ISASE 2020

Characteristics Extraction from Lifestyle Questionnaire Data in Specific Health Checkup Based on Health State Transition of Metabolic Syndrome

Norihiko HASHIMOTO*, Yoshiaki MIYAUCHI ** and Haruhiko NISHIMURA*

**Graduate School of Applied Informatics, University of Hyogo, 7-1-28 Minatojima-minami, Chuo-ku, Kobe, 650-0047, ab17d503@ai.u-hyogo.ac.jp, haru@ai.u-hyogo.ac.jp*
***Graduate School of Nursing, Nagoya City University, 1-Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya, 467-8601, y-miya@med.nagoya-cu.ac.jp*

Abstract: In this study, we extracted the Specific Health Checkup (SHC) stratification of Metabolic Syndrome (MS), the quantification of the questionnaire data and lifestyle factors by using of the data in 5,423 males examined the SHC in two consecutive years (2006, 2007) when SHC was initially started. Based on these data we analyzed the differences in lifestyle habits among Bad group (MS to MS), Good group (non-MS to non-MS), Worse group (non-MS to MS), and Improved group (MS to non-MS) for the state movement from 2006 to 2007, and extracted the characteristics of lifestyle factors specifying MS. As a result, we could grasp the lifestyle habits characterized by the difference between non-MS and MS. In addition, the lifestyle factors that affect the transition between MS and non-MS. It is important for Specific Health Guidance (SHG) to recognize the existence of lifestyle factors that characterize MS and non-MS to take preventive countermeasures for MS. In the future, we would like to connect our findings to the construction of a health support system that uses big data for SHC by linking to more accurate SHG.
Keywords: *Health Checkup, Metabolic Syndrome, Lifestyle*

1. INTRODUCTION

Metabolic Syndrome (MS) has become a significant problem worldwide, and the Specific Health Checkup (SHC) and Specific Health Guidance (SHG) aimed at preventing this condition were initiated in 2008 in Japan [1, 2]. Throughout the country, it is expected that the national health database, which stores huge data of tens of million people, will be linked to effective metabolic management for individual subjects through evaluation and analysis. However, there have been few reports analyzing the relationship between questionnaire data and examination data in the SHC so far [3].

In this study, we focused on the state transition of MS in two consecutive years (2006, 2007) and extracted each characteristics of lifestyle habits among the Bad group (MS to MS), Good group (non-MS to non-MS), Worse group (non-MS to MS), and Improved group (MS to non-MS). Coping with the emerged lifestyle-related factors leads to MS breakaway.

2. MATERIALS AND METHODS

2.1 Target data

We studied the medical examination findings and results of a lifestyle-related questionnaire of 5,423 anonymous male subjects who underwent a health examination in a certain establishment for two consecutive years (2006, 2007). The following SHC items were evaluated: Waist circumference (cm), Body Mass Index (BMI; kg/m^2), Systolic and Diastolic blood pressure (mmHg), Neutral fat (mg/dl), High-Density Lipoprotein (HDL) Cholesterol (mg/dl), Fasting blood glucose level (mg/dl), and Glycated hemoglobin (HbA1c; %). The questionnaire included 36 multiple-choice questions (4-5 possible answers each), with 12 questions each related to the exercise, nutrition, and lifestyle habits of the subject. One question about smoking habits had only two possible answers [4].

2.2 Stratification of examinees

The procedure of stratification based on the reference value of the inspection item is as follows [1]:

Step 1: Judge the risk of visceral fat accumulation by the Waist circumference and BMI.
I. Waist circumference ≥ 85 cm in male
II. Waist circumference < 85 cm in male and BMI ≥ 25

Step 2: Count additional risks from the results of physical tests and questionnaire.
(1) Blood glucose: Fasting blood glucose ≥ 100 mg/dl or HbA1c ≥ 5.6 %, or Receiving drug treatment (-checked from questionnaire)

