IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

IEEE Engineering in Medicine and Biology Society Section

Received 31 May 2024, accepted 24 June 2024, date of publication 17 July 2024, date of current version 26 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3429521

RESEARCH ARTICLE

A Study of the Relationship Between Driving and Health Based on Large-Scale Data Analysis Using PLSA and t-SNE

MITSUGU MERA[1], NANAE MICHIDA[1], MASANORI HONDA[1], KAZUO SAKAMOTO[1], YOSHINORI TAMADA[2], TATSUYA MIKAMI[3], AND SHIGEYUKI NAKAJI[4]

[1]Technical Research Center, Mazda Motor Corporation, Hiroshima 730-8670, Japan
[2]Research Center for Health-Medical Data Science, Hirosaki University Graduate School of Medicine, Aomori 036-8562, Japan
[3]Department of Preemptive Medicine, Innovation Center for Health Promotion, Hirosaki University Graduate School of Medicine, Aomori 036-8562, Japan
[4]Department of Social Medicine, Hirosaki University Graduate School of Medicine, Aomori 036-8562, Japan

Corresponding authors: Mitsugu Mera (mera.mi@mazda.co.jp), Nanae Michida (michida.n@mazda.co.jp), Masanori Honda (honda.masa@mazda.co.jp), Kazuo Sakamoto (sakamoto.kaz@mazda.co.jp), Yoshinori Tamada (y.tamada@hirosaki-u.ac.jp), Tatsuya Mikami (tmika@hirosaki-u.ac.jp), and Shigeyuki Nakaji (nakaji@hirosaki-u.ac.jp)

This work was supported in part by Iwaki Health Promotion Project; and in part by the Center of Innovation (COI) Program of Japan Science and Technology Agency (JST) under Grant JPMJCE1311, Grant JPMJCE1302, and Grant JPMJCA2201.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Hirosaki University School of Medicine.

ABSTRACT The purpose of this study is to facilitate knowledge discovery about the relationship between driving and health among the elderly. In the Iwaki Health Promotion Project that is an annual project conducted by Hirosaki University, we have included a survey on driving for the first time in 2019. After linking the data obtained from the survey with four years of health data for 2016–2019, we have utilized PLSA as a machine learning method to cluster those data in an integrated manner. As a result, we have found latent classes broadly classified according to whether the health level has been generally high or low. Also, when we have focused on a specific health item, for example, cognitive function, we have found some people with higher and lower maintenance of cognitive function over four years, even if they have belonged to a same latent class. To characterize these differences in detail, we have utilized t-SNE as a machine learning method. As a result, we have found that ''I like driving'' as a factor related to the Kansei (sensitivity) may characterize the high maintenance of cognitive function. For those who like driving, it is considered that the high maintenance of cognitive function may be occurred because they enjoy driving, have a wider range of activities, and increase the possibility of multitasking.

INDEX TERMS Cognitive function, driving, health, machine learning, PLSA, t-SNE.

I. INTRODUCTION

Japan is now facing a super-aged society. According to the World Health Organization (WHO), Japan has the highest life expectancy (LE) at birth in 2019 with 84.26 years [1]. Also, according to the Abridged Life Tables For Japan 2022, LE at birth in Japan is 81.05 years for males and 87.09 years for females [2]. Therefore, it is an important social issue to realize the extension of healthy life expectancy by ensuring transportation that allows people to go where they want to go, so that they can live a healthy and rich life for a long time, even if they grow older. In order to support the realization of the above, on the premise of providing safe and reliable cars from the standpoint of automotive industries, we believe that the ability to drive a car with one's own hands and feet may be a positive factor in promoting physical and mental vitality and maintaining health.

In fact, previous studies have reported that the cessation of driving among the elderly is associated with various health problems, including depression [3], and an increased risk of

The associate editor coordinating the review of this manuscript and approving it for publication was Kin Fong Lei.

functional decline with psychological frailty [4]. In addition, a 6-week cognitive training program that simulates driving a car using a video game among the elderly has also been reported to be effective in maintaining cognitive function and improving mood [5].

Conversely, a comprehensive review of the relationship between transportation and health [6] has been reported from the standpoint that a lifestyle with low physical activity may have adverse effects on health and public health due to inactivity and reduced walking ability. In addition, it has also been reported that the risk of functional limitations may be alleviated among the elderly if they use public transportation or bicycles as an alternative means of transportation after driving cessation [7]. Therefore, it is difficult to simply link driving to health. However, we believe that a more detailed survey of daily driving, combined with a comprehensive health survey, may reveal factors related to driving that may trigger good health.

On the other hand, although a study on driving is not included, there are reports that, for example, having hobbies and purpose in life (PIL) are factors associated with maintaining survival among the community-dwelling elderly [8]. In other words, the above suggests that factors that stimulate the Kansei (sensitivity) associated with emotional needs may have a positive effect on health. The Kansei (sensitivity) is a Japanese word that does not have a direct translation in English, however, every translation (such as sensibility, sentiment, emotion, and feeling) captures just some of the aspects of Kansei [9]. Therefore, it may be effective to conduct a survey on driving while actively taking into account factors related to the Kansei (sensitivity) in order to understand the relationship between driving and health in detail. However, as the number of survey items increases, classical statistical analysis has limitations in taking into account all items in an integrated manner and overlooking overall trends while dealing with nonlinearities and interdependencies in the data. In order to break through these limitations, in recent years, there have been an increasing number of reports in healthcare and other fields that have actively utilized machine learning based on ensemble methods, Bayesian statistics, and other approaches for a variety of data, including not only images and text, but also questionnaires and other surveys [10], [11], [12], [13], [14], [15], [16]. More specifically, the reports in [10], [11], and [12] are all examples of efforts to improve the accuracy in classification and prediction by using machine learning. In contrast, the reports in [13], [14], [15], and [16] correspond to efforts to ensure possibilities of knowledge discovery. In detail, [13] reports an example of trying to find a source of information for classifying medical images from non-image knowledge (bag of words in the explanatory documents accompanying the images). In [14], an example of approaching knowledge discovery by ensuring explainability equivalent to the basis of classification results for medical images is reported. In [15], an example of approaching knowledge discovery by structuring the interdependent relationships among factors related to disease risk is reported. In [16], an example of approaching knowledge discovery by embedding and visualizing factors related to psychological motivation for sports participation on a 2-dimensional plane while dealing with nonlinearities is reported. The methods used in these examples need to be differentiated according to the type and amount of data, and the purpose of analysis, but all are effective options for knowledge discovery. Nevertheless, we have seen no previous examples of studies using machine learning as in the above approaches to handle large-scale data obtained from a set of health and driving surveys while taking the Kansei (sensitivity) into account.

In this study, we have conducted a field survey with a wide range of health survey items as well as detailed driving questionnaire survey items related to the Kansei (sensitivity), such as driving preferences or usages. Our goal is to realize a car lifestyle that can assist healthy longevity according to each individual's attributes. In order to achieve this goal, the objective is to construct an approach that can not only understand the overall features of the elderly, but also focus on their specific characteristics as individual differences and scoop them up as findings, rather than simply ignoring them. Toward this, the purpose of this study is to facilitate knowledge discovery about the relationship between driving and health among the elderly by using machine learning to analyze the potential relationships among variables in the large-scale data obtained from the above. The driving survey itself, including items related to the Kansei (sensitivity), is unparalleled in prior research, and to our knowledge, this is the first study to acquire and analyze such items in combination with a large-scale health survey. As an approach to analyzing the data, although [16] is the most helpful among the previous reports, since the number of items we handle is much larger, we have adopted an approach that analyzes overall features and specific characteristics by combining two different methods, which are PLSA [17] and t-SNE [18]. Compared to other machine learning methods, PLSA can find latent classes by clustering samples and factors simultaneously and can easily obtain the factor contributions incidentally. Besides, since it is not bound to a prior distribution, latent classes are less abstract and can maximally incorporate the influence of the data. Therefore, we have selected PLSA because of its relative advantage in deepening our understanding of the current state of the data. In addition, since t-SNE can preserve the local rather than the global structure of the data, it has a high potential to be able to focus on and characterize some small number of subjects from the clustering results. Therefore, we have selected t-SNE because of its relative advantage in obtaining in-depth perspectives that are difficult to cover with other machine learning methods.

The remainder of this paper is organized as follows. Section II describes the data and analysis methods we use in this study. Section III describes the results obtained through

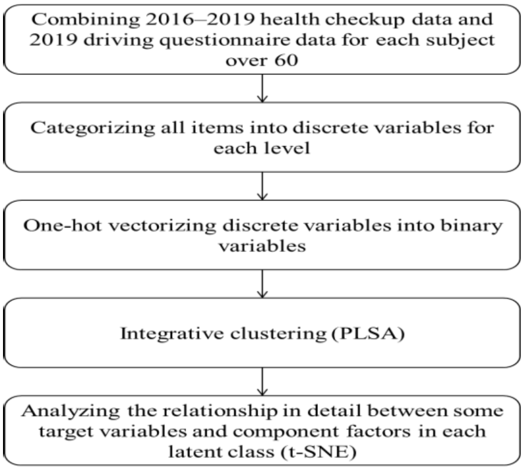

FIGURE 1. Overview of our analysis flow.

TABLE 1. Examples of basic statistics (maximum, minimum, mean and standard deviation) of the category "Personal record form" for subjects (1044 persons) in 2019.

Items	Max	Min	Mean	SD
Age [years old]	89	20	52.7	15.3
Body height [cm]	185.2	136.2	161.7	8.9
Body weight [kg]	133.7	31.9	60.6	12.5
Grip strength (Right hand) [kg]	68	6	31.2	9.6
Grip strength (Left hand) [kg]	60	2	30.5	9.6
Sitting trunk flexion [cm]	69	9	39.1	10.4
Whole body reaction time [msec]	1142	0	424.4	89.4
Sit-to-stand test [cm]	40	10	26.2	13
Timed up and go (TUG) test [sec]	13.06	2.92	4.8	1
Two-step test [cm]	350	132	249.2	34
Maximal walking speed of 10 meters [sec]	10.6	2.3	4.5	0.94

our data analysis. Section IV presents our discussion of the obtained results. Finally, Section V presents our conclusions of this study and our future works.

II. RESEARCH METHODS

The general flow of our large-scale data analysis in this study is shown in Fig. 1.

A. DATA FOR ANALYSIS

In this study, we utilize data accumulated in the Iwaki Health Promotion Project [19], a large-scale health survey conducted by Hirosaki University in Aomori Prefecture (Iwaki District, Hirosaki City) for approximately 1000 residents every year since 2005. Specifically, we link a single year's worth of data consisting of the driving questionnaire survey items obtained for the first time in 2019 to four years' worth of data consisting of the health survey items from the above project spanning 2016–2019. We then use our data restricted to the elderly (160 males and 252 females over 60 in 2019). As an overview of the data for all subjects, we take the subjects (1044 persons) in 2019 and show each aggregate result for "Personal record form", a kind of health checkup categories,

TABLE 2. Examples of aggregated questionnaire responses regarding driving frequency, driving time, transmission (AT/MT) and likes/dislikes about driving for subjects (1044 persons) in 2019.

Items of driving questionnaire	Response options	Percentage [%]
Driving frequency	Every day	77
	5–6 days a week	9.3
	3–4 days a week	5.7
	Less than 2 days a week	2.7
	None	5.3
Driving time	Less than 30 minutes	30.1
	30 minutes to 1 hour	42
	1–2 hours	14.5
	More than 2 hours	8.3
Transmission: Automatic (AT) or Manual (MT)	Only AT	67.3
	Only MT	16.5
	Both AT and MT	10.9
Likes or dislikes about driving	Like	36.6
	Rather like	40.8
	Rather dislike	16.1
	Dislike	1.1

and some items of the driving questionnaire in Tables 1 and 2, respectively.

Here we describe the details of the items handled in our analysis. First, the health survey items are shown in Table 3. We integrally handle items corresponding to the various health checkup categories of "Personal record form", "Cognitive function", "Health questionnaire", "Vision test", "Blood examination", "Blood pressure measurement of the extremities", "Pulmonary function testing (Spirometry)", "Heel bone density test", "Body composition measurement", and "Gravic body sway test (Otorhinolaryngology)". Although there are multiple items of cognitive function in the actual health surveys (see [19]), in this study, we only handle the MMSE for our analysis. In the following description, RHH and RLH are described as relatively high and low levels of health, respectively.

1) PERSONAL RECORD FORM

For the following 3 items, we do not define any association with RHH or RLH. We define the level divisions for those data as follows.

- "Age [years old]": We define 7 divisions into 10-year increments from 20s to 80s.
- "Body height [cm]": We define 5 equal divisions between the maximum and minimum.
- "Body weight [kg]": We define 5 equal divisions between the maximum and minimum.

For the following 8 items, there are no explicit criteria for any of RHH and RLH. Therefore, as those analysis criteria, we define 2 sets of thresholds excluding the neutral center when divided into 5 equal divisions between the maximum and minimum. The specific threshold values corresponding to RHH and RLH are defined as follows.

- "Grip strength (Right hand) [kg]": We define $>=$ 43.2 and $<$ 30.8 for RHH and RLH criteria, respectively.
- "Grip strength (Left hand) [kg]": We define $>=$ 36.8 and $<$ 25.2 for RHH and RLH criteria, respectively.

- "Sitting trunk flexion [cm]": We define >= 45 and < 33 for RHH and RLH criteria, respectively. Incidentally, see [20] for a detailed methodology on this item.
- "Whole body reaction time [msec]": We define < 456.8 and >= 685.2 for RHH and RLH criteria, respectively. Incidentally, see [21] for a detailed methodology on this item.
- "Sit-to-stand test [cm]": We define >= 28 and < 22 for RHH and RLH criteria, respectively. Incidentally, see [22] for a detailed methodology on this item.
- "Timed up and go (TUG) test [sec]": We define < 6.98 and >= 9 for RHH and RLH criteria, respectively. Incidentally, see [23] for a detailed methodology on this item.
- "Two-step test [cm]": We define >= 262.8 and < 219.2 for RHH and RLH criteria, respectively.
- "Maximal walking speed of 10 meters [sec]": We define < 5.62 and >= 7.28 for RHH and RLH criteria, respectively.

2) COGNITIVE FUNCTION

For the following 1 item, we divide each value into 3 divisions based on Level 3 (27–30 points), Level 2 (22–26 points), which is neutral, and Level 1 (0–21 points). Then, referring to [24], the specific threshold values corresponding to RHH and RLH are defined as follows.

- "Mini mental state examination (MMSE) score (0–30 points) [-]": We define Level 3 and Level 1 for RHH and RLH criteria, respectively.

3) HEALTH QUESTIONNAIRE

For the following 2 items, we divide each value into 4 divisions based on 0, 1–5, 6–10, and 11 or more people, while including dummies. Then, referring to [25], the specific threshold values corresponding to RHH and RLH are defined as follows.

- "Social connections (Number of friends who feel comfortable) [-]": We define non-zero and zero for RHH and RLH criteria, respectively.
- "Social connections (Number of family members and relatives who feel comfortable) [-]": We define non-zero and zero for RHH and RLH criteria, respectively.

For the following 1 item, we divide each value into 2 divisions based on 0–5 and 6–21 points. Then, referring to [26], the specific threshold values corresponding to RHH and RLH are defined as follows.

- "Pittsburgh sleep quality index (PSQI) total score (0–21 points) [-]": We define 0–5 and 6–21 points for RHH and RLH criteria, respectively. Incidentally, see [27] for a detailed methodology on this item.

For the following 1 item, we divide each value into 6 divisions based on 0–9, 10–19, 20–29, 30–39, 40–49, and 50–60 points, while including dummies. Then, referring to [28] (or [29]), the specific threshold values corresponding to RHH and RLH are defined as follows. Here a division which intersects both RHH and RLH is considered neutral.

- "Depressive state: 20-item center for epidemiologic studies depression scale (CESD-20) score (0–60 points) [-]": We define 0–15 and 16–60 points for RHH and RLH criteria, respectively. Incidentally, see [30] for a detailed methodology on this item.

For the following 1 item, we divide each value into 10 divisions based on 0–9, 10–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80–89, and 90–100 points, while including dummies. Then, referring to [31], the specific threshold values corresponding to RHH and RLH are defined as follows. Here a division which intersects both RHH and RLH is considered neutral.

- "Locomotive syndrome: 25-question geriatric locomotive function scale (GLFS-25) score (0–100 points) [-]": We define 0–15 and 16–60 points for RHH and RLH criteria, respectively. Incidentally, see [32] for a detailed methodology on this item.

For the following 8 items, we divide each value into 10 divisions based on 0–9, 10–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80–89, and 90–100 points, while including dummies. Then, referring to [33], the specific threshold values corresponding to RHH and RLH are defined as follows. Here a division which intersects both RHH and RLH is considered neutral.

- "Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Physical functioning (PF)": We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.
- "Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Role physical (RP)": We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.
- "Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Bodily pain (BP)": We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.
- "Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: General health (GH)": We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.
- "Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Vitality (VT)": We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.
- "Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Social functioning (SF)": We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.
- "Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points)

[-] - Subscale: Role emotional (RE)'': We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.
- ''Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Mental health (MH)'': We define 51–100 and 0–49 points for RHH and RLH criteria, respectively.

4) VISION TEST

The compact vision meter ''CA-1000'' (TOMEY Corporation, Nagoya City, Japan) [34] is used for this test. For the following 6 items, we divide each value into 2 divisions based on more than or less than 1.0. Then, referring to [35], the specific threshold values corresponding to RHH and RLH are defined as follows.

- ''Distance vision (Right eye) [-]'': We define $>= 1.0$ and others for RHH and RLH criteria, respectively.
- ''Distance vision (Left eye) [-]'': We define $>= 1.0$ and others for RHH and RLH criteria, respectively.
- ''Distance vision (Both eyes) [-]'': We define $>=$ 1.0 and others for RHH and RLH criteria, respectively.
- ''Near vision (Right eye) [-]'': We define $>= 1.0$ and others for RHH and RLH criteria, respectively.
- ''Near vision (Left eye) [-]'': We define $>= 1.0$ and others for RHH and RLH criteria, respectively.
- ''Near vision (Both eyes) [-]'': We define $>= 1.0$ and others for RHH and RLH criteria, respectively.

5) BLOOD EXAMINATION

For the following 15 items, we divide each value into 2 divisions referring to public information [36] (LSI Medience Corporation, Tokyo, Japan) (but only for Fischer's ratio, see [37]). The specific threshold values corresponding to RHH and RLH are defined as follows.

- ''Total bilirubin [mg/dL] (for liver diagnosis)'': We define 0.2–1.2 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Asparate aminotransferase (AST) (Glutamic oxaloacet ic transaminase (GOT)) [U/L] (for liver diagnosis)'': We define 10–40 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Alanine aminotransferase (ALT) (Glutamic pyruvate transaminase (GPT)) [U/L] (for liver diagnosis)'': We define 5–45 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Gamma-glutamyl transferase (γ-GT) (Gamma-glutam yl transpeptidase (γ-GTP)) [U/L] (for liver diagnosis)'': For males, we define $<= 80$ and others (abnormal) for RHH and RLH criteria, respectively. For females, we define $<= 30$ and others (abnormal) for RHH and RLH criteria, respectively.
- ''Total protein [g/dL] (for nutrition diagnosis)'': We define 6.7–8.3 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Creatinine [mg/dL] (for kidney diagnosis)'': For males, we define 0.61–1.04 and others (abnormal) for RHH and RLH criteria, respectively. For females, we define 0.47–0.79 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Uric acid [mg/dL] (for gout diagnosis)'': For males, we define 3.8–7.0 and others (abnormal) for RHH and RLH criteria, respectively. For females, we define 2.5–7.0 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Triglyceride (TG) [mg/dL] (for lipid diagnosis)'': We define 30–149 and others (abnormal) for RHH and RLH criteria, respectively.
- ''High-density lipoprotein (HDL) cholesterol [mg/dL] (for lipid diagnosis)'': For males, we define 40–85 and others (abnormal) for RHH and RLH criteria, respectively. For females, we define 40–95 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Albumin (ALB) by improved bromcresol purple (BCP) method [g/dL] (for nutrition diagnosis)'': We define 3.8–5.2 and others (abnormal) for RHH and RLH criteria, respectively.
- ''White blood cell count [/μL] (for immunodiagnosis)'': We define 3300–9000 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Hemoglobin [g/dL] (for anemia diagnosis)'': For males, we define 13.5–17.5 and others (abnormal) for RHH and RLH criteria, respectively. For females, we define 11.5–15.0 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Hemoglobin A1c (HbA1c) based on the national glycohemoglobin standardization program (NGSP) [%] (for diabetes diagnosis)'': We define 4.6–6.2 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Fischer's ratio [-] (for amino acid diagnosis)'': We define 2.31–4.29 and others (abnormal) for RHH and RLH criteria, respectively.
- ''Cortisol [μg/dL] (for stress diagnosis)'': We define 3.7–19.4 and others (abnormal) for RHH and RLH criteria, respectively.

For the following 1 item, we divide each value into 5 divisions based on 0 to less than 0.1, 0.1 to less than 0.2, 0.2 to less than 0.3, 0.3 to less than 0.4, and more than 0.4, while including dummies. Then, referring to [38], the specific threshold values corresponding to RHH and RLH are defined as follows.

- ''High sensitive C-reactive protein (CRP) [mg/dL] (for inflammation diagnosis)'': We define < 0.4 and $>= 0.4$ for RHH and RLH criteria, respectively.

6) BLOOD PRESSURE MEASUREMENT OF THE EXTREMITIES

For the following 4 items, we divide each value into 2 divisions referring to [39] (or [40]). The specific threshold values corresponding to RHH and RLH are defined as follows.

- ''Right brachial-ankle pulse wave velocity (RbaPWV) [cm/s] (for atherosclerosis diagnosis)'': We define

$<$ 1400 and others (abnormal) for RHH and RLH criteria, respectively.

- "Left brachial-ankle pulse wave velocity (LbaPWV) [cm/s] (for atherosclerosis diagnosis)": We define $<$ 1400 and others (abnormal) for RHH and RLH criteria, respectively.
- "Right ankle-brachial index (Rabi) [-] (for artery occlusion diagnosis)": We define $>$ 0.9 and others (abnormal) for RHH and RLH criteria, respectively.
- "Left ankle-brachial index (Labi) [-] (for artery occlusion diagnosis)": We define $>$ 0.9 and others (abnormal) for RHH and RLH criteria, respectively.

7) PULMONARY FUNCTION TESTING (SPIROMETRY)

For the following 1 item, we divide each value into 2 divisions referring to [41] (or [35]). The specific threshold values corresponding to RHH and RLH are defined as follows.

- "Gaensler's forced expiratory volume in one second percent (FEV1.0%G) [%] (for obstructive pulmonary disease diagnosis)": We define $>=$ 80 and others (abnormal) for RHH and RLH criteria, respectively.

8) HEEL BONE DENSITY TEST

For the following 1 item, we divide each value into 2 divisions referring to [42] (or [43]). The specific threshold values corresponding to RHH and RLH are defined as follows.

- "T-score (YAM) derived from the osteo sono-assessment index (OSI) [%] (for osteoporosis diagnosis)": We define $>=$ 80 and others (abnormal) for RHH and RLH criteria, respectively.

9) BODY COMPOSITION MEASUREMENT

For the following 1 item, we divide each value into 3 divisions referring to [44] (or [45], [46]). The specific threshold values corresponding to RHH and RLH are defined as follows.

- "Body fat percentage [%]": For males, we define 10–19 and others (non-standard) for RHH and RLH criteria, respectively. For females, we define 20–29 and others (non-standard) for RHH and RLH criteria, respectively.

10) GRAVIC BODY SWAY TEST (OTORHINOLARYNGOLOGY)

The Gravicorder (ANIMA Corporation, Tokyo, Japan) [47] is used for this test. For the following 1 item, there are no explicit criteria, and for this one, the actual diagnostic levels (common or large) divided by otorhinolaryngologists in clinical testing are directly used for divisions. Incidentally, see [48] for a detailed methodology on this item.

- "Evaluation of large sway (1: Large or 0: Common) [-]": We define 0 and 1 for RHH and RLH criteria, respectively.

We show the main items for 2019 in Table 3, while we also handle the same items for 2016–2018. Although some items are continuous variables and some are discrete variables, in this study, we handle each item as a discrete variable categorized into several levels. As a threshold for categorization, we establish a criterion to separate relatively high or low levels of health in the numerical values of each item. In addition, we also calculate the four-year slopes for each item over 2016–2019 for each subject, categorize them into discrete variables according to a level divided into 5 equal parts between the maximum and minimum values, and handle them together.

Next, the driving questionnaire survey items are shown in Table 4. We handle the items "Driving frequency", "Driving time", "Transmission: Automatic (AT) or Manual (MT)", "Likes or dislikes about driving", "Moments when you feel better while driving", "Scenes where you have felt better while driving", "Excitements ("Waku-Waku" in Japanese) that you can relate to", and "Purpose of driving". Since all of the above items have response options, they are all discrete variables.

The items in Table 4 and their response options are detailed below. These are all of our own design. However, only in the design of the item "Purpose of driving", some factors such as shopping, working, leisure and hobby, which are considered in [49] in relation to health, are referred to as response options. On the other hand, there are no previous studies that incorporate the other items in detail. For example, although there is a known study using driving style questionnaires (DSQ) designed to analyze the association with traffic accidents [50], they do not fit our purpose of actively analyzing the association with health and the Kansei (sensitivity). Therefore, we have designed generally all of our own, with the exception of the item "Purpose of driving".

11) DRIVING QUESTIONNAIRE

The following 3 items are general questionnaires about daily driving.

- "Driving frequency": We define the response options as "Every day", "5–6 days a week", "3–4 days a week", "Less than 2 days a week", and "None".
- "Driving time": We define the response options as "Less than 30 minutes", "30 minutes to 1 hour", "1–2 hours", and "More than 2 hours".
- "Transmission: Automatic (AT) or Manual (MT)": We define the response options as "Only AT", "Only MT", and "Both AT and MT".

The following 1 item is a questionnaire that takes into account a key factor that may be strongly related to the Kansei (sensitivity) leading to active driving.

- "Likes or dislikes about driving": We define the response options as "Like", "Rather like", "Rather dislike", and "Dislike".

The following 4 items are questionnaires that take into account other factors that may be related to the Kansei (sensitivity) for driving. For each item, the response options allows multiple answers.

- "Moments when you feel better while driving": We define the response options as "When hearing the sound

TABLE 3. Overview of our health checkup categories, items, their analysis criteria, and division explanations for subjects (1044 persons) in 2019.

Health checkup categories	Items	Analysis criteria as relatively high level of health	Analysis criteria as relatively low level of health	Number of level divisions	Supplementary explanations on how to divide
Personal record form	Age [years old]	-	-	7	Divisions into 10-year increments from 20s to 80s.
	Body height [cm]	-	-	5	5 equal divisions between the maximum and minimum.
	Body weight [kg]	-	-	5	
	Grip strength (Right hand) [kg]	>= 43.2	< 30.8	5	Each criterion represents 2 sets of thresholds excluding the center when divided into 5 equal divisions between the maximum and minimum.
	Grip strength (Left hand) [kg]	>= 36.8	< 25.2	5	
	Sitting trunk flexion [cm]	>= 45	< 33	5	
	Whole body reaction time [msec]	< 456.8	>= 685.2	5	
	Sit-to-stand test [cm]	>= 28	< 22	5	
	Timed up and go (TUG) test [sec]	< 6.98	>= 9	5	
	Two-step test [cm]	>= 262.8	< 219.2	5	
	Maximal walking speed of 10 meters [sec]	< 5.62	>= 7.28	5	
Cognitive function	Mini mental state examination (MMSE) score (0–30 points) [-]	Level 3 (27–30 points)	Level 1 (0–21 points)	3	Each criterion represents one of the thresholds when divided by Level 3 (27–30 points), Level 2 (22–26 points), and Level 1 (0–21 points).
Health questionnaire	Social connections (Number of friends who feel comfortable) [-]	>= 1	0	4	Each criterion represents a threshold of non-zero or not when divided by 0, 1–5, 6–10, and 11 or more people.
	Social connections (Number of family members and relatives who feel comfortable) [-]	>= 1	0	4	
	Pittsburgh sleep quality index (PSQI) total score (0–21 points) [-]	0–5 points (Without sleep disturbance)	6–21 points (With sleep disturbance)	2	Each criterion represents whether the score is 0–5 or not.
	Depressive state: 20-item center for epidemiologic studies depression scale (CESD-20) score (0–60 points) [-]	0–15 points (Without depression)	16–60 points (With depression)	6	Each criterion represents a subset of 0–60 scores divided into 10-point intervals (0–9, ..., 40–49, 50–60).
	Locomotive syndrome: 25-question geriatric locomotive function scale (GLFS-25) score (0-100 points) [-]	0–15 points (Without disability)	16–100 points (With disability)	10	Each criterion represents a subset of 0–100 scores divided into 10-point intervals (0–9, ..., 80–89, 90–100).
	Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Physical functioning (PF)	51–100 points	0–49 points	10	Each criterion represents a subset of 0–100 scores divided into 10-point intervals (0–9, ..., 80–89, 90–100).
	Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Role physical (RP)	51–100 points	0–49 points	10	
	Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Bodily pain (BP)	51–100 points	0–49 points	10	
	Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: General health (GH)	51–100 points	0–49 points	10	
	Medical outcomes study 36-item short-form health	51–100 points	0–49 points	10	

TABLE 3. ***(Continued.)*** **Overview of our health checkup categories, items, their analysis criteria, and division explanations for subjects (1044 persons) in 2019.**

	survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Vitality (VT)				
	Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Social functioning (SF)	51–100 points	0–49 points	10	
	Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Role emotional (RE)	51–100 points	0–49 points	10	
	Medical outcomes study 36-item short-form health survey version 2 (SF-36v2) score (0–100 points) [-] - Subscale: Mental health (MH)	51–100 points	0–49 points	10	
Vision test	Distance vision (Right eye) [-]	>= 1.0	Otherwise	2	Each criterion represents whether the visual acuity is more than or less than 1.0.
	Distance vision (Left eye) [-]	>= 1.0	Otherwise	2	
	Distance vision (Both eyes) [-]	>= 1.0	Otherwise	2	
	Near vision (Right eye) [-]	>= 1.0	Otherwise	2	
	Near vision (Left eye) [-]	>= 1.0	Otherwise	2	
	Near vision (Both eyes) [-]	>= 1.0	Otherwise	2	
Blood examination	Total bilirubin [mg/dL] (for liver diagnosis)	>= 0.2 and <= 1.2	Otherwise (Abnormal)	2	Each criterion refers to thresholds used in clinical testing.
	Asparate aminotransferase (AST) (Glutamic oxaloacetic transaminase (GOT)) [U/L] (for liver diagnosis)	>= 10 and <= 40	Otherwise (Abnormal)	2	
	Alanine aminotransferase (ALT) (Glutamic pyruvate transaminase (GPT)) [U/L] (for liver diagnosis)	>= 5 and <= 45	Otherwise (Abnormal)	2	
	Gamma-glutamyl transferase (γ-GT) (Gamma-glutamyl transpeptidase (γ-GTP)) [U/L] (for liver diagnosis)	<= 80 (for males), <= 30 (for females)	Otherwise (Abnormal)	2	
	Total protein [g/dL] (for nutrition diagnosis)	>= 6.7 and <= 8.3	Otherwise (Abnormal)	2	
	Creatinine [mg/dL] (for kidney diagnosis)	>= 0.61 and <= 1.04 (for males), >= 0.47 and <= 0.79 (for females)	Otherwise (Abnormal)	2	
	Uric acid [mg/dL] (for gout diagnosis)	>= 3.8 and <= 7.0 (for males), >= 2.5 and <= 7.0 (for females)	Otherwise (Abnormal)	2	
	Triglyceride (TG) [mg/dL] (for lipid diagnosis)	>= 30 and <= 149	Otherwise (Abnormal)	2	
	High-density lipoprotein (HDL) cholesterol [mg/dL] (for lipid diagnosis)	>= 40 and <= 85 (for males), >= 40 and <= 95 (for females)	Otherwise (Abnormal)	2	
	Albumin (ALB) by improved bromcresol purple (BCP) method [g/dL] (for nutrition diagnosis)	>= 3.8 and <= 5.2	Otherwise (Abnormal)	2	
	White blood cell count [/μL] (for immunodiagnosis)	>= 3300 and <= 9000	Otherwise (Abnormal)	2	

TABLE 3. ***(Continued.)*** **Overview of our health checkup categories, items, their analysis criteria, and division explanations for subjects (1044 persons) in 2019.**

	Hemoglobin [g/dL] (for anemia diagnosis)	>= 13.5 and <= 17.5 (for males), >= 11.5 and <= 15.0 (for females)	Otherwise (Abnormal)	2	
	Hemoglobin A1c (HbA1c) based on the national glycohemoglobin standardization program (NGSP) [%] (for diabetes diagnosis)	>= 4.6 and <= 6.2	Otherwise (Abnormal)	2	
	Fischer's ratio [-] (for amino acid diagnosis)	>= 2.31 and <= 4.29	Otherwise (Abnormal)	2	
	Cortisol [μg/dL] (for stress diagnosis)	>= 3.7 and <= 19.4	Otherwise (Abnormal)	2	
	High sensitive C-reactive protein (CRP) [mg/dL] (for inflammation diagnosis)	< 0.4	>= 0.4	5	Each criterion represents a subset of thresholds when divided into 5 divisions based on 0 to less than 0.1, 0.1 to less than 0.2, 0.2 to less than 0.3, 0.3 to less than 0.4, and more than 0.4.
Blood pressure measurement of the extremities	Right brachial-ankle pulse wave velocity (RbaPWV) [cm/s] (for atherosclerosis diagnosis)	< 1400	Otherwise (Abnormal)	2	Each criterion refers to thresholds used in clinical testing.
	Left brachial-ankle pulse wave velocity (LbaPWV) [cm/s] (for atherosclerosis diagnosis)	< 1400	Otherwise (Abnormal)	2	
	Right ankle-brachial index (Rabi) [-] (for artery occlusion diagnosis)	> 0.9	Otherwise (Abnormal)	2	
	Left ankle-brachial index (Labi) [-] (for artery occlusion diagnosis)	> 0.9	Otherwise (Abnormal)	2	
Pulmonary function testing (Spirometry)	Gaensler's forced expiratory volume in one second percent (FEV1.0%G) [%] (for obstructive pulmonary disease diagnosis)	>= 80	Otherwise (Abnormal)	2	Each criterion refers to thresholds used in clinical testing.
Heel bone density test	T-score (YAM) derived from the osteo sono-assessment index (OSI) [%] (for osteoporosis diagnosis)	>= 80	Otherwise (Abnormal)	2	Each criterion refers to thresholds used in clinical testing.
Body composition measurement	Body fat percentage [%]	>= 10 and <= 19 (for males), >= 20 and <= 29 (for females)	Otherwise (Non-standard)	3	Each criterion refers to thresholds used in clinical testing.
Gravic body sway test (Otorhinolaryngology)	Evaluation of large sway (1: Large or 0: Common) [-]	0	1	2	Each criterion refers to thresholds used in clinical testing.

of the engine running'', ''When putting your foot on the accelerator'', ''When hearing the engine accelerating'', ''When feeling enough acceleration'', ''When driving around a series of curves'', ''When turning the curves as you want'', ''When seeing the shining body of your car'', ''When feeling the texture of the interior'', ''When feeling that your car's design stands out from others'', ''When achieving fuel-efficient driving'', and ''None''.

- ''Scenes where you have felt better while driving'': We define the response options as ''Where feeling that you have your own space'', ''Where discovering a new route'', ''Where chatting and laughing with everyone'', ''Where traveling or driving with everyone'', ''Where spending time in the car with your family'', ''Where driving alone'', ''Where smelling the natural scent of the sea or greenery'', ''Where seeing the natural scenery of the sea or greenery'', and ``None''.

- ``Excitements (``Waku-Waku'' in Japanese) that you can relate to'': We define the response options as ``The sense of speed'', ``Handling'', ``The harmony with nature'', ``Exploring new routes'', ``The standout design'', ``Achieving fuel efficiency'', ``Your own space'', and ``Chatting and laughing with everyone''.
- ``Purpose of driving'': We define the response options as ``Commuting'', ``Shopping'', ``Working'', ``Farming'', ``Picking up and droping off'', ``Going for the leisure'', ``Going to the hospital'', ``Going for the hobby'', ``Going for the entertainment'', and ``Driving''.

We transform the discrete variables for driving and health described above into one-hot vectorized variables (i.e., binary variables with 1s for the corresponding attributes and 0s otherwise) according to their levels, and use them in an integrated manner in our analysis. Even if the original discrete variables have some missing values, since the binary variables corresponding to the levels representing the missing values can be generated simultaneously by one-hot vectorization, we remove the missing values as part of all binary variables. Each overview of the set of binary variables used for the analysis corresponding to males and females over 60 is shown in Tables 5 and 6, respectively (partially displayed due to the large number of binary variables).

B. ANALYSIS METHOD (STEP 1: LATENT CLASSES EXTRACTION)

In order to extract the relationships between driving and health, we cluster the data of over 1000 variables (one-hot vectorized binary variables) as shown in the previous section in an integrated manner. For this purpose, we use the Probabilistic Latent Semantic Analysis (PLSA) method [17], a machine learning technique (see Appendix A). The PLSA is a method originally proposed to extract several latent topics (latent classes) in the field of natural language processing for text clustering. Since the essence of PLSA is to soft cluster data structured as a co-occurrence matrix of documents (rows) and words (columns), it can be applied not only to text but also to questionnaires, images, and other similar data [12], [13]. The hyperparameters in PLSA include the number of latent classes K and the threshold ε related to the termination condition of the EM algorithm, which are used as $K = 5$ (for easy interpretability) and $\varepsilon = 10^{-8}$ in this study, respectively. The C language is used for our program implementations of PLSA.

Moreover, although PLSA is a soft clustering method, in this study, to facilitate the interpretation of each latent class, we derive K latent classes $C^{(k)}$ $(k = 1, \ldots, K)$ obtained from samples (subjects)

$$d \in D = \{d_1, \ldots, d_N\}$$
$$\text{with } N = \begin{cases} 160 & \text{(for males over 60)}, \\ 252 & \text{(for females over 60)} \end{cases} \quad (1)$$

TABLE 4. Overview of our items of driving questionnaire and their response options for subjects (1044 persons) in 2019.

Items of driving questionnaire	Response options
Driving frequency	Every day
	5–6 days a week
	3–4 days a week
	Less than 2 days a week
	None
Driving time	Less than 30 minutes
	30 minutes to 1 hour
	1–2 hours
	More than 2 hours
Transmission: Automatic (AT) or Manual (MT)	Only AT
	Only MT
	Both AT and MT
Likes or dislikes about driving	Like
	Rather like
	Rather dislike
	Dislike
Moments when you feel better while driving (Multiple answers allowed)	When hearing the sound of the engine running
	When putting your foot on the accelerator
	When hearing the engine accelerating
	When feeling enough acceleration
	When driving around a series of curves
	When turning the curves as you want
	When seeing the shining body of your car
	When feeling the texture of the interior
	When feeling that your car's design stands out from others
	When achieving fuel-efficient driving
	None
Scenes where you have felt better while driving (Multiple answers allowed)	Where feeling that you have your own space
	Where discovering a new route
	Where chatting and laughing with everyone
	Where traveling or driving with everyone
	Where spending time in the car with your family
	Where driving alone
	Where smelling the natural scent of the sea or greenery
	Where seeing the natural scenery of the sea or greenery
	None
Excitements (``Waku-Waku'' in Japanese) that you can relate to (Multiple answers allowed)	The sense of speed
	Handling
	The harmony with nature
	Exploring new routes
	The standout design
	Achieving fuel efficiency
	Your own space
	Chatting and laughing with everyone
Purpose of driving (Multiple answers allowed)	Commuting
	Shopping
	Working
	Farming
	Picking up and droping off
	Going for the leisure
	Going to the hospital
	Going for the hobby
	Going for the entertainment
	Driving

and factors (binary variables)

$$w \in W = \{A_1, \ldots, A_M\}$$
$$\text{with } M = \begin{cases} 1038 & \text{(for males over 60)}, \\ 1088 & \text{(for females over 60)} \end{cases} \quad (2)$$

as a pair $C^{(k)} = \left(C_d^{(k)}, C_w^{(k)}\right)$ of the set $C_d^{(k)}$ of component samples d and the set $C_w^{(k)}$ of component factors w with

TABLE 5. **Overview of a set of binary variables by one-hot vectorization (for males over 60).**

Binary variables *w* (for males over 60)	Names of variables (corresponding to the 2016–2019 health checkup results and their four-year slopes, and the 2019 driving questionnaire results, one-hot vectorized by those levels)
A1	Y19<Personal record form> Age [years old] _"60-69"
A2	Y19<Personal record form> Age [years old] _"70-79"
A3	Y19<Personal record form> Age [years old] _"80-89"
A4	Y19<Personal record form> Grip strength (Right hand) [kg] _">=18.4 and <30.8"
A5	Y19<Personal record form> Grip strength (Right hand) [kg] _">=30.8 and <43.2"
A6	Y19<Personal record form> Grip strength (Right hand) [kg] _">=43.2 and <55.6"
A7	Y19<Personal record form> Grip strength (Right hand) [kg] _">=(Min)6 and <18.4"
A8	Y19<Personal record form> Grip strength (Left hand) [kg] _">=25.2 and <36.8"
A9	Y19<Personal record form> Grip strength (Left hand) [kg] _">=36.8 and <48.4"
A10	Y19<Personal record form> Grip strength (Left hand) [kg] _">=13.6 and <25.2"
A11	Y19<Personal record form> Grip strength (Left hand) [kg] _">=48.4 and <=(Max)60"
A12	Y19<Personal record form> Grip strength (Left hand) [kg] _">=(Min)2 and <13.6"
⋮	⋮
A719	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=90 and <=(Max)100"
A720	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=60 and <70"
A721	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=70 and <80"
A722	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=50 and <60"
A723	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=80 and <90"
A724	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=30 and <40"
A725	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=40 and <50"
A726	Slope<Personal record form> Grip strength (Right hand) [kg] _">=-4.4 and <-0.6"
A727	Slope<Personal record form> Grip strength (Right hand) [kg] _">=-0.6 and <3.2"
A728	Slope<Personal record form> Grip strength (Right hand) [kg] _">=3.2 and <=(Max)7"
A729	Slope<Personal record form> Grip strength (Right hand) [kg] _">=-8.2 and <-4.4"
A730	Slope<Personal record form> Grip strength (Right hand) [kg] _">=(Min)-12 and <-8.2"
⋮	⋮
A980	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=-7.4 and <3.2"
A981	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=3.2 and <13.8"
A982	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=(Min)-18 and <-7.4"
A983	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=13.8 and <24.4"
A984	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=24.4 and <=(Max)35"
A985	Y19Car<Driving questionnaire> Driving frequency _"Every day"
A986	Y19Car<Driving questionnaire> Driving frequency _"5-6 days a week"
A987	Y19Car<Driving questionnaire> Driving frequency _"3-4 days a week"
A988	Y19Car<Driving questionnaire> Driving frequency _"Less than 2 days a week"
A989	Y19Car<Driving questionnaire> Driving frequency _"None"
⋮	⋮
A1029	Y19Car<Driving questionnaire> Purpose of driving _"Commuting"
A1030	Y19Car<Driving questionnaire> Purpose of driving _"Shopping"
A1031	Y19Car<Driving questionnaire> Purpose of driving _"Working"

TABLE 5. ***(Continued.)*** **Overview of a set of binary variables by one-hot vectorization (for males over 60).**

A1032	Y19Car<Driving questionnaire> Purpose of driving _"Farming"
A1033	Y19Car<Driving questionnaire> Purpose of driving _"Picking up and droping off"
A1034	Y19Car<Driving questionnaire> Purpose of driving _"Going for the leisure"
A1035	Y19Car<Driving questionnaire> Purpose of driving _"Going to the hospital"
A1036	Y19Car<Driving questionnaire> Purpose of driving _"Going for the hobby"
A1037	Y19Car<Driving questionnaire> Purpose of driving _"Going for the entertainment"
A1038	Y19Car<Driving questionnaire> Purpose of driving _"Driving"

maximum membership probabilities. That is, we derive $C_d^{(k)}$ and $C_w^{(k)}$ as

$$C_d^{(k)} = \left\{ d \in D | P(z_k|d) = \max_{z \in Z} P(z|d) \right\},$$
$$C_w^{(k)} = \left\{ w \in W | P(z_k|w) = \max_{z \in Z} P(z|w) \right\}. \quad (3)$$

C. ANALYSIS METHOD (STEP 2: ADDITIONAL ANALYSIS FOCUSED ON CHARACTERISTIC LATENT CLASSES)

As an additional analysis for the latent classes extracted by PLSA, when one of the latent classes is selected, we integrate its main component factors (binary variables) and several other target variables (binary variables) defined to focus on as a data set, and analyze the relationships between the component factors and target variables. For this purpose, we use the t-distributed Stochastic Neighbor Embedding (t-SNE) method [18], a machine learning (especially a manifold learning) technique (see Appendix B), as an approach that can evaluate similarities among all variables even if the target variables are defined as sparse (consisting of a few 1s) one-hot vectors. The t-SNE is a method that allows dimensional compression and visualization of data in a way that maximally satisfies the preservation of local structures of data (i.e., keeping data close together at close distances), while also ensuring that global structures of data are as intact as possible. Since the essence of t-SNE is to characterize data with distance-based similarities, it can be widely applied to a variety of data [14], [16]. The hyperparameters in t-SNE include the dimension n' after compression, the perplexity, the maximum number of iterations $t_{\max}$, the learning rate η, and the momentum $\alpha(t)$, which are used as $n' = 2$, $perplexity = 5$, $t_{\max} = 1000$, $\eta = 200$, and $\alpha(t) = 0.5(t \leq 250)$ or $0.8(t > 250)$ in this study, respectively. We have determined these values by referring to typical values in [18] as a guide. The R language is used for our program implementations of t-SNE.

TABLE 6. Overview of a set of binary variables by one-hot vectorization (for females over 60).

Binary variables *w* (for females over 60)	Names of variables (corresponding to the 2016–2019 health checkup results and their four-year slopes, and the 2019 driving questionnaire results, one-hot vectorized by those levels)
A1	Y19<Personal record form> Age [years old] _"70-79"
A2	Y19<Personal record form> Age [years old] _"60-69"
A3	Y19<Personal record form> Age [years old] _"80-89"
A4	Y19<Personal record form> Grip strength (Right hand) [kg] _">=18.4 and <30.8"
A5	Y19<Personal record form> Grip strength (Right hand) [kg] _">=30.8 and <43.2"
A6	Y19<Personal record form> Grip strength (Right hand) [kg] _">=(Min)6 and <18.4"
A7	Y19<Personal record form> Grip strength (Left hand) [kg] _">=13.6 and <25.2"
A8	Y19<Personal record form> Grip strength (Left hand) [kg] _">=25.2 and <36.8"
A9	Y19<Personal record form> Grip strength (Left hand) [kg] _">=(Min)2 and <13.6"
⋮	⋮
A758	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=70 and <80"
A759	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=80 and <90"
A760	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=90 and <=(Max)100"
A761	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=50 and <60"
A762	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=40 and <50"
A763	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=60 and <70"
A764	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=30 and <40"
A765	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=(Min)0 and <10"
A766	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=20 and <30"
A767	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=10 and <20"
A768	Slope<Personal record form> Grip strength (Right hand) [kg] _">=0.76 and <2.78"
A769	Slope<Personal record form> Grip strength (Right hand) [kg] _">=-1.26 and <0.76"
A770	Slope<Personal record form> Grip strength (Right hand) [kg] _">=-3.28 and <-1.26"
A771	Slope<Personal record form> Grip strength (Right hand) [kg] _">=2.78 and <=(Max)4.8"
A772	Slope<Personal record form> Grip strength (Right hand) [kg] _">=(Min)-5.3 and <-3.28"
⋮	⋮
A1030	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=4 and <12"
A1031	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=-4 and <4"
A1032	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=-12 and <-4"
A1033	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=12 and <=(Max)20"
A1034	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=(Min)-20 and <-12"
A1035	Y19Car<Driving questionnaire> Driving frequency _"Every day"
A1036	Y19Car<Driving questionnaire> Driving frequency _"5-6 days a week"
A1037	Y19Car<Driving questionnaire> Driving frequency _"3-4 days a week"
A1038	Y19Car<Driving questionnaire> Driving frequency _"Less than 2 days a week"
A1039	Y19Car<Driving questionnaire> Driving frequency _"None"
⋮	⋮
A1079	Y19Car<Driving questionnaire> Purpose of driving _"Commuting"
A1080	Y19Car<Driving questionnaire> Purpose of driving _"Shopping"
A1081	Y19Car<Driving questionnaire> Purpose of driving _"Working"
A1082	Y19Car<Driving questionnaire> Purpose of driving _"Farming"
A1083	Y19Car<Driving questionnaire> Purpose of driving _"Picking up and droping off"
A1084	Y19Car<Driving questionnaire> Purpose of driving _"Going for the leisure"
A1085	Y19Car<Driving questionnaire> Purpose of driving _"Going to the hospital"
A1086	Y19Car<Driving questionnaire> Purpose of driving _"Going for the hobby"
A1087	Y19Car<Driving questionnaire> Purpose of driving _"Going for the entertainment"
A1088	Y19Car<Driving questionnaire> Purpose of driving _"Driving"

III. RESULTS

A. RESULTS OF PRELIMINARY ANALYSIS TO DETERMINE THE NUMBER OF PLSA'S LATENT CLASSES, INCLUDING COMPARISON WITH CLASSIC STATISTICAL PLOTS USING EXISTING ITEMS

We first show the results of the preliminary analysis using PLSA described in the previous section and simple statistical plots (stratified scatter plots). Before analyzing the relationship between driving and health among the elderly over 60, we have examined whether there are driving questionnaire items that can simply stratify health, once using only 2018 gender-inclusive data for example, which also includes those aged 20–60. However, due to large variances in most health item data, we have been able to find no cases where they can be simply stratified by any of the driving items. As some examples, in Fig. 2, we show preliminary analysis results with scatter plots for all males and females over 20 in 2018, stratified by ''Driving frequency'', ''Driving time'', ''Transmission: Automatic (AT) or Manual (MT)'', and ''Likes or dislikes about driving'', respectively. In this figure, for each stratification, the left side plots age on the horizontal axis and CESD-20 score on the vertical axis, while the right side plots age on the horizontal axis and GLFS-25 score on the vertical axis. From this figure, it can be immediately understood that the data is difficult to adequately capture relationships using average values or conventional linear analysis (e.g., correlation analysis). Therefore, in order to efficiently extract the relationship with driving while also setting a large number of other health items, it is essential to deal with uncertainty and cluster the items and subjects simultaneously to find some stratifiable categories in a flexible manner other than the predefined driving items. This has led us to use PLSA as a kind of probabilistic modeling method.

We note here that it is necessary to consider how many latent classes to set in PLSA. As a reference, it has long been known that the Akaike information criterion (AIC) can be applied to regular models that are guaranteed to have asymptotically Gaussian distributions in the maximum likelihood estimators of model parameters [51]. However, it is not strictly applicable in PLSA involving latent class models that do not necessarily have such guarantees. Therefore, in this study, we have decided to determine the number of latent classes by making several specific changes. As a result, as shown in Fig. 3, we have confirmed that increasing the number of latent classes up to approximately $K = 10$ makes stratification by these latent classes relatively easier. In fact, for example, at least in the upper left and lower right parts of the plot (i.e., the parts easily understood as declining health from young to old), we can see that relatively different latent classes (especially $C^{(7)}$ and $C^{(2)}$) appear and support stratification, respectively. Based on these results, we have believed that if we focus on subjects over 60, the target of our analysis, we should set the number of latent classes to approximately $K = 5$ and analyze the data by gender to capture the reality of the data relatively more easily, although this is a qualitative judgment. Although this does not correspond to an objective approach to determining the number of latent classes, it has been our priority to ensure ease of understanding and interpretability in accordance with the actual data situation.

