-risk predictive model and a predictive scorecard was suggested for high-risk drinkers using a developed model.

Results

1. Differences in variables by target groups

Table 2 shows the key input variables used in the analysis by target groups. Of the 149,592 Korean adults, 27,830 (18.6%) participants were in the high-risk group and 121,762 (81.4%) participants were in the normal group. Significant differences between the 2 groups were observed in sociodemographic factors. The percentage of male participants in the high-

risk group was 83.7% and 49.7% in the normal group. In the high-risk group, there were significantly higher numbers of participants who were between 40–49 years old and who were high school graduates compared with the normal group. The biggest proportion of the high-risk group was participants employed in elementary work (30.9%), while the biggest proportion of the normal group had a different employment status [unemployed, full-time student, soldier (33.3%)]. Table 2 also provides a breakdown of the proportion in both groups depending on whether they received health check-ups during the prior 2 years, or they participated in over-moderate physical activity, or they were stressed. In addition, the percentage of current smokers in the high-risk group was higher than in the normal group [21.3% normal group, 51.5% high-risk group (Table 2)].

2. Model building and performance

The data set was divided into the training set (60%) and the validation set (40%). The models on the training set were

Table 2. Descriptive characteristics for the variables in the analysis.

Variable	Normal			High-risk			χ^2
	N	Weighted (N)	%	N	Weighted (N)	%	
Gender							
Male	57,155	12,117,462	49.7	23,099	4,765,414	83.7	8001.2402**
Female	64,607	12,254,590	50.3	4,731	929,837	16.3	
Age (y)							
19~29	18,337	5,270,497	21.6	3,266	965,149	16.9	1323.9462**
30~39	22,538	5,220,387	21.4	5,563	1,288,907	22.6	
40~49	27,158	5,650,642	23.2	7,988	1,639,864	28.8	
50~59	24,813	4,586,044	18.8	6,642	1,257,353	22.1	
60~69	16,337	2,224,066	9.1	3,009	402,031	7.1	
≥ 70	12,579	1,420,417	5.8	1,362	141,947	2.5	
Educational level							
Uneducated	6,428	617,121	2.5	671	74,932	1.3	415.7543**
Elementary school	13,854	1,563,923	6.4	2,373	289,721	5.1	
Middle school	12,379	1,855,573	7.6	2,985	471,388	8.3	
High school	46,141	9,828,631	40.3	12,546	2,609,866	45.8	
University or higher	42,960	10,506,805	43.1	9,255	2,249,344	39.5	
Occupation							
Administrative officer	16,547	4,144,491	17.0	3,585	906,076	15.9	2598.3654**
Clerical officer	13,423	3,209,875	13.2	3,623	874,896	15.4	
Service and sales worker	17,246	3,519,665	14.4	4,452	971,675	17.1	
Farmer and fisher	11,609	690,024	2.8	2,775	193,923	3.4	
Elementary work	22,987	4,683,069	19.2	8,534	1,761,094	30.9	
Others	39,950	8,124,927	33.3	4,861	987,588	17.3	
Monthly household income (million won)							
< 0.5	6,354	725,663	3.0	962	127,833	2.2	147.624**
0.5~1.0	11,466	1,432,910	5.9	1,971	264,739	4.6	
1.0~2.0	19,363	3,197,301	13.1	4,407	737,821	13.0	
2.0~3.0	24,489	4,881,712	20.0	6,103	1,226,208	21.5	
3.0~4.0	22,208	4,837,854	19.9	5,526	1,213,428	21.3	
4.0~5.0	15,220	3,541,153	14.5	3,718	860,796	15.1	
5.0~6.0	9,429	2,336,429	9.6	2,164	534,510	9.4	
≥ 6.0	13,233	3,419,029	14.0	2,979	729,916	12.8	
Marital status							
Married	36,924	8,458,594	34.7	7,721	1,803,399	31.7	60.7085**
Others	84,838	15,913,458	65.3	20,109	3,891,852	68.3	
Health check-up							
No	38,756	8,466,523	34.7	9,423	2,023,597	35.5	4.097*
Yes	83,006	15,905,529	65.3	18,407	3,671,655	64.5	
Experience of depression							
No	114,136	22,812,512	93.6	26,048	5,312,065	93.3	2.5624
Yes	7,626	1,559,540	6.4	1,782	383,186	6.7	
Subjective health status							
Bad	71,228	13,453,291	55.2	15,836	3,162,638	55.5	0.6535
Good	50,534	10,918,761	44.8	11,994	2,532,613	44.5	
Over-moderate physical activity							
No	93,719	18,985,090	77.9	20,358	4,229,206	74.3	104.8473**
Yes	28,043	5,386,962	22.1	7,472	1,466,045	25.7	
Subjective stress recognition							
No	90,439	17,625,857	72.3	18,880	3,713,607	65.2	335.6205**
Yes	31,323	6,746,195	27.7	8,950	1,981,645	34.8	
Currently smoking							
No	97,373	19,184,785	78.7	13,821	2,762,415	48.5	4779.969**
Yes	24,389	5,187,267	21.3	14,009	2,932,836	51.5	

Table 2. (Continued).

Variable	Normal			High-risk			χ^2
	N	Weighted (N)	%	N	Weighted (N)	%	
Obesity							
No	92,515	18,672,470	76.6	18,393	3,753,286	65.9	785.9349**
Yes	29,247	5,699,582	23.4	9,437	1,941,965	34.1	
Eating a breakfast regularly							
No	34,231	8,329,315	34.2	9,501	2,281,622	40.1	212.7888**
Yes	87,531	16,042,737	65.8	18,329	3,413,629	59.9	
Total	121,762	24,372,052	100.0	27,830	5,695,251	100.0	

* $p < 0.05$, ** $p < 0.01$

Table 3. AUC, Kolmogorov-Smirnov statistics, and misclassification rate for the predictive model.

	AUC	KS	Misclassification rate
Train data (60%)	0.7530	0.3967	0.1865
Validation data (40%)	0.7527	0.3999	0.1864

AUC = area under the curve; KS = Kolmogorov-Smirnov.

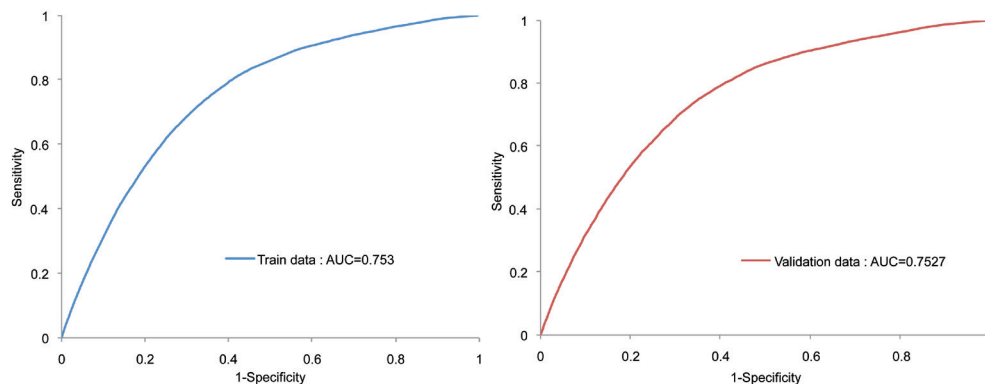


Figure 2. ROC curve for the predictive model.
AUC = area under the curve; ROC = receiver operating curve.

built and the validity of the models on the validation set were tested. The performance of the developed model was evaluated with respect to discrimination using the area under a receiver operating characteristic (ROC) curve, misclassification rate, and Kolmogorov-Smirnov statistics (Table 3).

The ROC charts are graphical displays that give the global measure of the predictive accuracy of the model (Figure 2). They display the sensitivity against 1-specificity of a classifier for a range of cut-offs. Sensitivity is a measure of accuracy for predicting events that is equal to the true positive divided by the total actual positive. 1-specificity is a measure of accuracy for predicting non-events that is equal to the true negative divided by the total actual negative. The performance of the models is demonstrated by the degree to which the ROC curves push up and to the left. The area under the curves can provide

a quantitative performance measure. The area will range from 0.5, for a worthless model, to 1, for a perfect classifier. The shapes of the ROC curves indicate that the predictive power of the model for predicting high-risk and normal is reasonably good (Figure 2).

Table 4 provides the parameter estimates of the risk prediction model for falling into high-risk group. The weighted logistic regression estimates revealed that men were significantly more likely to belong to high-risk group than women ($p < 0.01$). The parameter estimate for age groups showed that participants between the ages of 30 and 59 years old were significantly more likely to belong to a high-risk group than those aged under 29 ($p < 0.01$). Participants who had graduated from high school or had a lower level of education, were significantly more likely to belong to the high-risk group

than those who had graduated from university or had a higher qualification ($p < 0.01$). Participants who worked in business, sales and related occupations were significantly more likely to belong to the high-risk group than administrative employees ($p < 0.01$). The ORs of those participating in over-moderate

physical activity was higher than those without over-moderate physical activity. The ORs of person who had at least one of the following factors (smoking, depression experience, and stress relative to their reference group) were significantly higher [$p < 0.01$ (Table 4)].

Table 4. Result of weighted logistic regression analysis.

Variable	Category	$\hat{\beta}$	OR
Intercept		-3.0029**	
Gender (ref: female)	Male	1.2774**	3.59
Age (y, ref: 19-29)	30-39	0.1767**	1.19
	40-49	0.3399**	1.41
	50-59	0.3**	1.35
	60-69	-0.0289	0.97
	≥ 70	-0.6204**	0.54
Educational level (ref: University or higher)	Uneducated	0.5667**	1.76
	Elementary school	0.3488**	1.42
	Middle school	0.4102**	1.52
	High school	0.2919**	1.34
Occupation (ref: Administrative officer)	Clerical officer	0.23**	1.26
	Service and sales worker	0.1907**	1.21
	Farmer and fisher	0.0759	1.08
	Elementary work	0.1111**	1.12
	Others	-0.2007**	0.82
Monthly household income (million won, ref < 0.5)	0.5-1.0	-0.0398	0.96
	1.0-2.0	0.0564	1.07
	2.0-3.0	-0.00791	0.99
	3.0-4.0	0.0608	1.06
	4.0-5.0	0.0305	1.03
	5.0-6.0	0.0429	1.04
	≥ 6.0	0.0284	1.03
Marital status (ref: Other)	Married	0.0408	1.04
Health check-up (ref: No)	Yes	-0.0879**	0.92
Experience of depression (ref: No)	Yes	0.1715**	1.19
Subjective health status (ref: Bad)	Good	-0.0398	0.96
Over-moderate physical activity (ref: No)	Yes	0.0781**	1.08
Subjective stress recognition (ref: No)	Yes	0.2036**	1.23
Current smoking (ref: No)	Yes	0.7305**	2.08
Obesity (ref: Normal)	Obesity	0.2637**	1.30
Eating a breakfast regularly (ref: No)	Yes	-0.1995**	0.82

* $p < 0.05$, ** $p < 0.01$

3. Scorecard development

Scorecards for high-risk drinking were evaluated using the developed prediction model. In this study, the concept of point to double the odds (PDO), which is the most widely used scaling in the credit risk industry. For example, if PDO is set at 20, the odds of the person who receives 520 points through this method is twice as likely as those of the person who has 500 points. To make the scorecard, the adjusted coefficient was calculated by subtracting the smallest regression coefficient estimate from the assumed coefficient estimates of each variable to make the adjusted coefficient greater than or equal to zero. Then, the appropriate PDO was determined and the corrected regression coefficient transformed linearly into a single score as shown in Equation 1 [15,16].

$$\text{Score} = \text{adjusted coefficient} \times [\text{PDO}/\log(2)] \quad (\text{Equation 1})$$

In this study, PDO was set at 58.43994, and Table 5 showed the result of the scorecard (Table 5).

