

Original Article



Artificial Intelligence Approaches to Social Determinants of Cognitive Impairment and Its Associated Conditions

Kwang-Sig Lee ,¹ Kun Woo Park ²

¹AI Center, Korea University College of Medicine, Seoul, Korea

²Department of Neurology, Korea University College of Medicine, Seoul, Korea

OPEN ACCESS

Received: May 12, 2020

Revised: Aug 25, 2020

Accepted: Sep 6, 2020

Correspondence to

Kun Woo Park

Department of Neurology, Korea University College of Medicine, 73 Incheon-ro, Seongbuk-gu, Seoul 02841, Korea.
E-mail: kunu@korea.ac.kr

© 2020 Korean Dementia Association

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Kwang-Sig Lee

<https://orcid.org/0000-0002-0576-0098>

Kun Woo Park

<https://orcid.org/0000-0001-6526-614X>

Funding

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2018-0-01405) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

Conflict of Interest

The authors have no potential conflicts of interest.

ABSTRACT

Background and Purpose: This study uses an artificial-intelligence model (recurrent neural network) for evaluating the following hypothesis: social determinants of disease association in a middle-aged or old population are different across gender and age groups. Here, the disease association indicates an association among cerebrovascular disease, hearing loss and cognitive impairment.

Methods: Data came from the Korean Longitudinal Study of Ageing (2014–2016), with 6,060 participants aged 53 years or more, that is, 2,556 men, 3,504 women, 3,640 aged 70 years or less (70–), 2,420 aged 71 years or more (71+). The disease association was divided into 8 categories: 1 category for having no disease, 3 categories for having 1, 3 categories for having 2, and 1 category for having 3. Variable importance, the effect of a variable on model performance, was used for finding important social determinants of the disease association in a particular gender/age group, and evaluating the hypothesis above.

Results: Based on variable importance from the recurrent neural network, important social determinants of the disease association were different across gender and age groups: 1) leisure activity for men; 2) parents alive, income and economic activity for women; 3) children alive, education and family activity for 70–; and 4) brothers/sisters cohabiting, religious activity and leisure activity for 70+.

Conclusions: The findings of this study support the hypothesis, suggesting the development of new guidelines reflecting different social determinants of the disease association across gender and age groups.

Keywords: Cerebrovascular Disease; Hearing Loss; Cognitive Impairment; Social Determinant; Age; Gender

INTRODUCTION

Cerebrovascular disease, hearing loss and cognitive impairment are the major causes of disease burden on a global level.¹⁻⁶ Cerebrovascular disease was the third cause of death in South Korea (Korea hereafter) for Year 2016 (Y2016 hereafter) (45.8 per 100,000).¹ Otitis media ranked seventh among the leading causes of disease burden in the nation for Y2010, i.e., 294 disability-adjusted life years (DALYs) per 100,000.² And the disease burden of

Author Contributions

Conceptualization: Lee KS, Park KW; Data curation: Lee KS, Park KW; Formal analysis: Lee KS, Park KW; Funding acquisition: Lee KS, Park KW; Investigation: Lee KS, Park KW; Methodology: Lee KS, Park KW; Project administration: Lee KS, Park KW; Resources: Lee KS, Park KW; Software: Lee KS, Park KW; Supervision: Lee KS, Park KW; Validation: Lee KS, Park KW; Visualization: Lee KS, Park KW; Writing - original draft: Lee KS, Park KW; Writing - review & editing: Lee KS, Park KW.

dementia for the nation in Y2010 (274,849 DALYs) is predicted to almost triple by Y2050 (814,629 DALYs).³ This local pattern in Korea is consistent with its global counterpart. Stroke was the second cause of global deaths in Y2013, i.e., 12% (7 million) of 54 million deaths in the world.⁴ The global prevalence of hearing loss increased by 757% from 42 million in Y1985 to 360 million in Y2011.⁵ And the number of people with dementia in the world is expected to nearly double every 20 years, i.e., from 36 million in Y2010 to 66 million (or 115 million) in Y2030 (or Y2050).⁶