B. RESULTS OF LATENT CLASSES EXTRACTION USING PLSA

The results for five latent classes corresponding to males and females over 60 are shown in Figs. 4 to 8 and Figs. 9 to 13, respectively. In each figure, the component factors $w \in C_w^{(k)}$ related to one latent class $C^{(k)}$ are plotted as a bar chart based on the descending order of $P(w|z_k)$, which corresponds to the contribution of w in $C_w^{(k)}$, with an enlarged display of top 30 plots. At the same time, a bar chart corresponding to the partial extraction of at most 10 factors related to driving only is also displayed. These factors appearing in (1) and (2) of Figs. 4 to 13 are listed in text format in Appendix C.

C. PREPARATION FOR ADDITIONAL ANALYSIS USING T-SNE

Based on the PLSA results, we focus here on cases where differences appear in other attributes that we want to focus on, even if they belong to the same latent class. Specifically, we focus on cognitive function (MMSE) as a representative health item and define

- "MMSE_slope_good": a binary variable representing subjects who drive every day and have non-negative four-year slopes for MMSE over 2016–2019

and

- "MMSE_slope_bad": a binary variable representing subjects who drive every day and have negative four-year slopes for MMSE over 2016–2019

TABLE 7. Number of subjects corresponding to "MMSE_slope_good" and "MMSE_slope_bad" in each latent class, where the upper and lower tables are for males and females over 60, respectively.

Latent classes for males over 60	Number of component subjects $\left\|C_d^{(k)}\right\|$	Number of component factors $\left\|C_w^{(k)}\right\|$	**Number of subjects corresponding to "MMSE_slope _good"**	**Number of subjects corresponding to "MMSE_slope _bad"**
$C^{(1)}$	30	336	**1**	**0**
$C^{(2)}$	36	95	**1**	**0**
$C^{(3)}$	37	130	**2**	**1**
$C^{(4)}$	29	118	**2**	**0**
$C^{(5)}$	28	359	**5**	**3**

Latent classes for females over 60	Number of component subjects $\left\|C_d^{(k)}\right\|$	Number of component factors $\left\|C_w^{(k)}\right\|$	**Number of subjects corresponding to "MMSE_slope _good"**	**Number of subjects corresponding to "MMSE_slope _bad"**
$C^{(1)}$	49	297	**1**	**1**
$C^{(2)}$	55	165	**1**	**0**
$C^{(3)}$	58	191	**0**	**1**
$C^{(4)}$	45	194	**2**	**3**
$C^{(5)}$	45	241	**0**	**0**

as two different target variables that consider driving and MMSE maintenance simultaneously. Then, if we count the number of subjects corresponding to "MMSE_slope_good" and "MMSE_slope_bad" in each latent class, we obtain the results shown in Table 7. In particular, we can find a small number of subjects corresponding to each of the target variables in latent classes $C^{(3)}$ and $C^{(5)}$ for males over 60, and in latent classes $C^{(1)}$ and $C^{(4)}$ for females over 60, respectively. Therefore, we characterize those differences through an additional analysis using t-SNE according to the following approach.

First, in each of the above latent classes $C^{(k)}$, we restrict the set of subjects $C_d^{(k)}$ to be analyzed to a subset

$$\begin{aligned} C_{d'}^{(k)} &= \left\{ d' \in C_d^{(k)} | \text{A subject } d' \text{drives every day.} \right\} \\ &= \left\{ d'_1, \ldots, d'_{N_k} \right\} \text{ with } N_k = \left| C_{d'}^{(k)} \right| \end{aligned} \tag{4}$$

consisting only of those who drive every day. Next, for each binary variable $w = A_i \in C_w^{(k)}$, we also restrict the N-dimensional one-hot vector

$$\boldsymbol{a}_i^{(k)} = [n(d_1, A_i), \ldots, n(d_N, A_i)]^T \in \{0, 1\}^N \tag{5}$$

corresponding to A_i to an N_k-dimensional partial one-hot vector

$$\boldsymbol{a}'^{(k)}_i = \left[n\left(d'_1, A_i\right), \ldots, n\left(d'_{N_k}, A_i\right) \right]^T \in \{0, 1\}^{N_k}, \tag{6}$$

where the symbol T denotes the transposition. Then, since $N_k \leq \left| C_d^{(k)} \right|$, there is a possibility to be formed as $\boldsymbol{a}'^{(k)}_i = \mathbf{0}$ (the zero vector) for some $\boldsymbol{a}'^{(k)}_i$. Therefore, we process to remove those zero vectors and let

$$C_{w'}^{(k)} = \left\{ w' \in C_w^{(k)} | \left[n\left(d'_1, w'\right), \ldots, n\left(d'_{N_k}, w'\right) \right]^T \neq \mathbf{0} \right\}$$

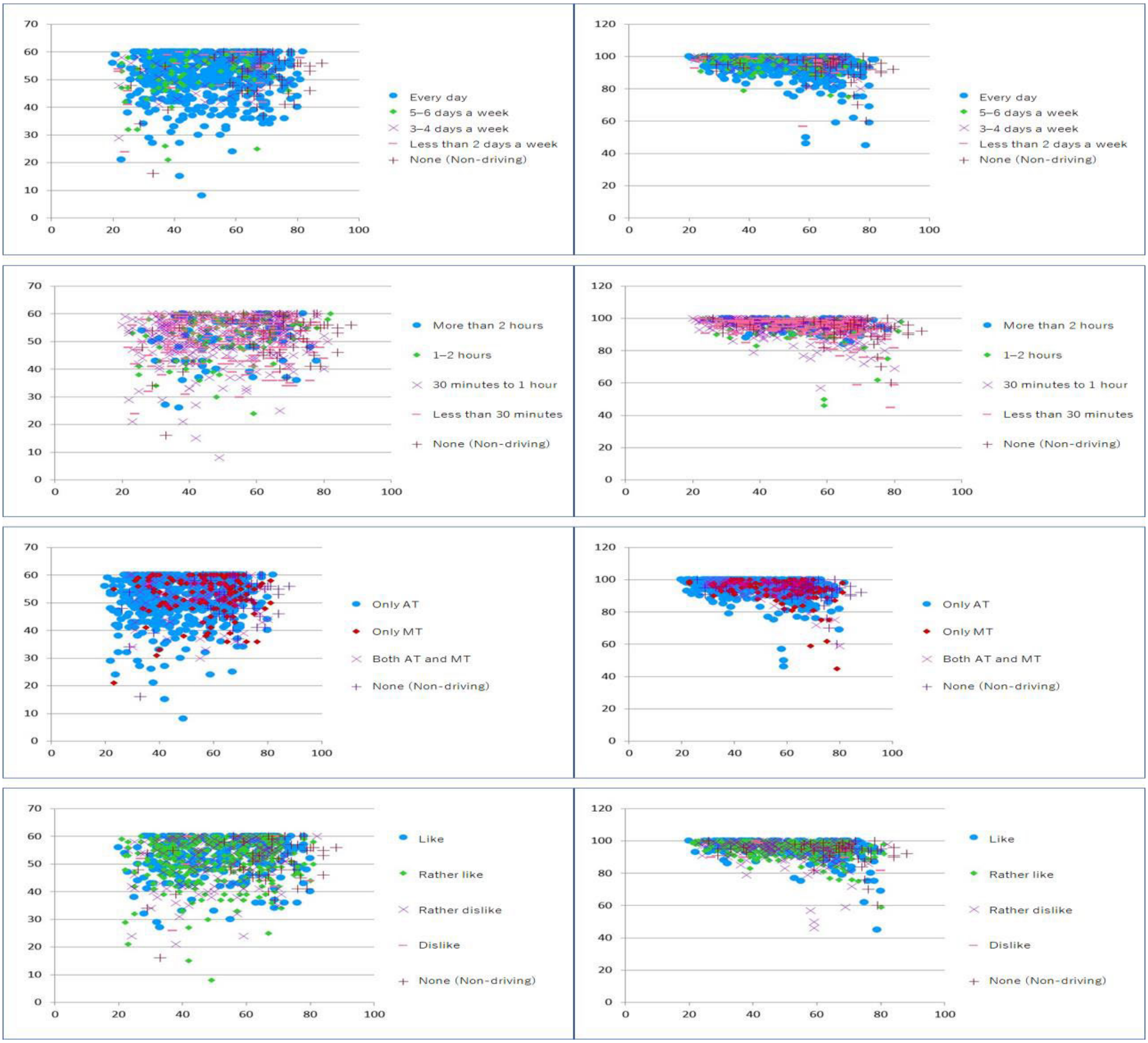

FIGURE 2. Examples of preliminary analysis results with scatter plots for all males and females over 20 in 2018, stratified by "Driving frequency", "Driving time", "Transmission: Automatic (AT) or Manual (MT)", and "Likes or dislikes about driving", respectively (on the left side, horizontal axis: age, vertical axis: CESD-20 score; on the right side, horizontal axis: age, vertical axis: GLFS-25 score).

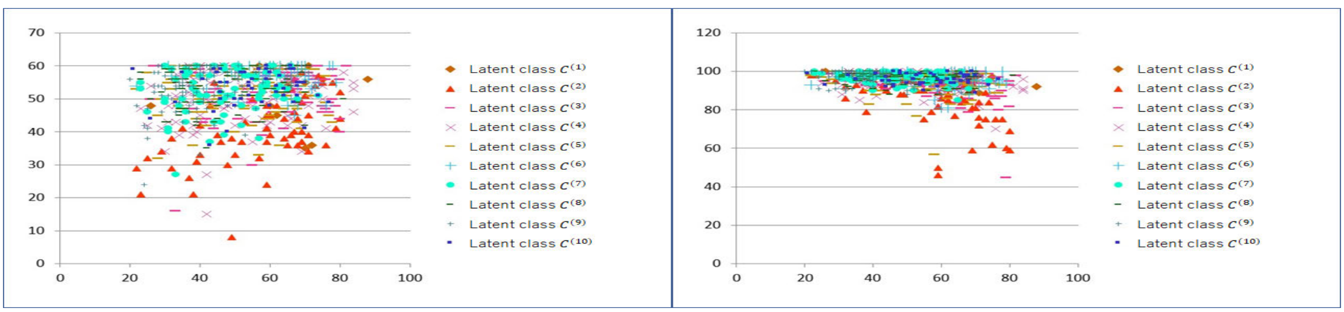

FIGURE 3. Examples of preliminary analysis results with scatter plots for all males and females over 20 in 2018, stratified by ten latent classes based on PLSA with $K = 10$ (on the left side, horizontal axis: age, vertical axis: CESD-20 score; on the right side, horizontal axis: age, vertical axis: GLFS-25 score).

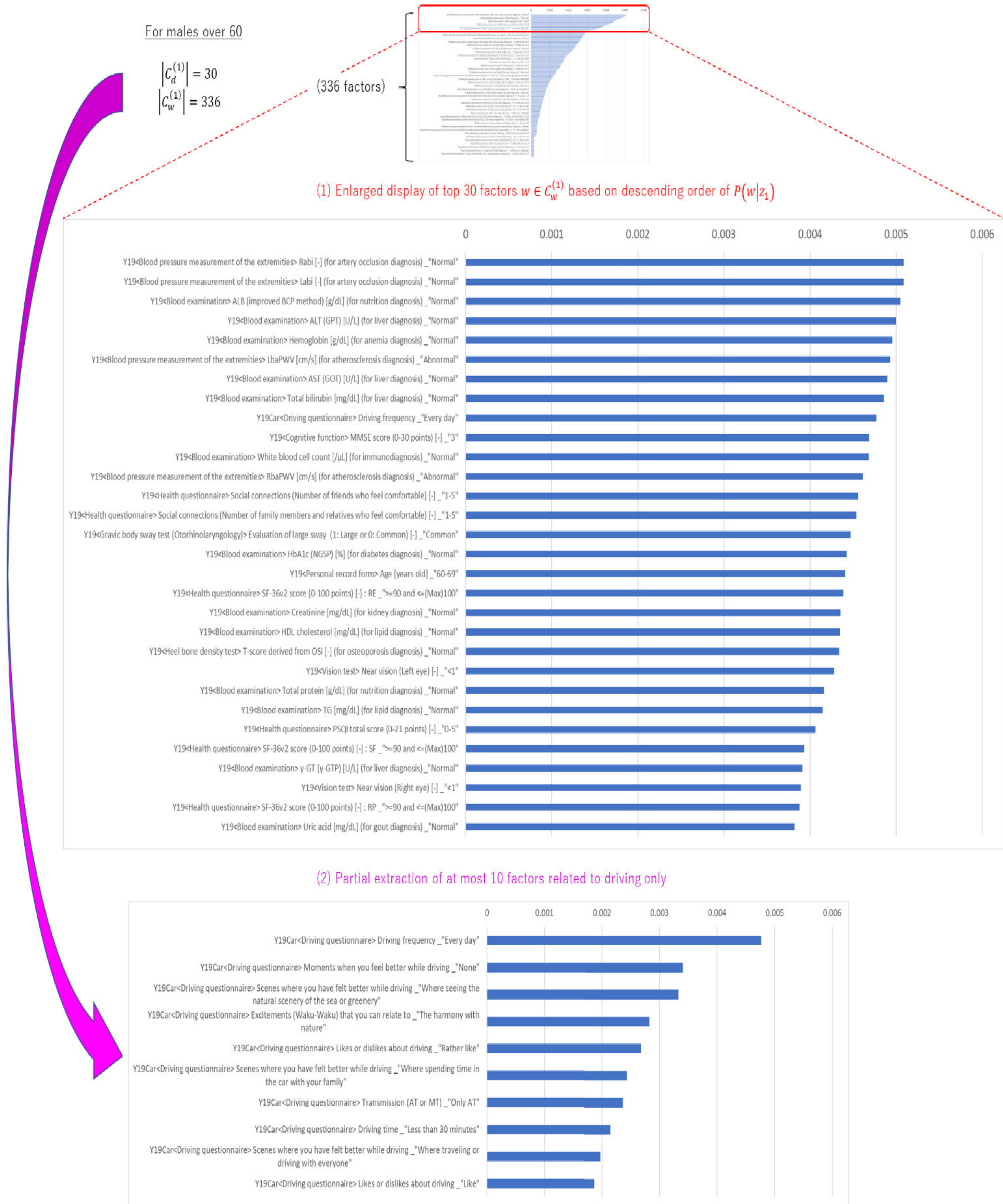

FIGURE 4. Results for the latent class $C^{(1)}$ (for males over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(1)}$ based on descending order of $P(w|z_1)$. (2) Partial extraction of at most 10 factors related to driving only.

$$= \{w'_1, \ldots, w'_{M_k}\} \text{ with } M_k = \left|C^{(k)}_{w'}\right| \tag{7}$$

be the set of binary variables to be analyzed. That is, the data $X^{(k)}$ to be input to t-SNE is given as an $N_k \times (M_k + L_k)$

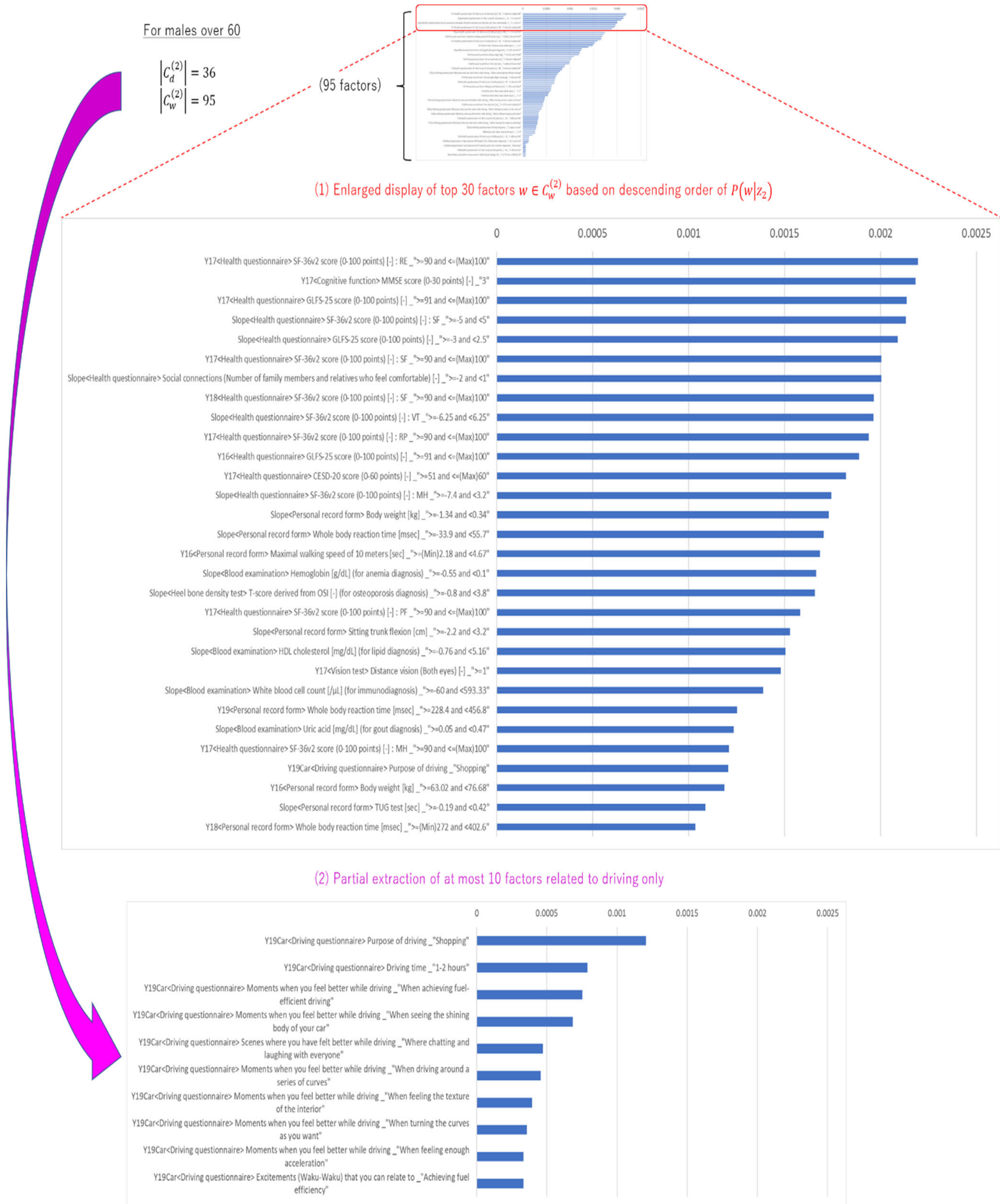

FIGURE 5. Results for the latent class $C^{(2)}$ (for males over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(2)}$ based on descending order of $P(w|z_2)$. (2) Partial extraction of at most 10 factors related to driving only.

matrix

$$X^{(k)} = \left[x_1^{(k)}, \ldots, x_{M_k}^{(k)}, s_1^{(k)}, \ldots, s_{L_k}^{(k)}\right] \tag{8}$$

that integrates M_k one-hot vectors

$$x_i^{(k)} = \left[n\left(d'_1, w'_i\right), \ldots, n\left(d'_{N_k}, w'_i\right)\right]^T \in \{0, 1\}^{N_k} \ (i = 1, \ldots, M_k) \tag{9}$$

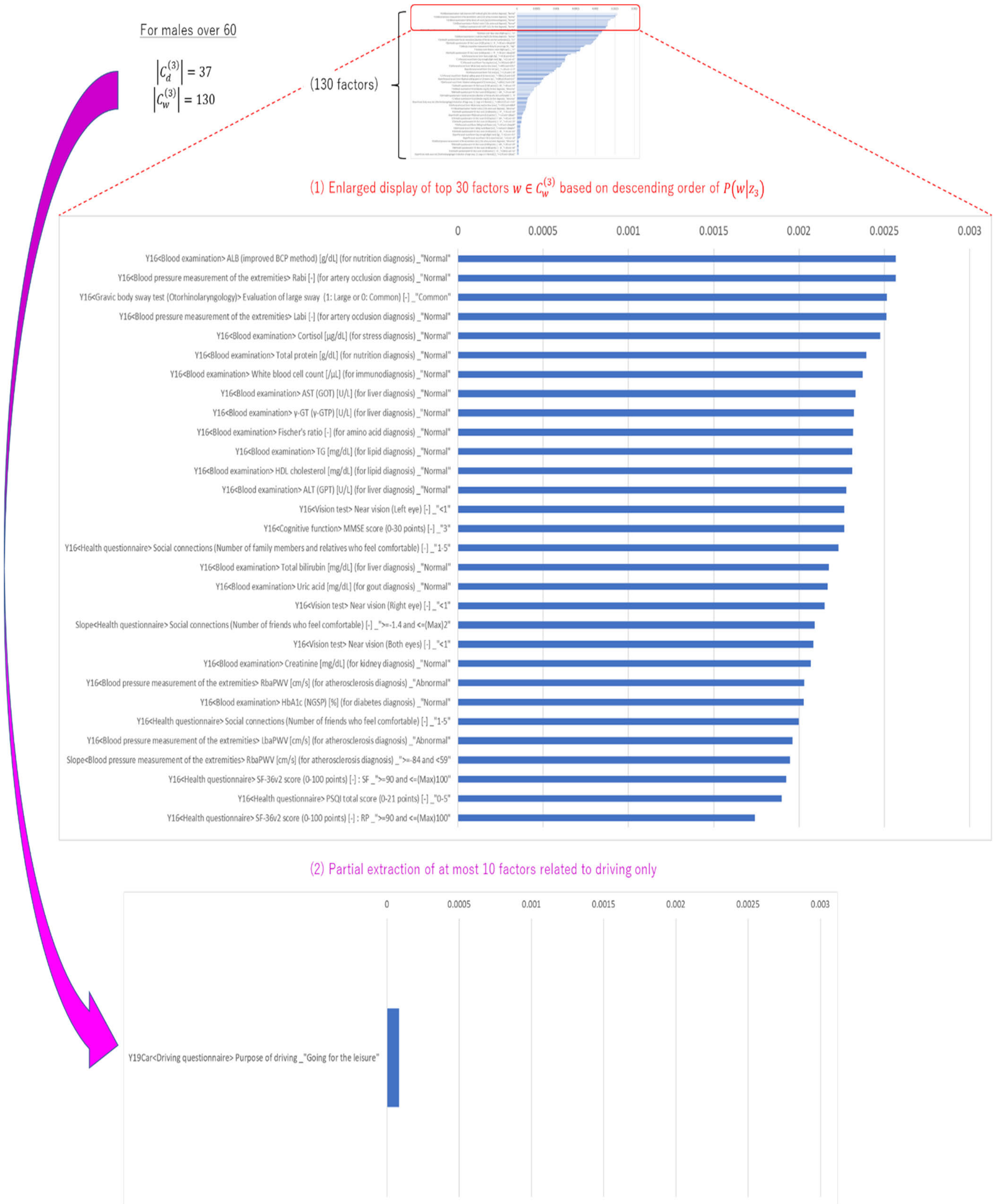

FIGURE 6. **Results for the latent class $C^{(3)}$ (for males over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(3)}$ based on descending order of $P(w|z_3)$. (2) Partial extraction of at most 10 factors related to driving only.**

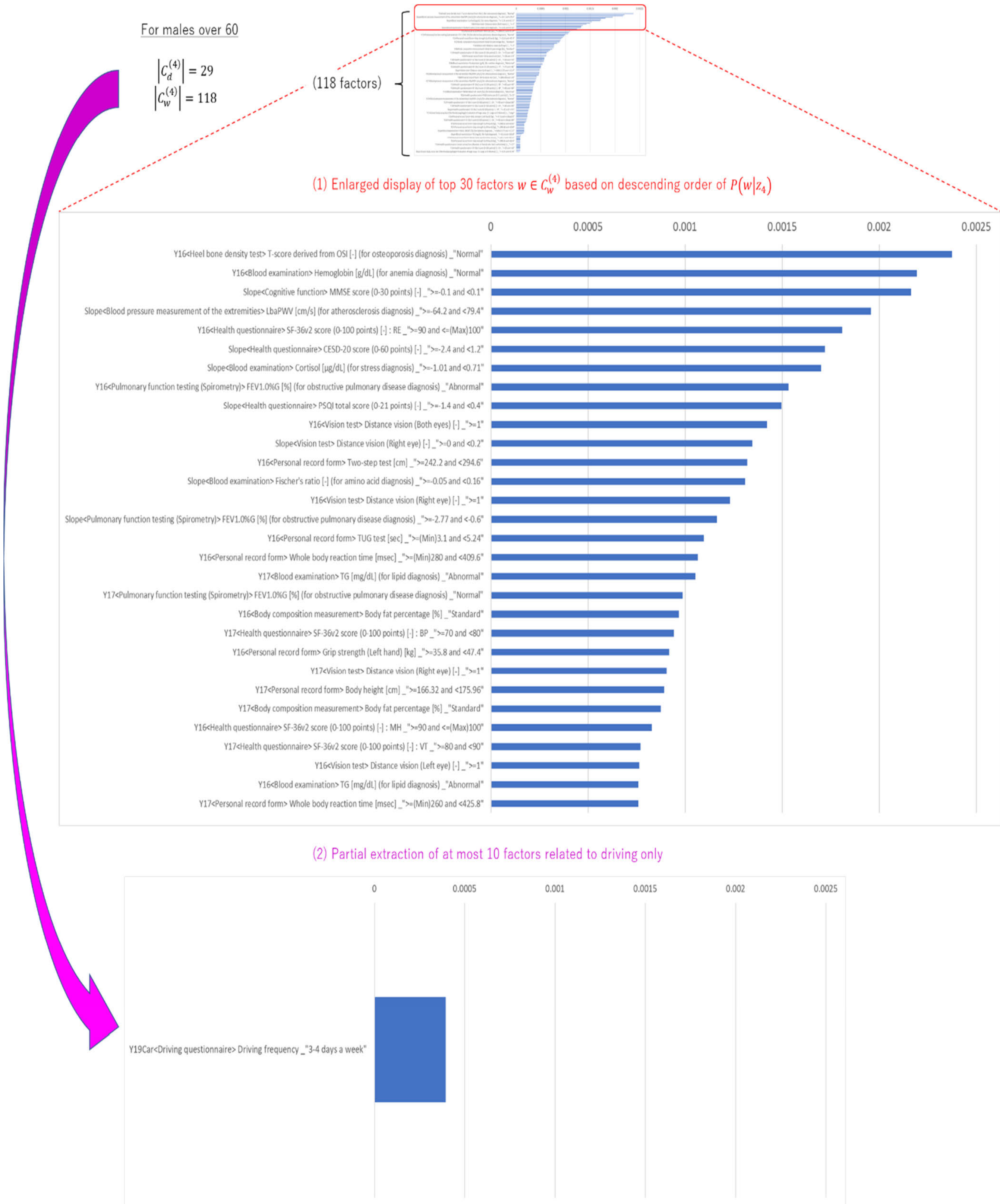

FIGURE 7. Results for the latent class $C^{(4)}$ (for males over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(4)}$ based on descending order of $P(w|z_4)$. (2) Partial extraction of at most 10 factors related to driving only.

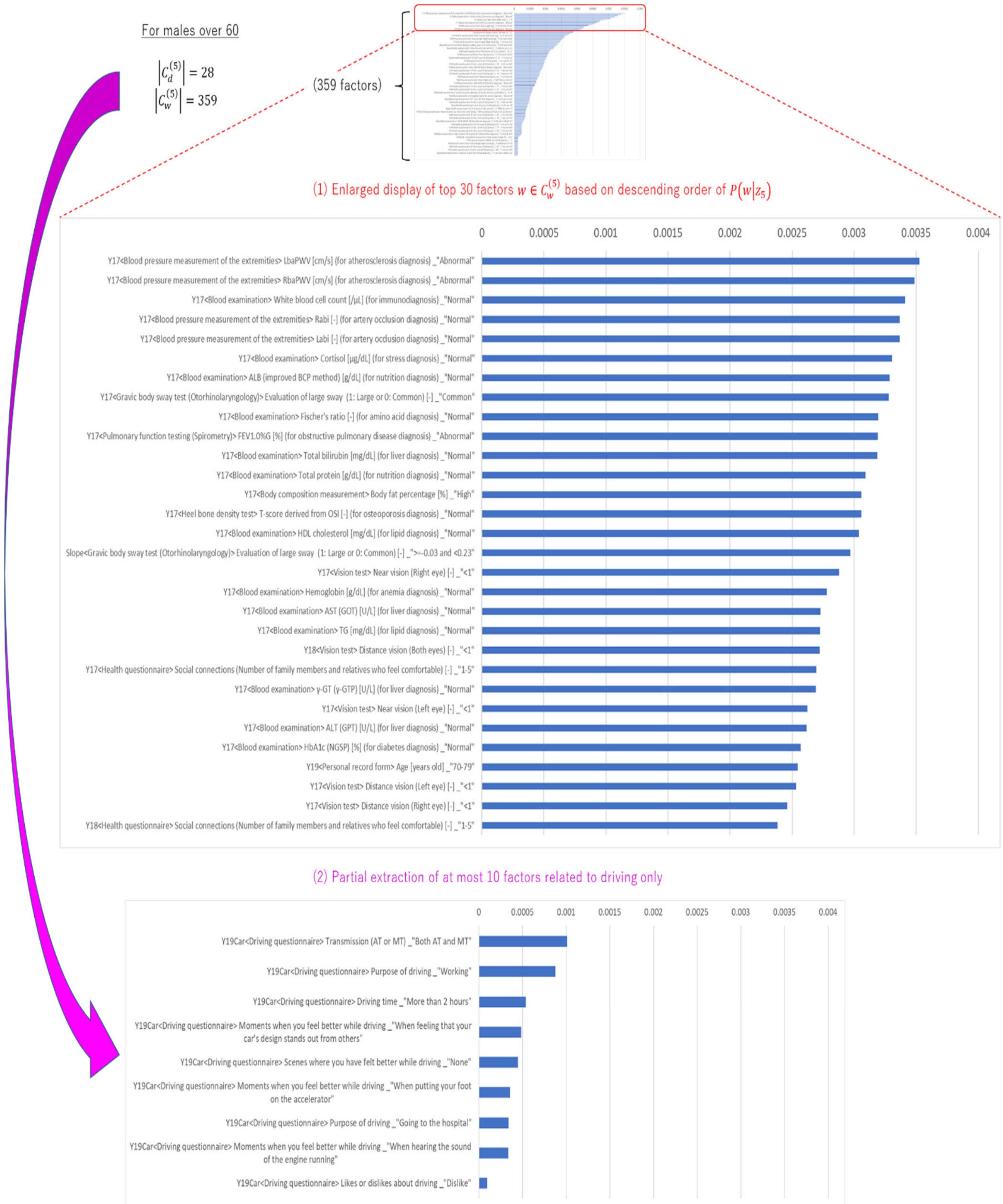

FIGURE 8. **Results for the latent class $C^{(5)}$ (for males over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(5)}$ based on descending order of $P(w|z_5)$. (2) Partial extraction of at most 10 factors related to driving only.**

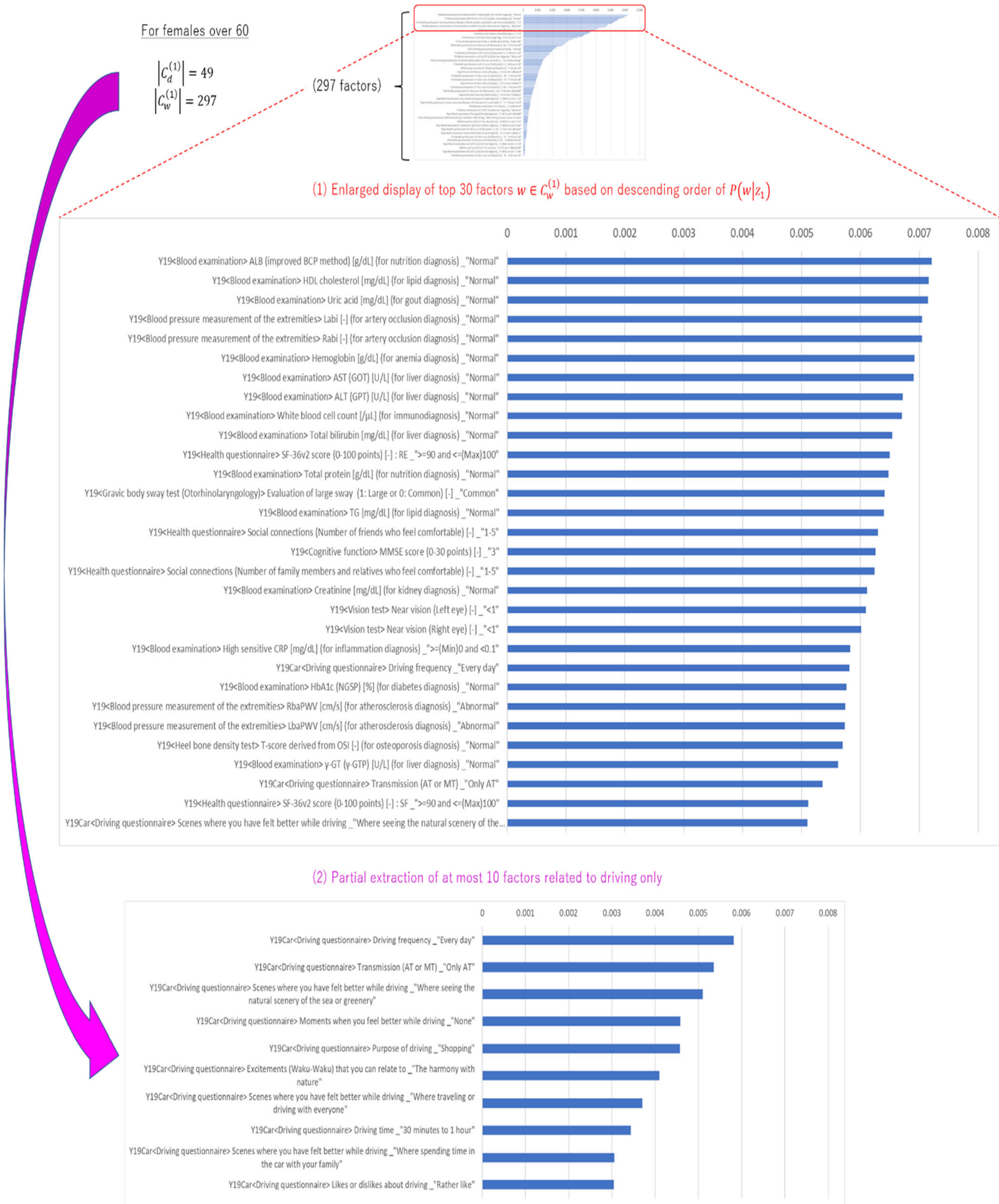

FIGURE 9. Results for the latent class $C^{(1)}$ (for females over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(1)}$ based on descending order of $P(w|z_1)$. (2) Partial extraction of at most 10 factors related to driving only.

corresponding to the component factors, and several other one-hot vectors

$$s_j^{(k)} \in \{0, 1\}^{N_k} \; (j = 1, \ldots, L_k) \tag{10}$$

corresponding to the defined target variables. Furthermore, in order to take the Kansei (sensitivity) factor into account actively, we set the number of target variables to $L_k \in \{2, 3\}$ by including not only "MMSE_slope_good" and "MMSE_slope_bad" but also a binary variable

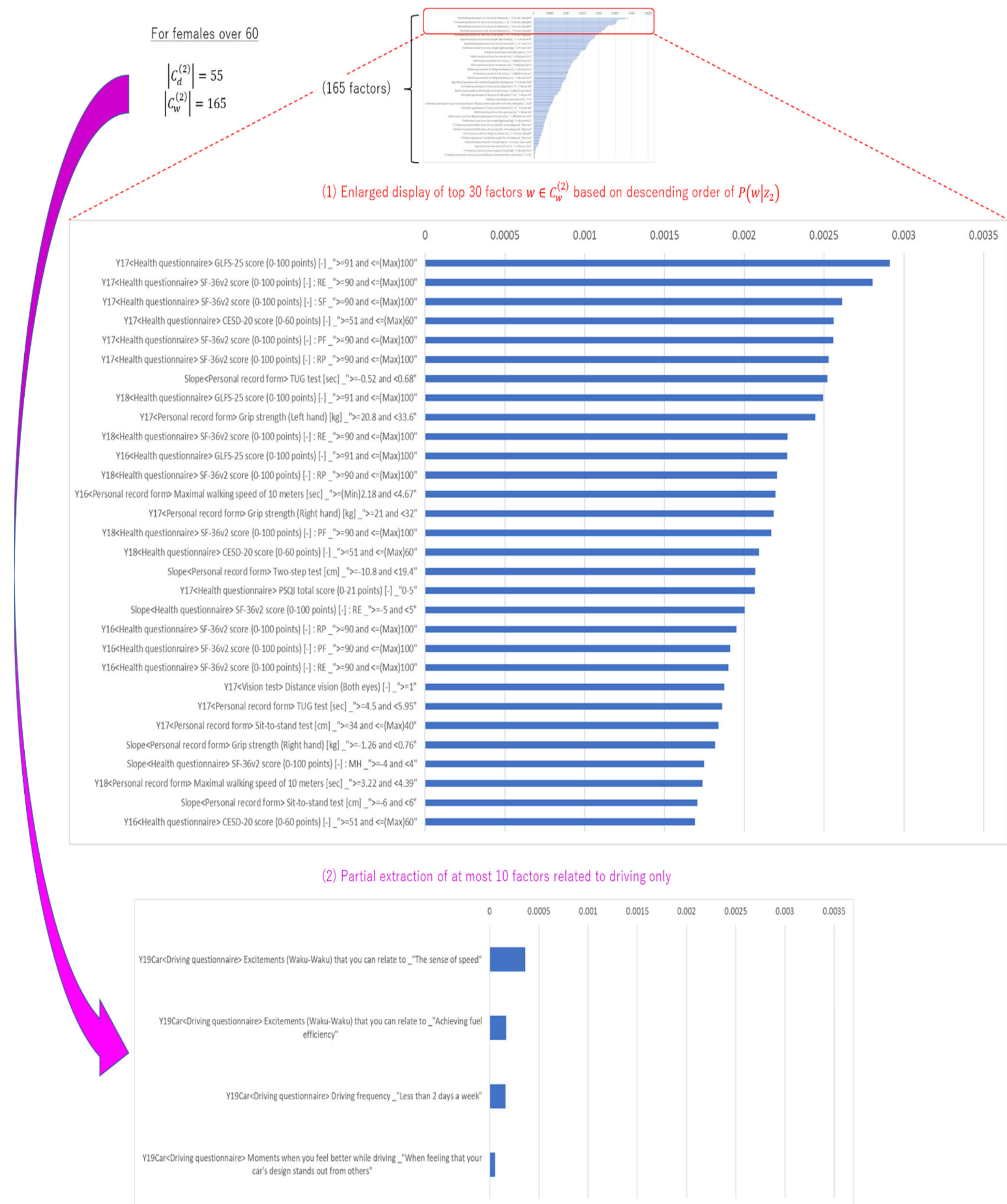

FIGURE 10. Results for the latent class $C^{(2)}$ (for females over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(2)}$ based on descending order of $P(w|z_2)$. (2) Partial extraction of at most 10 factors related to driving only.

- Y19Car<Driving questionnaire> Likes or dislikes about driving _"Like"

representing the subjects corresponding to "I like driving" (corresponding to A_{997} for males over 60 and A_{1047} for

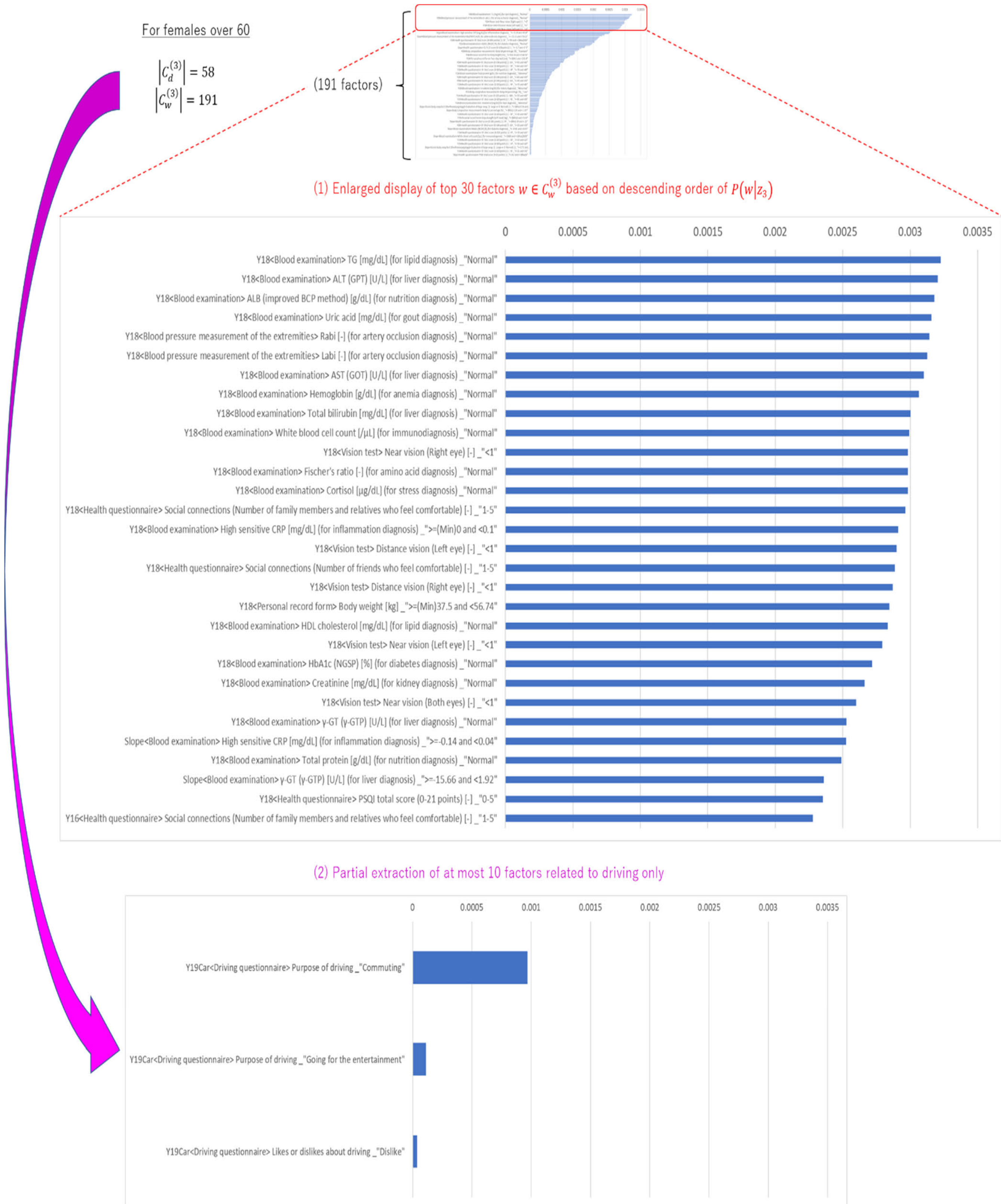

FIGURE 11. **Results for the latent class $C^{(3)}$ (for females over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(3)}$ based on descending order of $P(w|z_3)$. (2) Partial extraction of at most 10 factors related to driving only.**

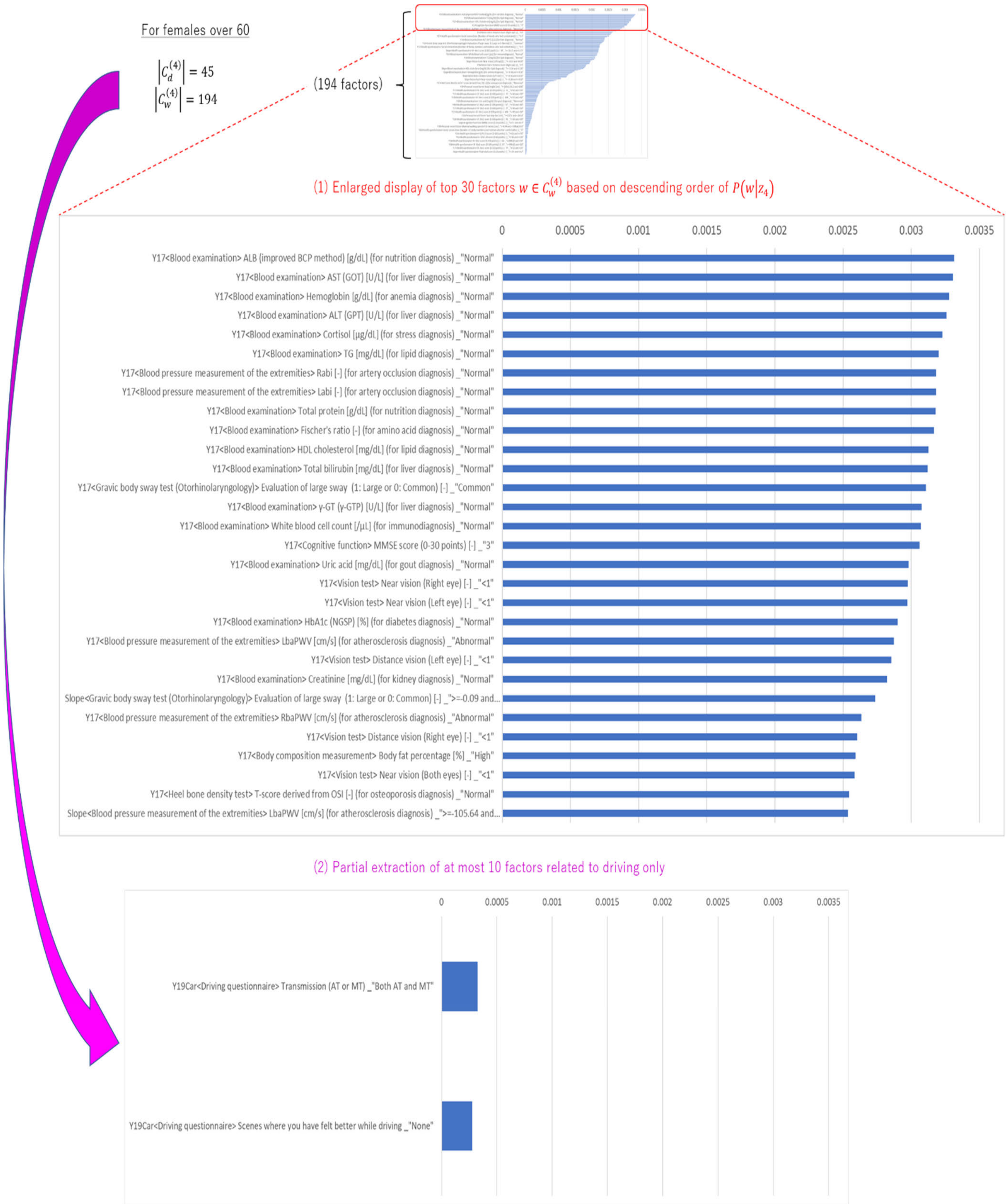

FIGURE 12. **Results for the latent class $C^{(4)}$ (for females over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(4)}$ based on descending order of $P(w|z_4)$. (2) Partial extraction of at most 10 factors related to driving only.**

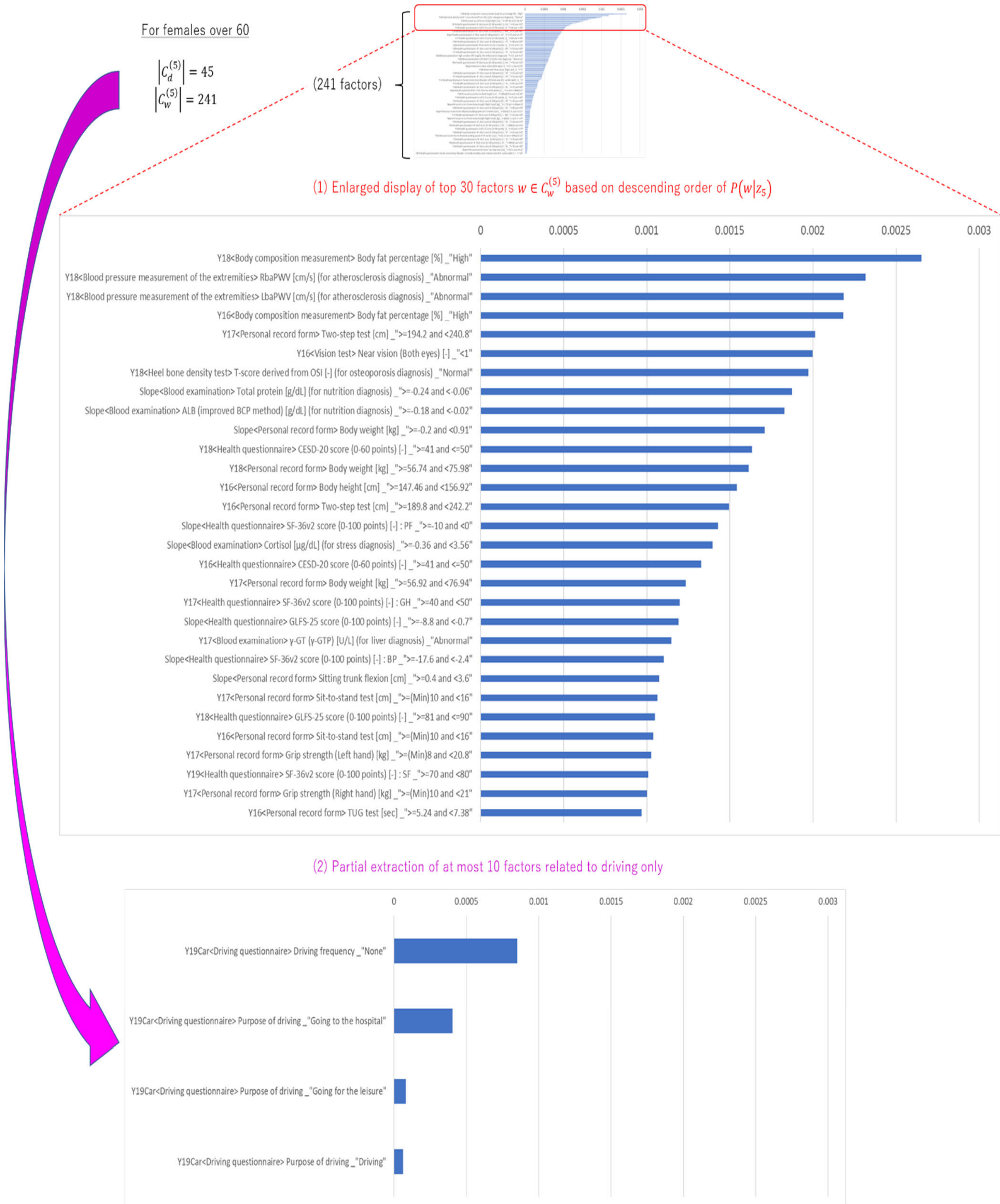

FIGURE 13. Results for the latent class $C^{(5)}$ (for females over 60). (1) Enlarged display of top 30 factors $w \in C_w^{(5)}$ based on descending order of $P(w|z_5)$. (2) Partial extraction of at most 10 factors related to driving only.

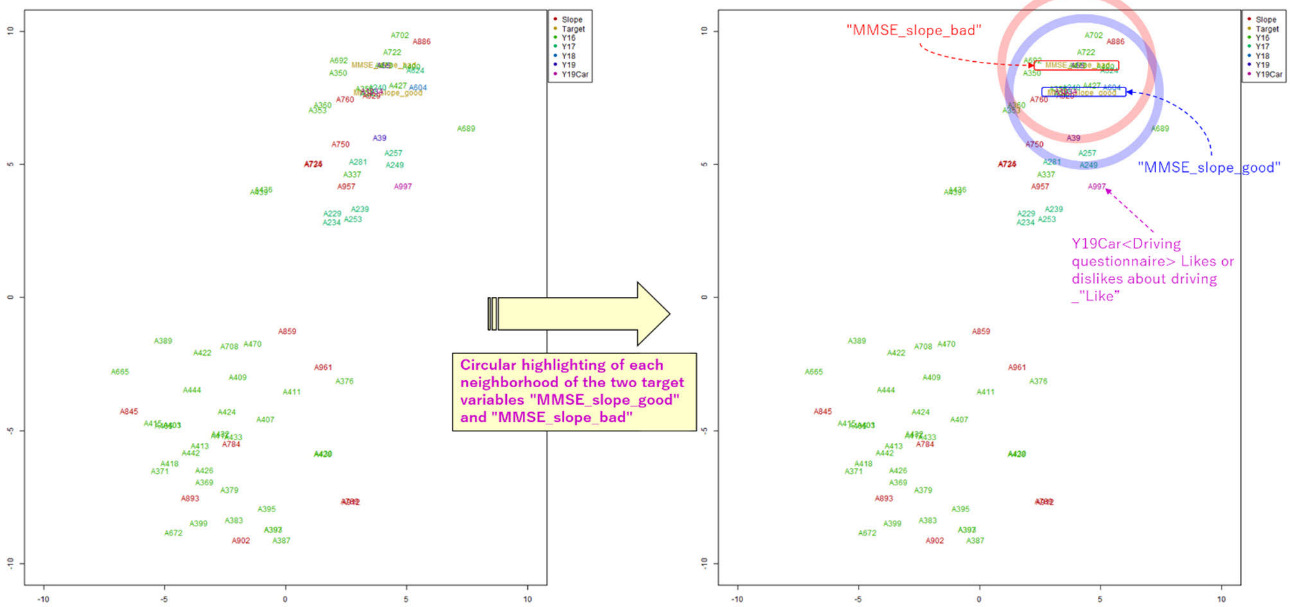

FIGURE 14. Characterization results for component factors and target variables corresponding to the latent class $C^{(3)}$ (for males over 60): The original on the left and the same on the right with target variables highlighted.

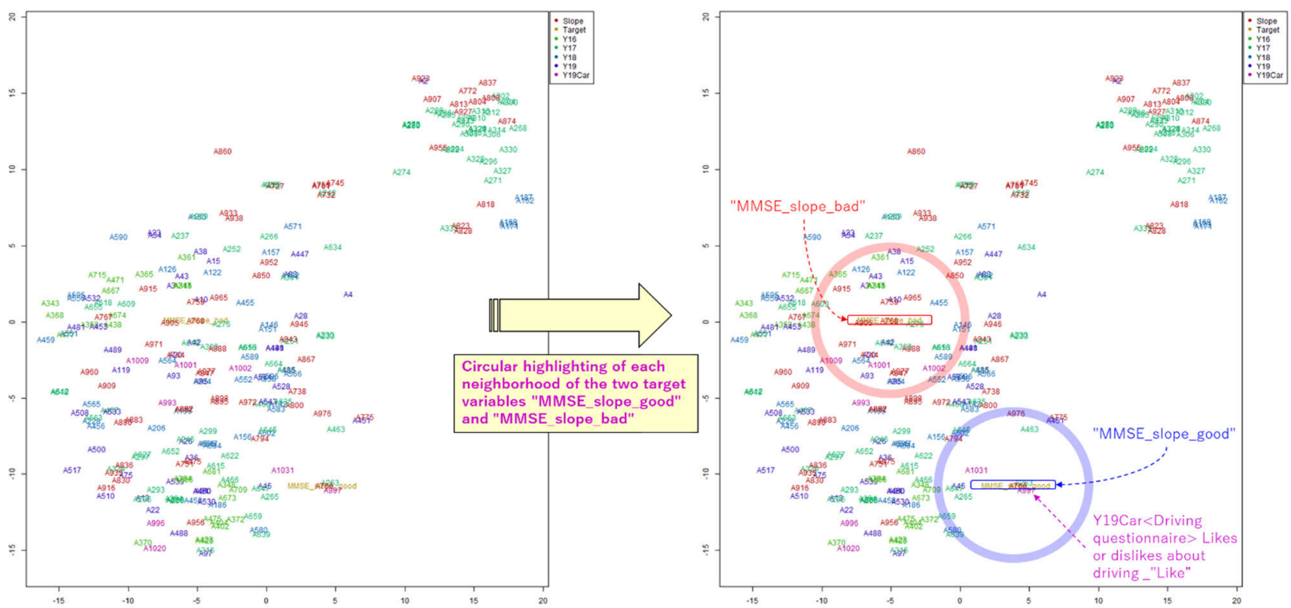

FIGURE 15. Characterization results for component factors and target variables corresponding to the latent class $C^{(5)}$ (for males over 60): The original on the left and the same on the right with target variables highlighted.

females over 60) as one of the target variables (related factors) if it is a variable not belonging as a component factor and not forming a zero vector.

D. ADDITIONAL ANALYSIS RESULTS USING T-SNE

According to the approach described in the previous section, Figs. 14 and 15 and Figs. 16 and 17 show the results of characterizing similarities among component factors and target variables for the latent classes $C^{(3)}$ and $C^{(5)}$ for males over 60 and $C^{(1)}$ and $C^{(4)}$ for females over 60, respectively. In particular, on the right side of each figure, the standing positions of the two target variables "MMSE_slope_good" and "MMSE_slope_bad" are clearly highlighted with their rough circular neighborhoods.

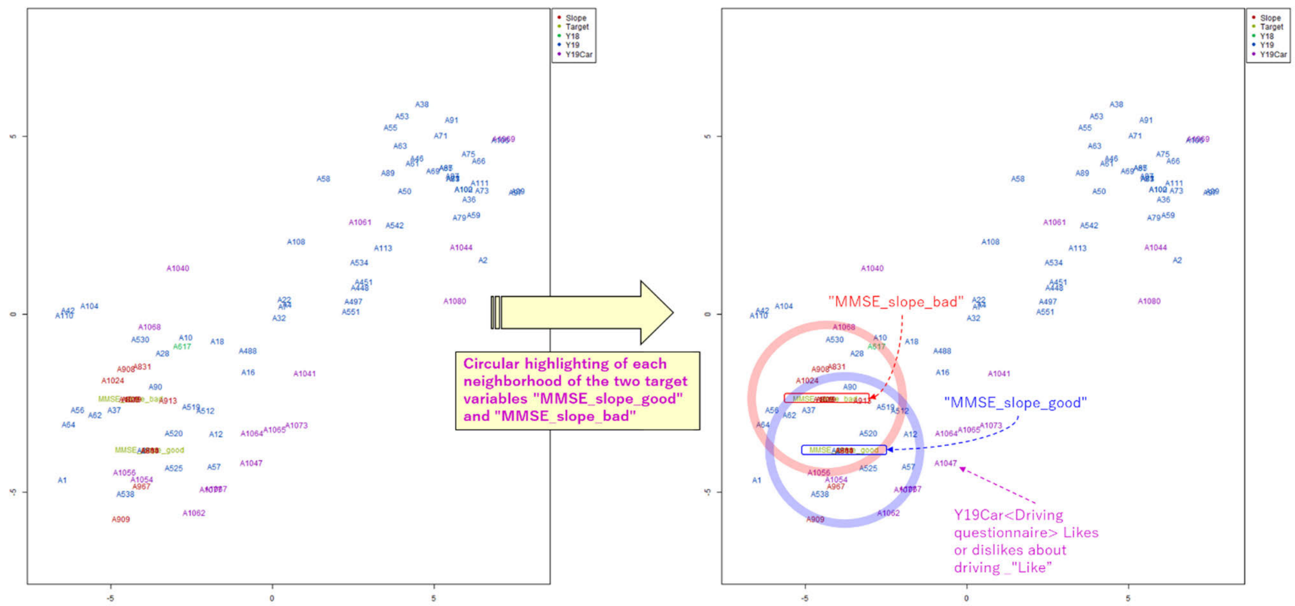

FIGURE 16. Characterization results for component factors and target variables corresponding to the latent class $C^{(1)}$ (for females over 60): The original on the left and the same on the right with target variables highlighted.

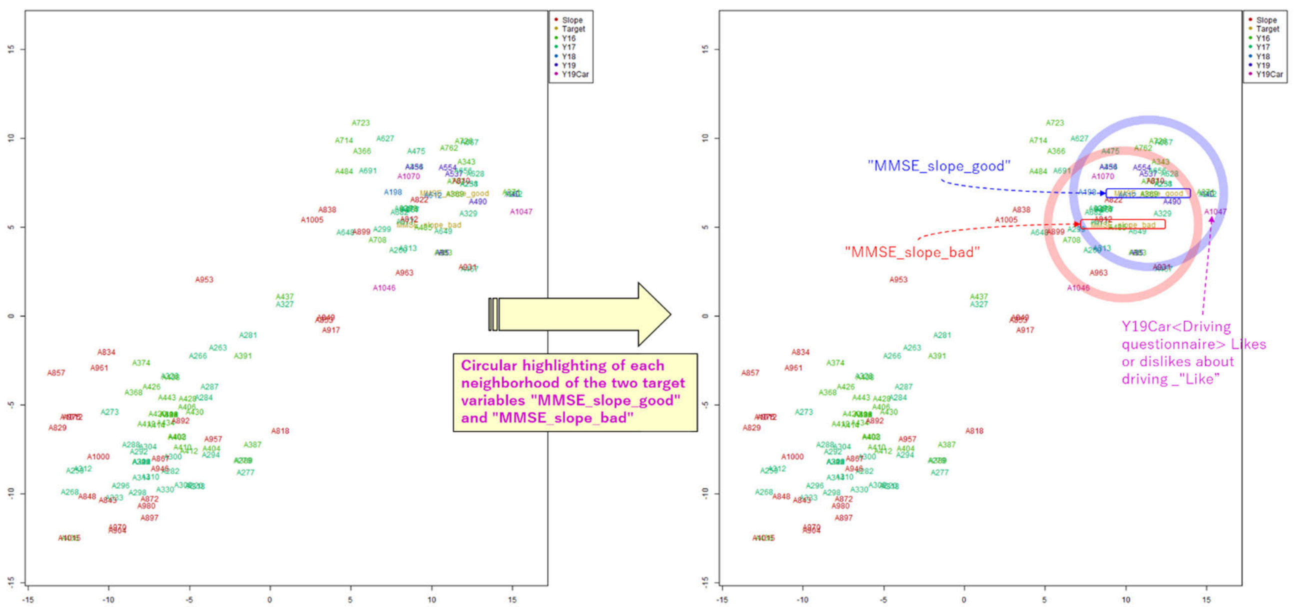

FIGURE 17. Characterization results for component factors and target variables corresponding to the latent class $C^{(4)}$ (for females over 60): The original on the left and the same on the right with target variables highlighted.

IV. DISCUSSION AND LIMITATIONS

A. INTERPRETATION AND DISCUSSION FOR PLSA RESULTS

Looking over the component factors of each latent class obtained by PLSA, we can interpret that $C^{(1)}$, $C^{(2)}$, $C^{(3)}$ and $C^{(4)}$ are broadly classified as relatively high health classes and $C^{(5)}$ as a relatively low health class, for both males over 60 and females over 60. More specifically, we can interpret the results as follows.

1) THE CASE OF MALES OVER 60

- The latent class $C^{(1)}$ can be interpreted as a relatively high health class that has the most principal feature with artery occlusion diagnosed as normal in 2019, and also has many items diagnosed as normal on the blood examination. Focusing on features with respect to driving, subjects who drive every day appear relatively in $C^{(1)}$.

TABLE 8. Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $C^{(3)}$ (for males over 60).

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _good"	**A flag variable representing subjects with non-negative MMSE slope for 4 years (2016-2019)**	**0**	**1.058576362**
A427	**Y16<Blood examination> Cortisol [µg/dL] (for stress diagnosis) _"Abnormal"**	**0.486281122**	**0.955881926**
A240	**Y17<Personal record form> Sitting trunk flexion [cm] _">=51.6 and <=(Max)64"**	**0.496415626**	**0.861061361**
A1034	**Y19Car<Driving questionnaire> Purpose of driving _"Going for the leisure"**	**0.675637015**	**1.067861974**
A929	**Slope<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Normal) [-] _">=(Min)-0.29 and <-0.03"**	**0.701093399**	**1.24122379**
A243	**Y17<Personal record form> Whole body reaction time [msec] _">=591.6 and <757.4"**	**0.796413963**	**1.207606353**
A356	**Y16<Personal record form> Sit-to-stand test [cm] _">=16 and <22"**	**0.997162146**	**1.172310703**
"MMSE_slope _bad"	A flag variable representing subjects with negative MMSE slope for 4 years (2016-2019)	1.058576362	0
A55	Y19<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _">=11"	1.058576362	0
A473	Y16<Health questionnaire> CESD-20 score (0-60 points) [-] _">=21 and <=30"	1.058576362	0
A604	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=70 and <80"	1.230909691	1.636979427
A400	Y16<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Abnormal"	1.37356348	1.158257491
A624	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=80 and <90"	1.403522479	1.323394683
A722	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=50 and <60"	1.547806682	0.612191897
A39	Y19<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)2.3 and <3.96"	1.711822066	2.719474472
A760	Slope<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)-0.49 and <0.07"	1.81123405	2.040137589
A350	Y16<Personal record form> Sitting trunk flexion [cm] _">=30.4 and <42.6"	2.223559195	1.905122595
A702	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=50 and <60"	2.238011358	1.331813881
A257	Y17<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)2.25 and <3.58"	2.253038878	3.307620531
A886	Slope<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _">=-0.26 and <-0.05"	2.390642528	1.827361088
A692	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=60 and <70"	2.396157122	1.862662077
A249	Y17<Personal record form> TUG test [sec] _">=(Min)3.04 and <4.5"	2.695725716	3.750880318
A360	Y16<Personal record form> TUG test [sec] _">=5.24 and <7.38"	2.752655201	2.923058996
A750	Slope<Personal record form> TUG test [sec] _">=-0.8 and <-0.19"	2.754775964	3.44657456
A281	Y17<Vision test> Distance vision (Left eye) [-] _">=1"	2.876800409	3.776172219
A353	Y16<Personal record form> Whole body reaction time [msec] _">=409.6 and <539.2"	3.003024769	3.202881288
A337	Y16<Personal record form> Grip strength (Right hand) [kg] _">=28.8 and <40.2"	3.39980788	4.293833932
A689	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=50 and <60"	3.473962427	4.15151389
A997	**Y19Car<Driving questionnaire> Likes or dislikes about driving _"Like"**	**3.552268929**	**4.610708248**
A957	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=-14.86 and <-1.14"	3.911601312	4.795119516

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _bad"	**A flag variable representing subjects with negative MMSE slope for 4 years (2016-2019)**	**1.058576362**	**0**
A55	**Y19<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _">=11"**	**1.058576362**	**0**
A473	**Y16<Health questionnaire> CESD-20 score (0-60 points) [-] _">=21 and <=30"**	**1.058576362**	**0**
A722	**Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=50 and <60"**	**1.547806682**	**0.612191897**
A240	Y17<Personal record form> Sitting trunk flexion [cm] _">=51.6 and <=(Max)64"	0.496415626	0.861061361
A427	Y16<Blood examination> Cortisol [µg/dL] (for stress diagnosis) _"Abnormal"	0.486281122	0.955881926
"MMSE_slope _good"	A flag variable representing subjects with non-negative MMSE slope for 4 years (2016-2019)	0	1.058576362
A1034	Y19Car<Driving questionnaire> Purpose of driving _"Going for the leisure"	0.675637015	1.067861974
A400	Y16<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Abnormal"	1.37356348	1.158257491
A356	Y16<Personal record form> Sit-to-stand test [cm] _">=16 and <22"	0.997162146	1.172310703
A243	Y17<Personal record form> Whole body reaction time [msec] _">=591.6 and <757.4"	0.796413963	1.207606353
A929	Slope<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Normal) [-] _">=(Min)-0.29 and <-0.03"	0.701093399	1.24122379
A624	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=80 and <90"	1.403522479	1.323394683
A702	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=50 and <60"	2.238011358	1.331813881
A604	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=70 and <80"	1.230909691	1.636979427
A886	Slope<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _">=-0.26 and <-0.05"	2.390642528	1.827361088
A692	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=60 and <70"	2.396157122	1.862662077

TABLE 8. ***(Continued.)*** **Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $C^{(3)}$ (for males over 60).**

A350	Y16<Personal record form> Sitting trunk flexion [cm] _">=30.4 and <42.6"	2.223559195	1.905122595
A760	Slope<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)-0.49 and <0.07"	1.81123405	2.040137589
A39	Y19<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)2.3 and <3.96"	1.711822066	2.719474472
A360	Y16<Personal record form> TUG test [sec] _">=5.24 and <7.38"	2.752655201	2.923058996
A353	Y16<Personal record form> Whole body reaction time [msec] _">=409.6 and <539.2"	3.003024769	3.202881288
A257	Y17<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)2.25 and <3.58"	2.253038878	3.307620531
A750	Slope<Personal record form> TUG test [sec] _">=-0.8 and <-0.19"	2.754775964	3.44657456
A249	Y17<Personal record form> TUG test [sec] _">=(Min)3.04 and <4.5"	2.695725716	3.750880318
A281	Y17<Vision test> Distance vision (Left eye) [-] _">=1"	2.876800409	3.776172219
A689	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=50 and <60"	3.473962427	4.15151389
A337	Y16<Personal record form> Grip strength (Right hand) [kg] _">=28.8 and <40.2"	3.39980788	4.293833932
A997	**Y19Car<Driving questionnaire> Likes or dislikes about driving _"Like"**	**3.552268929**	**4.610708248**
A726	Slope<Personal record form> Grip strength (Right hand) [kg] _">=-4.4 and <-0.6"	4.064653815	4.681423358

- The latent class $C^{(2)}$ can be interpreted as a relatively high health class that has the most principal feature with high SF-36v2 scores (especially Quality of Life related to Role Emotional) in 2017, and also has many items with high levels on the health questionnaire. Focusing on features with respect to driving, subjects who drive for the purpose of shopping appear relatively in $C^{(2)}$.
- The latent class $C^{(3)}$ can be interpreted as a relatively high health class that has the most principal feature with nutrition diagnosed as normal in 2016, and also has many items diagnosed as normal on the blood examination. Focusing on features with respect to driving, subjects who drive for the purpose of going for the leisure appear relatively in $C^{(3)}$.
- The latent class $C^{(4)}$ can be interpreted as a relatively high health class that has the most principal feature with osteoporosis diagnosed as normal in 2016, and also has subjects with high visual acuity (distance vision) or standard body fat percentage. Focusing on features with respect to driving, subjects who drive 3–4 days a week appear relatively in $C^{(4)}$.
- The latent class $C^{(5)}$ can be interpreted as a relatively low health class that has the most principal feature with atherosclerosis diagnosed as abnormal in 2017, and also has subjects with low visual acuity (distance vision) or high body fat percentage. Focusing on features with respect to driving, subjects who drive both AT and MT cars appear relatively in $C^{(5)}$. $C^{(5)}$ may also include a minority of subjects who drive reluctantly for the purpose of going to the hospital.

2) THE CASE OF FEMALES OVER 60

- The latent class $C^{(1)}$ can be interpreted as a relatively high health class that has the most principal feature with nutrition diagnosed as normal in 2019, and also has many items diagnosed as normal on the blood examination. Focusing on features with respect to driving, subjects who drive every day appear relatively in $C^{(1)}$.
- The latent class $C^{(2)}$ can be interpreted as a relatively high health class that has the most principal feature with high GLFS-25 scores in 2017, and also has many items with high levels on the health questionnaire. Focusing on features with respect to driving, subjects who can relate to the excitement for the sense of speed appear relatively in $C^{(2)}$.
- The latent class $C^{(3)}$ can be interpreted as a relatively high health class that has the most principal feature with lipid diagnosed as normal in 2018, and also has many items diagnosed as normal on the blood examination. Focusing on features with respect to driving, subjects who drive for the purpose of commuting appear relatively in $C^{(3)}$.
- The latent class $C^{(4)}$ can be interpreted as a relatively high health class that has the most principal feature with nutrition diagnosed as normal in 2017, and also has many items diagnosed as normal on the blood examination (but with high body fat percentage). Focusing on features with respect to driving, subjects who drive both AT and MT cars appear relatively in $C^{(4)}$.
- The latent class $C^{(5)}$ can be interpreted as a relatively low health class that has the most principal feature with high body fat percentage in 2018, and also has relatively low SF-36v2 scores and declining grip strength. Focusing on features with respect to driving, subjects who do not drive appear relatively in $C^{(5)}$. $C^{(5)}$ may also include a minority of subjects who drive reluctantly for the purpose of going to the hospital.

Based on the above interpretations, we can consider that there is support for a relatively positive relationship between driving and health for both males and females over 60.

B. INTERPRETATION AND DISCUSSION FOR T-SNE RESULTS

Based on the t-SNE results, in order to characterize "MMSE_slope_good" and "MMSE_slope_bad" in

TABLE 9. Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $C^{(5)}$ (for males over 60).

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _good"	**A flag variable representing subjects with non-negative MMSE slope for 4 years (2016-2019)**	**0**	**14.20499594**
A766	**Slope<Cognitive function> MMSE score (0-30 points) [-] _">=0.3 and <=(Max)0.5"**	**0.1602569**	**14.27880927**
A263	**Y17<Cognitive function> MMSE score (0-30 points) [-] _"2"**	**0.628476868**	**14.38590632**
A997	**Y19Car<Driving questionnaire> Likes or dislikes about driving _"Like"**	**0.802166576**	**14.88977486**
A1031	**Y19Car<Driving questionnaire> Purpose of driving _"Working"**	**3.014797228**	**11.65127117**
A265	**Y17<Personal record form> Body height [cm] _">=147.04 and <156.68"**	**3.843083671**	**12.70572136**
A463	**Y17<Health questionnaire> CESD-20 score (0-60 points) [-] _">=41 and <=50"**	**3.882624459**	**12.37970127**
A45	**Y19<Personal record form> Body height [cm] _">=146 and <155.8"**	**4.111713532**	**11.94079455**
A647	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=80 and <90"**	**4.451064897**	**11.95300799**
A976	**Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=-22.5 and <-7.5"**	**4.783540752**	**11.00263829**
A451	**Y19<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=81 and <=90"**	**5.196128076**	**13.64178865**
A602	**Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=50 and <60"**	**5.360923961**	**8.941835109**
A659	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=60 and <70"**	**5.386194925**	**13.46102774**
A639	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=50 and <60"**	**5.389263918**	**14.78306264**
A580	**Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=50 and <60"**	**5.403358602**	**14.4522602**
A645	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=30 and <40"**	**5.433143197**	**8.827227016**
A794	**Slope<Vision test> Distance vision (Left eye) [-] _">=0.07 and <0.28"**	**5.474150337**	**9.02137692**
A775	**Slope<Personal record form> Body weight [kg] _">=0.34 and <2.02"**	**5.491223513**	**13.76854173**
A800	**Slope<Vision test> Distance vision (Both eyes) [-] _">=0.06 and <0.28"**	**5.715444475**	**9.022048986**
A709	**Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=70 and <80"**	**6.07583269**	**11.50100974**
A583	**Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=30 and <40"**	**6.090991866**	**8.189589789**
A635	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=40 and <50"**	**6.261875628**	**8.233775583**
A142	**Y18<Personal record form> TUG test [sec] _">=6.16 and <7.71"**	**6.280023069**	**8.072448718**
A738	**Slope<Personal record form> Sitting trunk flexion [cm] _">=-7.6 and <-2.2"**	**6.495230187**	**8.668893803**
A372	**Y16<Personal record form> Body height [cm] _">=147.46 and <156.92"**	**6.62166639**	**13.32513485**
A156	**Y18<Personal record form> Body height [cm] _">=143.92 and <154.24"**	**6.625911268**	**8.317275959**
A466	**Y17<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=61 and <=70"**	**6.709748716**	**10.69214426**
A543	**Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=40 and <50"**	**6.822561119**	**7.431496753**
A673	**Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=50 and <60"**	**6.914386202**	**11.84250131**
A348	**Y16<Personal record form> Sitting trunk flexion [cm] _">=(Min)6 and <18.2"**	**6.940813895**	**10.9793205**

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _bad"	**A flag variable representing subjects with negative MMSE slope for 4 years (2016-2019)**	**14.20499594**	**0**
A768	**Slope<Cognitive function> MMSE score (0-30 points) [-] _">=(Min)-0.5 and <-0.3"**	**14.20476172**	**0.000622243**
A759	**Slope<Personal record form> Two-step test [cm] _">=-25.2 and <-10.4"**	**15.18297399**	**1.255242021**
A42	**Y19<Cognitive function> MMSE score (0-30 points) [-] _"2"**	**13.2231948**	**1.391964258**
A619	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=80 and <90"**	**13.43028672**	**1.467733448**
A10	**Y19<Personal record form> Grip strength (Left hand) [kg] _">=13.6 and <25.2"**	**15.06859793**	**1.472095506**
A276	**Y17<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _"6-10"**	**12.94248759**	**1.849643752**
A905	**Slope<Blood pressure measurement of the extremities> Rabi [-] (for artery occlusion diagnosis) _">=3.2 and <7.6"**	**15.41370827**	**1.880427856**
A303	**Y17<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _"Abnormal"**	**12.29621523**	**1.946739718**
A965	**Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=6.25 and <18.75"**	**14.49845483**	**2.292242899**
A888	**Slope<Blood examination> Cortisol [μg/dL] (for stress diagnosis) _">=0.71 and <2.43"**	**11.79935728**	**2.406228205**
A717	**Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=80 and <90"**	**16.61894408**	**2.518320312**

TABLE 9. *(Continued.)* Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $C^{(5)}$ (for males over 60).

