Preprints.org

Article

Not peer-reviewed version

Explainable Artificial Intelligence on Social Support, Poor Self-Rated Health and Dementia

Hye Gyeong Jeong , Chae-Won Kim , Hyuntae Park * , Kwang-Sig Lee *

Posted Date: 28 September 2023

doi: 10.20944/preprints202309.2007.v1

Keywords: self-rated health; dementia; family support; friendship activity; artificial intelligence

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Explainable Artificial Intelligence on Social Support, Poor Self-Rated Health and Dementia

Hye Gyeong Jeong [1,†], Chae-Won Kim [2,3,4,†], Hyuntae Park [1,*] and Kwang-Sig Lee [2,*]

[1] Obstetrics & Gynecology, Korea University Anam Hospital, Seoul 02841, Korea: psyche622@naver.com (H.G.J.)
[2] AI Center, