Doctoral Dissertation

Analyzing health status using questionnaires: Assessment of glucose metabolism status and water intake

Tomoki Uchida

Program of Information Science and Engineering
Graduate School of Science and Technology
Nara Institute of Science and Technology

Supervisor: Professor Satoshi Nakamura
Augmented Human Communication Lab.
(Division of Information Science)

Submitted on January 30, 2023

A Doctoral Dissertation
submitted to Graduate School of Science and Technology,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of Engineering

Tomoki Uchida

Thesis Committee:

Professor Satoshi Nakamura	(Supervisor)
Professor Keiichi Yasumoto	(Co-supervisor)
Professor Tomonori Okamura	(Keio University)
Assistant Professor Hiroki Tanaka	(Co-supervisor)

Analyzing health status using questionnaires: Assessment of glucose metabolism status and water intake*

Tomoki Uchida

Abstract

Increasing national medical expenses in a super-aging society is a severe issue. Therefore, everyone should monitor their own health status and prevent disease. Then, a simple method to assess health status is needed. This study addressed the following two challenges.

The first was water intake assessment. Maintaining adequate water intake is essential for physical and mental health. Then, a previous study validated a descriptive dietary record method to assess water intake. However, the challenge with this method was that it took much time to fill out the dietary survey and re-input the written text data into numerical data. Therefore, we aimed to establish a simpler method: the selective recall method. In a clinical trial, participants recorded the food and beverages consumed via a multiple-choice questionnaire. Then, we multiplied the obtained data by the water conversion factor for cooking to calculate water intake. At the same time, we assessed the same water intake by the descriptive dietary record method. As a result, there was a strong correlation between the water intakes by the two methods ($r = 0.94$, $p < 0.0001$). In addition, water intakes by the two methods from non-alcoholic beverages ($r = 0.94$, $p < 0.0001$), alcoholic drinks ($r = 1.00$, $p < 0.0001$), and food ($r = 0.72$, $p < 0.0001$) were also strongly correlated. Therefore, the selective recall method was shown to assess water intake accurately.

The second was identifying glucose metabolism status. Worldwide, 463 million people have diabetes. It is necessary to understand one's glucose metabolism status to take appropriate measures to prevent diabetes. Therefore, we aimed to identify the glucose metabolism statuses using a questionnaire. In a clinical trial, participants underwent an oral glucose tolerance test (OGTT) and completed a lifestyle and physical characteristics questionnaire. In the OGTT, participants intake 75 g glucose solution. Then, blood glucose and insulin levels were measured before and 30, 60, 90, and 120 minutes after glucose intake. We classified them into four glycometabolic categories based on the OGTT results: category 1: best glucose metabolism, category 2: low insulin sensitivity, category 3: low insulin secretion, and category 4: low insulin sensitivity and secretion. We developed machine learning models using questionnaire responses to identify the glycometabolic category. As a result, the AUCs

to classify category 1 and others, 2 and others, 3 and others, and 4 and others were 0.68 (95%CI: 0.62–0.75), 0.66 (0.58–0.73), 0.61 (0.51–0.70), and 0.70 (0.62–0.77). Furthermore, several selected variables were new lifestyle factors related to glucose metabolism status that have not been reported.

Keywords:
health status, questionnaire, random forest, glucose metabolism status, water intake

*Doctoral Dissertation, Graduate School of Science and Technology, Nara Institute of Science and Technology, January 30, 2023.

質問票による健康状態の分析:
糖代謝状態および水分摂取量の評価*

内田 朋希

要約

超高齢社会において国民医療費の増大が深刻な問題となっている。そのため、病気にかかる前に、誰もが自分の健康状態を把握し、病気を予防する必要がある。そこで、簡便な健康状態の評価方法が求められている。本研究では、以下の 2 つの課題に取り組んだ。

1 つ目は、水分摂取量の評価である。適切な水分摂取量を維持することは、心身の健康にとって重要である。そこで、先行研究において、水分摂取量を正確に評価できる descriptive dietary record method が確立された。しかし、この方法の課題は、被験者が食事調査票を手書きで記入し、調査者が記入データを数値データに再入力する必要があるため、多くの労力がかかることであった。そこで本研究では、より簡便な方法である selective recall method の構築と検証を目的とした。本臨床試験において、被験者は摂取した食品および飲料を選択式のアンケートを用いて回答した。そうして得られたデータに調理水分換算係数を乗じることで水分摂取量を算出した。加えて、比較のために descriptive dietary record method を用いて水分摂取量を算出した。その結果、それぞれの手法で算出した水分摂取量には強い相関が認められた（$r = 0.94$、$p < 0.0001$）。また、非アルコール飲料（$r = 0.94$, $p < 0.0001$）、アルコール飲料（$r = 1.00$, $p < 0.0001$）、食品（$r = 0.72$, $p < 0.0001$）由来の水分摂取量においても強い相関があった。したがって、selective recall method により、水分摂取量を由来別に高精度で評価できることが示された。

2 つ目は、糖代謝状態の評価である。現在、全世界の糖尿病患者は 4 億 6,300 万人にのぼる。糖尿病を予防するためには、自分の糖代謝状態を把握し、適切な対策を講じることが重要である。そこで本研究では、アンケートから糖代謝状態を把握することを目的とした。本臨床試験において、被験者は経口ブドウ糖負荷試験（OGTT）を受け、生活習慣と身体的特徴に関するアンケートに回答した。OGTT では、被験者は 75g ブドウ糖溶液を摂取し、摂取前、摂取 30、60、90、120 分後に血糖値およびインスリン値測定を行った。OGTT の結果に基づき、被験者は 4 つの glycometabolic category に分類された。カテゴリー1: 良好な糖代謝状態、2: インスリン感受性低下、3: インスリン分泌不足、4: インスリン

感受性および分泌低下である。次に、アンケート回答結果を用いて glycometabolic category を予測する機械学習モデルを構築した。その結果、カテゴリー1 とその他、2 とその他、3 とその他、4 とその他を分類するための AUC はそれぞれ、0.68（95％信頼区間：0.62-0.75）、0.66（0.58-0.73）、0.61（0.51-0.70）、0.70（0.62-0.77）であった。なお、モデルの変数のいくつかは、糖代謝との関連がこれまで報告されていない新たな生活習慣因子であった。

Keywords:
健康状態、アンケート、ランダムフォレスト、糖代謝状態、水分摂取

*Doctoral Dissertation, Graduate School of Science and Technology, Nara Institute of Science and Technology, January 30, 2023.

Contents

List of figures

List of tables

List of abbreviations

OGTT, oral glucose tolerance test

AUC, area under the receiver operating characteristic curve

WCFC, the water conversion factor for cooking

FPG, fasting plasma glucose level

30 mPG, 30-min post-load plasma glucose level during the OGTT

120 mPG, 120-min post-load plasma glucose level during the OGTT

HbA1c, hemoglobin A1c

BMI, body mass index

RF, random forest

GBM, gradient boosting machine

LR, logistic regression

ANN, artificial neural network

GLMM, generalized linear mixed model

LASSO, least absolute shrinkage and selection operator

EN, elastic net

XGB; XGBoost

DL, deep learning

DT, decision tree

SVM, support vector machine

AQP2, aquaporin-2

Cre, creatinine

ROC curve. the receiver operating characteristic curve

PCA, principal component analysis

1 Introduction

1.1 Background

Japan is a super-aging society, with 29% of its population aged 65 and over [1]. Accordingly, national medical expenses in 2021 were 44 trillion yen and are expected to increase [2]. In the future, it may be challenging to maintain the current health and medical social systems. Traditionally, health conditions were examined at hospitals, and diseases were detected and cured. While from now on, everyone, including those who don't have a disease, should monitor their own health status and prevent disease. Therefore, a method of assessing health status is required, which does not need specialized equipment, skills, or knowledge on the part of the user. Thus, this study addressed the following two issues. The first is water intake assessment. A simple method of assessing water intake helps maintain adequate water intake. It is essential for physical and mental health. The second is identifying glucose metabolism status. It is required to implement appropriate measures to prevent diabetes. Previous studies suggest that both indicators are predictable from dietary and lifestyle surveys. Therefore, we utilized questionnaires for these assessments. The questionnaires are simple and can be answered on the spot. Therefore, they can be helpful for the general public to understand their health status and prevent disease.

1.1.1 Importance of assessing water intake

Water constitutes 60% of the human body and is a vital component of life [3]. Water is essential for maintaining homeostasis; it is a medium for delivering oxygen, nutrients, and hormones throughout the body [4]. It takes part in biochemical reactions inside the cells and contributes to temperature regulation [5]. Hence, adequate water intake is essential for physical and mental health. Inadequate water intake increases the risk of renal and cardiovascular diseases and contributes to metabolic diseases [6-8]. A randomized clinical study examined the effects of increasing water intake by 1.1 L/day for 12 weeks in healthy adults. The study resulted in decreased systolic blood pressure, increased basal body temperature, reduced blood urea nitrogen concentration, suppressed reduction of the glomerular filtration rate, and changes in the intestinal microbiome [9]. In addition, water intake affects mood and cognitive function [10-12]. Therefore, there is a need for an accurate and simple method to assess daily water intake. It helps people to intake an adequate amount of water and maintain their health. Furthermore, a simple validated method can be used for epidemiological studies of the relationship between water intake and health.

1.1.2 Importance of identifying glucose metabolism status

The number of people with diabetes is increasing globally. 463 million people worldwide had diabetes as of 2019, and this is estimated to rise to 700 million by 2045 [13]. Fortunately, however, lifestyle modifications and pharmacological interventions can reduce the risk of developing diabetes [14-18].

Insulin is a hormone produced in pancreatic β-cells that lowers blood glucose levels. When blood glucose levels rise, insulin promotes the uptake of glucose in the blood into the whole-body tissue. As a result, blood glucose levels drop to normal. Therefore, decreased insulin sensitivity and impaired insulin secretion play significant roles in the pathogenesis of diabetes [19, 20]. Currently available data suggest that decreased insulin sensitivity is primarily due to obesity and low muscle mass, whereas impaired insulin secretion is mainly due to aging and genetic factors [21-24]. Hence, it is essential for individuals without diabetes to understand their glucose metabolism status, i.e., insulin sensitivity and insulin secretion. Then, it is necessary to take appropriate measures to prevent diabetes that suit each status.

The oral glucose tolerance test (OGTT) is a standard method for measuring glucose metabolism. It is also used for diagnosing diabetes and pre-diabetes [25]. In this test, a patient is loaded with glucose solution, and multiple blood samples are drawn to measure changes in blood glucose levels. While because of its laboriousness and invasiveness, it is rarely performed on individuals without diabetes. Therefore, there is a need for a method to identify glucose metabolism status more easily than OGTT.

1.2 Previous studies and their problems

1.2.1 Water intake assessment

A descriptive dietary record method is one of the most accurate diet survey methods. It needs subjects to describe all food and beverages consumed and their amounts [26-28]. A 24-hour dietary recall method is also an international standard method. In this method, an investigator asks subjects how much food and beverages they have consumed in the past 24 hours [28, 29]. These methods can accurately calculate water intake from beverages. In comparison, the diet survey data alone cannot calculate water intake from food. A comprehensive database of the water content of various dishes is needed to calculate it. Food and Nutrient Database for Dietary Studies by the U.S. Department of Agriculture provides the nutrient and water content of various foods consumed in the United States [30]. This database and these diet survey methods have been used to assess water intake in the National Health and Nutrition

Examination Survey in the United States [31]. However, the challenge with this method is that it takes time to write down the participants' diet surveys or interviews. In addition, it is necessary to convert the text data of the diet survey into numerical data [26-29].

Cooking involves various processes, such as boiling, baking, frying, and steaming. Thus, there are many variations of dishes in different countries and regions. Therefore, water intake analysis, including food intake, has rarely been reported [32]. Japan has unique dietary habits compared with other countries. Traditional Japanese food consists of cooked rice, soup, and side dishes, including the main dish of meat or fish and other dishes of fresh vegetables or simmered food. Therefore, previously, only one study assessed habitual water intake in Japanese people [33]. However, the challenge of this study was to calculate water intake based on the water content of ingredients rather than dishes. Thus, changes in the water content of dishes due to cooking may have been ignored. Therefore, Murakami et al. [34] established the database of water conversion factor for cooking (WCFC) for Japanese food, which accounts for the effect of cooking on the water content of food. Then, combining this WCFC with the descriptive dietary record method, they developed the method to calculate total water intake from food and beverages [9]. However, the descriptive dietary record method is not well suited to large-scale epidemiological studies because collecting participant responses and interpreting data are time-consuming processes [26, 27].

In this thesis, we aim to develop a simpler web-based method of assessing water intake. No previous study on water intake assessment took such an approach. Alternatively, previous studies on dietary surveys have worked on making questionnaires web-based and simplifying. The web-based dietary surveys did not merely replace existing methods but had various advantages. One of the most significant advantages was automatically calculating nutritional values from input data [35]. In the web-based method by Subar AF, participants reported their meals through searching or browsing for foods in a hierarchical list. Braekman E et al. [36] compared the results of the paper-and-pencil questionnaire and the web-based questionnaire. They found that the web-based questionnaire had fewer errors and missing data due to the alert function. Furthermore, simplified questionnaires were developed to reduce the effort of respondents and researchers. Hotz C et al. [37] generated a list of commonly consumed foods and portion sizes appropriate to their survey region. They narrowed the list of commonly consumed foods by interviewing randomly sampled subjects and key informants knowledgeable about local foods. Portion size distributions were also calculated through interviews and market products survey. Such a list is helpful as choices for the questionnaire. Yamaoka K et al. [38] developed a self-administered semi-quantitative food frequency questionnaire for the nutritional education of patients with diabetes. The

questionnaire choices consisted of 65 foods based on a food substitution table for diabetes patients. Taru C et al. [39] developed a simple dietary questionnaire for estimating energy intake. To assess energy intake particularly accurately, they subdivided the staple food and confectionery options and added an option for sweetened beverages. Sasaki et al. [40, 41] developed a brief-type self-administered diet history questionnaire to assess the nutrient amount habitually consumed. It was designed to assess habitual nutrient intake rather than dietary intake over specific days. Therefore, for example, a question item on the intensity of the dish's flavor was included to assess salt intake. In summary, choice-based questionnaires are less accurate compared to open-ended questionnaires due to their lower information content. Therefore, the questionnaire options should consist of dishes commonly consumed in the region. Using options with a high contribution to the evaluation index is also necessary.

1.2.2 Pre-diabetes screening

Methods of screening for pre-diabetes using laboratory values and questionnaires have been developed. Table 1.1 reviews the important recent studies on pre-diabetes screening tools.

De Silva et al. [42] identified predictors of individuals with high fasting plasma glucose level (FPG), high hemoglobin A1c (HbA1c), or high plasma glucose level during OGTT. Combined use of feature selection and machine learning, including random forests (RF), gradient boosting machine (GBM), logistic regression (LR), and artificial neural network (ANN), selected 25 socio-economic, clinical, and biochemical factors. They used the National Health and Nutrition Examination Survey (NHANES) dataset. However, it may incur effort and cost due to the large amount of input data required for screening.

Birk et al. [43] developed a tool for screening individuals with high FPG using the Global Diet Quality Score and lifestyle questionnaire responses. In this study, RF, generalized linear mixed model (GLMM), least absolute shrinkage and selection operator (LASSO), and elastic net (EN) were used. They showed that dietary factors were important for pre-diabetes screening. However, well-trained interviewers were needed to obtain dietary information in the Global Diet Quality Score.

Abbas et al. [44] reported a risk score for screening individuals with high HbA1c. They used only non-invasively measured factors, including age, sex, BMI, waist circumference, and blood pressure. The algorithms utilized RF, GBM, XGBoost (XGB), LR, and deep learning (DL). Moreover, Dong et al. [45] developed a risk assessment model to detect individuals with high FPG and HbA1c. Eight non-invasively measured risk factors were selected, including age, BMI, waist-to-hip ratio, systolic blood pressure, waist circumference,

sleep duration, smoking, and recreational activity time. The XGB model showed superior performance than the LR model. The study characteristically used indicators of sleep and exercise in addition to clinical factors. However, in these two studies, the screening target did not include individuals with high blood glucose levels after glucose loading. In addition, some of the factors could not be evaluated on the spot and may require laboratory data.

Tian et al. [46] developed a risk score for pre-diabetes and diabetes using questionnaires and blood test results using the LR model. The factors are age, sex, BMI, smoking, FPG, fasting plasma triglyceride level, and history of high FPG. However, research participants were limited to the staff of an oil field in China. Furthermore, invasive measurement factors were required for screening.

Shen et al. [47] analyzed the association between dietary patterns and pre-diabetes risk using the validated semi-quantitative food frequency questionnaire. Multivariate logistic regression analysis showed that the dietary Western pattern score and grains-vegetables pattern score predicted pre-diabetes risk. However, clinical and anthropometric measurements were also needed for adjustment. In addition, well-trained interviewers were required to obtain dietary information. In these previous studies, machine learning models were used more often than ANN, which is less interpretable, to analyze the relationship between those factors and the pathogenesis of pre-diabetes.

Table 1.1 Review of the important recent studies on pre-diabetes screening.

Ref. No.	Screening targets	Factors	Models	Tool challenges
42	FPG 100-125 mg/dL, 120 mPG 100-125 mg/dL, and HbA1c 5.7 6.4%	25 socio-economic, clinical, and biochemical factors	RF, GBM, LR, and ANN	Invasive measurement factors were required for screening.
43	FPG ≥ 100 mg/dL	Global diet quality score, age, smoking, alcohol drinking, unable to walk, use of rations card, and time spent in sedentary activities	RF, GLMM, LASSO, and EN	Well-trained interviewers were needed to obtain dietary information.
44	HbA1c 5.7-6.4%	Age, sex, BMI, waist circumference, and blood pressure	RF, GBM, XGB, LR, and DL	Lack of individuals with high blood glucose levels from screening targets. Some of the factors could not be answered on the spot and may require the linkage of the laboratory data.
45	FPG 110-125 mg/dL and HbA1c 5.7-6.4%	Age, BMI, waist-to-hip ratio, systolic blood pressure, waist circumference, sleep duration, smoking, and vigorous recreational activity time	XGB and LR	Lack of individuals with hyperglycemia after glucose loading from screening targets. Some of the factors could not be answered on the spot and may require the linkage of the laboratory data.
46	FPG ≥ 110 mg/dL and 120 mPG ≥ 140 mg/dL	Age, sex, BMI, smoking, FPG, fasting plasma triglyceride level, and history of high FPG	LR	Research participants were limited to staff in an oil field in China. Invasive measurement factors were required for screening.
47	FPG 100-125 mg/dL, HbA1c 5.7-6.4%, and 120 mPG 140-199 mg/dL	Semi-quantitative food frequency questionnaire answers and clinical and anthropometric measurements scores	LR	Well-trained interviewers were needed to obtain dietary information. Invasive measurement factors were required for screening.

Furthermore, these previous studies had common challenges. As described in Section 1.1.2, it is crucial to understand one's glucose metabolism status, i.e., insulin sensitivity and secretion, to take appropriate measures to prevent diabetes that suit each status. However, pre-diabetes present overlapping pathophysiology. It exhibits both decreased insulin sensitivity and impaired insulin secretion [48, 49]. Therefore, another screening category may be required to clearly distinguish between decreased insulin sensitivity and secretion.

1.3 Approaches in this thesis

1.3.1 A selective recall method for assessing water intake

We aim to validate a selective recall method for assessing water intake from both food and beverages in Japanese people. In a clinical trial, we assess participants' water intake by the selective recall method and an established method, the descriptive dietary record method. The validity of the selective recall method is verified by showing the correlation between the water intakes from the two methods. In addition, we analyze the relationship between water intake and hydration status by urinalysis.

The new method requires less labor for participants than the descriptive dietary record method. In the selective recall method, participants record food and beverages consumed via a multiple-choice questionnaire on Google forms. On the other hand, in the descriptive dietary record method, participants fill out a paper-based questionnaire for all food and beverages consumed. Furthermore, the new method reduces much labor for researchers. The selective recall method automatically calculates water intake by multiplying the ingested food and beverage data by the respective water content [34]. On the other hand, the descriptive dietary record method must require well-trained dietitians to manually enter the paper-based questionnaire results into the calculation software [28].

1.3.2 Identifying glucose metabolism status in non-diabetic Japanese adults using machine learning models with a simple questionnaire

We aim to identify the glucose metabolism status of non-diabetic Japanese adults using machine learning models with a questionnaire. We use clinical trial data of Japanese adult volunteers. In the clinical trial, we perform OGTT on the participants to identify their precise glucose metabolism status. Participants also complete a questionnaire about lifestyle and physical characteristics. Then, we develop machine learning models with questionnaire responses as explanatory variables and glucose metabolism status as objective variables.

Decision tree, support vector machine, random forest, and XGBoost are developed. Their performances are evaluated and compared. In addition, we search for new lifestyle factors related to glucose metabolism status based on the importance of the factors in the developed model.

1.4 Contribution of this thesis

In water intake assessment, this study has two unique contributions:

- This is the first study to validate a selective recall method for assessing Japanese people's water intake from both food and beverages. This method makes assessing water intake more accessible than previous methods.
- Using this method, we report for the first time the accurate habitual water intake of the Japanese. In addition, we report the relationship between water intake and hydration status by urinalysis.

In identifying glucose metabolism status, this study has three unique contributions:

- The model's factors include only lifestyle and physical information that can be answered on the spot. Because invasive measurement or excessive factors are not needed, it can be easily and widely used by the general population.
- Through cross-sectional analysis of glucose metabolism status and lifestyle questionnaire, we discover new lifestyle factors related to glucose metabolism status.
- We identify glucose metabolism status rather than pre-diabetes. Glucose metabolism status of non-diabetic individuals can be classified into four categories based on OGTT results [50]. Each category has clearly different characteristics of insulin sensitivity and secretion: category 1 – best glucose metabolism, category 2 – low insulin sensitivity, category 3 – low insulin secretion, and category 4 – combined characteristics of both categories 2 and 3. Then, in this study, we develop a model to identify these four categories of glucose metabolism status. It can clearly distinguish between decreased insulin sensitivity and secretion.

1.5 Outline of this thesis

In Chapter 2, we develop and validate a selective recall method to assess water intake. In addition, we will analyze the relationship between participants' habitual water intake and

hydration status. In Chapter 3, we identify the glucose metabolism statuses of non-diabetic Japanese adults using machine learning models with a questionnaire. In Chapter 5, we summarize the overall results and discuss existing limitations and future works.

2 A selective recall method for assessing water intake

2.1 Methods

2.1.1 Data collection

In this observational cross-sectional study, participants aged 20 to 64 were recruited from employees working at the Suntory World Research Center (Kyoto, Japan). 32 participants, comprising eight men aged 20 to 44 years (young men group), eight women aged 20 to 44 years (young women group), eight men aged 45 to 64 years (middle-aged men group), and eight women aged 45 to 64 years (middle-aged women group), were included (figure 2.1). The rationales for the sample size are as follows [51]: 21 participants are required to achieve a correlation coefficient of 0.7 at power = 0.85 and $\alpha = 0.01$.

As the survey days, participants were assigned four days within the study period, including three working days and one non-working day. On the day after each survey day, participants were asked to (1) collect first-morning urine, (2) answer the selective recall method questionnaire, (3) fill in the descriptive dietary record method questionnaire, and (4) answer the other questionnaire on daily life and background information. The selective recall method questionnaire was submitted in Google forms, and the participants could not correct it once submitted. The two methods were performed on the same day consecutively to minimize recall bias and compare water intake calculated from each method.

This study was conducted in accordance with the guidelines in the Helsinki Declaration (as revised by the Fortaleza General Meeting of the World Medical Association, Brazil, 2013). All participants provided written informed consent. This study complied with the Ethical Guidelines for Medical Research Involving Human Subjects (2014 Ministry of Education, Culture, Sports, Science and Technology and the Ministry of Health, Government of Japan, Labour and Welfare Ministerial notification No. 3). The ethics committees of Suntory Holdings Limited approved all procedures. The study was registered at the University Hospital Medical Information Network-Clinical Trials Registry (UMIN-CTR).

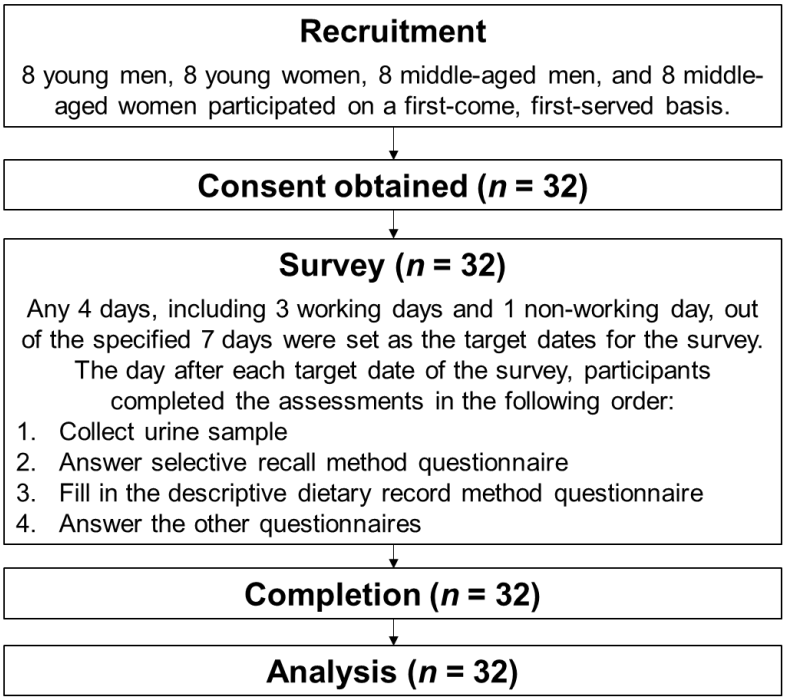

Figure 2.1 Flowchart of the clinical study.

2.1.2 Water intake calculated from a selective recall method

The questionnaire for beverages accounted for the following items: water, tea, coffee, milk beverages, and other beverages. The intake times were recorded as (1) from waking up to breakfast, (2) at breakfast, (3) from breakfast to lunch, (4) at lunch, (5) from lunch to dinner, (6) at dinner, (7) from dinner to 30 min before bedtime, (8) 30 min before bedtime, and (9) from bedtime to waking up. The amount consumed was selected from five options: approximately 100 mL, 200 mL, 350 mL, 500 mL, and more than 700 mL.

For alcoholic beverages, the participants chose beer, whiskey, sake, shochu, wine, chuhai (a fruit-flavored carbonated alcoholic beverage), sour, and other alcoholic beverages. Participants did not provide an answer to this question if they did not drink any alcoholic beverages. The quantity options were set to five easy-to-understand amounts, considering the commercial beverage products and cups.

For food, the participants chose the amount of 40 food categories consumed. These items were based on the frequently consumed food from approximately 2.3 million dishes, reported by ~28,000 people in the Japanese application FoodLog (foo.log Inc., Tokyo, Japan). Intake options were: 0.5, 1, 1.5, and more than 2 servings. The intake times were recorded as for (1)

breakfast, (2) lunch, (3) dinner, and (4) nighttime/others. Table 2.2 shows the 40 food categories, approximate one-person servings, and water contents (mL per one-person serving).

The survey data of non-alcoholic and alcoholic beverages were multiplied by each water content to calculate the water intake. These water contents per 100 g were obtained from the Standard Tables of Food Composition in Japan 2010 (table 2.1).

The food survey data was multiplied by the WCFC to calculate the water intake. Each cooked dish's water content was calculated by multiplying the raw ingredients' water content [52] by a cooking factor. The median water content of the dishes in each food category was used as the water content for the category. Metabolic water was included in the water content of food. It was calculated using a standard method based on the number of carbohydrates (g), proteins (g), and fat (g) in the food [53].

Table 2.1 Water content coefficients for non-alcoholic beverages and alcoholic beverages.

Type	Category	Water content (mL per 100 g)
Non-alcoholic beverages	Tea	99.7
	Water	100.0
	Coffee	93.5
	Milk beverage	86.8
	Other Non-alcoholic beverages	88.8
Alcoholic beverages	Beer	91.2
	Whiskey	90.0
	Rice wine (sake)	83.1
	Shochu	85.2
	Wine	88.7
	Chuhai and sour	81.8
	Other alcoholic beverages	85.4

Table 2.2 Water content coefficients for food categories.

Type	Category [one-person serving]	Water content (mL per one-person serving)
Rice dishes	Rice ball [1 piece]	96.2
	Rice porridge, boiled rice with tea, rice and vegetable porridge, and risotto [1 bowl]	271.9
	Rice boiled with added ingredients [1 bowl]	172.8
	Boiled rice (white rice, unpolished rice, barley rice, etc.) [1 bowl]	113.9
	Fried rice, pilaf, omelette with rice [1 dish]	169.0
	Sushi [1 dish: around 7 pieces]	201.7
	Rice bowl with a topping [1 bowl]	334.0
	Curry and rice, hashed meat and rice [1 dish]	509.5
Bread	Bread [1 slice]	48.4
	Bread roll, croissant [1 piece], French bread [1 slice]	26.4
	Confectionary, dressed bread, hamburger [1 dish], pizza [1 slice]	97.9
Noodle dishes and others	Cup noodles (ramen, udon, soba) [1 cup]	335.1
	Ramen noodles (with broth) [1 bowl]	770.1
	Udon, soba noodles (with broth) [1 bowl]	631.8
	Udon, soba, somen, cold noodles (dipping soup) [1 bowl]	345.9
	Pasta, spaghetti, gratin [1 dish]	258.3
	Yakisoba, yaki udon [1 dish]	295.8
	Food made from flour (okonomiyaki, takoyaki, etc.) [1 dish]	308.8
	Cereal [1 dish]	222.0
Soups	Soup (miso soup, soup, etc.) [1 bowl]	175.7
	Stew [1 dish]	306.0
	Hot pot [1 dish]	367.1
Side dishes to accompany rice	Simmered food (main dish: meat and potato stew, pot-au-feu, etc.) [1 dish]	185.3
	Simmered food (side dish: braised dried daikon, hijiki salad, etc.) [1 dish]	88.7
	Fried food [1 dish]	37.8
	Grilled food (grilled fish, grilled meat, fried egg, etc.) [1 dish]	84.9

	Stir-fried food (stir-fried vegetables, mapo tofu, scrambled egg, etc.) [1 dish]	83.0
	Boiled/steamed food (steamed vegetable, boiled egg, steamed meat dumpling, etc.) [1 dish]	80.4
	Raw food (sashimi, natto, cold tofu, cod roe, etc.) [1 dish]	57.2
	Salads and raw vegetables [1 dish]	104.7
	Vinegared food [1 dish]	87.2
	Boiled greens with soy sauce [1 dish]	69.5
	Pickles [1 dish]	17.0
Fruits, sweets, and snacks	Cheese, dried fruit, and nuts [1 dish]	6.9
	Fruit (mandarin orange, apple, etc.) [1 dish; half an apple]	102.5
	Small fruit (strawberry, grape, etc.) [1 dish; around 5 pieces]	63.5
	Yogurt, set custard pudding, jelly, popsicle [1 piece]	99.5
	Cake, sweet pie, tart [1 piece]	71.0
	Japanese confectionary [1 piece]	40.1

2.1.3 Water intake calculated from a descriptive dietary record method

Participants completed a paper-based descriptive dietary record method questionnaire for food and beverages consumed. Figure 2.2 shows an example of filling out the descriptive dietary record method questionnaire that was distributed to the participants. Dietitians inputted the data from the completed questionnaire into the original nutritional calculation software (foo.log.Inc., Tokyo, Japan). Then, the dietary and beverage intake data was multiplied by each WCFC to calculate the total water intake.

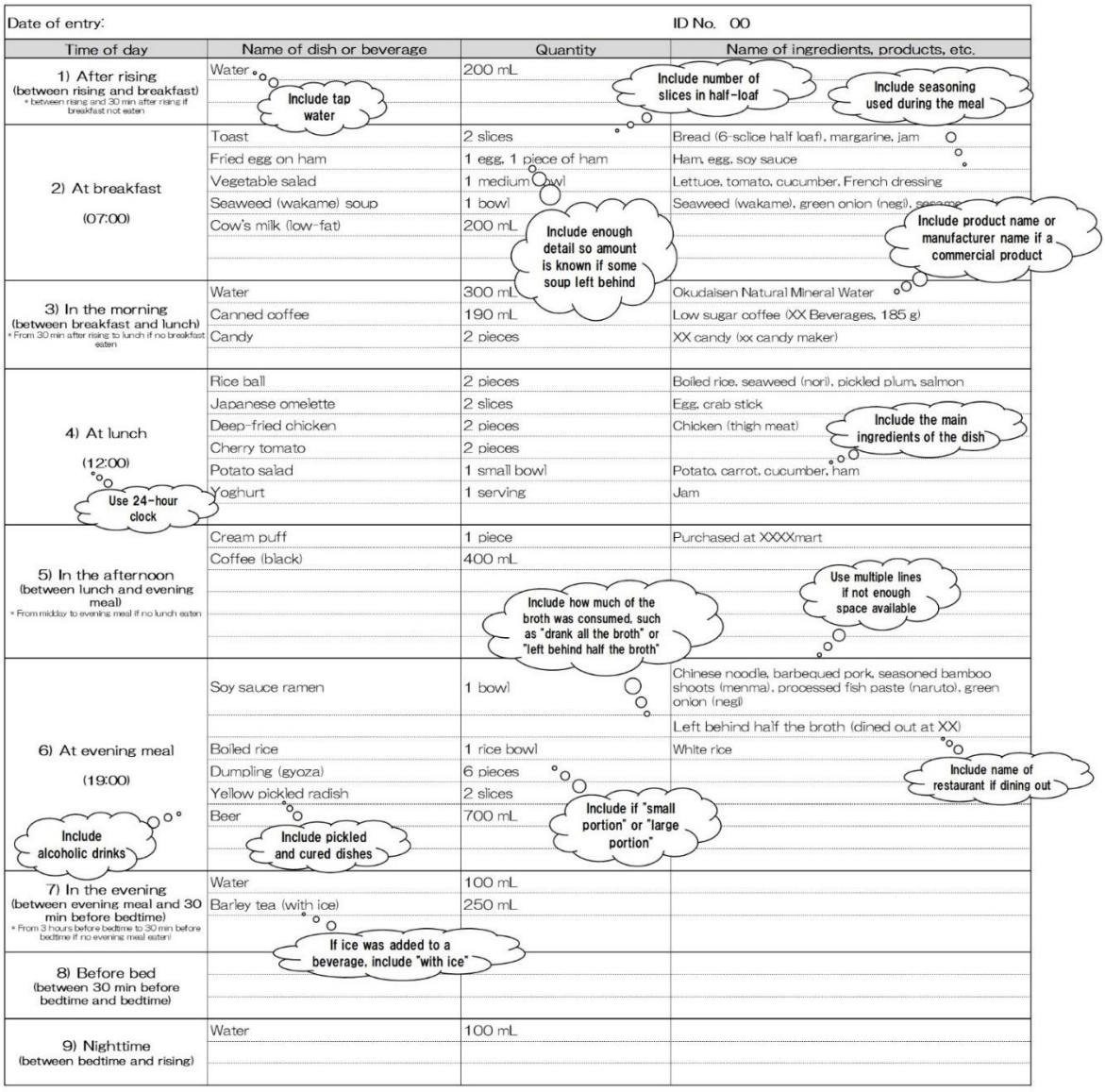

Figure 2.2 Example of filling out the descriptive dietary record method questionnaire.

2.1.4 Other survey items

Participants' first-morning urine was collected on the next day of each survey day. Urine osmolality, the concentration of aquaporin 2 (AQP2), creatinine (Cre), sodium, and potassium were measured. Osmolality was measured using the freezing point depression method. AQP2 was measured using the enzyme-linked immunosorbent assay (Otsuka Pharmaceutical Co., Ltd., Tokyo, Japan). Cre was measured using the enzymatic method.

Sodium and potassium were measured using the electrode method. The measurements were carried out by LSI Medience Co. (Tokyo, Japan). AQP2/Cre, an index of renal water reabsorption, was calculated. 24 h urinary sodium excretion was calculated using the following Tanaka method formula [54].

$$\text{24 h urinary sodium excretion (g/day)} = 21.98\left(\frac{10 \times \text{sodium (mEq/L)}}{\text{creatinine (mg/dL)}} \times \text{predicted creatinine excretion} *\right)^{0.392}$$

$$* \text{ Predicted creatinine excretion (mg/day)} = 14.89(\text{weight (kg)}) + 16.14(\text{height (cm)}) - 2.04(\text{age}) - 2244.45$$

The number of toilet visits on the survey day was assessed via a questionnaire. Participants also self-reported their gender, age, height, and weight on a background information form. BMI was calculated using the formula.

$$\text{BMI} = \text{weight (kg)}/\text{height (m)}^2$$

2.1.5 Statistical Analysis

Correlation analysis was performed using simple linear regression analysis for the water intakes calculated from the selective recall method and the descriptive dietary record method. The mean of each participant's water intake over the four-days was used in the correlation analysis of the two methods. Pearson's correlation test was used to calculate the relationships between the data. Williams's test was applied for the difference between two dependent correlations that share one variable [55]. A p-value < 0.05 was considered to indicate statistical significance. Data aggregation and statistical analysis were performed using Microsoft Excel 2013, JMP ver. 14.0.0 (SAS Institute Inc., Cary, NC, USA), and R ver. 4.0.1 (R Foundation for Statistical Computing, Vienna, Austria).

2.2 Results

2.2.1 Participant characteristics

All participants completed the study. Table 2.3 shows the baseline characteristics and survey results. The mean total water intake of all participants from the selective recall method was 2949 mL/day. The mean total water intake of all participants from the descriptive dietary record method was 2970 mL/day. The mean number of night-time toilet visits was 0.1 times per night. Since the urine samples were collected immediately after waking up, the number of night-time toilet visits may affect the urinalysis results. However, in the result, the mean number of night-time toilet visits during bedtime was 0.1 ± 0.3 times/night. Therefore, its effect on urinalysis parameters was negligible.

Table 2.3 Baseline characteristics and water intake of the participants.

Parameter		All (n = 32)	Young men (n = 8)	Young women (n = 8)	Middle-aged men (n = 8)	Middle-aged women (n = 8)
Baseline characteristic	Age (years)	42.1 (12.6)	33.3 (6.1)	28.1 (2.9)	55.4 (5.5)	51.6 (4.0)
	Height (cm)	167.1 (8.1)	172.4 (3.8)	160.5 (6.1)	174.3 (5.4)	161.3 (5.0)
	Weight (kg)	59.3 (9.8)	63.1 (8.9)	50.0 (4.9)	69.4 (6.3)	54.8 (4.7)
	BMI (kg/m^2)	21.1 (2.2)	21.2 (2.7)	19.4 (1.3)	22.8 (1.4)	21.1 (1.7)
Water intake from a selective recall method	Total (mL/day)	2949 (828)	3474 (1022)	2647 (772)	3102 (824)	2576 (334)
	From beverage (mL/day)	1825 (720)	2413 (936)	1670 (703)	1667 (510)	1549 (360)
	From food (mL/day)	963 (256)	1010 (252)	870 (237)	1006 (240)	968 (312)
	From alcoholic beverage (mL/day)	161 (274)	51 (93)	107 (198)	428 (403)	60 (106)
Water intake from a descriptive dietary record method)	Total (mL/day)	2970 (1059)	3534 (1488)	2513 (798)	3312 (968)	2520 (459)
	From beverage (mL/day)	1775 (912)	2501 (1375)	1581 (667)	1617 (557)	1401 (429)
	From food (mL/day)	1032 (329)	989 (183)	830 (216)	1251 (420)	1060 (305)
	From alcoholic beverage (mL/day)	162 (289)	44 (92)	103 (225)	444 (442)	59 (103)
Urinalysis	Urine osmolality (mOsm/kg)	604 (247)	610 (236)	649 (270)	663 (226)	493 (227)

	Na (mEq/L)	102.4 (54.2)	105.4 (59.7)	110.3 (48.5)	118.3 (56.1)	75.4 (43.6)
	K (mEq/L)	28.9 (16.4)	28.8 (16.7)	27.8 (18.2)	30.4 (13.6)	28.4 (17.3)
	Creatinine (mg/dL)	133.6 (64.9)	153.2 (67.3)	143.7 (74.6)	141.0 (50.0)	96.7 (52.1)
	AQP2 (ng/mL)	5.01 (5.45)	4.24 (4.29)	6.14 (6.82)	6.74 (6.36)	2.92 (2.65)
	AQP2/Cre	0.03 (0.02)	0.02 (0.02)	0.04 (0.02)	0.04 (0.03)	0.03 (0.01)
	Na/K	4.1 (2.1)	4.5 (2.4)	4.9 (2.3)	4.1 (1.8)	2.9 (1.1)
	Na excretion (g/day)	0.129 (0.035)	0.104 (0.028)	0.146 (0.031)	0.118 (0.027)	0.147 (0.033)
Toilet visits	Day time (times/day)	6.6 (1.6)	7.1 (1.5)	6.5 (1.3)	6.5 (1.3)	6.5 (2.0)
	Night time (times/day)	0.1 (0.3)	0.1 (0.3)	0.0 (0.0)	0.2 (0.4)	0.2 (0.4)

Data are presented as mean (SD)

2.2.2 Correlation between water intake calculated from a selective recall method and a descriptive dietary record method

A strong positive correlation was observed between the total water intake calculated from the selective recall method and that calculated from the descriptive dietary record method ($r = 0.94$; $p < 0.0001$; linear regression, $y = 1.20x - 568$; figure 2.3.a). The mean of each participant's water intake over the four-days was used in the correlation analysis of the two methods. In addition, the water intake from non-alcoholic beverages ($r = 0.94$; $p < 0.0001$; linear regression, $y = 1.20x - 407$; figure 2.3.b), alcoholic beverages ($r = 1.00$; $p < 0.0001$; linear regression, $y = 1.05x - 7.66$; figure 2.3.c), and food ($r = 0.72$; $p < 0.0001$; linear regression, $y = 0.90x + 166$; figure 2.3.d) calculated from the selective recall method and the descriptive dietary record method exhibited positive correlations.

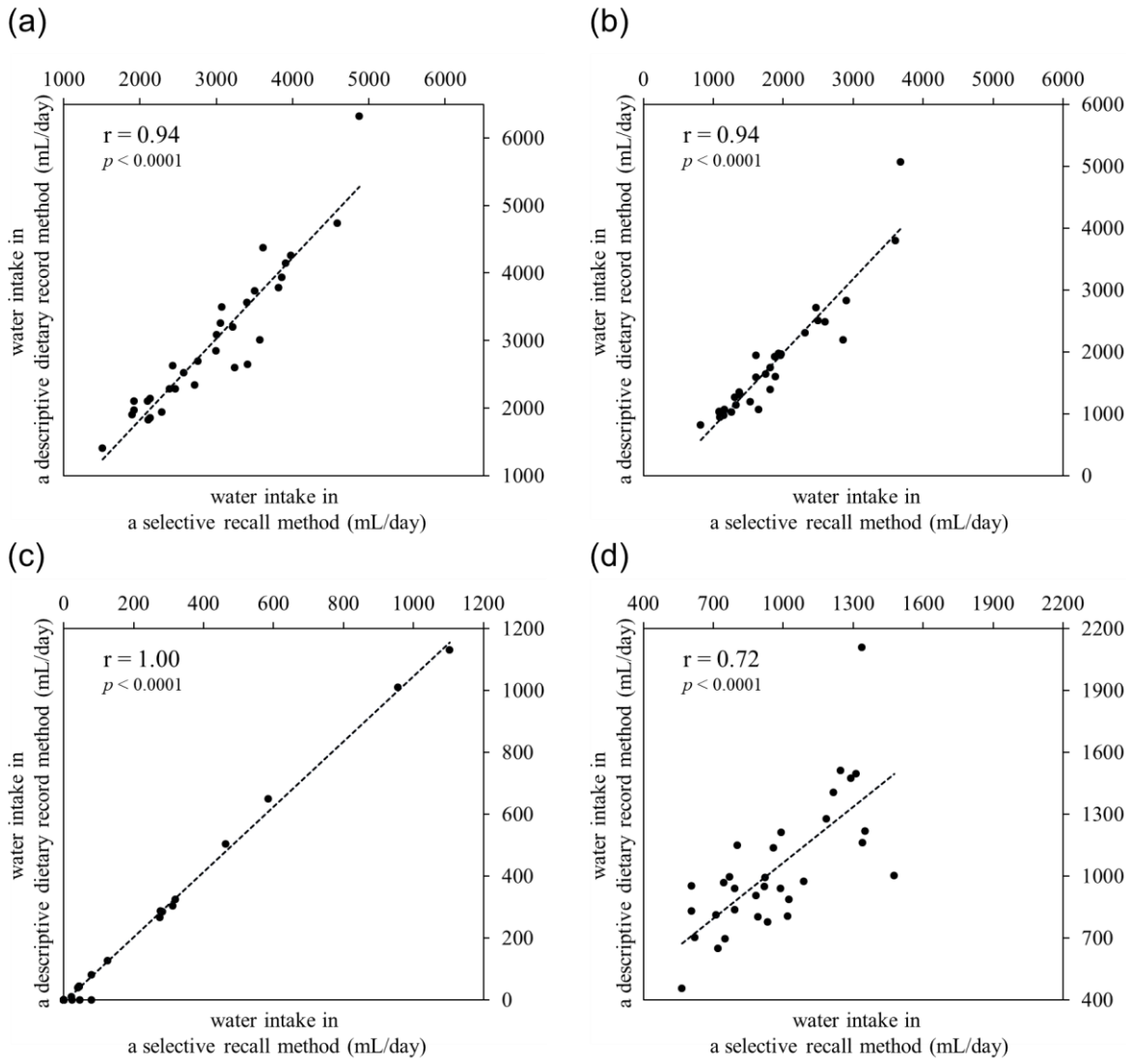

Figure 2.3 Correlation between total water intake calculated from the selective recall method and the descriptive dietary record method. Correlations for (a) total water intake, (b) water intake from non-alcoholic beverages, (c) water intake from alcoholic beverages, (d) and water intake from food. The data were each participant's water intake over the four days.

2.2.3 Relationship between water intake and urine osmolality

Urine osmolality is an established indicator of dehydration status. It increases during dehydration. Therefore, we analyzed the correlation between urine osmolality and water intake. We also examined whether the correlations differed by sex and age. At first, we

confirmed if either method could assess the correlation between water intake and urine osmolality similarly. There was a significant correlation between urine osmolality and water intake calculated from the selective recall method (r = -0.27; p = 0.0018; figure 2.4.a). Similarly, there was a significant correlation between them calculated from the descriptive dietary record method (r = -0.31; p = 0.0004; figure 2.4.b). Williams's test was applied for the difference between two dependent correlations that share one variable. We tested the significance of the difference between the two correlations using Williams's test [55]. There was no significant difference in these two correlation coefficients (p = 0.70).

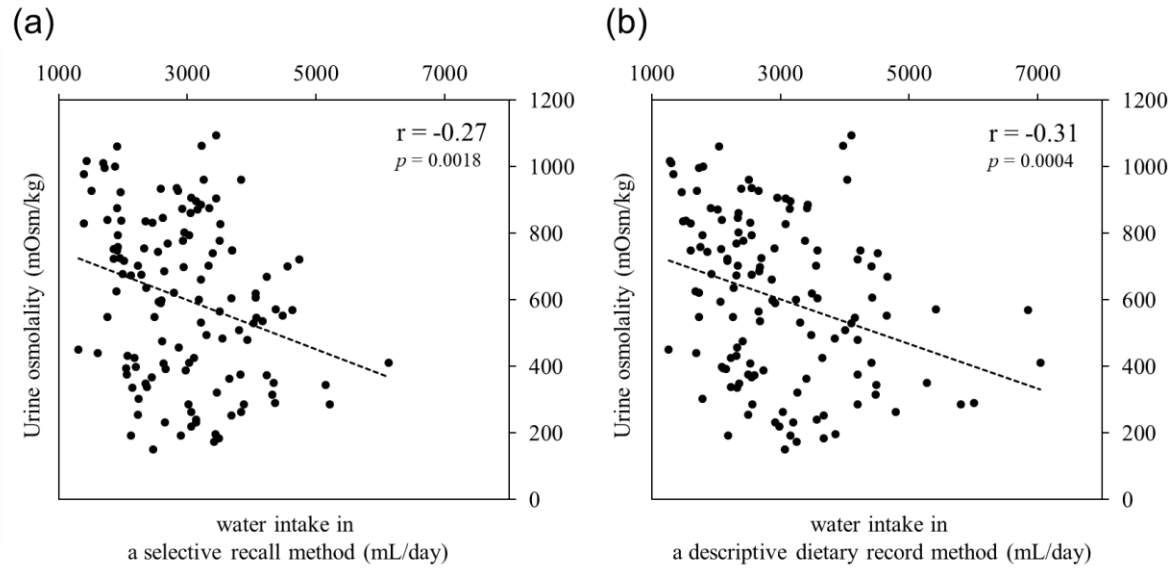

Figure 2.4 Correlation between water intake and urine osmolality for all participants. Figure 2.4.a shows water intake from the selective recall method. Figure 2.4.b shows water intake from the descriptive dietary record method. The points represent daily data for the participants.

We next analyzed the relationship between water intake and urine osmolality by sex and age group. There were significant correlations between water intake and urine osmolality for young (20 to 44 years old) men (figure 2.5.a) and women (figure 2.5.b). However, there was no significant correlation for middle-aged (45 to 64 years old) men (figure 2.5.c) and women (figure 2.5.d).

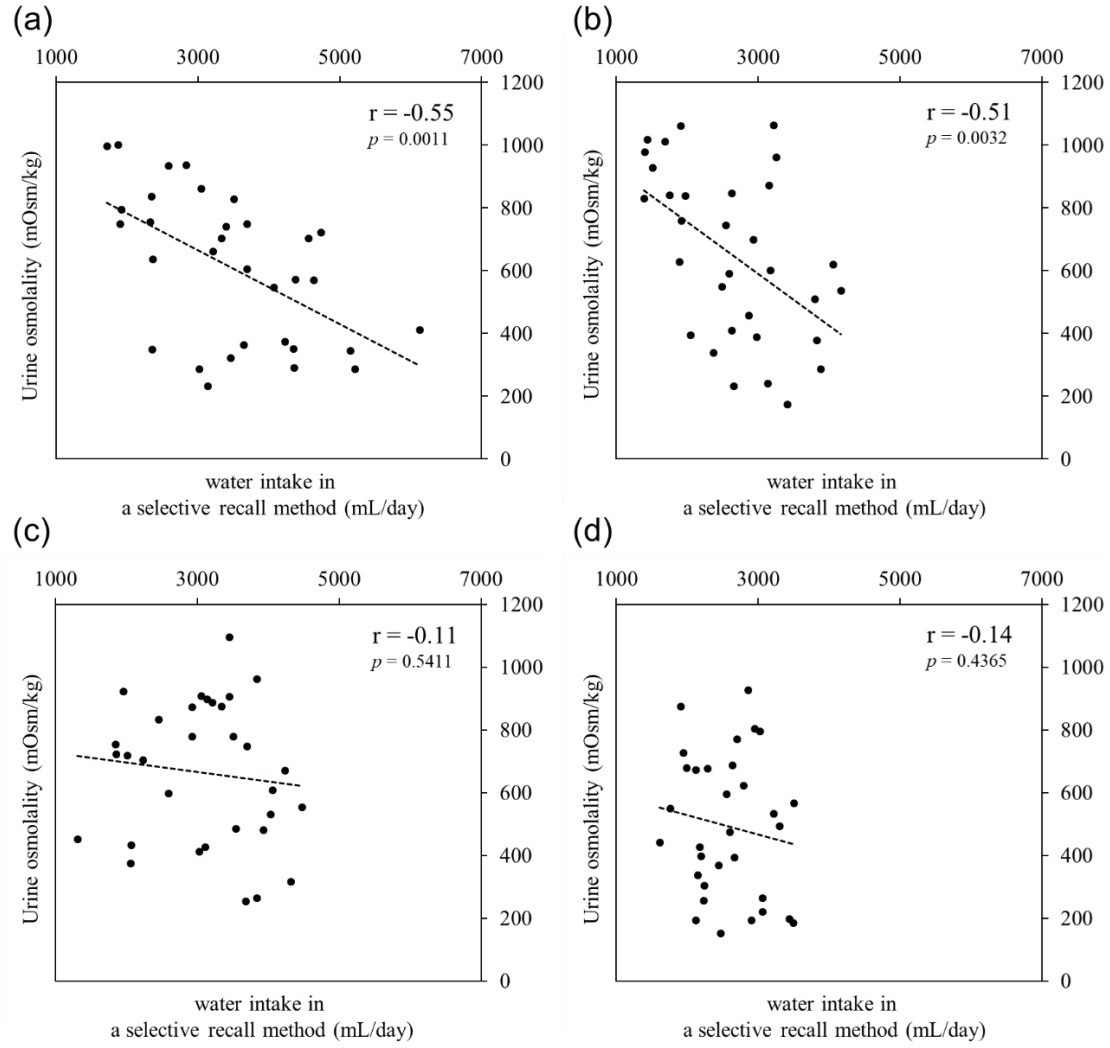

Figure 2.5 Correlation between urine osmolality and water intake for each age and sex group. Correlations for (a) young men (20 to 44 years old); (b) young women (20 to 44 years old); (c) middle-aged men (45 to 60 years old); and (d) middle-aged women (45 to 60 years old). The points represent daily data for the participants.

2.2.4 Relationship between urine osmolality and aquaporin 2/creatinine

The correlation between urine osmolality and AQP2/Cre, an indicator of reabsorption, were examined to determine the body's responses to hydration status. Analysis of all participants revealed a positive correlation between urine osmolality and urine AQP2/Cre ($r = 0.67$; $p < 0.0001$; figure 2.6.a). In addition, a positive correlation was observed in each of the four groups (figure 2.6.b-e).

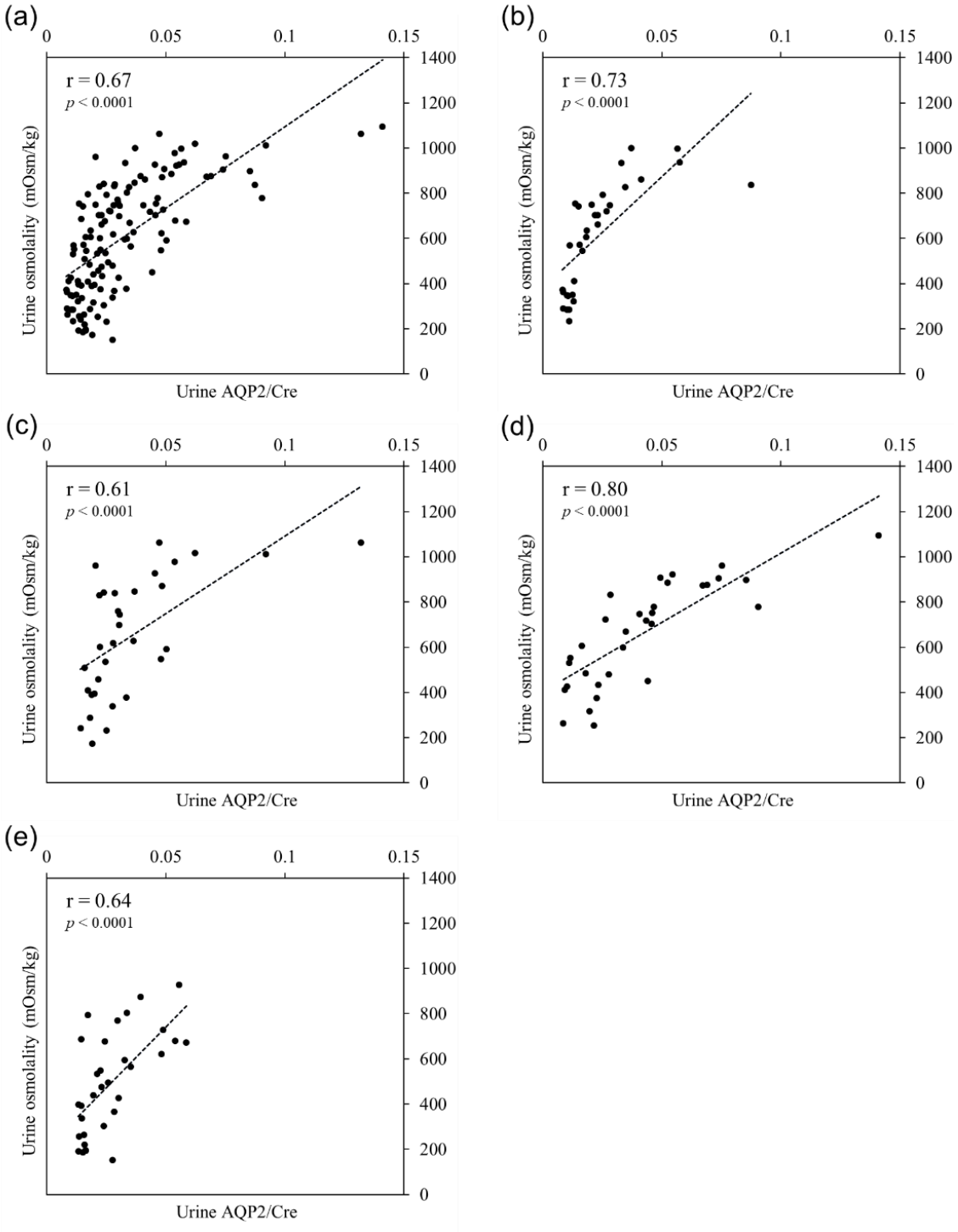

Figure 2.6 Correlation between urine osmolality and AQP2/creatinine. Correlation for (a) all participants; (b) young men; (c) young women; (d) middle-aged men; and (e) middle-aged women. The points represent daily data for the participants.

2.3 Discussion

2.3.1 Validation of a selective recall method for assessing water intake

This study verified the validity of a selective recall method to assess water intake. The total water intakes, calculated from the selective recall method and the descriptive dietary record method, showed a strong correlation. This result demonstrated the validity of the selective recall method. In addition, water intake volumes by the two methods from non-alcoholic beverages, alcoholic beverages, and food were also strongly correlated. Therefore, the selective recall method provides a highly accurate assessment of the water intake for each source. The descriptive dietary record method documents all food and beverages consumed, and it is one of the most precise diet survey methods [26, 27]. Murakami et al. developed a method to rigorously assess water intake by multiplying dietary information collected from the descriptive dietary record method by the WCFC [34]. However, the method is time and labor-intensive because a dietitian manually enters the paper-based description into the nutritional calculation software. The selective recall method saves time and effort because the survey data is collected online and processed directly.

The food-derived water intakes calculated from the selective recall method and the descriptive dietary record method exhibited a positive correlation. However, the correlation coefficient was lower than that of beverages. The reason may be that the selective recall method has 40 food category choices, whereas the descriptive dietary record method can be described all food. Further research exploring more suitable food categories for the choices could increase the correlation.

2.3.2 Habitual water intake of the participants

Since the daily diet is highly variable, validation for the water intake assessment method must be carried out over several days. Diet surveys are generally conducted for three to seven days [55]. This study accounted for variation in the diet by surveying for four days consisting of three working days and one non-working day. We assessed water intake from food, beverages, and alcoholic beverages. The ratio of water intake from beverages to that from food was 2.1:1. The previous Japanese survey reported that the proportion was 1:1. The study suggested that the Japanese ingest more water from food than people in other countries [33]. In comparison, participants in this study consumed more of their water intake from beverages than food. It may be because this survey was conducted in the summer.

2.3.3 Analysis of hydration status by urinalysis

Urine osmolality is an established indicator of hydration status. It increases when dehydrated [56, 57]. Urine osmolality in the normal hydration status was reported to be 818-924 mOsm/kg, with water intake ranging from 2049-2453 mL/day. While in this study, the mean urine osmolality of all participants was 604 mOsm/kg. The mean total water intake of all participants was 2949 mL/day. These results showed that most participants were in average to high hydration status.

Water intake significantly correlated with urine osmolality among young men and women. On the other hand, among the middle-aged groups, no significant correlation was observed between water intake and urine osmolality. Hydration status is controlled by thirst-induced water uptake promotion and renal water conservation [58, 59]. These mechanisms of the participants may have fluctuated with aging. Alternatively, the diuretic effect of alcohol may affect the hydration status. Among the middle-aged men, water intake from alcoholic beverages was 444 ml/day, which was higher than that of other groups.

We, therefore, checked the renal reabsorption function of the participants. Dehydration induces the pituitary gland to secrete vasopressin. It triggers AQP2 translocation to the plasma membrane on the luminal surface of the kidney collecting duct. Water is reabsorbed via AQP2 in a kidney, thereby regulating the body fluid volume. The quantity of AQP2 in urine is correlated with the amount of AQP2 translocated to the renal plasma membrane. Therefore, urine AQP2 is an indicator of renal resorption [60-62]. In addition, AQP2/Cre is the urine AQP2 concentration divided by the urine creatinine concentration, independent of the urine concentration [63]. This study found positive correlations between urine osmolality and urine AQP2/Cre among young and middle-aged groups. This result was reasonable, indicating that the renal water reabsorption mechanism was enhanced in response to dehydration [61]. Furthermore, these results showed that the mechanism worked normally in all four groups.

3 Identifying glucose metabolism status using machine learning models

3.1 Methods

3.1.1 Data collection

In this cross-sectional study, we recruited non-diabetic Japanese adults aged 20 to 64 in Tokyo. Those with cardiovascular, liver, and kidney disorders, those taking medication, pregnant women, and lactating women were excluded. Diabetes was defined as a fasting plasma glucose level ≥ 126 mg/dL, 120-min post-load plasma glucose level during the OGTT (120 mPG) ≥ 200 mg/dL, and the use of anti-diabetic medications [25]. Participants underwent height and weight measurements and OGTT. In the OGTT, participants were loaded with 75 g glucose solution. Then, blood sampling was performed before and 30, 60, 90, and 120 minutes after glucose loading. The sampled blood was used for the measurement of blood glucose levels and blood insulin levels. Participants also completed a questionnaire on lifestyle and physical characteristics. Participants with a survey response rate of less than 90% were excluded from the analysis. A total of 977 participants were suitable for the study.

For external verification data, we recruited non-diabetic Japanese adults aged 20 to 64 in Hokkaido. The selection and exclusion criteria were the same. The same examinations and a questionnaire were conducted on them. A total of 452 participants were suitable for the study.

These two studies were conducted in accordance with the guidelines in the Helsinki Declaration. All participants provided written informed consent. These two studies complied with the Ethical Guidelines for Medical Research Involving Human Subjects. The ethics committees of Nihonbashi Egawa Clinic or Fukuhara Clinic approved all procedures. The studies were registered at the UMIN-CTR.

3.1.2 Classification of glycometabolic category

We classified the participants into the four glycometabolic categories based on 30 min post-load plasma glucose level (30 mPG), 120 mPG, and the Matsuda index during the OGTT. These values are associated with insulin sensitivity and secretion [48, 64, 65]. The Matsuda index was calculated as follows.

$$\text{Matsuda index} = \frac{10{,}000}{\sqrt{(\text{fasting glucose})(\text{fasting insulin})(\text{mean glucose in OGTT})(\text{mean insulin in OGTT})}}$$

The classification criteria were as follows: condition A – 30 mPG <157 mg/dL and condition B – 120 mPG <126 mg/dL and Matsuda index > 4.97, category 1 satisfies conditions A and B, category 2 satisfies condition A but not condition B, category 3 satisfies condition B but not condition A, and category 4 satisfies neither condition A nor condition B. The four categories were the objective variables of the models in this study.

	120 mPG < 126 mg/dL and Matsuda index > 4.97	120 mPG ≥ 126 mg/dL or Matsuda index ≤ 4.97
30 mPG < 157 mg/dL	**Category 1** best glucose metabolism	**Category 2** low insulin sensitivity
30 mPG ≥ 157 mg/dL	**Category 3** low insulin secretion	**Category 4** low insulin sensitivity and low insulin secretion

Figure 3.1 Classification and characteristics of glycometabolic category.

Each category has clearly different insulin sensitivity and secretion characteristics: category 1 – best glucose metabolism, category 2 – low insulin sensitivity, category 3 – low insulin secretion, and category 4 – combined characteristics of both categories 2 and 3. In this study, we developed models to identify these four categories of glucose metabolism status. The rationale for the categorization and characteristics of each category were explained in a previous study [50].

3.1.3 Model development

For the explanatory variables, we obtained a dataset that included age, sex, height, BMI, and the questionnaire responses. The questionnaire consisted of 309 questions that did not require clinical examination data and could be answered on the spot (appendix A). The topics of the

questions included exercise, sleep, dietary habits, water intake, alcohol intake, physical condition, constitution, family history, workstyle, and lifestyle. As pretreatments for the analysis, we replaced missing answers with the mode. Questions with answers in the nominal variable were split, and each answer was converted to a dummy variable. Then, we evaluated the correlations between all variables to identify multicollinearity. If two variables had a Spearman's correlation coefficient greater than 0.7, one was excluded. Therefore, we obtained 283 explanatory variables. Before inputting the data into the SVM, training and testing datasets were standardized (mean of 0 and variance of 1), respectively.

We developed four-class classification models to identify the glycometabolic category. We considered decision tree, support vector machine, random forest, and XGBoost. We used these models rather than deep learning models to interpret the variables' importance and develop a simpler model by narrowing down the variables. We used the rpart package of R ver. 4.1.0 for the decision tree. The tuned hyperparameters were the minimum number of observations in a node and the maximum depth of trees. We used the randomForest package of R ver. 4.1.0 for the random forest. The tuned hyperparameters were the number of variables randomly sampled at each tree, the minimum size of terminal nodes, and the number of trees to grow. We used the kernlab package of R ver. 4.1.0 for the support vector machine. The tuned hyperparameter was the cost of constraints violation. The kernel function was set to the linear kernel. We used the xgboost package of R ver. 4.1.0 for XGBoost. The tuned hyperparameters were the subsample ratio to all variables at each tree, maximum depth of trees, and learning rate. Figure 3.2 shows the models' training, testing, and validation processes. We randomly split the original dataset into training (70%) and testing (30%) datasets. Undersampling of the training dataset was performed because four categories of the dataset were imbalanced. We conducted 5-fold cross-validation to find the optimal hyperparameters using the training dataset. Then, we trained the models with the optimal hyperparameters using the training dataset. Finally, we assessed the model performances using the testing dataset.

In addition, the top ten most important variables in the random forest model were selected. The importance of each variable was assessed by the mean decrease in the Gini coefficient. It is the mean of the total decrease in node impurity by a variable, weighted by the proportion of samples reaching that node in each decision tree in the random forest. Then, we trained another random forest model using only these ten variables as explanatory variables using the training dataset. Finally, we assessed the model performance using the testing dataset and verified using the external verification dataset.

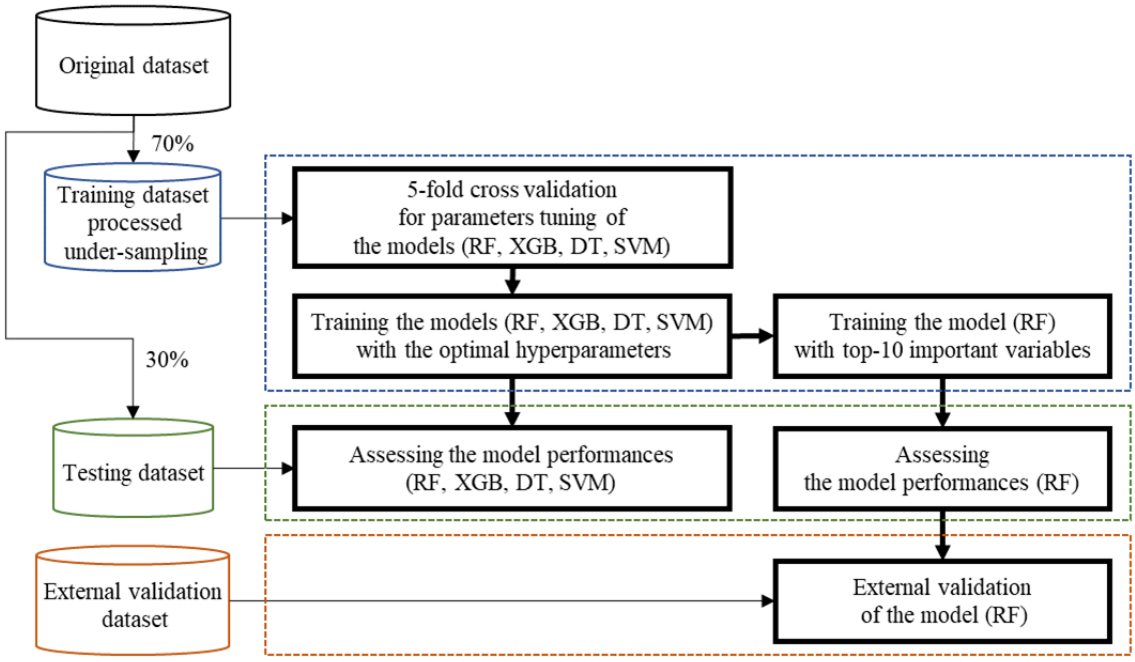

Figure 3.2 Training, testing, and validation processes of the models.

The performances were considered based on the areas under the receiver operating characteristic curve (AUCs) for classifying category 1 and others, category 2 and others, category 3 and others, category 4 and others, and the mean of these AUCs. The 95% confidence interval of the AUCs was computed with 2,000 stratified bootstrap replicates. We used Delong's method to calculate *p*-values to compare the AUCs [66].

The characteristics of each glycometabolic category and the OGTT values were compared using analysis of variance (ANOVA) with Dunnett's test for multiple comparisons [67]. Spearman's correlation test calculated the relationships between the selected ten variables and 30 mPG, 120 mPG, and the Matsuda index. A p-value < 0.05 was considered to indicate statistical significance. Statistical analysis was performed using the statistical software package R ver. 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).

3.1.4 Analysis of important factors in each category

We developed four random forest models of binary classification: a model classifying category 1 and others, a model classifying category 2 and others, a model classifying category 3 and others, and a model classifying category 4 and others. Then, we consider the variables

with high importance in identifying each category. Figure 3.3 shows the training and testing processes of the models. First, we randomly split the original dataset into training (70%) and testing (30%) datasets. Undersampling of the training dataset was performed. Then, we conducted 5-fold cross-validation to find the optimal hyperparameters using the training dataset. Then, we trained the models with the optimal hyperparameters using the training dataset. Finally, we assessed the model performances using the testing dataset.

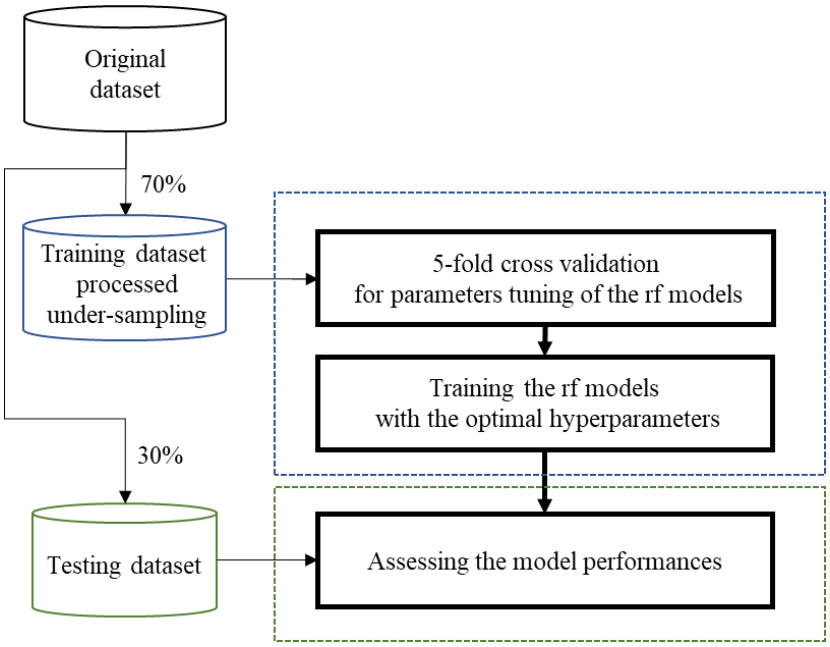

Figure 3.3 Training and testing processes of the models.

3.1.5 Regression analysis of glucose metabolism status

The glucose metabolism status is originally a continuous status rather than four discrete categories. Therefore, we developed models to predict the continuous values of glucose metabolism status. We predicted 30 mPG, 120 mPG, and the Matsuda index, which are the classification criteria of the glycometabolic category. Figure 3.4 shows the training and testing processes of the multiple regression analysis. We randomly split the original dataset into training (70%) and testing (30%) datasets. Using the training dataset, we calculated multiple regression equations for 30 mPG, 120 mPG, and Matsuda index. The variables were selected with backward selection. In the method, first, we developed a multiple linear regression model with all the original dataset variables. Then, we removed the variable with

the highest p-value from the model and fit a new model. We repeated this process until all variables in the model had a *p*-value below 0.05. Then, we assessed the prediction performance of the multiple regression equations using the testing dataset. The performances were evaluated by the correlation between predicted and reference 30 mPG, 120 mPG, and Matsuda index.

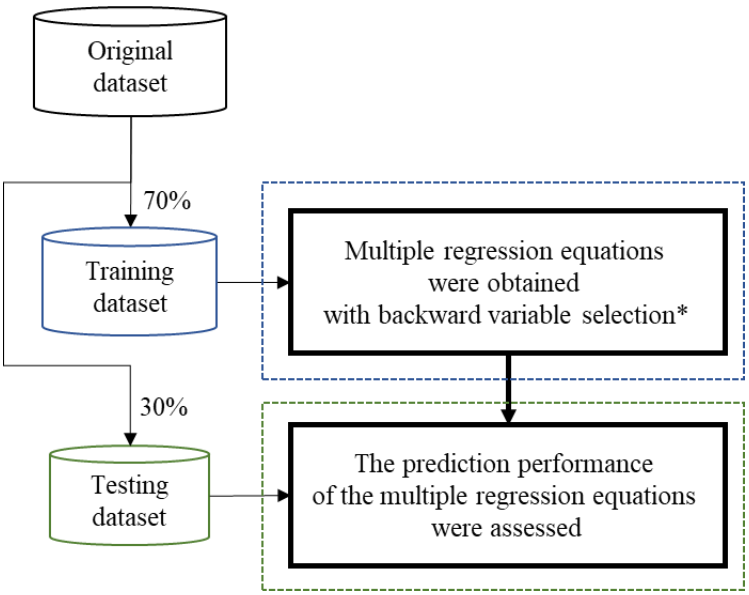

Figure 3.4 Training and testing processes of the multiple regression analysis.

3.2 Results

3.2.1 Participant characteristics

Of the total of 977 eligible participants in the original dataset, the glycometabolic categories 1, 2, 3, and 4 accounted for 46% (*n* = 448), 21% (*n* = 206), 14% (*n* = 133), and 19% (*n* = 190), respectively (table 3.1). Regarding age, categories 3 and 4 were significantly higher than category 1. Regarding the BMI, categories 2 and 4 were significantly higher than category 1. The questionnaire answers were obtained from 977 participants. None of the subjects had more than 1% of missing answers. The characteristics of the preprocessed questionnaire answers for each category of the participants are shown in appendix B.

Table 3.1. Characteristics of the participants in each glycometabolic category.

Parameter	Category 1	Category 2	Category 3	Category 4
n	448	206	133	190
Sex (% women)	53.1	56.3	40.6	44.2
Age (years)	42.3 (41.2–43.4)	43.8 (42.3–45.3)	46.7 (44.7–48.6)*	48.9 (47.4–50.4)*
Height (cm)	164.8 (164.1–165.6)	164.8 (163.6–166.0)	165.6 (164.2–166.9)	166.0 (164.9–167.1)
BMI (kg/m^2)	21.4 (21.1–21.6)	23.5 (23.1–24.0)*	21.6 (21.2–22.0)	23.4 (22.9–23.8)*
30 mPG (mg/dL)	129.5 (127.9–131.1)	139.5 (137.7–141.2)*	171.9 (169.7–174.2)*	178.1 (175.5–180.8)*
120 mPG (mg/dL)	94.4 (92.9–96.0)	127.6 (124.4–130.8)*	99.3 (96.3–102.3)	134.6 (130.5–138.6)*
Matsuda index	9.8 (9.4–10.2)	5.8 (5.4–6.3)*	7.7 (7.3–8.1)*	5.0 (4.6–5.4)*

Data are presented as mean (95% confidence interval), percentage, or the number of individuals. *$p < 0.05$ vs. category 1.

3.2.2 Model performances

Table 3.2 shows the performances of the models for identifying glycometabolic category. The random forest model had the highest performance among the models. Its AUCs (95% confidence intervals) to classify category 1 and others, category 2 and others, category 3 and others, and category 4 and others were 0.69 (0.63–0.75), 0.68 (0.61–0.75), 0.63 (0.55-0.72), 0.67 (0.59-0.74). There was no statistically significant difference from the AUCs of the other models.

Table 3.2 Performances of the models for identifying glycometabolic category (95% confidence intervals).

Model	AUC to classify category 1 and others	AUC to classify category 2 and others	AUC to classify category 3 and others	AUC to classify category 4 and others	Mean of AUCs
Decision tree	0.63 (0.58-0.70)	0.68 (0.60-0.75)	0.56 (0.45-0.66)	0.61 (0.53-0.70)	0.62
Support vector machine	0.64 (0.57-0.70)	0.65 (0.57-0.73)	0.58 (0.47-0.68)	0.55 (0.48-0.64)	0.61
Random forest	0.69 (0.63-0.74)	0.68 (0.61-0.76)	0.63 (0.55-0.72)	0.67 (0.59-0.74)	0.67
XGBoost	0.62 (0.56-0.68)	0.58 (0.50-0.66)	0.59 (0.49-0.69)	0.60 (0.52-0.68)	0.60

In the random forest model, the top ten most important variables were age, height, BMI, and the following questions: “Do you wake up in the middle of the night”, “Which do you usually eat: rice or bread”, “Frequency of tea intake per week at lunch”, “Do you wake up late on a non-working day”, “Frequency of mobile phone and tablet computer use at bedtime”, “Frequency of soup intake”, and “Frequency of toothbrush replacement”.

3.2.3 Model performance using ten variables

We next developed another random forest model using these ten variables. Table 3.3 shows the performance of the model. Its AUCs (95% confidence intervals) to classify category 1 and others, category 2 and others, category 3 and others, and category 4 and others were 0.68 (0.62–0.75), 0.66 (0.58–0.73), 0.61 (0.51–0.70), and 0.70 (0.62–0.77), respectively. Its AUC to classify category 1 and others was not significantly lower than that of the previous random forest model shown in table 3.2 (p-value was 0.86). Its AUCs to classify category 2 and others, category 3 and others, and category 4 and others were also not significantly lower than those of the previous random forest model (p-values were 0.33, 0.51, 0.11).

Table 3.3 Performance of the random forest model using the ten variables (95% confidence interval).

Model	AUC to classify category 1 and others	AUC to classify category 2 and others	AUC to classify category 3 and others	AUC to classify category 4 and others	Mean of AUCs
Random forest using ten variables	0.68 (0.62–0.75)	0.66 (0.58–0.73)	0.61 (0.51–0.70)	0.70 (0.62–0.77)	0.66

Figure 3.5 shows the model's receiver operating characteristic (ROC) curves. The AUC for classifying category 4 and others was the highest among the AUCs for classifying each category.

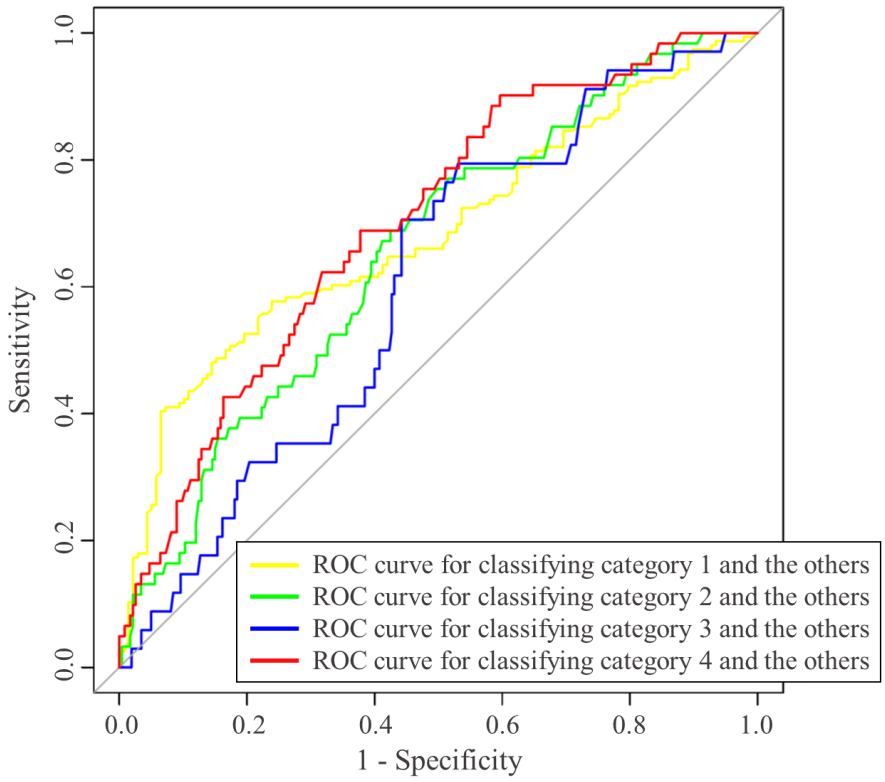

Figure 3.5 ROC curves of the random forest model using the ten variables.

Table 3.4 shows the importance of the ten variables in the model. Table 3.5 shows correlation coefficients between the variables and 30 mPG, 120 mPG, and the Matsuda index, which are the classification criteria of the glycometabolic category. We partially modified the variables' wording to make the correlations' positives and negatives easier to understand. There were significant correlations between most variables and classification criteria. The variable "Frequency of eating bread" did not significantly correlate to 120 mPG, but its *p*-value was less than 0.1.

Table 3.4 Variables of the model and their importance.

Question type	Variable	Mean decrease in Gini coefficient
Physical characteristics	BMI	10.3
Physical characteristics	Age	8.1
Physical characteristics	Height	3.3
Sleep habits and drowsiness	Do you wake up in the middle of the night?	3.1
Dietary habits	Which do you usually eat: rice or bread?	2.5
Water intake	Frequency of tea intake per week at lunch	2.1
Sleep habits and drowsiness	Do you wake up late on a non-working day?	1.9
Workstyle and lifestyle	Frequency of mobile phone and tablet computer use at bedtime	1.4
Dietary habits	Frequency of soup intake	1.4
Workstyle and lifestyle	Frequency of toothbrush replacement	0.8

Table 3.5 Correlation coefficients between the variables and the classification criteria of the glycometabolic category.

Variable	Correlation coefficient		
	30 mPG	120 mPG	Matsuda index
Body mass index	0.14*	0.17*	-0.50*
Age	0.26*	0.22*	-0.05#
Height	0.10*	-0.07*	-0.07*
Frequency of waking up in the middle of the night	0.11*	0.09*	-0.04
Frequency of eating bread, not rice	<0.00	0.06#	-0.04
Frequency of tea intake per week at lunch	<0.00	0.06#	-0.07*
Frequency of waking up late on a non-working day	-0.07*	-0.06#	-0.05
Frequency of mobile phone and tablet computer use at bedtime	-0.02	<0.00	-0.06*
Frequency of soup intake	0.06*	-0.03	0.09*
Frequency of toothbrush replacement	-0.07*	-0.02	-0.01

*$p < 0.05$ and $0.5 \leq {}^{\#}p < 0.1$ in Spearman's correlation test.

3.2.4 Model performance in the external validation

Of the total of 452 eligible participants in the external validation dataset, the glycometabolic categories 1, 2, 3, and 4 accounted for 47% (n = 213), 30% (n = 135), 7% (n = 32), and 16% (n = 72), respectively (table 3.6). The characteristics of the questionnaire answers used for the model are shown in appendix C.

Table 3.6 Characteristics of the participants in each glycometabolic category.

Parameters	Category 1	Category 2	Category 3	Category 4
n	213	135	32	72
Sex (% women)	0.48	0.49	0.38	0.47
Age (years)	38.6 (37.0–40.2)	43.7 (42.0–45.5)*	44.3 (40.6–47.9)*	46.4 (44.2–48.6)*
Height (cm)	165.9 (164.7–167.1)	165.4 (163.9–166.9)	167.2 (164.4–170.1)	165.2 (163.3–167.0)
BMI (kg/m^2)	22.0 (21.7–22.4)	23.8 (23.2–24.4)*	22.0 (21.0–23.1)	23.9 (23.2–24.7)*
30 mPG (mg/dL)	124.3 (121.9–126.7)	138.1 (135.8–140.4)*	166.4 (163.5–169.4)*	178.3 (173.6–182.9)*
120 mPG (mg/dL)	100.6 (98.5–102.8)	142.0 (137.5–146.5)*	100.4 (95.0–105.8)	156.5 (148.9–164.2)*
Matsuda index	10.7 (10.2–11.3)	7.5 (6.7–8.2)*	8.5 (7.1–9.8)*	5.4 (4.7–6.2)*

Data are presented as mean (95% confidence interval), percentage, or the number of individuals. *$p < 0.05$ vs. category 1.

The performance of the previous random forest model using ten variables was verified using the external validation dataset. Its AUCs (95% confidence intervals) to classify category 1 and others, category 2 and others, category 3 and others, and category 4 and others were 0.66 (0.61–0.71), 0.57 (0.51–0.62), 0.60 (0.50–0.69), 0.64 (0.57–0.71), respectively (table 3.7). Its AUC to classify category 1 and others was not significantly lower than that of the previous random forest model shown in table 3.3 (*p*-value was 0.63). Its AUCs to classify category 2 and others, category 3 and others, and category 4 and others were also not significantly lower than those of the previous random forest model (*p*-values were 0.06, 0.26, 0.89). Figure 3.6 shows the ROC curves of the model. The AUC for classifying category 1 and others was the highest among the AUCs for classifying each category.

Table 3.7 Performance of the random forest model using the ten variables in the external validation (95% confidence interval).

Model	AUC to classify category 1 and the others	AUC to classify category 2 and the others	AUC to classify category 3 and the others	AUC to classify category 4 and the others	Mean of AUCs
Random forest using ten variables	0.66 (0.61–0.71)	0.57 (0.51–0.62)	0.60 (0.50–0.69)	0.64 (0.57–0.71)	0.62

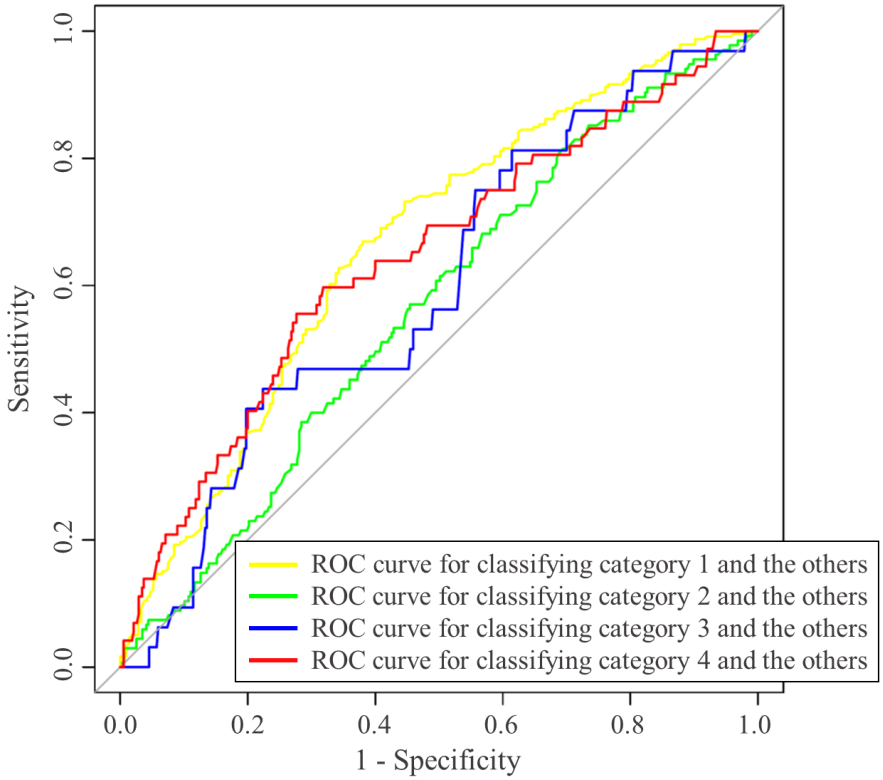

Figure 3.6 ROC curves of the random forest model using the ten variables in the external validation.

3.2.5 Analysis of important factors for each category using binary classification models

Table 3.8 shows the performance of the binary classification random forest models. The model classifying category 1 and others had the best performance. Category 3 was the most

difficult to classify.

Table 3.8 Performance of the binary classification random forest models (95% confidence interval).

Model	Model classifying category 1 and others	Model classifying category 2 and others	Model classifying category 3 and others	Model classifying category 4 and others
AUC	0.69 (0.63-0.75)	0.62 (0.54-0.70)	0.59 (0.49-0.70)	0.64 (0.56-0.72)

Tables 3.9-12 show the top ten important variables of each model. They suggest important factors in the identification of each category. Variables in category 2 were primarily related to diet and water intake habits. Variables in category 3 were mainly related to physical condition and constitution. Variables in category 1 and 4 were related to various types of factors. Alcohol intake is also typical for category 4.

Table 3.9 Variables of the model classifying category 1 and others.

Question Type	Variable	Mean decrease in Gini coefficient
Physical characteristics	BMI	9.58
Physical characteristics	Age	6.46
Physical characteristics	Height	3.17
Dietary habits	Frequency of intake of sweet bread, side dish bread, sandwiches, and pizza.	3.08
Physical condition and constitution	Gained 10 kg or more since age 20	2.83
Sleep habits and drowsiness	Stay up late	2.32
Water intake	Frequency of tea intake at lunch	2.03
Dietary habits	Frequency of cheese, dried fruits, and nuts intake	1.66
Physical condition and constitution	Greasy forehead	1.65
Workstyle and lifestyle	Screen time on working day	1.65

Table 3.10 Variables of the model classifying category 2 and others.

Question Type	Variable	Mean decrease in Gini coefficient
Physical characteristics	BMI	1.68
Physical characteristics	Height	1.09
Water intake	Frequency of tea intake at lunch	1.05
Physical characteristics	Age	0.98
Water intake	Frequency of milk and yogurt drink at breakfast	0.75
Dietary habits	Frequency of fruits (except berries) intake	0.73
Dietary habits	Which do you usually eat: rice or bread?	0.69
Dietary habits	Frequency of simmered food intake	0.65
Workstyle and lifestyle	Take time to soak in a bath	0.65
Dietary habits	Frequency of soup intake	0.64

Table 3.11 Variables of the model classifying category 3 and others.

Question Type	Variable	Mean decrease in Gini coefficient
Physical characteristics	BMI	1.52
Dietary habits	Frequency of eating lunch	0.84
Physical condition and constitution	Nausea and abdominal bloating	0.61
Physical condition and constitution	Coldness in areas other than hands and feet	0.59
Exercise habits	Duration of walking on working days	0.58
Dietary habits	Amount of soup left in a noodle dish	0.57
Water intake	Frequency of milk and yogurt drink intake at lunch	0.56
Physical condition and constitution	Food or drink stings your teeth	0.52
Alcohol intake	Frequency of wine intake	0.50
Water intake	Frequency of fruit and vegetable drinks intake	0.48

Table 3.12 Variables of the model classifying category 4 and others.

Question Type	Variable	Mean decrease in Gini coefficient
Physical characteristics	BMI	2.74
Physical characteristics	Age	2.66
Alcohol intake	Frequency of shochu and awamori intake	1.95
Water intake	Frequency of water intake at breakfast	1.10
Water intake	Frequency of fruit and vegetable drinks intake in the morning	0.80
Exercise habits	Duration of walking	0.77
Sleep habits and drowsiness	Wake up refreshed in the morning	0.73
Physical condition and constitution	Sweat even though not doing anything	0.68
Alcohol intake	Frequency of whiskey, brandy, gin, and vodka intake	0.64
Water intake	Frequency of milk and yogurt drink intake in the morning	0.63

3.2.6 Regression analysis of glucose metabolism status

Figure 3.7.a shows the correlation between the predicted 30 mPG by the multiple regression model and the referenced 30 mPG. There was a significant but weak correlation ($r = 0.30$). In addition, the predicted 30 mPG values were not distributed over the high and low values, which were difficult to predict. Figure 3.7.b shows the correlation between the predicted 120 mPG by the multiple regression model and the referenced 120 mPG. There was also a significant but weak correlation ($r = 0.20$). The predicted 120 mPG values were also concentrated around the midpoint. Figure 3.7.c shows the correlation between the predicted Matsuda index by the multiple regression model and the referenced Matsuda index. There was a significant correlation ($r = 0.53$). The predicted Matsuda index values were not distributed over the high values

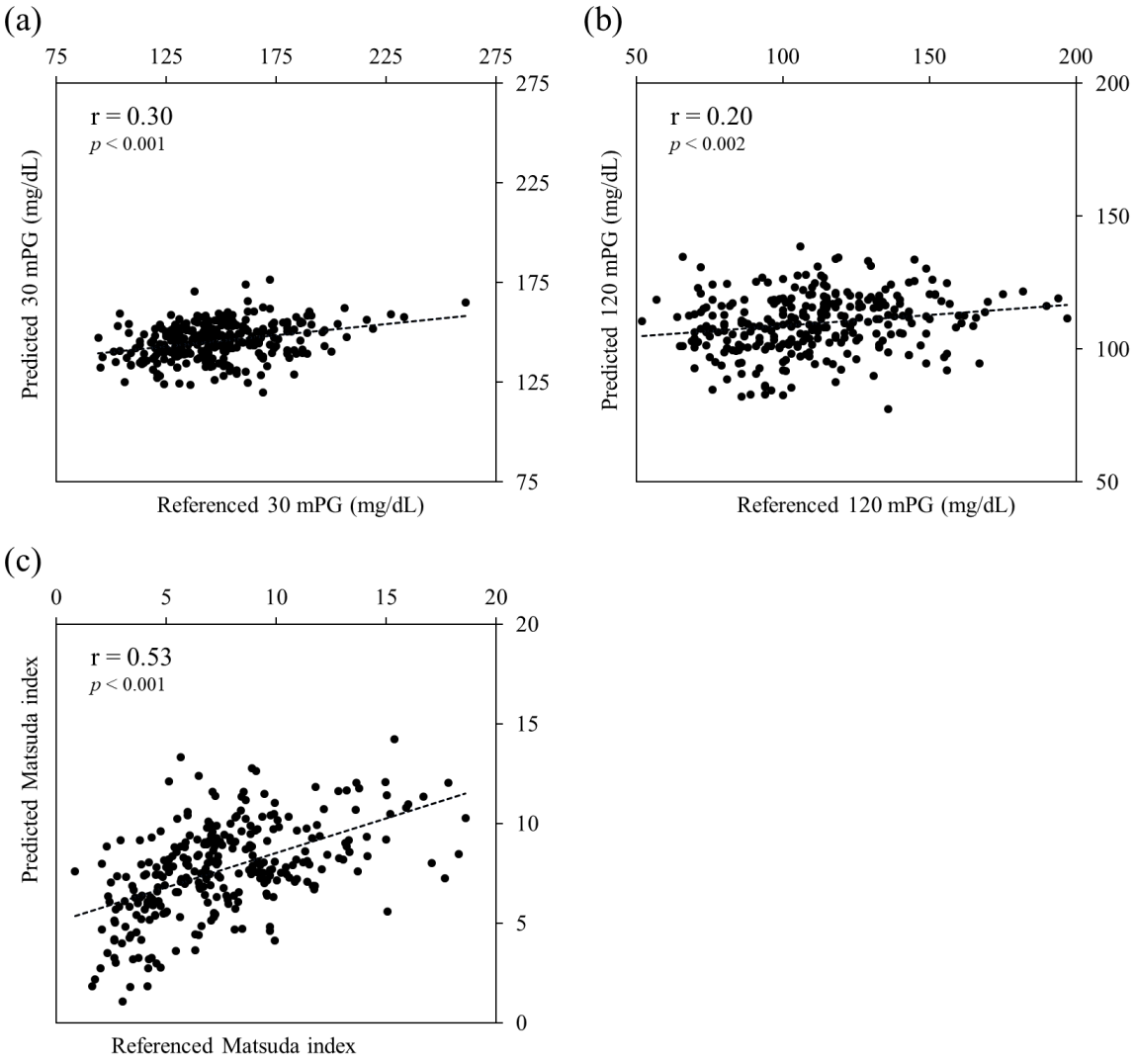

Figure 3.7 Correlations between the referenced and predicted values by the multiple regression models. (a) 30 mPG, (b) 120 mPG, (c) Matsuda index.

3.3 Discussion

3.3.1 Strengths of this study to identify glucose metabolism status

In this study, we identify the glucose metabolism status of non-diabetic Japanese using machine learning models with a questionnaire. This study had two unique features. First, we identified the glucose metabolism status rather than pre-diabetes. Our previous study classified the non-diabetic Japanese population into four glycometabolic categories. Each category had distinctly different insulin sensitivity and secretory characteristics [50].

Although screening tools for pre-diabetes have been developed [42-47], pre-diabetes presents overlapping pathophysiology of impaired insulin sensitivity and secretion [68, 69]. This is the first study to develop a model to distinguish between decreased insulin sensitivity and secretion. This model encourages individuals to understand their glucose metabolism status. It may help them to change their lifestyle to prevent diabetes.

Second, the model requires only ten questions about lifestyle and physical information that can be answered on the spot. Unlike patients with diabetes, non-diabetic people have no strong motivation or compulsion to undergo screening tests. Clinical measurement values, such as fasting plasma glucose, are valid predictors of glucose metabolism status [42-47]. However, the need to link to clinical laboratory data may limit their scope of use. Moreover, in the questionnaire tool, variables that the user cannot remember may reduce the user's motivation to use. The simplicity of our tool helps individuals expand their opportunities to know their glucose metabolism status.

Notably, this model should be used for screening, and a clinical test should make an accurate diagnosis by a doctor [25]. Despite the increasing number of patients with diabetes worldwide [13], impaired glucose metabolism is being overlooked because of its asymptomatic nature [70]. Understanding one's glucose metabolism status may provoke stronger behavioral motivation than vague lifestyle-related improvement suggestions.

3.3.2 Comparison with previous studies

A systematic review of risk assessment tools for pre-diabetes reported a mean AUC of 0.7, ranging from 0.66 to 0.75 [71]. Meanwhile, in this study, the AUC for classifying category 1 (the best glucose metabolism group) and the other categories (impaired glucose metabolism groups) was 0.68 in the random forest model using ten variables. However, this model has its own advantages, as it differs from previous studies. Furthermore, some previous studies used the history of hyperglycemia and hypertension as variables [39, 72-75]. These factors have a clear association with diabetes risk. Thus, these factors may have contributed to improving model performance. On the other hand, the participants of this study excluded patients with hyperglycemia and hypertension. Therefore, it may have affected the model's performance.

We used a random forest model for variable selection. We employed ten variables as a nice round number that was manageable to answer for a respondent. The 95% confidence intervals for the AUCs in our model were wide; therefore, we could not determine the number of questions from a threshold at which the model performance would be statistically

significantly decreased. Choosing the optimal number of questions with a larger sample size is a topic for future study. The AUC for classifying category 1 and the AUC for classifying category 4 were higher than those of categories 2 and 3. Category 1 has the best glucose metabolism status, while category 4 has the worst glucose metabolism status with low insulin secretion and sensitivity [50]. Categories 2 and 3 are intermediate. Therefore, categories 1 and 4 may have been easier to identify.

3.3.3 Lifestyle factors related to glucose metabolism status

We selected ten variables using a random forest. There were significant correlations between most variables and 30 mPG, 120 mPG, and the Matsuda index, which are the classification criteria of the glycometabolic category. Furthermore, these variables have been reported to be associated with glucose metabolism status and diabetes risk. The variables included age, BMI, and height. Aging diminishes the ability to secrete insulin [23, 76], whereas obesity decreases insulin sensitivity [22, 77-79]. Cohort studies in Europe and Israel reported that height and risk of type 2 diabetes are inversely correlated [80, 81].

Sleep, diet, and lifestyle factors were also employed. Laboratory interventions of circadian disruption were found to attenuate insulin sensitivity and secretion [82-64]. Sleep duration is related to the risk of developing diabetes. Sleeping 7 to 8 hours per day has the lowest diabetes risk [85]. Insomnia disorder with short sleep duration is associated with a higher risk of diabetes [86]. In addition, bedtime mobile phone use decreases sleep quality [87, 88]. Therefore, screen time at bedtime may be associated with glucose metabolism status. In fact, our correlation analysis showed that the more frequently mobile phone and tablet computer were used at bedtime, the worse the Matsuda Index was. However, the association has not been reported. This study is the first to suggest an association between glucose metabolism status and mobile phone and tablet computer use at bedtime.

Preference for rice or bread reflects an individual's dietary style. Rice is the primary source of carbohydrates for Asians. There is a positive association between rice intake and the development of diabetes [89]. Replacing refined grains with whole grains is recommended for diabetes prevention [90]. The relationship between diabetes and various dietary styles, such as Mediterranean and vegetarian diets, has been studied [91, 92]. However, there is no one dietary style that is best for diabetes prevention [93, 94]. Further research is needed to determine the appropriate dietary style for each person. The frequency of soup intake may also be a factor related to dietary style.

Tea is rich in polyphenols and caffeine. Several *in vitro* studies have shown that tea

components enhance insulin sensitivity and secretion [95-98]. In addition, multiple epidemiological studies have shown that chronic tea consumption decreases the risk of diabetes [99-102]. On the contrary, intervention trials have reported inconsistent results regarding the effects of tea on glucose metabolism [103-106].

Periodontal disease and oral inflammation worsen glycemic control and increase diabetes risk [107, 108]. Therefore, oral hygiene habits, such as the frequency of toothbrush replacement, may be necessary for maintaining glucose metabolism status. However, their association has not been reported [109]. This study suggests for the first time that oral hygiene habits may be associated with glucose metabolism status.

Furthermore, we developed four random forest models of binary classification to consider the factors with high importance for each category. Each model's top ten important variables suggest essential factors in identifying each category. The results indicated that category 2 was related to diet and water intake habits. Category 2 has impaired insulin sensitivity. Obesity and impaired lipid metabolism are reported to cause impaired insulin sensitivity [110, 111]. The result provided additional evidence for a link between dietary habits and insulin sensitivity. Category 3 was related to physical condition and constitution. Category 3 has decreased insulin secretion. A genetic factor is one of the causes of insulin secretion deficiency [21]. Therefore, the physical constitution may be important in identifying category 3. Categories 1 and 4 were associated with various types of factors. Category 4 is a state in which both insulin sensitivity and secretion are deficient. The result reflected that the accumulation of diverse lifestyle habits causes the deterioration of glucose metabolism. Further study should clarify the causal relationship and molecular mechanism between them. It may offer new lifestyle habits to prevent diabetes.

Besides, principal component analysis (PCA) may contribute to considering important factors in the dataset. Then, we performed PCA on the original dataset. As a result, the contribution ratio of principal component 1 was very high at 0.986. Height, age, and BMI factors accounted for most of the PC1 principal component loading. Furthermore, the principal component scores could not separate participants into categories 1, 2, 3, and 4. Then, we further excluded height, age, and BMI from the original dataset and performed PCA. However, we could not separate participants into categories 1, 2, 3, and 4. Therefore, these results showed that PCA is difficult to identify glucose metabolism status with the questionnaire (data not shown).

3.3.4 Regression analysis of glucose metabolism status

We attempted to predict glucose metabolism status with continuous values by multiple regression analysis. However, there were weak correlations between referenced and predicted values by the multiple regression models of 30 mPG and 120 mPG. In addition, high and low values were difficult to predict. As to the Matsuda index, the correlation coefficient between referenced and predicted values was 0.53. However, to detect categories 2 and 4, i.e., decreased insulin sensitivity, it is essential to predict a range of Matsuda index below 4.97. In this range, the correlation coefficient was 0.21 (figure 3.8).

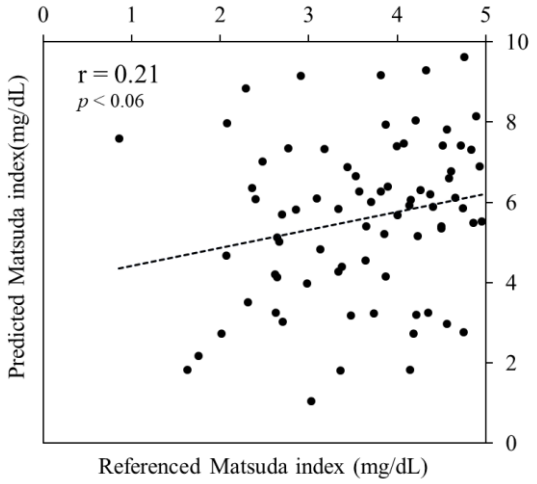

Figure 3.8 Correlation between the referenced and the predicted Matsuda index by the multiple regression model in the range of 0 to 4.97.

The lack of training data on high and low values may cause low prediction performance. Data augmentation may help solve this issue. The synthetic minority oversampling technique (SMOTE) is one of the data augmentation techniques. This technique generates new synthetic samples of the minority class by applying the nearest neighbor method [112]. Furthermore, blood glucose levels range from approximately 70 to 250 mg/dL. On the other hand, there are up to six options for questionnaire responses. Therefore, there may be a limitation in predicting continuous values of the glucose metabolism status rather than the four glycometabolic categories. The results showed that the prediction performance of the multiple regression models was poor, and the range of predicted values was skewed. Therefore, we did not classify glycometabolic categories based on the predicted 30 mPG, 120 mPG, and Matsuda index.

4 Conclusion

4.1 Conclusion of this thesis

In Chapter 2, we validated a selective recall method for assessing water intake in Japanese people. In this method, intake data for food and beverages is collected through a multiple-choice questionnaire. Then, the data is multiplied by each WCFC to calculate the total water intake. In a clinical trial, we assessed participants' water intake using the selective recall method and an established method, the descriptive dietary record method. Then we analyzed the correlation between the water intakes from the two methods. As a result, the two methods had a strong positive correlation ($r = 0.94$, $p < 0.0001$). In addition, water intakes by the two methods from non-alcoholic beverages ($r = 0.94$, $p < 0.0001$), alcoholic beverages ($r = 1.00$, $p < 0.0001$), and food ($r = 0.72$, $p < 0.0001$) were also strongly correlated. Therefore, we demonstrated that the selective recall method accurately assesses the total water intake for each source. Thus, the method will help people to intake an adequate amount of water and maintain their health. It will also contribute to future epidemiological studies to examine the relationship between water intake habits and health.

In addition, we reported for the first time the accurate habitual water intake of Japanese people using this method. The survey was conducted on four days, including three working days and one non-working day. The mean total water intake of all participants was 2949 mL/day. The ratio of water intake from beverages to food was 2.1:1. We also analyzed the relationship between water intake and hydration status by urinalysis. A significant, negative correlation was observed between total water intake and urine osmolality in men ($r = -0.55$, $p = 0.0011$) and women ($r = -0.51$, $p = 0.0032$) aged 20 to 44. On the other hand, no correlation was observed among the participants aged 45 to 64. The body's response to water intake may have fluctuated with aging.

In Chapter 3, we developed a model to identify the glucose metabolism status of non-diabetics using a simple questionnaire. The AUC for identifying categories with impaired glucose metabolism was 0.68. In the external validation, the AUC was 0.66, and the robustness of the model was demonstrated. The model had the following two features. First, it identified glucose metabolism status, i.e., insulin sensitivity and secretion, rather than pre-diabetes. Then, it may help us to take appropriate measures to prevent diabetes that suit each status. Second, it required only ten factors, only questions about lifestyle and physical information that could be answered on the spot. Because invasive measurement factors or excessive factors are not needed, it can be easily and widely used by the general population. This model provides an opportunity for many non-diabetic individuals to identify their

glucose metabolism status. That may lead them to start improving their lifestyle to reduce their diabetes risk.

The variables of the model were selected using a random forest. They are age, height, BMI, and the following questions: “Do you wake up in the middle of the night”, “Which do you usually eat: rice or bread”, “Frequency of tea intake per week at lunch”, “Do you wake up late on a non-working day”, “Frequency of mobile phone and tablet computer use at bedtime”, “Frequency of soup intake”, and “Frequency of toothbrush replacement”. Several factors have not been reported to be associated with glucose metabolism status. Thus, this study suggested new factors related to glucose metabolism status and diabetes risk.

This thesis analyzed water intake and glucose metabolism status using questionnaires. These methods are simple and do not require special examinations at hospitals. Therefore, they will help the general population to monitor their own health status and prevent disease. In addition, these methods can be used in future studies to elucidate the relationship between lifestyle and health status. It may offer a new lifestyle modification approach to prevent disease.

4.2 Remaining problems and future directions

In the water intake assessment, this study had some limitations. First, the surveys were conducted only in the summer. The ratio of water intake from beverages to that from food was 2.1:1 in this study. However, a season of a survey may affect water intake volume. Second, the participants were employees who worked at Suntory World Research Center in Kyoto. The results of this study might be subject to the potential influence of participant demographics and seasonality. Third, Participants were limited to those under 65 years of age. Older people are at increased risk of dehydration due to low muscle mass, impaired kidney function, physical and cognitive impairments, and dull thirst [113]. Therefore, it is essential to verify the validity of this method in the elderly. A large-scale study conducted throughout the four seasons in multiple regions is needed. A trial with elderly subjects is also required. It will provide more general and robust evidence of the effectiveness of this method.

Furthermore, this study did not search for the best choices of survey instrument. The correlation coefficient of the water intakes from food was 0.72. The descriptive dietary record method can describe all foods, while the selective recall method limits the options to 40 food categories. This is one reason the water intake from the two methods did not match. Further research to explore the optimal choices will improve the accuracy of the selective recall method.

In identifying glucose metabolism status, this study had some limitations. First, the participants recruited volunteers rather than randomly selected population-based samples. Second, the questionnaire was not validated, so measurement errors may have occurred. Nevertheless, the model was validated by external validation using data from people in other regions of Japan. Therefore, the robustness of the model was confirmed. Third, the questionnaire was aimed at Japanese people. In particular, the dietary questions were based on Japanese food. Therefore, the model's application to other countries and ethnic groups may be limited. Further study is needed to acquire and analyze data from different populations to expand the application scope.

In addition, we analyzed the relationship between glucose metabolism status and a wide range of lifestyle information, including unknown diabetes risk factors. The ten variables were selected for the random forest model. Some of these lifestyles have not been reported to be associated with glucose metabolism status. This result may suggest meaningful information for future diabetes studies. Further study should clarify the causal relationship and molecular mechanism. They may suggest new lifestyle habits to prevent diabetes.

This thesis assessed water intake and glucose metabolism status using questionnaires. We aimed to establish methods and improve these accuracies in each. Further study is required to find variables common to the two health statuses. It will develop a method to predict both conditions with a common questionnaire. Furthermore, there are many other health indicators and diseases besides these two. This study cannot cover them all. For example, there were 3.3 million diabetes patients, while there were 9.9 million hypertensive disease patients and 1.8 million malignant neoplasms patients in Japan in 2017 [114]. For these diseases as well, health status assessment methods are needed to reduce the risk. A future prospect of this study is to develop a questionnaire that can predict multiple health conditions with common variables. Therefore, more people will be able to assess their health status from various perspectives. We believe this study can be a first step toward constructing such a beneficial assessment method.

Appendix

A. Questionnaire on lifestyle and physical characteristics

Please answer the following questions about your lifestyle and physical condition over the last month.

Please provide only one option per question except for the last question in this survey, for which you can provide multiple answers.

[Exercise habits]

◆ Activity intensity

How did you spend your working days?	1. Mainly static activities (Most of the work was done sitting down. There was little movement, or it was a static way to move while sitting down, such as driving).	2. Exercise moderately (You spent a lot of time sitting, but you had the opportunity to exercise for commuting, moving in the workplace, standing work / customer service, shopping, housework, etc.)	3. Actively moving (You spent a lot of time moving and standing for work and housework. You spent a lot of time on farm work, carrying heavy luggage, going up and down stairs, etc.)
How did you spend your non-working days?	1. Mainly static activities (Most of the time was spent sitting, and most of the activities were static, such as watching TV, reading, and driving).	2. Light exercise and walking	3. Active engagements such as mountain climbing, running, swimming, tennis, and soccer.

◆ Travel method and time required (working days)

On foot	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
Running	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
By bicycle	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
By train / bus	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
By car / motorcycle	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min

◆ Travel method and time required (non-working days)

On foot	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
Running	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
By bicycle	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
By train / bus	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
By car / motorcycle	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min

◆ Duration of exercise (weekly total)

Walking	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
Running	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
Swimming	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
Strength training	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min
Other exercise	1. 90 min or more	2. 60–89 min	3. 30–59 min	4. 5 – 29 min	5. Less than 5 min

◆ Please select the one that applies to your exercise situation and awareness of exercise.

You joined a sports gym	1. Yes	2. No			
You own athletic shoes or gym clothes	1. Yes	2. No			
Time of exercise	1. before meals	2. after meals	3. both before and after meals	4. not decided in particular	5. No exercise habits

[Sleep habits and drowsiness]

◆ Sleep habits

Average sleeping time	1. 8 hours or more	2. 7 hours to less than 8 hours	3. 6 hours to less than 7 hours	4. 5 hours to less than 6 hours	5. less than 5 hours
Variation in bedtime	1. almost the same time every day.	2. different time on working days and non-working days	3. Different time each day		
Falling asleep	1. You don't have trouble falling asleep	2. You have trouble falling asleep			
Variation in sleep duration	1. almost the same duration every day	2. Different duration on working days and non-working days	3. Different duration each day		

◆ Sleep habits and drowsiness

Feel that you get enough sleep	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Stay up late	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Take sleeping pills	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Drink alcohol to sleep	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
You fall asleep and wake up right away	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Frequency of mobile phone and tablet computer use at bedtime	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Spend in a bright room until just before going to bed	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Wake up in the middle of the night	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Wake up during the night because you feel the need to go to the toilet.	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Wake up in the middle of the night and have difficulty sleeping afterwards	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Feel suffocation during sleep	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
When you wake up in the morning still feel sleepy and tired	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Wake up refreshed in the morning	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Feel sticky in the mouth when you wake up in the morning	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Not hungry when you wake up in the morning.	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Become drowsy after a meal	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Long-lasting drowsiness after a meal	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Feel too sleepy to stay awake, even while working or doing household chores	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Wake up late on non-working day	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
You cannot wake up on non-working day without an alarm clock.	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Take a nap on non-working day	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Have habits to sleep well	1. Yes	2. No			
Have been told that you have stopped breathing during sleep	1. Yes	2. No			
Have been told you snores	1. Yes	2. No			
Have been told you are bruxing your teeth during sleeping	1. Yes	2. No			

[Dietary habits]

◆ Frequency and duration

Frequency of eating breakfast per week	1. Every day	2. 4–6 times a week	3. 2–3 times a week	4. Once a week	5. Less than once a week
Time required for breakfast	1. 30 min or more	2. 20 min to less than 30 min	3. 10 min to less than 20 min	4. 5 min to less than 10 min	5. less than 5 min
Frequency of eating lunch per week	1. Every day	2. 4–6 times a week	3. 2–3 times a week	4. Once a week	5. Less than once a week
Time required for lunch	1. 30 min or more	2. 20 min to less than 30 min	3. 10 min to less than 20 min	4. 5 min to less than 10 min	5. less than 5 min
Frequency of eating supper per week	1. Every day	2. 4–6 times a week	3. 2–3 times a week	4. Once a week	5. Less than once a week
Time required for supper	1. 30 min or more	2. 20 min to less than 30 min	3. 10 min to less than 20 min	4. 5 min to less than 10 min	5. less than 5 min
Frequency of eating after-supper snacks per week	1. Every day	2. 4–6 times a week	3. 2–3 times a week	4. Once a week	5. Less than once a week
Time required for after-supper snacks	1. 30 min or more	2. 20 min to less than 30 min	3. 10 min to less than 20 min	4. 5 min to less than 10 min	5. less than 5 min
Frequency of snacking per week (between breakfast and lunch or between lunch and dinner)	1. Every day	2. 4–6 times a week	3. 2–3 times a week	4. Once a week	5. Less than once a week
Time required for snacks (between breakfast and lunch or between lunch and dinner)	1. 30 min or more	2. 20 min to less than 30 min	3. 10 min to less than 20 min	4. 5 min to less than 10 min	5. less than 5 min

◆ Meal timing

Eat breakfast at about the same time	1. Yes	2. No	3. You eat breakfast less than once a week
Eat lunch at about the same time	1. Yes	2. No	3. You eat lunch less than once a week
Eat supper at about the same time	1. Yes	2. No	3. You eat supper less than once a week
Eat after-supper snacks at about the same time	1. Yes	2. No	3. You eat after-supper snacks less than once a week
Eat snacks (between breakfast and lunch or between lunch and dinner) at about the same time	1. Yes	2. No	3. You eat snacks less than once a week

◆ Dietary habits

Frequency of consuming yogurt and probiotics drinks	1. Every day	2. 4–6 times a week	3. 2–3 times a week	4. once a week	5. less than once a week
Which do you usually eat: rice or bread?	1. usually eat rice	2. eat rice more often than bread	3. about the same frequency	4 eat bread more often than rice	5. usually eat bread
Amount of rice you usually ate compared to one serving of rice at a restaurant	1. more	2. somewhat more	3. about the same	4. somewhat less	5. less
Amount of main dish you usually ate compared to one serving of main dish at a restaurant	1. more	2. somewhat more	3. about the same	4. somewhat less	5. less
Frequency of soup intake	1. more	2. somewhat more	3. about the same	4. somewhat less	5. less
Frequency of noodles intake	1. more	2. somewhat more	3. about the same	4. somewhat less	5. less
Frequency of vegetable intake	1. more	2. somewhat more	3. about the same	4. somewhat less	5. less
Amount of soup left in noodle dish	1. very little left	2. 20–30%	3. About half	4. 70–80%	5. almost all left

◆ Characteristics of your diet

Try to eat well-balanced diet	1. Yes	2. No
Try not to overeat	1. Yes	2. No
Try to choose low-carbohydrate diet	1. Yes	2. No
faster to eat than others	1. Yes	2. No
Try to choose low-salt diet	1. Yes	2. No
Try to eat vegetables first	1. Yes	2. No
Often eat sweets	1. Yes	2. No

Often eat seafood, tofu, and beans.	1. Yes	2. No
Eat more than other people	1. Yes	2. No
Often have a large serving and extra serving	1. Yes	2. No
Feel unsatisfied without rice dish or noodles after drinking alcohol	1. Yes	2. No
Finish supper two hours before going to bed	1. Yes	2. No
Meals served one by one	1. Yes	2. No
Often eat after-meal dessert	1. Yes	2. No
Often add salt or soy sauce to dishes.	1. Yes	2. No
Try to choose meat with less fat and remove excess fat.	1. Yes	2. No
Often eat until you're full.	1. Yes	2. No
Often eat rice dish or noodle or sweets after drinking alcohol.	1. Yes	2. No
Hardly ever cook at home	1. Yes	2. No
Try to chew well and eat	1. Yes	2. No
After supper, go to bed on a full stomach.	1. Yes	2. No
Try to eat a lot of vegetables	1. Yes	2. No

◆ Meal details

	Days of intake per week					
	Don't eat	More than one day a month but less than one day a week	One day a week	2 to 3 days a week	4 to 6 days a week	Almost every day a week
■ Rice						
Rice ball	0	1	2	3	4	5
Rice porridge and risotto	0	1	2	3	4	5
Japanese seasoned rice with vegetables	0	1	2	3	4	5
Rice (white rice, brown rice, barley rice, etc.)	0	1	2	3	4	5
Fried rice, pilaf, and omelet rice	0	1	2	3	4	5
Sushi	0	1	2	3	4	5
Bowl of rice with topping	0	1	2	3	4	5
Curry rice, hashed beef rice	0	1	2	3	4	5
■ Bread						
Plain bread	0	1	2	3	4	5
Bread rolls, croissants, French bread	0	1	2	3	4	5
Sweet bread, side dish bread, sandwiches, hamburgers, and pizza.	0	1	2	3	4	5
■ Noodles, etc.						
Noodles (ramen, udon, soba, pasta, etc.)	0	1	2	3	4	5
Okonomiyaki (savory Japanese-style of pancake.)	0	1	2	3	4	5
Cereal	0	1	2	3	4	5
■ Soup						
Soup	0	1	2	3	4	5
Stew	0	1	2	3	4	5
Japanese hot pot dish	0	1	2	3	4	5
■ Main and side dishes						
Simmered Food (main dishes: pot-au-feu, meat potato, etc.)	0	1	2	3	4	5
Simmered Food (side dishes: hijiki seaweed, dried strips of radish, etc.)	0	1	2	3	4	5

Deep-fried food	0	1	2	3	4	5
Grilled fish (grilled fish, grilled meat, hamburger steak, fried egg, etc.)	0	1	2	3	4	5
Stir-fried food (stir-fried vegetables, mapo tofu, roasted eggs, etc.)	0	1	2	3	4	5
Boiled and steamed food (hot vegetables, boiled eggs, chawanmushi, shumai, etc.)	0	1	2	3	4	5
Raw food (sashimi, natto, tofu, cod roe, etc.)	0	1	2	3	4	5
Salad, raw vegetables	0	1	2	3	4	5
Vinegared food	0	1	2	3	4	5
Ohitashi (boiled spinach seasoned with soy sauce)	0	1	2	3	4	5
Japanese pickles	0	1	2	3	4	5
■ Cheese, fruits, desserts, sweets, etc.						
Cheese, dried fruits, nuts	0	1	2	3	4	5
Fruits (except berries)	0	1	2	3	4	5
Fruits (berries)	0	1	2	3	4	5
Yogurt, pudding, jelly, ice cream	0	1	2	3	4	5
Cakes, pies, tarts	0	1	2	3	4	5
Traditional Japanese sweets	0	1	2	3	4	5
Fried confectionery	0	1	2	3	4	5

[Water intake]

◆ Drink intake

Time	Beverage	frequency						amount				
		Don't drink	More than one day a month but less than one day a week	One day a week	2 to 3 days a week	4 to 6 days a week	Almost every day a week	Don't drink	about 100 mL	about 200 mL	About 350 mL	500 mL or more
1. After waking up–before breakfast	Tea	0	1	2	3	4	5	A	B	C	D	E
	Water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E
2. At breakfast	Tea	0	1	2	3	4	5	A	B	C	D	E
	Water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E
3. After breakfast–before lunch	Tea	0	1	2	3	4	5	A	B	C	D	E
	Water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E
4. At lunch	Tea	0	1	2	3	4	5	A	B	C	D	E
	Water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E
5. After lunch–before dinner	Tea	0	1	2	3	4	5	A	B	C	D	E
	Water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E
6. At supper	Tea	0	1	2	3	4	5	A	B	C	D	E
	Water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E
7. After supper to 30 minutes before bedtime	Tea	0	1	2	3	4	5	A	B	C	D	E
	Water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E

8. 30 minutes before bedtime to bedtime	Tea	0	1	2	3	4	5	A	B	C	D	E
	water	0	1	2	3	4	5	A	B	C	D	E
	Coffee	0	1	2	3	4	5	A	B	C	D	E
	Milk and yogurt drink	0	1	2	3	4	5	A	B	C	D	E
	Fruit and vegetable drinks	0	1	2	3	4	5	A	B	C	D	E
	Other soft drinks	0	1	2	3	4	5	A	B	C	D	E

◆ Water intake habits

Your awareness of water intake	1. Try to drink water even if you aren't thirsty	2. Try to drink when you feel thirsty	3. Not particularly conscious
How do you drink water	1. Little by little	2. Drink much at once	
Your water intake compared to that of others	1. More than others	2. As much as anyone else.	3. Less than others
Wake up and drink water while in bed	1. Always	2. Sometimes	3. Never
Bring drinks from home when going out for long periods of time	1. Always	2. Sometimes	3. Never
Put sugar in coffee and tea	1. Always	2. Sometimes	3. Never

[Alcohol intake]

◆ Alcohol intake

Types of alcoholic beverages [definition of one drink].	frequency						amount				
	Do not drink	More than one day a month but less than one day a week	One day a week	2 to 3 days a week	4 to 6 days a week	Almost every day a week	Do not drink	1 drink	2 drinks	3 drinks	4 drinks or more
Beer [500 mL or 1 medium bottle or 1 beer mug]	0	1	2	3	4	5	A	B	C	D	E
Chu-hai and sour [350 mL]	0	1	2	3	4	5	A	B	C	D	E
Shochu, and Awamori [100 mL or 1 cup]	0	1	2	3	4	5	A	B	C	D	E
Sake [180 mL]	0	1	2	3	4	5	A	B	C	D	E
Whiskey, brandy, gin, vodka [30 mL or single cup]	0	1	2	3	4	5	A	B	C	D	E
Wine [120 mL or 1 glass]	0	1	2	3	4	5	A	B	C	D	E

◆ Alcohol intake habits

Have many opportunities to drink alcohol	1. Yes	2. No
Drink alcohol during the day on non-working day	1. Yes	2. No
Eat too much when drinking alcohol	1. Yes	2. No
Eat low-calorie snacks when drinking alcohol	1. Yes	2. No
Not drink alcohol more than two days a week	1. Yes	2. No

[Physical condition and constitution]

◆ Regarding your family (grandparents, parents, siblings)

Someone in your family has/had diabetes	1. Yes	2. No
Someone in your family is/was obesity	1. Yes	2. No
Someone in your family has/had dementia	1. Yes	2. No
Someone in your family has/had thinning hair	1. Yes	2. No

◆ Physical condition

Feel energetic	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Get tired easily	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Feel out of breath easily	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Feel your heart pounding	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Prefer a quiet environment and find it troublesome to talk to others	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Nausea and abdominal bloating	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Cold hands and feet even in warm places	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Coldness in areas other than hands and feet (back, abdomen, hips, knees, etc.)	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
sweat even though not doing anything	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Greasy forehead	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Dry skin and lips	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Greasy nose	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Acne and pimples	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Sticky mouth	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Feel thirsty	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Diarrhea after intake cold food	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Sticky stool and feel of incomplete defecation	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Hard stool and constipation	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Bleed when brushing teeth	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Gingival recession and wide tooth gaps	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never
Food or drink stings your teeth.	1. Always	2. Frequently	3. Sometimes	4. Seldom	5. Never

◆Your constitution

Easy to get fat	1. Yes	2. No	
Hard to build muscle even after exercising	1. Yes	2. No	
Cannot go through with a diet	1. Yes	2. No	3 have never been on a diet
Regained weight after a diet	1. Yes	2. No	3 have never been on a diet
Gained more than 10 kg after the age of 20	1. Yes	2. No	
Gained more than 10 kg after the age of 20	1. Yes	2. No	
Weight has changed by more than 3 kg in the last year	1. Yes	2. No	
Go to the bathroom more often than others	1. Yes	2. No	
Lower abdomen sticks out	1. Yes	2. No	
Concerned about hair loss	1. Yes	2. No	
Hair getting thinner	1. Yes	2. No	
Flabby belly	1. Yes	2. No	
Regular teeth	1. Yes	2. No	
Tooth decay	1. Yes	2. No	
Periodontal disease or alveolar pyorrhea	1. Yes	2. No	
Often wear warm clothes in winter	1. Yes	2. No	
Sensitive to cold	1. Yes	2. No	

[Workstyle and lifestyle]

◆ Workstyle

Workstyle	1. Full-time work	2. Part-time work	3. Other than 1, 2 or unemployed			
Do you have a night shift?	1. Yes	2. No	3. Unemployed			
Job description	1. Sedentary work	2. Standing work	3. Physical work	4. Unemployed		
Requires complicated thinking	1. Yes	2. No	3. Unemployed			
Overtime in a month	1. 80 hours or more	2. 60 hours or more and less than 80 hours	3. 40 hours or more and less than 60 hours	4. 20 hours or more and less than 40 hours	5 Less than 20 hours	6. Unemployed
Often go home after 20:00	1. Yes	2. No	3. Unemployed			

◆ Oral hygiene habits

Select all the times you brush your teeth	1. After waking up	2. After breakfast	3. After lunch	4. Within 1 hour after supper	5. Before going to bed (after 1 hour after supper)
Use a dental floss or an interdental brush	1. Every day	2. 4-6 times a week	3. 2-3 times a week	4. Once a week	5. Less than once a week
Use mouth rinse	1. Every day	2. 4-6 times a week	3. 2-3 times a week	4. Once a week	5. Less than once a week
Frequency of toothbrush replacement	1. Within 1 month	2. Every 2 months	3. Every 3 months	4. Every 4 months	5. Do not replace for more than 4 months

◆ Lifestyle

Walk faster than others	1. Yes	2. No			
Often need to take stairs	1. Yes	2. No			
Try to use stairs instead of elevator or escalator	1. Yes	2. No			
Take time to soak in bath	1. 30 min or more	2. About 20 min	3. About 10 min	4. About 5 min	5. shower only
Often lie down right after eating.	1. Yes	2. No			
Often spend non-working days at home	1. Yes	2. No			

◆ Screen time (TV, computer, tablet, smart phone, etc.)

On working day (if you don't work, on weekdays)	1. 7 hours or more	2. 5 hours or more and less than 7 hours	3. 3 hours or more and less than 5 hours	4. 1 hour or more and less than 3 hours	5. less than one hour
On non-working day (if you don't work, on weekend)	1. 7 hours or more	2. 5 hours or more and less than 7 hours	3. 3 hours or more and less than 5 hours	4. 1 hour or more and less than 3 hours	5. less than one hour

◆ Please select all the benefits and target organs of supplements that you take at least 4 times a week.

1. Beauty and skin	2. Health maintenance and improvement	3. Joint	4. Fatigue recovery	5. Nutrition
6. Eye	7. Antioxidant and anti-aging	8. Weight loss	9. Slimming (becomes slim with good style)	10. Body fat suppression
11. Stiffness and pain in the neck, shoulders, and back	12. Bone	13. High blood pressure	14. Muscle	15. Bowel control and constipation
16. Anticoagulant	17. Others	18. Do not take supplements		

B. Characteristics of the preprocessed questionnaire answers for each glycometabolic category

	Category 1	Category 2	Category 3	Category 4
◆ Activity intensity				
How did you spend your working days?	1.83 (1.78–1.88)	1.79 (1.71–1.86)	1.89 (1.79–1.98)	1.79 (1.71–1.86)
How did you spend your non-working days?	1.51 (1.45–1.57)	1.42 (1.35–1.50)	1.51 (1.41–1.61)	1.51 (1.43–1.59)
◆ Travel method and time required (working days)				
On foot	3.10 (3.00–3.20)	3.15 (3.02–3.27)	3.14 (2.95–3.33)	3.17 (3.03–3.32)
Running	4.84 (4.79–4.89)	4.84 (4.77–4.91)	4.74 (4.63–4.85)	4.81 (4.71–4.90)
By bicycle	4.52 (4.44–4.59)	4.55 (4.45–4.65)	4.50 (4.35–4.64)	4.54 (4.43–4.65)
By train/bus	3.49 (3.37–3.61)	3.52 (3.33–3.71)	3.46 (3.22–3.70)	3.51 (3.31–3.70)
By car/motorcycle	4.65 (4.57–4.73)	4.46 (4.31–4.61)	4.57 (4.41–4.74)	4.58 (4.45–4.71)
◆ Travel method and time required (non-working days)				
On foot	3.48 (3.39–3.57)	3.53 (3.40–3.66)	3.61 (3.44–3.77)	3.48 (3.34–3.62)
Running	4.76 (4.70–4.82)	4.80 (4.72–4.89)	4.77 (4.67–4.88)	4.78 (4.69–4.87)
By train/bus	4.39 (4.31–4.47)	4.33 (4.21–4.45)	4.26 (4.08–4.45)	4.35 (4.22–4.49)
By car/motorcycle	4.50 (4.41–4.58)	4.17 (4.00–4.33)	4.37 (4.18–4.55)	4.23 (4.06–4.39)
◆ Duration of exercise (weekly total)				
Walking	4.00 (3.87–4.12)	4.24 (4.08–4.39)	4.10 (3.87–4.33)	3.83 (3.62–4.04)
Swimming	4.93 (4.89–4.97)	4.96 (4.91–5.00)	4.90 (4.82–4.99)	4.94 (4.88–4.99)
Strength training	4.52 (4.44–4.60)	4.60 (4.49–4.71)	4.61 (4.47–4.75)	4.57 (4.44–4.71)
Other exercises	4.43 (4.32–4.53)	4.56 (4.43–4.70)	4.59 (4.41–4.76)	4.58 (4.44–4.71)
◆ Please select the one that applies to your exercise situation and awareness of exercise.				
You joined a sports gym	1.87 (1.83–1.90)	1.86 (1.82–1.91)	1.86 (1.81–1.92)	1.89 (1.84–1.93)
You own athletic shoes or gym clothes	1.28 (1.24–1.32)	1.32 (1.25–1.38)	1.33 (1.25–1.41)	1.29 (1.22–1.35)
Time of exercise (before meals)	1.83 (1.80–1.87)	1.88 (1.84–1.93)	1.86 (1.80–1.92)	1.86 (1.81–1.91)
Time of exercise (after meals)	1.90 (1.87–1.93)	1.92 (1.89–1.96)	1.89 (1.83–1.94)	1.93 (1.90–1.97)
Time of exercise (both before and after meals)	1.98 (1.96–1.99)	1.99 (1.98–2.00)	1.98 (1.96–2.01)	1.99 (1.97–2.00)
Time of exercise (not decided in particular)	1.65 (1.61–1.70)	1.67 (1.60–1.73)	1.72 (1.64–1.80)	1.59 (1.52–1.66)
Time of exercise (No exercise habits)	1.63 (1.59–1.68)	1.54 (1.47–1.61)	1.55 (1.46–1.63)	1.63 (1.56–1.70)

◆ Sleep habits				
Average sleeping time	2.95 (2.86–3.03)	2.85 (2.74–2.97)	2.98 (2.84–3.11)	2.88 (2.75–3.01)
Variation in bedtime	1.31 (1.26–1.37)	1.34 (1.25–1.43)	1.26 (1.17–1.36)	1.34 (1.25–1.43)
Falling asleep	1.25 (1.21–1.29)	1.29 (1.23–1.35)	1.25 (1.17–1.32)	1.25 (1.19–1.31)
Variation in sleep duration	1.40 (1.34–1.45)	1.43 (1.34–1.52)	1.40 (1.29–1.51)	1.35 (1.26–1.44)
◆ Sleep habits and drowsiness				
Feel that you get enough sleep	2.54 (2.46–2.63)	2.56 (2.41–2.70)	2.67 (2.50–2.84)	2.58 (2.42–2.74)
Stay up late	3.27 (3.17–3.36)	3.35 (3.20–3.51)	3.53 (3.35–3.72)	3.49 (3.32–3.66)
Take sleeping pills	4.95 (4.93–4.98)	4.90 (4.84–4.96)	4.97 (4.93–5.01)	4.93 (4.88–4.99)
Drink alcohol to sleep	4.78 (4.72–4.84)	4.75 (4.66–4.84)	4.55 (4.37–4.73)	4.63 (4.50–4.75)
You fall asleep and wake up right away	4.40 (4.32–4.48)	4.35 (4.23–4.47)	4.33 (4.18–4.48)	4.21 (4.06–4.35)
Frequency of mobile phone and tablet computer use at bedtime	2.16 (2.06–2.26)	2.09 (1.95–2.23)	2.33 (2.13–2.53)	2.27 (2.10–2.44)
Spend in a bright room until just before going to bed	2.43 (2.31–2.55)	2.44 (2.27–2.61)	2.52 (2.29–2.74)	2.42 (2.22–2.61)
Wake up in the middle of the night	3.83 (3.73–3.92)	3.68 (3.53–3.83)	3.54 (3.35–3.74)	3.49 (3.33–3.66)
Wake up during the night because you feel the need to go to the toilet.	3.97 (3.87–4.07)	3.87 (3.73–4.02)	3.63 (3.42–3.84)	3.63 (3.44–3.81)
Wake up in the middle of the night and have difficulty sleeping afterward	4.32 (4.25–4.40)	4.28 (4.16–4.40)	4.25 (4.09–4.40)	4.16 (4.02–4.30)
Feel suffocation during sleep	4.90 (4.87–4.93)	4.90 (4.85–4.95)	4.91 (4.86–4.96)	4.81 (4.73–4.90)
When you wake up in the morning still feel sleepy and tired	3.14 (3.03–3.24)	3.23 (3.07–3.39)	3.35 (3.18–3.53)	3.28 (3.11–3.45)
Wake up refreshed in the morning	2.84 (2.75–2.93)	2.86 (2.70–3.02)	2.74 (2.57–2.92)	2.61 (2.46–2.76)
Feel sticky in the mouth when you wake up in the morning	3.59 (3.48–3.70)	3.70 (3.54–3.87)	3.66 (3.46–3.87)	3.47 (3.29–3.66)
Not hungry when you wake up in the morning.	3.50 (3.39–3.61)	3.50 (3.33–3.66)	3.52 (3.32–3.72)	3.44 (3.27–3.62)
Become drowsy after a meal	2.82 (2.71–2.93)	2.98 (2.83–3.13)	2.97 (2.77–3.17)	3.04 (2.88–3.19)
Wake up late on a non-working day	3.48 (3.38–3.59)	3.47 (3.31–3.62)	3.48 (3.28–3.68)	3.69 (3.51–3.87)
You cannot wake up on a non-working day without an alarm clock.	4.03 (3.92–4.14)	4.00 (3.84–4.17)	4.21 (4.02–4.41)	4.21 (4.03–4.38)
Take a nap on a non-working day	3.62 (3.51–3.72)	3.54 (3.40–3.68)	3.74 (3.56–3.91)	3.49 (3.32–3.67)
Have habits to sleep well	1.83 (1.79–1.86)	1.86 (1.82–1.91)	1.85 (1.79–1.91)	1.81 (1.75–1.86)

Have been told that you have stopped breathing during sleep	1.94 (1.91–1.96)	1.89 (1.85–1.94)	1.94 (1.90–1.98)	1.88 (1.84–1.93)
Have been told you snore	1.48 (1.43–1.52)	1.40 (1.34–1.47)	1.47 (1.38–1.55)	1.30 (1.23–1.37)
Have been told you are bruxing your teeth during sleeping	1.68 (1.63–1.72)	1.69 (1.63–1.75)	1.77 (1.70–1.85)	1.68 (1.61–1.75)
◆ Frequency and duration				
Frequency of eating breakfast per week	1.67 (1.57–1.78)	1.57 (1.44–1.70)	1.74 (1.55–1.94)	1.57 (1.42–1.73)
Time required for breakfast	3.25 (3.16–3.33)	3.18 (3.06–3.31)	3.15 (2.99–3.31)	3.25 (3.12–3.38)
Frequency of eating lunch per week	1.20 (1.15–1.25)	1.16 (1.10–1.22)	1.33 (1.20–1.46)	1.23 (1.15–1.31)
Time required for lunch	2.48 (2.40–2.56)	2.39 (2.28–2.51)	2.51 (2.37–2.66)	2.52 (2.39–2.65)
Frequency of eating supper per week	1.08 (1.05–1.12)	1.08 (1.03–1.12)	1.13 (1.04–1.22)	1.16 (1.09–1.24)
Time required for supper	1.97 (1.89–2.05)	1.81 (1.70–1.92)	1.86 (1.71–2.01)	1.92 (1.79–2.04)
Frequency of eating after-supper snacks per week	4.37 (4.27–4.47)	4.35 (4.20–4.50)	4.47 (4.30–4.64)	4.49 (4.36–4.63)
Time required for after-supper snacks	3.21 (3.14–3.28)	3.06 (2.98–3.15)	3.11 (2.98–3.23)	3.13 (3.04–3.23)
Frequency of snacking per week (between breakfast and lunch or between lunch and dinner)	3.40 (3.27–3.53)	3.24 (3.04–3.44)	3.56 (3.32–3.79)	3.63 (3.43–3.82)
Time required for snacks (between breakfast and lunch or between lunch and dinner)	4.07 (4.00–4.14)	3.98 (3.87–4.09)	4.05 (3.93–4.18)	4.04 (3.95–4.13)
◆ Meal timing				
Eat breakfast at about the same time	1.05 (1.03–1.07)	1.09 (1.05–1.13)	1.05 (1.01–1.08)	1.08 (1.04–1.12)
Eat lunch at about the same time	1.13 (1.10–1.16)	1.11 (1.06–1.15)	1.11 (1.05–1.16)	1.13 (1.08–1.18)
Eat supper at about the same time	1.19 (1.16–1.23)	1.15 (1.10–1.20)	1.16 (1.10–1.22)	1.14 (1.09–1.19)
Eat after-supper snacks at about the same time	1.13 (1.10–1.17)	1.13 (1.08–1.17)	1.14 (1.08–1.20)	1.12 (1.07–1.17)
Eat snacks (between breakfast and lunch or between lunch and dinner) at about the same time	1.31 (1.27–1.36)	1.27 (1.21–1.33)	1.29 (1.21–1.36)	1.25 (1.19–1.31)
◆ Dietary habits				
Frequency of consuming yogurt and probiotics drinks	3.70 (3.58–3.83)	3.51 (3.32–3.71)	3.63 (3.39–3.87)	3.55 (3.34–3.75)
Which do you usually eat: rice or bread?	1.99 (1.88–2.11)	2.15 (1.97–2.33)	1.81 (1.62–2.00)	2.14 (1.95–2.32)

Amount of rice you usually eat compared to one serving of rice at a restaurant	3.22 (3.12–3.31)	3.17 (3.04–3.31)	3.11 (2.92–3.29)	3.21 (3.05–3.36)
Amount of the main dish you usually eat compared to one serving of the main dish at a restaurant	3.01 (2.92–3.10)	2.92 (2.80–3.04)	2.81 (2.65–2.97)	2.84 (2.71–2.97)
Frequency of soup intake	3.14 (3.04–3.24)	3.20 (3.07–3.33)	3.01 (2.84–3.18)	3.11 (2.96–3.26)
Frequency of noodles intake	3.22 (3.12–3.31)	3.13 (2.99–3.26)	3.16 (2.99–3.33)	3.18 (3.04–3.33)
Frequency of vegetable intake	2.85 (2.76–2.95)	2.93 (2.80–3.07)	2.84 (2.67–3.01)	2.79 (2.64–2.94)
Amount of soup left in a noodle dish	2.95 (2.83–3.08)	3.06 (2.88–3.25)	2.55 (2.32–2.78)	2.87 (2.67–3.07)
◆ Characteristics of your diet				
Try to eat a well-balanced diet	1.24 (1.20–1.28)	1.23 (1.17–1.29)	1.18 (1.11–1.25)	1.20 (1.14–1.26)
Try not to overeat	1.46 (1.41–1.50)	1.47 (1.40–1.53)	1.49 (1.40–1.57)	1.40 (1.33–1.47)
Try to choose a low-carbohydrate diet	1.68 (1.64–1.72)	1.62 (1.55–1.69)	1.65 (1.57–1.74)	1.62 (1.55–1.69)
faster to eat than others	1.54 (1.50–1.59)	1.44 (1.37–1.51)	1.54 (1.46–1.63)	1.44 (1.37–1.51)
Try to choose a low-salt diet	1.54 (1.49–1.58)	1.49 (1.42–1.56)	1.61 (1.53–1.69)	1.55 (1.48–1.62)
Try to eat vegetables first	1.27 (1.23–1.31)	1.31 (1.25–1.37)	1.26 (1.19–1.34)	1.26 (1.20–1.32)
Often eat sweets	1.41 (1.36–1.45)	1.41 (1.34–1.48)	1.51 (1.43–1.60)	1.51 (1.43–1.58)
Often eat seafood, tofu, and beans.	1.40 (1.35–1.44)	1.39 (1.33–1.46)	1.35 (1.27–1.44)	1.32 (1.25–1.39)
Eat more than other people	1.64 (1.60–1.69)	1.61 (1.54–1.67)	1.63 (1.55–1.71)	1.55 (1.48–1.62)
Often have a large serving and extra serving	1.76 (1.72–1.80)	1.78 (1.72–1.84)	1.77 (1.69–1.84)	1.76 (1.70–1.82)
Feel unsatisfied without rice dishes or noodles after drinking alcohol	1.83 (1.80–1.87)	1.87 (1.82–1.92)	1.81 (1.75–1.88)	1.85 (1.80–1.90)
Finish supper two hours before going to bed	1.29 (1.25–1.33)	1.29 (1.23–1.35)	1.28 (1.20–1.36)	1.23 (1.17–1.29)
Meals served one by one	1.28 (1.24–1.33)	1.29 (1.22–1.35)	1.22 (1.15–1.29)	1.24 (1.18–1.30)
Often eat after-meal dessert	1.20 (1.16–1.23)	1.21 (1.16–1.27)	1.22 (1.15–1.29)	1.27 (1.21–1.34)
Often add salt or soy sauce to dishes.	1.82 (1.79–1.86)	1.84 (1.79–1.89)	1.80 (1.74–1.87)	1.76 (1.70–1.82)
Try to choose meat with less fat and remove excess fat.	1.52 (1.48–1.57)	1.54 (1.48–1.61)	1.59 (1.51–1.68)	1.58 (1.51–1.65)
Often eat until you're full.	1.43 (1.38–1.47)	1.35 (1.28–1.42)	1.44 (1.35–1.52)	1.48 (1.41–1.55)
Often eat rice dishes or noodles, or sweets after drinking alcohol.	1.74 (1.70–1.78)	1.76 (1.70–1.82)	1.79 (1.72–1.86)	1.77 (1.71–1.83)
Hardly ever cook at home	1.65 (1.61–1.69)	1.65 (1.58–1.72)	1.72 (1.64–1.80)	1.67 (1.61–1.74)
Try to chew well and eat	1.48 (1.44–1.53)	1.48 (1.41–1.55)	1.55 (1.46–1.63)	1.51 (1.43–1.58)

After supper, go to bed on a full stomach.	1.75 (1.71–1.79)	1.71 (1.65–1.78)	1.71 (1.63–1.78)	1.74 (1.68–1.80)
Try to eat a lot of vegetables	1.35 (1.31–1.39)	1.33 (1.27–1.40)	1.32 (1.24–1.40)	1.31 (1.24–1.37)
◆ Meal details (days of intake per week)				
Rice ball	1.83 (1.71–1.95)	1.92 (1.74–2.11)	1.80 (1.58–2.02)	1.75 (1.57–1.94)
Rice porridge and risotto	0.64 (0.57–0.71)	0.70 (0.59–0.82)	0.59 (0.46–0.71)	0.69 (0.57–0.81)
Japanese seasoned rice with vegetables	0.80 (0.73–0.87)	0.83 (0.73–0.93)	0.82 (0.70–0.94)	0.84 (0.75–0.94)
Rice (white rice, brown rice, barley rice, etc.)	4.24 (4.15–4.34)	4.19 (4.04–4.33)	4.35 (4.18–4.52)	4.08 (3.93–4.24)
Fried rice, pilaf, and omelet rice	1.34 (1.26–1.42)	1.24 (1.13–1.36)	1.26 (1.11–1.41)	1.27 (1.15–1.39)
Sushi	1.00 (0.94–1.07)	1.01 (0.91–1.10)	1.04 (0.91–1.16)	0.94 (0.85–1.04)
Bowl of rice with a topping	1.38 (1.29–1.46)	1.38 (1.26–1.51)	1.37 (1.22–1.51)	1.26 (1.14–1.38)
Curry rice, hashed beef rice	1.31 (1.25–1.38)	1.23 (1.14–1.33)	1.40 (1.28–1.51)	1.26 (1.16–1.36)
Plain bread	2.22 (2.08–2.36)	2.22 (2.01–2.44)	2.13 (1.86–2.40)	2.18 (1.95–2.40)
Bread rolls, croissants, French bread	1.15 (1.05–1.24)	1.17 (1.02–1.33)	1.11 (0.94–1.27)	1.18 (1.03–1.33)
Sweet bread, side dish bread, sandwiches, hamburgers, and pizza.	2.03 (1.91–2.14)	1.88 (1.72–2.04)	1.83 (1.63–2.02)	1.87 (1.69–2.04)
Noodles (ramen, udon, soba, pasta, etc.)	2.11 (2.02–2.20)	2.23 (2.08–2.37)	2.28 (2.12–2.43)	2.03 (1.89–2.18)
Okonomiyaki (savory Japanese style of pancake.)	0.78 (0.71–0.84)	0.75 (0.65–0.85)	0.73 (0.62–0.84)	0.75 (0.64–0.85)
Cereal	0.82 (0.70–0.94)	0.79 (0.61–0.97)	0.71 (0.50–0.93)	0.69 (0.53–0.86)
Japanese hot pot dish	3.32 (3.18–3.45)	3.09 (2.89–3.30)	3.60 (3.38–3.83)	3.39 (3.19–3.60)
Stew	0.76 (0.70–0.83)	0.69 (0.59–0.78)	0.70 (0.59–0.80)	0.73 (0.63–0.84)
Other soup	1.03 (0.95–1.11)	0.91 (0.78–1.04)	1.03 (0.87–1.19)	1.02 (0.89–1.14)
Simmered Food (main dishes: pot-au-feu, meat potato, etc.)	1.87 (1.77–1.97)	1.90 (1.73–2.06)	1.86 (1.67–2.05)	1.87 (1.71–2.04)
Simmered Food (side dishes: hijiki seaweed, dried radish strips, etc.)	1.79 (1.69–1.90)	1.70 (1.54–1.86)	1.68 (1.48–1.87)	1.74 (1.57–1.92)
Deep-fried food	2.14 (2.05–2.23)	2.08 (1.94–2.21)	2.23 (2.06–2.40)	2.15 (2.02–2.29)
Grilled dish (grilled fish, meat, hamburger steak, fried egg, etc.)	2.57 (2.49–2.66)	2.52 (2.39–2.66)	2.80 (2.65–2.96)	2.69 (2.55–2.83)
Stir-fried food (stir-fried vegetables, mapo tofu, roasted eggs, etc.)	2.56 (2.47–2.64)	2.45 (2.31–2.59)	2.56 (2.41–2.71)	2.47 (2.34–2.60)
Boiled and steamed food (hot vegetables, boiled eggs, chawanmushi, shumai, etc.)	1.91 (1.81–2.01)	1.90 (1.74–2.06)	2.08 (1.89–2.28)	1.88 (1.71–2.04)

Raw food (sashimi, natto, tofu, cod roe, etc.)	2.43 (2.32–2.55)	2.46 (2.29–2.63)	2.63 (2.41–2.85)	2.42 (2.25–2.59)
Salad, raw vegetables	3.29 (3.18–3.40)	3.25 (3.08–3.43)	3.44 (3.25–3.64)	3.35 (3.17–3.54)
Vinegared food	1.18 (1.08–1.28)	1.21 (1.05–1.37)	1.26 (1.06–1.46)	1.32 (1.16–1.48)
Ohitashi (boiled spinach seasoned with soy sauce)	1.32 (1.22–1.43)	1.27 (1.12–1.43)	1.32 (1.15–1.50)	1.30 (1.14–1.46)
Japanese pickles	1.61 (1.49–1.73)	1.59 (1.41–1.77)	1.68 (1.44–1.93)	1.61 (1.41–1.81)
Cheese, dried fruits, nuts	2.44 (2.30–2.57)	2.57 (2.35–2.79)	2.49 (2.22–2.76)	2.47 (2.27–2.67)
Fruits (except berries)	1.99 (1.86–2.11)	1.98 (1.77–2.19)	1.76 (1.53–1.99)	2.07 (1.87–2.27)
Fruits (berries)	1.35 (1.23–1.46)	1.43 (1.27–1.60)	1.25 (1.05–1.45)	1.35 (1.19–1.51)
Yogurt, pudding, jelly, ice cream	2.41 (2.27–2.54)	2.44 (2.24–2.64)	2.32 (2.05–2.60)	2.46 (2.24–2.68)
Cakes, pies, tarts	1.19 (1.11–1.27)	1.16 (1.04–1.28)	1.12 (0.96–1.28)	1.02 (0.91–1.13)
Traditional Japanese sweets	1.11 (1.03–1.18)	1.13 (1.00–1.25)	1.07 (0.91–1.23)	1.11 (0.98–1.24)
Fried confectionery	2.14 (2.02–2.25)	2.27 (2.10–2.44)	2.18 (1.96–2.40)	2.10 (1.90–2.30)
◆ Drink intake (after waking up-before breakfast)				
Tea (frequency)	1.78 (1.59–1.96)	1.85 (1.57–2.14)	1.59 (1.23–1.94)	1.69 (1.39–1.99)
Water (frequency)	3.03 (2.83–3.22)	3.19 (2.90–3.47)	3.05 (2.69–3.42)	3.23 (2.92–3.53)
Coffee (frequency)	1.81 (1.61–2.01)	2.05 (1.75–2.34)	1.87 (1.50–2.25)	1.84 (1.54–2.14)
Milk and yogurt drink (frequency)	1.29 (1.13–1.45)	1.21 (0.99–1.44)	0.92 (0.66–1.18)	1.10 (0.87–1.33)
Fruit and vegetable drinks (frequency)	0.77 (0.65–0.88)	0.86 (0.67–1.05)	0.56 (0.37–0.74)	0.82 (0.62–1.02)
Other soft drinks (frequency)	0.66 (0.55–0.77)	0.64 (0.48–0.80)	0.62 (0.42–0.83)	0.54 (0.38–0.69)
◆ Drink intake (at breakfast)				
Tea (frequency)	1.82 (1.63–2.01)	1.97 (1.68–2.25)	1.86 (1.51–2.21)	1.82 (1.52–2.12)
Water (frequency)	2.11 (1.91–2.32)	2.14 (1.83–2.44)	1.92 (1.56–2.29)	2.03 (1.72–2.35)
Coffee (frequency)	2.23 (2.03–2.43)	2.42 (2.13–2.72)	2.44 (2.07–2.81)	2.33 (2.03–2.62)
Milk and yogurt drink (frequency)	1.39 (1.22–1.55)	1.53 (1.29–1.78)	1.38 (1.08–1.67)	1.35 (1.12–1.59)
Fruit and vegetable drinks (frequency)	0.79 (0.67–0.91)	0.90 (0.71–1.09)	0.68 (0.48–0.89)	0.99 (0.78–1.20)
Other soft drinks (frequency)	0.40 (0.31–0.50)	0.43 (0.29–0.56)	0.38 (0.22–0.54)	0.44 (0.28–0.59)
◆ Drink intake (after breakfast-before lunch)				
Tea (frequency)	2.00 (1.82–2.19)	2.28 (2.00–2.56)	1.94 (1.60–2.28)	2.29 (2.00–2.59)
Water (frequency)	2.40 (2.20–2.60)	2.37 (2.08–2.66)	2.24 (1.88–2.60)	2.46 (2.15–2.77)
Coffee (frequency)	2.06 (1.87–2.24)	2.02 (1.75–2.30)	2.10 (1.74–2.45)	2.01 (1.72–2.29)
Milk and yogurt drink (frequency)	0.62 (0.50–0.74)	0.59 (0.41–0.76)	0.49 (0.29–0.69)	0.66 (0.49–0.84)

Fruit and vegetable drinks (frequency)	0.44 (0.35–0.54)	0.50 (0.36–0.64)	0.37 (0.23–0.50)	0.59 (0.43–0.75)
Other soft drinks (frequency)	0.50 (0.40–0.60)	0.56 (0.41–0.71)	0.64 (0.44–0.84)	0.69 (0.51–0.87)
◆ Drink intake (at lunch)				
Tea (frequency)	2.71 (2.53–2.88)	3.08 (2.82–3.35)	2.92 (2.60–3.25)	2.82 (2.55–3.09)
Tea (amount)	1.55 (1.44–1.65)	1.72 (1.56–1.88)	1.70 (1.51–1.89)	1.56 (1.40–1.73)
Water (frequency)	2.27 (2.08–2.45)	2.39 (2.11–2.67)	2.21 (1.88–2.55)	2.29 (2.01–2.58)
Coffee (frequency)	1.37 (1.21–1.53)	1.40 (1.15–1.64)	1.41 (1.13–1.70)	1.29 (1.06–1.53)
Milk and yogurt drink (frequency)	0.49 (0.39–0.60)	0.49 (0.34–0.64)	0.53 (0.34–0.71)	0.61 (0.44–0.78)
Fruit and vegetable drinks (frequency)	0.49 (0.40–0.58)	0.45 (0.31–0.58)	0.48 (0.31–0.65)	0.62 (0.46–0.78)
Other soft drinks (frequency)	0.49 (0.40–0.59)	0.49 (0.34–0.63)	0.49 (0.32–0.65)	0.57 (0.42–0.73)
◆ Drink intake (after lunch-before dinner)				
Tea (frequency)	2.29 (2.12–2.47)	2.51 (2.25–2.78)	2.15 (1.82–2.48)	2.38 (2.12–2.65)
Coffee (frequency)	2.23 (2.06–2.41)	2.22 (1.96–2.48)	2.38 (2.05–2.72)	2.11 (1.83–2.39)
Milk and yogurt drink (frequency)	0.58 (0.47–0.68)	0.58 (0.41–0.75)	0.41 (0.25–0.58)	0.54 (0.38–0.70)
Fruit and vegetable drinks (frequency)	0.44 (0.37–0.52)	0.48 (0.35–0.60)	0.41 (0.27–0.56)	0.53 (0.38–0.67)
Other soft drinks (frequency)	0.72 (0.61–0.83)	0.70 (0.54–0.86)	0.76 (0.55–0.97)	0.78 (0.60–0.96)
◆ Drink intake (at supper)				
Tea (frequency)	2.42 (2.23–2.61)	2.73 (2.45–3.02)	2.58 (2.20–2.95)	2.41 (2.10–2.71)
Water (frequency)	2.03 (1.84–2.22)	2.13 (1.85–2.42)	2.03 (1.68–2.38)	2.05 (1.75–2.34)
Coffee (frequency)	0.47 (0.36–0.58)	0.47 (0.30–0.64)	0.38 (0.19–0.56)	0.58 (0.40–0.77)
Milk and yogurt drink (frequency)	0.29 (0.22–0.37)	0.28 (0.17–0.40)	0.26 (0.12–0.39)	0.35 (0.21–0.49)
Fruit and vegetable drinks (frequency)	0.31 (0.24–0.39)	0.34 (0.21–0.48)	0.31 (0.16–0.46)	0.35 (0.23–0.47)
Other soft drinks (frequency)	0.52 (0.42–0.63)	0.44 (0.29–0.58)	0.48 (0.30–0.66)	0.54 (0.37–0.70)
◆ Drink intake (after supper to 30 minutes before bedtime)				
Tea (frequency)	1.98 (1.79–2.16)	1.87 (1.61–2.13)	1.78 (1.43–2.13)	1.68 (1.40–1.96)
Water (frequency)	2.49 (2.30–2.68)	2.61 (2.32–2.90)	2.14 (1.77–2.50)	2.38 (2.08–2.69)
Coffee (frequency)	0.93 (0.78–1.07)	0.99 (0.76–1.22)	0.86 (0.60–1.12)	0.99 (0.74–1.23)
Milk and yogurt drink (frequency)	0.56 (0.46–0.67)	0.62 (0.44–0.79)	0.37 (0.21–0.53)	0.48 (0.33–0.63)
Fruit and vegetable drinks (frequency)	0.37 (0.29–0.44)	0.35 (0.23–0.48)	0.35 (0.21–0.49)	0.36 (0.24–0.49)
Other soft drinks (frequency)	0.56 (0.46–0.67)	0.48 (0.34–0.62)	0.62 (0.42–0.83)	0.59 (0.43–0.75)
◆ Drink intake (30 minutes before bedtime to bedtime)				
Tea (frequency)	0.85 (0.70–1.00)	0.74 (0.54–0.95)	0.94 (0.65–1.23)	0.67 (0.46–0.88)