A341	Y16<Personal record form> Grip strength (Right hand) [kg] _">=(Min)6 and <17.4"	16.61888639	2.51853447
A345	Y16<Personal record form> Grip strength (Left hand) [kg] _">=12.6 and <24.2"	16.61811288	2.518545674
A20	Y19<Personal record form> Whole body reaction time [msec] _">=913.6 and <=(Max)1142"	13.67813537	2.64066257
A744	Slope<Personal record form> Whole body reaction time [msec] _">=234.9 and <=(Max)324.5"	13.67797418	2.641611516
A3	Y19<Personal record form> Age [years old] _"80-89"	17.1290507	2.92717643
A1001	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When hearing the sound of the engine running"	12.67829095	2.971022735
A43	Y19<Cognitive function> MMSE score (0-30 points) [-] _"1"	17.13440544	3.115779573
A564	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=70 and <80"	13.92618841	3.3104195
A977	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=7.5 and <22.5"	11.27555345	3.374877722
A971	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=5 and <15"	15.36134399	3.399541855
A122	Y18<Personal record form> Grip strength (Right hand) [kg] _">=18.4 and <28.8"	16.14427203	3.405733065
A847	Slope<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _">=(Min)-1.23 and <-0.8"	11.28613196	3.483246233
A455	Y18<Health questionnaire> CESD-20 score (0-60 points) [-] _">=41 and <=50"	13.31170694	3.766904634
A93	Y19<Blood examination> White blood cell count [/µL] (for immunodiagnosis) _"Abnormal"	12.96220716	3.934885101
A126	Y18<Personal record form> Grip strength (Left hand) [kg] _">=13.8 and <25.6"	18.1170232	3.952968246
A204	Y18<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Abnormal"	11.1101406	3.970169329
A95	Y19<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Abnormal"	11.14537756	3.985373693
A915	Slope<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _">=1.58 and <3.75"	18.09133234	4.048145704

particular, the top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance to them are shown in Tables 8 to 11. In each table, the factors in the highlighted (light blue) cells in the upper table are characterized as clearly closer distances for "MMSE_slope_good", while the factors in the highlighted (light blue) cells in the lower table are characterized as clearly closer distances for "MMSE_slope_bad". Considering the interpretations of the PLSA results described in the previous section, we can see that the factors characterizing "MMSE_slope_good" and "MMSE_slope_bad" are not well differentiated in the relatively high health latent classes (i.e., $C^{(3)}$ for males over 60 and $C^{(1)}$ and $C^{(4)}$ for females over 60). Conversely, we can see that they are clearly differentiated in the relatively low health latent class (i.e., $C^{(5)}$ for males over 60). In particular, we can see that the related factor corresponding to "I like driving", which is included as a driving item taking the Kansei (sensitivity) into account, emerges as a high ranking factor that characterizes "MMSE_slope_good". This can be interpreted to mean that even if health declines are similar, those who like driving may have less MMSE decline than those who do not. As a natural reason, for those who like driving, we can consider that the high maintenance of cognitive function (MMSE) may be occurred because they enjoy driving, have a wider range of activities, and increase the possibility of multitasking through choosing to drive themselves rather than alternative transportation.

C. SIGNIFICANCE OF THE RESULTS OF THIS STUDY

The results of this study suggest that cognitive decline is more likely to be suppressed in the elderly who like to drive, and that they are more likely to continue social participation and improve their quality of life through driving, leading to self-actualization and community revitalization. In other words, we believe that we have suggested that it is important to design automobiles that allow the elderly to enjoy safe, secure, and comfortable transportation according to their health status, and to create opportunities for them to come to like driving through public health transportation policies that should be tailored to individuals to promote their health, taking into account the balance between the convenience of automobiles and their effects on health.

On the other hand, we cannot rule out the possibility that factors not included in this dataset, such as differences in diet, medical history, exercise habits, and residential area, as well as factors extracted from actual driving behavior rather than from questionnaires, may be confounding factors with large contributions in some of the latent classes. Therefore, we believe that it is important to integrate such factors into the analysis in order to further increase the confidence for the latent classes.

D. LIMITATIONS OF THIS STUDY

In this study, the number of latent classes in PLSA has determined based on preliminary analysis and interpretability.

TABLE 10. Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $c^{(1)}$ (for females over 60).

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _good"	**A flag variable representing subjects who drive every day and have a non-negative MMSE slope for 4 years (2016-2019)**	**0**	**1.555886127**
A811	**Slope<Cognitive function> MMSE score (0-30 points) [-] _">=0.3 and <=(Max)0.5"**	**0**	**1.555886127**
A964	**Slope<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _">=6.77 and <=(Max)12.03"**	**0.002054522**	**1.554043554**
A800	**Slope<Personal record form> Two-step test [cm] _">=49.6 and <79.8"**	**0.002408186**	**1.555952424**
A21	**Y19<Personal record form> Sit-to-stand test [cm] _">=28 and <34"**	**0.176678069**	**1.497160993**
A1054	**Y19Car<Driving questionnaire> Moments when you feel better while driving _"When feeling enough acceleration"**	**0.849012165**	**2.280174002**
A520	**Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=90 and <=(Max)100"**	**0.874361334**	**1.64754686**
A525	**Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=70 and <80"**	**0.879386672**	**2.340596681**
A1056	**Y19Car<Driving questionnaire> Moments when you feel better while driving _"When turning the curves as you want"**	**0.983399353**	**2.057841612**
A967	**Slope<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _">=1.56 and <4.78"**	**1.034057449**	**2.462910273**
A538	**Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=80 and <90"**	**1.43420523**	**2.659166801**
A913	Slope<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _">=(Min)-1000 and <-280"	1.51462263	1.160091869
A827	Slope<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _">=(Min)-3 and <-1.8"	1.553294195	0.00453053
A1019	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=(Min)-22.92 and <-14.58"	1.55380672	0.003515625
"MMSE_slope _bad"	A flag variable representing subjects who drive every day and have a negative MMSE slope for 4 years (2016-2019)	1.555886127	0
A809	Slope<Cognitive function> MMSE score (0-30 points) [-] _">=(Min)-0.5 and <-0.3"	1.555886127	0
A37	Y19<Cognitive function> MMSE score (0-30 points) [-] _"2"	1.584920305	0.582040204
A519	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=60 and <70"	1.771990853	1.891274221
A90	Y19<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Abnormal"	1.791599182	0.852642012
A62	Y19<Vision test> Near vision (Left eye) [-] _">=1"	1.966483482	1.185810154
A512	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=90 and <=(Max)100"	2.038321014	2.331301694
A57	Y19<Vision test> Distance vision (Both eyes) [-] _"<1"	2.039673288	3.209195085
A12	Y19<Personal record form> Sitting trunk flexion [cm] _">=45 and <57"	2.114947751	2.832482994
A909	Slope<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _">=-0.02 and <0.14"	2.118479017	3.372519295
A1077	Y19Car<Driving questionnaire> Excitements (Waku-Waku) that you can relate to _"Your own space"	2.188776224	3.554368902
A1062	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where feeling that you have your own space"	2.237215667	3.758850627
A1024	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=23.25 and <=(Max)37.5"	2.268961462	0.749778092
A1067	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where driving alone"	2.312480802	3.648722601
A831	Slope<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _">=-1.36 and <-0.52"	2.377793106	1.001773213
A908	Slope<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _">=-0.34 and <-0.18"	2.397563468	0.857489554

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _bad"	**A flag variable representing subjects who drive every day and have a negative MMSE slope for 4 years (2016-2019)**	**1.555886127**	**0**
A809	**Slope<Cognitive function> MMSE score (0-30 points) [-] _">=(Min)-0.5 and <-0.3"**	**1.555886127**	**0**
A1019	**Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=(Min)-22.92 and <-14.58"**	**1.55380672**	**0.003515625**
A827	**Slope<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _">=(Min)-3 and <-1.8"**	**1.553294195**	**0.00453053**
A37	**Y19<Cognitive function> MMSE score (0-30 points) [-] _"2"**	**1.584920305**	**0.582040204**
A1024	**Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=23.25 and <=(Max)37.5"**	**2.268961462**	**0.749778092**
A90	**Y19<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Abnormal"**	**1.791599182**	**0.852642012**

TABLE 10. ***(Continued.)*** **Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $C^{(1)}$ (for females over 60).**

A908	**Slope<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _">=-0.34 and <-0.18"**	**2.397563468**	**0.857489554**
A831	**Slope<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _">=-1.36 and <-0.52"**	**2.377793106**	**1.001773213**
A913	**Slope<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _">=(Min)-1000 and <-280"**	**1.51462263**	**1.160091869**
A62	**Y19<Vision test> Near vision (Left eye) [-] _">=1"**	**1.966483482**	**1.185810154**
A21	Y19<Personal record form> Sit-to-stand test [cm] _">=28 and <34"	0.176678069	1.497160993
A964	Slope<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _">=6.77 and <=(Max)12.03"	0.002054522	1.554043554
"MMSE_slope _good"	A flag variable representing subjects who drive every day and have a non-negative MMSE slope for 4 years (2016-2019)	0	1.555886127
A811	Slope<Cognitive function> MMSE score (0-30 points) [-] _">=0.3 and <=(Max)0.5"	0	1.555886127
A800	Slope<Personal record form> Two-step test [cm] _">=49.6 and <79.8"	0.002408186	1.555952424
A28	Y19<Personal record form> Two-step test [cm] _">=175.6 and <219.2"	2.766648538	1.640002706
A520	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=90 and <=(Max)100"	0.874361334	1.64754686
A56	Y19<Vision test> Distance vision (Left eye) [-] _">=1"	2.494446619	1.651724023
A530	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=60 and <70"	3.139347786	1.714627778
A519	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=60 and <70"	1.771990853	1.891274221
A64	Y19<Vision test> Near vision (Both eyes) [-] _">=1"	2.617458352	2.037977035
A1056	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When turning the curves as you want"	0.983399353	2.057841612
A1068	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where smelling the natural scent of the sea or greenery"	3.481947884	2.13075672
A617	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=90 and <=(Max)100"	3.097443906	2.186144504
A1054	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When feeling enough acceleration"	0.849012165	2.280174002
A512	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=90 and <=(Max)100"	2.038321014	2.331301694
A525	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=70 and <80"	0.879386672	2.340596681
A10	Y19<Personal record form> Sitting trunk flexion [cm] _">=33 and <45"	3.36841072	2.442084246
A967	Slope<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _">=1.56 and <4.78"	1.034057449	2.462910273

Therefore, it is difficult to completely remove arbitrariness, and there is a limit in universality that can be widely deployed unless a more objective method of determination is taken into account. In addition, even though it is possible in the case of document analysis to validate latent classes by separately setting up other documents that contain a limited number of words and common topics, because the target of analysis in this study is humans, the difficulty of completely covering attribute differences, such as unlimited individual and regional differences, makes it impossible to set up a separate population for validation, which limits the analysis to an understanding of the current situation. Furthermore, note that the above analysis by t-SNE is only a characterization of a small number of subjects. Therefore, it is a limitation of this study that it is impossible to say whether the above findings represent general features or not, while it is certainly possible to facilitate knowledge discovery without wasting a small number of subjects. In other words, in order to avoid simply ignoring subjects corresponding to singular values or outliers for which statistical significance is impossible to discuss, t-SNE can be a useful approach if we limit its use to actively extracting hypotheses specific to minority events, which are difficult to capture with classical statistics, by locally characterizing target variables as much as possible in terms of which factors appear in closest proximity.

As mentioned above, including in the introduction, both PLSA and t-SNE have unique advantages in terms of ease of understanding and characterizing the current state of the data, and it would be difficult to replicate them using other methods. On the other hand, as something that cannot be understood by PLSA in the first place, we cannot find a parent-child relationship to whether people drive because they are healthy, or whether they are healthy because they drive. Therefore, we believe that such findings may be found complementary by utilizing a Bayesian network, for example, together with a data set that can track changes over time for a longer period of time.

V. CONCLUSION

In this study, to facilitate knowledge discovery about the relationship between driving and health among the elderly, we have analyzed large-scale data obtained from a set of health and driving surveys using the machine learning methods PLSA and t-SNE for males and females over 60. The PLSA results show that there are broad categories of latent classes that can be interpreted as having a generally high or low level of health. In particular, a relatively positive

TABLE 11. Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $c^{(4)}$ (for females over 60).

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _good"	**A flag variable representing subjects who drive every day and have a non-negative MMSE slope for 4 years (2016-2019)**	**0**	**2.383599945**
A369	**Y16<Cognitive function> MMSE score (0-30 points) [-] _"2"**	**0.000491672**	**2.383618621**
A707	**Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=70 and <80"**	**0.713313291**	**2.976768398**
A810	**Slope<Cognitive function> MMSE score (0-30 points) [-] _">=0.1 and <0.3"**	**0.827016039**	**3.187840189**
A238	**Y17<Personal record form> Whole body reaction time [msec] _">=591.6 and <757.4"**	**1.02136143**	**3.379418987**
A251	**Y17<Personal record form> Two-step test [cm] _">=147.6 and <194.2"**	**1.021498656**	**3.379549357**
A537	**Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=60 and <70"**	**1.129453493**	**3.290428257**
A612	**Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=60 and <70"**	**1.366282826**	**1.714246774**
A329	**Y17<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Abnormal"**	**1.379916785**	**2.51596913**
A656	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=30 and <40"**	**1.38929497**	**3.713310976**
A490	**Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=40 and <50"**	**1.476945995**	**3.295720326**
A554	**Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=50 and <60"**	**1.54314431**	**3.449959595**
A628	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=50 and <60"**	**1.678276064**	**4.05575867**
A343	**Y16<Personal record form> Sitting trunk flexion [cm] _">=18.2 and <30.4"**	**1.935518676**	**4.249337735**
A822	Slope<Personal record form> Body weight [kg] _">=(Min)-3.54 and <-2.43"	2.217297351	1.58018456
A649	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=50 and <60"	2.223612311	0.913057186
"MMSE_slope _bad"	A flag variable representing subjects who drive every day and have a negative MMSE slope for 4 years (2016-2019)	2.383599945	0
A762	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=40 and <50"	2.603609088	4.50739156
A485	Y16<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=51 and <=60"	2.712558873	0.374980316
A267	Y17<Personal record form> Body weight [kg] _">=76.94 and <96.96"	2.912582379	1.456503012
A476	Y17<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=41 and <=50"	2.945751702	1.531548821
A456	Y19<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=41 and <=50"	2.956717716	3.416040874
A334	Y17<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Normal) [-] _"Large"	2.95679325	3.415709318
A726	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=30 and <40"	3.005313147	5.191777197
A378	Y16<Personal record form> Body weight [kg] _">=76.68 and <90.34"	3.009848745	1.605279317
A1070	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"None"	3.020868491	3.034915438
A667	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=30 and <40"	3.035697029	5.294399846
A812	Slope<Cognitive function> MMSE score (0-30 points) [-] _">=-0.3 and <-0.1"	3.170313604	1.284012693
A635	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=10 and <20"	3.319182942	1.889179751
A475	Y17<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=61 and <=70"	3.389804679	4.237102647

Binary variables	Description of variables	"MMSE_slope_good"	"MMSE_slope_bad"
"MMSE_slope _bad"	**A flag variable representing subjects who drive every day and have a negative MMSE slope for 4 years (2016-2019)**	**2.383599945**	**0**
A485	**Y16<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=51 and <=60"**	**2.712558873**	**0.374980316**
A649	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=50 and <60"**	**2.223612311**	**0.913057186**
A812	**Slope<Cognitive function> MMSE score (0-30 points) [-] _">=-0.3 and <-0.1"**	**3.170313604**	**1.284012693**
A267	**Y17<Personal record form> Body weight [kg] _">=76.94 and <96.96"**	**2.912582379**	**1.456503012**
A677	**Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=30 and <40"**	**3.451335076**	**1.496480083**
A476	**Y17<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=41 and <=50"**	**2.945751702**	**1.531548821**
A822	**Slope<Personal record form> Body weight [kg] _">=(Min)-3.54 and <-2.43"**	**2.217297351**	**1.58018456**
A378	**Y16<Personal record form> Body weight [kg] _">=76.68 and <90.34"**	**3.009848745**	**1.605279317**
A612	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=60 and <70"	1.366282826	1.714246774
A363	Y16<Personal record form> Two-step test [cm] _">=137.4 and <189.8"	3.390589769	1.751830673
A35	Y19<Personal record form> Maximal walking speed of 10 meters [sec] _">=8.94 and <=(Max)10.6"	3.405445748	1.755575877
A3	Y19<Personal record form> Age [years old] _"80-89"	3.400917132	1.759818369
A313	Y17<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Abnormal"	4.2345141	1.852961016
A635	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=10 and <20"	3.319182942	1.889179751
A682	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=80 and <90"	3.609163273	1.996891943
A260	Y17<Cognitive function> MMSE score (0-30 points) [-] _"2"	4.73154807	2.373057341

TABLE 11. ***(Continued.)*** **Top 30 factors in ascending order of the 2-dimensional compressed Euclidean distance for each of "MMSE_slope_good" (see upper table) and "MMSE_slope_bad" (see lower table), corresponding to the latent class $c^{(4)}$ (for females over 60).**

"MMSE_slope _good"	A flag variable representing subjects who drive every day and have a non-negative MMSE slope for 4 years (2016-2019)	0	2.383599945
A369	Y16<Cognitive function> MMSE score (0-30 points) [-] _"2"	0.000491672	2.383618621
A329	Y17<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Abnormal"	1.379916785	2.51596913
A198	Y18<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Abnormal"	3.829541164	2.912351789
A299	Y17<Blood examination> Creatinine [mg/dL] (for kidney diagnosis) _"Abnormal"	4.952151346	2.947806606
A707	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=70 and <80"	0.713313291	2.976768398
A1070	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"None"	3.020868491	3.034915438
A963	Slope<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _">=1.51 and <6.77"	5.43481484	3.081396834
A810	Slope<Cognitive function> MMSE score (0-30 points) [-] _">=0.1 and <0.3"	0.827016039	3.187840189
A537	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=60 and <70"	1.129453493	3.290428257
A708	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=80 and <90"	5.442325467	3.295596216
A490	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=40 and <50"	1.476945995	3.295720326
A931	Slope<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _">=-0.19 and <0.1"	4.208461526	3.356408379

relationship is found between driving and health. In addition, the t-SNE results show that the factors characterizing subjects who drive every day but maintain or decrease cognitive function (MMSE) are more clearly differentiated in the relatively low health class. In particular, "I like driving" is found to be a notable related factor characterizing the high maintenance of cognitive function. This finding has been effectively obtained because of a detailed analysis while handling items that take the Kansei (sensitivity) into account. Our future works include examining fewer qualitative decision rules for the number of latent classes, clarifying more detailed dependencies between driving and health based on changes over time, and increasing confidence in our findings on the overall features and specific characteristics between driving and health by measuring and integrating data on actual driving behavior, not just on questionnaires. In order to strongly promote healthy longevity in the future, we believe that it is an important theme to provide cars that stimulate the Kansei (sensitivity) and make driving enjoyable, while facilitating knowledge discovery through integrated data analysis that incorporates further relevant factors and longer-term data.

APPENDIX

A. OVERVIEW OF PLSA

In this appendix, we give an overview of PLSA [17]. Let the data

$$\boldsymbol{A} = \begin{bmatrix} n(d_1, w_1) & \cdots & n(d_1, w_M) \\ \vdots & \ddots & \vdots \\ n(d_N, w_1) & \cdots & n(d_N, w_M) \end{bmatrix} \tag{11}$$

consist of an $N \times M$ matrix based on the co-occurrence frequency $n(d, w) \in \{0, 1, 2, \ldots\}$ for pairs of documents (corresponding to samples representing subjects in this study) $d \in D = \{d_1, \ldots, d_N\}$ and words (corresponding to factors representing survey items in this study) $w \in W = \{w_1, \ldots, w_M\}$, associated with topics (latent variables) $z \in Z = \{z_1, \ldots, z_K\}$. Denote the probability that each document d is chosen as $P(d)$, the conditional probability that each topic z is chosen given document d as $P(z|d)$, and the conditional probability that each word w is chosen given topic z as $P(w|z)$. Then, the simultaneous probability $P(d, w)$ of each document d and word w is denoted by

$$P(d, w) = P(d)\, P(w|d) = P(d) \sum_{z \in Z} P(w|z)\, P(z|d). \tag{12}$$

This can be rewritten as

$$P(d, w) = \sum_{z \in Z} P(z)\, P(d|z)\, P(w|z). \tag{13}$$

Using this, the parameters $P(d|z)$, $P(w|z)$ and $P(z)$ are obtained by maximizing the log-likelihood function

$$L = \sum_{d \in D, w \in W} n(d, w) \log P(d, w) \tag{14}$$

according to the following EM algorithm.

1) INITIALIZATION ($t = 0$)

The initial values of $P(d|z)$, $P(w|z)$ and $P(z)$ are generated by uniform random numbers.

2) E-STEP ($t \geq 0$)

$$P(z|d, w) = \frac{P(z)\, P(d|z)\, P(w|z)}{\sum_{z' \in Z} P(z')\, P(d|z')\, P(w|z')}. \tag{15}$$

3) M-STEP ($t \geq 0$)

$$P(d|z) = \frac{\sum_{w \in W} n(d, w)\, P(z|d, w)}{\sum_{d' \in D, w \in W} n(d', w)\, P(z|d', w)},$$
$$P(w|z) = \frac{\sum_{d \in D} n(d, w)\, P(z|d, w)}{\sum_{d \in D, w' \in W} n(d, w')\, P(z|d, w')},$$
$$P(z) = \frac{\sum_{d \in D, w \in W} n(d, w)\, P(z|d, w)}{\sum_{d \in D, w \in W} n(d, w)}. \tag{16}$$

TABLE 12. List of factors $w \in C_w^{(1)}$ corresponding to (1) and (2) in Fig. 4.

	Binary variables w	Description of variables	$P(w\|z_1)$
(1) Top 30 factors based on descending order of $P(w\|z_1)$	A107	Y19<Blood pressure measurement of the extremities> Rabi [-] (for artery occlusion diagnosis) _"Normal"	0.00508876
	A108	Y19<Blood pressure measurement of the extremities> Labi [-] (for artery occlusion diagnosis) _"Normal"	0.00508876
	A90	Y19<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _"Normal"	0.0050486
	A76	Y19<Blood examination> ALT (GPT) [U/L] (for liver diagnosis) _"Normal"	0.005000545
	A94	Y19<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Normal"	0.004956808
	A105	Y19<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.004932231
	A74	Y19<Blood examination> AST (GOT) [U/L] (for liver diagnosis) _"Normal"	0.004895892
	A72	Y19<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Normal"	0.004856915
	A985	Y19Car<Driving questionnaire> Driving frequency _"Every day"	0.004767429
	A41	Y19<Cognitive function> MMSE score (0-30 points) [-] _"3"	0.004686438
	A92	Y19<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _"Normal"	0.004680756
	A103	Y19<Blood pressure measurement of the extremities> RbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.004610529
	A56	Y19<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _"1-5"	0.004556443
	A52	Y19<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _"1-5"	0.004538345
	A116	Y19<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Common) [-] _"Common"	0.004472047
	A96	Y19<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Normal"	0.004426731
	A1	Y19<Personal record form> Age [years old] _"60-69"	0.004408062
	A529	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=90 and <=(Max)100"	0.004386401
	A82	Y19<Blood examination> Creatinine [mg/dL] (for kidney diagnosis) _"Normal"	0.004353822
	A88	Y19<Blood examination> HDL cholesterol [mg/dL] (for lipid diagnosis) _"Normal"	0.00434652
	A111	Y19<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Normal"	0.004339511
	A68	Y19<Vision test> Near vision (Left eye) [-] _"<1"	0.00427879
	A80	Y19<Blood examination> Total protein [g/dL] (for nutrition diagnosis) _"Normal"	0.00415919
	A86	Y19<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Normal"	0.004147996
	A118	Y19<Health questionnaire> PSQI total score (0-21 points) [-] _"0-5"	0.004065784
	A524	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=90 and <=(Max)100"	0.003929904
	A78	Y19<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _"Normal"	0.003911227
	A67	Y19<Vision test> Near vision (Right eye) [-] _"<1"	0.003891358
	A487	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=90 and <=(Max)100"	0.003877966
	A84	Y19<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _"Normal"	0.003816855
(2) At most 10 factors related to driving only	A985	Y19Car<Driving questionnaire> Driving frequency _"Every day"	0.004767429
	A1011	Y19Car<Driving questionnaire> Moments when you feel better while driving _"None"	0.003404221
	A1019	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where seeing the natural scenery of the sea or greenery"	0.003327118
	A1023	Y19Car<Driving questionnaire> Excitements (Waku-Waku) that you can relate to _"The harmony with nature"	0.002823064
	A998	Y19Car<Driving questionnaire> Likes or dislikes about driving _"Rather like"	0.002679084
	A1016	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where spending time in the car with your family"	0.002431759
	A994	Y19Car<Driving questionnaire> Transmission (AT or MT) _"Only AT"	0.002362463
	A990	Y19Car<Driving questionnaire> Driving time _"Less than 30 minutes"	0.002144223
	A1015	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where traveling or driving with everyone"	0.001969074
	A997	Y19Car<Driving questionnaire> Likes or dislikes about driving _"Like"	0.001867135

TABLE 13. List of factors $w \in C_w^{(2)}$ corresponding to (1) and (2) in Fig. 5.

