

Identifying New Risk Factors of Diabetes Based Japan Census Data

PEI JIANG[†] 鈴木 裕之[‡] 小尾 高史[†]
 Pei Jiang Hiroyuki Suzuki Takashi Obi

1. Backgrounds

World Health Organization (WHO) reports that 422 million adults have diabetes mellitus (DM) around the world, which makes one of every eleven people a DM patient [1]. Some risk factors of diabetes have been identified: age [2], gender [2], obesity [3], smoking status [4], alcohol consumption [5], occupation [6], work stress [7], work years [6] and weekly work hours [8], etc. However, some unclear risk factors perhaps still exist due to the complexity of our modern society. Finding more risk factors of diabetes mellitus can help us preventing diabetes more efficiently. To identify new risk factors associated with DM, we used logistic analysis to analyze the Japanese life census data, which were collected in order to describe the daily life of the citizens by the Ministry of Healthcare, Labor and Welfare (MHLW) [9] in Japan. Because the census data are not designed for diabetes, some new risk factors of DM perchance can be found.

2. Methodology

Logistic analysis is one robust methods in medical and healthcare field, which can be used to identify risk factors of diseases and clarify the risk difference caused by each factor. However, for category data, logistic analysis cannot show the difference of each category. Therefore, Hayashi's quantification method 2 (qt2) [10] were used to classify the importance of identified factors. Moreover, we also used linear discrimination analysis (LDA) and linear regression to verify the associated risk factors of DM again. Meanwhile, Qt2 method also can show the connection of category factors and DM, which can help us clear the effects caused by each factor. Finally, new risk factors of DM were identified and verified in this study.

3. Results

After logistic analysis, LDA and Linear regression analysis, 15 risk factor of DM were identified, which is shown as Table 1 and Table 2. Especially, 5 new risk factors of diabetes were identified. The importance of factors was also identified using qt2 (Table 2). The relationship of each factor's category was also find which is shown as Fig 1.

3.1 Identified risk factors

After analysis, 15 risk factors of diabetes are identified, which were shown as Table 1 and Table2. Meanwhile, the results of qt2 and liner regression analysis (Table2) verified that identified risk factors in logistic analysis are risk factors of DM.

Similar as previous researches, we found that male is facing around 2.8 higher risk than female (Table1) and people who have obesity have over 6 times higher risk of diabetes than general people. Meanwhile, with age growing, the risk of diabetes become higher (OR=1.273).

Health investigation status (have or no investigation), health awareness level, public pension type in Japan, insurance type in Japan and room numbers were identified as risk factors of diabetes for the first time. We found that having health investigation will release the risk of diabetes in Japan. Meanwhile, Data of Table 2 show that new identified factors: health awareness level, public pension and insurance type have more significant connection of diabetes mellitus in Japan. Especially, the factor health awareness

Table1: Identified Factors by logistic analysis

Factors		P-value	Odds ratio of Diabetes
Smoking situation	4 situations	0.0324	Fig1
Number of family members	7 categories	0.0037	Fig1
Obesity	Have/no	< 0.001	6.098
Hyperlipidemia	Have/no	< 0.001	1.371
Hypertension	Have/no	< 0.001	1.260
Insurance type	5 types	0.003	Fig1
Public pension	4 categories	< 0.001	Fig1
Health awareness	5 levels	< 0.001	Fig1
Profession	12 kinds	0.0005	Fig1
work days	0~7days	0.0069	Fig1
Stress	Yes/no	0.002	1.035
Age	15~	< 0.001	1.273
Gender	Male/female	< 0.001	2.793
Health investigation	yes / no	0.089	0.899
Room numbers	9 categories	0.000	Fig1

Table 2: Identified factor by other methods

Category Factors	LDA factor contributions	qt2 factor importance	Linear regression	
			Coefficients	P-values
Smoking situation	0.065	0.61	0.006	0.021
Health investigation	-0.072	0.15	-0.013	0.006
Health awareness	0.219	1.01	0.018	0
Stress	-0.043	0.07	-0.006	0.134
Hypertension	-0.086	0.22	-0.015	0.003
Hyperlipidemia	0.115	0.29	0.026	0
Obesity	0.557	4.33	0.341	0
Profession	-0.062	0.74	-0.001	0.031
Weekly work day	0.001	0.34	-	-
Public pension type	-0.061	0.31	-0.005	0.043
Insurance type	-0.019	0.6	-	-
Gender	0.472	0.83	0.069	0
Age	0.594	2.95	0.016	0
Room numbers	0.046	0.4	0.002	0.039
Family members	0.024	0.33	-	-