Then, is there a high degree of comorbidity (or association) among the 3 diseases above (i.e., cerebrovascular disease, hearing loss and cognitive impairment)? A previous study used an artificial-intelligence model (RNN, recurrent neural network) for evaluating 1) whether social determinants are major determinants of association among the 3 diseases in a middle-aged or old population (Hypothesis 1) and 2) whether association among the 3 diseases is very strong in the middle-aged or old (Hypothesis 2).⁷ The source of data for this research was the Korean Longitudinal Study of Ageing (2014–2016), with 6,060 participants aged 53 years or more. The dependent variable of this research, association among the 3 diseases, had 8 categories, that is, 1 category for having no disease, 3 categories for having 1, 3 categories for having 2, and 1 category for having 3. Variable importance from the RNN, an accuracy difference between a complete model and a model excluding a certain variable, was introduced for evaluating the 2 hypotheses of this research: 1) Hypothesis 1 was tested based on whether family support (such as parents/children alive), socioeconomic status (such as education/income) and social activity (such as friendship/leisure activity) in Y2014 are top-10 determinants of the association in Y2016; and 2) Hypothesis 2 was tested based on whether cerebrovascular disease, hearing loss and cognitive impairment in Y2014 are top-5 determinants of the association in Y2016 (this would be one approach to evaluate correlation among the 3 diseases over time). Here, the greater “accuracy decrease” can be considered to be the greater variable importance. According to variable importance from the RNN, cerebrovascular disease, cognitive impairment and hearing loss in Y2014 were top-3 determinants of the association in Y2016. Children alive, education, income and friendship activity in Y2014 were top-10 determinants of the association in Y2016. The findings of this research supported the 2 hypotheses, drawing a policy implication that the promotion of preventive measures, family support, socioeconomic status and friendship activity would be needed for managing association among the 3 diseases in a middle-aged or old population in Korea. These results were consistent with existing literature on social determinants of chronic diseases in a middle-aged or old population.^{8,10}

However, a previous study reported a sub-group variation in old Koreans in terms of the determinants of chronic diseases.¹¹ Data for this cross-sectional study came from the 2005 Korean National Health and Nutrition Examination Survey with 3609 participants aged 65–89. In this previous study, a sub-group variation was found between a positive linkage between physical disability and a chronic disease: for example, a positive relationship between physical disability and stroke was stronger 1) for men than women and 2) for those aged 70–79 than those aged 65–69. This line of research would help to develop the new guidelines across gender and age groups in Korea. But little literature is available and more effort is needed in this direction. In this context, this study tests the following hypotheses from the literature and discussion above:

Hypothesis 1: There exist gender and age-group differences regarding the importance of social determinants in association among cerebrovascular disease, hearing loss and cognitive impairment in a middle-aged or old population.

Hypothesis 2: There exist gender and age-group differences regarding the degree of association among cerebrovascular disease, hearing loss and cognitive impairment in a middle-aged or old population.

METHODS

The data source of this study was the Korean Longitudinal Study of Ageing (KLoSA) in Y2014 and Y2016 (See the previous study⁷ for more description of data). Data were publicly available and de-identified. The final sample of this study consisted of 6,060 subjects aged 53 or more, i.e., 2,556 men, 3,504 women, 3,640 aged 70 years or less (70-), 2,420 aged 71 years or more (71+). The KLoSA question on cerebrovascular disease in Y2014 and Y2016 was "Since the last survey, have you ever been diagnosed by a doctor as cerebrovascular disease? 1. Yes. 5. No." [C038]. The inquiry on hearing loss in Y2014 and Y2016 was "Do you experience difficulty in daily activity because of hearing loss? 1. Yes. 5. No." [C092]. Mini-Mental State Examination score [C401-C419 in the KLoSA inquiries] was recoded as no cognitive impairment (at least as high as a cutoff) vs. yes (lower than the cutoff) with the different cutoffs based on age, education and gender¹² (**Supplementary Table 1**). Disease association (association among cerebrovascular disease, hearing loss and cognitive impairment) in Y2016 was divided into 8 categories: 1) "0" for having no disease; 2) "1," "2" and "3" for having cerebrovascular disease only, hearing loss only and cognitive impairment only, respectively; 3) "4," "5," and "6" for having cerebrovascular disease & hearing loss, cerebrovascular disease & cognitive impairment and hearing loss & cognitive impairment, respectively; and "7" for having all the 3 diseases (See the previous study⁷ for more description of the variables).

Six machine learning approaches were applied for the prediction of the disease association: RNN, logistic regression, decision tree, naïve Bayes, random forest and support vector machine.^{7,13,19} The disease association in Y2016 served as the dependent variable of the models. Cerebrovascular disease, hearing loss and cognitive impairment in Y2014 and the demographic, socioeconomic and health-related factors in Y2014 served as the independent variables of the models. Data on 2,556/3,504 men/women were divided into training and validation sets with a 50:50 ratio. Likewise, data on 3,640/2,420 for 70-/71+ were divided into training and validation sets with a 50:50 ratio. The models were built (or trained) based on the training set with 1,278/1,752 observations for men/women (or 1,820/1,210 observations for 70-/70+) then the models trained were validated based on the validation set with 1,278/1,752 observations for men/women (or 1,820/1,210 observations for 70-/70+).

