Ph.D. Dissertation of Nayeon Kwon

Stress-sensitive personality and genetic risks as predictors of hypertension

– Focusing on the relationship among Type A, stress and genetic risk score–

고혈압의 예측 인자로서 스트레스에 민감한 성격과 유전적 위험도

August 2022

Graduate School of Humanity
Seoul National University
Cognitive Science Major

Nayeon Kwon

Stress-sensitive personality and genetic risks as predictors of hypertension –

Focusing on the relationship among Type A, stress and genetic risk score–

Hong-Gee Kim

Submitting a Ph.D. Dissertation of Science

August 2022

Graduate School of Humanity
Seoul National University
Cognitive Science Major

Nayeon Kwon

Confirming the Ph.D. Dissertation written by
Nayeon Kwon
August 2022

Chair Chung-Taek Kim (Seal)
Vice Chair Hong-Gee Kim (Seal)
Examiner Jae-Il Lee (Seal)
Examiner Jeong-Jun Im (Seal)
Examiner Soo-Kyoung Lee (Seal)

Abstract

In this study, Type A, stress level, genetic The influence on the development of hypertension through risk was analyzed by cross-section and longitudinal.

First, through cross-sectional data, Type A and perceived stress levels for the onset of hypertension, and the causal structure of obesity (BMI) estimated to mediate them, were analyzed by structural equation modeling. As a result, Type A in both men and women was observed only as an indirect effect on hypertension. Type A had a positive relationship with perceived stress level, and perceived stress had a negative effect on BMI, which had a significant positive effect on blood pressure. As a result, the higher the stress level, the lower the BMI level and the lower the blood pressure. This seems to be related to the lower food intake in the high-stress group compared to the low-stress group.

Also, in Cox regression analysis, which analyzed the effects of Type A, stress level, and genetic risk on hypertension with data from 2001 to 2006, and their interaction, only genetic risk and Type A were significantly risky. ratio was increased, and their interaction was not observed.

Finally, through Decision Trees and its extension algorithms, Random Forest and XGBoost, covariates and genetic risk are used as basic models, Type A and perceived stress are added to improve the performance of the predictive model. analyzed. As a result, overall predictive performance in female data was good, and the performance improvement was the highest when Type A was added. In the male data, there was an improvement when both Type A and perceived stress were added. In the distribution of feature importance through SHAP, a method of XAI, the contribution of Type A was high in the case of female data, and the perceived stress was mixed, preventing prediction. For male data, perceived stress contributed slightly more than Type A. In the end, it was confirmed that the genetic risk and Type A, which were

significant variables in the survival analysis, were used significantly in the predictive model with good performance, and did not follow when the predictive model did not perform well.

Through this, it can be seen that, in cross-section, Type A personality indirectly negatively affects hypertension, but in the long term, it has a significant effect as much as the genetic risk in the development of hypertension, and it is confirmed that it contributes significantly in the predictive model. .

Through this study, it was possible to identify a part of the causal structure of personality, stress, obesity and blood pressure, which can be used as a basis for a customized proposal to prevent hypertension not only at the genetic level but also at the behavioral level. That is, through this study, it is possible to understand the relationship between the factors of complex diseases and to contribute to the development of personalized medicine according to personality.

Keyword : Hypertension, Complex disease, Type A, personality, stress, SEM, survival analysis, CoxPH, NRI, IDI, XAI, SHAP

Student Number : 2017-37906

Table of Contents

List of Tables

List of Illustrations

Abbreviations

CAD	Coronary Artery Disease
CVD	Cardiovascular Disease
CHD	Coronary Heart Disease
HTN	Hypertension
BP	Blood Pressure
SBP	Systolic Blood Pressure
DBP	Diastolic Blood Pressure
HPA	Hypothalamic-Pituitary-Adrenal
CRH	Corticotropin Releasing Hormone
BMI	Body Mass Index
SNP	Single Nucleotide Polymorphism
SNV	Single-Nucleotide Variant
DNA	DeoxyriboNucleic Acid
GWAS	Genome Wide Association Study
CFA	Confirmatory Factor Analysis
SEM	Structural Equation Modeling
CR	Composite Reliability
AVE	Average Variance Extracted
PWI-SF	Psychosocial Wellbeing Index-short form
HWE	Hardy-Weinberg Equilibrium
MAF	Minor Allele Frequency
GCR	Genotype Call Rate
QC	Quality Control
NRI	Net Reclassification Improvement
IDI	Integrated Discrimination Improvement
XAI	eXplainable Artificial Intelligence
SHAP	SHapley Additive exPlanations

Chapter 1. Introduction

1.1. Study Background

1.1.1 Complex disease and Essential hypertension

Specific variants or mutations of single genes can cause diseases such as sickle cell disease, cystic fibrosis, or Huntington disease. These diseases are referred to as single-gene disorder or Mendelian disorder.

On the other hand, a number of common diseases are caused by not a single factor but a combination of genetic, environmental, and lifestyle factors. These diseases such as asthma, diabetes, hypertension, kidney diseases, autoimmune diseases, etc. are called complex diseases.

Complex disease is a disease that develops from a variety of causes due to the interaction of genetics, environment, and lifestyle rather than a single, distinct cause (Hunter, 2005). They are influenced by multiple genes and their interactions, lifestyles such as exercise, diet, drink, and environment like exposure to pollutant environments or job conditions.

Although complex diseases often run in families, they do not have a clear-cut pattern of inheritance. Since families not only share the composition of genes but also share the tastes or tendencies stemming from genes and lifestyles mostly under the same environment. That's why complex diseases are difficult to predict and treat. Moreover, genes involved in the disease are also complex and have interactions with different contributions.

Recently, methods such as polygenic analysis for finding a cluster of genes that have a statistically very small contribution but clearly affect them are being studied (Crouch & Bodmer, 2020; Khoury, Janssens, & Ransohoff, 2013). At the same time, gene by environment interaction studies are also actively conducted for

examining how they affect each other in the actual development of disease (Manuck & McCaffery, 2014; Ottman, 1996).

Hypertension is a representative complex disease. Hypertension or high blood pressure is defined as average systolic pressure is higher than 140mmHg or average diastolic pressure is higher than 90mmHg. Essential hypertension is hypertension whose causes are not apparent, in other words, its causes are not from mendelian forms such as renovascular disease, renal failure, aldosteronism etc. Essential hypertension accounts for 90~95% of all cases of hypertension.

Hypertension might be considered lightly because it has no specific symptoms. However, it is the first ranked risk factor for diverse diseases (WHO, 2019) in that it affects the homeostasis of other organs not only blood vessels but heart, brain or kidney, leading to various complications. Especially, hypertension is a fatal factor for cardiovascular disease (CVD) such as stroke, myocardial infarction, or heart failure (Mills, Stefanescu, & He, 2020), and for a long time CVD is the top leading factor for death in the world (WHO, 2021). Therefore, identifying not only the causes of hypertension but also the relationship among the causes can have the effect of preventing and managing more diseases, not just hypertension.

As a complex disease, the causes of hypertension are diverse and different for individuals. First of all, since there is a familial tendency, if both parents are hypertensive, 80% of their children can become hypertensive in the future, and if one parent is hypertensive, 25−49% of the children can become hypertensive (Robinson, Batisky, Hayes, Nahata, & Mahan, 2005). Also, essential hypertension is heterogeneous, with different patients having different causal factors depending on lifestyles and environment.

According to Korea hypertension fact sheet 2021, from 1998 to 2019, hypertension has a prevalence of about 28% (12 million) among adults over the age of 20 in Korea (H. C. Kim et al., 2022), and this rate is similar to that of the world accounting for about 31%

of the global population (Mills et al., 2020).

As described above, exploring the causative factors of hypertension and understanding the relationship among them is valuable in that it can prevent and manage hypertension itself, as well as prevent other diseases that hypertension affects.

1.1.2. Etiological factors in Essential hypertension

1.1.2.1. Genetic factors

Genetic factors or predispositions are defined as a relatively high likelihood of developing a particular disease based on a person's genetic traits. Genetic factors result from specific genetic variations that are often inherited from a parent.

The genetic variants of someone can be found through the comparison of the nucleic sequence of DNA and that of the reference DNA. Whether the variants, referred to as SNV (single nucleotide variation) or SNP (single nucleotide polymorphism), are associated with a specific disease can be determined through association studies such as Genome Wide Association Study (GWAS). The GWAS results indicate the possibility that some variants of the specific DNA position could be related to the disease.

An individual's risk for developing the disease can be calculated by these associated variants, which is called Genetic Risk Score (GRS) (Janssens & van Duijn, 2008).

It is difficult to say with certainty the genes with mutations are the cause of disease in that the genes with the mutations vary by race and are difficult to calculate for the entire population. However, causative gene information has been accumulated through repeated research results, it has become possible to predict the onset of individual diseases.

The known genes affecting the onset of hypertension are KCNK3

(Manichaikul et al., 2016), Adamts16 (Joe et al., 2009), TBX4 (Hernandez-Gonzalez et al., 2020), KCNJ1, ROMK, CASR, NR3C2, SCNN1B, and SCNN1G (Tobin et al., 2008), especially for Asians, ATP2B1 (Tabara et al., 2010), CYP2C9 (Yu et al., 2004) etc.

Of course, these variants of genes are not the only factor for hypertension but it contributes to a higher odds ratio than those who don't have the variants. Currently, genetic factors could explain 30% to 50% of interindividual variability in HTN (Lalouel, 2003; Levy et al., 2000). In effect, this factor can interact with an individual's environment or lifestyle, which can affect the development of the disease comprehensively.

1.1.2.2. Non-Genetic factors

It is known that a number of non-genetic factors increase Blood Pressure, including obesity, insulin resistance, high alcohol intake, high salt intake (in salt-sensitive patients), low potassium intake, aging, sedentary lifestyle, stress etc. Moreover, many of these factors are additive, such as obesity and alcohol intake (Adrogué & Madias, 2007; Sever & Poulter, 1989).

First of all, diet is an important factor. According to the research results of The International Study of Salt and Blood Pressure (INTERSALT), a general daily sodium intake was between 50 and 100 mmol, for adults aged 25-55 years, for every 50 mmol increase in daily sodium intake, systolic blood pressure increased by 5 mm Hg and diastolic blood pressure increased by 3 mm Hg (Intersalt Cooperative Research Group, 1988).

Also, it is well-known that potassium intake in a diet is also important in controlling blood pressure. Potassium works to excrete sodium remaining in our body and suppresses the activity of renin, a hormone that raises blood pressure. In the same study of INTERSALT as above, decreasing daily potassium excretion by 50 mmol increasing systolic pressure of 3.4 mm Hg and diastolic

pressure of 1.9 mm Hg (Intersalt Cooperative Research Group, 1988)

Gender is another factor for developing and taking care of hypertension. According to the meta-analysis results of Song et. al., the prevalence of hypertension was higher in men and awareness and treatment are higher in women in many societies including the USA and China (J.-J. Song, Ma, Wang, Chen, & Zhong, 2020).

Recently, socioeconomic status (SES) including education level is considered as an important factor since a volume of studies revealed that low SES is associated with higher blood pressure (Grotto, Huerta, & Sharabi, 2008; Leng, Jin, Li, Chen, & Jin, 2015; Wu et al., 2013).

In addition, as an environmental factor, resident area could also influence hypertension. Although the results were not consistent, hypertension was mostly more prevalent or increasing in rural areas than in urban areas (Gupta, 2016; Shen et al., 2019). The researchers estimated the reason is the difference of awareness, access to treatment or diagnosis, and changes of dietary life.

Moreover, psychological stress has long been recognized as an environmental factor that can control blood pressure. Stressful stimulation induces awakenings of a protective alert along with secreting adrenomedullary hormone repeatedly, which makes individuals vulnerable to high blood pressure (Henry & Cassel, 1969).

To sum up, the non-genetic factors affecting hypertension are various from the traditional factors to natural environment and social environment factors. The good point is those factors can be controlled by an individual's lifestyle.

1.1.2.3. Gene by Environment Interaction

Although environmental factors such as diet, exercise, or psychological stress play an important role in the pathogenesis of hypertension, individual reactions may vary depending on the differences in the genes that determine the physiological system. In

other words, even if people have the same environmental factors, whether the disease occurs or not can be depending on the genetic makeup of their own not even directly but also indirectly through other chain reactions to the stimuli.

As a representative example, there is a meta-research on the interaction between the genes involved in asthma and the environment. This study suggested that mutations in certain genes involved in asthma can be functionally silenced until carriers of those mutations are exposed to relevant environmental stimuli that promote cell differentiation (allergens and/or parasites in the case of IL13). Both functional changes and environmental stimuli that modify the cellular environment are required for the interaction to occur (Ober & Vercelli, 2011).

It is clear that the physiological circuitries of populations with vulnerable genetic mutations are also more sensitively affected by prenatal and postnatal environmental factors including dietary adequacy and endocrine disruptive chemicals (Jackson, 2014), also including metabolic related traits such as insulin resistance (Pausova, Tremblay, & Hamet, 1999).

In the case of hypertension, a population who has genes sensitive to sodium may be more prone to developing it with high sodium intake (Jin et al., 2010; Nakagawa et al., 2006). Especially under the sodium rich diet culture of the society like Korea, individuals having the gene makeup would be more vulnerable to the onset of hypertension.

1.2. Questions and Purpose of the Research

1.2.1. The motivation and Research Questions

As described above, complex diseases are affected by the complex interaction of genetic factors and non-genetic factors such as environmental factors and lifestyle. As a representative complex disease, hypertension is also influenced by genetic factors and non-

genetic factors.

Thus, although it is important to discover genes that are sensitive to hypertension or vulnerable environments, if we know the relationship and flow of how and to what extent they are affected by environmental factors, the onset of hypertension can be effectively prevented by controlling the environment or lifestyle. It is also helpful to delay or manage the disease.

With this motivation, this study will analyze the interaction effects of psychosocial and environmental factors such as psychological stress on people who are genetically vulnerable to hypertension. And then, by inferring causal relationships among non-genetic factors, we want to lay the foundation for prevention through effectively adjusting lifestyles.

In detail, the first question of the current study is how non-genetic factors affect each other on the development of hypertension. In particular, a personality which is easy to get stressed can accumulate frequent stressful experiences. psychological stress is an emotional state through an individual's cognitive arousal, and its effect on the disease will be different depending on the individual's cognitive sensitivity. Such individual differences lead to the question of whether a stress-sensitive or vulnerable personality may be more susceptible to stress-mediated diseases. Even in the same situation, people with a more sensitive personality or tendency to stress are likely to react to stress more easily or more frequently, and thus it is plausible that the effect of stress would increase for these people.

Subordinately, examining whether psychological factors represented by stress actually affect the onset of physical diseases would tell the no separation of mind and body, which has influenced modern philosophy and cognitive science (Jaegwon Kim, 1999). Therefore, it would be meaningful as an example of biological observation of monism.

In addition, not only the direct relationship with hypertension, stress is known to be highly correlated with obesity, and obesity is a major risk factor for hypertension. Considering these relationships,

it is also necessary to investigate whether obesity is involved between stress and hypertension.

Furthermore, this study will investigate whether those non genetic factors interact with genetic factors based on genetic risk scores.

Lastly, this study will confirm whether those factors improve the prediction of the development of hypertension through machine learning algorithms.

To sum up, this study will analyze the relationships among factors related to a personality on the development of hypertension cross-sectionally as well as longitudinally, and their interaction with genetic factors. Then, we will try to figure out whether personality has value as a predictor of hypertension.

First, we will identify the definitions and characteristics of selected non-genetic risk factors – stress, personality, and obesity, and examine studies related to these factors on hypertension.

1.2.2. Stress as a risk factor for HTN

The concept of stress has been utilized and studied since its first use in physiological and biomedical research by Selye (Selye, 1957). Although the definition of stress still has been ambiguous (Romero, Dickens, & Cyr, 2009), it can be explained by the relationship between the stressor and stress response in general. A stressor means a stimulus that threatens homeostasis such as dangerous situations or unusual events and the stress response is the reaction of the stressor psychologically or physiologically (Chrousos, 2009).

The stress response can be evaluated from three perspectives; physiological responses, psychological responses, and behavioral responses.

The physiological response to stress is related to the activation of Hypothalamus-Pituitary-Adrenal (HPA) Axis and Autonomic

Nervous System (ANS) (Selye, 1976). Negative sensations of stress are transmitted to the hypothalamus to elicit autonomic and endocrine responses. As a result, along with corticotropin-releasing hormone (CRH), catecholamines such as norepinephrine and adrenocortical hormones such as cortisol are released into the blood, affecting cardiovascular, digestive and endocrine organs. In detail, these hormones stimulate the sympathetic nerve, causing the body's response to increase heart rate and blood pressure.

The psychological response is referred to as psychological distress or emotional distress (Brotman, Golden, & Wittstein, 2007) and used as a psychological phenomenon. In this concept, self-perception of feeling discomfort to a specific stressor is an important aspect (Ridner, 2004). Since psychological distress is based on individual perception, it can be measured by questionnaires (Shin, 2013).

The behavioral response means specific behaviors as a response to a stressor. Verbal and facial expression for humans as well as scratching or lip biting also appeared for non-human primates (Troisi, 2002) as responses.

The stressful circumstances that elicit these responses may be temporary, but they can persist or recur. Transient stress reactions can be classified as acute stress, and sustained or recurrent stress as chronic stress. More officially, the definition of Acute stress is the normal short-term physiological response to the perception of major threats or demands, and that of Chronic stress is the abnormal ongoing physiological response to the continuing perception of unresolvable major threats or demands (Eggers, 2007). In detail, anxiety, depression, uneasiness, anger, apathy, and alienation are emotions that commonly accompany chronic stress (Cohen, 2000).

Namely, chronic stress indicates that the individual experiences persisting stress responses and the accumulation of the responses affects the onset of complex diseases such as depression, the

metabolic syndrome, and essential hypertension etc. (Sparrenberger et al., 2009) and those diseases also influence each other. In short, persistent or recurrent stress responses can not only directly influence the onset of hypertension but also be a mediator leading to hypertension.

Particularly, since daily life hassles or continuous stressful events depends on individuals' perception and their cognitive appraisal, it is valid to investigate the relationships among perceived stress and hypertension (Bhelkar, Deshpande, Mankar, & Hiwarkar, 2018; Gerin et al., 2012).

For this reason, it is valid to stress refers to the perceived stress in daily life in the current study.

1.2.3. Personality type as a risk factor for HTN

Since the study result that Type A behavior pattern was related to coronary heart diseases were reported in 1959 (Friedman, 1974; Friedman & Rosenman, 1971; Friedman & Rosenman, 1959), there have been many studies investigating the relationship between Type A and heart related diseases such as cardiovascular disease (CVD)(Pollock, Chen, Harville, & Bazzano, 2017), coronary heart disease (CHD)(Myrtek, 2001), essential hypertension (HTN)(Diamond, 1982; Irvine, Garner, Craig, & Logan, 1991) and acute coronary syndrome (ACS)(Čatipović-Veselica, Glavaš, Kristek, & Šram, 2001).

According to Friedman, Type A behavior pattern is described as an individual whose characteristics include impatience, exaggerated competitiveness, a chronic sense of time urgency, aggressive drive and hostility. These characteristics, especially their hostility and time-urgency, seem to stem from poor self-esteem (Friedman & Rosenman, 1959).

Type B behavior pattern is referred to as the one which is not Type A. Thus, individuals with Type B are relatively easygoing or

generous. According to Houston (Houston, 1983), individuals with Type A comparing those with Type B showed greater psychological arousal under the situations in which other people annoy or harass them, which could say they tend to be vulnerable to stress.

On the same direction, the 8.5-year-follow up study of Western Collaborative Group Study, Type A was associated with CHD (Rosenman, Brand, Sholtz, & Friedman, 1976) as well as angina and myocardial infarction (Haynes, Feinleib, & Kannel, 1980). Especially Type A with poor self-control is vulnerable to stress and at risk for CHD in the study of Heilbrun & Friedberg (Heilbrun Jr & Friedberg, 1988).

As following research suggested that the subcomponents of Type A, such as anger-in, hostility, and impatience, are considered as toxic components which play a significant role in CVDs and HTN, the effect of Type A has been focused on those subcomponents in particular (Diamond, 1982; Larkin & Zayfert, 2004; Sehgal, 2000).

Nevertheless, as the different or insignificant study results from those of Friedman were also reported, the relationships have been inconsistent. For example, in the meta-analysis of studies conducted from 1966 to 1998, the relationship between Type A and CHD was not significant. Although hostility was related to CHD, the effect size was low (Myrtek, 2001).

Likewise, the study results about the association between Type A and HTN have been inconsistent.

In a study of 221 hypertensive patients and 221 controls by Al Asadi, the proportion of Type A behavior was significantly higher in hypertensive group than the control group (Al Asadi, 2010). In addition, in the study of 100 hypertensive patients and 100 normotensive people in India, hypertensive patients had higher scores in Type A as well as stress, depression and anxiety (Sharma, 2012). On the other hand, there are also the opposite results. In the study of Mann & Brennan, they didn't find the association between Type A and CVD (Mann & Brennan, 1987), and

recent studies show the similar results (Ducher, Fauvel, & Cerutti, 2006).

Nevertheless, it is still valuable in that specific personality traits could negatively influence cardiac health including HTN (Molinari, Bellardita, & Compare, 2006).

Other personality types which affect cardiac diseases are Type C and Type D (Denollet, 1991). According to Denollet, Type C individuals are relatively more cooperative, unassertive, tending to suppress negative emotions. In contrast, Type D or "Distressed" personality is based on two traits, negative affectivity and social inhibition (Denollet, 1998). Type D individuals are described as often worried without reasons, have a pessimistic perspective, easily feel depressed and irritated, and rarely experience positive feelings. Diverse studies along with Denollet's have argued that Type D personality is related to stress and cardiac diseases including hypertension (Denollet, 2000; Grande, Romppel, & Barth, 2012; Sher, 2005).

Type D is associated with depression and anxiety, as inferred from the word "Distressed", and the relationships between Type D and diverse diseases including essential hypertension and CVDs etc. have been examined actively until recently (Breik & Elbedour, 2021; Oliva et al., 2016; Svansdottir et al., 2012).

Especially with regard to stress, the hypothesis that some specific personalities, e.g. Type A or Type D could be sensitive to stressors has been confirmed by experiments, which suggests that the personality can be a predictive factor for stress-induced diseases such as essential hypertension.

1.2.4. Obesity as a risk factor for HTN

Obesity or being overweight is one strong risk factor for essential hypertension and other comorbid diseases including cardiovascular disease (Jiang, Lu, Zong, Ruan, & Liu, 2016; Rahmouni, Correia, Haynes, & Mark, 2005). A number of studies have been conducted for investigating the association between obesity and hypertension in various views.

A few causes of high arterial pressure in obesity are known as the abnormal activation of the sympathetic nervous system and the renin–angiotensin–aldosterone system. Also, obesity is related to endothelial dysfunction and renal function, which can affect the onset of hypertension. This circularity starts from the structural changes inside of the kidney and increased abdominal pressure due to obesity, which causes renal sodium reabsorption disorder. The disorder increases arterial blood pressure and metabolic abnormalities as well as other factors including inflammation, oxidative stress, and lipotoxicity etc. and is positively reinforced via vicious circle (Kincaid–Smith, 2004).

Although obesity as an independent disease has genetic causes, the main cause is the imbalance between energy intake and expenditure (Jiang et al., 2016). Especially, the continuing discovery of mechanisms regulating appetite and metabolism can explain obesity–induced hypertension (Rahmouni et al., 2005).

In addition, many paths connect stress to obesity. Mostly the pathways explain stress induce obesity by impeding self–regulation, by overeating, by ingesting high calorie foods, by changing reward processing in the brain, and by reducing sleep time affecting secretion of hormone and peptide such as leptin or ghrelin which balances appetite (Koch, Sepa, & Ludvigsson, 2008; Tomiyama, 2019).

Although a huge volume of research found the positive relationship between stress and food–intake, counterfactual results especially for animals also exist. For example, when yellow rats

mutated to dominate obesity and normal rats were exposed to repeated emotional stress, stress increased basal and stress-induced concentrations of corticosterone in both rats, and decreased food intake to hamper development of obesity (Bazhan, Makarova, AIu, & Iakovleva, 2007; Solomon, Foster, Bartness, & Huhman, 2007).

In research on human, acute stress suppressed the increase of appetite in non-obese participants (Nakamura et al., 2020) and some people lost their appetite when got emotionally stressed (Kandiah, Yake, & Willett, 2008).

The reason why it was mentioned as 'some people' is because there have been many recent studies that consider this to be the effect of individual differences (Emond et al., 2016; Maniam & Morris, 2012).

According to the study of Maniam & Morris, the mechanism of people increasing appetite when they are stressed and that of people who are losing appetite under a similar situation is different. They suggest the stress responses are regulated by the reward pathway involving serotonin and dopamine, whose degree of activation can be different.

Therefore, it is necessary to investigate the relationship between stress and obesity in that not only both affect the development of hypertension, but also the two factors closely influence each other.

1.2.5. Purpose and Expected contribution of the study

As we have seen in each of the above risk factors, these factors may influence hypertension independently as well as they interact with each other by giving and receiving effects.

As summarizing the studies on the relationship among the above variables affecting hypertension, first, there are many studies that personality such as Type A is more vulnerable to stress, and it was

significantly higher in the hypertensive group.

In particular, acute stress stimulates the sympathetic nervous system and results in an increase in blood pressure, and the fact that general stress is higher in the hypertensive group shows the possibility of having effects on the development of hypertension when the stress response leads to chronicity.

Meanwhile, stress is also related to eating habits. In most cases, when individuals get psychologically stressed at work or in daily life, they eat a lot of comfort food and highly palatable foods for avoidance or compensation, which easily leads to obesity (Sinha & Jastreboff, 2013; Torres & Nowson, 2007).

Of course, these findings are not consistent. There are studies showing that Type A personality did not have a direct relationship with blood pressure, or studies showing that only anger, irritability, and hostility, which are sub-components of Type A, have toxic effects on cardiovascular diseases including hypertension.

In addition, there are also studies arguing stress does not just lead to overeating, but also causes loss of appetite depending on the person, so that the higher the stress level, the weight loss. Also, there are studies in which the relationship between stress and hypertension is not significant.

While the studies that have been conducted are numerous and inconsistent, there are few studies on the overall relationship among the factors, and no in-depth studies have been investigated on how these lifestyle-related environmental factors interact with genetic factors.

Thus, in order to expand these studies, this study will analyze the relationship among the key variables mentioned above and how they interact with genetic risk.

This is also connected to the research problem presented in Section 1.2.1, and this study has the following research objectives.

1. By investigating how much personality, stress, and obesity can affect essential hypertension, the current study will reveal relationships among the factors in terms of personality and perceived stress in the short-term period.