Table 2 also showed that males (score: 248.0), uneducated participants (score: 110.0), participants under 69 years of age (score: 114.8-186.4), and current smokers (score: 141.8) had scores higher than 100. For example, if an individual belongs in the following categories:

-
- ✓ a male in his forties (40-49 years)
 - ✓ uneducated
 - ✓ worked in a clerical office (business and financial operations occupations)
 - ✓ monthly household income of 1.0~2.0 million won married
 - ✓ without receiving health check-up
 - ✓ experience of depression
 - ✓ poor health status
 - ✓ experience of over-moderate physical activity stressed
 - ✓ current smoking
 - ✓ obesity
 - ✓ no regular breakfast
-

The total score will be 1,000 (248.0 + 186.4 + 110.0 + 83.6 + 19.5 + 7.9 + 17.1 + 33.3 + 7.7 + 15.2 + 39.5 + 141.8 + 51.2 + 38.7) and he will belong to the most high-risk drinking group.

Discussion

The high-risk drinking predictive model was developed in Korea using cross-sectional data from KCHS (2014). A total of 149,592 individuals were included in this study, and the weighted multiple logistic regression model was employed to develop the high-risk drinking predictive model. In addition, a scorecard for high-risk drinking can be used that was designed using the developed prediction model.

This study found that the major influencing factors for being a high-risk drinker were gender, age, educational level, occupation, whether they received health check-up, depressive symptoms, over-moderate physical activity, mental stress, smoking status, obese status, and regular breakfast. These findings were largely consistent with previous studies [17-19]. High-risk drinkers were more likely to be men in their thirties to fifties (30-59 years), or were office or sales workers. In particular, current smokers had an increased likelihood of high-risk drinking. However, monthly household income, marital status, and health status were not significantly related to the risk of falling into the high-risk drinking group in this study.

The results from the scorecard showed that the largest score range were found in the following factors: gender, age, educational level, and smoking status. In addition, the uneducated participants had the highest risk factor score according to the education level, and that of clerical officers was the highest according to the occupation category. For example, a male who is in his forties (40~49 years), uneducated, worked in a clerical office, and currently smoking, will score at least 769.9 points for high-risk drinking. In Korea, individuals with the above-mentioned factors are more likely to become involved in social relationships, which could increase the likelihood of high-risk drinking [11].

A scorecard is mainly used by credit rating agencies to measure consumers' credit so that the company prevent losses caused by consumers' activities such as taking loans, issuing credit cards and buying insurance, etc. This scorecard system can be applied, to quantify the scoring method, not only to help medical service providers to understand the meaning, but also to help the general public to understand the dangers of high-risk drinking more easily. In this respect, this study is meaningful. In addition, it can provide a basis for more effective healthcare services such as education to prevent high-risk drinking. In addition to the data used in this study, further refinement of the model by reflecting the local, social environment and geographical factors related to drinking is expected to enable the setting of various measures to solve and prevent high-risk drinking.

Finally, the scorecard modeling methodology will be helpful in measuring and understanding the level of health risk

Table 5. Result of scorecard.

Variable	Category	Score	Max score
Gender	Male	248.0	248.0
	Female	0.0	
Age (y)	19-29	120.4	186.4
	30-39	154.7	
	40-49	186.4	
	50-59	178.7	
	60-69	114.8	
	≥ 70	0.0	
	Educational level	Uneducated	
Elementary school		67.7	
Middle school		79.6	
High school		56.7	
University or higher		0.0	
Occupation	Administrative officer	39.0	83.6
	Clerical officer	83.6	
	Service and sales worker	76.0	
	Farmer and fisher	53.7	
	Elementary work	60.5	
	Others	0.0	
Monthly household income (million won)	< 0.5	7.7	18.7
	0.5-1.0	0.0	
	1.0-2.0	18.7	
	2.0-3.0	6.2	
	3.0-4.0	19.5	
	4.0-5.0	13.6	
	5.0-6.0	16.1	
	≥ 6.0	13.2	
Marital status	Married	7.9	7.9
	Other	0.0	
Health check-up	Yes	0.0	17.1
	No	17.1	
Experience of depression	Yes	33.3	33.3
	No	0.0	
Subjective health status	Good	0.0	7.7
	Bad	7.7	
Over-moderate physical activity	Yes	15.2	15.2
	No	0.0	
Subjective stress recognition	Yes	39.5	39.5
	No	0.0	
Current smoking	Yes	141.8	141.8
	No	0.0	
Obesity	Obesity	51.2	51.2
	Normal	0.0	
Eating a breakfast regularly	Yes	0.0	38.7
	No	38.7	

behaviors measured by various statistical models of health education program providers and users besides drinking.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

References

- [1] Maclure M. Demonstration of deductive meta-analysis: ethanol intake and risk of myocardial infarction. *Epidemiol Rev* 1993;15(2):328-51.
- [2] Chick J. Alcohol, health, and the heart: implications for clinicians. *Alcohol Alcohol* 1998;33(6):576-91.
- [3] Brewer RD, Swahn MH. Binge drinking and violence. *JAMA* 2005;294(5):616-8.
- [4] Green CA, Polen MR. The health and health behaviors of people who do not drink alcohol. *Am J Prev Med* 2001;21(4):298-305.
- [5] Rehm J, Room R, Graham K et al. The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease: an overview. *Addiction* 2003;98(9):1209-28.
- [6] World Health Organization. *Global status report on alcohol and health 2014*. Geneva (Switzerland): 2014. 376 p.
- [7] Hahm BJ, Cho MJ. Prevalence of alcohol use disorder in a South Korean community: changes in the pattern of prevalence over the past 15 years. *Soc Psychiatry Psychiatr Epidemiol* 2005;40(2):114-9.
- [8] Kim MK, Ko MJ, Han JT. Alcohol consumption and mortality from all-cause and cancers among 1.34 million Koreans: the results from the Korea national health insurance corporation's health examinee cohort in 2000. *Cancer Cause Control* 2010;21(12):2295-302.
- [9] Jang H, Jang WM, Park JH, et al. Alcohol consumption and the risk of type 2 diabetes mellitus: effect modification by hypercholesterolemia: the Third Korea National Health and Nutrition Examination Survey (2005). *Asia Pac J Clin Nutr* 2012;21(4):588-93.
- [10] Byeon H. Prediction modeling of high risk drinking in Korea using CRT method. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology* 2015;5(3):99-108. [in Korean].
- [11] Ryu SY, Crespi CM, Maxwell AE. Drinking patterns among Korean adults: results of the 2009 Korean Community Health Survey. *J Prev Med Public Health* 2013;46(4):183-91.
- [12] Choi R, Hwang BD. Use Characteristics of Health Examinations Services from Health Insurance Subscribers. *J Korea Content Assoc* 2011;11(2):331-440. [in Korean].
- [13] Joo AR, Park IH, Choi IH, et al. A Study of Health Concerns, Health Behavior, and Related Factors in Elders in Senior Centers. *J Korean Acad Fundam Nurs* 2007;14(4):500-6. [in Korean].
- [14] Kim RB, Park KS, Hong DY, et al. Factors Associated with Cancer Screening Intention in Eligible Persons for National Cancer Screening Program. *J Prev Med Public Health* 2010;43(1):62-72. [in Korean].
- [15] Kang HC, Han ST, Choi JH, et al. Methodology of data mining for big data analysis: A case study on SAS Enterprise Miner. Seoul (Korea): Free Academy; 2014. 154 p.
- [16] Siddiqi N. *Credit risk scorecards*. New Jersey (NJ): John Wiley & Sons; 2006. 113 p.
- [17] Balsa AI, Homer JF, Fleming MF, et al. Alcohol consumption and health among elders. *Gerontol* 2008;48(5):622-36.
- [18] Naimi TS, Brewer RD, Mokdad A, et al. Binge drinking among US adults. *JAMA* 2003;289(1):70-5.
- [19] Li Y, Jiang Y, Zhang M, et al. Drinking behaviour among men and women in China: the 2007 China Chronic Disease and Risk Factor Surveillance. *Addiction* 2011;106(11):1946-56.