

1. Development of schizophrenia classification model

The schizophrenia classification model (SZ-classifier) was developed in our previous study [1]. The sample set for developing the SZ-classifier was collected by an internet research agency's pooled panel (Rakuten Insight, Inc.) [2]. It comprised 223 schizophrenia patients and 1776 healthy controls. The following information of each sample was collected: the existence of schizophrenia which is the response and 76 features, including demographic, health-related backgrounds, physical comorbidities, psychiatric comorbidities, and social comorbidities. The inclusion criteria are described in section **1.1 Data collection and inclusion criteria**. The feature definitions are described in section **1.2 Feature definitions**.

The SZ-classifier was constructed based on artificial neural network considering the significance lies in capturing non-linear relationship and higher order interaction. Specifically, it was structured with 5 hidden layers: 128-64-32-16-8 neurons in each layer, HeNormal weight initializer, ReLU activation function in the hidden layers, and sigmoid activation in the output layer. The programming was performed using Python (version 3.7), and computational environment used was Jupyter Notebook. Details of the development has been published elsewhere [1].

1.1 Data collection and inclusion criteria [3]

Data were collected from an internet research agency's pooled panels (Rakuten Insight, Inc., incorporated approximately 2.3 million panelists by 2022).[2] A prevalence case-control study was conducted using those panels. Subjects' age was limited from 20- to 75 years old.

For involving subjects who currently have schizophrenia, 5,584 individuals who self-reported schizophrenia were sampled in the Rakuten Insight disease panel. 3,256 respondents answered the following four questions in advance of the survey: (1) Are you currently suffering from schizophrenia only, schizophrenia and migraine, schizophrenia and a sleep disorder, or schizophrenia, migraine, and a sleep disorder? (2) Have you experienced auditory hallucinations lasting more than one month? (3) Have you never used stimulants or other illegal drugs and never been an alcoholic? (4) Have you experienced the first auditory hallucination lasting more than one month at less than 60 years of age? Those who answered "yes" to all four questions were considered to currently have schizophrenia. Therefore, 223 respondents were involved in the survey.

For involving subjects who do not currently have schizophrenia, 28,000 participants in the Japan COVID-19 and Society Internet Survey (JASTIS, which was also conducted using the Rakuten Insight Panel)[4-6] were sampled. 6,656 respondents answered the following four questions in advance of the survey: (1) Are you currently suffering from mental illness? (2) Have you had mental illness in the past? (3) Have you experienced auditory hallucinations? (4) Have you ever used stimulants or other illegal drugs, been

an alcoholic, or received psychiatric treatment? Those who answered “no” to all four questions were considered not to have schizophrenia. Therefore, 1,776 respondents were involved in the survey.

Finally, 223 Subjects aged 20-75 years who currently have schizophrenia and 1776 health controls were included in this study.

1.2 Feature definitions [3]

Feature variables were formatted from a self-administered questionnaire assessed 1) demographic, health-related backgrounds, and physical comorbidities, 2) psychiatric comorbidities, and 3) social comorbidities.

1) Demographic, health-related backgrounds, and physical comorbidities

Demographic, health-related backgrounds, and physical comorbidities include **age**, **body mass index (BMI)**, **smoking status** [current, past, never], **number of cigarettes per day**, **alcohol drinking frequency** [never, past, current: <1 day in one month; 1-3 days in one month; 1-2 days in one week; 3-4 days in one week; 5-6 days in one week; everyday.], **sports** [frequency: < once in one month; 1-3 times in one month; 1-2 times in one week; 3-4 times in one week; almost every day. Intensity: each time < 30 minutes; 30-59 minutes; 1-2 hours; 2-3 hours; 3-4 hours; >4 hours], **eating behaviors** [tendency to overeat: yes/no; eating speed: very fast, fast, normal, slow, very slow; having breakfast, eating out, and eating instant foods: < once in one month; 1-3 times in one month; 1-2 times in one week; 3-4 times in one week; almost every day], **bowel movement** [frequency: <3times per week, 3-4 times per week, 5-6 times per week, once a day, > once a day], **Stool** [soft, normal, hard, recurrent diarrhea and constipation], **restriction in functional capacity** [7], **self-rated health status (SRHS)** [8], and **physical comorbidities** [overweight, cancer, cardiovascular disease, heart failure, hypertension, diabetes, dyslipidemia, gout, sleep apnea syndrome, and fracture: have been treated or not].

Restriction in functional capacity

To evaluate functional capacity restrictions, we used the Scale of Independence in Daily Living for the Disabled Elderly published by the Ministry of Health, Labour and Welfare, Japan [7]. The term “restrictions in functional capacity” refers to a multidimensional concept that involves sensory loss, impaired mobility, vascular disorders, gait impairments, problems with daily living (ADLs), and changes in body systems. Participants self-assessed restrictions by choosing one of the following options: (1) “I have no physical disabilities,” (2) “I go out alone, using transportation,” (3) “I can only go out alone in my neighborhood,” (4) “I go out with help and live mostly out of bed during the day,” (5) “I can go out with help, but I go out infrequently, and I spend most of the daytime sleeping on and off in bed,” (6) “I can ride in a wheelchair by myself

and eat and toilet away from the bed,” (7) “I can ride in a wheelchair with assistance,” (8) “I can roll over in bed,” and (9) “I cannot roll over in bed.”