Accuracy, a ratio of correct predictions among 1,278/1,752 (or 1,820/1,210) observations, was introduced as a criterion for validating the models trained. Variable importance, the effect of a variable on model performance, was used for evaluating the 2 hypotheses: Hypothesis 1 was evaluated based on whether there exist gender and age-group differences regarding the importance rankings of family support, socioeconomic status and social activity in Y2014 as the determinants of the disease association in Y2016; and Hypothesis 2 was evaluated based on whether there exist gender and age-group differences regarding the importance rankings of cerebrovascular disease, hearing loss and cognitive impairment in Y2014 as the determinants of the disease association in Y2016. Here, the greater "accuracy decrease" leads to the greater variable importance. Python 3.52 was employed for the analysis on May 2020.

Ethics statement

Need for ethics approval and consent to participate was waived (data were publicly available and de-identified).

RESULTS

Supplementary Table 2 is a frequency table for participants' disease association and categorical attributes. Among the 4 sub-groups (2,556 men, 3,504 women, 3,640 for 70– and 2,420 for 71+) in Y2016, 552 (22%), 847 (24%), 514 (14%) and 885 (37%) were diagnosed as at least one of cerebrovascular disease, hearing loss and cognitive impairment, whereas 98 (4%), 83 (2%), 37 (1%) and 144 (6%) were characterized by the diagnosis of the 2 or 3 diseases, respectively. Among the men in Y2014, 160 (6%), 53 (2%) and 380 (15%) were diagnosed as cerebrovascular disease, hearing loss and cognitive impairment, while their respective female counterparts were 161 (5%), 48 (1%) and 677 (20%). Among the 70– participants in Y2014, 138 (4%), 25 (1%) and 366 (10%) were diagnosed as cerebrovascular disease, hearing loss and cognitive impairment, while their respective 71+ counterparts were 183 (7%), 76 (3%) and 691 (29%). Descriptive statistics for participants' continuous attributes are shown in **Supplementary Table 3**. All (or 75%) of the men in Y2014 were older than 53 (or 60) and it was also true for the women. The average ages of the 4 sub-groups (men, women, 70– and 71+) were 67.5, 68.1, 61.3 and 77.7, whereas their respective monthly incomes were \$1,941, \$774, \$1,596 and \$769.

Based on the results of this study, the accuracy was higher for the RNN than for logistic regression: 0.8224 vs. 0.7848 for men, 0.8059 vs. 0.7837 for women, 0.8835 vs. 0.8593 for 70–, 0.7053 vs. 0.6926 for 71+ (**Table 1**). The top-10 variables of the disease association in terms of variable importance from the RNN are presented for all participants and the sub-groups in **Table 2** and **Figs. 1-4**. The top-10 variables regarding Hypothesis 1 (or Hypothesis 2) were in orange (or sea) background in the table. Here, variable importance from the RNN was measured as an accuracy gap between a complete model (RNN-Full) and a model excluding a certain variable (e.g., RNN-Cognitive Impairment) (**Supplementary Table 4**). Firstly, cerebrovascular disease and cognitive impairment in Y2014 were top-5 determinants of the disease association in Y2016 for all sub-groups. These results would highlight the significance of preventive measures for cerebrovascular disease and cognitive impairment in a middle-aged or old population. However, hearing loss in Y2014 ranked in the top 3 for men but it was out of the top-10 list for the other sub-groups. This indicates that the degree of the disease association would be stronger for men than for the other sub-groups (This supports Hypothesis 1).

Secondly, the findings of this study suggest that important social determinants of the disease association would be different across gender and age groups. Leisure activity in Y2014 was a top-10 determinant of the disease association in Y2016 for men but it was not the case for

Table 1. Accuracy measures of machine learning methods

Model/sub-group	Men	Women	70–	71+
Logistic regression	0.7848	0.7848	0.8593	0.6926
Decision tree	0.6862	0.6862	0.7560	0.5463
Naive bayes	0.1471	0.1471	0.0637	0.0264
Random forest-1,000 trees	0.7786	0.7786	0.8549	0.6992
Support vector machine	0.7746	0.7746	0.8522	0.6397
Recurrent neural network	0.8224	0.8059	0.8835	0.7050

Social Determinants of Cognitive Impairment

Table 2. Top-10 variables for disease association in all participants and 4 sub-groups