2. By verifying those factors affect in the long-term period and whether those factors interact with genetic factors, the current study will investigate a part of gene by environment interactions

3. By analyzing short term and long term data, whether a personality type and perceived stress can be significant predictors for the development of hypertension will be confirmed.

The conceptual model for this is as follows.

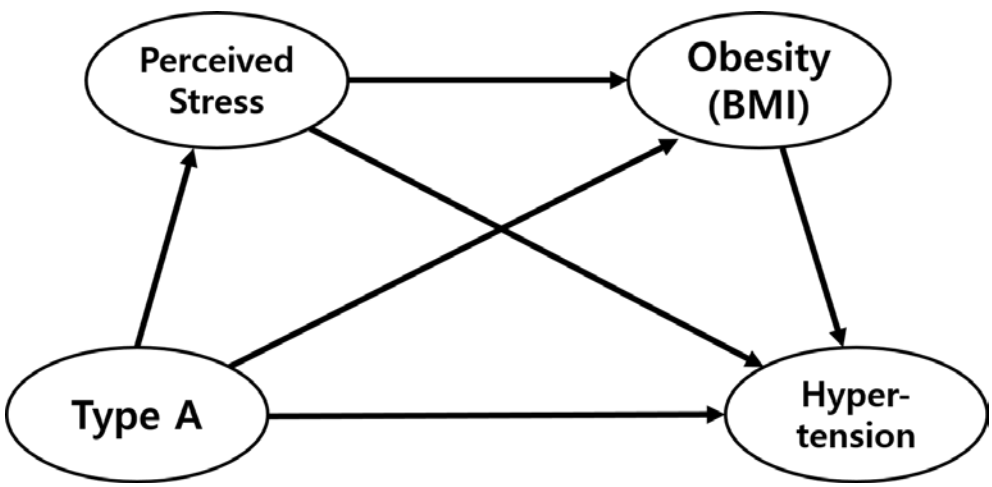


Figure 1.1 The conceptual model of the relationship for blood pressure

By achieving this objective, the current study can provide lifestyle information that can be adjusted to prevent hypertension according to individual differences such as personality. This contributes to prevention medicine in two aspects.

First, if it is found that individual differences, such as a

personality vulnerable to stress, affect disease, it may qualify as a factor to consider such characteristics in personalized medicine.

Second, by revealing how stress occurring at work or at home interacts with obesity or blood pressure, it can be helpful to know the relationship with eating habits and to control them at the environmental or personal level.

All of these have the characteristics of prevention and management. Therefore, through the results of this study, it will be possible to present a valid basis for diseases and management of lifestyle-related complex diseases.

1.3. Methodologies and Organization

The research plan to achieve the research purpose is as follows.

Data. Received data from the Korea Centers for Disease Control and Prevention. The data include epidemiological information (general, Type A score, stress score), biochemical measurements, daily food intake (including salt) from food intake frequency survey, lifestyle information such as smoking, drinking and exercise, socioeconomic status data like education level, monthly income, etc. and their genome information.

1. Conducting structural equation model analysis to infer the causal relationship between key factors, Type A – perceived stress – obesity – hypertension.

2. SNPs with high association with hypertension are extracted using GWAS with the genome information. Then, calculating the odds ratio of each SNP related to hypertension as determined after association analysis with the GWAS results.

3. Computing weighted genetic risk score (wGRS) based on the result of 2.

4. Analyzing main effect and interaction effect of perceived stress as well as Type A and wGRS on the developing hypertension through Survival Analysis (Kaplan Meier estimation and Cox proportional hazards model analysis) with the biennial data from the baseline (2001−2002) to the 7th period (2015−2016).

5. Evaluating what extent to improve the prediction model performance by Perceived Stress and Type A (Tree based models).

After the analysis, the results will be discussed about the relationships among the non−genetic key factors and the validity of using personality for forecasting the onset of hypertension.

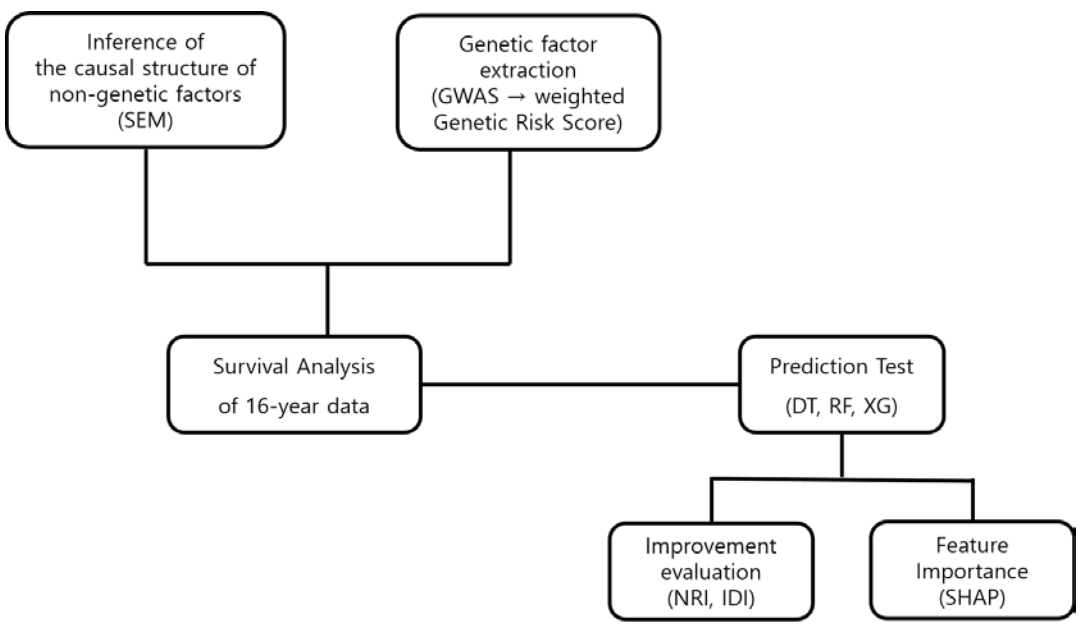


Figure 1.2 The overall structure and process of the study

Chapter 2. Methods

2.1. Community cohort data of Korean genome research project

Korean Genome and Epidemiology Study; KoGES is a cohort project conducted by the National Institutes of Health, Korea Centers for Disease Control and Prevention, to establish a scientific basis for the implementation of customized preventive medicine by identifying risk factors for chronic common diseases to Koreans.

KoGES consists of a 'population-based cohort' and a 'gene-environment model cohort' to identify risk factors for genetic-environment interaction in chronic diseases. Since 2001, approximately 235,000 participants have been recruited for the basic survey, and repeated follow-up surveys every two or four years for tracking.

Community cohort is a part of population-based cohort consisting of residents in An-seong and An-san. Using the colony extraction method, a representative sample of residents aged 40 to 69 was recruited starting from 2001 and 2002 (total of 10,038 people) and tracked every two years.

Along with health and lifestyle-related surveys, human materials were collected such as blood, urine, and DNA etc. The KoGES data basically include questionnaire items like general matters, disease history, disease treatment status, drug history, family history, lifestyle habits such as drinking and smoking, and physical activity, as well as clinical trials such as blood pressure and pulse measurement, body composition analysis, blood and urine tests, etc. It also consists of examination items that include electrocardiograms, chest x-rays, and lung function tests.

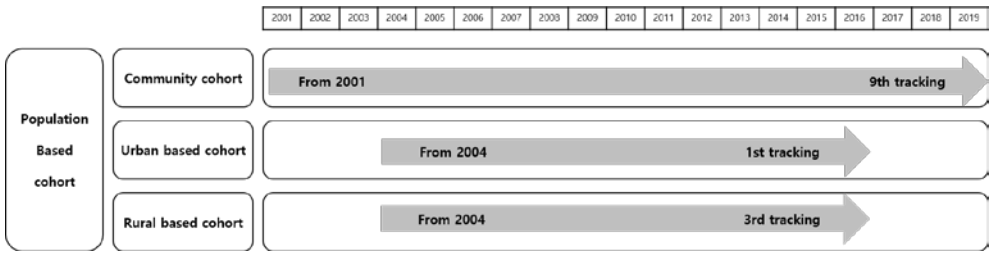


Figure 2.1 The community cohort data composition and tracking status

The current study used the initial (2001−2002) data which include personality information such as Type A and third (2005−2006) period data, which include psychological status information such as psychosocial well−being index short form (PWI−SF) and blood pressure information. Since the data of this cohort have been accumulated for the general health information of the two areas, not for the specific diseases, the disease diagnosed samples were relatively smaller than the normal ones. Therefore, in this study, the groups were not divided into the patient and the normal, and the continuous measurement values of SBP and DBP were used.

Among the total of 10,030 initial samples, the samples missing blood pressure values were removed and there remained 4308. In the remaining data, every missing valued sample for variates and covariates were got rid of during the tracking period. Those who were diagnosed as hypertension and who's blood pressure was over 140 mmHg for SBP or 90 mmHg for DBP and taking drugs for hypertension were excluded from the base year samples and the second period (2001−2004) samples. Taking these steps could not only limit the effects of already hypertensive people but also focus on the effect of perceived stress on the development of hypertension after getting the stress information.

Hence, variables were used from the third period data except for Type A (initial); perceived stress, drinking, smoking, exercising, food and sodium intake, socioeconomic data and food intake. These were used for the cross−sectional analysis i.e. SEM (for the third

period) as well as for the longitudinal analysis, Survival analysis, with tracking blood pressure and BMI information.

In the longitudinal analysis, hypertension was defined as over 140 mmHg for SBP or 90 mmHg for DBP, diagnosed as hypertension or taking drugs for hypertension from the third period.

The total of used samples was 3579.

2.2. Measurements and Statistical Analysis

2.2.1. General information

The range of the age in the cohort was from 40 to 70. The mean age of the male group and the female group was 50.85 (sd=8.01) and 51.75 (sd=8.45), respectively. The number of males was 1668 and one of females was 1911.

In terms of socioeconomic status, there were 4 classes depending of monthly income; below ₩1,000,000, ₩1,000,000~₩3,000,000, ₩3,000,000~₩6,000,000, and over ₩6,000,000. We categorized these into Over Middle Group (over ₩3,000,000) and Under Middle Group (under ₩3,000,000).

Final education status was composed of 5 categories, from lack of schooling, below high school graduation, high school graduation, college graduation, to university graduation or higher. These were categorized into Over Middle Group (high school graduation or higher) and Under Middle Group (less than high school graduation) based on the sample aged generation in Korea.

2.2.2 Type A Behavior Pattern

The questionnaire for testing Type A consisted of 10 questions based on Jenkins Activity Survey (Begley & Boyd, 1985; Matteson & Ivancevich, 1980). The contents of questions were whether being impatient, competitive, short-tempered, compulsive to time, easy to

anger, etc. For example, “I am impatient waiting in line” or “Others rate me as hard-driving and competitive”, etc. The answer range was 4-scale from Never (1) – Barely (2) – Often (3) – All the time (4), total range is from 0 to 30. When dividing into Type A and Non-Type A categorical types, the median value between the minimum and maximum values (Begley & Boyd, 1985), which was 14 points. The questionnaire items are attached in the appendix.

2.2.3. Perceived Stress

To measure perceived stress, Psychosocial Well-being Index – Short-Form (PWI-SF) was used (Jang, 2000). PWI-SF is an 18-item-short form of PWI developed to measure the degree of perceived stress for the Korean population based on Goldberg's GHQ (General Health Questionnaire)-60, approved of reliability and validity (Goldberg, 1978; Lee & Lee, 1996). Although well-being and perceived stress are not in the perfect counterpart, considering the definition of stressor (threatening one’s psychological well-being), this measure could reflect perceived stress as the degree of damage to their well-being. Besides, according to the study of Winefield et. al., it is obvious that psychological well-being is negatively associated with psychological distress (Winefield, Gill, Taylor, & Pilkington, 2012), hence it is feasible that the opposite scale of the well-being index is considered as a measure of perceived stress.

The contents of the questionnaire consist of vitality, sleep disorder, social role performance, self-confidence, and depression, which are similar to those of Perceived Stress Scale (S. Cohen, Kamarck, & Mermelstein, 1994). This questionnaire was developed to assess the socio-psychological distress degree of workers or residents and has been used as a tool for examining the correlation between perceived stress risk factors and disease risk factors (Jang, 2000; D. Kim et al., 2017; J. Y. Kim et al., 2021; Seo, Kim, & Ha, 2010).

The answer was 4-point Likert, 0 (Never) – 1 (Barely) – 2 (Mostly) – 3 (Always), total range is from 0 to 54. In case of using categorical data, the data were divided based on the total score. Although The proposed cutoff values for the level of stress were healthy group (score ≤ 8), potential stress group (score 9–26), and stress group (score ≥ 27)(Jang, 2000), only two groups were divided in the current study; high stress group (score ≥ 27) and low stress group(score ≤ 26). The questionnaire items used are attached in the appendix.

2.2.4. Anthropometry and Biochemical measurement

Before blood pressure measurement, the participants were let rest at least for 5 minutes, the average value was used by measuring the left and right arms twice with a time difference of at least 5 minutes in the sitting position. Systolic blood pressure (SBP) and Diastolic blood pressure (DBP) were measured by an experienced nurse with a mercury sphygmomanometer (Baumanometer) using a cuff of an appropriate size around the subject's arm.

Obesity is diagnosed by measuring the weight in relation to the height of an individual, thereby determining or calculating the body mass index. The National Institutes of Health have defined 30 kg/m2 as the limit over which an individual is qualified as obese (Jiang et al., 2016). Therefore, BMI was used as an index to determine the degree of obesity in the current study.

Height and weight were measured to the 0.1 cm and 0.1 kg without shoes, and BMI was calculated with weight in kilograms divided by height in meters squared.

The total cholesterol was extracted from serum using ADVIA 1650 equipment.

2.2.5. Food and sodium intake

The amount of sodium intake per day was computed based on the refined food intake frequency survey data. Also, the total amount of food per day was calculated from the refined food intake frequency survey data for weight(g) from the third follow-up data, which is the same time point as perceived stress measurement.

2.2.6. smoking, drinking, exercise

Items of the questionnaires to investigate smoking were 1) Never 2) Past 3) Present sometimes 4) Present heavily. The measure of drinking was 1) Never 2) Past 3) Present.

In terms of exercise, the question was how often the individual exercised vigorously and the options were following; 1) Never 2) Once a week 3) twice to three times 4) More than five times.

2.2.7. Statistics

For the group comparison test by gender, continuous variables were used such as age, perceived stress, Type A, BMI, SBP, DBP, Sodium intake, Total cholesterol level and Amount of food intake.

The normality of the distribution was confirmed through the Shapiro test. Since all variables did not follow a normal distribution, Wilcoxon rank-sum test, a nonparametric test, was performed.

Likewise, continuous variables were used in the analysis of Structural Equation Model (SEM) and Prediction Model Tests.

Meanwhile, Survival analysis was conducted with categorical data divided into two levels according to the criteria for each variable.

To conduct GWAS and calculate weighted genetic risk score, PLINK 1.7 (Purcell et al., 2007) was used on Ubuntu 18.04 server.

For the survival analysis, Survival (Therneau & Lumley, 2013),

Survminer (Kassambara, Kosinski, Biecek, & Fabian, 2017), and gtsummary (Sjoberg, Whiting, Curry, Lavery, & Larmarange, 2021) were used in R.

CFA and SEM were computed by Lavaan (Rosseel, 2012) and semTools (Jorgensen et al., 2018) in R.

For computing statistics and Machine Learning Models, Scipy (Virtanen et al., 2020), scikit-learn (Pedregosa et al., 2011), and shap (Lundberg & Lee, 2017) were used in python.

2.3. The inference of the causal structure among non-genetic psychological relevant risk factors

Structural Equation Modeling consists of two steps; the measurement modeling and the structural modeling.

In the measurement modeling, confirmatory factor analysis was performed to determine whether the model fits to the data and screen out invalid observation variables.

In the structural modeling, path analysis was conducted. The influence of the independent variable on the dependent variable (total effect) can be identified by analyzing whether the independent variable directly affects the dependent variable (direct effect) or is mediated through other variables (indirect effect).

If the fit indices of this model are valid, it has validity that the causal structure of each variable follows the model shape.

2.3.1 Concept model

As mentioned in the section of the Purpose and Expected contribution of the study, the concept of the current model is the relationships among the non-genetic and psychological relevant factors affecting hypertension.

First of all, based on the review of related research papers, Type A people are more likely to develop cardiovascular disease, including

high blood pressure (Čatipović-Veselica et al., 2001; Diamond, 1982; Irvine et al., 1991). Besides, Type A people respond more sensitively to stressful situations than those who do not (Heilbrun Jr & Friedberg, 1988; Houston, 1983), so it can be assumed that their response threshold is low and the frequency of recognizing stress is high .

Secondly, there is a huge volume of research that psychological stress induces binge eating or overeating (Tomiyama, 2019), which is the cause of obesity. However, the results are not consistent, there are also studies showing psychological stress reduces appetite or food intake (Kandiah et al., 2008; Nakamura et al., 2020). Either way, it is presumable to assume that perceived stress influences obesity.

Lastly, it is obvious that obesity influences the development of hypertension (Jiang et al., 2016).

Therefore, the hypotheses are following;

Hypothesis1: Type A would affect the degree and frequency of perceived stress and the onset of hypertension.

Hypothesis2: Perceived stress would influence the onset of hypertension and obesity.

Hypothesis3: Obesity would mediate perceived stress and the onset of hypertension as well as influence blood pressure itself.

The graphical hypothesis about the relationships by hypotheses is illustrated in Figure 2.2.

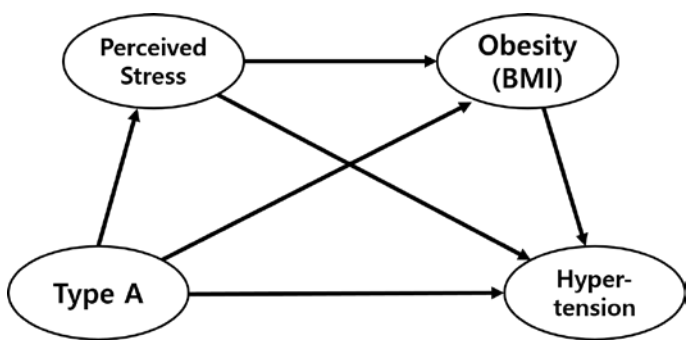


Figure 2.2 Concept model for the relationships among Type A, perceived stress, obesity and the onset of hypertension. Arrows mean the direction of influence

2.3.2 Confirmatory Factor Analysis

In the measurement modeling, Confirmatory factor analysis (CFA) was performed to determine whether the model fits to the data and to select proper observed variables representing latent variables of perceived stress (through PWI score) and Type A, which were measured by Likert-scale surveys.

The validity of the measured variables for the latent variables can be obtained by reliability and construct validity consists of convergent validity and discriminant validity.

Reliability is an index used to evaluate the internal agreement of observed variables to explain latent variables. The coefficient of this internal agreement is called Cronbach's Alpha, and it is that the closer to 1, the higher the reliability. It is judged as good if it is 0.7 or more (Bland & Altman, 1997).

Construct validity consists of convergent validity and discriminant validity.

Convergent validity is a method of judging validity based on the criterion that the correlation between each measurement should be high when the same concept is measured with different items. Convergent validity of the measurements was evaluated with composite reliability (CR), Average Variance Extracted (AVE). The criteria to CR is 0.6 or more and that of AVE is 0.5 or more (Fornell & Larcker, 1981).

Discriminant validity judges the validity based on the criterion that the correlation between each measurement should be low when measurements are performed using the same method for different concepts. That is, unrelated measures should not be related in reality to the data. Discriminant validity is accepted when correlation coefficients are lower than the square root of AVE.

The CFA was performed with the 10 items constituting Type A, the items impeding the explanatory power of Type A, that is, items with a factor loading value less than 0.5 were removed.

The CFA for PWI-SF was also implemented by removing the items that hindered the explanatory power of Perceived stress from the 18 items constituting it.

To evaluate the validity of the CFA result, several fit indices were used including the Tucker–Lewis Index (TLI), Comparative Fit Index (CFI), Goodness of Fit Index (GFI), and the Root Mean Square Error of Approximation (RMSEA) value. The known acceptable cutoff value of each Goodness-of-Fit was CFI $\geq$ 0.9, GFI $\geq$ 0.9, TLI $\geq$ 0.9, and RMSEA be below 0.06 to 0.08 indicates good fit (Hu & Bentler, 1999; MacCallum, Browne, & Sugawara, 1996; Schermelleh-Engel, Moosbrugger, & Müller, 2003).

2.3.3. Structural Equation Modeling

In the structural modeling, the serial multiple mediation hypothesis (Preacher & Hayes, 2008) for Type A, perceived stress, BMI and BP (SBP and DBP) was tested by path analysis.

In path analysis, the direct effect refers to the amount of change in Y relative to the amount of change in X while controlling all covariates so that no parameters exist between the variables X and Y (Pearl, 2009). In this case, if X and Y have a linear relationship, the path coefficient is the same as the regression coefficient.

In a model without interactive effects, the indirect effect of X on Y is multiplied by the direct effects constituting the indirect path. In this case, if there are multiple indirect paths, the total indirect effect is calculated as the sum of the individual indirect effects.

The total effect of X on Y is the sum of the indirect and direct effects of a change in Y when X changes by one unit.

Thus, in the current study, direct effect is the influence of Type A on BP, and indirect effect means the multiplication of the coefficient

from Type A to perceived stress, perceived stress to BMI, and from BMI to BP. The total effect is the sum of direct and indirect effects. The total effects including indirect effects and direct effects were computed with 5000 bootstrapping and maximum likelihood estimation was used for parameter estimation.

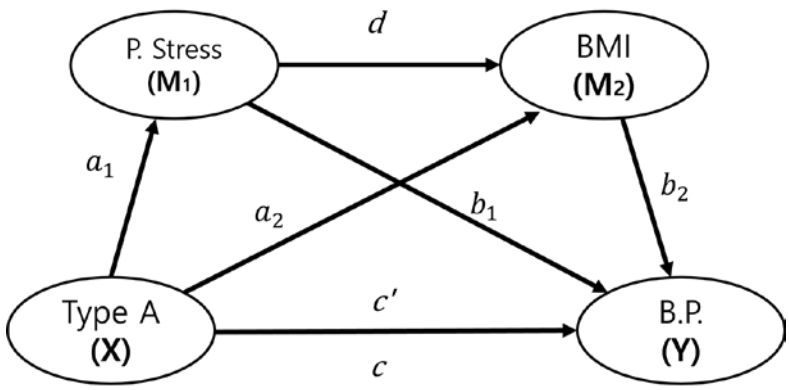


Figure 2.3 Direct effect, indirect effect, and total effect in the model

Figure 2.3 shows the direct effect and indirect effects X on Y via M1 and M2 in the model. a1 is the direct effect of the independent variable X (Type A) on M1 (Perceived Stress) and a2 is the direct effect of X on M2 (BMI). b1 is the direct effect of M1 on Y (Blood Pressure), and b2 is the direct effect of M2 on Y. d is the direct effect M1 on M2. c' means the direct effect of X on Y and c is the total effect of X on Y via Ms.

That is, the direct effect of X on Y is c.

The indirect effect of X on Y is the multiplication of a1, b1, and d.

The total effect of X on Y, c' is c + (a1 * b1 * d).

The specific model presented in lavaan with regression equations for the calculation of the above effect is as follows.

M1 ~ a1 * X
M2 ~ a2 * X + d * M1
Y ~ c' * X + b1 * M1 + b2 * M2
indirect := a1 * b1 * d

As with the CFA results, the goodness of fit of the SEM model is also evaluated by the chi-square test, TLI, CFI, GFI, RMSEA.

2.3.4. Post-hoc Test

When a post-hoc comparison is required after the analysis result, the average comparison method according to the data distribution was used for group comparison.

In the present study, to interpret the relationship between stress and BMI, the Wilcoxon rank-sum test was used for food intake according to the stress group (high and low), since they did not follow the normal distribution.

2.4. Genetic risk factors through weighted genetic risk score from GWAS

2.4.1. Introduction of GWAS

GWAS (Genome-wide association study) was first tried to identify candidate chromosome locus associated with myocardial fraction in Ozaki group in The Institute of Physical and Chemical Research (RIKEN) in Japan (Ozaki et al., 2002).

Based on the premise that the diversity of traits is the cause of gene polymorphism, GWAS analyzes genetic polymorphism present in DNA to determine the presence or absence of disease.

Genetic polymorphisms, which underlie trait diversity, are present in about 1% of about 3 billion base pairs (sequences of bases A, T, G, and C) in the 30,000 genes in humans, and among the various

indicators to express the polymorphism, GWAS mainly uses single nucleotide polymorphism (SNP).

As seen in Figure 2.4, the place where this SNP exists is called an SNP locus, and in most cases, one SNP locus is made up of a combination of two bases. Since humans inherit one base from each parent, the combination of these two bases can express genetic factors, which is called a genotype. For example, in a SNP composed of bases A and T, each individual has one of AA, AT, and TT genotypes. Also, a target that is passed on to the next generation in a stable form, such as a base, is more broadly called an allele. Bases A and T in the previous example correspond to each allele, and the genotype can be said to be a combination of two alleles. Therefore, genetic locus is also a term for the place where the allele is present and depending on the loci, several SNPs are contained in one gene. SNPs are usually expressed by numbering them starting with rs.

GWAS analysis can identify variants by comparing the reference human SNP with that of an individual, and whether this variant is related to a specific trait or disease can be known through the accumulated database.

Alternatively, as shown in Figure 2.5, SNPs associated with a specific disease (hypertension in the current study) can be discovered by comparing SNPs of the hypertensive case group and the control group.