(2) Lipid: Triglyceride ≥150 mg/dl or
HDL cholesterol < 40 mg/dl, or Receiving drug treatment
(3) Blood pressure: Systolic BP ≥130 mmHg or
Diastolic BP ≥85 mmHg or Receiving drug treatment
(4) Smoking: Experienced (-checked from questionnaire)
*(4) is included only if at least one risk exists
throughout (1) - (3)

Step 3: Classify into four health guidance levels for the examinees who correspond to the case I in Step 1. According to the number of additional risks in (1) - (4) (abbreviate to #ARs):
If #ARs ≥ 2, then [Positive support] level.
If #ARs = 1, then [Motivation support] level.
If #ARs = 0, then [Information provision] level.

For the examinees who correspond to the case II in Step 1,
If #ARs ≥ 3, then [Positive support] level.
If #ARs = 1 or 2, then [Motivation support] level.
If #ARs = 0, then [Information provision] level.

And the others are classified into [Excluded health guidance] level.

2.3 MS state transition based on support level

1. At each support level in the Stratification described in section 2.2, the Positive support and Motivation support are classified as MS, and the Information provision and Excluded health guidance are classified as non-MS.
2. The group was divided into four groups (Bad, Improved, Worse, and Good), depending on the MS and non-MS states based on the support level [6,7].

2.4 Binary representation of the examination data

To express the health status based on the reference value, binarization is performed of the examination data in Section 2.1. The following pairs of factors included in our inspection data are known to contribute to the risk of metabolic syndrome: (1) waist circumference and BMI, (2) fasting blood glucose and HbA1c, (3) neutral fat and HDL cholesterol, and (4) systolic and diastolic blood pressures. Therefore, as an expression of health according to the inspection data, these data (body shape, blood sugar, lipids, and blood pressure) can be represented by 4-bit representation of the 16 potential outcomes, ranging from the (0000) state to (1111) state, where 1 is defined as one or both of the two items of each factor being outside the reference value/range, and 0 is defined as both items being inside the reference (Fig.1) [4].

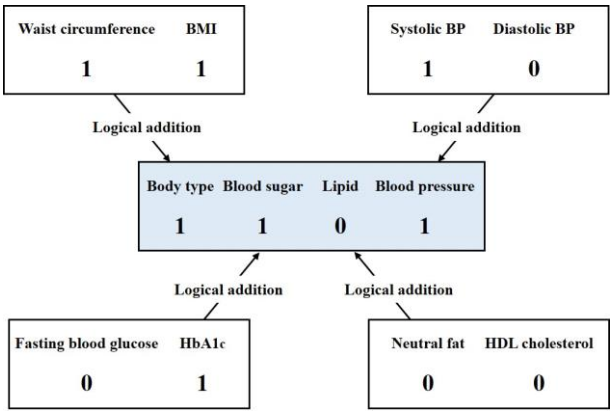

Figure 1: Example of the binary representation for a set of examination data

2.5 Questionnaire data scoring

We allocated a numerical value of 1 - 5 points to the questionnaire data based on the appropriate degree of choice [4]. The followings are examples of scoring options.

Q 1: Do you care the balance between the amount of eating (energy intake) and exercise (energy consumption)?
A 1. I always care. (5 point)
A 2. I frequently care. (4 point)
A 3. I sometimes care. (3 point)
A 4. I occasionally care. (2 point)
A 5. I hardly ever care. (1 point)

Q14: How much do you eat?
A 1. I eat considerably less. (3 point)
A 2. I always eat moderately. (5 point)
A 3. I eat not so much. (4 point)
A 4. I eat a little too much. (2 point)
A 5. I always eat to feel full. (1 point)

The potential sum of all 36 questions ranged from 36 to 180 points. For the alternative Q34 for smoking, we quantified a different pattern from the others.

2.6 Extraction of lifestyle from questionnaire data

Lifestyle factors are extracted by principal component analysis by using the points for all 36 questions. The number of factors was set to the number of principal components with a cumulative contribution ratio of 0.6 or more, and selected a representative question, which factor load after varimax rotation is 0.4 or more [5].