Rank	All*	Men	Women	70–	71+
1	Cerebrovascular disease [†]	Cerebrovascular disease	Cerebrovascular disease	Cerebrovascular disease	Cognitive impairment
2	Cognitive impairment	Cognitive impairment	Cognitive impairment	Life satisfaction - overall	Cerebrovascular disease
3	Hearing loss	Hearing loss	Life satisfaction - overall	Activity - political	Age
4	Age	Body mass index	Age	Cognitive impairment	Activity - friend
5	# Children alive ^b	Activity - political	Drinking	Meeting - friend	# Brothers/sisters cohabiting
6	Education	Subjective health	Economic activity	Activity - family	Drinking
7	Income	Life satisfaction - overall	Parents alive	Subjective health	Activity - religious
8	Activity - friend	Activity - leisure	Income	# Children alive	Activity - political
9	Gender	Smoking	Activity - political	Life satisfaction - economic	Activity - leisure
10	Marriage	Age	Region	Education	# Children alive

*Reference: †; †Top-10 variables for disease association regarding Hypothesis 1 (or Hypothesis 2) are presented as orange (or sea) background. Variables in bold characters ranked among top 10 for sub-groups only (they were out of the top-10 list for all participants).

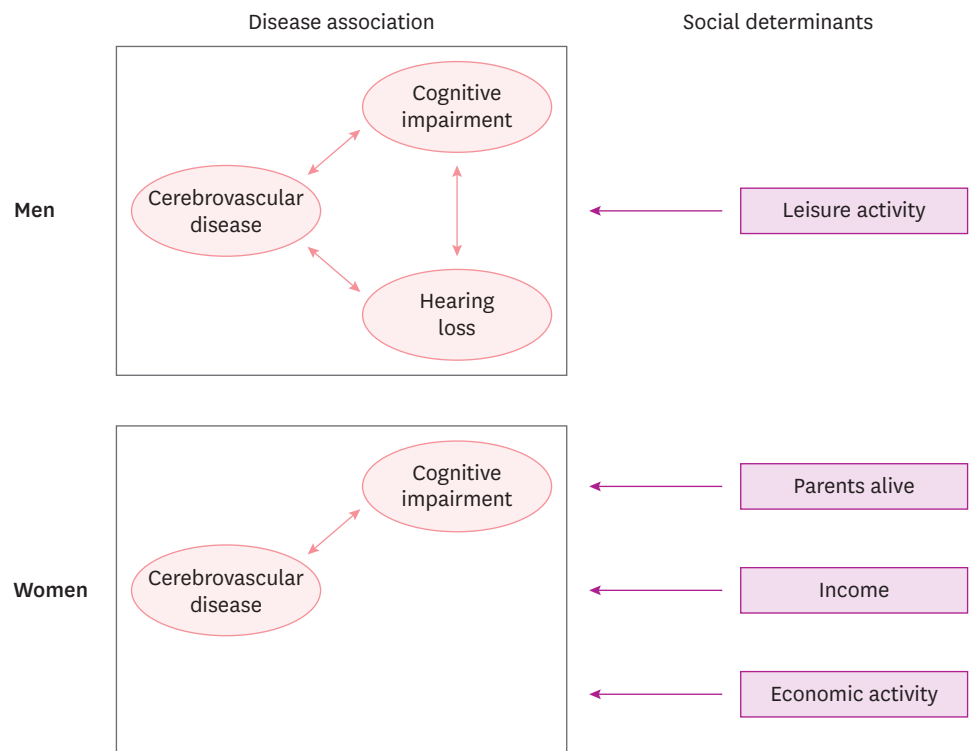


Fig. 1. Social determinants of disease association: gender variation.

women. Parents alive, income and economic activity in Y2014 ranked in the top 10 for women but not for men. Children alive, education and family activity in Y2014 were the top-10 variables for 70– but it was not true for 71+. And the importance rankings of brothers/sisters cohabiting, religious activity and leisure activity in Y2014 were higher than 10 for 71+ but not for 70–. In other words, the following social determinants were relatively more important for the following sub-group: 1) leisure activity for men; 2) parents alive, income and economic activity for women; 3) children alive, education and family activity for 70–; and 4) brothers/sisters cohabiting, religious activity and leisure activity for 70+. Finally, the logistic regression results (**Supplementary Table 5**) provide useful information about the sign and magnitude for the effect of the major determinant on the disease association. For example, the odds of cerebrovascular disease in Y2016 is 100 times as high for men with the disease in Y2014 as for men without the disease in Y2014. Similarly, the odds of cerebrovascular disease and one

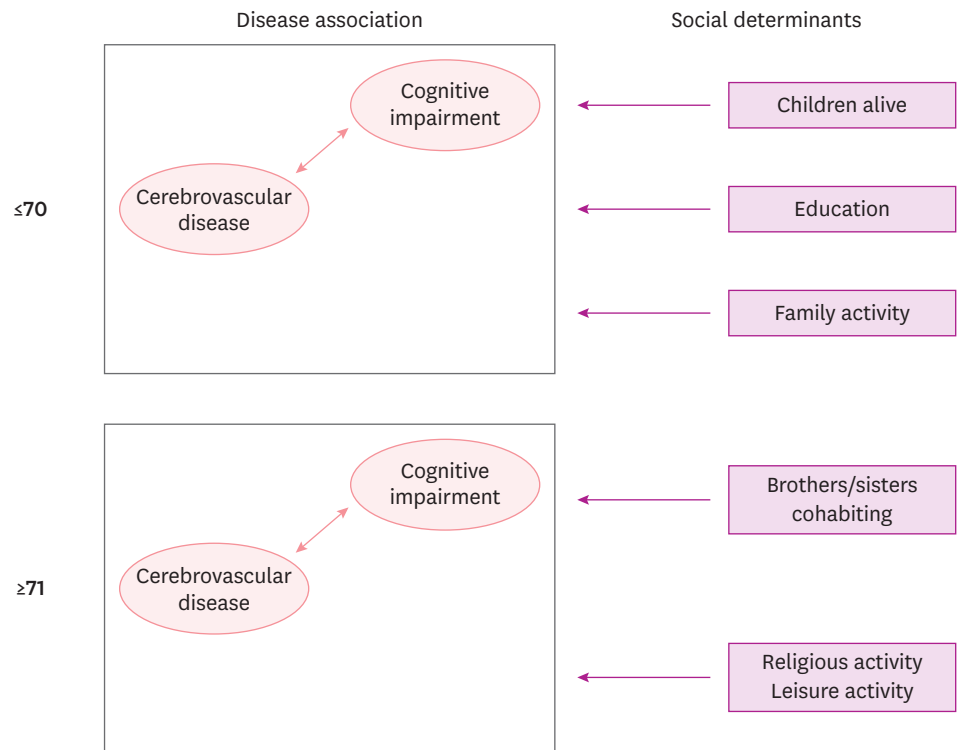


Fig. 2. Social determinants of disease association: age variation.

more disease in Y2016 is 100 times as high for men with cerebrovascular disease in Y2014 as for men without cerebrovascular disease in Y2014.

DISCUSSION

The existing literature on social determinants of chronic diseases drew a policy implication that the promotion of preventive measures, family support (e.g., children alive), socioeconomic status (e.g., education, income) and social activity (e.g., friendship activity) would be needed for managing association among chronic diseases such as cerebrovascular disease, hearing loss and cognitive impairment in a middle-aged or old population (“All” column in **Table 2**).⁷⁴⁰ In general, the results of this study agree with those of the previous studies, highlighting the importance of family support, socioeconomic status and social activity in association among cerebrovascular disease, hearing loss and cognitive impairment in a middle-aged or old population. However, the findings of this study also register significant variations across gender and age groups regarding important social determinants of the disease association: 1) leisure activity for men; 2) parents alive, income and economic activity for women; 3) children alive, education and family activity for 70-; and 4) brothers/sisters cohabiting, religious activity and leisure activity for 70+ (**Table 2, Figs. 1-4**).

In this vein, this study presents the following policy implications for managing association among cerebrovascular disease, hearing loss and cognitive impairment in a middle-aged or old population. Firstly, promotional efforts for leisure activity request more attention as central health policy for middle-age or old men. Secondly, social policy needs to be modified for strengthening the economic activity and income of middle-aged or old women. Middle-