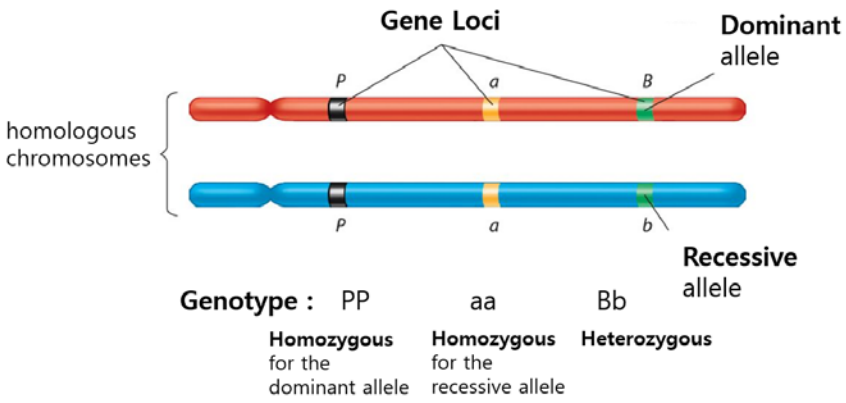


Figure 2.4 Allele, Gene locus, and Genotype on a chromosome

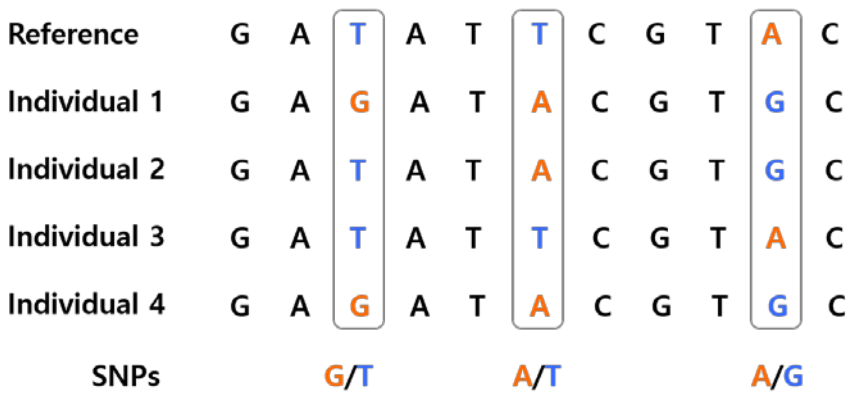


Figure 2.5 SNPs for individuals and comparison with the reference allele

This means that GWAS is a data access analysis method that filters out genes related to a specific disease by targeting all genes, rather than targeting specific gene candidates. By this reason, the P-value of the SNP for a single disease can become very small. Therefore, a strict filtering criterion is required, and it is usually judged based on from 5.0 × e-8 to 5.0 × e-5. SNPs that have passed these criteria or genes represented by those SNPs can be

identified with a Manhattan plot as shown in Figure 2.6. A Manhattan plot is a type of scatter plot and is usually used to represent large data points with non-zero values. It is usually used to identify and display statistically significant SNPs in GWAS (Gibson, 2010).

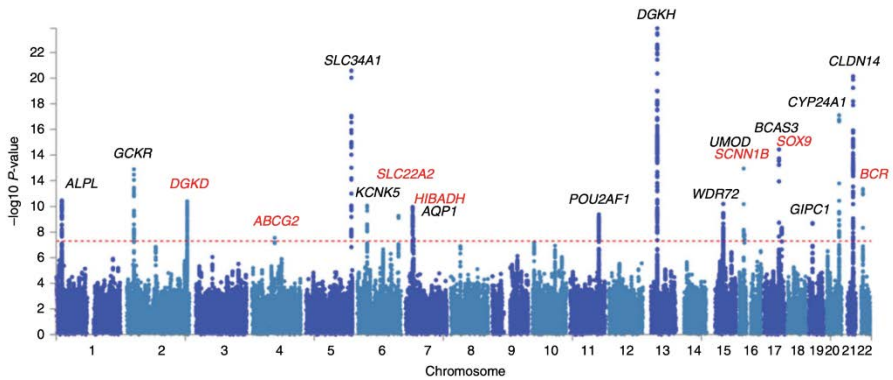


Figure 2.6 An example of a Manhattan plot for associated genes with the kidney stone disease and the loci from GWAS results

Basically, the main purpose of statistical analysis in GWAS is estimation and testing. In GWAS, the odds ratio is estimated for the difference in genotype frequency of SNP loci between hypertensive people and normal ones. If there is a difference in the environmental factors between the comparison group, a regression model can be adjusted using environmental factors such as age and gender in addition to the SNP loci.

2.4.2. Procedure of GWAS

In this study, the de-identified KoGES genome data of 8,842 individuals were received and utilized from the Korea Centers for Disease Control and Prevention.

The data were in the format of .map and .ped files of the PLINK

program, and were quality-controlled with a genotype call rate (GCR) > 96% and heterozygosity ≤ 30%. Also, for the marker, the quality was controlled as GCR > 95%, MAF (Minor Allele Frequency) > 0.01, and HWE (Hardy-Weinberg Equilibrium) p-value > 1.0×10^{-6}.

These data were read by Affymetrix Genome-Wide Human SNP Array 5.0 and 6.0. Since Affymetrix's genotyping data are presented as unique SNP IDs (SNP_A-*), they must be converted into rs-number format (rs99999) using the annotation file provided by the company. In addition, the rs number, chromosome number, and physical position must be corrected through the latest annotation file because the annotation file updates not only the rs number but also the physical location and chromosome information due to new research results.

Therefore, the overall pre-processing of genome analysis for this study is as follows.

1) Convert to binary format BED/BIM/FAM file instead of map or ped file with large memory capacity.
2) After quality control (QC) for this study,
3) Using the Affymetrix annotation file,
4) Change to RS number from SNP_A-*,
5) update the chromosome number,
6) also updating the physical position.
7) Align the DNA strands forward.

The criteria of quality control (for removal) are HWE ≤ $1x10^{-6}$, MAF < 0.01, SNP GCR ≥ 95%, and individual GCR ≥ 95%.

Based on these data, association analysis was performed by logistic regression with the dependent variable as the onset of hypertension and the independent variable as SNPs under the additive genetic model. The additive genetic model assumes that the contribution of all the three genotypes such as “AA”, “Aa”, and “aa”

to the phenotype (or the case) is different from each other (Zhao, Song, Wang, & Wang, 2016).

The significant SNPs were selected by applying Bonferroni correction (p-value = $5.0e^{-5}$).

2.4.3. Weighted Genetic Risk Score

After the GWAS has screened for SNPs associated with a particular disease (hypertension, in this study), the genetic risk score can be calculated from the sum of the risk-causing alleles.

In addition, the odds ratio can be calculated through the association analysis (basically logistic regression analysis with additive genetic effect) between the selected SNP and the disease, and this can be used as a weighting factor for the genetic risk score.

Therefore, the weighted genetic risk can be calculated as the sum of the values multiplied by the odds ratios of the related SNPs.

Expressing this as a formula is as follows:

$$\mathrm{w}GRS = \sum_{i=1}^{k} \beta_i N_i$$

where k is for the number of known independent genetic variants with strong association, β_i for is log odds ratio, and N_i for is the number of risk alleles (Cooke Bailey & Igo Jr, 2016). When dividing into the risk group and the non-risk group by categorization, a score of 0.65 was used as the standard.

In the current study, GWAS, association analysis and the calculation of wGRS were conducted by PLINK program (Purcell et al., 2007).

2.5. Interaction of Perceived Stress and Type A on wGRS

2.5.1. Survival analysis

In the present study, survival analysis was used in order to determine the extent to which independent variables such as perceived stress influence the onset of hypertension. Survival analysis is a statistical technique for estimating the survival time from the start of observation to death. In particular, the method of accumulating and estimating the survival probability for each event is called the Kaplan-Meier estimation method (Efron, 1988).

Instead of death, the target event was the onset of hypertension and survival probability of the present study means the time an individual hasn't had the disease yet. Since data were collected every two years from 2001 to 2016, the time in this study is the collection period.

Analysis results can usually be intuitively expressed through survival curves. However, because the Kaplan-Meier estimation cannot control more than one independent variable, it is necessary to use the Cox proportional hazards model to analyze the influence of multiple variables.

2.5.1.1. Kaplan-Meier Estimation

The Kaplan Meier estimate calculates the interval survival rate at each event point during the entire study period, resulting in a final cumulative survival rate. For this reason, it is also called the Product Limit Method (Goel, Khanna, & Kishore, 2010).

In this estimation, the survival rate is computed for each event time point. After sorting the data in the order of the observation period of the event, the interval survival rate P(t) is calculated as the

ratio of the number of survivors among the number of observations for each interval. If one person dies during the observation period, the interval survival rate is $\frac{n-1}{n}$. The cumulative survival rate S(t) can be calculated by sequentially multiplying the survival rates for each section, and the exact formula is as follows, the onset of hypertension instead of Death in the current study.

$$S_{t+1} = S * \frac{(N_{t+1} - D_{t+1})}{N_{t+1}}$$

S: *Survival Probability*
N: *Number at Risk*
D: *Number at Death*
t : *time*

2.5.1.2 Cox Proportional Hazards Model Analysis

In this study, when looking at the difference in the onset of hypertension according to the perceived stress, other variables may affect the dependent variable.

Particularly, it is necessary to control for variables (potential confounders) that directly or indirectly affect the onset, such as lifestyle habits including drinking as well as variables that generally influence such as gender and age.

Since survival analysis deals with hazard ratios similar to odds ratios, cox regression can be used using the algorithm of logistic regression analysis.

The definition of hazard rate *h(t)* and survival rate *S(t)* in the Cox proportional hazards model are as follows.

$$h(t) = \frac{f(t)}{S(t)} = \lim_{\Delta t \to 0} \frac{\Pr(t < T < t + \Delta t | T > t)}{\Delta t}$$

$$S(t) = 1 - F(t) \;\; where \;\; F(t) = \int_0^t f(u)du = Pr(T < t)$$

In this formula, T is for the time of the event occurrence, t is for mediated time.

The survival rate *S(t)* means the cumulative probability that an event does not occur until period t by avoiding all the probabilities of occurrence of an event in each session. In other words, it means that hypertension did not develop until stage t in this study.

The hazard rate *h(t)* is the probability that the event will occur at the time t, given that the event does not occur until the time t.

Based on these, the hazard function of the Cox proportional hazards model is as follows:

$$h(t|x) = \;\; h_0(t)e^{x_1\beta_1 + x_2\beta_2 + \cdots + x_i\beta_i}$$

where *h(0)* is the time-varying baseline hazard function, x_i are the variables and β_i are the coefficients to be estimated.

The formula indicates that hazard ratio changes proportionally according to the independent variables X_i affecting the onset of the event. In case that there are multiple variables, it is possible to know the actual effect of the variable or the interaction by stratifying the covariates and estimating the common beta value of the variable to be observed in Cox PH model.

2.4.2.1. Variates and Covariates

The variate for event used in the Kaplan-Meier survival analysis and the Cox proportional hazards model was the onset of hypertension, and the variate for time was the data collection time. The independent variables for this were weighted GRS, perceived stress, Type A, respectively, and their interactions were considered, especially for wGRS.

Known risk factors and collected variables on hypertension were included in simple cox regression. The valid covariates were AGE, Education level, Monthly income, Drink, Sodium intake, and total serum cholesterol and BMI. Since the analysis was performed separately depending on Gender, it was not included as a covariate.

As seen in Figure 2.7, smoking and exercising are not valid as covariates, which are probably biased. The result indicates people who had smoked have a lower hazard ratio than Never smoked people. Also, the hazard ratio of exercising people was higher than not exercising people.

That's because people tended to quit smoking after being diagnosed with some diseases and started exercising after being diagnosed with obesity. Thus, Smoke and Exercise were excluded as they could lead to misinterpretation of the overall results.

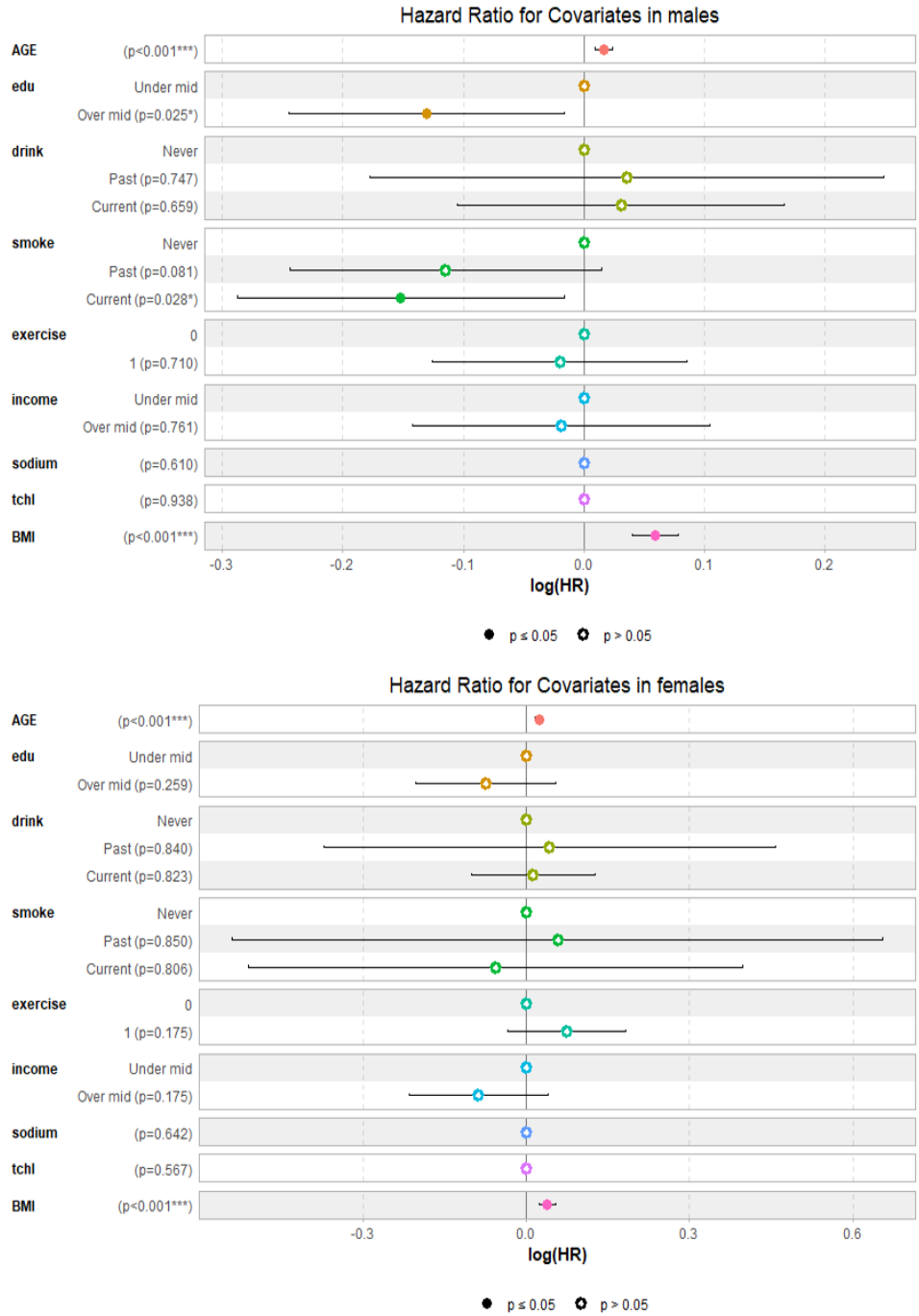


Figure 2.7 The result of Cox regression analysis with covariates

2.6. The evaluation of the contribution of Type A and Perceived Stress on Hypertension prediction

2.6.1. Tree Models

The purpose of this study is to investigate whether perceived stress, which is a psychosocial factor, and Type A personality type, which can affect the extent of perceived stress, can be predictors contributing to the direct prediction of hypertension. To determine this, a decision tree model was used to identify whether there was a change in the performance of the prediction results.

2.6.1.1. Decision Tree

The decision tree method is a widely used powerful statistical tool for classification, prediction, interpretation, and data manipulation (Y.-Y. Song & Ying, 2015). Also, it is a very simple model that predicts the outcome by making a split based on the predictor that can best reduce the RSS (sum of squared error). In particular, it has the advantage of good explanatory power because it is possible to see at a glance which variables are important and what the prediction results are according to the values of the variables.

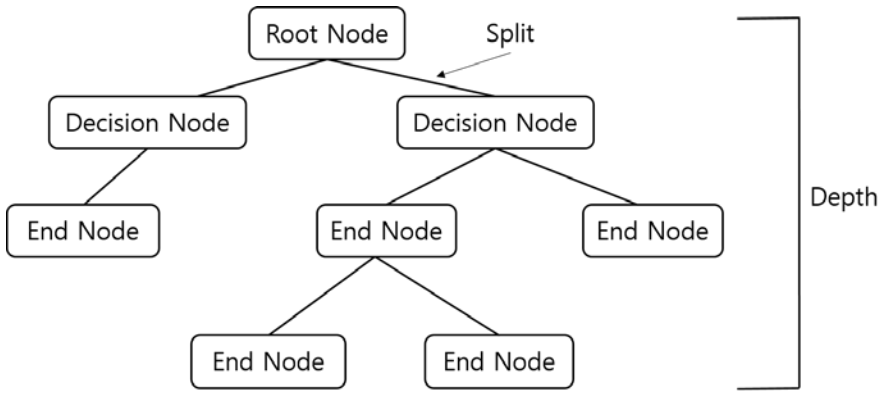


Figure 2.8 The structure of the decision tree model

The structure of the decision tree model is as seen in Figure 2.8. From the start of the Root Node, data can be classified by the criterion of the Decision Node. For example, if the input data has an SBP value over 140, it is classified as hypertension in the Terminal Node. It is expressed that the closer the classified variable is to the target variable, the higher the purity. In other words, child nodes split by specific criteria from the parent node have higher purity.

Basically, these characteristics or criteria related to the degree of the purity can be entropy, Gini index, classification error, information gain, gain ratio, and twoing criteria. This splitting procedure continues until the end nodes are homogeneous or ending criteria are met. Along with this way, the tree model learns to increase the accuracy by modifying parameters such as the number of nodes and that of depth.

2.6.1.2. Random Forest

Random forest (Breiman, 2001) is an ensemble model as a method of bagging a decision tree. It was more generalized to mitigate the shortcomings of the decision tree which is easily overfitted to the data.

Among Ensemble models combining multiple models, Random forest is a technique that creates multiple decision trees, passes data through each tree at the same time, and selects the final classification model with the most results.

Random forest uses bagging to create multiple trees, which is a method to build a tree by using a subset of the data training set. For example, if there are 1000 pieces of data in the training set, only 100 of them are randomly selected and utilized when creating a tree. Therefore, all trees are built using different data, but these data are all subsets of the entire training set.

Not only data but features are also selected using a subset of existing features when building a tree. Figure 2.9 shows the architecture of Random Forest.

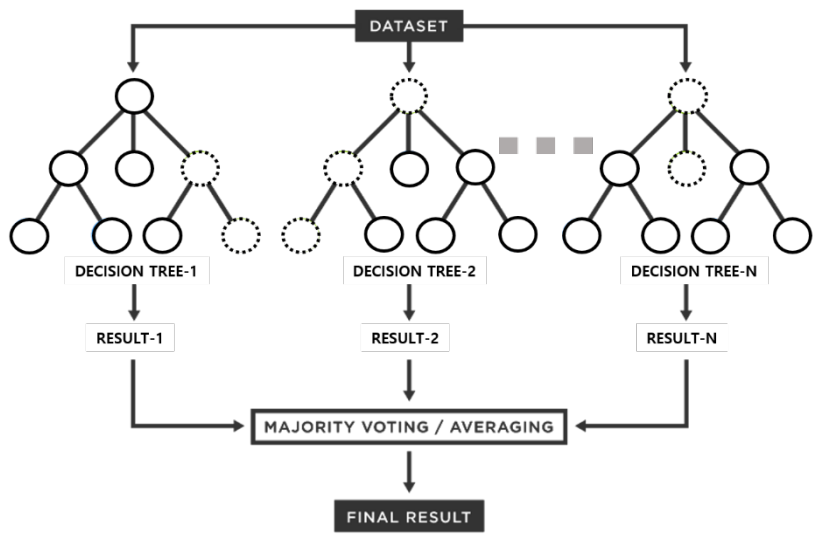


Figure 2.9 The structure of Random Forest Model by Bagging to decision trees

2.6.1.3. XGBoost

XGBoost or eXtreme Gradient Boosting (Chen & Guestrin, 2016) is another ensemble model using boosting. Boosting is a strategy to lower bias by training subsequent trees using what errors the previous tree made and reiterating this process to following trees.

The way of Gradient Boosting uses residuals to train the following tree. Figure 2.10 illustrates the architecture of XGBoost.

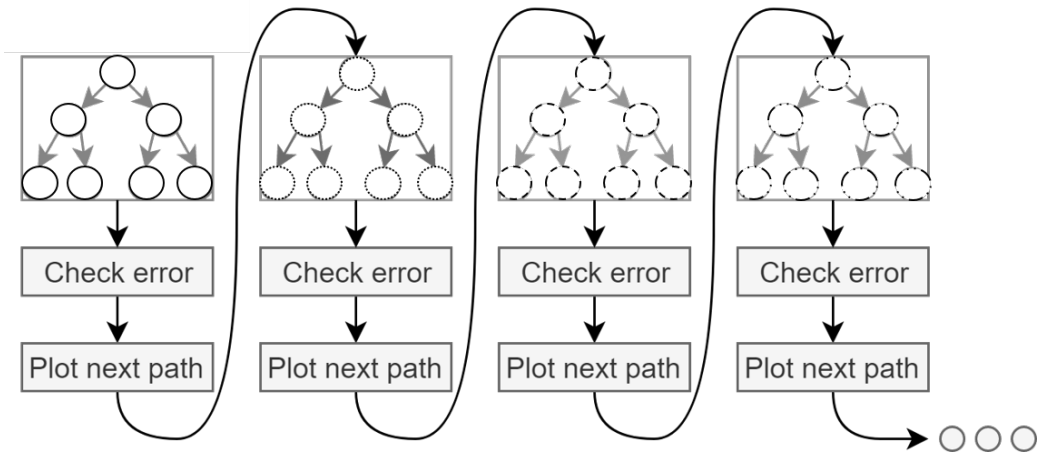


Figure 2.10 The structure of XGBoost

2.6.2. Criteria for Performance Improvement

2.6.2.1. Reclassification (NRI and IDI)

Although there are various methods for evaluating model performance, in the current study, the reclassification method was used to determine whether the contribution of risk factors affects the performance improvement.

Reclassification measures how much better the new model is than the old model, and Net Reclassification Improvement (NRI) and Integrated Discrimination Improvement (IDI) are often used.

NRI, for example, when a new model (test 2) is created by adding important risk factors to one model (test 1), the degree of improvement of the model is evaluated by determining how well reclassification was performed (Pencina, D'Agostino Sr, D'Agostino

Jr, & Vasan, 2008).

$$\mathrm{NRI} = \mathrm{P(up|event)} - \mathrm{P(down|event)} + P(down|non-event) - P(up|non-event)$$

In this formula, *event* means the occurrence of a disease, and *non−event* means normal. Also, *up* is the number of improved classification items, and *down* is the number of degraded classification items. The sum of these two becomes the NRI value, and dividing this value by the number of samples gives you an idea of how much you've improved. If the NRI value of the new model with the added risk factor increases compared to the old model, this risk factor can be meaningful as a predictive factor.

IDI (Pencina et al., 2008) is defined as follow:

$$\mathrm{IDI} = (\mathrm{IS_{new}} - \mathrm{IS_{old}}) - (\mathrm{IP_{new}} - \mathrm{IP_{old}})$$

IS means the integral of sensitivity over all possible cut−off values from the (0, 1) interval and IP means the corresponding integral of one minus specificity.

$$\widehat{IDI} = \left(\overline{\hat{P}}_{new,events} - \overline{\hat{P}}_{old,events}\right) - (\overline{\hat{P}}_{new,non-events} - \overline{\hat{P}}_{new,non-events})$$

This formula indicates the result of subtracting the calculated value of the old model from the calculated value of the new model, and it is the predicted average of the probability that an event will occur.

That is, the idea of IDI is that if it is a useful factor, it will increase the risk of the case and lower the risk of the control.

If the risk factor added to the new model contributes to the prediction of risk, the first term has a large positive value and the second term has a large negative value, so the difference between the two, the IDI value, will increase.

The significance of the value of IDI can be checked depending on whether or not the null hypothesis of IDI=0 is rejected, and the test formula is as follows.

$$\text{zIDI} = \frac{\widehat{IDI}}{\sqrt{(SE_{events}{}^2) + \widehat{(SE_{non-events}{}^2)}}}$$

In the current study, the validity as risk factors of perceived stress and Type A was tested by comparing NRI and IDI of the Decision Tree models with and without them.

2.6.2.2. Feature Importance by XAI (SHAP)

Van Lent, Fisher and Mancuso pointed out that computer systems and artificial intelligence systems had become increasingly complex, but the intermediate process between input and output was still unknown. They modified the NPC (Non-Player Character) artificial intelligence in the military simulation combat program to suggest an architecture that explains the NPC's behavioral reasons, and used the term XAI (Explainable Artificial Intelligence)(Van Lent, Fisher, & Mancuso, 2004).

SHapley Additive exPlanations or SHAP (Lundberg & Lee, 2017) is one of the popular XAI techniques, based on Lloyd Shapley's idea for a solution in cooperative game theory (Shapley, 1951). Shapley value was named in honor of him (Roth, 1988).

Cooperative gaming indicates that positive cooperation is the best option if the individual gains from playing cooperatively are greater than the individual gains from playing non-cooperatively (Owen, 2013). Based on cooperative game theory, the Shapley value is obtained through the average expected marginal contribution according to the presence or absence of the interesting feature after all possible combinations of several features have been considered

(Winter, 2002).

Expressing this as a formula, it is as follows:

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!}(v(S \cup \{i\}) - v(S))$$

$\emptyset_i$: shapely value for i
n : number of features
S : set of total features – i feature
v(S) : contribution of S
v(S ∪ {i}) :contribution of S + i

In the above formula, the contribution of the I^{th} feature is subtracted from the total contribution minus the sum of contributions excluding the I^{th} feature.

Entering features into the model returns prediction labels and accuracy. For example, if there is a model with 4 features (x_1, x_2, x_3, x_4), the value S contributed by one feature to the entire model is calculated as follows.

$$\emptyset_2(v) = \frac{1}{4!}\sum_R \left[v\left(P_2{}^R \cup \{2\}\right) - v\left(P_2{}^R\right)\right]$$
$$= \frac{1}{4!}\sum_R [v(\{x_1, x_3, x_4\}^R \cup \{x_2\}) - v(\{x_1, x_3, x_4\}^R)]$$

That is, the average is obtained by subtracting all combinations that the model can represent and combinations of features except for feature x_2.

As mentioned above, SHAP is an algorithm for calculating these Shapley values. Tree SHAP decomposes the output of the model into the contribution of each feature, samples over all possible ordered pairs of features, and computes an average value. In the case that the

Shapley value may be a negative number, it may be interpreted that the corresponding feature has a negative effect on prediction.

Therefore, the advantage of the contribution estimated by SHAP is that it not only calculates the influence considering the dependence between features, but also calculates the features that negatively affect the prediction, unlike other methods of calculating the importance of features.

In the present study, Tree SHAP was used as proposed by Lundberg (Lundberg, Erion, & Lee, 2018) for tree-based machine learning models such as decision trees, random forests and gradient boosted trees. Tree SHAP calculates the Shapley value quickly and accurately even when the features are dependent on each other, as well as accurately estimates the Shapley value.