Water (frequency)	2.04 (1.85–2.24)	2.06 (1.77–2.35)	2.06 (1.69–2.43)	2.13 (1.82–2.43)
Coffee (frequency)	0.27 (0.19–0.35)	0.25 (0.13–0.38)	0.16 (0.04–0.28)	0.28 (0.15–0.42)
Milk and yogurt drink (frequency)	0.27 (0.18–0.35)	0.17 (0.08–0.27)	0.14 (0.05–0.23)	0.24 (0.12–0.36)
Fruit and vegetable drinks (frequency)	0.13 (0.08–0.17)	0.11 (0.04–0.17)	0.17 (0.06–0.27)	0.16 (0.07–0.24)
Other soft drinks (frequency)	0.25 (0.18–0.32)	0.22 (0.12–0.32)	0.31 (0.14–0.47)	0.23 (0.12–0.34)
◆ Water intake habits				
Try to drink water even if you aren't thirsty	1.57 (1.52–1.62)	1.57 (1.50–1.64)	1.64 (1.56–1.72)	1.54 (1.47–1.61)
Try to drink when you feel thirsty	1.64 (1.59–1.68)	1.63 (1.56–1.70)	1.65 (1.56–1.73)	1.67 (1.61–1.74)
Not particularly conscious	1.79 (1.75–1.83)	1.80 (1.75–1.86)	1.71 (1.64–1.79)	1.79 (1.73–1.85)
How do you drink water	1.25 (1.21–1.29)	1.24 (1.18–1.30)	1.32 (1.24–1.40)	1.29 (1.22–1.35)
Your water intake compared to that of others	1.90 (1.83–1.96)	1.95 (1.86–2.04)	1.85 (1.73–1.97)	1.74 (1.65–1.84)
Wake up and drink water while in bed	2.70 (2.65–2.75)	2.77 (2.69–2.84)	2.60 (2.49–2.71)	2.62 (2.53–2.71)
Bring drinks from home when going out for long periods of time	1.85 (1.78–1.92)	1.84 (1.73–1.96)	1.89 (1.74–2.03)	1.81 (1.69–1.92)
Put sugar in coffee and tea	2.42 (2.35–2.49)	2.41 (2.31–2.52)	2.29 (2.14–2.43)	2.51 (2.40–2.62)
◆ Alcohol intake				
Beer (frequency)	1.06 (0.93–1.18)	0.97 (0.78–1.15)	1.29 (1.02–1.57)	1.26 (1.04–1.48)
Chu-hai and sour (frequency)	0.89 (0.78–1.00)	0.82 (0.66–0.98)	1.20 (0.93–1.46)	0.98 (0.78–1.17)
Shochu and Awamori (frequency)	0.25 (0.19–0.32)	0.15 (0.07–0.23)	0.38 (0.21–0.56)	0.53 (0.37–0.69)
Sake (frequency)	0.26 (0.20–0.32)	0.22 (0.13–0.30)	0.19 (0.09–0.29)	0.24 (0.15–0.32)
Whiskey, brandy, gin, vodka (frequency)	0.22 (0.16–0.28)	0.12 (0.06–0.18)	0.31 (0.15–0.47)	0.29 (0.18–0.40)
Wine (frequency)	0.51 (0.43–0.60)	0.52 (0.39–0.66)	0.42 (0.27–0.57)	0.49 (0.37–0.62)
◆ Alcohol intake habits				
Have many opportunities to drink alcohol	1.82 (1.78–1.85)	1.82 (1.76–1.87)	1.75 (1.68–1.83)	1.76 (1.70–1.82)
Drink alcohol during the day on a non-working day	1.92 (1.89–1.94)	1.87 (1.82–1.92)	1.88 (1.82–1.94)	1.88 (1.83–1.93)
Overeat when drinking alcohol	1.87 (1.84–1.90)	1.83 (1.78–1.89)	1.83 (1.76–1.89)	1.86 (1.81–1.91)
Eat low-calorie snacks when drinking alcohol	1.67 (1.62–1.71)	1.68 (1.62–1.75)	1.78 (1.71–1.85)	1.65 (1.58–1.72)
Not drink alcohol more than two days a week	1.23 (1.19–1.26)	1.23 (1.17–1.29)	1.35 (1.26–1.43)	1.21 (1.15–1.27)
◆ Regarding your family (grandparents, parents, siblings)				
Someone in your family has/had diabetes	1.73 (1.69–1.77)	1.63 (1.56–1.70)	1.76 (1.69–1.83)	1.66 (1.60–1.73)
Someone in your family is/was obesity	1.77 (1.73–1.81)	1.77 (1.71–1.83)	1.74 (1.66–1.81)	1.72 (1.66–1.79)

Someone in your family has/had dementia	1.85 (1.81–1.88)	1.82 (1.77–1.87)	1.86 (1.81–1.92)	1.79 (1.74–1.85)
Someone in your family has/had thinning hair	1.66 (1.61–1.70)	1.65 (1.58–1.72)	1.54 (1.46–1.63)	1.62 (1.55–1.69)
◆ Physical condition				
Feel energetic	2.13 (2.05–2.22)	2.16 (2.03–2.28)	2.20 (2.04–2.35)	2.20 (2.07–2.33)
Get tired easily	3.23 (3.14–3.33)	3.14 (3.00–3.28)	3.49 (3.32–3.66)	3.39 (3.24–3.55)
Feel out of breath easily	4.52 (4.45–4.60)	4.29 (4.15–4.42)	4.53 (4.39–4.68)	4.35 (4.22–4.49)
Feel your heart pounding	4.65 (4.59–4.71)	4.60 (4.50–4.71)	4.67 (4.55–4.79)	4.64 (4.54–4.74)
Prefer a quiet environment and find it troublesome to talk to others	3.70 (3.59–3.81)	3.61 (3.45–3.77)	4.01 (3.83–4.18)	3.86 (3.70–4.02)
Nausea and abdominal bloating	4.34 (4.26–4.42)	4.18 (4.05–4.32)	4.52 (4.39–4.65)	4.27 (4.14–4.40)
Cold hands and feet even in warm places	4.35 (4.25–4.45)	4.31 (4.16–4.45)	4.52 (4.36–4.68)	4.55 (4.41–4.69)
Coldness in areas other than hands and feet (back, abdomen, hips, knees, etc.)	4.44 (4.35–4.54)	4.41 (4.28–4.54)	4.59 (4.43–4.76)	4.57 (4.44–4.69)
sweat even though not doing anything	4.63 (4.56–4.71)	4.59 (4.49–4.68)	4.71 (4.60–4.81)	4.51 (4.37–4.64)
Greasy forehead	4.16 (4.06–4.27)	3.83 (3.67–3.99)	4.12 (3.92–4.32)	3.91 (3.73–4.08)
Dry skin and lips	3.66 (3.55–3.77)	3.71 (3.56–3.86)	3.60 (3.39–3.82)	3.73 (3.54–3.91)
Greasy nose	3.77 (3.66–3.88)	3.37 (3.19–3.55)	3.74 (3.53–3.96)	3.63 (3.45–3.81)
Acne and pimples	4.13 (4.03–4.23)	4.00 (3.85–4.14)	4.32 (4.15–4.49)	4.29 (4.16–4.43)
Feel thirsty	3.57 (3.48–3.66)	3.45 (3.31–3.60)	3.61 (3.44–3.78)	3.62 (3.48–3.77)
Diarrhea after intake of cold food	4.46 (4.37–4.54)	4.37 (4.24–4.50)	4.39 (4.24–4.54)	4.52 (4.41–4.64)
Sticky stool and feeling of incomplete defecation	4.42 (4.34–4.49)	4.36 (4.23–4.49)	4.58 (4.46–4.70)	4.42 (4.30–4.53)
Hard stool and constipation	4.02 (3.91–4.12)	3.94 (3.77–4.10)	4.11 (3.92–4.29)	4.28 (4.14–4.43)
Bleed when brushing teeth	4.15 (4.06–4.24)	4.09 (3.96–4.23)	4.21 (4.03–4.39)	4.19 (4.05–4.34)
Gingival recession and wide tooth gaps	3.96 (3.84–4.07)	3.96 (3.79–4.13)	3.99 (3.78–4.20)	3.86 (3.67–4.05)
Food or drink stings your teeth.	4.29 (4.20–4.37)	4.22 (4.07–4.37)	4.49 (4.35–4.62)	4.24 (4.09–4.39)
◆Your constitution				
Easy to get fat	1.51 (1.47–1.56)	1.30 (1.24–1.36)	1.56 (1.48–1.65)	1.39 (1.32–1.46)
Hard to build muscle even after exercising	1.56 (1.51–1.60)	1.50 (1.43–1.56)	1.59 (1.51–1.68)	1.53 (1.46–1.60)
Have been on a diet	1.35 (1.31–1.39)	1.24 (1.18–1.30)	1.41 (1.32–1.49)	1.32 (1.25–1.39)
Cannot go through with a diet	1.25 (1.21–1.29)	1.17 (1.12–1.22)	1.20 (1.13–1.27)	1.16 (1.11–1.22)
Regained weight after a diet	1.26 (1.22–1.30)	1.22 (1.16–1.28)	1.23 (1.16–1.31)	1.18 (1.13–1.24)

Gained more than 10 kg after the age of 20	1.73 (1.68–1.77)	1.48 (1.41–1.54)	1.66 (1.58–1.74)	1.48 (1.41–1.55)
Gained more than 10 kg after the age of 20	1.62 (1.58–1.67)	1.52 (1.46–1.59)	1.64 (1.56–1.72)	1.59 (1.52–1.66)
Weight has changed by more than 3 kg in the last year	1.38 (1.34–1.43)	1.37 (1.31–1.44)	1.36 (1.28–1.44)	1.27 (1.21–1.34)
Go to the bathroom more often than others	1.59 (1.54–1.63)	1.59 (1.52–1.66)	1.54 (1.46–1.63)	1.52 (1.45–1.59)
Lower abdomen sticks out	1.56 (1.52–1.61)	1.41 (1.34–1.48)	1.44 (1.36–1.53)	1.47 (1.40–1.55)
Concerned about hair loss	1.69 (1.65–1.73)	1.61 (1.54–1.68)	1.62 (1.54–1.71)	1.52 (1.45–1.59)
Hair getting thinner	1.67 (1.63–1.71)	1.58 (1.51–1.65)	1.54 (1.46–1.63)	1.50 (1.43–1.57)
Flabby belly	1.54 (1.50–1.59)	1.33 (1.27–1.39)	1.50 (1.42–1.59)	1.36 (1.29–1.43)
Regular teeth	1.48 (1.43–1.52)	1.46 (1.39–1.53)	1.49 (1.40–1.57)	1.52 (1.45–1.59)
Tooth decay	1.90 (1.87–1.93)	1.94 (1.91–1.97)	1.89 (1.84–1.95)	1.95 (1.92–1.98)
Periodontal disease or alveolar pyorrhea	1.79 (1.75–1.83)	1.78 (1.72–1.83)	1.71 (1.63–1.78)	1.73 (1.66–1.79)
Often wear warm clothes in winter	1.62 (1.58–1.67)	1.66 (1.59–1.72)	1.65 (1.57–1.74)	1.73 (1.67–1.79)
Sensitive to cold	1.49 (1.44–1.54)	1.56 (1.49–1.63)	1.57 (1.49–1.66)	1.63 (1.56–1.70)
◆ Workstyle				
Workstyle (full-time work)	1.39 (1.35–1.44)	1.44 (1.37–1.51)	1.36 (1.28–1.44)	1.43 (1.36–1.50)
Workstyle (part-time work)	1.77 (1.73–1.81)	1.70 (1.64–1.77)	1.78 (1.71–1.85)	1.75 (1.69–1.81)
Did you work?	1.14 (1.11–1.17)	1.14 (1.09–1.18)	1.13 (1.07–1.19)	1.17 (1.11–1.22)
Did you have a night shift?	1.96 (1.94–1.98)	1.97 (1.95–1.99)	1.98 (1.96–2.01)	1.97 (1.95–2.00)
Job description	1.40 (1.34–1.46)	1.32 (1.23–1.40)	1.40 (1.28–1.52)	1.36 (1.27–1.45)
Requires complicated thinking	1.60 (1.56–1.65)	1.57 (1.50–1.64)	1.61 (1.53–1.69)	1.58 (1.51–1.65)
Overtime in a month	4.67 (4.60–4.74)	4.62 (4.50–4.74)	4.64 (4.51–4.77)	4.65 (4.53–4.76)
Often go home after 20:00	1.70 (1.66–1.74)	1.75 (1.69–1.81)	1.71 (1.64–1.79)	1.76 (1.70–1.82)
◆ Oral hygiene habits				
Select all the times you brush your teeth (after waking up)	1.63 (1.58–1.67)	1.61 (1.54–1.67)	1.53 (1.45–1.62)	1.55 (1.48–1.62)
Select all the times you brush your teeth (after breakfast)	1.40 (1.35–1.44)	1.46 (1.39–1.53)	1.50 (1.42–1.59)	1.42 (1.35–1.49)
Select all the times you brush your teeth (after lunch)	1.58 (1.53–1.63)	1.67 (1.61–1.74)	1.67 (1.59–1.75)	1.62 (1.55–1.69)
Select all the times you brush your teeth (within 1 hour after supper)	1.77 (1.74–1.81)	1.82 (1.76–1.87)	1.81 (1.75–1.88)	1.75 (1.69–1.81)

Select all the times you brush your teeth (Before going to bed)	1.23 (1.19–1.27)	1.24 (1.18–1.30)	1.25 (1.17–1.32)	1.32 (1.25–1.39)
Use dental floss or an interdental brush	3.82 (3.68–3.96)	3.58 (3.36–3.80)	3.61 (3.32–3.90)	3.69 (3.46–3.92)
Use mouth rinse	4.01 (3.87–4.15)	4.00 (3.79–4.20)	3.83 (3.56–4.10)	3.77 (3.54–4.01)
Frequency of toothbrush replacement	2.61 (2.50–2.72)	2.67 (2.50–2.84)	2.52 (2.32–2.72)	2.45 (2.28–2.61)
◆ Lifestyle				
Walk faster than others	1.33 (1.29–1.38)	1.43 (1.36–1.50)	1.37 (1.29–1.45)	1.34 (1.27–1.40)
Often need to take stairs	1.52 (1.48–1.57)	1.58 (1.51–1.65)	1.59 (1.50–1.67)	1.58 (1.51–1.65)
Try to use stairs instead of an elevator or escalator	1.56 (1.51–1.61)	1.55 (1.48–1.62)	1.56 (1.48–1.65)	1.58 (1.51–1.65)
Take time to soak in a bath	3.19 (3.06–3.32)	3.11 (2.94–3.29)	3.32 (3.10–3.55)	3.47 (3.28–3.66)
Often lie down right after eating.	1.60 (1.55–1.64)	1.56 (1.49–1.63)	1.64 (1.56–1.72)	1.65 (1.58–1.72)
Often spend non-working days at home	1.60 (1.56–1.65)	1.58 (1.51–1.65)	1.60 (1.52–1.69)	1.52 (1.45–1.59)
◆ Screen time (TV, computer, tablet, smartphone, etc.)				
On a working day (if you don't work, on weekdays)	2.64 (2.53–2.75)	2.42 (2.25–2.59)	2.52 (2.31–2.73)	2.66 (2.48–2.83)
On a non-working day (if you don't work, on the weekend)	2.76 (2.67–2.86)	2.55 (2.40–2.70)	2.68 (2.50–2.85)	2.76 (2.60–2.93)
◆ Please select all the benefits and target organs of supplements that you take at least 4 times a week.				
Beauty and skin	1.94 (1.92–1.96)	1.92 (1.88–1.96)	1.96 (1.93–2.00)	1.92 (1.88–1.96)
Health maintenance and improvement	1.90 (1.87–1.93)	1.96 (1.93–1.98)	1.91 (1.86–1.96)	1.91 (1.87–1.95)
Joint	1.99 (1.99–2.00)	2.00 (2.00–2.00)	1.99 (1.98–2.01)	1.97 (1.94–1.99)
Fatigue recovery	1.95 (1.93–1.97)	1.97 (1.95–1.99)	1.93 (1.89–1.98)	1.97 (1.95–2.00)
Nutrition	1.94 (1.91–1.96)	1.97 (1.94–1.99)	1.94 (1.90–1.98)	1.90 (1.86–1.94)
Eye	1.98 (1.97–1.99)	1.99 (1.98–2.00)	1.99 (1.98–2.01)	1.97 (1.95–2.00)
Antioxidant and anti-aging	1.97 (1.96–1.99)	1.99 (1.98–2.00)	1.98 (1.95–2.00)	1.98 (1.96–2.00)
Weight loss	2.00 (1.99–2.00)	1.99 (1.98–2.00)	1.98 (1.96–2.01)	1.99 (1.98–2.01)
Slimming (becomes slim with good style)	1.99 (1.98–2.00)	2.00 (2.00–2.00)	1.99 (1.98–2.01)	1.99 (1.98–2.01)
Body fat suppression	1.99 (1.99–2.00)	1.98 (1.95–2.00)	1.98 (1.96–2.01)	1.99 (1.97–2.00)
Stiffness and pain in the neck, shoulders, and back	2 .00(1.99–2.00)	2.00 (2.00–2.00)	1.99 (1.98–2.01)	1.99 (1.98–2.01)
Bone	2 .00(1.99–2.00)	2.00 (2.00–2.00)	1.99 (1.98–2.01)	1.99 (1.97–2.00)

High blood pressure	2.00 (2.00–2.00)	2.00 (2.00–2.00)	1.99 (1.98–2.01)	1.99 (1.97–2.00)
Muscle	1.98 (1.96–1.99)	1.99 (1.97–2.00)	1.98 (1.95–2.00)	1.98 (1.97–2.00)
Bowel control and constipation	1.98 (1.97–1.99)	1.99 (1.97–2.00)	1.98 (1.96–2.01)	1.96 (1.93–1.99)
Anticoagulant	1.99 (1.98–2.00)	1.99 (1.98–2.00)	1.99 (1.98–2.01)	1.99 (1.98–2.01)
Others	1.99 (1.98–2.00)	2.00 (1.99–2.00)	2.00 (2.00–2.00)	1.99 (1.98–2.01)
Do not take supplements	1.16 (1.13–1.20)	1.16 (1.11–1.21)	1.17 (1.10–1.23)	1.23 (1.17–1.29)

C. Characteristics of the questionnaire answers for each glycometabolic category used in the external validation of the random forest model.

	Category 1	Category 2	Category 3	Category 4
Frequency of mobile phone and tablet computer use at bedtime	2.28 (2.14–2.42)	2.27 (2.09–2.46)	2.59 (2.31–2.88)	2.40 (2.14–2.66)
Do you wake up in the middle of the night?	3.83 (3.69–3.96)	3.84 (3.68–4.01)	3.97 (3.68–4.26)	3.74 (3.51–3.96)
Do you wake up late on a non-working day?	3.90 (3.77–4.03)	3.98 (3.82–4.14)	4.13 (3.76–4.49)	4.01 (3.80–4.23)
Which do you usually eat: rice or bread?	1.88 (1.73–2.03)	1.72 (1.56–1.88)	1.78 (1.34–2.22)	1.63 (1.39–1.86)
Frequency of soup intake	2.98 (2.79–3.17)	3.05 (2.80–3.30)	3.56 (3.13–3.99)	3.43 (3.12–3.74)
Frequency of tea intake per week at lunch	2.84 (2.57–3.10)	2.95 (2.59–3.31)	3.47 (2.78–4.16)	3.33 (2.90–3.77)
Frequency of toothbrush replacement	2.42 (2.28–2.57)	2.30 (2.14–2.46)	2.25 (1.91–2.59)	2.51 (2.27–2.76)

Data are presented as mean (95% confidence interval)

Acknowledgments

I would like to express my gratitude to Professor Satoshi Nakamura for giving me an opportunity to research in his laboratory. Furthermore, I thank him for his qualified guidance and advice on my research. I would like to thank Professor Keiichi Yasumoto for his qualified advice and for suggesting a new method for my research. I also would like to thank Professor Tomonori Okamura of Keio University as a thesis committee for giving valuable advice and suggestion. I am very grateful to Assistant Professor Hiroki Tanaka for his attentive guidance in my research and his support in classes and seminars. I also would like to thank all Augmented Human Communication Laboratory staff and members. Discussions with them in the lab seminar helped me promote my research. Finally, I am grateful to my co-researchers Norihito Murayama, Yuji Nonaka, Takanori Teramoto, Hiroshi Watanabe, Yumi Nakamura, and Takeshi Kanamori of Suntory Global Innovation Center Limited.

References

[1] Statistics Bureau, Ministry of Internal Affairs and Communications, Japan. Population estimates. https://www.stat.go.jp/data/jinsui/index.html (2022.10.15)

[2] Ministry of Health, Labour and Welfare, Japan. Medical expenditure trends in 2021. https://www.mhlw.go.jp/topics/medias/year/21/index.html (2022.10.15)

[3] Kleiner SM. Water: an essential but overlooked nutrient. J Am Diet Assoc 1999; 99: 200-206.

[4] Häussinger D, Lang F, Gerok W. Regulation of cell function by the cellular hydration state. Am J Physiol 1994; 267: E343-E355

[5] Jéquier E, Constant F. Water as an essential nutrient: the physiological basis of hydration. Eur J Clin Nutr 2010; 64: 115-123.

[6] Cotter JD, Thornton SN, Lee JK, Laursen PB. Are we being drowned in hydration advice? Thirsty for more? Extrem Physiol Med 2014; 3: 18.

[7] Nishikawa T, Miyamatsu N, Higashiyama A, Hojo M, Nishida Y, Fukuda S, Hirata T, Ichiura K, Kubota Y, Kubo S, Ueba T, Kadota A, Sugiyama D, Okamura T. Daily habit of water intake in patients with cerebral infarction before its onset; comparison with a healthy population: A cross-sectional study. Cerebrovasc Dis 2019; 47: 143-150.

[8] Cui R, Iso H, Eshak ES, Maruyama K, Tamakoshi A; JACC Study Group. Water intake from foods and beverages and risk of mortality from CVD: the Japan Collaborative Cohort (JACC) Study. Public Health Nutr 2018; 21: 3011-3017.

[9] Nakamura Y, Watanabe H, Tanaka A, Yasui M, Nishihira J, Murayama N. Effect of increased daily water intake and hydration on health in Japanese adults. Nutrients 2020; 12: 1191.

[10] Armstrong LE, Ganio MS, Casa DJ, Lee EC, McDermott BP, Klau JF, Jimenez L, Le Bellego L, Chevillotte E, Lieberman HR. Mild dehydration affects mood in healthy young women. J Nutr 2012; 142: 382-388.

[11] Benton D, Burgess N. The effect of the consumption of water on the memory and

attention of children. Appetite 2009; 53: 143-146.

[12] Pross N, Demazières A, Girard N, Barnouin R, Metzger D, Klein A, Perrier E, Guelinckx I. Effects of changes in water intake on mood of high and low drinkers. PLoS One 2014; 9: e94754.

[13] Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition.

[14] Gillies CL, Abrams KR, Lambert PC, Cooper NJ, Sutton AJ, Hsu RT, Khunti K. Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis. BMJ. 2007 Feb 10;334(7588):299.

[15] Crandall JP, Knowler WC, Kahn SE, Marrero D, Florez JC, Bray GA, Haffner SM, Hoskin M, Nathan DM; Diabetes Prevention Program Research Group. The prevention of type 2 diabetes. Nat Clin Pract Endocrinol Metab. 2008 Jul;4(7):382-93.

[16] Paulweber B, Valensi P, Lindström J, Lalic NM, Greaves CJ, McKee M, Kissimova-Skarbek K, Liatis S, Cosson E, Szendroedi J, Sheppard KE, Charlesworth K, Felton AM, Hall M, Rissanen A, Tuomilehto J, Schwarz PE, Roden M, Paulweber M, Stadlmayr A, Kedenko L, Katsilambros N, Makrilakis K, Kamenov Z, Evans P, Gilis-Januszewska A, Lalic K, Jotic A, Djordevic P, Dimitrijevic-Sreckovic V, Hühmer U, Kulzer B, Puhl S, Lee-Barkey YH, AlKerwi A, Abraham C, Hardeman W, Acosta T, Adler M, AlKerwi A, Barengo N, Barengo R, Boavida JM, Charlesworth K, Christov V, Claussen B, Cos X, Cosson E, Deceukelier S, Dimitrijevic-Sreckovic V, Djordjevic P, Evans P, Felton AM, Fischer M, Gabriel-Sanchez R, Gilis-Januszewska A, Goldfracht M, Gomez JL, Greaves CJ, Hall M, Handke U, Hauner H, Herbst J, Hermanns N, Herrebrugh L, Huber C, Hühmer U, Huttunen J, Jotic A, Kamenov Z, Karadeniz S, Katsilambros N, Khalangot M, Kissimova-Skarbek K, Köhler D, Kopp V, Kronsbein P, Kulzer B, Kyne-Grzebalski D, Lalic K, Lalic N, Landgraf R, Lee-Barkey YH, Liatis S, Lindström J, Makrilakis K, McIntosh C, McKee M, Mesquita AC, Misina D, Muylle F, Neumann A, Paiva AC, Pajunen P, Paulweber B, Peltonen M, Perrenoud L, Pfeiffer A, Pölönen A, Puhl S, Raposo F, Reinehr T, Rissanen A, Robinson C, Roden M, Rothe U, Saaristo T, Scholl J, Schwarz PE, Sheppard KE,

Spiers S, Stemper T, Stratmann B, Szendroedi J, Szybinski Z, Tankova T, Telle-Hjellset V, Terry G, Tolks D, Toti F, Tuomilehto J, Undeutsch A, Valadas C, Valensi P, Velickiene D, Vermunt P, Weiss R, Wens J, Yilmaz T. A European evidence-based guideline for the prevention of type 2 diabetes. Horm Metab Res. 2010 Apr;42 Suppl 1:S3-36.