Social Determinants of Cognitive Impairment

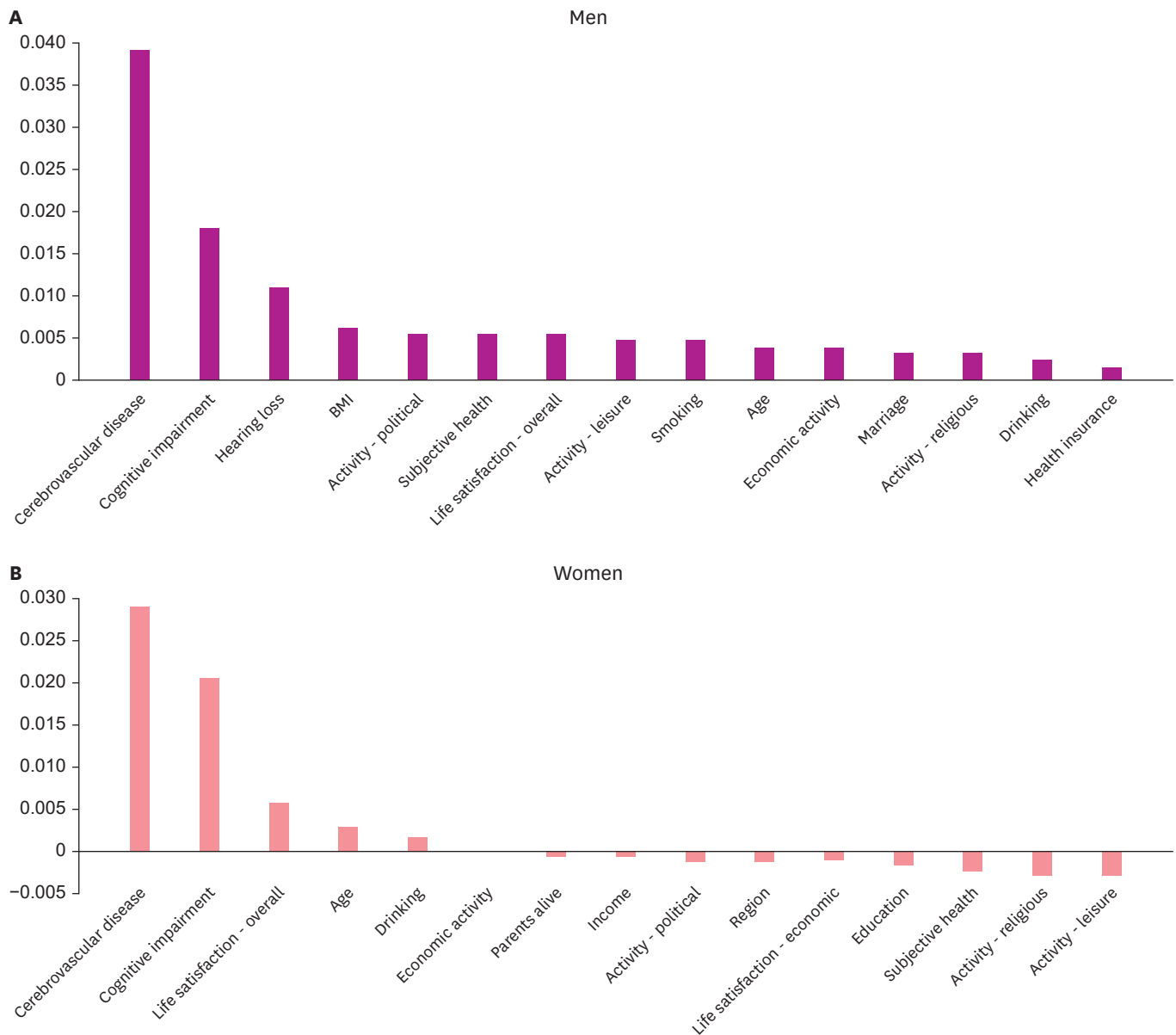


Fig. 3. Variable importance from the random forest: men vs. women. BMI: body mass index.

aged or old women are more likely to be in poverty than their male counterparts in a nation like Korea²⁰ and social policy needs to be revised so that it can reduce this disparity. Thirdly, it would be necessary to encourage brothers/sisters cohabiting and religious activity among those aged 71 or more. Different types of family support (e.g., parents/children alive vs. brothers/sisters cohabiting) are expected to have different effects on association among chronic diseases across gender and age groups but little research has been available on this topic. This study would be a good starting point for further research. In addition, existing literature focuses on a relationship between religious activity and mental health in a middle-aged or old population¹⁰ but the findings of this study suggest that this line of research can be extended to other chronic diseases including cerebrovascular disease, hearing loss and cognitive impairment.