Chapter 3. Results

3.1. Baseline characteristics of the data

Table 3.1 shows variable-specific characteristics of groups by gender and the difference. When the prevalence was defined as the number of people by age group and the proportion of those who had met the definition of hypertension during the follow-up period, the prevalence was high in males in their 40s and 50s, and high in females in their 60s. This is similar to the trend shown in the National Health and Nutrition Survey, in which the prevalence rates of males in their 40s, 50s, and 60s increase to about 30%, 45%, and 55%, respectively, and for females, it increases in a pattern of about 13%, 30%, and 55%, respectively (H. C. Kim et al., 2022).

BMI was not significantly different between the male and the female. However, Type A score (p<.001), sodium intake (p<.001), SBP (p<.001), and DBP (p<.001) were higher in the male group, and perceived stress (p<.001) and total cholesterol level (p<.001) were higher in the female group. According to this result, the analysis was performed by gender.

Table 3.1 General characteristics of the variables

Characteristics	**Male** M (SD)	**Female** M (SD)	**W** (p-value)	**unit**
Numbers	1668	1911		N
Age	50.85 (8.01)	51.75 (8.45)	-2.73 (.01)*	
40-49 (HTN)	885 (0.89)	919 (0.7))	2.33 (.02)*	N
50-59 (HTN)	474 (0.90)	549 (0.87)	-2.52 (.01)*	N
60-69 (HTN)	309 (0.92)	443 (0.96)	0.47 (.64)	N
P.stress (PWI)	15.15 (7.77)	17.77 (8.45)	-9.24 (<.001)***	

BMI	24.47 (2.78)	24.89 (3.16)	-2.59(.01)*	kg/m^2
Type A	9.99 (5.56)	9.31 (5.59)	3.61 (<.001)***	
SBP	116.67 (15.43)	114.83 (16.62)	3.96 (< .001)***	mmHg
DBP	79.11 (9.80)	76.03 (10.13)	8.62 (< .001)***	mmHg
Sodium/day	2859.17 (1499.96)	2587.18 (1600.91)	7.44 (< .001)***	mg
Cholesterol	187.99 (32.82)	195.88 (34.16)	-6.84 (< .001)***	mg/dL
Amount of Food intake/day	1652.46 (541.82)	1553.80 (582.35)	6.69 (< .001)***	g
Exercise				
Yes	736	678		n
No	932	1233		n
Drink				
Never	310	1376		n
Past	138	28		n
Current	1220	507		
Smoke				
Never	415	1876		n
Past	691	12		n
Current	562	23		n
Education				
Under middle	626	1258		n
Over middle	1042	653		n
Income				
Under middle	730	1174		n
Over middle	938	737		n

P. stress: Perceived stress, BMI: Body Mass Index, Sodium/day: daily sodium intake, Cholesterol: total cholesterol, Education middle: high school, Income: monthly income; middle $3000 *p<.05 **p<.01 ***p<.001

3.2. GWAS and wGRS for hypertension

Figure 3.1 shows significant SNPs for the hypertensive phenotype shown on the Manhattan plot. Among 189,813 SNPs, 26 SNPs with significance e−5 or less in the association analysis with hypertension were summarized in Table 3.2.

In the table, OR or Odds Ratio was used as a weight in the genetic risk score, and the genetic risk score was calculated according to the number of risk (usually minor) alleles.

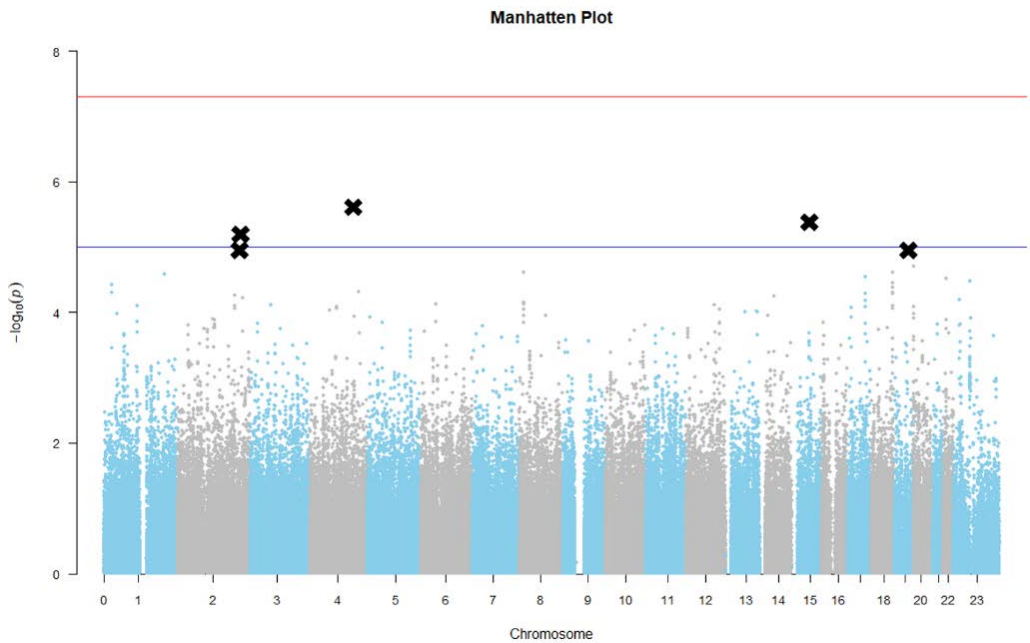


Figure 3.1 Manhattan plot after GWAS for hypertension

Table 3.2 Significant SNPs for hypertension

Gene	Chr.	SNP	Function	D/R	OR	P
ADC	1p35.1	rs16835244	Ala288Ser	G/T	1.07	7.5×10-5
SLC25A5P2	2p22.3	rs280749	intergen	G/T	1.04	4.3×10-5
PLCL1	2q33	rs921465	intergen	T/G	1.04	1.5×10-5
ALDOAP1	3p21.2	rs411457	intergen	T/A	1.03	9.4×10-5
FAM47E	4q21.1	rs6532316	intron	A/G	1.04	4.9×10-5
TTC1	5q33.3	rs6556466	intron	T/C	1.05	9.1×10-5
ATP10B	5q34	rs2010725	intergen	T/G	1.08	1.4×10-5
MANEA	6q16.1	rs494801	intergen	T/G	1.04	3.4×10-5
FUT9	6q16	rs9404150	intron	G/A	1.03	6.1×10-5
IMMP2L		rs2966417	intergen	G/A	1.04	4.2×10-5
MKLN1	7q32	rs13228664	intron	A/G	1.03	7.9×10-5
RAD23B	9q31.2	rs10739236	intron	T/C	1.04	3.9×10-5
COL27A1	9q32	rs4978577	intron	C/T	1.04	3.9×10-5
NPS	10q26.2	rs7912681	intergen	C/T	1.04	2.5×10-5
GLRX3	10q26	rs1176440	intergen	T/C	1.04	6.2×10-5
LOC646388	11p12	rs10768467	intergen	G/A	1.18	4.1×10-5
DLG2	11q14.1	rs1515089	intron	C/T	1.05	2.5×10-5
ATP2B1	12q21.3	rs1724975	intergen	A/G	1.04	7.5×10-7
CUX2	12q24.12	rs12229654	intergen	G/T	1.05	2.8×10-5
HECTD4	12q24.13	rs2074356	intron	A/G	1.06	1.1×10-6
MBIP	14q13.3	rs2254613	intergen	T/G	1.04	2.8×10-5
LOC283584	14q31.3	rs1362721	intergen	T/C	1.04	6.6×10-5
CSK	15q24.1	rs1378942	intron	A/C	1.05	3.5×10-6
CCL2	17q11.2	rs2097761	intergen	A/G	1.04	5.2×10-5
VAPA	18p11.22	rs12966494	intergen	A/G	1.04	5.9×10-5
RPL12P4	20q13.2	rs11905645	intergen	C/T	1.04	6.7×10-5

Figure 3.2 shows the distribution of weighted genetic risk scores for the total sample of 3579. To confirm the explanatory power of the weighted genetic risk score for the onset of hypertension, logistic

regression analysis and linear regression analysis on systolic and diastolic blood pressure were performed.

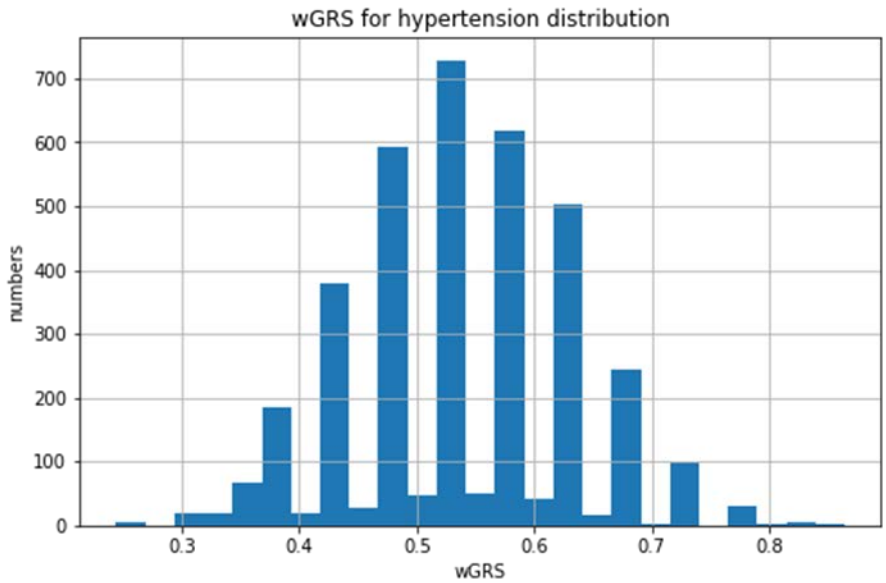


Figure 3.2 The distribution of wGRS for hypertension

As a result of logistic regression analysis, B = 1.69, EXP(B) = 5.42 (p = .004) when adjusted for age, sex, salt intake, and BMI terms. In other words, as the weighted genetic risk score increases by 0.1, the likelihood of developing hypertension increases by 5.42 times. The pseudo R2 for this analysis was 0.04, and the p value of the log likelihood ratio (LLR) test was 2.14e-19, which was significantly explainable than the LL-Null model.

In a linear regression analysis with systolic blood pressure (SBP) and diastolic blood pressure (DBP) as dependent variables, respectively, when adjusted for age, sex, salt intake, and BMI, B = 8.62 (p=.002), R2 = 0.144 for SBP, B = 4.27 (p=.016), R2 = 0.092 for DBP, indicating that weighted genetic risk score significantly affects the development of hypertension.

3.3. Causal structure inference of hypertension risk factors through structural equation model

3.3.1. The Results of CFA

The Measurement Modeling Results

The summary of the CFA (Confirmatory Factor Analysis) results is presented in Table 3.3.

Reliability

The internal reliability was measured by Cronbach's alpha, and as shown in Table 3.3, the reliabilities for Type A and Perceived stress were .7 or higher, which was proper for the both.

Goodness-of-Fit

Type A

The results of the current study for males were CFI = 0.908, TLI = 0.871, RMSEA = 0.069. Although TLI was slightly unreached to the standard, other indices were acceptable. The results for females were CFI = 0.938, TLI = 0.914, RMSEA = 0.054, all indices met the criteria.

Perceived Stress

The results for males were CFI = 0.914, TLI = 0.890, RMSEA = 0.091. Since RMSEA is less than 1, the index of this model can be regarded as just acceptable and TLI was close to the criterion. In addition, as CFI met the criteria, the indices were overall acceptable.

The results for females were CFI = 0.914, TLI = 0.889, RMSEA = 0.085. As with the result of males, all indices indicate the model is acceptable.

Validity

Convergent validity

As the summary results were shown in Table 3.3, C.R for males on Type A was 0.7 and that for females was also 0.71, which are suitable. C.R for males on Perceived Stress was 0.87 and that of females was 0.85, which are also adequate.

Discriminant validity (Correlation and AVE)

AVE for males on Type A is 0.38 and that of females is 0.33, meanwhile AVE for males on Perceived Stress is 0.47 and that of females is 0.35. Considering that the criterion of AVE is over 0.5, the values are inadequate. However, as Composite reliabilities (CRs) of Type A and Perceived Stress are higher than .60 for both genders, the convergent validity of Type A and Perceived Stress are appropriate (Fornell & Larcker, 1981).

Table 3.3 The reliability, model-fit, and the validity of the measurements

			Male		**Female**	
		criteria	Type A	Perceived Stress	Type A	Perceived Stress
Reliability	Cronbach's α	>.70	.67	.87	.69	.85
Goodness-of-Fit Indices	χ^2 (df)		39.877(5)	522.972(35)	15.126(5)	523.180(35)
	p	>.05	.000	.000	.000	.000
	TLI	>.90	.873	.890	.969	.889
	CFI	>.90	.937	.914	.984	.914
	RMSEA	<.10	.065	.091	.033	.085
Validity	C.R	>.70	.71	.87	.73	.85
	AVE	>.50	.38	.47	.39	.35

df : Degree of freedom, TLI : Turker-Lewis Index, CFI : Comparative Fit Index, RMSEA : Root Mean Error of Approximation, C.R : Composite Reliability, AVE : Average Variance Extracted

The valid items in each questionnaire were selected with a factor loading of 0.5 or more, and items less than 0.5 were removed.

In Table 3.4, the non-standardized path coefficient means the raw score of Y when X increases by 1 point while controlling for the covariates predicting Y. The standardized path coefficient indicates how many standard deviations Y increases when X increases by 1 standard deviation in the same state.

The standardized regression coefficients for each item also exceeded the minimum standard of 0.6, and the C.R value for each factor also exceeded the criterion of 1.96, resulting in a suitable result for concentrated validity. The selected items are attached to the appendix.

Table 3.4 The results of CFA and selected items

Variable	Path	Unstandardized estimates		S.E.		C.R. Ratio		Standardized estimates	
		Male	Female	Male	Female	Male	Female	Male	Female
Type A	→ No. 1	0.454	0.483	0.03	0.029	15.133	16.655	14.978	16.45
	→ No. 3	0.479	0.476	0.027	0.025	17.741	19.040	17.995	19.16
	→ No. 4	0.497	0.482	0.027	0.027	18.407	17.852	18.132	17.93
	→ No. 5	0.458	0.54	0.029	0.028	15.793	19.286	15.993	19.56
	→ No. 6	0.352	0.278	0.022	0.021	16.000	13.238	15.97	13.52
	→ No. 8	0.537	0.485	0.027	0.025	19.889	19.400	19.791	19.80
	→ No. 9	0.532	0.553	0.03	0.028	17.733	19.750	17.985	19.45
	→ No. 10	0.504	0.508	0.023	0.022	21.913	23.091	22.246	23.11
Perceived Stress	→ No. 1	0.506	0.538	0.021	0.022	24.095	24.455	23.912	24.02
	→ No. 5	0.552	0.567	0.022	0.022	25.091	25.773	25.255	25.40
	→ No. 6	0.555	0.491	0.022	0.023	25.227	21.348	25.122	21.71
	→ No. 8	0.574	0.623	0.021	0.021	27.333	29.667	27.381	29.13
	→ No. 9	0.642	0.687	0.02	0.019	32.100	36.158	32.471	35.57
	→ No. 10	0.62	0.666	0.018	0.019	34.444	35.053	33.659	35.26
	→ No. 11	0.533	0.536	0.02	0.022	26.650	24.364	26.166	24.55
	→ No. 12	0.562	0.546	0.018	0.02	31.222	27.300	30.863	27.98
	→ No. 15	0.257	0.31	0.016	0.019	16.063	16.316	15.886	16.23
	→ No. 17	0.556	0.616	0.019	0.02	29.263	30.800	29.697	30.57

*** : p<.001. C.R. ratio : Critical ratio(=estimates/SE), Fixed Index, SE : Standard error

3.3.2. The Results of SEM

Figure 3.3 shows the results of path analysis among hypertension, Type A, Perceived Stress, and BMI by gender. Also, specific coefficients and p values are summarized in Table 3.5, and the effects of variables on each other were reported in Table 3.6.

All models were adjusted for covariates (age, education, income, and drink, sodium intake) satisfying the criteria of the goodness-of-fit indices. The male model for HTN met GFI=0.935, CFI=0.918, TLI=0.906, RMSEA=0.048. The female model for HTN was acceptable with GFI=0.933, CFI=0.912, TLI=0.899, RMSEA=0.05.

Regardless of gender, only indirect effects were identified for Type A on HTN (β=−0.001, p<.03). First, for males, Type A had a positive effect on perceived stress (β=0.14, p<.001), and it indicates that stress had a negative effect on BMI (β=−0.48, p<.001), and BMI had a positive effect on HTN (β=0.02 p<.001).

Also for women, Type A had a positive effect on perceived stress (β=0.13, p<.001), stress had a negative effect on BMI (β=−0.51, p<.001), and BMI had a positive effect on HTN (β=0.02, p<.001).

In both genders, the direct influences of perceived stress and Type A were not significant on HTN.

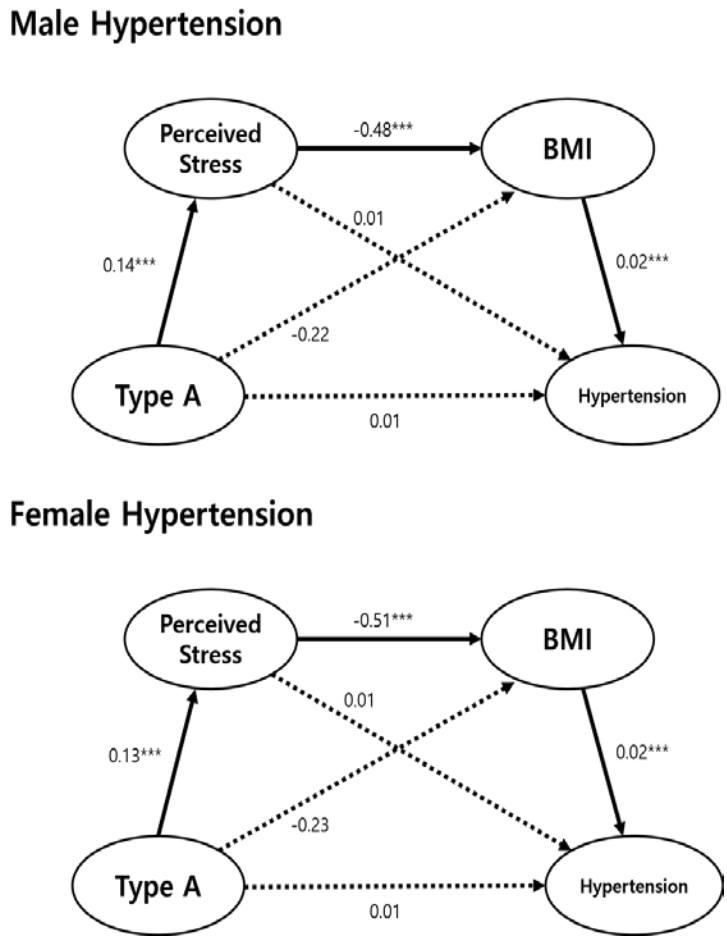


Figure 3.3 The results of SEM for hypertension by gender

Table 3.5 Path coefficients in SEM

Gender	DV	IV	β	Z	p
Male	**P. stress**	**Type A**	**0.140**	**3.555**	**.000*****
	BMI	Type A	-0.219	-1.287	.198
	BMI	**P. stress**	**-0.479**	**-3.337**	**.001*****
	H.T	Type A	0.007	0.401	.689
	H.T	P.stress	0.010	0.658	.511
	H.T	**BMI**	**0.020**	**8.296**	**.000*****
Female	**P. stress**	**Type A**	**0.132**	**3.486**	**.000*****
	BMI	Type A	-0.225	-1.299	.194
	BMI	**P.stress**	**-0.508**	**-3.439**	**.001*****
	H.T	Type A	0.008	0.446	.656
	H.T	P.stress	0.010	0.662	.508
	H.T	**BMI**	**0.020**	**8.495**	**.000*****

Table 3.6 The effects Type A for hypertension through mediators

Gender	Effect	β (95% CI)	p
male	direct	0.007 (-0.028 - 0.040)	.689
	indirect	**-0.001 (-0.003 - 0.000)**	**.027***
	total	0.006 (-0.029 - 0.038)	.741
female	direct	0.008 (-0.027 - 0.041)	.656
	indirect	**-0.001 (-0.003 - 0.000)**	**.025***
	total	0.007 (-0.028 - 0.040)	.706

3.3.3. The Results of Post-hoc Test

From the results of SEM, it was identified that perceived stress had a negative effect on BMI. In other words, the higher the stress, the lower the obesity rate. To confirm this phenomenon, an average comparison was made to see if there was a difference in actual food intake based on stress in the 3rd period data, which is the same cross-sectional data.

Since the Shapiro test result did not follow a normal distribution, Wilcoxon rank sum test was performed on the food intake of the high and low stressed groups, and the results presented significant differences as shown in Table 3.7 and Figure 3.4.

Food intake was significantly higher in the low-stressed group than that in the high-stressed group, which was the same in both genders (W=3.32, p<.001 for males, W=6.59, p<.001 for females).

It is ambiguous to say that food intake is a criterion for obesity. Nevertheless, given that high food intake is associated with a higher risk of obesity, the difference between the two groups suggests that the high-stressed group has a lower food intake and, therefore, is less likely to be obese.

Hence, the result could explain the negative relationship between Perceived Stress and BMI, a result of SEM.

Table 3.7 The comparison of food-intake between high stress group vs. low stress group by gender

Gender	Food intake/day (g)		W	P-value
	Stress High Mean(sd)	Stress Low Mean(sd)		
Male	1497.95 (441)	1666.18 (548)	3.32	< .001***
Female	1382.41 (537)	1584.34 (585)	6.59	< .001***

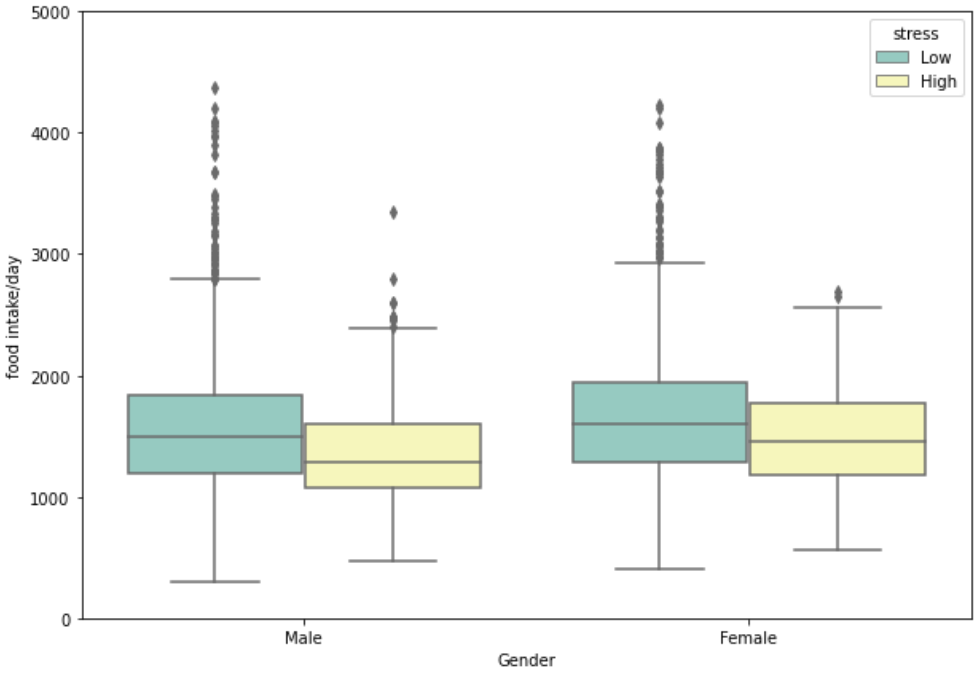


Figure 3.4 The comparison of food-intake between high stress vs. low stress groups by gender

3.4. Survival analysis

3.4.1. Kaplan-Meier Estimation

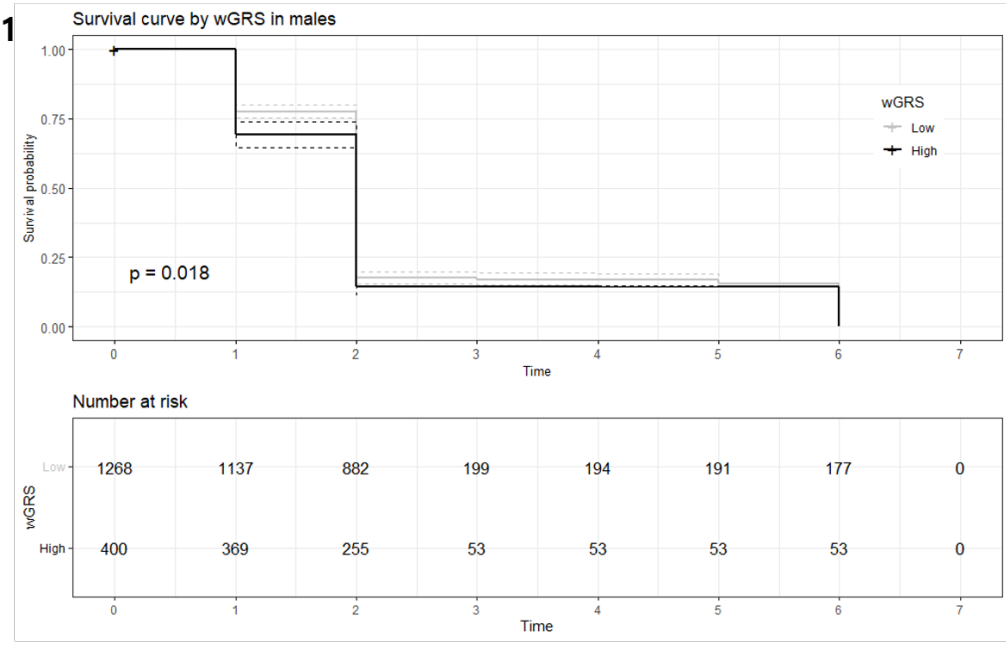
Data summaries in terms of raw Kaplan−Meier curve and relative risk estimations for each risk factor (wGRS, Perceived Stress, Type A) for hypertension are illustrated in Figure 3.5, and analyzed by gender. The comparison groups were divided into two, for each wGRS high and wGRS low (score over 0.65 out of 1 and the rest), Perceived Stress high and low (score over 27 out of 54 and the rest), Type A and Non−Type A (score over 14 out of 27 and the rest).

Figure 3.5 A shows the survival curve for the onset of hypertension by wGRS. In both genders, the group with higher wGRS showed a significantly higher incidence ($p<.02$ for males, $p<.001$ for females).