Self-rated health status (SRHS)

SRHS is a self-reported measure of health status that incorporates a person’s biological, mental, social, and functional aspects, including individual and cultural beliefs and health behaviors. It is a strong predictor of all-cause mortality in general populations [8]. Participants responded to the question “What do you think of your general health status during the previous month?” by choosing one of the following options: “great,” “pretty good,” “good,” “not so good,” and “bad.”

2) Psychiatric comorbidities

Psychiatric comorbidities include **frequency of sleep medication use** [never, < 1 day per week, 1-2 days per week, 3-4 days per week, 5-6 days per week, every day], **bedtime** [before 7:00pm, around 8:00pm, around 9:00pm, around 10:00pm, around 11:00pm, around 0:00am, around 1:00am, around 2:00am, around 3:00am, after 4:00am, not fixed], **sleep hours** [≤ 5 hours, 6 hours, 7 hours, 8 hours, 9 hours, ≥ 10 hours], **hypnagogic disorder frequency** [almost never, < 1 time per week, 1-2 times per week, 3-4 times per week, 5-6 times per week, almost every day], **deep sleep disorder frequency** (waking up tired) [almost never, < 1 time per week, 1-2 times per week, 3-4 times per week, 5-6 times per week, almost every day], **middle waking, or early waking frequency** [almost never, < 1 time per week, 1-2 times per week, 3-4 times per week, 5-6 times per week, almost every day], **depressive symptoms** [9,10], **perceived stress** [11], **ikigai** (a Japanese term which means positive reason for living), **happiness**, and **internet use hours per week**.

Depressive symptoms (CES-D)

We used a modified 11-item Center for Epidemiological Studies Depression (CES-D) Scale in this study [9,10]. The existence of depressive symptoms was defined as a score of 8 or higher.

Perceived stress (PSS-4)

We assessed perceived stress with a 4-item Perceived Stress Scale (PSS-4) [11]. Scores are on a 16-point scale, with higher total scores indicating more severe perceived stress. Perceived stress was defined as being present when the score was higher than 7, the median of the PSS-4 scores for non-schizophrenic participants.

Ikigai

The Japanese term “Ikigai” is a positive reason for living. Participants were asked, “Do you have any positive reasons to live?” and answered the question from four options: “very much so,” “yes,” “no,” or “not at all.” Participants who answered “no” or “not at all” were defined as absent of ikigai.

Happiness

Participants were asked, “How happy do you feel about your life?” and answered the question from four options: “very happy,” “happy,” “neither happy nor unhappy,” and “unhappy.”

Participants who answered “neither happy nor unhappy” or “unhappy” were defined as absence of happiness.

3) Social comorbidities

Social comorbidities include **taking regular health checkups** [yes/no], **educational background** [junior/senior high school, university, junior college, or vocational school], **type of occupation** [unemployed, homemaker, white-collar workers, or blue-collar workers], **type of employment** [regular, irregular, or self-employed/business people], **household income** [< 3, 3–6, 6–9, 9–12–15, ≥ 15 million Japanese Yen], **marital status** [unmarried, married, divorced, widowed, others], **family structure** [living alone, living with parents, living with spouse, living with children, and living with other people; how many people living with], **social support** [12,13], and **social capital** (cognitive and structural dimensions) [14,15].

Social support

We assessed social support with the ENRICH Social Support Instrument (ESSI) [12,13]. A higher total score indicates higher availability of social support.

Social capital

To assess social capital, we partly referred to the Integrated Questionnaire for the Measurement of Social Capital (SC-IQ) [14]. It is common practice in social capital research to distinguish between structural and cognitive dimensions [15].

We focused on cognitive social capital using the following SC-IQ items: (1) “Can most people be trusted?”; (2) “Does one have to be alert or is someone likely to take advantage of you?”; and (3) “Are most people willing to help if you need it?” Responses were selected from four categories: “strongly disagree,” “disagree,” “agree,” and “strongly agree.” For the three questions, cognitive social capital was defined as high when there were two or more responses of “agree” or “strongly agree” to question (1), “disagree” or “strongly disagree” to question (2), and “agree” or “strongly agree” to question (3).

We also focused on structural social capital using the following SC-IQ items: “How often do you participate in community organizations, self-help groups, charities, volunteer groups, or religious gatherings?” The response was selected from four categories: “not at all/ very seldom,” “sometimes,” “less than once a week,” and “more than once a week.” Structural social capital was defined as high when the response was “more than once a week.”

Table S1. Characteristics of schizophrenia, bipolar, and major depressive disorder samples, among top 5 important features.

| | Schizophrenia N=61 | Bipolar N=32 | Major depressive disorder N=56 |
|---|-----------------------|-----------------|---|
| Frequency of sleep medication use, n(%) | | | |
| Never | 24 (39.3) | 11 (34.4) | 19 (33.9) |
| <1 day, 1-2 days, 3-4 days, 5-6days per week | 7 (11.5) | 7 (21.8) | 6 (10.7) |
| Everyday | 30 (49.2) | 14 (43.8) | 31 (55.4) |
| Age, years, mean±SD | 41.2±13.4 | 46.3±11.8 | 48.5±14.5 ** |
| Household income, n(%) | | | |
| 0-299 million JPY | 24 (39.3) | 10 (31.3) | 16 (28.6) |
| 300-599 million JPY | 19 (31.1) | 14 (43.8) | 21 (37.5) |
| 600-899 million JPY | 10 (16.4) | 5 (15.6) | 12 (21.4) |
| ≥900 million JPY | 8 (13.2) | 3 (9.3) | 7 (22.5) |
| Type of employment, n(%) | | | |
| Unemployed | 36 (59.0) | 14 (43.8) ** | 20 (35.7) ** |
| Regular | 1 (1.6) | 11 (34.4) ** | 26 (46.4) ** |
| Non-regular, others | 24 (39.4) | 7 (21.8) ** | 10 (17.9) ** |
| Bedtime, n(%) | | | |
| Before 10pm | 30 (49.2) | 11 (34.4) | 18 (32.1) * |
| 10pm-12pm | 26 (42.6) | 18 (56.3) | 24 (42.9) * |
| Over 12pm, not regular | 5 (8.2) | 3 (9.3) | 14 (25.0) * |

Bipolar and major depressive disorder are compared with schizophrenia, respectively. The *p*-values are indicated by superscripts: * denotes *p*-value < 0.05, and ** denotes *p*-value < 0.01.

References

1. He Y, Matsunaga M, Li Y, et al. Classifying Schizophrenia Cases by Artificial Neural Network Using Japanese Web-Based Survey Data: Case-Control Study. *JMIR Form Res.* 2023 Nov 15;7:e50193.
2. Rakute disease panel https://insight.rakuten.com/wordpress/wp-content/uploads/Rakuten_Insight_Panel_Book.pdf [Accessed Feb 29.2024]

3. Matsunaga M, Li Y, He Y, et al. Physical, Psychiatric, and Social Comorbidities of Individuals with Schizophrenia Living in the Community in Japan. *Int J Environ Res Public Health*. 2023; 20(5):4336.
4. Kusama T, Kiuchi S, Takeuchi K, Ikeda T, Nakazawa N, Kinugawa A, et al. Information usage and compliance with preventive behaviors for COVID-19: A longitudinal study with data from the JACSIS 2020/JASTIS 2021. *Healthcare*. 2022; 10: 521.
5. Wakabayashi M, Sugiyama Y, Takada M, Kinjo A, Iso H, Tabuchi T. Loneliness and increased hazardous alcohol use: Data from a nationwide Internet survey with 1-year follow-up. *Int J Environ Res Public Health*. 2022; 19: 12086.
6. Sasaki R, Ota A, Yatsuya H, Tabuchi T. Gender difference in fear and anxiety about and perceived susceptibility to COVID-19 in the third wave of pandemic among the Japanese general population: A nationwide web-based cross-sectional survey. *Int. J. Environ. Res. Public Health*. 2022; 19: 16239.
7. The Ministry of Health, Labour and Welfare. The Standards for the Degree of Independent Living for Elderly And Disabled People. 1991. <https://www.mhlw.go.jp/file/06-Seisakujouhou-12300000-Roukenkyoku/0000077382.pdf>. Accessed Oct 17 2023 (in Japanese).
8. DeSalvo KB, Bloser N, Reynolds K, He J, Muntner P. Mortality prediction with a single general self-rated health question. *J Gen Intern Med*. 2006;21(3):267-275. doi:10.1111/j.1525-1497.2005.00291.x
9. Radloff LS. The CES-D scale. *Applied Psychological Measurement*. 1977;1(3):385-401. doi:10.1177/014662167700100306
10. Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D Depression Symptoms index. *J Aging Health*. 1993;5(2):179-193. doi:10.1177/089826439300500202
11. Warrtig SL, Forshaw M, South J, White A. New, normative, English-sample data for the Short Form Perceived Stress Scale (PSS-4). *J Health Psychol*. 2013;18(12):1617-1628. doi:10.1177/1359105313508346
12. The ENRICHD Investigators. Enhancing recovery in coronary heart disease (ENRICHD) study intervention: rationale and design. *Psychosom Med*. 2001; 63: 747–755. PMID: 11573023.
13. Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD): Study design and methods. *American Heart Journal*. 2000;139(1):1-9. doi:10.1016/s0002-8703(00)90301-6
14. Grootaert G, Narayan D, Jones VN, Woolcock M. Measuring Social Capital : an Integrated Questionnaire. *World Bank Publications*. January 2004. <https://ideas.repec.org/b/wbk/wbpubs/15033.html>.
15. Murayama H, Fujiwara Y, Kawachi I. Social Capital and Health: A review of Prospective Multilevel studies. *J Epidemiol*. 2012;22(3):179-187. doi:10.2188/jea.je20110128