Social Determinants of Cognitive Impairment

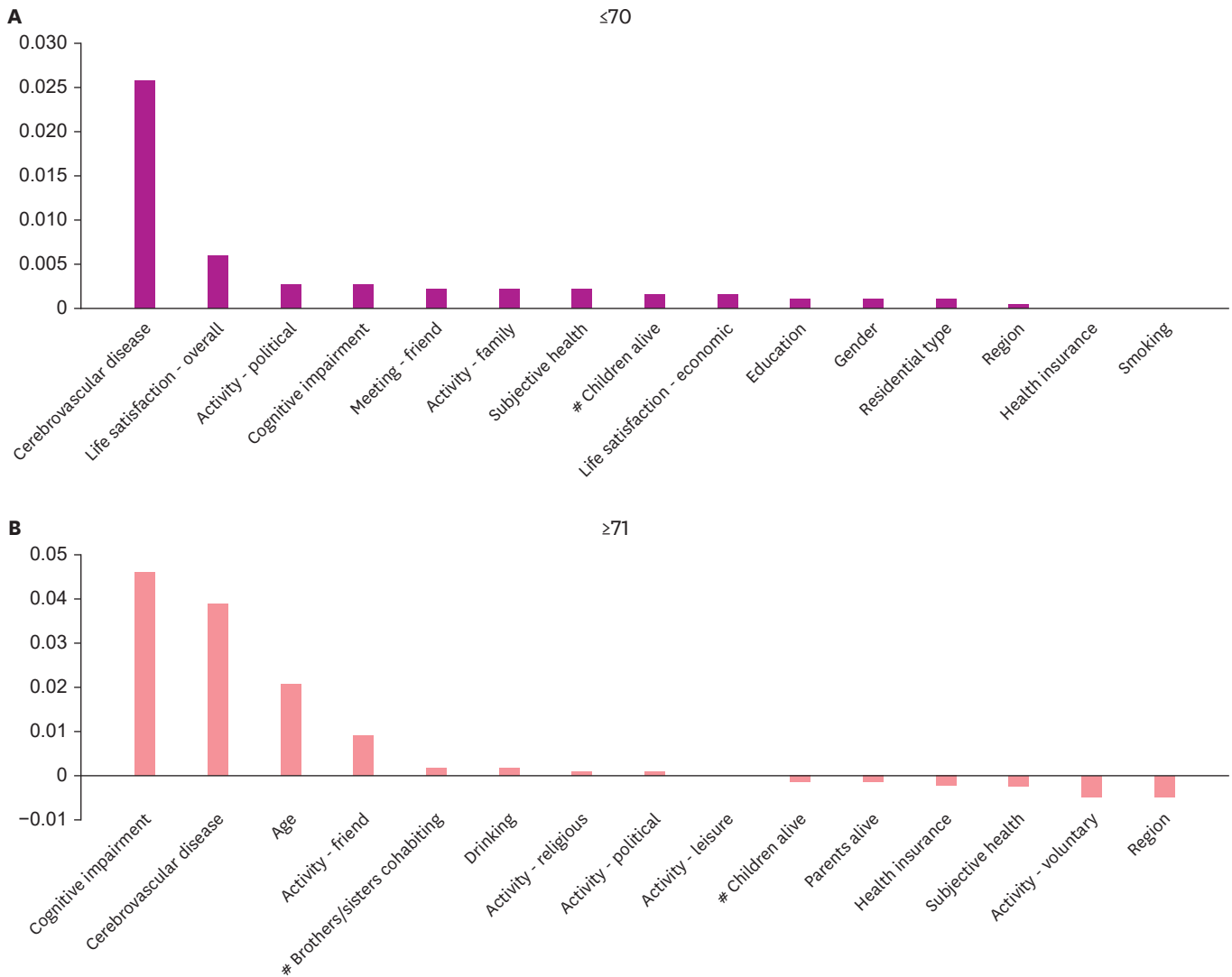


Fig. 4. Variable importance from the random forest: 70- vs. 71+.

This study still had some limitations. Firstly, this study used a weak form of the longitudinal design because of constraints on memory capacity. Employing data in all six waves with a strong form of the longitudinal design is expected to improve the accuracy of the RNN much more. Secondly, expanding this study to other chronic diseases and other determinants of association such as health utility usage might add a great contribution to this line of research. Thirdly, this study did not consider possible relationships or mediating effects among independent variables. Fourthly, combining gender and age in 4 sub-groups, e.g., 70- men, 70- women, 71+ men, 71+ women, would offer additional insight on the major determinants of the association among the 3 diseases. Fifthly, it was beyond the scope of this study to address and resolve the issue of class imbalance in this study: 2004 participants belonged to the category of NNN (cerebrovascular disease no, hearing loss no, cognitive impairment no) but only 3 belonged to YYY (cerebrovascular disease yes, hearing loss yes, cognitive impairment yes) in Y2016. Finally, comparing different types of family support regarding their effects on the association among the 3 diseases would be a good topic for further research.

In conclusion, social determinants of disease association in a middle-aged or old population are different across gender and age groups. For managing association among the 3 diseases, it would be necessary to develop the new guidelines reflecting different social determinants of the disease association across gender and age groups.