2.7 Characteristics extraction of MS lifestyle based on state transition

1. From section 2.4, we confirmed the 4-bit health state transition from 2006 to 2007. In addition, in order to clearly grasp the characteristics of the MS and non-MS, the states (1000, 1001) for the Waist circumference and BMI, and the Systolic and Diastolic blood pressures, which correspond to the boundary region among the groups (Bad, Improved, Worse, and Good) were excluded.
2. To check the average difference of each question point between Good group and Bad group, the average difference rate (ADR) was calculated by the following equation:

   {(Good group average point of Q#) - (Bad group average point of Q#)} / (Bad group average point of Q#).

   The ADR for each lifestyle factor was evaluated from the sum of the ADRs of the representative question items selected in section 2.6:

   (Total ADR of representative question items for a lifestyle factor) / (the number of representative question items).

   A similar evaluation was made between Improved group and Worse group. We extracted lifestyle factors affecting MS through the both cases (Good vs Bad, Improved vs Worse).

The data were processed using the software R (ver. 3.6.1).

3. RESULTS

3.1 MS state transition based on support level data

Table 2 shows the result of classifying the state transitions between MS and non-MS into four groups according to section 2.5. The order of the number of subjects became to be Good > Bad > Improved > Worse.

Table 2: MS state transition from 2006 to 2007

2006 \ 2007	MS 1,741	non-MS 3,682
MS 1,794	Bad 1,347	Improved 447
non-MS 3,629	Worse 394	Good 3,235

3.2 Characteristics extraction of MS lifestyle based on state transition

From sections 2.4 and 2.7, the subjects in each group on the transition table between the 4-bit health state (16 × 16) are shown as the colored areas in Table 3; Bad (upper left: 1,061), Improved (upper right: 195), Worse (lower left: 151), and Good (lower right: 2,700), excluding the boundary states (1000, 1001) for the Waist circumference and BMI, and Systolic and Diastolic blood pressure.

Table 3: Transitions between 4-bit health states and the number of appropriate subjects

2006 \ 2007	1111	1110	1101	1011	1100	1010	1001	1000	0111	0110	0101	0011	0100	0010	0001	0000
1111	76	27	34	26	11	7	11	6	9	4	8	2	4	1	3	3
1110	26	96	7	4	33	15	5	11	2	6	3	0	6	5	0	6
1101	25	8	103	4	35	0	36	11	4	0	16	0	6	1	6	7
1011	20	8	11	65	3	35	34	15	0	0	1	2	1	2	3	3
1100	7	22	40	2	108	7	7	33	0	1	3	0	13	1	2	13
1010	9	27	7	27	5	121	9	62	0	2	0	0	1	24	5	16
1001	6	2	31	23	8	11	103	51	0	0	7	0	4	1	30	21
1000	4	6	13	17	56	55	58	290	0	1	5	0	9	10	15	102
0111	6	1	2	1	0	0	0	0	16	9	12	3	3	5	6	1
0110	3	6	1	0	0	1	0	2	4	26	5	2	13	13	3	8
0101	1	1	9	2	5	0	4	1	7	2	105	3	31	4	42	27
0011	3	0	1	8	0	1	0	2	4	2	2	22	3	12	12	8
0100	2	6	6	1	22	2	3	10	6	16	47	4	209	9	25	153
0010	0	4	0	2	0	11	0	4	4	13	0	10	10	83	9	51
0001	1	1	3	3	2	4	13	14	2	1	23	13	8	7	148	116
0000	1	3	6	2	7	10	14	70	6	10	11	8	120	60	104	999

Table 4 shows the ADR of each lifestyle factor in the cases of Good vs Bad and Improved vs Worse. The factor with the largest value in both cases was “Amount of daily exercise” and it was recognized from the representative questions that MS is influenced by the frequency of exercise such as gymnastics, jogging and swimming, and by the degree of awareness on changes in the amount of exercise and weight, and on maintaining and improving health. The second one was “Smoking”. The number of smokers in MS was about 7 % higher than in non-MS, but smoking does not affect the judgement whether MS or non-MS in the Stratification [1]. From Good vs Bad “Regular eating habits” ranks third and it was recognized that the amount of meal had a great effect. From Improved vs Worse “Leisure time exercise” ranks second and it was found that frequency of free time, and exercises and sports interacting with nature have a powerful effect.