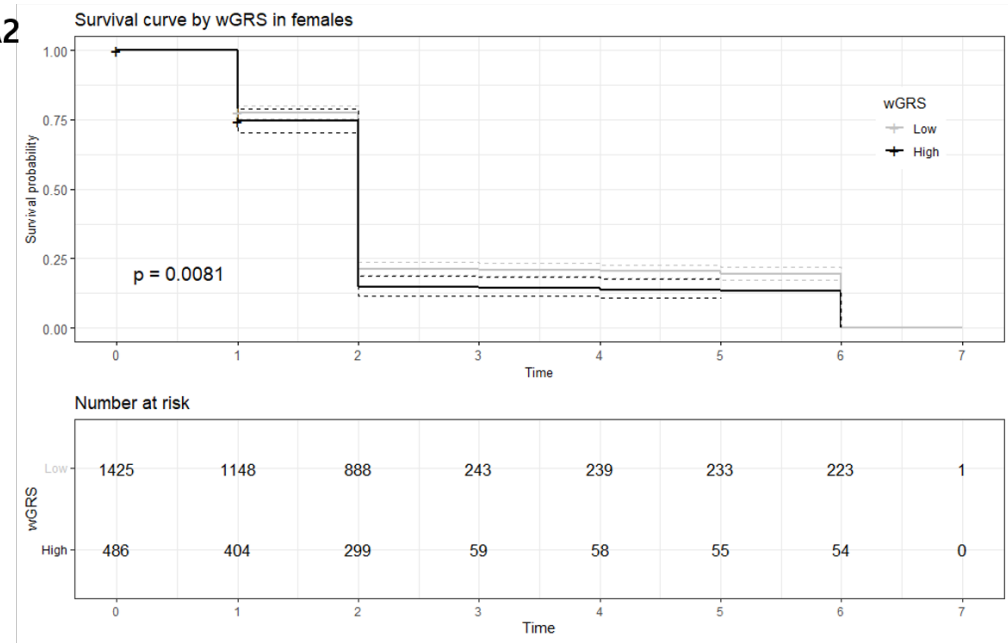
Figure 3.5 B is a survival curve according to perceived stress, which was not significant for males, and significantly increased the incidence rate for females ($p=.02$).

The survival curve in Figure 3.5 C illustrates that the rate of hypertensive risk of the Type A group was significantly higher in females ($p<.001$) and also higher in males as well, though marginally significant ($p=.06$).

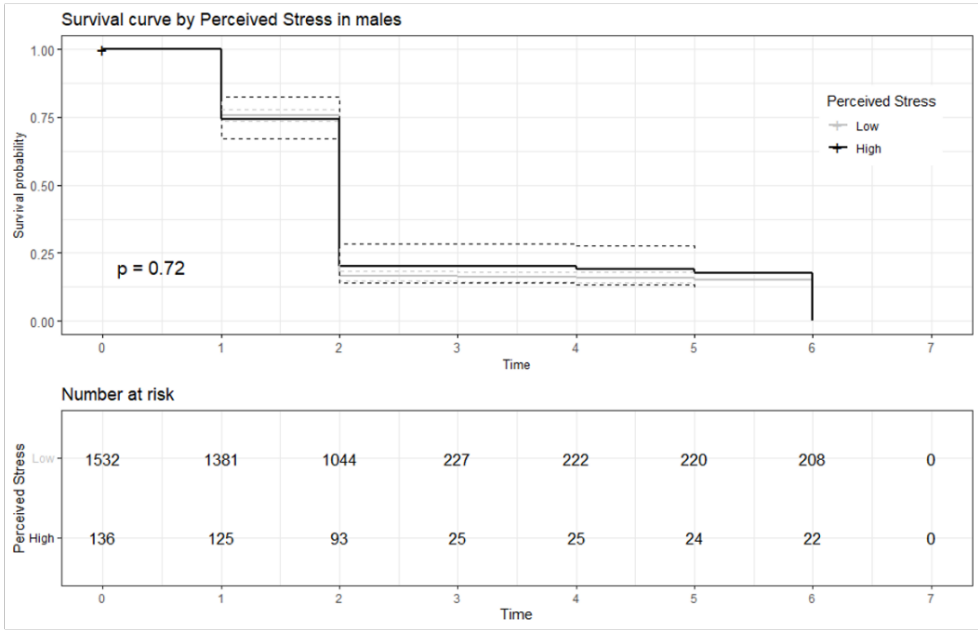
A1


Survival curve by wGRS in males
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.018
wGRS
Low
High
0
1
2
3
4
5
6
7
Time
Number at risk
wGRS
Low
High
1268
1137
882
199
194
191
177
0
400
369
255
53
53
53
53
0
Time

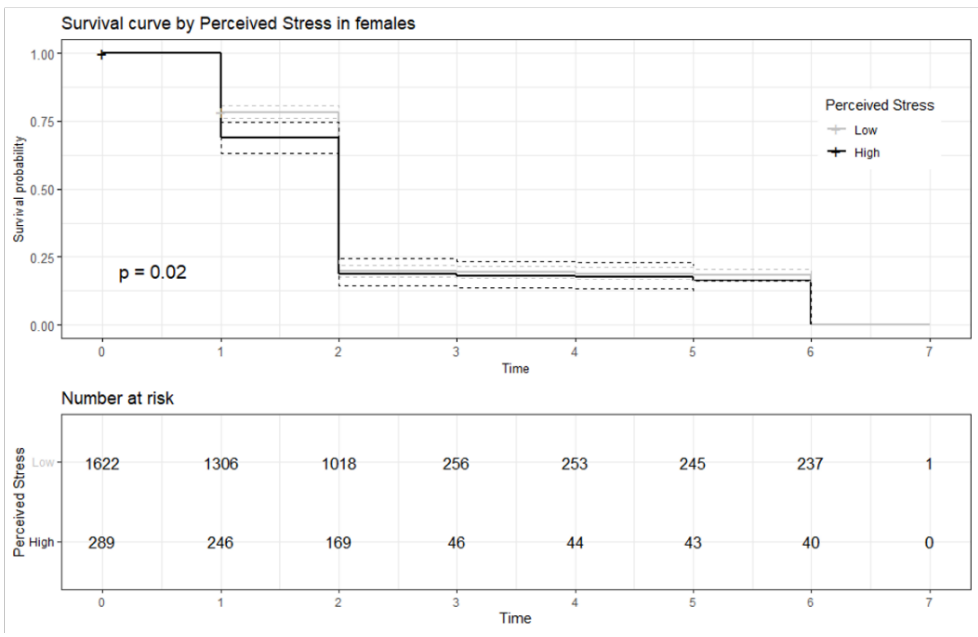
A2


Survival curve by wGRS in females
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.0081
wGRS
Low
High
0
1
2
3
4
5
6
7
Time
Number at risk
wGRS
Low
High
1425
1148
888
243
239
233
223
1
486
404
299
59
58
55
54
0
Time

B1


Survival curve by Perceived Stress in males
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.72
Perceived Stress
Low
High
0
1
2
3
4
5
6
7
Time
Number at risk
Perceived Stress
Low
High
1532
1381
1044
227
222
220
208
0
136
125
93
25
25
24
22
0
Time

B2


Survival curve by Perceived Stress in females
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.02
Perceived Stress
Low
High
0
1
2
3
4
5
6
7
Time
Number at risk
Perceived Stress
Low
High
1622
1306
1018
256
253
245
237
1
289
246
169
46
44
43
40
0
Time

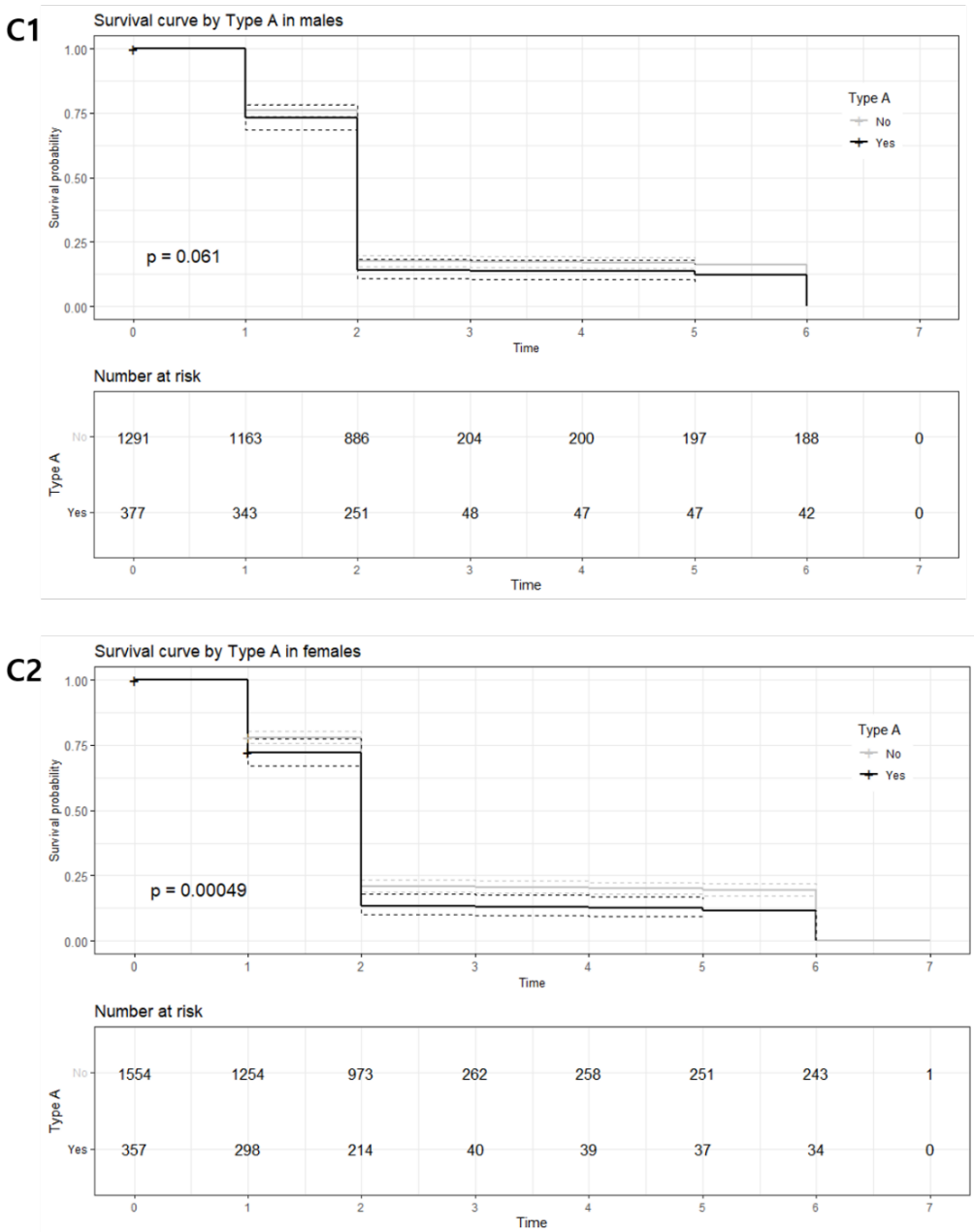


Figure 3.5 Survival curves by Kaplan-Meier estimation. A) for wGRS B) for Perceived Stress C) for Type A 1) for males 2) for females

3.4.2. Cox Proportional Hazard Ratio Model Analysis

The results of the Cox hazard ratio proportional model are summarized in Table 3.8 and Figure 3.6 The model was controlled for covariates, the data collection period used as the time variable, and the onset of hypertension used as the dependent variable.

For wGRS, the risk of the high-scoring group was 1.13 times higher (p=.04) for males and 1.11 (p=.04) times higher for females when the low-scoring group was referenced.

In the case of Perceived stress, there was no significant effect in both males and females.

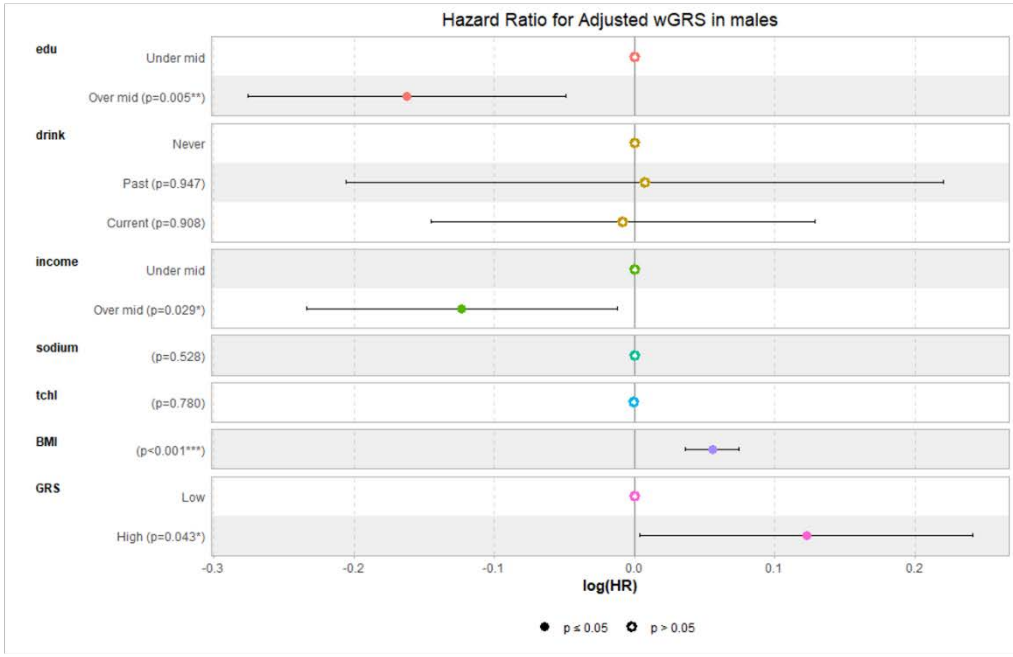
When non-Type A personality was the reference, Type A personality has 1.15 times (p=.02) higher hazard ratio for males and 1.19 times (p=.01) higher one for females.

As a result of analysis by adding the interaction term, there was no significant interaction between the variables, and the hazard ratio for Type A increased from 1.15 to 1.16 (p=.04) for males and from 1.19 to 1.22 (p=.01) for females.

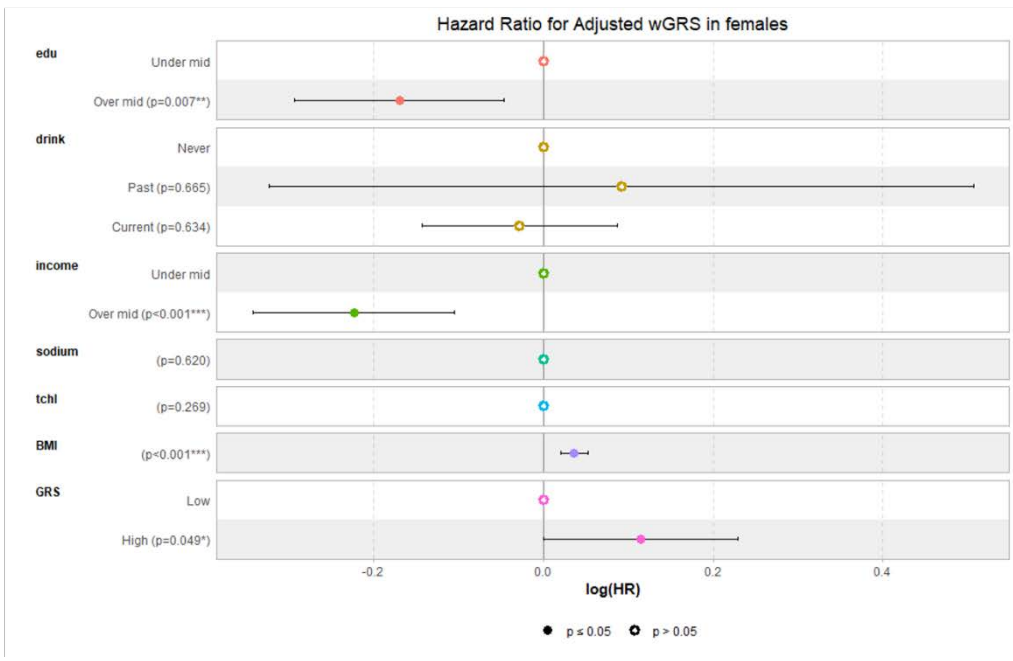
Table 3.8 The results of Cox PH model for adjustment with and without interaction

	Multivariate (for adjustment)						Multivariate with Interaction (for adjustment)					
Gender	**Male**			**Female**			**Male**			**Female**		
Variables	**β**	**HR (95% CI)**	**p**	**β**	**HR (95% CI)**	**p**	**β**	**HR (95% CI)**	**p**	**β**	**HR (95% CI)**	**p**
wGRS												
Low		Reference			Reference			Reference			Reference	
High	**0.13**	**1.13 (1.00-1.27)**	**0.04 ***	**0.11**	**1.11 (1.00-1.26)**	**0.04***	**0.13**	**1.14 (1.00-1.30)**	**0.04***	**0.12**	**1.14 (0.99-1.27)**	**0.05†**
Perceived stress												
Low		Reference			Reference			Reference			Reference	
High	-0.02	0.98 (0.82-1.18)	0.85	0.06	1.06 (0.93-1.22)	0.4	-0.006	1.00 (0.81-1.23)	0.95	0.07	1.07 (0.91-1.25)	0.41
Type A												
No		Reference			Reference			Reference			Reference	
Yes	**0.14**	**1.15 (1.02-1.30)**	**0.02 ***	**0.18**	**1.19 (1.05-1.35)**	**0.01** ***	**0.15**	**1.16 (1.01-1.33)**	**0.04***	**0.20**	**1.22 (1.05-1.41)**	**0.01****
wGRS * P.stress												
GRS low:stress low								Reference			Reference	
GRS high:stress high							-0.04	0.96 (0.62-1.50)	0.87	-0.03	0.99 (0.72-1.36)	.94
wGRS*Type A												
GRS low:Type A(not)								Reference			Reference	
GRS high:Type A							-0.02	0.98 (0.74-1.31)	0.91	-0.08	0.92 (0.68-1.24)	0.59

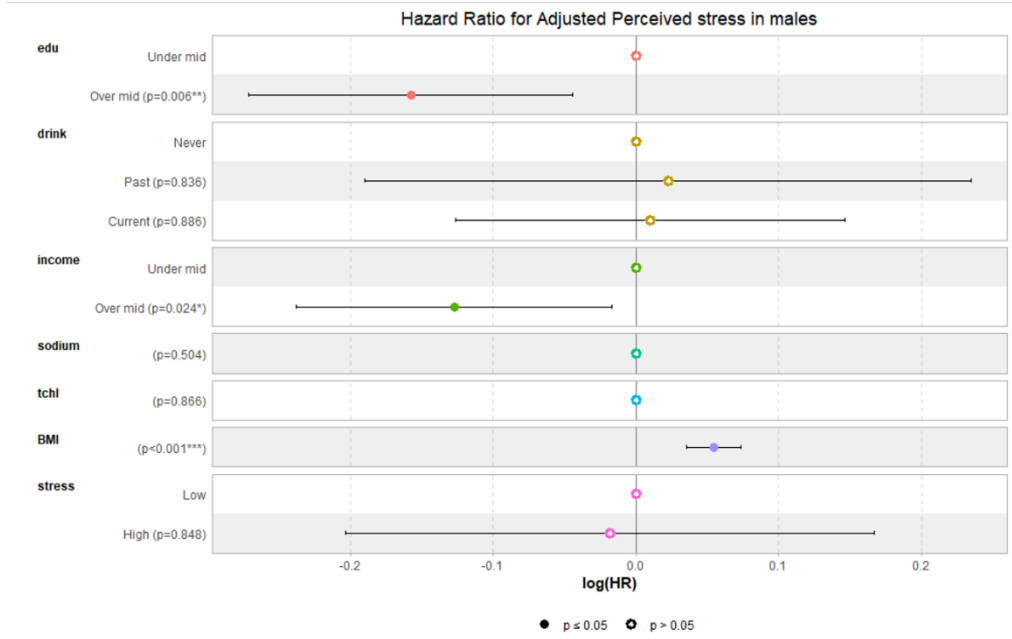
A1


Hazard Ratio for Adjusted wGRS in males
edu
Under mid
Over mid (p=0.005**)
drink
Never
Past (p=0.947)
Current (p=0.908)
income
Under mid
Over mid (p=0.029*)
sodium
(p=0.528)
tchl
(p=0.780)
BMI
(p<0.001***)
GRS
Low
High (p=0.043*)
-0.3
-0.2
-0.1
0.0
0.1
0.2
log(HR)
p ≤ 0.05
p > 0.05

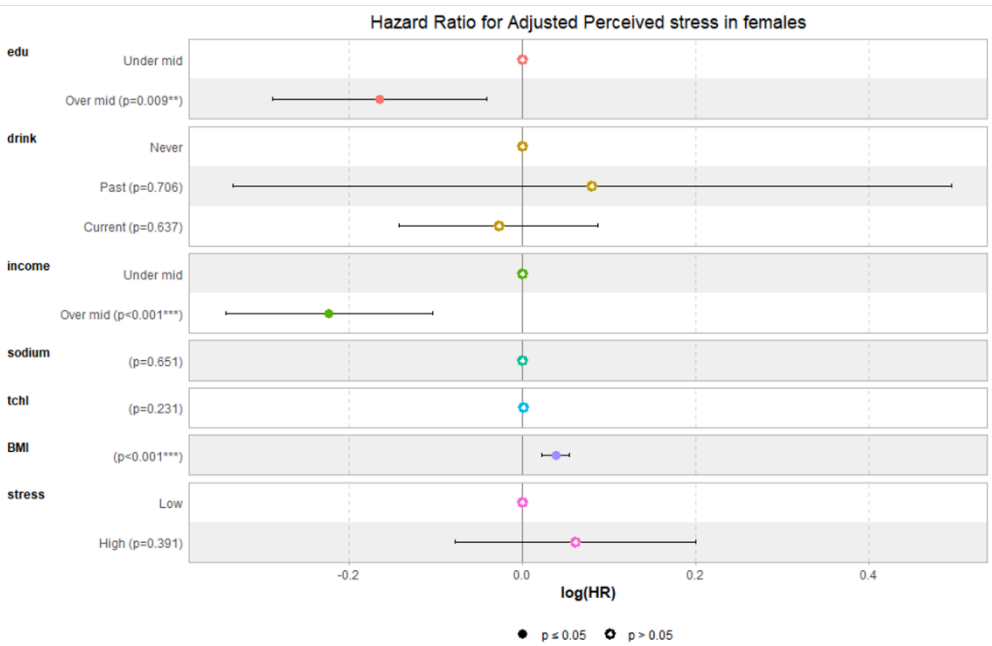
A2


Hazard Ratio for Adjusted wGRS in females
edu
Under mid
Over mid (p=0.007**)
drink
Never
Past (p=0.665)
Current (p=0.634)
income
Under mid
Over mid (p<0.001***)
sodium
(p=0.620)
tchl
(p=0.269)
BMI
(p<0.001***)
GRS
Low
High (p=0.049*)
-0.2
0.0
0.2
0.4
log(HR)
p ≤ 0.05
p > 0.05

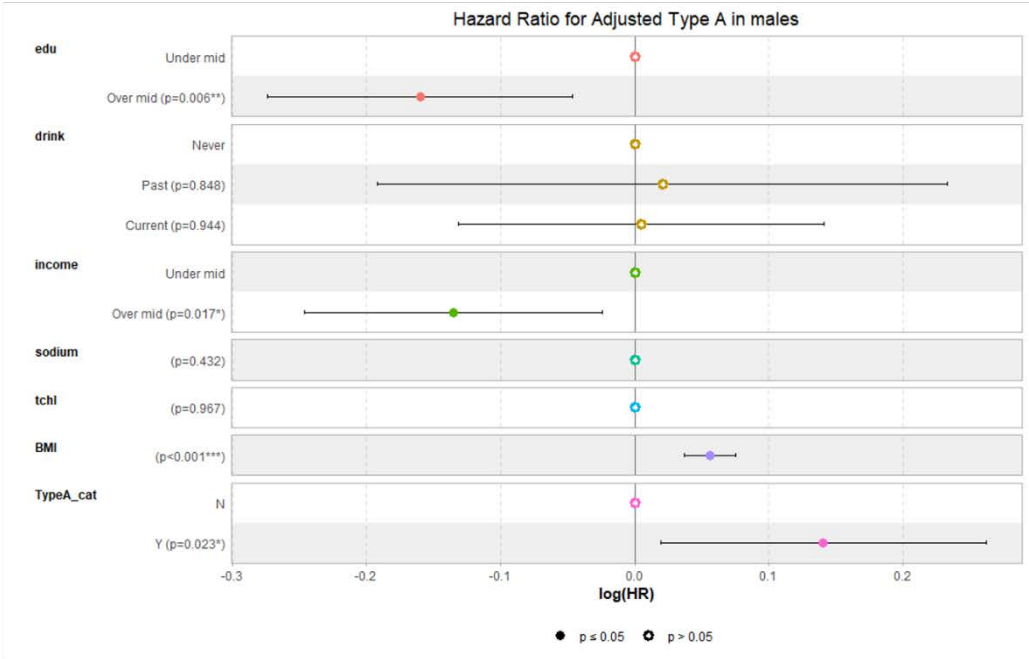
B1


Hazard Ratio for Adjusted Perceived stress in males
edu
Under mid
Over mid (p=0.006**)
drink
Never
Past (p=0.836)
Current (p=0.886)
income
Under mid
Over mid (p=0.024*)
sodium
(p=0.504)
tchl
(p=0.866)
BMI
(p<0.001***)
stress
Low
High (p=0.848)
-0.2
-0.1
0.0
0.1
0.2
log(HR)
p ≤ 0.05
p > 0.05