SUPPLEMENTARY MATERIALS

Supplementary Table 1

Mini-Mental State Examination score cutoff for cognitive impairment

[Click here to view](#)

Supplementary Table 2

Frequency tables for participants' disease association and categorical attributes

[Click here to view](#)

Supplementary Table 3

Descriptive statistics for participants' continuous attributes

[Click here to view](#)

Supplementary Table 4

Accuracy and VI from the RNN

[Click here to view](#)

Supplementary Table 5

Multinomial logistic regression results: odds ratio

[Click here to view](#)

REFERENCES

1. Statistics Korea. Year 2016 Statistics on Causes of Death in Korea. Sejong: Statistics Korea; 2017.
2. Lee KS, Park JH. Burden of disease in Korea during 2000–10. *J Public Health (Oxf)* 2014;36:225-234.
[PUBMED](#) | [CROSSREF](#)
3. Park JH, Eum JH, Bold B, Cheong HK. Burden of disease due to dementia in the elderly population of Korea: present and future. *BMC Public Health* 2013;13:293.
[PUBMED](#) | [CROSSREF](#)
4. Feigin VL, Norrving B, Mensah GA. Global burden of stroke. *Circ Res* 2017;120:439-448.
[PUBMED](#) | [CROSSREF](#)
5. Olusanya BO, Neumann KJ, Saunders JE. The global burden of disabling hearing impairment: a call to action. *Bull World Health Organ* 2014;92:367-373.
[PUBMED](#) | [CROSSREF](#)
6. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement* 2013;9:63-75.e2.
[PUBMED](#) | [CROSSREF](#)

7. Lee KS, Park KW. Social determinants of the association among cerebrovascular disease, hearing loss and cognitive impairment in a middle-aged or older population: recurrent neural network analysis of the Korean Longitudinal Study of Aging (2014–2016). *Geriatr Gerontol Int* 2019;19:711-716.
[PUBMED](#) | [CROSSREF](#)
8. Ministry of Health and Welfare (KR). Plan for Ageing Society and Population. Seoul: MHWK; 2011.
9. Lee YK, Jung KH, Oh YH, Yeom JH, Kim HA. Plans for Improving Leisure and Welfare Service among the Old in Korea. Seoul: Korea Institute for Health and Social Affairs; 2012.
10. Choi Y, Park EC, Kim JH, Yoo KB, Choi JW, Lee KS. A change in social activity and depression among Koreans aged 45 years and more: analysis of the Korean Longitudinal Study of Aging (2006–2010). *Int Psychogeriatr* 2015;27:629-637.
[PUBMED](#) | [CROSSREF](#)
11. Kim IH. Age and gender differences in the relation of chronic diseases to activity of daily living (ADL) disability for elderly South Koreans: based on representative data. *J Prev Med Public Health* 2011;44:32-40.
[PUBMED](#) | [CROSSREF](#)
12. Kim KW, Kim MH, Kim BJ, Kim JR, Kim TH, Moon SW, et al. The Standardization of the Dementia Diagnosis Instrument. Seongnam: Seoul National University Bundang Hospital; 2009.
13. Han J, Micheline K. Data Mining: Concepts and Techniques. 2nd ed. San Francisco, CA: Elsevier; 2006.
14. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc* 2017;24:361-370.
[PUBMED](#)
15. Pham T, Tran T, Phung D, Venkatesh S. Predicting healthcare trajectories from medical records: a deep learning approach. *J Biomed Inform* 2017;69:218-229.
[PUBMED](#) | [CROSSREF](#)
16. Suo Q, Ma F, Canino G, Gao J, Zhang A, Veltri P, et al. A multi-task framework for monitoring health conditions via attention-based recurrent neural networks. *AMIA Annu Symp Proc* 2018;2017:1665-1674.
[PUBMED](#)
17. Wang T, Qiu RG, Yu M. Predictive modeling of the progression of Alzheimer's disease with recurrent neural networks. *Sci Rep* 2018;8:9161.
[PUBMED](#) | [CROSSREF](#)
18. Sousa RT, Pereira LA, Soares AS. Predicting diabetes disease evolution using financial records and recurrent neural networks. *arXiv*. Forthcoming 2018.
19. Liu J, Zhang Z, Razavian N. Deep HER: chronic disease prediction using medical notes. *arXiv*. Forthcoming 2018.
20. Yoon HS. Korea: balancing economic growth and social protection for older adults. *Gerontologist* 2013;53:361-368.
[PUBMED](#) | [CROSSREF](#)