	Binary variables w	Description of variables	$P(w\|z_2)$
(1) Top 30 factors based on descending order of $P(w\|z_2)$	A651	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=90 and <=(Max)100"	0.002194803
	A262	Y17<Cognitive function> MMSE score (0-30 points) [-] _"3"	0.002182164
	A465	Y17<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=91 and <=(Max)100"	0.002134845
	A970	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=-5 and <5"	0.002130881
	A942	Slope<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=-3 and <2.5"	0.00209072
	A646	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=90 and <=(Max)100"	0.002003453
	A780	Slope<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _">=-2 and <1"	0.002003425
	A588	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=90 and <=(Max)100"	0.001963945
	A966	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=-6.25 and <6.25"	0.001963272
	A614	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=90 and <=(Max)100"	0.001938916
	A474	Y16<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=91 and <=(Max)100"	0.001887722
	A462	Y17<Health questionnaire> CESD-20 score (0-60 points) [-] _">=51 and <=(Max)60"	0.00182089
	A980	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=-7.4 and <3.2"	0.001743838
	A776	Slope<Personal record form> Body weight [kg] _">=-1.34 and <0.34"	0.001729206
	A741	Slope<Personal record form> Whole body reaction time [msec] _">=-33.9 and <55.7"	0.001703292
	A367	Y16<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)2.18 and <4.67"	0.001683081
	A869	Slope<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _">=-0.55 and <0.1"	0.001663685
	A918	Slope<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _">=-0.8 and <3.8"	0.001656987
	A607	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=90 and <=(Max)100"	0.001580127
	A737	Slope<Personal record form> Sitting trunk flexion [cm] _">=-2.2 and <3.2"	0.001528935
	A855	Slope<Blood examination> HDL cholesterol [mg/dL] (for lipid diagnosis) _">=-0.76 and <5.16"	0.001503276
	A282	Y17<Vision test> Distance vision (Both eyes) [-] _">=1"	0.001480564
	A864	Slope<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _">=-60 and <593.33"	0.001387137
	A18	Y19<Personal record form> Whole body reaction time [msec] _">=228.4 and <456.8"	0.001251933
	A844	Slope<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _">=0.05 and <0.47"	0.001234297
	A658	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=90 and <=(Max)100"	0.001210293
	A1030	Y19Car<Driving questionnaire> Purpose of driving _"Shopping"	0.001206429
	A375	Y16<Personal record form> Body weight [kg] _">=63.02 and <76.68"	0.001186442
	A752	Slope<Personal record form> TUG test [sec] _">=-0.19 and <0.42"	0.001087022
	A135	Y18<Personal record form> Whole body reaction time [msec] _">=(Min)272 and <402.6"	0.001034146
(2) At most 10 factors related to driving only	A1030	Y19Car<Driving questionnaire> Purpose of driving _"Shopping"	0.001206429
	A992	Y19Car<Driving questionnaire> Driving time _"1-2 hours"	0.000789832
	A1010	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When achieving fuel-efficient driving"	0.000752151
	A1007	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When seeing the shining body of your car"	0.000683626
	A1014	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where chatting and laughing with everyone"	0.000471942
	A1005	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When driving around a series of curves"	0.000456598
	A1008	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When feeling the texture of the interior"	0.000394631
	A1006	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When turning the curves as you want"	0.000357239
	A1004	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When feeling enough acceleration"	0.000333964
	A1026	Y19Car<Driving questionnaire> Excitements (Waku-Waku) that you can relate to _"Achieving fuel efficiency"	0.000332321

TABLE 14. List of factors $w \in C_w^{(3)}$ corresponding to (1) and (2) in Fig. 6.

	Binary variables w	Description of variables	$P(w\|z_3)$
(1) Top 30 factors based on descending order of $P(w\|z_3)$	A417	Y16<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _"Normal"	0.002564693
	A432	Y16<Blood pressure measurement of the extremities> Rabi [-] (for artery occlusion diagnosis) _"Normal"	0.002564693
	A442	Y16<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Common) [-] _"Common"	0.002515375
	A433	Y16<Blood pressure measurement of the extremities> Labi [-] (for artery occlusion diagnosis) _"Normal"	0.00251292
	A426	Y16<Blood examination> Cortisol [μg/dL] (for stress diagnosis) _"Normal"	0.002476155
	A407	Y16<Blood examination> Total protein [g/dL] (for nutrition diagnosis) _"Normal"	0.002395529
	A418	Y16<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _"Normal"	0.002371947
	A401	Y16<Blood examination> AST (GOT) [U/L] (for liver diagnosis) _"Normal"	0.002331188
	A405	Y16<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _"Normal"	0.002321063
	A424	Y16<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _"Normal"	0.002315398
	A413	Y16<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Normal"	0.002311487
	A415	Y16<Blood examination> HDL cholesterol [mg/dL] (for lipid diagnosis) _"Normal"	0.00231138
	A403	Y16<Blood examination> ALT (GPT) [U/L] (for liver diagnosis) _"Normal"	0.002275733
	A395	Y16<Vision test> Near vision (Left eye) [-] _"<1"	0.002265468
	A369	Y16<Cognitive function> MMSE score (0-30 points) [-] _"3"	0.002264882
	A379	Y16<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _"1-5"	0.002230707
	A399	Y16<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Normal"	0.002174749
	A411	Y16<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _"Normal"	0.002167497
	A393	Y16<Vision test> Near vision (Right eye) [-] _"<1"	0.002148683
	A784	Slope<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _">=-1.4 and <=(Max)2"	0.002091152
	A397	Y16<Vision test> Near vision (Both eyes) [-] _"<1"	0.002082404
	A409	Y16<Blood examination> Creatinine [mg/dL] (for kidney diagnosis) _"Normal"	0.002068646
	A428	Y16<Blood pressure measurement of the extremities> RbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.002030527
	A422	Y16<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Normal"	0.00202786
	A383	Y16<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _"1-5"	0.001997905
	A430	Y16<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.001960724
	A893	Slope<Blood pressure measurement of the extremities> RbaPWV [cm/s] (for atherosclerosis diagnosis) _">=-84 and <59"	0.001946219
	A708	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=90 and <=(Max)100"	0.001925283
	A444	Y16<Health questionnaire> PSQI total score (0-21 points) [-] _"0-5"	0.001896383
	A672	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=90 and <=(Max)100"	0.001739822
(2) At most 10 factors related to driving only	A1034	Y19Car<Driving questionnaire> Purpose of driving _"Going for the leisure"	0.000083521

4) TERMINATION CONDITION ($t \geq 1$)

Compute the log-likelihood function $L = L[t]$ and the error rate $E = \left|(L[t] - L[t-1]) / L[t]\right| (t \geq 1)$ for the iteration number $t = 0, 1, 2, \ldots$ of E-step and M-step. Under a sufficiently small threshold $\varepsilon > 0$, if $E < \varepsilon$ then terminate, otherwise repeat E-step and M-step.

B. OVERVIEW OF T-SNE

In this appendix, we give an overview of t-SNE [18]. Suppose we have the data

$$\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m] \tag{17}$$

consisting of an $n \times m$ matrix with mn-dimensional (high-dimensional) vectors $\boldsymbol{x}_i = [x_{i1}, \ldots, x_{in}]^T$ $(i = 1, \ldots, m)$.

TABLE 15. List of factors $w \in C_w^{(4)}$ corresponding to (1) and (2) in Fig. 7.

	Binary variables w	Description of variables	$P(w\|z_4)$
(1) Top 30 factors based on descending order of $P(w\|z_4)$	A437	Y16<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Normal"	0.002374536
	A420	Y16<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Normal"	0.002193019
	A765	Slope<Cognitive function> MMSE score (0-30 points) [-] _">=-0.1 and <0.1"	0.002164032
	A897	Slope<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _">=-64.2 and <79.4"	0.001956053
	A714	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=90 and <=(Max)100"	0.001808116
	A937	Slope<Health questionnaire> CESD-20 score (0-60 points) [-] _">=-2.4 and <1.2"	0.001719924
	A889	Slope<Blood examination> Cortisol [µg/dL] (for stress diagnosis) _">=-1.01 and <0.71"	0.001700305
	A435	Y16<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _"Abnormal"	0.001534111
	A932	Slope<Health questionnaire> PSQI total score (0-21 points) [-] _">=-1.4 and <0.4"	0.001496355
	A391	Y16<Vision test> Distance vision (Both eyes) [-] _">=1"	0.001423048
	A788	Slope<Vision test> Distance vision (Right eye) [-] _">=0 and <0.2"	0.001346757
	A363	Y16<Personal record form> Two-step test [cm] _">=242.2 and <294.6"	0.001319203
	A885	Slope<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _">=-0.05 and <0.16"	0.001309264
	A388	Y16<Vision test> Distance vision (Right eye) [-] _">=1"	0.001229863
	A913	Slope<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _">=-2.77 and <-0.6"	0.00116274
	A359	Y16<Personal record form> TUG test [sec] _">=(Min)3.1 and <5.24"	0.001095794
	A352	Y16<Personal record form> Whole body reaction time [msec] _">=(Min)280 and <409.6"	0.001066831
	A305	Y17<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Abnormal"	0.001051635
	A326	Y17<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _"Normal"	0.000988072
	A440	Y16<Body composition measurement> Body fat percentage [%] _"Standard"	0.000968698
	A623	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=70 and <80"	0.000942889
	A342	Y16<Personal record form> Grip strength (Left hand) [kg] _">=35.8 and <47.4"	0.000918278
	A278	Y17<Vision test> Distance vision (Right eye) [-] _">=1"	0.000903884
	A264	Y17<Personal record form> Body height [cm] _">=166.32 and <175.96"	0.000892474
	A331	Y17<Body composition measurement> Body fat percentage [%] _"Standard"	0.000875073
	A719	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=90 and <=(Max)100"	0.000828203
	A638	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : VT _">=80 and <90"	0.000769385
	A390	Y16<Vision test> Distance vision (Left eye) [-] _">=1"	0.000763573
	A414	Y16<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Abnormal"	0.000759613
	A244	Y17<Personal record form> Whole body reaction time [msec] _">=(Min)260 and <425.8"	0.000759485
(2) At most 10 factors related to driving only	A987	Y19Car<Driving questionnaire> Driving frequency _"3-4 days a week"	0.000395085

Consider embedding each $\boldsymbol{x}_i$ as a dimension-compressed vector $\boldsymbol{y}_i = [y_{i1}, \ldots, y_{in'}]^T$ in an n'-dimensional (low-dimensional) space through a transformation into the conditional probabilities representing certain similarities based on the Euclidean distances among $\boldsymbol{x}_i$ $(i = 1, \ldots, m)$.

Using a Gaussian distribution, in the high-dimensional space, define the conditional probability $p_{j|i}$ and the simultaneous probability p_{ij} by

$$p_{j|i} = \frac{\exp\left(-\left\|\boldsymbol{x}_i - \boldsymbol{x}_k\right\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq i} \exp\left(-\left\|\boldsymbol{x}_i - \boldsymbol{x}_k\right\|^2 / 2\sigma_i^2\right)},$$

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}, \tag{18}$$

where $p_{i|i} = 0$ and $p_{ii} = 0$. The value of each σ_i, which corresponds to the standard deviation of the Gaussian distribution, is obtained through a binary search to fit the following predefined perplexity:

$$perplexity = Perp\left(P_i\right) = 2^{H(P_i)}, \tag{19}$$

where

$$H\left(P_i\right) = -\sum_j p_{j|i} \log_2 p_{j|i} \tag{20}$$

TABLE 16. List of factors $w \in C_w^{(5)}$ corresponding to (1) and (2) in Fig. 8.

	Binary variables w	Description of variables	$P(w \mid z_5)$
(1) Top 30 factors based on descending order of $P(w \mid z_5)$	A322	Y17<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.003525208
	A320	Y17<Blood pressure measurement of the extremities> RbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.003483934
	A310	Y17<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _"Normal"	0.003409987
	A324	Y17<Blood pressure measurement of the extremities> Rabi [-] (for artery occlusion diagnosis) _"Normal"	0.003367009
	A325	Y17<Blood pressure measurement of the extremities> Labi [-] (for artery occlusion diagnosis) _"Normal"	0.003367009
	A318	Y17<Blood examination> Cortisol [μg/dL] (for stress diagnosis) _"Normal"	0.003304661
	A308	Y17<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _"Normal"	0.003286067
	A333	Y17<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Common) [-] _"Common"	0.003277375
	A316	Y17<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _"Normal"	0.003192598
	A327	Y17<Pulmonary function testing (Spirometry)> FEV1.0%G [%] (for obstructive pulmonary disease diagnosis) _"Abnormal"	0.003190056
	A290	Y17<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Normal"	0.003188289
	A298	Y17<Blood examination> Total protein [g/dL] (for nutrition diagnosis) _"Normal"	0.003091832
	A330	Y17<Body composition measurement> Body fat percentage [%] _"High"	0.00305949
	A328	Y17<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Normal"	0.003057675
	A306	Y17<Blood examination> HDL cholesterol [mg/dL] (for lipid diagnosis) _"Normal"	0.00303726
	A927	Slope<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Common) [-] _">=-0.03 and <0.23"	0.002968519
	A285	Y17<Vision test> Near vision (Right eye) [-] _"<1"	0.002878419
	A312	Y17<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Normal"	0.002781233
	A292	Y17<Blood examination> AST (GOT) [U/L] (for liver diagnosis) _"Normal"	0.002728347
	A304	Y17<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Normal"	0.002725633
	A173	Y18<Vision test> Distance vision (Both eyes) [-] _"<1"	0.002723274
	A271	Y17<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _"1-5"	0.002693591
	A296	Y17<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _"Normal"	0.002692158
	A286	Y17<Vision test> Near vision (Left eye) [-] _"<1"	0.002621846
	A294	Y17<Blood examination> ALT (GPT) [U/L] (for liver diagnosis) _"Normal"	0.002615505
	A314	Y17<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Normal"	0.002569527
	A2	Y19<Personal record form> Age [years old] _"70-79"	0.00254386
	A280	Y17<Vision test> Distance vision (Left eye) [-] _"<1"	0.002530479
	A279	Y17<Vision test> Distance vision (Right eye) [-] _"<1"	0.002460001
	A162	Y18<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _"1-5"	0.002382582
(2) At most 10 factors related to driving only	A996	Y19Car<Driving questionnaire> Transmission (AT or MT) _"Both AT and MT"	0.001007407
	A1031	Y19Car<Driving questionnaire> Purpose of driving _"Working"	0.000875267
	A993	Y19Car<Driving questionnaire> Driving time _"More than 2 hours"	0.000536143
	A1009	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When feeling that your car's design stands out from others"	0.000483354
	A1020	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"None"	0.000447385
	A1002	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When putting your foot on the accelerator"	0.00035449
	A1035	Y19Car<Driving questionnaire> Purpose of driving _"Going to the hospital"	0.000341443
	A1001	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When hearing the sound of the engine running"	0.00033523
	A1000	Y19Car<Driving questionnaire> Likes or dislikes about driving _"Dislike"	0.000093913

TABLE 17. List of factors $w \in C_w^{(1)}$ corresponding to (1) and (2) in Fig. 9.

	Binary variables w	Description of variables	$P(w\mid z_1)$
(1) Top 30 factors based on descending order of $P(w\mid z_1)$	A83	Y19<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _"Normal"	0.007213253
	A81	Y19<Blood examination> HDL cholesterol [mg/dL] (for lipid diagnosis) _"Normal"	0.007159492
	A77	Y19<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _"Normal"	0.007146575
	A102	Y19<Blood pressure measurement of the extremities> Labi [-] (for artery occlusion diagnosis) _"Normal"	0.007046807
	A100	Y19<Blood pressure measurement of the extremities> Rabi [-] (for artery occlusion diagnosis) _"Normal"	0.007046807
	A87	Y19<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Normal"	0.006917934
	A67	Y19<Blood examination> AST (GOT) [U/L] (for liver diagnosis) _"Normal"	0.006899577
	A69	Y19<Blood examination> ALT (GPT) [U/L] (for liver diagnosis) _"Normal"	0.006721101
	A85	Y19<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _"Normal"	0.006700795
	A66	Y19<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Normal"	0.006540452
	A542	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=90 and <=(Max)100"	0.006499329
	A73	Y19<Blood examination> Total protein [g/dL] (for nutrition diagnosis) _"Normal"	0.00647646
	A111	Y19<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Common) [-] _"Common"	0.006413272
	A79	Y19<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Normal"	0.006399523
	A50	Y19<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _"1-5"	0.006296187
	A36	Y19<Cognitive function> MMSE score (0-30 points) [-] _"3"	0.006251205
	A46	Y19<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _"1-5"	0.006239237
	A75	Y19<Blood examination> Creatinine [mg/dL] (for kidney diagnosis) _"Normal"	0.006112807
	A61	Y19<Vision test> Near vision (Left eye) [-] _"<1"	0.00609334
	A59	Y19<Vision test> Near vision (Right eye) [-] _"<1"	0.006010095
	A91	Y19<Blood examination> High sensitive CRP [mg/dL] (for inflammation diagnosis) _">=(Min)0 and <0.1"	0.005825363
	A1035	Y19Car<Driving questionnaire> Driving frequency _"Every day"	0.005812322
	A89	Y19<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Normal"	0.005759896
	A97	Y19<Blood pressure measurement of the extremities> RbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.005737529
	A99	Y19<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.005730138
	A106	Y19<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Normal"	0.005697965
	A71	Y19<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _"Normal"	0.005619868
	A1044	Y19Car<Driving questionnaire> Transmission (AT or MT) _"Only AT"	0.005354118
	A534	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=90 and <=(Max)100"	0.00511006
	A1069	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where seeing the natural scenery of the sea or greenery"	0.005101324
(2) At most 10 factors related to driving only	A1035	Y19Car<Driving questionnaire> Driving frequency _"Every day"	0.005812322
	A1044	Y19Car<Driving questionnaire> Transmission (AT or MT) _"Only AT"	0.005354118
	A1069	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where seeing the natural scenery of the sea or greenery"	0.005101324
	A1061	Y19Car<Driving questionnaire> Moments when you feel better while driving _"None"	0.004584596
	A1080	Y19Car<Driving questionnaire> Purpose of driving _"Shopping"	0.004569997
	A1073	Y19Car<Driving questionnaire> Excitements (Waku-Waku) that you can relate to _"The harmony with nature"	0.004100509
	A1065	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where traveling or driving with everyone"	0.003705238
	A1041	Y19Car<Driving questionnaire> Driving time _"30 minutes to 1 hour"	0.003437928
	A1066	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"Where spending time in the car with your family"	0.003056245
	A1048	Y19Car<Driving questionnaire> Likes or dislikes about driving _"Rather like"	0.003040914

TABLE 18. List of factors $w \in C_w^{(2)}$ corresponding to (1) and (2) in Fig. 10.

	Binary variables w	Description of variables	$P(w\|z_2)$
(1) Top 30 factors based on descending order of $P(w\|z_2)$	A471	Y17<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=91 and <=(Max)100"	0.002912184
	A679	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE ">=90 and <=(Max)100"	0.00280419
	A672	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=90 and <=(Max)100"	0.002612774
	A468	Y17<Health questionnaire> CESD-20 score (0-60 points) [-] _">=51 and <=(Max)60"	0.002559955
	A626	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=90 and <=(Max)100"	0.002557728
	A636	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP ">=90 and <=(Max)100"	0.002528522
	A794	Slope<Personal record form> TUG test [sec] _">=-0.52 and <0.68"	0.002520588
	A461	Y18<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=91 and <=(Max)100"	0.002494173
	A227	Y17<Personal record form> Grip strength (Left hand) [kg] _">=20.8 and <33.6"	0.002443814
	A609	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=90 and <=(Max)100"	0.00227158
	A482	Y16<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=91 and <=(Max)100"	0.002269042
	A567	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=90 and <=(Max)100"	0.002204886
	A365	Y16<Personal record form> Maximal walking speed of 10 meters [sec] _">=(Min)2.18 and <4.67"	0.002195856
	A224	Y17<Personal record form> Grip strength (Right hand) [kg] _">=21 and <32"	0.002183084
	A559	Y18<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=90 and <=(Max)100"	0.002169873
	A457	Y18<Health questionnaire> CESD-20 score (0-60 points) [-] _">=51 and <=(Max)60"	0.002092263
	A798	Slope<Personal record form> Two-step test [cm] _">=-10.8 and <19.4"	0.002068646
	A335	Y17<Health questionnaire> PSQI total score (0-21 points) [-] _"0-5"	0.002066626
	A1025	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=-5 and <5"	0.00200243
	A705	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RP _">=90 and <=(Max)100"	0.001951017
	A695	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=90 and <=(Max)100"	0.001911573
	A749	Y16<Health questionnaire> SF-36v2 score (0-100 points) [-] : RE _">=90 and <=(Max)100"	0.001900498
	A280	Y17<Vision test> Distance vision (Both eyes) [-] _">=1"	0.00187551
	A244	Y17<Personal record form> TUG test [sec] _">=4.5 and <5.95"	0.001861292
	A241	Y17<Personal record form> Sit-to-stand test [cm] _">=34 and <=(Max)40"	0.001837808
	A769	Slope<Personal record form> Grip strength (Right hand) [kg] _">=-1.26 and <0.76"	0.001815562
	A1031	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : MH _">=-4 and <4"	0.001749146
	A145	Y18<Personal record form> Maximal walking speed of 10 meters [sec] _">=3.22 and <4.39"	0.001738834
	A789	Slope<Personal record form> Sit-to-stand test [cm] _">=-6 and <6"	0.001705076
	A478	Y16<Health questionnaire> CESD-20 score (0-60 points) [-] _">=51 and <=(Max)60"	0.001690965
(2) At most 10 factors related to driving only	A1071	Y19Car<Driving questionnaire> Excitements (Waku-Waku) that you can relate to _"The sense of speed"	0.000360357
	A1076	Y19Car<Driving questionnaire> Excitements (Waku-Waku) that you can relate to _"Achieving fuel efficiency"	0.00016775
	A1038	Y19Car<Driving questionnaire> Driving frequency _"Less than 2 days a week"	0.000160173
	A1059	Y19Car<Driving questionnaire> Moments when you feel better while driving _"When feeling that your car's design stands out from others"	0.000053489

is the Shannon entropy. In addition, using a Student t-distribution with one degree of freedom, in the low-dimensional space, define the simultaneous probability q_{ij} by

$$q_{ij} = \frac{\left(1 + \left\|\boldsymbol{y}_i - \boldsymbol{y}_j\right\|^2\right)^{-1}}{\sum_{k \neq l}\left(1 + \left\|\boldsymbol{y}_k - \boldsymbol{y}_l\right\|^2\right)^{-1}}, \tag{21}$$

where $q_{ii} = 0$. Using the Kullback–Leibler divergence, define the cost function C by

$$C = KL\left(P \| Q\right) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}. \tag{22}$$

TABLE 19. List of factors $w \in C_w^{(3)}$ corresponding to (1) and (2) in Fig. 11.

	Binary variables w	Description of variables	$P(w\|z_3)$
(1) Top 30 factors based on descending order of $P(w\|z_3)$	A189	Y18<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Normal"	0.003226399
	A179	Y18<Blood examination> ALT (GPT) [U/L] (for liver diagnosis) _"Normal"	0.00320495
	A193	Y18<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _"Normal"	0.003177986
	A187	Y18<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _"Normal"	0.003158461
	A214	Y18<Blood pressure measurement of the extremities> Rabi [-] (for artery occlusion diagnosis) _"Normal"	0.00314138
	A215	Y18<Blood pressure measurement of the extremities> Labi [-] (for artery occlusion diagnosis) _"Normal"	0.003125825
	A177	Y18<Blood examination> AST (GOT) [U/L] (for liver diagnosis) _"Normal"	0.003101233
	A197	Y18<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Normal"	0.00306324
	A176	Y18<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Normal"	0.002999881
	A195	Y18<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _"Normal"	0.002991895
	A169	Y18<Vision test> Near vision (Right eye) [-] _"<1"	0.002982328
	A206	Y18<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _"Normal"	0.002981597
	A208	Y18<Blood examination> Cortisol [μg/dL] (for stress diagnosis) _"Normal"	0.002981013
	A156	Y18<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _"1-5"	0.00296381
	A201	Y18<Blood examination> High sensitive CRP [mg/dL] (for inflammation diagnosis) _">=(Min)0 and <0.1"	0.002910405
	A165	Y18<Vision test> Distance vision (Left eye) [-] _"<1"	0.002899539
	A160	Y18<Health questionnaire> Social connections (Number of friends who feel comfortable) [-] _"1-5"	0.002884851
	A163	Y18<Vision test> Distance vision (Right eye) [-] _"<1"	0.002870487
	A153	Y18<Personal record form> Body weight [kg] _">=(Min)37.5 and <56.74"	0.002844952
	A191	Y18<Blood examination> HDL cholesterol [mg/dL] (for lipid diagnosis) _"Normal"	0.00283316
	A171	Y18<Vision test> Near vision (Left eye) [-] _"<1"	0.002791267
	A199	Y18<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Normal"	0.002717207
	A186	Y18<Blood examination> Creatinine [mg/dL] (for kidney diagnosis) _"Normal"	0.002660787
	A173	Y18<Vision test> Near vision (Both eyes) [-] _"<1"	0.00259852
	A181	Y18<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _"Normal"	0.002528376
	A926	Slope<Blood examination> High sensitive CRP [mg/dL] (for inflammation diagnosis) _">=-0.14 and <0.04"	0.002525193
	A183	Y18<Blood examination> Total protein [g/dL] (for nutrition diagnosis) _"Normal"	0.002490572
	A877	Slope<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _">=-15.66 and <1.92"	0.00235759
	A222	Y18<Health questionnaire> PSQI total score (0-21 points) [-] _"0-5"	0.002351933
	A379	Y16<Health questionnaire> Social connections (Number of family members and relatives who feel comfortable) [-] _"1-5"	0.002278417
(2) At most 10 factors related to driving only	A1079	Y19Car<Driving questionnaire> Purpose of driving _"Commuting"	0.000968019
	A1087	Y19Car<Driving questionnaire> Purpose of driving _"Going for the entertainment"	0.00011286
	A1050	Y19Car<Driving questionnaire> Likes or dislikes about driving _"Dislike"	0.000038163

Then, each $\boldsymbol{y}_i$ that minimizes C is obtained by using a gradient descent method. That is, for each iteration $t = 2, 3, \ldots, t_{\max}$, compute the gradient

$$\frac{\partial C}{\partial \boldsymbol{y}_i} = 4 \sum_j \left(p_{ij} - q_{ij}\right) \left(\boldsymbol{y}_i - \boldsymbol{y}_j\right) \left(1 + \left\|\boldsymbol{y}_i - \boldsymbol{y}_j\right\|^2\right)^{-1} \tag{23}$$

and update the t-th solution $\boldsymbol{Y}^{(t)} = \left[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\right]$ with

$$\boldsymbol{Y}^{(t)} = \boldsymbol{Y}^{(t-1)} + \eta \frac{\partial C}{\partial \boldsymbol{Y}^{(t-1)}} + \alpha(t) \left(\boldsymbol{Y}^{(t-1)} - \boldsymbol{Y}^{(t-2)}\right), \tag{24}$$

TABLE 20. List of factors $w \in C_w^{(4)}$ corresponding to (1) and (2) in Fig. 12.