Table 4: Average difference rate (ADR) of lifestyle factors in the cases of Good vs Bad and Improved vs Worse

Good vs Bad		Improved vs Worse	
Lifestyle Factor	**ADR [%]**	**Lifestyle Factor**	**ADR [%]**
Amount of daily exercise	**12.90**	**Amount of daily exercise**	**14.91**
Smoking	**10.08**	**Leisure time exercise**	**8.77**
Regular eating habits	**6.42**	**Smoking**	**7.00**
Exercise awareness	**6.38**	**Exercise awareness**	**5.44**
Salt and fat intake	**5.74**	**Regular eating habits**	**5.43**
Leisure time exercise	**5.58**	**Salt and fat intake**	**3.25**
Nutritional balance of diet	**2.47**	**Alcohol intake**	**3.12**
Alcohol intake	**2.39**	**Nutritional balance of diet**	**2.61**
Others	**2.37**	**Others**	**1.36**
Healthy fulfillment	**2.09**	**Healthy fulfillment**	**-1.92**

4. SUMMARY AND FUTURE CONCERNS

In this study, we focused on the changes in MS status in two consecutive years regarding the characteristics of MS in questionnaire data in SHC. The characteristics of lifestyle could be extracted from the subjects in each group (Bad, Improved, Worse, and Good) in transitions between 4-bit health states.

In specific health guidance, it is expected that examinees will be aware of the existence of characteristic lifestyle factors, and this will accelerate guidance on MS prevention countermeasures.

In the future, by linking to more accurate health guidance, we would like to construct a health support system that utilizes big data for SHC and SHG.

ACKNOWLEDGMENTS

This work was supported by Grant-in-Aid for Scientific Research of Japan.

REFFERENCES

[1] Specific Health Checkups and Specific Health Guidance, The Health Service Bureau of the Ministry of Health, Labor and Welfare, 2018.

[2] Guide for smooth implementation of Specific Health checkups and Specific Health guidance (3rd edition), The Health Service Bureau of the Ministry of Health, Labor and Welfare, 2018.

[3] E. Ashizawa, S. Katano, A. Harada, *et al*., Exploring the link between standard lifestyle questionnaires administered during specific medical check-ups and incidence of metabolic syndrome in Chiba Prefecture. Japanese Journal of Public Health. Vol.61 No.4 pp.176-185, 2014. (in Japanese)

[4] Y. Miyauchi and H. Nishimura, Construction and Evaluation of Bayesian Networks Related to the Specific Health Checkup and Guidance on Metabolic Syndrome, Innovation in Medicine and Healthcare 2015 (Smart Innovation, Systems and Technologies, Vol.45) Y.W. Chen *et al*. (Eds.), pp. 183–193, Springer International Publishing, 2015.

[5] Y. Miyauchi, H. Nishimura, Y. Nakano, A Study of Bayesian Network Model Related to the Specific Health Checkup based on Lifestyle Factor Analysis. Transactions of Japan Society of Kansei Engineering 15(7), 693–701 (2016). (in Japanese)

[6] N. Hashimoto, Y. Miyauchi, H. Nishimura, Analysis of Distinctive Lifestyle Factors for Metabolic Syndrome Using the Specific Health Checkup Data, The 38th Joint Conference of Medical Informatics, 2018. (in Japanese)

[7] N. Hashimoto, Y. Miyauchi, H. Nishimura, Analysis of Lifestyle Habit Characterizing Metabolic Syndrome in the Specific Health Checkup, The 21st Conference of Japan Society of Kansei Engineering, 2019. (in Japanese)