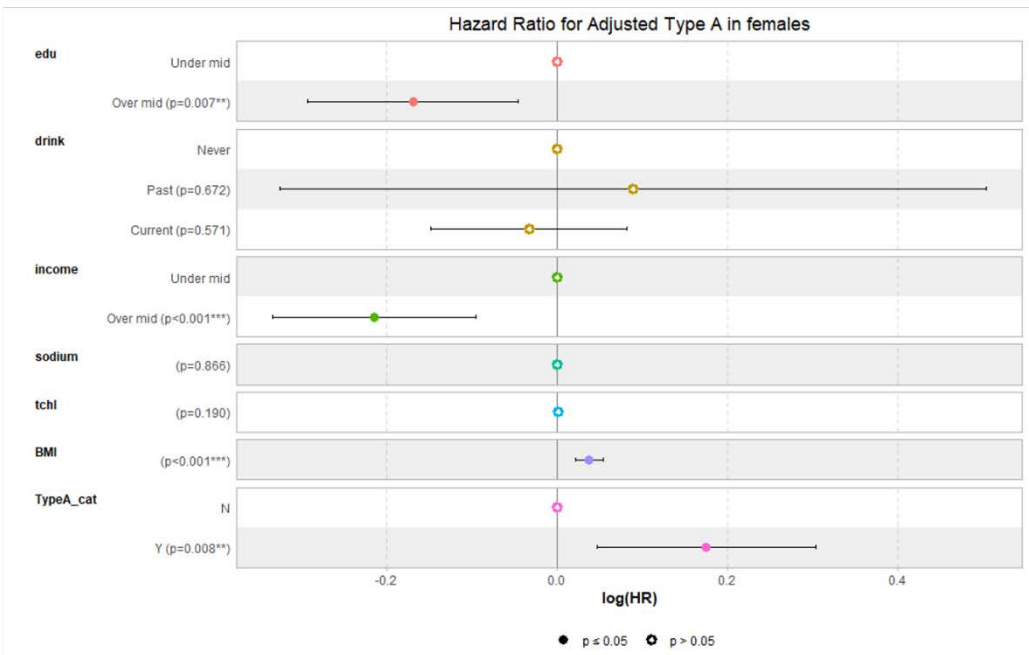
B2


Hazard Ratio for Adjusted Perceived stress in females
edu
Under mid
Over mid (p=0.009**)
drink
Never
Past (p=0.706)
Current (p=0.637)
income
Under mid
Over mid (p<0.001***)
sodium
(p=0.651)
tchl
(p=0.231)
BMI
(p<0.001***)
stress
Low
High (p=0.391)
-0.2
0.0
0.2
0.4
log(HR)
p ≤ 0.05
p > 0.05

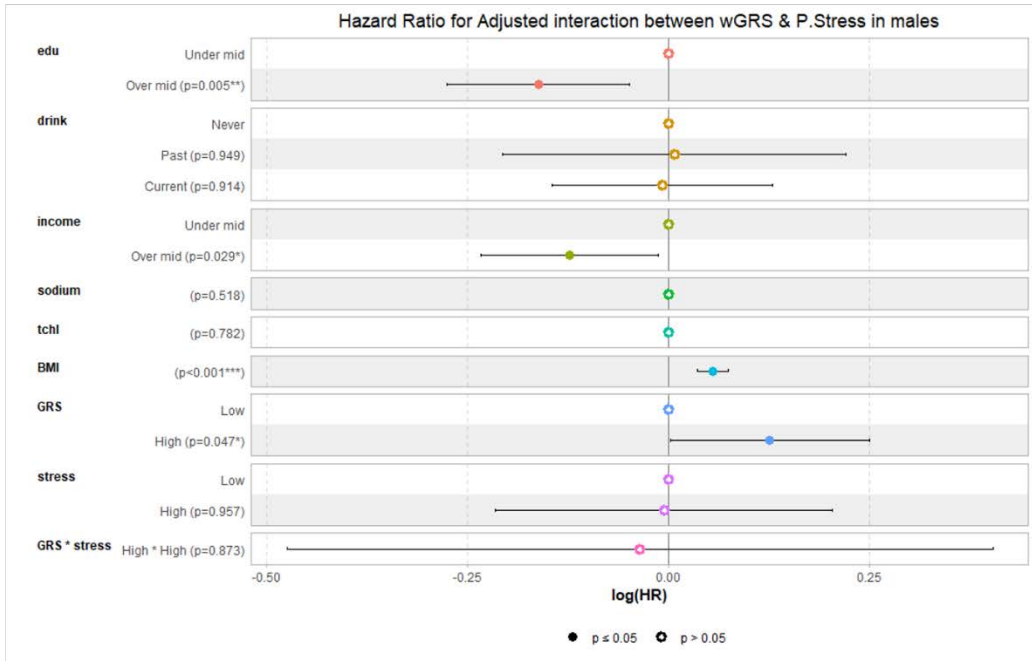
C1


Hazard Ratio for Adjusted Type A in males
edu
Under mid
Over mid (p=0.006**)
drink
Never
Past (p=0.848)
Current (p=0.944)
income
Under mid
Over mid (p=0.017*)
sodium
(p=0.432)
tchl
(p=0.967)
BMI
(p<0.001***)
TypeA_cat
N
Y (p=0.023*)
-0.3
-0.2
-0.1
0.0
0.1
0.2
log(HR)
p ≤ 0.05
p > 0.05

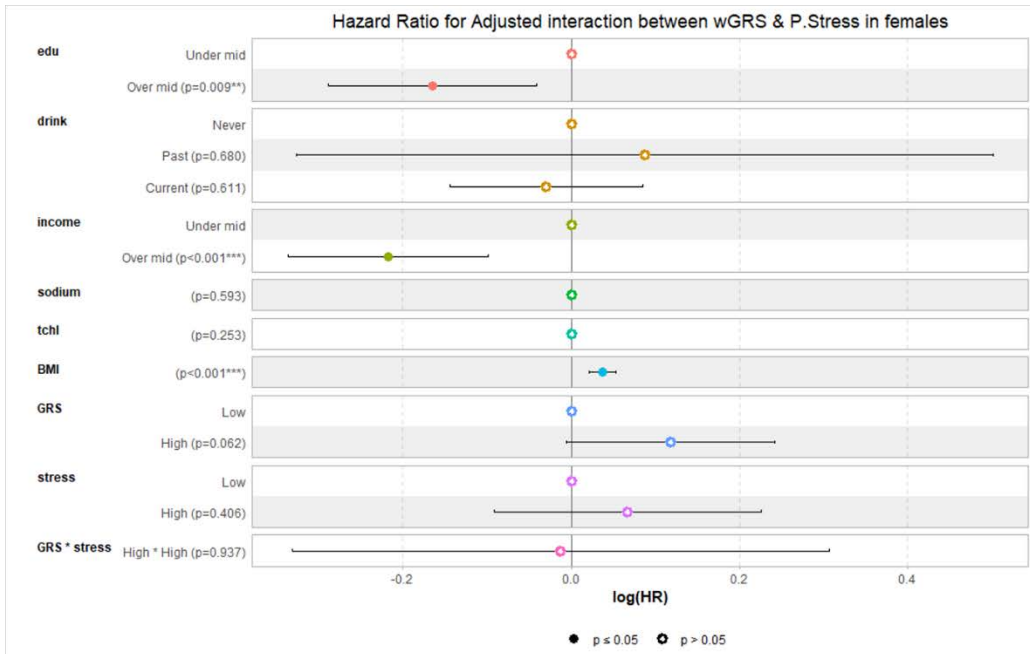
C2


Hazard Ratio for Adjusted Type A in females
edu
Under mid
Over mid (p=0.007**)
drink
Never
Past (p=0.672)
Current (p=0.571)
income
Under mid
Over mid (p<0.001***)
sodium
(p=0.866)
tchl
(p=0.190)
BMI
(p<0.001***)
TypeA_cat
N
Y (p=0.008**)
-0.2
0.0
0.2
0.4
log(HR)
p ≤ 0.05
p > 0.05

D1


Hazard Ratio for Adjusted interaction between wGRS & P.Stress in males
edu
Under mid
Over mid (p=0.005**)
drink
Never
Past (p=0.949)
Current (p=0.914)
income
Under mid
Over mid (p=0.029*)
sodium
(p=0.518)
tchl
(p=0.782)
BMI
(p<0.001***)
GRS
Low
High (p=0.047*)
stress
Low
High (p=0.957)
GRS * stress
High * High (p=0.873)
-0.50
-0.25
0.00
0.25
log(HR)
p ≤ 0.05
p > 0.05

D2


Hazard Ratio for Adjusted interaction between wGRS & P.Stress in females
edu
Under mid
Over mid (p=0.009**)
drink
Never
Past (p=0.680)
Current (p=0.611)
income
Under mid
Over mid (p<0.001***)
sodium
(p=0.593)
tchl
(p=0.253)
BMI
(p<0.001***)
GRS
Low
High (p=0.062)
stress
Low
High (p=0.406)
GRS * stress
High * High (p=0.937)
-0.2
0.0
0.2
0.4
log(HR)
p ≤ 0.05
p > 0.05

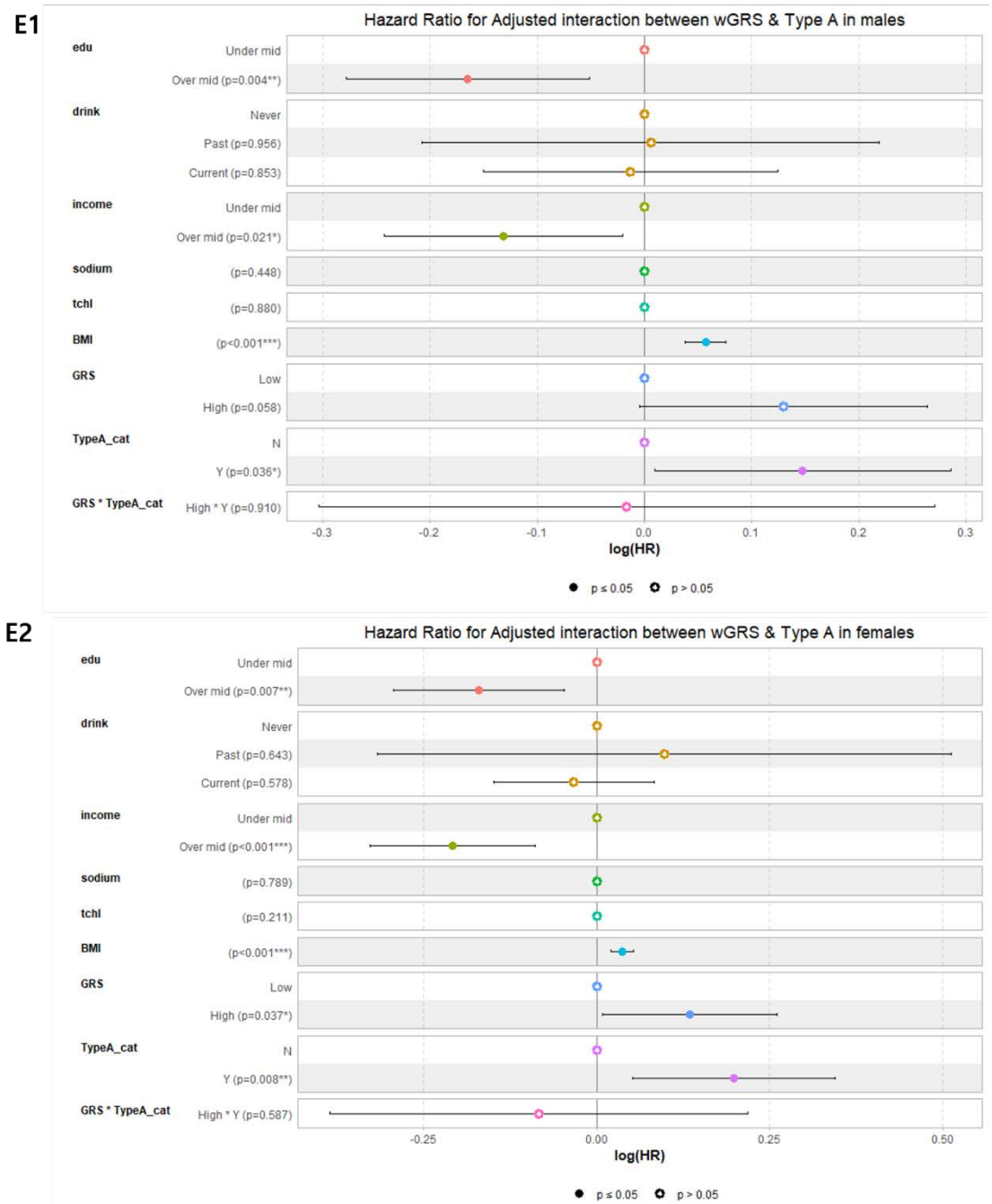


Figure 3.6 Hazard Ratio of each variable with or without interaction A) for wGRS B) for Perceived Stress C) for Type A D) for wGRS **x** Perceived Stress interaction E) for wGRS **x** Type A interaction 1) for males 2) for females

3.5. The Contribution of Perceived Stress and Type A

3.5.1. Performance Improvement of prediction models

Table 3.9 summarizes the performance changes of the model with features added to the basic model. The indicated value is the average of the values obtained in 5 runs.

The reference model is a model featuring covariates and wGRS, Model 1 is the reference model plus Type A, Model 2 is model 1 plus perceived stress, Model 3 adds both Type A and perceived stress to Model 1.

As shown in Table 3.9, overall, the AUC was better when using XGBoost with female data than Decision Tree or Random Forest.

As for the predictive performance of Decision Tree with male data, both NRI and IDI values decreased when only Type A was added (NRI=−0.02, IDI=−0.39), and increased when perceived stress was added (NRI=0.02, IDI=0.85). The performance was improved the most when both Type A and perceived stress were added (NRI=0.03, IDI=1.73).

In Random Forest, performance was improved only when Type A was added, and in XGBoost, only perceived stress was the feature that improved performance in all values of NRI and IDI.

For female data, all algorithms performed poorly in the model in which only perceived stress was added. Meanwhile, in all algorithms, the performance was the most improved only when Type A was added.

Figure 3.7 shows NIR and IDI, which are predictive performance improvement indicators for each algorithm for data divided by gender.

For male data, IDI of Decision Tree and XGBoost showed the highest improvement, while in the female data, it showed that the IDI values of Decision Tree, Random Forest, and XGBoost were stably improved in Model 1 in which only Type A was added.

NRI is the sum of the predicted improvement of the occurrence of an event and the improvement of the prediction of the non-event. IDI is an indicator of whether the added factor increases the probability of occurrence of an event. Therefore, the fact that the IDI value is relatively higher than that of the NRI in this result can be interpreted that the added factor is more closely related to the onset of hypertension, which is consistent with the previous Cox PH model results.

Table 3.9 The improvement of performance for features by algorithms

Gender	Model	Decision Tree			Random Forest			XGBoost		
		AUC	NRI	IDI	AUC	NRI	IDI	AUC	NRI	IDI
Male	Reference	0.52	-0.02	-0.39	0.69	**0.04**	**0.45**	0.70	-0.11	1.0
	Model 1	0.51			0.69			0.67		
	Reference	0.51	**0.02**	**0.85**	0.53	-0.08	-0.53	0.57	**0.07**	**0.47**
	Model 2	0.52			0.53			0.60		
	Reference	0.59	**0.03**	**1.73**	0.63	-0.02	-0.05	0.60	-0.03	1.48
	Model 3	0.60			0.61			0.58		
Female	Reference	0.59	**0.28**	**1.41**	0.77	**0.01**	**0.84**	0.73	**0.01**	**1.81**
	Model 1	0.60			0.78			0.77		
	Reference	0.57	-0.04	-1.95	0.76	-0.03	-0.58	0.70	-0.04	1.83
	Model 2	0.55			0.75			0.71		
	Reference	0.55	**0.004**	**0.47**	0.73	-0.02	0.46	0.69	**0.04**	**1.22**
	Model 3	0.55			0.74			0.70		

Reference: covariates + wGRS, Model 2: covariates + wGRS + Type A, Model 3: covariates + wGRS + Perceived Stress, Model 4: covariates + wGRS + TypeA + Perceived Stress

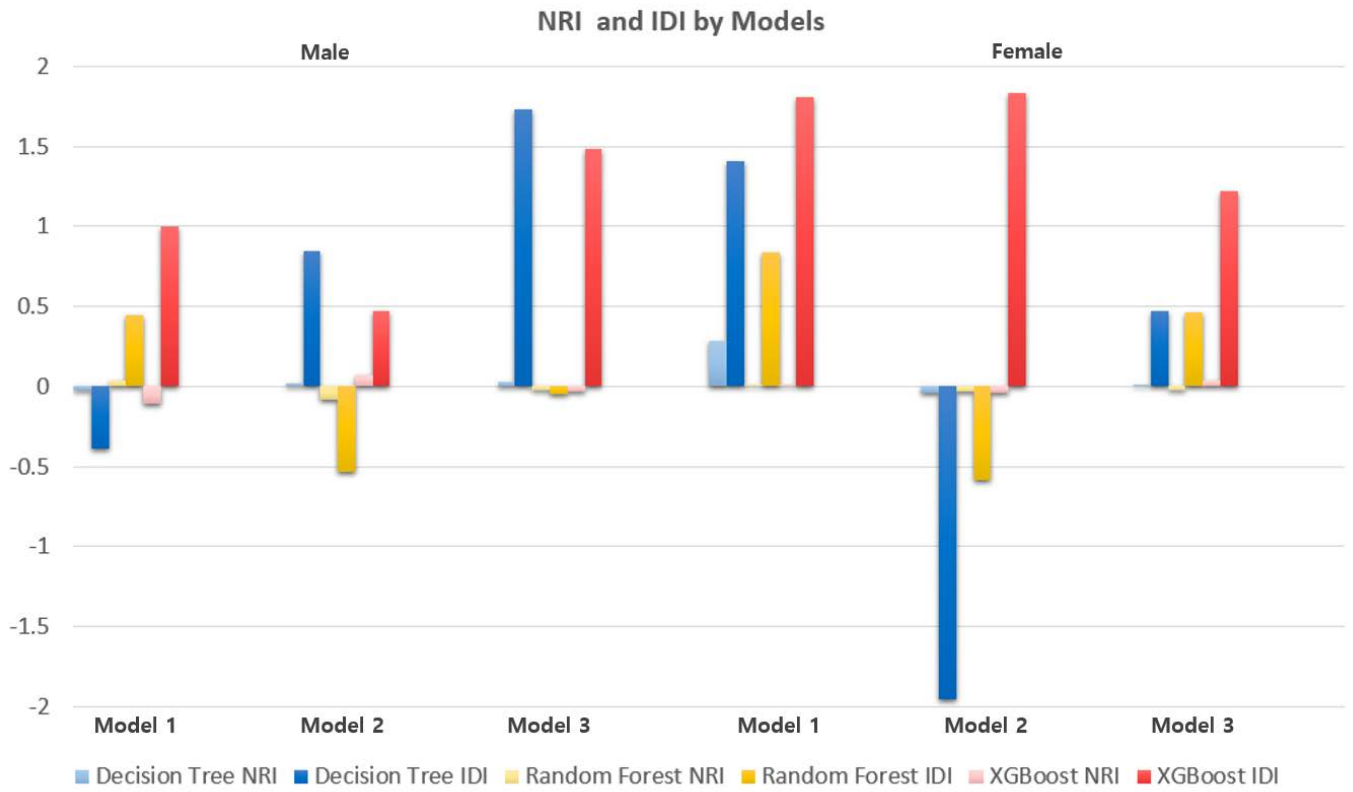


Figure 3.7 The degree of performance improvements for features by algorithms

3.5.2. SHAP as Feature Importance

Briefly explaining SHAP, it is the average of all combinations that the model can express by subtracting the combination of features except for the feature whose contribution is needed to know.

In the current study, the SHAP values were derived from the training result of XGBoost, which had the best performance.

Figure 3.8 shows the absolute value of the calculated influence of each feature.

In male data, wGRS had the greatest influence other than covariates (0.67), perceived stress influenced 0.47, and Type A influenced 0.36.

In the female data, the influence of wGRS was as high as 0.63, that of Type A had 0.4, and that of the perceived stress had the effect of 0.34.

Therefore, it can be interpreted that Type A and Perceived Stress contribute as risk factors to the predictive model. However, since this value is absolute, it is not known whether it positively or negatively affects the development of hypertension. In order to know the direction of the influence, it is necessary to visualize the influence by dividing it into positive and negative effects.

Figure 3.9 indicates how each feature affects the Shapley value distribution. Each data point means the value of a sample for the feature. Simply put, if the SHAP value is positive, the probability of predicting it as 1 is high, and if it is negative, the probability of predicting it as 0 is high.

In the case of male data, it is difficult to say that the trend of perceived stress is clearly displayed because it is mainly concentrated at 0 of the SHAP value and looks mixed. However, the SHAP values are closer to the positive side when the feature values are low (blue), and the SHAP values are closer to the negative side when the feature values are high (red). This means that when the

stress score is low, it contributes more to the prediction of the onset of hypertension as 1, and in the opposite case, the probability of predicting it as 0 increases.

Meanwhile, in Type A, the higher the feature value (red), the higher the SHAP value, which positively contributes to the prediction of the onset of hypertension as 1. Also, the lower the feature value (blue), the more negative the SHAP value, suggesting that it contributes to negatively predicting the onset of hypertension as 0.

Also in female data, when the feature value of Type A was high (red), the SHAP value had a positive value, and when the feature value was low (blue), the SHAP value had a negative value. This indicates that the stronger the tendency of Type A, the more positively it affects the prediction of hypertension as 1.

However, in the case of perceived stress, it can be inferred that the feature values could not have contributed significantly to the classification since low feature values are at both extremes of the SHAP value and are mixed, which is not consistent.

For both male and female data, wGRS and Type A showed positive SHAP values as the feature value increased, suggesting that they had a positive relationship in predicting the onset of hypertension.

This is consistent with the analysis results of the Cox PH model, where wGRS and Type A had positive effects on the onset of hypertension, but perceived stress had no significant effect.

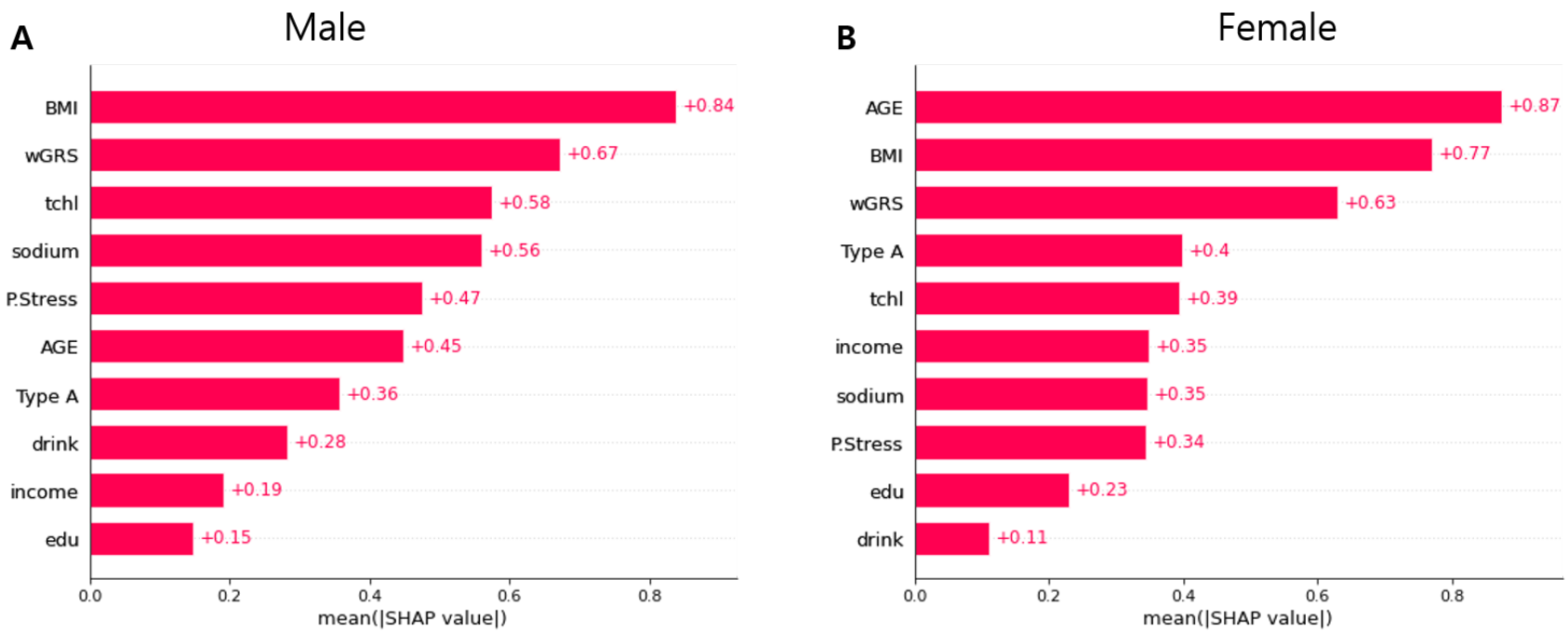


Figure 3.8 The averaged absolute SHAP values for features by gender

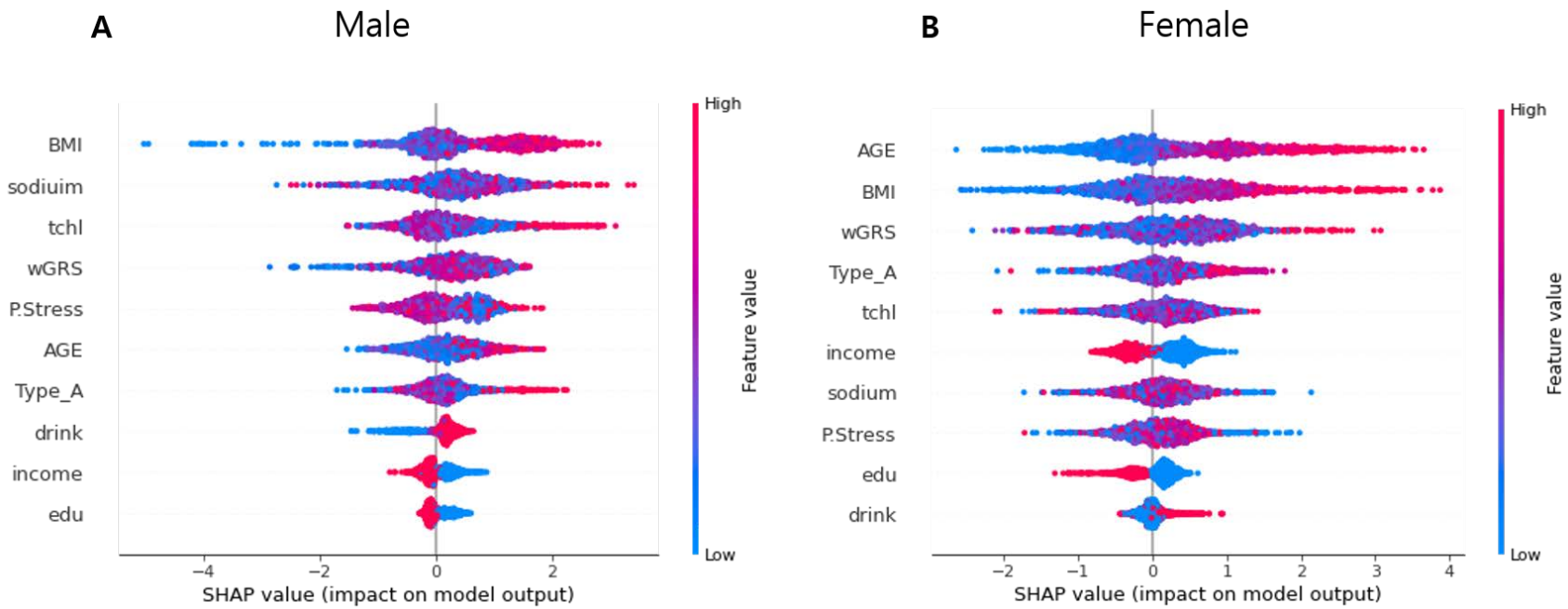


Figure 3.9 The distribution of SHAP values of each feature by gender

Chapter 4. Discussion

The purpose of this study was to investigate whether personality is valuable as a predictor of hypertension by analyzing it from multiple angles.

First, the causal structure was looked into of non-genetic risk factors with cross-sectional data (KoGES 3rd period data, 2005-2006) according to the hypothesis that 'Type A personality is vulnerable to stress, and stress will influence hypertension by mediating BMI'. As a result of the analysis, Type A personality did not directly affect hypertension in both males and females. However, as hypothesized, Type A showed a positive relationship with perceived stress, and high stress had a negative effect on BMI, and consequently lowered the likelihood of developing hypertension.

Second, survival analysis was conducted with longitudinal data from 2001 to 2016 to examine weighted genetic risk scores (wGRS), perceived stress, Type A, and their interactions in the onset of hypertension. As a result of the analysis of the Cox PH model adjusted for the covariates, there was no interaction between each variable, and wGRS and Type A significantly affected the onset of hypertension respectively. Perceived stress did not show a significant effect.

Third, whether these variables actually contribute to the improvement of the hypertension onset prediction model was confirmed through the tree-based model, which are Decision Tree, Random Forest, and XGBoost. Prediction results were better in female data than in male ones. Also, the performance was the best when using XGBoost.

As a result of comparing the performance by adding Type A and perceived stress with the covariates and wGRS as references, NRI and IDI were shown to increase to positive values, suggesting that

they contribute to prediction. In addition, in the result of calculating the SHAP value in the same model, Type A showed an effective contribution to the prediction of hypertension.

4.1. Causal structure of psychosocial factors involved in the onset of hypertension

According to the results of the current study, the properties of Type A such as anger and impatience might make it easier to feel psychological stress in the daily lives in both male and female groups.

For the both groups, type A could not affect SBP and DBP directly but indirectly could affect via perceived stress and BMI. In our structural model, Type A personality positively affects perceived stress and perceived stress negatively influences BMI. In the sense that BMI are positively associated with the onset of hypertension, it may explain high perceived stress might lower the blood pressure through lowering BMI. After all, the effect of Type A on the development of hypertension was only significant indirectly.

These results are different from those of the major early studies and similar to recent studies (Ducher et al., 2006). In early studies, it has been known people of Type A behavior pattern tend to have increased cholesterol level, delayed blood coagulation time or coronary artery disease (CAD) risk (Friedman & Rosenman, 1959) as well as CVD (P. P. Chang, Ford, Meoni, Wang, & Klag, 2002; Gallacher, Sweetnam, Yarnell, Elwood, & Stansfeld, 2003; Suls & Bunde, 2005). In recent studies, however, evidence that Type A directly affects HTN or CVD is mixed or feeble (Sahoo, Padhy, Padhee, Singla, & Sarkar, 2018).

As a result of this study using cross-sectional data, the fact that Type A does not directly affect HTN is the same as relatively recent results. At least for a short period of time, Type A has no effect on hypertension. Rather, according to this structural model, Type A had

an indirectly negative effect on the onset of hypertension by mediating perceived stress and BMI.

Meanwhile, considering the time-compulsive, competitive, and easily irritable Type A traits, it is reasonable to speculate that these traits are susceptible to stress (Byrne & Rosenman, 1986; Lecic-Tosevski, Vukovic, & Stepanovic, 2011), as significantly represented in SEM of the present study.

What is unusual in this study is that perceived stress has a negative effect on BMI and has no significant effect on hypertension, which seems quite different from the flow of many studies.

In fact, a majority of studies claim perceived stress is positively related to increasing BMI behaviors and exposure to stressful situations raises the development of HTN. However, the study results of relationships between perceived stress and BMI and NTN have been inconsistent.

For instance, perceived stress in the workplaces positively influences BMI due to overeating or sedentary behaviors (Kouvonen, Kivimäki, Cox, Cox, & Vahtera, 2005), in addition to losing self-regulation especially in women (Moore & Cunningham, 2012; O'Neill, Kamper-DeMarco, Chen, & Orom, 2020). In these studies, the authors indicated highly perceived stressed people with their job control or strain tended to eat energy-dense foods as comfort food to relieve their perceived stress and feel better (Dallman et al., 2003; Kouvonen et al., 2005).

On the other hand, there are shreds of evidence to support negative emotions caused by perceived stress or anxiety that were associated with low blood pressure. These studies suggest the possibility that perceived stress may inhibit appetite and reduce eating behavior (Stone & Brownell, 1994).

Especially according to the animal models (Dallman & Bhatnagar, 2010; Willner, 1997), the rats exposed to repeated stressors showed hypophagia (reduced food intake), and in the studies of Kivimaki et. al, non-obese people had lower BMIs due to perceived stress (Kivimäki et al., 2006). These inhibited appetites by negative

emotions occur through activation of the sympathetic nervous system which suppresses upper gastrointestinal motility (Wardle & Gibson, 2002).

The relationship between perceived stress and hypertension is also inconsistent. It is generally known that stressors make blood pressure higher (Sparrenberger et al., 2009; Vrijkotte, Van Doornen, & De Geus, 2000) and exposure to stressful environments increase the risk of HTN, but the opposite results also exist. For instance, in the recent study of Miguet et. al., after adjusting confounders, perceived stress lowers blood pressure in a Swedish cohort (Miguet et al., 2021). Similarly, psychological distress had marginally negative association with SBP and DBP in the lean people (Toyoshima, Otsuka, Hashimoto, Tamakoshi, & Yatsuya, 2014).

Therefore, one possible explanation for this result is that the negative emotions and anxiety that stress causes are likely to suppress appetite. Actually, in the post-hoc test comparing the daily food intake between the high-stress group and the low-stress group, the high-stress group had a lower food intake, which was significant for both males and females. This pattern is still significant after adjusting with covariates such as age, education, income, drink, and total cholesterol.

Of course, the effect of psychologically perceived stress may vary depending on individuals. According to Simmons et al., found depression subgroups increasing appetite were associated with hyper-activation of putative meso-corticolimbic reward (such as comfort food) circuitry, while those decreasing appetites were associated with hypo-activation of insula regions, which contributed to individual differences (Simmons et al., 2020).

In the study of Hong & Hong (Hong & Hong, 2019), among the 11,782 Korean participants, the underweight group of the four categorizations depending on BMI had the highest odds ratio for depression. The authors explained that fasting, reducing food

consumption, and skipping a meal were associated with depression symptoms. This trend was also observed in other studies on the Korean population. Kim et.al, reported that the high correlation between lean people and depression in the Korean population (Jinseok Kim, Noh, Park, & Kwon, 2014). Or chronic stress may not be significant for obesity in Koreans. (A. K. Chang & Choi, 2015; Ha & Park, 2012). According to the study of Chang & Choi, perceived stress was higher in the normal weight group than under or over weight group. She suggested obesity is less related to mental health in Koreans.

Thus, considering depression is one of the main responses of psychological distress and the tendency of Koreans to stress and obesity, it could be suggested Koreans might be prone to avoid meals or intake less when they feel stressed.

To sum up, considering the characteristics of Type A such as impatient, competitive, short-tempered, compulsive to time, easy to anger, etc., it is feasible that Type A personality people would be easy to be perceived stress in daily lives including jobs (Orpen, 1982), which increases the likelihood of experiencing negative emotions (Cohen, 2000). Although further studies should support this view, Type A people might tend to avoid eating or take in less in the stressful situation, which links to low BMI and the onset of hypertension at least in the short term.

4.2. Interaction of genetic and psychosocial factors from a long-term perspective

In the cross-sectional data, Type A and perceived stress did not have a direct effect on HTN, but considering that personality affects lifestyle throughout life, it is necessary to analyze whether it is involved in the development of HTN in the long term aspect.

Also, it is worth analyzing in the long term whether Type A and perceived stress, which are involved in psychosocial factors as non-

genetic factors, interact with genetic risk to change the incidence rate.

Therefore, survival analysis was performed using the Cox proportional Hazard model with longitudinal (2001–2016) data adjusting the covariates. As a result, it was found that higher wGRS and Type A increased the HR of hypertension in both gender groups, respectively. The effect of perceived stress was not significant and no interaction effect was observed.

In a cross–sectional study using only 3rd phase data, perceived stress had an indirect negative effect on the development of hypertension through BMI, but it was not an independent risk factor in survival analysis using longitudinal data.

In contrast to the inborn genetic risk (wGRS) or congenital personality traits (Type A) which are likely to change little throughout life, the perception of stress is highly variable and highly influenced by circumstances. Since this study used stress measurements for one period, it is unreasonable to assume that this stress continued to affect later life.

Therefore, it may be more reasonable to assume that stress–sensitive personalities (Type A) more frequently experience stress response rather than one period of stress experience in a long term aspect. In other words, Type A people are more likely to experience increased catecholamines by stimulating their sympathetic nerves with anger, anxiety, and competition even in situations that are not perceived as stressful, and are more likely to have high blood pressure as their daily lives continue for the rest of their lifetime (Flaa, Eide, Kjeldsen, & Rostrup, 2008).

Meanwhile, the interaction of psychological stress with genetic risk for cardiovascular disease is not much clear. In the study of Svensson et.al., psychological stress did not have any significant interaction on GRS for Myocardial infarction (MI), coronary arterial disease (CAD), and cardiovascular death. However, when constructing stress–sensitive GRS including SNPs which interact unfavorably with stress for CVD, they found significant interaction

with stress and CAD, MI and cardiovascular death (Svensson, Kitlinski, Engström, & Melander, 2017). They found perceived stress is not independently associated with the end points of CVDs as well as does not interact with any GRS of CVDs. Instead, perceived stress only interacts with specific SNP variants.

Although the research of Svensson et.al. shed light on the relationship between CVD and stress, not blood pressure, but considering that hypertension is a close predictor of CVD, the results can be considered to be in line with the results of the current study.

Thus, it cannot be ruled out that the results on the interaction of the current study could be different if the risk score would be recalculated by constructing stress-sensitive SNPs among 26 SNPs from being used for wGRS for HTN. If this is confirmed, perceived stress is more likely to interact with individual genetic mutations than with just wGRS, the sum of variants in terms of simple polygenicity.

wGRS and Type A had only the main effect on the onset of HTN, respectively, but had also no interaction effect.

As for wGRS, when the low-risk group was used as a reference, the HR in the high-risk group was 1.12 fold significantly higher in males and 1.14 fold meaningfully higher in females. In short, albeit wGRS influenced the development of HTN significantly, there was minimal change in HR values between the high and low groups.

Therefore, considering the stability of wGRS that significantly affects even after being adjusted for the crucial covariates, it may be the intervention of other environmental factors or some physiological events such as menopause for females that make a clinical difference (Davis et al., 2007; Polotsky & Polotsky, 2010).

4.3. Relevance of Type A and Stress as Predictors of Hypertension

To determine whether Type A personality and perceived stress have suitability as predictors of hypertension, in this study,

performance improvement and SHAP values were compared through tree-based algorithms.

The variables of the reference model consisted of age, BMI, total cholesterol, salt intake, monthly income, education level, drinking and wGRS, and performance was compared by adding Type A or perceived stress to each or all of them. As improvement indicators, not only AUC, but NRI (Net reclassification improvement), and IDI (Integrated Discrimination Improvement) were used, which are types of reclassification methods.

The SHAP values were calculated with all features including Type A and perceived stress through the results of XGBoost.

Although the overall prediction performance was not very good, there are some confirmed facts.

First, when only Type is added, or when Type A and perceived stress are both added as variables, NRI and IDI are expressed as positive numbers, which means improvement in actual performance. On average, the model that included both Type A and perceived stress performed the best in the male data, and the model that included only Type A performed the best in female data.

Second, not only was the overall predictive performance better, but the improvement of the model was also greater in the female data.

Third, in the calculation results of average absolute SHAP values, the contribution of Type A and perceived stress was different in the male and female data. In male data, the SHAP value of perceived stress (0.47) was higher than that of Type A (0.36), but in female data, the SHAP value of Type A (0.4) was higher than that of perceived stress (0.34).

As shown in Figure 3 in the results of Chapter 3, which visualizes the SHAP value and its influence for all features, it is relatively clear that the SHAP value of Type A has a positive relationship with the feature value in the data of both genders, which means high feature values of Type A contribute on predicting hypertension as 1.

However, In the case of perceived stress, male data has a large distribution and looks to have a negative relationship, whereas in female data, low feature values occupy both extremes of SHAP values and have a static relationship in a part close to 0, which appears mixed. Since the direction of the data for perceived stress is not clear and mixed, the contribution to the model that classifies 1 and 0 would have been low. This can explain why the performance was the best when only Type A was added in the female data when comparing the model performance.

However, in case that the distribution is small or the magnitude of the absolute value is small with being offset by 0, the influence for the accurate prediction seems insignificant.

Although the influence of perceived stress was higher than that of Type A in the male data, the relationship between the perceived stress feature value and the SHAP value does not look distinct enough to contribute to the accurate classification of 1 and 0. Perhaps the low XGBoost prediction accuracy of male data might be due to this unconvincing influence of these features, which was improved after adding Type A data.

To summarize, the predictive performance of female data was better than that of male data. In addition, in the female data, type A had a positive effect on accurate prediction, and perceived stress actually interfered. In the male data, the prediction performance was the most improved when both type A and perceived stress were added.

Also, as a result of calculating the SHAP value and comparing the distribution by features, a positive relationship between Type A and SHAP values was clearly found, and in the case of perceived stress, a mixed appearance was observed.

If we interpret this in a clinical sense, it can be said that the female data have a clearer pattern of factors predicting the onset of hypertension at least in these data. In particular, in the case of

females, the prevalence in their 40s is very low, but in their 60s, the proportion rises to a greater extent than that of males. Although this is not yet established evidence, hormonal changes such as menopause may have an effect (Polotsky & Polotsky, 2010). These hormonal changes can affect the stress perception levels (Bauld & Brown, 2009; Nosek et al., 2010) or eating patterns (Duval et al., 2014). Therefore, it seems reasonable that age has the highest contribution to the SHAP value in females, and was helpful for the prediction.

Chapter 5. Conclusion & Limitations

The current study examined the effect and structure of personality (Type A) and perceived stress, which are psychosocial factors among non-genetic factors, and their influence on the onset of hypertension as well as investigated whether those factors have an interaction with genetic factors and can be valuable predictors for hypertension. This approach has significance in that it explores the interaction relationship as well as the independent influence of genetic and psychosocial factors in the explanation of the pathogenesis of complex diseases such as hypertension.

As a result of the SEM, cross-sectionally, Type A personality increased perceived stress, which lowered BMI, and consequently, indirectly affected the development of hypertension in a negative direction. However, in survival analysis using longitudinal data, perceived stress had no significant effect, but weighted genetic risk scores and Type A personality were found to significantly affect the development of hypertension for both males and females.

In addition, when Type A for females and Type A and perceived stress for males were both used as features, the performance of the predictive model was improved. In other words, they were confirmed to have value as predictors.

However, this study also has some limitations.

The first limitation is that the tools to measure Type A and perceived stress were self-report questionnaires.

For Type A, it can be more accurate that a trained expert consults and observes than self-report questionnaires (Day, Therrien, & Carroll, 2005). Moreover, Type A needs to be measured under the premise that the emphasized part or personality is different depending on the culture.

Also, the stress measurement tool used in the present study was initially developed to measure socio-psychological well-being (Lee

& Lee, 1996). In this study, perceived stress was operationally defined and used as the opposite concept of well-being (Winefield et al., 2012) in daily lives. Therefore, there is a need for more valid rationales as to whether the two extremes of mental health, well-being and stress, can be used as opposite concepts. Furthermore, it also became a limitation that not only were stress levels measured not directly but also the failure to utilize physiological evidence for stress, such as blood cortisol concentration or IL-6.

In terms of weighted genetic risk score, obtained as the linear sum of alleles known as risk alleles, is a simple indicator of the genetic risk of hypertension. In particular, complex diseases such as hypertension are multifactorial genetic diseases in which multiple genes are affected (Deng, 2007). Therefore, a method of calculating risk through multiple genes rather than a linear method may reflect more accurate genetic risk.

Despite these limitations, this study has significance in that it confirmed not only how much individual differences such as personality and psychosocial factors can influence the onset of complex diseases like hypertension, but also whether they interact with genetic factors, in the approach of Gene x Environment interaction.

In addition, the results of this study can be used as a basis for clinical development.

First of all, it helps to understand the mechanisms of complex disease by elucidating the relationships among non-genetic risk factors that are not much influential and dependent on each other.

Also, from a long-term perspective, it can be helpful in predicting or preventing disease through personality affecting lifestyle or repetitive behavioral responses. This approach could contribute as a starting point to the development of personalized medicine at the behavioral level.

In future research, it is necessary to confirm more accurate and

consistent results through more rigorous personality measurement tools and physiological indicators that can objectively check the level of stress.

This research approach will contribute to the development of personalized treatment and prevention using individual differences at the behavioral level as well as biological individual differences such as genomics.

Bibliography

Adrogué, H. J., & Madias, N. E. (2007). Sodium and potassium in the pathogenesis of hypertension. *New England journal of medicine, 356*(19), 1966-1978.

Al Asadi, J. (2010). Type A behaviour pattern: is it a risk factor for hypertension? *EMHJ-Eastern Mediterranean Health Journal, 16 (7), 740-745, 2010*.

Bauld, R., & Brown, R. F. (2009). Stress, psychological distress, psychosocial factors, menopause symptoms and physical health in women. *Maturitas, 62*(2), 160-165.

Bazhan, N., Makarova, E., AIu, S., & Iakovleva, T. (2007). Repeating of emotional stress prevents development of melanocortin obesity and type 2 diabetes in the mice with the Agouti yellow mutation. *Rossiiskii Fiziologicheskii Zhurnal Imeni IM Sechenova, 93*(11), 1237-1244.

Begley, T. M., & Boyd, D. P. (1985). The relationship of the Jenkins Activity Survey to Type A behavior among business executives. *Journal of Vocational Behavior, 27*(3), 316-328.

Bhelkar, S., Deshpande, S., Mankar, S., & Hiwarkar, P. (2018). Association between stress and hypertension among adults more than 30 years: A case-control study. *National Journal of Community Medicine, 9*(6), 430-433.

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj, 314*(7080), 572.

Breik, W., & Elbedour, S. (2021). The Predictive Ability of Type D Personality Pattern, Anxiety, and Depression in Cardiac Disease. *European Journal of Mental Health, 16*(2), 196-209.

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

Brotman, D. J., Golden, S. H., & Wittstein, I. S. (2007). The cardiovascular toll of stress. *The Lancet, 370*(9592), 1089-1100.

Byrne, D., & Rosenman, R. (1986). The Type A behaviour pattern as a precursor to stressful life-events: A confluence of coronary risks. *British Journal of Medical Psychology, 59*(1), 75-82.

Čatipović-Veselica, K., Glavaš, B., Kristek, J., & Šram, M. (2001). Components of type A behavior and two-year prognosis of patients with acute coronary syndrome. *Psychological reports, 89*(3), 467-475.

Chang, A. K., & Choi, J. Y. (2015). Factors influencing BMI classifications of Korean adults. *Journal of physical therapy science, 27*(5), 1565-1570.

Chang, P. P., Ford, D. E., Meoni, L. A., Wang, N.-Y., & Klag, M. J. (2002). Anger in young men and subsequent premature cardiovascular disease: the precursors study. *Archives of internal medicine, 162*(8), 901-906.

Chen, T., & Guestrin, C. (2016). *Xgboost: A scalable tree boosting system.* Paper presented at the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining.

Chrousos, G. P. (2009). Stress and disorders of the stress system. *Nature reviews endocrinology, 5*(7), 374-381.

Cohen, J. I. (2000). Stress and mental health: a biobehavioral perspective. *Issues in mental health nursing, 21*(2), 185-202.
Cooke Bailey, J. N., & Igo Jr, R. P. (2016). Genetic risk scores. *Current protocols in human genetics, 91*(1), 1.29. 21-21.29. 29.
Crouch, D. J., & Bodmer, W. F. (2020). Polygenic inheritance, GWAS, polygenic risk scores, and the search for functional variants. *Proceedings of the National Academy of Sciences, 117*(32), 18924-18933.
Dallman, M. F., & Bhatnagar, S. (2010). Chronic stress and energy balance: role of the hypothalamo-pituitary-adrenal axis. *Comprehensive physiology*, 179-210.
Dallman, M. F., Pecoraro, N., Akana, S. F., La Fleur, S. E., Gomez, F., Houshyar, H., . . . Manalo, S. (2003). Chronic stress and obesity: a new view of "comfort food". *Proceedings of the National Academy of Sciences, 100*(20), 11696-11701.
Davis, B., Arnett, D., Boerwinkle, E., Ford, C., Leiendecker-Foster, C., Miller, M., . . . Eckfeldt, J. H. (2007). Antihypertensive therapy, the α-adducin polymorphism, and cardiovascular disease in high-risk hypertensive persons: the Genetics of Hypertension-Associated Treatment Study. *The Pharmacogenomics Journal, 7*(2), 112-122.
Day, A. L., Therrien, D. L., & Carroll, S. A. (2005). Predicting psychological health: Assessing the incremental validity of emotional intelligence beyond personality, Type A behaviour, and daily hassles. *European Journal of personality, 19*(6), 519-536.
Deng, A. Y. (2007). Genetic basis of polygenic hypertension. *Human molecular genetics, 16*(R2), R195-R202.
Denollet, J. (1991). Negative affectivity and repressive coping: Pervasive influence on self-reported mood, health, and coronary-prone behavior. *Psychosomatic medicine, 53*(5), 538-556.
Denollet, J. (1998). Personality and risk of cancer in men with coronary heart disease. *Psychological Medicine, 28*(4), 991-995.
Denollet, J. (2000). Type D personality: A potential risk factor refined. *Journal of psychosomatic research, 49*(4), 255-266.
Diamond, E. L. (1982). The role of anger and hostility in essential hypertension and coronary heart disease. *Psychological Bulletin, 92*(2), 410.
Ducher, M., Fauvel, J.-P., & Cerutti, C. (2006). Risk profile in hypertension genesis: a five-year follow-up study. *American journal of hypertension, 19*(8), 775-780.
Duval, K., Prud'Homme, D., Rabasa-Lhoret, R., Strychar, I., Brochu, M., Lavoie, J., & Doucet, E. (2014). Effects of the menopausal transition on dietary intake and appetite: a MONET Group Study. *European journal of clinical nutrition, 68*(2), 271-276.
Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association, 83*(402), 414-425.
Eggers, A. E. (2007). Redrawing Papez'circuit: a theory about how acute stress becomes chronic and causes disease. *Medical hypotheses,*

69(4), 852-857.
Emond, M., Ten Eycke, K., Kosmerly, S., Robinson, A. L., Stillar, A., & Van Blyderveen, S. (2016). The effect of academic stress and attachment stress on stress-eaters and stress-undereaters. *Appetite, 100*, 210-215.
Flaa, A., Eide, I. K., Kjeldsen, S. E., & Rostrup, M. (2008). Sympathoadrenal stress reactivity is a predictor of future blood pressure: an 18-year follow-up study. *Hypertension, 52*(2), 336-341.
Friedman, M. (1974). *Type A behavior and your heart*: Fawcett.
Friedman, M., & Rosenman, R. (1971). Type A Behavior Pattern: its association with coronary heart disease. *Ann. Clin. Res., 3*(6), 300-312.
Friedman, M., & Rosenman, R. H. (1959). Association of specific overt behavior pattern with blood and cardiovascular findings: blood cholesterol level, blood clotting time, incidence of arcus senilis, and clinical coronary artery disease. *Journal of the American medical association, 169*(12), 1286-1296.
Gallacher, J. E. J., Sweetnam, P. M., Yarnell, J. W. G., Elwood, P. C., & Stansfeld, S. A. (2003). Is Type A Behavior Really a Trigger for Coronary Heart Disease Events? , *65*(3), 339-346. doi:10.1097/01.Psy.0000041608.55974.A8
Gerin, W., Zawadzki, M. J., Brosschot, J. F., Thayer, J. F., Christenfeld, N. J., Campbell, T. S., & Smyth, J. M. (2012). Rumination as a mediator of chronic stress effects on hypertension: a causal model. *International Journal of Hypertension, 2012*.
Gibson, G. (2010). Hints of hidden heritability in GWAS. *Nature genetics, 42*(7), 558-560.
Goel, M. K., Khanna, P., & Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *International journal of Ayurveda research, 1*(4), 274.
Grande, G., Romppel, M., & Barth, J. (2012). Association between type D personality and prognosis in patients with cardiovascular diseases: a systematic review and meta-analysis. *Annals of behavioral medicine, 43*(3), 299-310.
Grotto, I., Huerta, M., & Sharabi, Y. (2008). Hypertension and socioeconomic status. *Current opinion in cardiology, 23*(4), 335-339.
Gupta, R. (2016). Convergence in urban–rural prevalence of hypertension in India. *Journal of human hypertension, 30*(2), 79-82.
Ha, Y., & Park, H. (2012). Association between obesity and self-rated health in Korean males and females. *Journal of Korean Biological Nursing Science, 14*(3), 203-211.
Haynes, S., Feinleib, M., & Kannel, W. B. (1980). The relationship of psychosocial factors to coronary heart disease in the Framingham Study. III. Eight-year incidence of coronary heart disease. *American Journal of Epidemiology, 111*(1), 37-58.
Heilbrun Jr, A. B., & Friedberg, E. B. (1988). Type A personality, self-control, and vulnerability to stress. *Journal of Personality Assessment, 52*(3), 420-433.

Henry, J. P., & Cassel, J. C. (1969). Psychosocial factors in essential hypertension recent epidemiologic and animal experimental evidence. *American Journal of Epidemiology, 90*(3), 171-200.

Hernandez-Gonzalez, I., Tenorio, J., Palomino-Doza, J., Martinez Meñaca, A., Morales Ruiz, R., Lago-Docampo, M., . . . Perez-Olivares, C. (2020). Clinical heterogeneity of Pulmonary Arterial Hypertension associated with variants in TBX4. *PloS one, 15*(4), e0232216.

Hong, G. W., & Hong, S. M. (2019). Relationships among body mass index, body image, and depression in Korean adults: Korea National Health and nutrition examination survey 2014 and 2016. *Journal of obesity metabolic syndrome, 28*(1), 61.

Houston, B. K. (1983). Psychophysiological responsivity and the Type A behavior pattern. *Journal of Research in Personality, 17*(1), 22-39.

Hunter, D. J. (2005). Gene-environment interactions in human diseases. *Nature reviews genetics, 6*(4), 287-298.