	Binary variables w	Description of variables	$P(w\|z_4)$
(1) Top 30 factors based on descending order of $P(w\|z_4)$	A306	Y17<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _"Normal"	0.003314465
	A290	Y17<Blood examination> AST (GOT) [U/L] (for liver diagnosis) _"Normal"	0.003305411
	A310	Y17<Blood examination> Hemoglobin [g/dL] (for anemia diagnosis) _"Normal"	0.003276421
	A292	Y17<Blood examination> ALT (GPT) [U/L] (for liver diagnosis) _"Normal"	0.00325929
	A316	Y17<Blood examination> Cortisol [μg/dL] (for stress diagnosis) _"Normal"	0.003228009
	A302	Y17<Blood examination> TG [mg/dL] (for lipid diagnosis) _"Normal"	0.003200871
	A322	Y17<Blood pressure measurement of the extremities> Rabi [-] (for artery occlusion diagnosis) _"Normal"	0.003181471
	A324	Y17<Blood pressure measurement of the extremities> Labi [-] (for artery occlusion diagnosis) _"Normal"	0.003181471
	A296	Y17<Blood examination> Total protein [g/dL] (for nutrition diagnosis) _"Normal"	0.003179882
	A314	Y17<Blood examination> Fischer's ratio [-] (for amino acid diagnosis) _"Normal"	0.003166197
	A304	Y17<Blood examination> HDL cholesterol [mg/dL] (for lipid diagnosis) _"Normal"	0.003126633
	A288	Y17<Blood examination> Total bilirubin [mg/dL] (for liver diagnosis) _"Normal"	0.003120051
	A333	Y17<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Common) [-] _"Common"	0.003108015
	A294	Y17<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _"Normal"	0.003077983
	A309	Y17<Blood examination> White blood cell count [/μL] (for immunodiagnosis) _"Normal"	0.003071518
	A259	Y17<Cognitive function> MMSE score (0-30 points) [-] _"3"	0.003061469
	A300	Y17<Blood examination> Uric acid [mg/dL] (for gout diagnosis) _"Normal"	0.002981912
	A282	Y17<Vision test> Near vision (Right eye) [-] _"<1"	0.00297398
	A284	Y17<Vision test> Near vision (Left eye) [-] _"<1"	0.002970005
	A312	Y17<Blood examination> HbA1c (NGSP) [%] (for diabetes diagnosis) _"Normal"	0.002901345
	A320	Y17<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.002874015
	A279	Y17<Vision test> Distance vision (Left eye) [-] _"<1"	0.002853675
	A298	Y17<Blood examination> Creatinine [mg/dL] (for kidney diagnosis) _"Normal"	0.002822479
	A975	Slope<Gravic body sway test (Otorhinolaryngology)> Evaluation of large sway (1: Large or 0: Common) [-] _">=-0.09 and <0.19"	0.002736866
	A318	Y17<Blood pressure measurement of the extremities> RbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.002635095
	A277	Y17<Vision test> Distance vision (Right eye) [-] _"<1"	0.002604096
	A330	Y17<Body composition measurement> Body fat percentage [%] _"High"	0.002590865
	A287	Y17<Vision test> Near vision (Both eyes) [-] _"<1"	0.002585172
	A328	Y17<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Normal"	0.002544393
	A946	Slope<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _">=-105.64 and <140.48"	0.002534276
(2) At most 10 factors related to driving only	A1046	Y19Car<Driving questionnaire> Transmission (AT or MT) _"Both AT and MT"	0.000326749
	A1070	Y19Car<Driving questionnaire> Scenes where you have felt better while driving _"None"	0.000276265

where $t_{\max}$, η and $\alpha(t)$ represent the predefined values named the maximum number of iterations, the learning rate and the momentum, respectively. The initial value $\boldsymbol{Y}^{(1)} = \boldsymbol{Y}^{(0)}$ is generated by a normal random number with a mean 0 and a standard deviation 10^{-4}.

C. COMPONENT FACTORS APPEARING IN (1) AND (2) OF FIGS. 4 TO 13

In this appendix, we list component factors (binary variables) appearing in (1) and (2) of Figs. 4 to 13 as Tables 12 to 21 in text format, respectively.

TABLE 21. List of factors $w \in C_w^{(5)}$ corresponding to (1) and (2) in Fig. 13.

	Binary variables w	Description of variables	$P(w \mid z_5)$
(1) Top 30 factors based on descending order of $P(w \mid z_5)$	A219	Y18<Body composition measurement> Body fat percentage [%] _"High"	0.002652552
	A211	Y18<Blood pressure measurement of the extremities> RbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.002315342
	A212	Y18<Blood pressure measurement of the extremities> LbaPWV [cm/s] (for atherosclerosis diagnosis) _"Abnormal"	0.002184744
	A440	Y16<Body composition measurement> Body fat percentage [%] _"High"	0.002182334
	A250	Y17<Personal record form> Two-step test [cm] _">=194.2 and <240.8"	0.002013026
	A397	Y16<Vision test> Near vision (Both eyes) [-] _"<1"	0.001998066
	A217	Y18<Heel bone density test> T-score derived from OSI [-] (for osteoporosis diagnosis) _"Normal"	0.001973037
	A883	Slope<Blood examination> Total protein [g/dL] (for nutrition diagnosis) _">=-0.24 and <-0.06"	0.001873763
	A907	Slope<Blood examination> ALB (improved BCP method) [g/dL] (for nutrition diagnosis) _">=-0.18 and <-0.02"	0.00182702
	A819	Slope<Personal record form> Body weight [kg] _">=-0.2 and <0.91"	0.001709939
	A458	Y18<Health questionnaire> CESD-20 score (0-60 points) [-] _">=41 and <=50"	0.001633984
	A154	Y18<Personal record form> Body weight [kg] _">=56.74 and <75.98"	0.001612866
	A372	Y16<Personal record form> Body height [cm] _">=147.46 and <156.92"	0.001542413
	A362	Y16<Personal record form> Two-step test [cm] _">=189.8 and <242.2"	0.001495397
	A998	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : PF _">=-10 and <0"	0.001428976
	A936	Slope<Blood examination> Cortisol [μg/dL] (for stress diagnosis) _">=-0.36 and <3.56"	0.001394567
	A477	Y16<Health questionnaire> CESD-20 score (0-60 points) [-] _">=41 and <=50"	0.001328362
	A265	Y17<Personal record form> Body weight [kg] _">=56.92 and <76.94"	0.001234584
	A657	Y17<Health questionnaire> SF-36v2 score (0-100 points) [-] : GH _">=40 and <50"	0.001198635
	A991	Slope<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=-8.8 and <-0.7"	0.001190264
	A295	Y17<Blood examination> γ-GT (γ-GTP) [U/L] (for liver diagnosis) _"Abnormal"	0.001146688
	A1007	Slope<Health questionnaire> SF-36v2 score (0-100 points) [-] : BP _">=-17.6 and <-2.4"	0.00110331
	A780	Slope<Personal record form> Sitting trunk flexion [cm] _">=0.4 and <3.6"	0.001073857
	A240	Y17<Personal record form> Sit-to-stand test [cm] _">=(Min)10 and <16"	0.001064069
	A462	Y18<Health questionnaire> GLFS-25 score (0-100 points) [-] _">=81 and <=90"	0.001048649
	A352	Y16<Personal record form> Sit-to-stand test [cm] _">=(Min)10 and <16"	0.001038759
	A228	Y17<Personal record form> Grip strength (Left hand) [kg] _">=(Min)8 and <20.8"	0.001025616
	A536	Y19<Health questionnaire> SF-36v2 score (0-100 points) [-] : SF _">=70 and <80"	0.001007574
	A225	Y17<Personal record form> Grip strength (Right hand) [kg] _">=(Min)10 and <21"	0.001000375
	A356	Y16<Personal record form> TUG test [sec] _">=5.24 and <7.38"	0.000967239
(2) At most 10 factors related to driving only	A1039	Y19Car<Driving questionnaire> Driving frequency _"None"	0.000852574
	A1085	Y19Car<Driving questionnaire> Purpose of driving _"Going to the hospital"	0.000404522
	A1084	Y19Car<Driving questionnaire> Purpose of driving _"Going for the leisure"	0.000080476
	A1088	Y19Car<Driving questionnaire> Purpose of driving _"Driving"	0.000062799

ACKNOWLEDGMENT

The authors would like to express their sincere appreciation to the Technical Leader Takehisa Kohira at Mazda Motor Corporation's Technical Research Center and everyone involved, including the reviewers, for their valuable comments in writing this article.

REFERENCES

[1] World Health Organization (WHO). (2023). *Global Health Observatory Life Expectancy At Birth (Years)*. Accessed: Aug. 4, 2023. [Online]. Available: https://www.who.int/data/gho/data/themes/topics/indicator-groups/indicator-group-details/GHO/life-expectancy-and-healthy-life-expectancy

[2] *Ministry of Health, Labour and Welfare (MHLW) of Japan, Abridged Life Tables For Japan 2022*. Accessed: Aug. 4, 2023. [Online]. Available: https://www.mhlw.go.jp/english/database/db-hw/vs02.html

[3] S. Chihuri, T. J. Mielenz, C. J. DiMaggio, M. E. Betz, C. DiGuiseppi, V. C. Jones, and G. Li, "Driving cessation and health outcomes in older adults," *J. Amer. Geriatrics Soc.*, vol. 64, no. 2, pp. 332–341, Feb. 2016.

[4] H. Shimada, S. Lee, T. Doi, S. Bae, K. Tsutsumimoto, and H. Arai, "Prevalence of psychological frailty in japan: NCGG-SGS as a Japanese national cohort study," *J. Clin. Med.*, vol. 8, no. 10, p. 1554, Sep. 2019.

[5] R. Nouchi, A. Kobayashi, H. Nouchi, and R. Kawashima, "Newly developed TV-based cognitive training games improve car driving skills, cognitive functions, and mood in healthy older adults: Evidence from a randomized controlled trial," *Frontiers Aging Neurosci.*, vol. 11, pp. 1–15, May 2019.

[6] J. F. Sallis, L. D. Frank, B. E. Saelens, and M. K. Kraft, "Active transportation and physical activity: Opportunities for collaboration on transportation and public health research," *Transp. Res. A, Policy Pract.*, vol. 38, no. 4, pp. 249–268, May 2004.

[7] H. Hirai, M. Ichikawa, N. Kondo, and K. Kondo, "The risk of functional limitations after driving cessation among older Japanese adults: The JAGES cohort study," *J. Epidemiol.*, vol. 30, no. 8, pp. 332–337, 2020.

[8] K. Tomioka, N. Kurumatani, and H. Hosoi, "Relationship of having hobbies and a purpose in life with mortality, activities of daily living, and instrumental activities of daily living among community-dwelling elderly adults," *J. Epidemiol.*, vol. 26, no. 7, pp. 361–370, 2016.

[9] A. Camurri, S. Hashimoto, M. Ricchetti, K. Suzuki, R. Trocca, and G. Volpe, "KANSEI analysis of movement in dance/music interactive systems," *Presented at the 2nd Int. Symp. Humanoid Robotics*, Tokyo, Japan, 1999.

[10] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big data from healthcare communities," *IEEE Access*, vol. 5, pp. 8869–8879, 2017.

[11] W. Huang, T. W. Ying, W. L. C. Chin, L. Baskaran, O. E. H. Marcus, K. K. Yeo, and N. S. Kiong, "Application of ensemble machine learning algorithms on lifestyle factors and wearables for cardiovascular risk prediction," *Sci. Rep.*, vol. 12, no. 1, pp. 1–12, Jan. 2022.

[12] G. S. Tomar, M. Singh, S. Rai, A. Kumar, R. Sanyal, and S. Sanyal, "Probabilistic latent semantic analysis for unsupervised word sense disambiguation," *Int. J. Comput. Sci. Issues*, vol. 10, no. 5, pp. 127–133, 2013.

[13] C. H. Cao and H. L. Cao, "The research on medical image classification algorithm based on PLSA-BOW model," *Technol. Health Care*, vol. 24, pp. 665–674, Jun. 2016.

[14] R. Corizzo, Y. Dauphin, C. Bellinger, E. Zdravevski, and N. Japkowicz, "Explainable image analysis for decision support in medical healthcare," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2021, pp. 4667–4674.

[15] R. Myte, B. Gylling, J. Häggström, J. Schneede, P. M. Ueland, G. Hallmans, I. Johansson, R. Palmqvist, and B. Van Guelpen, "Untangling the role of one-carbon metabolism in colorectal cancer risk: A comprehensive Bayesian network analysis," *Sci. Rep.*, vol. 7, no. 1, pp. 1–12, Feb. 2017.

[16] J. Walsh, I. T. Heazlewood, M. DeBeliso, and M. Climstein, "Application of t-distributed stochastic neighbor embedding (t-SNE) to clustering of social affiliation and recognition psychological motivations in masters athletes," *Int. J. Sport, Exercise Health Res.*, vol. 4, no. 1, pp. 1–6, May 2020.

[17] T. Hofmann, "Probabilistic latent semantic analysis," in *Proc. Conf. Uncertainty Artif. Intell.*, 1999, pp. 289–296.

[18] L. J. P. van der Maaten and G. E. Hinton, "Visualizing high-dimensional data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

[19] S. Nakaji, K. Ihara, K. Sawada, S. Parodi, T. Umeda, I. Takahashi, K. Murashita, S. Kurauchi, and I. Tokuda, "Social innovation for life expectancy extension utilizing a platform-centered system used in the Iwaki health promotion project: A protocol paper," *SAGE Open Med.*, vol. 9, Jan. 2021, Art. no. 205031212110026.

[20] E. Sasaki, S. Sasaki, D. Chiba, Y. Yamamoto, A. Nawata, E. Tsuda, S. Nakaji, and Y. Ishibashi, "Age-related reduction of trunk muscle torque and prevalence of trunk sarcopenia in community-dwelling elderly: Validity of a portable trunk muscle torque measurement instrument and its application to a large sample cohort study," *PLoS ONE*, vol. 13, no. 2, Feb. 2018, Art. no. e0192687.

[21] A. Takano, Y. Iwama, T. Kanno, and T. Kan, "The relationship between whole body reaction time and physical fitness factors due to differences in exercise habits," *Nat. Inst. Technol., Ichinoseki*, vol. 52, pp. 9–14, Dec. 2017.

[22] J. Alcazar, J. Losa-Reyna, C. Rodriguez-Lopez, A. Alfaro-Acha, L. Rodriguez-Mañas, I. Ara, F. J. García-García, and L. M. Alegre, "The sit-to-stand muscle power test: An easy, inexpensive and portable procedure to assess muscle power in older people," *Experim. Gerontol.*, vol. 112, pp. 38–43, Oct. 2018.

[23] D. Podsiadlo and S. Richardson, "The timed 'up & go': A test of basic functional mobility for frail elderly persons," *J. Amer. Geriatrics Soc.*, vol. 39, no. 2, pp. 142–148, 1991.

[24] H. Kato, Y. Takahashi, C. Iseki, R. Igari, H. Sato, H. Sato, S. Koyama, M. Tobita, T. Kawanami, M. Iino, K. Ishizawa, and T. Kato, "Tooth loss-associated cognitive impairment in the elderly: A community-based study in Japan," *Internal Med.*, vol. 58, no. 10, pp. 1411–1416, 2019.

[25] R. M. Carpiano and P. W. Hystad, "'Sense of community belonging' in health surveys: What social capital is it measuring?" *Health Place*, vol. 17, no. 2, pp. 606–617, 2011.

[26] K. Ahrberg, M. Dresler, S. Niedermaier, A. Steiger, and L. Genzel, "The interaction between sleep quality and academic performance," *J. Psychiatric Res.*, vol. 46, no. 12, pp. 1618–1622, Dec. 2012.

[27] N. Okubo, M. Matsuzaka, I. Takahashi, K. Sawada, S. Sato, N. Akimoto, T. Umeda, and S. Nakaji, "Relationship between self-reported sleep quality and metabolic syndrome in general population," *BMC Public Health*, vol. 14, no. 1, pp. 1–7, Dec. 2014.

[28] Y. Y. Leung, T. C. P. Lee, S. M. Y. Ho, and L. K. Cheung, "Trigeminal neurosensory deficit and patient reported outcome measures: The effect on life satisfaction and depression symptoms," *PLoS ONE*, vol. 8, no. 8, Aug. 2013, Art. no. e72891.

[29] *Ministry of Health, Labour and Welfare (MHLW) of Japan, CES-D*. Accessed: Sep. 15, 2023. [Online]. Available: https://kokoro.mhlw.go.jp/glossaries/word-1568/

[30] N. Sugawara, A. Sasaki, N. Yasui-Furukori, S. Kakehata, T. Umeda, A. Namba, S. Nakaji, H. Shinkawa, and S. Kaneko, "Hearing impairment and cognitive function among a community-dwelling population in Japan," *Ann. Gen. Psychiatry*, vol. 10, no. 1, p. 27, 2011.

[31] H. Tsuji, T. Tetsunaga, T. Tetsunaga, H. Misawa, K. Nishida, and T. Ozaki, "Cognitive factors associated with locomotive syndrome in chronic pain patients: A retrospective study," *J. Orthopaedic Sci.*, vol. 26, no. 5, pp. 896–901, Sep. 2021.

[32] D. Chiba, E. Tsuda, K. Wada, G. Kumagai, E. Sasaki, A. Nawata, S. Nakagomi, I. Takahashi, S. Nakaji, and Y. Ishibashi, "Lumbar spondylosis, lumbar spinal stenosis, knee pain, back muscle strength are associated with the locomotive syndrome: Rural population study in Japan," *J. Orthopaedic Sci.*, vol. 21, no. 3, pp. 366–372, May 2016.

[33] S. M. Foster and E. H. Maples, "Occupational stress in veterinary support staff," *J. Veterinary Med. Educ.*, vol. 41, no. 1, pp. 102–110, Mar. 2014.

[34] *TOMEY, Compact Vision Meter*. Accessed: Nov. 16, 2023. [Online]. Available: http://www.ritz-med.co.jp/info/wp-content/uploads/2017/09/5CA-1000.pdf.

[35] *Japan Society of Ningen Dock, How To Read the Examination Chart*. Accessed: Sep. 15, 2023. [Online]. Available: https://www.ningen-dock.jp/public/method

[36] *LSI Medience, Explanation of Inspection Items*. Accessed: Sep. 13, 2023. [Online]. Available: https://data.medience.co.jp/compendium/toc.html

[37] *Kyoto Microbio Laboratory, Biochemical Tests*. Accessed: Sep. 15, 2023. [Online]. Available: http://www.kyobiken.or.jp/system/site_data/site_0/page_677/version_5/file/0037.pdf

[38] J. Morinaga, J. Zhao, M. Endo, T. Kadomatsu, K. Miyata, T. Sugizaki, Y. Okadome, Z. Tian, H. Horiguchi, K. Miyashita, N. Maruyama, M. Mukoyama, and Y. Oike, "Association of circulating ANGPTL 3, 4, and 8 levels with medical status in a population undergoing routine medical checkups: A cross-sectional study," *PLoS ONE*, vol. 13, no. 3, Mar. 2018, Art. no. e0193731.

[39] M. Takae, E. Yamamoto, T. Tokitsu, F. Oike, T. Nishihara, K. Fujisue, D. Sueta, H. Usuku, K. Motozato, M. Ito, and H. Kanazawa, "Clinical significance of brachial-ankle pulse wave velocity in patients with heart failure with reduced left ventricular ejection fraction," *Amer. J. Hypertension*, vol. 32, no. 7, pp. 657–667, Jun. 2019.

[40] C.-C. Lin, T.-Y. Chen, C.-I. Li, C.-S. Liu, C.-H. Lin, M.-C. Wang, S.-Y. Yang, and T.-C. Li, "Renal markers and risks of all cause and cardiovascular mortality from the Taichung community based cohort study," *Sci. Rep.*, vol. 11, no. 1, p. 14143, Jul. 2021.

[41] *Global Initiative for Chronic Obstructive Lung Disease (GOLD), 2020 Global Strategy for Prevention, Diagnosis and Management of COPD*. Accessed: Sep. 15, 2023. [Online]. Available: http://goldcopd.org/gold-2017-global-strategy-diagnosis-management-prevention-copd/

[42] S. Soen, M. Fukunaga, T. Sugimoto, T. Sone, S. Fujiwara, N. Endo, I. Gorai, M. Shiraki, H. Hagino, T. Hosoi, H. Ohta, T. Yoneda, and T. Tomomitsu, "Diagnostic criteria for primary osteoporosis: Year 2012 revision," *J. Bone Mineral Metabolism*, vol. 31, no. 3, pp. 247–257, May 2013.

[43] *GE HealthCare, Osteoporosis Prevention and Treatment Guidelines*. Accessed: Sep. 15, 2023. [Online]. Available: https://www.gehealthcare.co.jp/products/bone-and-metabolic-health/kotsumitsu/technicaltips00r

[44] L. Hu, X. Huang, C. You, J. Li, K. Hong, P. Li, Y. Wu, Q. Wu, Z. Wang, R. Gao, H. Bao, and X. Cheng, "Prevalence of overweight, obesity, abdominal obesity and obesity-related risk factors in southern China," *PLoS ONE*, vol. 12, no. 9, Sep. 2017, Art. no. e0183934.

[45] *Nagahama Red Cross Hospital, Body Fat Percentage*. Japanese Red Cross Soc. Accessed: Sep. 15, 2023. [Online]. Available: https://www.nagahama.jrc.or.jp/departments/topics/taishibo.html

[46] *Yokohama Sports Medical Center, Health & Fitness Information*. Accessed: Sep. 15, 2023. [Online]. Available: https://www.yspc-ysmc.jp/ysmc/column/health-fitness/diet-practice-2.html.

[47] ANIMA. *Introduction of Data on Healthy Values of Gravic Body Sway Meter*. Accessed: Nov. 16, 2023. [Online]. Available: https://anima.jp/products/healthy_value/

[48] S. Goto, A. Sasaki, I. Takahashi, Y. Mitsuhashi, S. Nakaji, and A. Matsubara, "Relationship between cognitive function and balance in a community-dwelling population in Japan," *Acta Oto-Laryngologica*, vol. 138, no. 5, pp. 471–474, May 2018.

[49] P. Katz, "Function, disability, and psychological well-being," *Adv. Psychosom. Med.*, vol. 25, pp. 41–62, Jan. 2004.

[50] D. J. French, R. J. West, J. Elander, and J. M. Wilding, "Decision-making style, driving style, and self-reported involvement in road traffic accidents," *Ergonomics*, vol. 36, no. 6, pp. 627–644, Jun. 1993.

[51] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proc. 2nd Int. Symp. Inf. Theory*, B. N. Petrov and F. Caski, Eds., Akademiai Kiado, Budapest, 1973, pp. 267–281.

MITSUGU MERA received the Ph.D. degree from the Graduate School for Mathematics, Kyushu University, Japan, in 2010. He is currently with Mazda Motor Corporation, Hiroshima, Japan. His current research interests include utilizations of mathematics related to data analysis and computer science in automotive industries.

NANAE MICHIDA received the Ph.D. degree from the Graduate School of Biosphere Science, Hiroshima University, Japan, in 2006. She is currently with Mazda Motor Corporation, Hiroshima, Japan. She specializes in psychophysiology. Her current research interest includes people's well-being with automobiles.

MASANORI HONDA received the Ph.D. degree from the Graduate School of Engineering, Hiroshima University, Japan, in 2021. He is currently with Mazda Motor Corporation, Hiroshima, Japan. He specializes in mechanical engineering. His current research interest includes people's well-being with automobiles.

KAZUO SAKAMOTO is currently with Mazda Motor Corporation, Hiroshima, Japan. His current research interest includes modeling of driver and passenger in vehicles.

YOSHINORI TAMADA received the Ph.D. degree in informatics from Kyoto University, in 2005. He is currently a Professor with the Department of Medical Data Intelligence and the Director of the Research Center for Health-Medical Data Science, Hirosaki University Graduate School of Medicine. His primary research interests include health and medical big data analysis, especially using with health and biological network estimation methods and supercomputers.

TATSUYA MIKAMI received the Ph.D. degree from Hirosaki University Graduate School of Medicine, Japan, in 1997. He is currently a Professor with the Department of Preemptive Medicine, Hirosaki University Graduate School of Medicine, conducting research and social activities to promote the health promotion of local residents.

SHIGEYUKI NAKAJI received the Ph.D. degree from Hirosaki University Graduate School of Medicine, Japan, in 1981. He is currently specially-appointed as a Professor with the Department of Social Medicine, Hirosaki University Graduate School of Medicine, conducting research and social activities to promote the health promotion of local residents.