[†] 東京工業大学 Tokyo Institute of Technology

[‡] 群馬大学 Gunma University

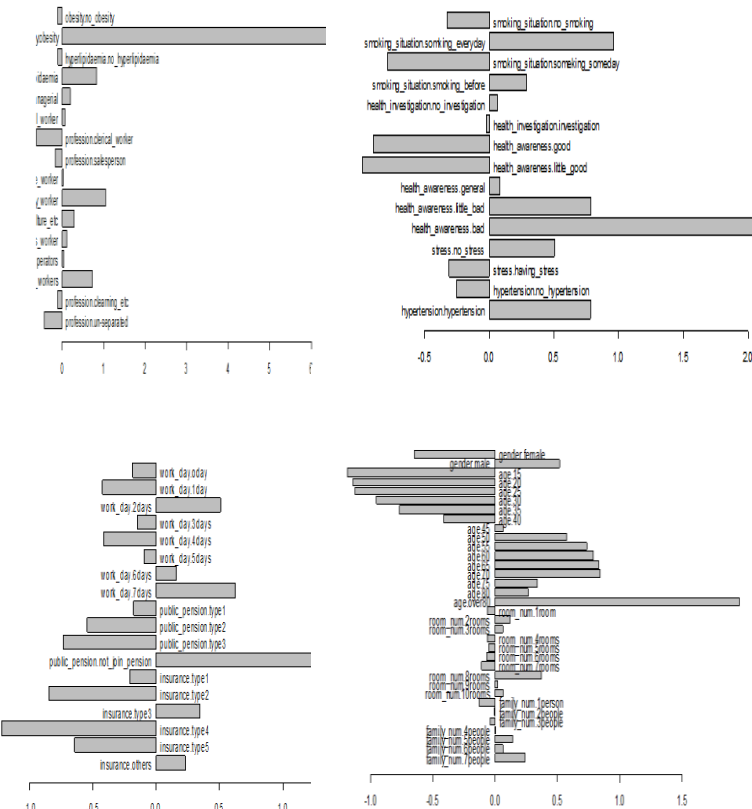


Figure 1. Category score of qt2

level, whose coefficient and qt2 factor importance both are bigger than smoking situation and stress. The higher qt2 factor importance of health awareness and qt2 category score of health awareness advises us that improving health awareness level perhaps can help to prevent the spreading diabetes in Japan. We also originally found that even though stress is risk factor of diabetes, the qt2 factor importance is the smallest. This indicates that risk factor: stress doesn't increase the risk of diabetes as so much as other risk factors, even though stress is risk factors of diabetes.

3.2 Category score of risk factors

For factors whose value is category data, such as smoking situation, public pension type, insurance type and health awareness level, clarifying the different of each category can help us take efficient measure to preventing diabetes.

Category scores of qt2 show the relationship among category data and diabetes. The highest category score of having obesity (Fig 1) indicates the serious risk caused by obesity. While smoking everyday will leads higher risk of diabetes compared with other kind of smoking situation. Considering health awareness level, people have general or bad level health awareness, they will have higher possibility of having diabetes. Similarly, people without joining public pension, they are facing higher risk of DM than people who joining public pension in Japan. Data of Fig 1 also shows that Japanese who are over 45 years old, they starting facing the risk of diabetes, especially for people who is over 80 years old. For Japanese who are over 80 years old, their category score of qt2 is over 2 times higher than other ages people. Meanwhile, we find

that people whose family have over 8 rooms and people whose family have over 4 members, they start facing the risk of diabetes. However, because of Japan's special public pension and insurance systems, continual studies are demanded to clarify the relationship among public pension, insurance, family members, ages, and diabetes. Nevertheless, the qt2 category score help us clarify the risk of diabetes for different people in Japan.

4. Conclusion

In this study, besides identified risk factors in previous research, 5 new risk factors of diabetes were identified: health investigation, health awareness level, public pension, insurance type and room numbers. Especially health awareness level, it has more deep relationship with diabetes than gender, smoking situation, hyperlipidemia and hypertension. The new identified factors paved the way to preventing diabetes in Japan more efficiently. Certainly, there are limits in our study. We haven't clarified the deep reasons of different effects caused by risk factors. Meanwhile, our analysis results only certificate the diabetes risk factors in Japan. In other country, the risk factors of diabetes may be different. However, our study still indicates the probability of identifying new risk factors of diabetes using new and complex data. Our analysis also widens the research of data science.

Acknowledgements

We would like to thank MHLW for supplying the necessary data for our study.

References

- [1] WHO. Global report on diabetes . [Online] 2016. [Cited: 07 29, 2019.] <https://apps.who.int/iris/handle/10665/204871?locale-attribute=en&>
- [2] Reversed gender distribution of diabetes in Northern Canada. Y.Chen, S.Wright, R.Westfal. s.l. : Public Health, 2009, Vol. 123, p. 782~786.
- [3] The magnitude of association between overweight and obesity and the risk of diabetes: A meta-analysis of prospective cohort studies. Asnawi Abdullah, Anna Peeters, Maximilian de Courten, Johannes Stoelewind. s.l. : Diabetes Research and Clinical Practice, 2010, Vol. 89, pp. 309-319.
- [4] Smoking and mortality among women with type 2 diabetes: The Nurses' Health Study cohort. AL-Delaimy WK, Willett WC, Manson JE, Speizer FE, Hu FB. s.l. : Diabetes care, 2001, Vol. 24(12), pp. 2043-8.
- [5] Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analysis. Vanesa, Bellou, Lazaros Belbasis, Loanna Tzoulaki, Evangelos Evangelou. s.l. : Plos One, 2018.
- [6] Overtime, psychosocial working condations, and occurrence of non-insulin dependent diabetes mellitus in Japanese men. Norito Kawakami, Shunichi Araki, Naoyoshi Takatsuka, Hiroyuki Shimizu, Hiroshi Ishibashi. s.l. : Epidemiol Communijty Health, 1999, Vol. 53, pp. 359-363.
- [7] Work stress, Sense of coherence and risk of type 2 diabetes in a prospective study of middle-aged Swedish men and women. Eriksson AK, van denDonk M, Hilding A, Ostenson CG. s.l. : Diabetes care, 2013, Vol. 36, pp. 2683-89.
- [8] Work-related stress and type2 diabetes: systematic review and meta-analysis. Cosgrove MP, Sargeant LA, Caleyachetty R, Griffin SJ. s.l. : Occup Med, 2010, Vol. 62, pp. 167-173.
- [9] Ministry of Health, Labor and Welfare. [Online] <http://www.mhlw.go.jp/toukei/itaku/tokumei.html>.
- [10] On the prediction of phenomena from qualitative data and the quantification of qualitative data from hte mathematico-statistical point of view. Haysshi, Chikio. s.l. : Annals of the Institute of Statistical Mathematics, 1951, Vol. 3, pp. 69-98