Intersalt Cooperative Research Group. (1988). Intersalt: an international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion. *BMJ: British Medical Journal*, 319-328.

Irvine, J., Garner, D. M., Craig, H. M., & Logan, A. G. (1991). Prevalence of Type A behavior in untreated hypertensive individuals. *Hypertension, 18*(1), 72-78.

Jackson, F. L. (2014). Gene-environment interactions in human health: case studies and strategies for developing new paradigms and research methodologies. *Frontiers in Genetics, 5*, 271.

Jang, S. (2000). A Study on the Development of Questionnaire for the Measurement of Stress in Korea. *Yonsei sociology, Yonsei Social Sciences Research Institute, Republic of Korea, 14*, 71-114.

Janssens, A. C. J., & van Duijn, C. M. (2008). Genome-based prediction of common diseases: advances and prospects. *Human molecular genetics, 17*(R2), R166-R173.

Jiang, S. Z., Lu, W., Zong, X. F., Ruan, H. Y., & Liu, Y. (2016). Obesity and hypertension. *Experimental and therapeutic medicine, 12*(4), 2395-2399.

Jin, H.-S., Hong, K.-W., Lim, J.-E., Hwang, S.-Y., Lee, S.-H., Shin, C., . . . Oh, B. (2010). Genetic variations in the sodium balance-regulating genes ENaC, NEDD4L, NDFIP2 and USP2 influence blood pressure and hypertension. *Kidney and Blood Pressure Research, 33*(1), 15-23.

Joe, B., Saad, Y., Lee, N. H., Frank, B. C., Achinike, O. H., Luu, T. V., . . . Yerga-Woolwine, S. (2009). Positional identification of variants of Adamts16 linked to inherited hypertension. *Human molecular genetics, 18*(15), 2825-2838.

Kandiah, J., Yake, M., & Willett, H. (2008). Effects of stress on eating practices among adults. *Family and Consumer Sciences Research Journal, 37*(1), 27-38.

Kassambara, A., Kosinski, M., Biecek, P., & Fabian, S. (2017). Package 'survminer'. *Drawing Survival Curves using 'ggplot2'(R package version 03 1)*.

Khoury, M. J., Janssens, A. C. J., & Ransohoff, D. F. (2013). How can polygenic inheritance be used in population screening for common diseases? *Genetics in Medicine, 15*(6), 437-443.

Kim, D., Shin, H.-J., Kim, S.-W., Hong, J.-M., Lee, K. S., & Lee, S.-H. (2017). Psychological problems of pneumothorax according to resilience, stress, and post-traumatic stress. *Psychiatry Investigation, 14*(6), 795.

Kim, H. C., Lee, H., Lee, H.-H., Seo, E., Kim, E., Han, J., & Kwon, J.-Y. (2022). Korea hypertension fact sheet 2021: analysis of nationwide population-based data with special focus on hypertension in women. *Clinical hypertension, 28*(1), 1-5.

Kim, J. (1999). Making sense of emergence. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 95*(1/2), 3-36.

Kim, J., Noh, J.-W., Park, J., & Kwon, Y. D. (2014). Body mass index and depressive symptoms in older adults: a cross-lagged panel analysis. *PloS one, 9*(12), e114891.

Kim, J. Y., Joo, Y. S., Jhee, J. H., Han, S. H., Yoo, T.-H., Kang, S.-W., & Park, J. T. (2021). Effect of Psychosocial Distress on the Rate of Kidney Function Decline. *Journal of General Internal Medicine, 36*(10), 2966-2974.

Kincaid-Smith, P. (2004). Hypothesis: obesity and the insulin resistance syndrome play a major role in end-stage renal failure attributed to hypertension and labelled 'hypertensive nephrosclerosis'. *Journal of Hypertension, 22*(6), 1051-1055.

Kivimäki, M., Head, J., Ferrie, J. E., Shipley, M. J., Brunner, E., Vahtera, J., & Marmot, M. G. (2006). Work stress, weight gain and weight loss: evidence for bidirectional effects of job strain on body mass index in the Whitehall II study. *International journal of obesity, 30*(6), 982-987.

Koch, F.-S., Sepa, A., & Ludvigsson, J. (2008). Psychological stress and obesity. *The Journal of pediatrics, 153*(6), 839-844. e833.

Kouvonen, A., Kivimäki, M., Cox, S. J., Cox, T., & Vahtera, J. J. P. m. (2005). Relationship between work stress and body mass index among 45,810 female and male employees. *67*(4), 577-583.

Lalouel, J.-M. (2003). Large-scale search for genes predisposing to essential hypertension. *American journal of hypertension, 16*(2), 163.

Larkin, K. T., & Zayfert, C. (2004). Anger expression and essential hypertension: Behavioral response to confrontation. *Journal of psychosomatic research, 56*(1), 113-118.

Lecic-Tosevski, D., Vukovic, O., & Stepanovic, J. (2011). Stress and personality. *Psychiatriki, 22*(4), 290-297.

Lee, C.-Y., & Lee, J.-Y. (1996). Reliability and validity of PWI (psychosocial wellbeing index). *Journal of Preventive Medicine Public Health, 29*(2), 255-264.

Leng, B., Jin, Y., Li, G., Chen, L., & Jin, N. (2015). Socioeconomic status and hypertension: a meta-analysis. *Journal of Hypertension, 33*(2), 221-229.

Levy, D., DeStefano, A. L., Larson, M. G., O'Donnell, C. J., Lifton, R. P., Gavras, H., . . . Myers, R. H. (2000). Evidence for a gene influencing blood

pressure on chromosome 17: genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham Heart Study. *Hypertension, 36*(4), 477-483.
Lundberg, S. M., Erion, G. G., & Lee, S.-I. (2018). Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*.
Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems, 30*.
Maniam, J., & Morris, M. J. (2012). The link between stress and feeding behaviour. *Neuropharmacology, 63*(1), 97-110.
Manichaikul, A., Rich, S. S., Allison, M. A., Guagliardo, N. A., Bayliss, D. A., Carey, R. M., & Barrett, P. Q. (2016). KCNK3 variants are associated with hyperaldosteronism and hypertension. *Hypertension, 68*(2), 356-364.
Mann, A., & Brennan, P. (1987). Type A behaviour score and the incidence of cardiovascular disease: a failure to replicate the claimed associations. *Journal of psychosomatic research, 31*(6), 685-692.
Manuck, S. B., & McCaffery, J. M. (2014). Gene-environment interaction. *Annual review of psychology, 65*, 41-70.
Miguet, M., Olivo, G., Ciuculete, D.-M., Elmståhl, S., Lind, L., & Schiöth, H. B. J. S. J. o. P. H. (2021). Perceived stress is related to lower blood pressure in a Swedish cohort. 14034948211030352.
Mills, K. T., Stefanescu, A., & He, J. (2020). The global epidemiology of hypertension. *Nature Reviews Nephrology, 16*(4), 223-237.
Molinari, E., Bellardita, L., & Compare, A. (2006). Clinical psychology for cardiac disease. In *Clinical psychology and heart disease* (pp. 5-18): Springer.
Moore, C. J., & Cunningham, S. A. (2012). Social position, psychological stress, and obesity: a systematic review. *Journal of the Academy of Nutrition Dietetics, 112*(4), 518-526.
Myrtek, M. (2001). Meta-analyses of prospective studies on coronary heart disease, type A personality, and hostility. *International journal of cardiology, 79*(2-3), 245-251.
Nakagawa, K., Holla, V. R., Wei, Y., Wang, W.-H., Gatica, A., Wei, S., . . . Price, E. (2006). Salt-sensitive hypertension is associated with dysfunctional Cyp4a10 gene and kidney epithelial sodium channel. *The Journal of clinical investigation, 116*(6), 1696-1702.
Nakamura, C., Ishii, A., Matsuo, T., Ishida, R., Yamaguchi, T., Takada, K., . . . Yoshikawa, T. (2020). Neural effects of acute stress on appetite: a magnetoencephalography study. *PloS one, 15*(1), e0228039.
Nosek, M., Kennedy, H. P., Beyene, Y., Taylor, D., Gilliss, C., & Lee, K. (2010). The effects of perceived stress and attitudes toward menopause and aging on symptoms of menopause. *Journal of midwifery & women's health, 55*(4), 328-334.
O'Neill, J., Kamper-DeMarco, K., Chen, X., & Orom, H. J. E. B. (2020). Too stressed to self-regulate? Associations between stress, self-reported executive function, disinhibited eating, and BMI in women. *39*, 101417.
Ober, C., & Vercelli, D. (2011). Gene–environment interactions in human

disease: nuisance or opportunity? *Trends in genetics, 27*(3), 107-115.
Oliva, F., Versino, E., Gammino, L., Colombi, N., Ostacoli, L., Carletto, S., . . . Picci, R. L. (2016). Type D personality and essential hypertension in primary care: A cross-sectional observational study within a cohort of patients visiting general practitioners. *The Journal of nervous and mental disease, 204*(1), 43-48.
Orpen, C. J. J. o. H. S. (1982). Type A personality as a moderator of the effects of role conflict, role ambiguity and role overload on individual strain. *8*(2), 8-14.
Ottman, R. (1996). Gene-environment interaction: definitions and study design. *Preventive medicine, 25*(6), 764-770.
Owen, G. (2013). *Game theory*: Emerald Group Publishing.
Ozaki, K., Ohnishi, Y., Iida, A., Sekine, A., Yamada, R., Tsunoda, T., . . . Nakamura, Y. (2002). Functional SNPs in the lymphotoxin-α gene that are associated with susceptibility to myocardial infarction. *Nature genetics, 32*(4), 650-654.
Pausova, Z., Tremblay, J., & Hamet, P. (1999). Gene-environment interactions in hypertension. *Current hypertension reports, 1*(1), 42-50.
Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics surveys, 3*, 96-146.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research, 12*, 2825-2830.
Pencina, M. J., D'Agostino Sr, R. B., D'Agostino Jr, R. B., & Vasan, R. S. (2008). Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine, 27*(2), 157-172.
Pollock, B. D., Chen, W., Harville, E. W., & Bazzano, L. A. (2017). Associations between hunter type A/B personality and cardiovascular risk factors from adolescence through young adulthood. *International journal of behavioral medicine, 24*(4), 593-601.
Polotsky, H. N., & Polotsky, A. J. (2010). *Metabolic implications of menopause.* Paper presented at the Seminars in reproductive medicine.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Daly, M. J. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics, 81*(3), 559-575.
Rahmouni, K., Correia, M. L., Haynes, W. G., & Mark, A. L. (2005). Obesity-associated hypertension: new insights into mechanisms. *Hypertension, 45*(1), 9-14.
Ridner, S. H. (2004). Psychological distress: concept analysis. *Journal of advanced nursing, 45*(5), 536-545.
Robinson, R. F., Batisky, D. L., Hayes, J. R., Nahata, M. C., & Mahan, J. D. (2005). Significance of heritability in primary and secondary pediatric hypertension. *American journal of hypertension, 18*(7), 917-921.
Romero, L. M., Dickens, M. J., & Cyr, N. E. (2009). The reactive scope model—a new model integrating homeostasis, allostasis, and stress.

Hormones and behavior, 55(3), 375-389.
Rosenman, R. H., Brand, R. J., Sholtz, R. I., & Friedman, M. (1976). Multivariate prediction of coronary heart disease during 8.5 year follow-up in the Western Collaborative Group Study. *The American journal of cardiology, 37*(6), 903-910.
Roth, A. E. (1988). *The Shapley value: essays in honor of Lloyd S. Shapley*: Cambridge University Press.
Sahoo, S., Padhy, S. K., Padhee, B., Singla, N., & Sarkar, S. (2018). Role of personality in cardiovascular diseases: An issue that needs to be focused too! *Indian heart journal, 70*, S471-S477.
Sehgal, M. (2000). Anger, anxiety and type A behaviour as determinants of essential hypertension and coronary heart disease. *Journal of the Indian Academy of Applied Psychology*.
Selye, H. (1957). *Stress*: Ed. Scientifiche Einaudi.
Selye, H. (1976). Stress without distress. In *Psychopathology of human adaptation* (pp. 137-146): Springer.
Seo, H. J., Kim, H. H., & Ha, M. H. (2010). The Influence of Stress on the Change of Health Behavior among Industrial Workers: A 4-Year Follow-up Study. *Study of Stress, 18*(2), 87-97.
Sever, P. S., & Poulter, N. R. (1989). A hypothesis for the pathogenesis of essential hypertension: the initiating factors. *Journal of hypertension. Supplement: official journal of the International Society of Hypertension, 7*(1), S9-12.
Shapley, L. S. (1951). Notes on the n-Person Game—II: The Value of an n-Person Game.(1951).
Sharma, A. (2012). Hypertension: psychological fallout of type A, stress, anxiety and anger. *Advances in Asian Social Science, 3*(4), 751-758.
Shen, C., Zhou, Z., Lai, S., Tao, X., Zhao, D., Dong, W., . . . Gao, J. (2019). Urban-rural-specific trend in prevalence of general and central obesity, and association with hypertension in Chinese adults, aged 18–65 years. *BMC Public Health, 19*(1), 1-8.
Sher, L. (2005). Type D personality: the heart, stress, and cortisol. *Qjm, 98*(5), 323-329.
Shin, H. C. (2013). Measuring stress with questionnaires. *Journal of the Korean Medical Association, 56*(6), 485-495.
Simmons, W. K., Burrows, K., Avery, J. A., Kerr, K. L., Taylor, A., Bodurka, J., . . . Drevets, W. C. (2020). Appetite changes reveal depression subgroups with distinct endocrine, metabolic, and immune states. *Molecular psychiatry, 25*(7), 1457-1468.
Sinha, R., & Jastreboff, A. M. (2013). Stress as a common risk factor for obesity and addiction. *Biological psychiatry, 73*(9), 827-835.
Sjoberg, D. D., Whiting, K., Curry, M., Lavery, J. A., & Larmarange, J. (2021). Reproducible Summary Tables with the gtsummary Package. *R Journal, 13*(1).
Solomon, M. B., Foster, M. T., Bartness, T. J., & Huhman, K. L. (2007). Social defeat and footshock increase body mass and adiposity in male Syrian hamsters. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 292*(1), R283-R290.

Song, J.-J., Ma, Z., Wang, J., Chen, L.-X., & Zhong, J.-C. (2020). Gender differences in hypertension. *Journal of cardiovascular translational research, 13*(1), 47-54.

Song, Y.-Y., & Ying, L. (2015). Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry, 27*(2), 130.

Sparrenberger, F., Cichelero, F., Ascoli, A., Fonseca, F., Weiss, G., Berwanger, O., . . . Fuchs, F. (2009). Does psychosocial stress cause hypertension? A systematic review of observational studies. *Journal of human hypertension, 23*(1), 12-19.

Stone, A. A., & Brownell, K. D. (1994). The stress-eating paradox: multiple daily measurements in adult males and females. *Psychology Health, 9*(6), 425-436.

Suls, J., & Bunde, J. J. P. b. (2005). Anger, anxiety, and depression as risk factors for cardiovascular disease: the problems and implications of overlapping affective dispositions. *131*(2), 260.

Svansdottir, E., Karlsson, H. D., Gudnason, T., Olason, D. T., Thorgilsson, H., Sigtryggsdottir, U., . . . Denollet, J. (2012). Validity of Type D personality in Iceland: association with disease severity and risk markers in cardiac patients. *Journal of Behavioral Medicine, 35*(2), 155-166.

Svensson, T., Kitlinski, M., Engström, G., & Melander, O. (2017). A genetic risk score for CAD, psychological stress, and their interaction as predictors of CAD, fatal MI, non-fatal MI and cardiovascular death. *PloS one, 12*(4), e0176029.

Tabara, Y., Kohara, K., Kita, Y., Hirawa, N., Katsuya, T., Ohkubo, T., . . . Miyata, T. (2010). Common variants in the ATP2B1 gene are associated with susceptibility to hypertension: the Japanese Millennium Genome Project. *Hypertension, 56*(5), 973-980.

Therneau, T., & Lumley, T. (2013). R survival package. *R Core Team*.

Tobin, M. D., Tomaszewski, M., Braund, P. S., Hajat, C., Raleigh, S. M., Palmer, T. M., . . . Samani, N. J. (2008). Common variants in genes underlying monogenic hypertension and hypotension and blood pressure in the general population. *Hypertension, 51*(6), 1658-1664.

Tomiyama, A. J. (2019). Stress and obesity. *Annual review of psychology, 70*, 703-718.

Torres, S. J., & Nowson, C. A. (2007). Relationship between stress, eating behavior, and obesity. *Nutrition, 23*(11-12), 887-894.

Toyoshima, H., Otsuka, R., Hashimoto, S., Tamakoshi, K., & Yatsuya, H. (2014). Body mass index-modified relationship of chronic mental stress with resting blood pressure during 5 years in Japanese middle-aged male workers. *Circ J, 78*(6), 1379-1386. doi:10.1253/circj.cj-13-1086

Troisi, A. (2002). Displacement activities as a behavioral measure of stress in nonhuman primates and human subjects. *Stress, 5*(1), 47-54.

Van Lent, M., Fisher, W., & Mancuso, M. (2004). *An explainable artificial intelligence system for small-unit tactical behavior.* Paper presented at the Proceedings of the national conference on artificial intelligence.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., . . . Bright, J. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods, 17*(3), 261-272.

Vrijkotte, T. G., Van Doornen, L. J., & De Geus, E. J. (2000). Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension, 35*(4), 880-886.

Wardle, J., & Gibson, E. L. (2002). Impact of stress on diet: Processes and implications.

WHO, W. (2019). The Global Burden of Disease. *World Health Organization*.

WHO, W. (2021). Leading causes of death. *World Health Organization*.

Willner, P. (1997). Validity, reliability and utility of the chronic mild stress model of depression: a 10-year review and evaluation. *Psychopharmacology, 134*(4), 319-329.

Winefield, H. R., Gill, T. K., Taylor, A. W., & Pilkington, R. M. (2012). Psychological well-being and psychological distress: is it necessary to measure both? *Psychology of Well-Being: Theory, Research and Practice, 2*(1), 1-14.

Winter, E. (2002). The shapley value. *Handbook of game theory with economic applications, 3*, 2025-2054.

Wu, W.-H., Yang, L., Peng, F.-H., Yao, J., Zou, L.-L., Liu, D., . . . Qu, J.-M. (2013). Lower socioeconomic status is associated with worse outcomes in pulmonary arterial hypertension. *American journal of respiratory and critical care medicine, 187*(3), 303-310.

Yu, B.-n., Luo, C.-H., Wang, D., Wang, A., Li, Z., Zhang, W., . . . Zhou, H.-H. (2004). CYP2C9 allele variants in Chinese hypertension patients and healthy controls. *Clinica Chimica Acta, 348*(1-2), 57-61.

Zhao, F., Song, M., Wang, Y., & Wang, W. (2016). Genetic model. *Journal of cellular and molecular medicine, 20*(4), 765.

국문초록

고혈압의 예측 인자로서
스트레스에 민감한 성격과 유전적 위험도

서울대학교 인문대학원 협동과학 인지과학
인지과학 전공

권 나 연

고혈압은 대표적인 복합질환으로서 95% 이상이 특별한 원인이 없는 본태성으로, 유전적 영향과 생활 습관과 같은 환경적 요인이 복합적으로 작용하여 발생한다. 30세 이상 한국인의 고혈압 유병률은 2019년을 기준으로 약 30%에 달하며, 고혈압이 관상동맥질환, 뇌혈관질환, 심부전, 신장질환 등 많은 합병증을 유발하는 것을 고려하면 사전에 고혈압을 예방하는 것이 특히 중요하다고 할 수 있다. 고혈압의 발병에는 유전적 원인 외에도 식습관이나 운동과 같은 생활 습관 및 사회심리적 스트레스도 영향을 미치는 것으로 알려져 있다. 그렇다면 스트레스를 자주, 쉽게 인지하는 성격적인 특징도 이에 관여할 수 있다고 추정할 수 있으나, 이러한 주요 요인들 간의 인과 관계 및 유전적 요인과의 상호작용에 대해서는 연구가 많이 이루어 지지 않고 있다. 따라서 본 연구에서는 고혈압 발병에 영향을 주는 요인 중, 심리적으로 인지하는 스트레스와, 스트레스에 민감한 성격인 Type A를 중심으로 고혈압의 발병에 대한 영향력을 다각도로 분석하고 또한 고혈압의 예측 인자로서 가치가 있는지를 확인하였다. Type A는 시간에 대한 강박, 경쟁심, 적개심이 높은 성격 유형으로 분류되며, 스트레스에 민감하다고 알려져 있다.

본 연구는 질병청에서 2001년부터 격년으로 수집하고 있는 지역사회 기반 코호트의 임상 자료와 한국인 유전체 역학 조사 사업(KoGES)을 통해 수집된 이들의 유전체 자료를 사용하여, Type A와 스트레스 수준, 유전적 위험도를 통해 고혈압 발병에 대한 영향력을 단면(cross-section) 및 종단(longitudinal)으로 분석하였다.

먼저 단면적 자료를 통해 고혈압의 발병에 대한 Type A와 인지된 스트레스 수준, 그리고 이들을 매개할 것으로 추정한 비만(BMI)의 인과

구조를 구조방정식(structural equation modeling)으로 분석한 결과, 남녀 모두에서 Type A는 고혈압에 간접적인 효과만 관찰되었다. Type A는 인지된 스트레스 수준과 정적인(positive) 관계가 있었으며, 인지된 스트레스는 혈압에 정적으로 크게 영향을 미치는 BMI에 부적인(negative) 영향을 미쳤다. 이로 인해, 스트레스 수준이 높을수록 BMI 수준이 낮고, 혈압도 낮은 경향을 보였다. 이는 스트레스 수준이 높은 그룹이 낮은 그룹에 비해 음식 섭취량이 낮은 것과도 관련이 있는 것으로 보인다.

또한 2001년부터 2016년까지의 데이터로 Type A, 스트레스 수준, 유전적 위험도가 고혈압에 각각 미치는 영향과 이들의 상호작용을 분석한 콕스 (Cox) 회귀 분석에서는 유전적 위험도와 Type A만이 유의미하게 위험 비율을 높였으며, 이들의 상호작용은 관찰되지 않았다.

마지막으로 결정 나무 (Decision Tress) 및 그 확장 알고리즘인 랜덤 포레스트 (Random Forest)와 XGBoost를 통해 공변인과 유전적 위험도를 기본 모델로, Type A와 인지된 스트레스를 추가하며 예측 모델의 성능 개선도를 분석하였다. 개선 여부의 지표는 AUC뿐 아니라 재분류 (Reclassification) 방법인 NRI (Net Reclassification Improvements)와 IDI (Integrated Discrimination Index)를 사용하였으며, XAI의 한 종류인 SHAP value를 통해 feature importance를 계산하였다.

그 결과, 전반적으로 여성 데이터에서의 예측 성능이 좋았으며, Type A를 추가했을 때 성능 개선도가 가장 높았다. 남성 데이터에서는 Type A와 인지된 스트레스를 모두 추가하였을 때 개선이 있었다. SHAP value의 분포에서는 여성 데이터의 경우 Type A의 기여도가 높았으며, 인지된 스트레스는 혼재되어 있어 예측을 방해하는 것으로 나타났다. 남성 데이터는 인지된 스트레스가 Type A보다 기여도가 조금 더 높았다. 결국 생존 분석에서 유의미한 변수였던 유전적 위험도와 Type A가 좋은 성능을 보이는 예측 모델에서 유의미하게 사용되고 있으며, 예측 모델의 성과가 좋지 않은 경우에는 이를 따르지 않은 것을 확인할 수 있었다.

이를 통해, 단면적으로는 Type A 성격이 고혈압에 간접적으로 부정적인 영향을 주지만 장기적으로는 고혈압 발병에 있어 유전적 위험도만큼 이나 유의미한 영향을 미친다는 것을 알 수 있으며, 예측 모델에서도 유의미하게 기여하는 것을 확인하였다.

본 연구를 통해 성격과 스트레스, 비만과 혈압에 대한 인과 구조의

한 부분을 확인할 수 있었으며, 이는 유전적인 수준에서의 예방뿐 아니라 행동 수준에서도 고혈압을 예방할 수 있는 맞춤형 제안의 근거로서 활용될 수 있다. 즉, 본 연구를 통해 복합 질환의 요인들 간의 관계를 이해할 뿐만 아니라, 성격에 따른 개인 맞춤형 의료의 발전에도 기여하는 바가 있을 것으로 사료된다.

Appendix

Type A Score (Jankins Activity Survey)

문항	응답	사용 여부
나는 약속시간에 늦는다든가 또는 일을 느리게 진행되는 것을 참지 못한다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√
나는 줄을 서서 기다리는 것을 싫어 한다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	
사람들은 내가 쉽게 흥분한다고 말한다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√
나는 나의 일과 오락을 경쟁적으로 하려고 한다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√
나는 내가 해야 할 일을 미루고 잠시 쉬고 있을 때 죄의식을 갖는다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√
나는 대화에서 다른 사람들의 말을 가로챈다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√
나는 심한 압력하에 있을 때 쉽게 흥분하고 화를 낸다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	
나는 시간을 정해놓고 강박적으로 일을 한다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√
나는 내가 하고 싶은 일이 다른 사람에 의해 좌우되는 것을 싫어한다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√
나는 현실적으로 그렇게 할 필요가 없을 때에도 나 자신을 몰아세운다	1=전혀 그렇지 않다, 2=좀처럼 그렇지 않다, 3=종종 그렇다, 4=거의 항상 그렇다	√

PWI-SF (Psychosocial Wellbeing Index – Short Form)

문항	응답	사용 여부
현재 매우 편안하며 건강하다고 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
잠자고 난 후에도 개운한 감이 없다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	
매우 피곤하고 지쳐 있어 먹는 것 조차도 힘들다고 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	
근심 걱정 때문에 편안하게 잠을 자지 못한다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	
정신이 맑고 깨끗하다고 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
기력(원기)이 왕성함을 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
밤이면 심란해지거나 불안해 진다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	
대다수의 사람들과 마찬가지로 나를 잘 관리해 나간다고 생각한다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
전체적으로 현재 내가 하고 있는 일은 잘 되어가고 있다고 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
내가 행한 일의 방법이나 절차에 만족한다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
어떤 일에 바로 착수(시작)할 수 있다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
정상적인 일상 생활을 즐길 수 있다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
안절부절 못하거나 성질이 심술궂게 된다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	

나에게 닥친 문제를 해결해 나갈 수 있다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	
불행하고 우울함을 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
나 자신에 대한 신뢰감이 없어지고 있다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	
모든 것을 고려해 볼 때 행복감을 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	√
삶을 살아갈 만한 가치가 있다고 느낀다.	1=항상그렇다, 2=대부분그렇다, 3=약간그렇다, 4=전혀그렇지 않다	