[17] Lindström J, Neumann A, Sheppard KE, Gilis-Januszewska A, Greaves CJ, Handke U, Pajunen P, Puhl S, Pölönen A, Rissanen A, Roden M, Stemper T, Telle-Hjellset V, Tuomilehto J, Velickiene D, Schwarz PE, Acosta T, Adler M, AlKerwi A, Barengo N, Barengo R, Boavida JM, Charlesworth K, Christov V, Claussen B, Cos X, Cosson E, Deceukelier S, Dimitrijevic-Sreckovic V, Djordjevic P, Evans P, Felton AM, Fischer M, Gabriel-Sanchez R, Gilis-Januszewska A, Goldfracht M, Gomez JL, Greaves CJ, Hall M, Handke U, Hauner H, Herbst J, Hermanns N, Herrebrugh L, Huber C, Hühmer U, Huttunen J, Jotic A, Kamenov Z, Karadeniz S, Katsilambros N, Khalangot M, Kissimova-Skarbek K, Köhler D, Kopp V, Kronsbein P, Kulzer B, Kyne-Grzebalski D, Lalic K, Lalic N, Landgraf R, Lee-Barkey YH, Liatis S, Lindström J, Makrilakis K, McIntosh C, McKee M, Mesquita AC, Misina D, Muylle F, Neumann A, Paiva AC, Pajunen P, Paulweber B, Peltonen M, Perrenoud L, Pfeiffer A, Pölönen A, Puhl S, Raposo F, Reinehr T, Rissanen A, Robinson C, Roden M, Rothe U, Saaristo T, Scholl J, Schwarz PE, Sheppard KE, Spiers S, Stemper T, Stratmann B, Szendroedi J, Szybinski Z, Tankova T, Telle-Hjellset V, Terry G, Tolks D, Toti F, Tuomilehto J, Undeutsch A, Valadas C, Valensi P, Velickiene D, Vermunt P, Weiss R, Wens J, Yilmaz T. Take action to prevent diabetes--the IMAGE toolkit for the prevention of type 2 diabetes in Europe. Horm Metab Res. 2010 Apr;42 Suppl 1:S37-55.

[18] Saito T, Watanabe M, Nishida J, Izumi T, Omura M, Takagi T, Fukunaga R, Bandai Y, Tajima N, Nakamura Y, Ito M; Zensharen Study for Prevention of Lifestyle Diseases Group. Lifestyle modification and prevention of type 2 diabetes in overweight Japanese with impaired fasting glucose levels: a randomized controlled trial. Arch Intern Med. 2011 Aug 8;171(15):1352-60.

[19] DeFronzo RA. Pathogenesis of type 2 diabetes mellitus. Med Clin North Am. 2004 Jul;88(4):787–835, ix.

[20] Park SY, Gautier JF, Chon S. Assessment of Insulin Secretion and Insulin Resistance

in Human. Diabetes Metab J. 2021 Sep;45(5):641–654.

[21] Szoke E, Gerich JE. Role of impaired insulin secretion and insulin resistance in the pathogenesis of type 2 diabetes mellitus. Compr Ther. 2005 Summer;31(2):106–12.

[22] Boden G. Pathogenesis of type 2 diabetes. Insulin resistance. Endocrinol Metab Clin North Am. 2001 Dec;30(4):801–15, v.

[23] Iozzo P, Beck-Nielsen H, Laakso M, Smith U, Yki-Järvinen H, Ferrannini E. Independent influence of age on basal insulin secretion in nondiabetic humans. European Group for the Study of Insulin Resistance. J Clin Endocrinol Metab. 1999 Mar;84(3):863-8.

[24] Sato M, Tamura Y, Nakagata T, Someya Y, Kaga H, Yamasaki N, Kiya M, Kadowaki S, Sugimoto D, Satoh H, Kawamori R, Watada H. Prevalence and Features of Impaired Glucose Tolerance in Young Underweight Japanese Women. J Clin Endocrinol Metab. 2021 Apr 23;106(5):e2053-e2062.

[25] World Health Organization & International Diabetes Federation. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. World Health Organization; 2006. Available from: https://apps.who.int/iris/handle/10665/43588.

[26] Freedman LS, Potischman N, Kipnis V, Midthune D, Schatzkin A, Thompson FE, Troiano RP, Prentice R, Patterson R, Carroll R, Subar AF. A comparison of two dietary instruments for evaluating the fat-breast cancer relationship. Int J Epidemiol 2006; 35: 1011-1021.

[27] Day N, McKeown N, Wong M, Welch A, Bingham S. Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. Int J Epidemiol 2001; 30: 309-317.

[28] Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health. 2014 Jul 22;36:e2014009.

[29] Timon CM, van den Barg R, Blain RJ, Kehoe L, Evans K, Walton J, Flynn A, Gibney ER. A review of the design and validation of web- and computer-based 24-h dietary recall tools. Nutr Res Rev. 2016 Dec;29(2):268-280.

[30] U.S. DEPARTMENT OF AGRICULTURE. Food and Nutrient Database for Dietary Studies. https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds/ (Last Modified: 1/7/2021)

[31] Rosinger A, Herrick K. Daily Water Intake Among U.S. Men and Women, 2009-2012. NCHS Data Brief. 2016 Apr;(242):1-8.

[32] EFSA Panel on Dietetic Products, Nutrition, and Allergies (NDA); Scientific Opinion on Dietary reference values for water. EFSA Journal 2010; 8(3):1459.

[33] Tani Y, Asakura K, Sasaki S, Hirota N, Notsu A, Todoriki H, Miura A, Fukui M, Date C. The influence of season and air temperature on water intake by food groups in a sample of free-living Japanese adults. Eur J Clin Nutr 2015; 69: 907-913.

[34] Murakami M, Mizuma K, Nakamura Y, Watanabe R. Estimation of water intake from food moisture in the Japanese diet using a cooking-based conversion factor for water content. J Food Sci 2021; 86: 266-275.

[35] Arens-Volland AG, Spassova L, Bohn T. Promising approaches of computer-supported dietary assessment and management-Current research status and available applications. Int J Med Inform. 2015 Dec;84(12):997-1008.

[36] Braekman E, Berete F, Charafeddine R, Demarest S, Drieskens S, Gisle L, Molenberghs G, Tafforeau J, Van der Heyden J, Van Hal G. Measurement agreement of the self-administered questionnaire of the Belgian Health Interview Survey: Paper-and-pencil versus web-based mode. PLoS One. 2018 May 21;13(5):e0197434.

[37] Hotz C, Abdelrahman L. Simple methods to obtain food listing and portion size distribution estimates for use in semi-quantitative dietary assessment methods. PLoS One. 2019 Oct 18;14(10):e0217379.

[38] Yamaoka K, Tango T, Watanabe M, Yokotsuka M. Validity and reproducibility of a semi-quantitative food frequency questionnaire for nutritional education of patients of diabetes mellitus (FFQW65). Nihon Koshu Eisei Zasshi. 2000 Mar;47(3):230-44.

[39] Taru C, Tsutou A, Miyawaki I. A modified simple questionnaire to estimate dietary energy intake for the Japanese. Kobe J Med Sci. 2012 Jan 16;57(3):E106-15.

[40] Sasaki S, Yanagibori R, Amano K. Self-administered diet history questionnaire developed for health education: a relative validation of the test-version by comparison with 3-day diet record in women. J Epidemiol. 1998 Oct;8(4):203-15.

[41] Kobayashi S, Honda S, Murakami K, Sasaki S, Okubo H, Hirota N, Notsu A, Fukui M, Date C. Both comprehensive and brief self-administered diet history questionnaires satisfactorily rank nutrient intakes in Japanese adults. J Epidemiol. 2012;22(2):151-9.

[42] De Silva K, Jönsson D, Demmer RT. A combined strategy of feature selection and machine learning to identify predictors of prediabetes. J Am Med Inform Assoc. 2020;27(3):396-406.

[43] Birk N, Matsuzaki M, Fung TT, Li Y, Batis C, Stampfer MJ, Deitchler M, Willett WC, Fawzi WW, Bromage S, Kinra S, Bhupathiraju SN, Lake E. Exploration of Machine Learning and Statistical Techniques in Development of a Low-Cost Screening Method Featuring the Global Diet Quality Score for Detecting Prediabetes in Rural India. J Nutr. 2021 Oct 23;151(12 Suppl 2):110S-118S.

[44] Abbas M, Mall R, Errafii K, Lattab A, Ullah E, Bensmail H, Arredouani A. Simple risk score to screen for prediabetes: A cross-sectional study from the Qatar Biobank cohort. J Diabetes Investig. 2021 Jun;12(6):988-997.

[45] Dong W, Tse TYE, Mak LI, Wong CKH, Wan YFE, Tang HME, Chin WY, Bedford LE, Yu YTE, Ko WKW, Chao VKD, Tan CBK, Lam LKC. Non-laboratory-based risk assessment model for case detection of diabetes mellitus and pre-diabetes in primary care. J Diabetes Investig. 2022 Aug;13(8):1374-1386.

[46] Tian X, Liu Y, Han Y, Shi J, Zhu T. Risk Score for Detecting Dysglycemia: A Cross-Sectional Study of a Working-Age Population in an Oil Field in China. Med Sci Monit. 2017 Jun 11;23:2833-2841.

[47] Shen XM, Huang YQ, Zhang XY, Tong XQ, Zheng PF, Shu L. Association between dietary patterns and prediabetes risk in a middle-aged Chinese population. Nutr J. 2020 Jul 30;19(1):77.

[48] DECODE study group. Will new diagnostic criteria for diabetes change phenotype of patients with diabetes? Reanalysis of European epidemiological data. BMJ 317:371–

375, 1998.

[49] Shaw JE, Zimmet PZ, de Courten M, Dowse GK, Chitson P, Gareeboo H, Hemraj F, Fareed D, Tuomilehto J, Alberti KGMM: Impaired fasting glucose or impaired glucose tolerance: what best predicts future diabetes in Mauritius? Diabetes Care 22:399-402, 1999.

[50] Uchida T, Teramoto T, Fukizawa S, Kato H, Nonaka Y, Suematsu M, Murayama N. Characteristics of the glycometabolic categories based on the oral glucose tolerance test results in Japanese adults without diabetes. Eur Rev Med Pharmacol Sci. 2022 Apr;26(8):2765-2774.

[51] Stephen BH, Steven RC, Warren SB, Deborah GG, Thomas BN, Kihara M, Kihara M. Designing Clinical Research 4th edition. Medical science international, 2014 (in Japanese).

[52] Date C, Fukui M, Yokoyama T, Yoshiike N, Matsumura Y, Tanaka H. Development of food frequency questionnaire. The Jpn J Nutr Diet 1998; 56: 313–325 (in Japanese).

[53] Bender DA. A Dictionary of Food and Nutrition 3rd ed. Oxford University Press, 2009.

[54] Tanaka T, Okamura T, Miura K, Kadowaki T, Ueshima H, Nakagawa H, Hashimoto T. A simple method to estimate populational 24-h urinary sodium and potassium excretion using a casual urine specimen. J Hum Hypertens 2002; 16: 97-103.

[55] Steiger, J. H. Tests for comparing elements of a correlation matrix. Psychological Bulletin 1980; 87(2), 245-251.

[56] Perrier E, Vergne S, Klein A, Poupin M, Rondeau P, Le Bellego L, Armstrong LE, Lang F, Stookey J, Tack I. Hydration biomarkers in free-living adults with different levels of habitual fluid consumption. Br J Nutr 2013; 109: 1678-1687.

[57] Armstrong LE, Kavouras SA, Walsh NP, Roberts WO. Diagnosing dehydration? Blend evidence with clinical observations. Curr Opin Clin Nutr Metab Care 2016; 19: 434-438.

[58] Hooper L, Bunn D, Jimoh FO, Fairweather-Tait SJ. Water-loss dehydration and aging. Mech Ageing Dev 2014; 136-137: 50-58.

[59] Lavizzo-Mourey RJ. Dehydration in the elderly: a short review. J Natl Med Assoc 1987; 79: 1033-1038.

[60] Nielsen S, Frøkiaer J, Marples D, Kwon TH, Agre P, Knepper MA. Aquaporins in the kidney: from molecules to medicine. Physiol Rev 2002; 82: 205-244.

[61] Ranieri M, Di Mise A, Tamma G, Valenti G. Vasopressin-aquaporin-2 pathway: recent advances in understanding water balance disorders. F1000Res 2019; 8: F1000 Faculty Rev-149.

[62] Pedersen RS, Bentzen H, Bech JN, Pedersen EB. Effect of water deprivation and hypertonic saline infusion on urinary AQP2 excretion in healthy humans. Am J Physiol Ren Physiol 2001; 280: F860-F867.

[63] Makabe S, Manabe S, Kataoka H, Akihisa T, Yoshida R, Ushio Y, Sato M, Tsuchiya K, Mochizuki T, Nitta K. Urinary Aquaporin 2 as a Potential Indicator Predicting Tolvaptan Response in Patients With ADPKD. Kidney Int Rep. 2021 Jul 14;6(9):2436-2444.

[64] Hirakawa Y, Hata J, Yoshinari M, Higashioka M, Yoshida D, Shibata M, Honda T, Sakata S, Kato H, Teramoto T, Maki H, Nishimoto S, Kitazono T, Ninomiya T. 30-minute postload plasma glucose levels during an oral glucose tolerance test predict the risk of future type 2 diabetes: the Hisayama Study. BMJ Open Diabetes Res Care 202; 8: e001156.

[65] Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22:1462–1470.

[66] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011 Mar 17;12:77.

[67] Frank Bretz, Torsten Hothorn, and Peter Westfall. On multiple comparisons in R. R News, 2(3):14--17, December 2002.

[68] Kodama K, Tojjar D, Yamada S, Toda K, Patel CJ, Butte AJ. Ethnic differences in the relationship between insulin sensitivity and insulin response: a systematic review and

meta-analysis. Diabetes Care. 2013 Jun;36(6):1789-96.

[69] Magkos F, Lee MH, Lim M, Cook AR, Chhay V, Loh TP, Chia KS, Baig S, Ang IYH, Tay JYY, Khoo CM, Halter JB, Toh SA. Dynamic assessment of insulin secretion and insulin resistance in Asians with prediabetes. Metabolism. 2022 Mar;128:154957.

[70] Tseng E, Greer RC, O'Rourke P, Yeh HC, McGuire MM, Clark JM, Maruthur NM. Survey of primary care providers' knowledge of screening for, diagnosing and managing prediabetes. J Gen Intern Med. 2017 Nov;32(11):1172-1178.

[71] Barber SR, Davies MJ, Khunti K, Gray LJ. Risk assessment tools for detecting those with pre-diabetes: a systematic review. Diabetes Res Clin Pract. 2014 Jul;105(1):1-13.

[72] Gray LJ, Taub NA, Khunti K, Gardiner E, Hiles S, Webb DR, Srinivasan BT, Davies MJ. The Leicester Risk Assessment score for detecting undiagnosed Type 2 diabetes and impaired glucose regulation for use in a multiethnic UK setting. Diabet Med. 2010 Aug;27(8):887-95.

[73] Gray LJ, Davies MJ, Hiles S, Taub NA, Webb DR, Srinivasan BT, Khunti K. Detection of impaired glucose regulation and/or type 2 diabetes mellitus, using primary care electronic data, in a multiethnic UK community setting. Diabetologia. 2012 Apr;55(4):959-66.

[74] Robinson CA, Agarwal G, Nerenberg K. Validating the CANRISK prognostic model for assessing diabetes risk in Canada's multi-ethnic population. Chronic Dis Inj Can. 2011 Dec;32(1):19–31.

[75] Tan C, Sasagawa Y, Kamo KI, Kukitsu T, Noda S, Ishikawa K, Yamauchi N, Saikawa T, Noro T, Nakamura H, Takahashi F, Sata F, Tada M, Kokai Y. Evaluation of the Japanese Metabolic Syndrome Risk Score (JAMRISC): a newly developed questionnaire used as a screening tool for diagnosing metabolic syndrome and insulin resistance in Japan. Environ Health Prev Med. 2016 Nov;21(6):470-479.

[76] Basu R, Breda E, Oberg AL, Powell CC, Dalla Man C, Basu A, Vittone JL, Klee GG, Arora P, Jensen MD, Toffolo G, Cobelli C, Rizza RA. Mechanisms of the age-associated deterioration in glucose tolerance: contribution of alterations in insulin secretion, action, and clearance. Diabetes. 2003 Jul;52(7):1738-48.

[77] Barazzoni R, Gortan Cappellari G, Ragni M, Nisoli E. Insulin resistance in obesity: an overview of fundamental alterations. Eat Weight Disord. 2018 Apr;23(2):149-157.

[78] Abdullah A, Peeters A, de Courten M, Stoelwinder J. The magnitude of association between overweight and obesity and the risk of diabetes: a meta-analysis of prospective cohort studies. Diabetes Res Clin Pract. 2010 Sep;89(3):309-19.

[79] Duval S, Vazquez G, Baker WL, Jacobs DR Jr; CODA study group. The Collaborative Study of Obesity and Diabetes in Adults (CODA) project: meta-analysis design and description of participating studies. Obes Rev. 2007 May;8(3):263-76.

[80] Wittenbecher C, Kuxhaus O, Boeing H, Stefan N, Schulze MB. Associations of short stature and components of height with incidence of type 2 diabetes: mediating effects of cardiometabolic risk factors. Diabetologia. 2019 Dec;62(12):2211-2221.

[81] Furer A, Afek A, Beer Z, Derazne E, Tzur D, Pinhas-Hamiel O, Reichman B, Twig G. Height at Late Adolescence and Incident Diabetes among Young Men. PLoS One. 2015 Aug 25;10(8):e0136464.

[82] Buxton OM, Cain SW, O'Connor SP, Porter JH, Duffy JF, Wang W, Czeisler CA, Shea SA. Adverse metabolic consequences in humans of prolonged sleep restriction combined with circadian disruption. Sci Transl Med. 2012 Apr 11;4(129):129ra43.

[83] McHill AW, Melanson EL, Higgins J, Connick E, Moehlman TM, Stothard ER, Wright KP Jr. Impact of circadian misalignment on energy metabolism during simulated nightshift work. Proc Natl Acad Sci U S A. 2014 Dec 2;111(48):17302-7.

[84] Bescos R, Boden MJ, Jackson ML, Trewin AJ, Marin EC, Levinger I, Garnham A, Hiam DS, Falcao-Tebas F, Conte F, Owens JA, Kennaway DJ, McConell GK. Four days of simulated shift work reduces insulin sensitivity in humans. Acta Physiol (Oxf). 2018 Jun;223(2):e13039.

[85] Shan Z, Ma H, Xie M, Yan P, Guo Y, Bao W, Rong Y, Jackson CL, Hu FB, Liu L. Sleep duration and risk of type 2 diabetes: a meta-analysis of prospective studies. Diabetes Care. 2015 Mar;38(3):529-37.

[86] Johnson KA, Gordon CJ, Chapman JL, Hoyos CM, Marshall NS, Miller CB, Grunstein RR. The association of insomnia disorder characterised by objective short sleep

duration with hypertension, diabetes and body mass index: A systematic review and meta-analysis. Sleep Med Rev. 2021 Oct;59:101456.

[87] Exelmans L, Van den Bulck J. Bedtime mobile phone use and sleep in adults. Soc Sci Med. 2016 Jan;148:93–101. doi: 10.1016/j.socscimed.2015.11.037. Epub 2015 Dec 2.

[88] He JW, Tu ZH, Xiao L, Su T, Tang YX. Effect of restricting bedtime mobile phone use on sleep, arousal, mood, and working memory: A randomized pilot trial. PLoS One. 2020 Feb 10;15(2):e0228756.

[89] Ren G, Qi J, Zou Y. Association between intake of white rice and incident type 2 diabetes - An updated meta-analysis. Diabetes Res Clin Pract. 2021 Feb;172:108651.

[90] Aune D, Norat T, Romundstad P, Vatten LJ. Whole grain and refined grain consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. Eur J Epidemiol. 2013 Nov;28(11):845-58.

[91] Olfert MD, Wattick RA. Vegetarian Diets and the Risk of Diabetes. Curr Diab Rep. 2018 Sep 18;18(11):101.

[92] Salas-Salvadó J, Bulló M, Babio N, Martínez-González MÁ, Ibarrola-Jurado N, Basora J, Estruch R, Covas MI, Corella D, Arós F, Ruiz-Gutiérrez V, Ros E; PREDIMED Study Investigators. Reduction in the incidence of type 2 diabetes with the Mediterranean diet: results of the PREDIMED-Reus nutrition intervention randomized trial. Diabetes Care. 2011 Jan;34(1):14-9.

[93] Sandouk Z, Lansang MC. Diabetes with obesity--Is there an ideal diet? Cleve Clin J Med. 2017 Jul;84(7 Suppl 1):S4–S14.

[94] Kolb H, Martin S. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. BMC Med. 2017 Jul 19;15(1):131.

[95] Wu LY, Juan CC, Hwang LS, Hsu YP, Ho PH, Ho LT. Green tea supplementation ameliorates insulin resistance and increases glucose transporter IV content in a fructose-fed rat model. Eur J Nutr. 2004 Apr;43(2):116-24.

[96] Yan J, Zhao Y, Suo S, Liu Y, Zhao B. Green tea catechins ameliorate adipose insulin resistance by improving oxidative stress. Free Radic Biol Med. 2012 May

1;52(9):1648-57.

[97] Wolfram S, Raederstorff D, Preller M, Wang Y, Teixeira SR, Riegger C, Weber P. Epigallocatechin gallate supplementation alleviates diabetes in rodents. J Nutr. 2006 Oct;136(10):2512-8.

[98] Sacramento JF, Ribeiro MJ, Yubero S, Melo BF, Obeso A, Guarino MP, Gonzalez C, Conde SV. Disclosing caffeine action on insulin sensitivity: effects on rat skeletal muscle. Eur J Pharm Sci. 2015 Apr 5;70:107-16.

[99] Jing Y, Han G, Hu Y, Bi Y, Li L, Zhu D. Tea consumption and risk of type 2 diabetes: a meta-analysis of cohort studies. J Gen Intern Med. 2009 May;24(5):557-62.

[100] Iso H, Date C, Wakai K, Fukui M, Tamakoshi A; JACC Study Group. The relationship between green tea and total caffeine intake and risk for self-reported type 2 diabetes among Japanese adults. Ann Intern Med. 2006 Apr 18;144(8):554-62.

[101] van Dieren S, Uiterwaal CS, van der Schouw YT, van der A DL, Boer JM, Spijkerman A, Grobbee DE, Beulens JW. Coffee and tea consumption and risk of type 2 diabetes. Diabetologia. 2009 Dec;52(12):2561-9.

[102] InterAct Consortium, van Woudenbergh GJ, Kuijsten A, Drogan D, van der A DL, Romaguera D, Ardanaz E, Amiano P, Barricarte A, Beulens JW, Boeing H, Bueno-de-Mesquita HB, Dahm CC, Chirlaque MD, Clavel F, Crowe FL, Eomois PP, Fagherazzi G, Franks PW, Halkjaer J, Khaw KT, Masala G, Mattiello A, Nilsson P, Overvad K, Ramón Quirós J, Rolandsson O, Romieu I, Sacerdote C, Sánchez MJ, Schulze MB, Slimani N, Sluijs I, Spijkerman AM, Tagliabue G, Tjønneland A, Tumino R, Forouhi NG, Sharp S, Langenberg C, Feskens EJ, Riboli E, Wareham NJ. Tea consumption and incidence of type 2 diabetes in Europe: the EPIC-InterAct case-cohort study. PLoS One. 2012;7(5):e36910.

[103] Yu J, Song P, Perry R, Penfold C, Cooper AR. The Effectiveness of Green Tea or Green Tea Extract on Insulin Resistance and Glycemic Control in Type 2 Diabetes Mellitus: A Meta-Analysis. Diabetes Metab J. 2017 Aug;41(4):251-262.

[104] Mackenzie T, Leary L, Brooks WB. The effect of an extract of green and black tea on glucose control in adults with type 2 diabetes mellitus: double-blind randomized

study. Metabolism 2007 Oct;56(10):1340–4.

[105] Hale PJ, Horrocks PM, Wright AD, Fitzgerald MG, Nattrass M, Bailey CJ. Xiaoke tea, a Chinese herbal treatment for diabetes mellitus. Diabet Med. 1989 Nov;6(8):675-6.

[106] Fuchs D, Nyakayiru J, Draijer R, Mulder TP, Hopman MT, Eijsvogels TM, Thijssen DH. Impact of flavonoid-rich black tea and beetroot juice on postprandial peripheral vascular resistance and glucose homeostasis in obese, insulin-resistant men: a randomized controlled trial. Nutr Metab (Lond). 2016 May 13;13:34.

[107] Baeza M, Morales A, Cisterna C, Cavalla F, Jara G, Isamitt Y, Pino P, Gamonal J. Effect of periodontal treatment in patients with periodontitis and diabetes: systematic review and meta-analysis. J Appl Oral Sci. 2020 Jan 10;28:e20190248.

[108] Polak D, Shapira L. An update on the evidence for pathogenic mechanisms that may link periodontitis and diabetes. J Clin Periodontol. 2018 Feb;45(2):150–166.

[109] Tagelsir A, Cauwels R, van Aken S, Vanobbergen J, Martens LC. Dental caries and dental care level (restorative index) in children with diabetes mellitus type 1. Int J Paediatr Dent. 2011 Jan;21(1):13-22.

[110] Yaribeygi H, Farrokhi FR, Butler AE, Sahebkar A. Insulin resistance: Review of the underlying molecular mechanisms. J Cell Physiol 2019; 234: 8152-8161.

[111] Yazıcı D, Sezer H. Insulin Resistance, Obesity and Lipotoxicity. Adv Exp Med Biol 2017; 960: 277-304.

[112] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research. 2002;16:321–57.

[113] Hooper L, Bunn D, Jimoh FO, Fairweather-Tait SJ. Water-loss dehydration and aging. Mech Ageing Dev. 2014 Mar-Apr;136-137:50-8.

[114] Ministry of Health, Labour and Welfare, Japan. Overview of patient survey in 2017. 2019 Mar 1.

Publication lists

Journal papers

1. Uchida T, Nakamura Y, Tanaka H, Nakamura S, Okamura T, Watanabe H, Murayama N. Validity of a selective recall method for assessing water intake and its relationship with hydration status. Eur Rev Med Pharmacol Sci. 2021 Nov;25(21):6623-6632.

2. Uchida T, Kanamori T, Teramoto T, Nonaka Y, Tanaka H, Nakamura S, Murayama N. Identifying Glucose Metabolism Status in Nondiabetic Japanese Adults Using Machine Learning Model with Simple Questionnaire. Comput Math Methods Med. 2022 Sep 9;2022:1026121.

International conference

1. Uchida T, Kanamori T, Teramoto T, Kato T, Nonaka Y, Suematsu M, Tanaka H, Nakamura S, Murayama N. Characteristics of glycometabolism in individuals without diabetes and a model to assess their glucometabolic category. 15th International Conference on Advanced Technologies & Treatments for Diabetes. 27-30 April 2022.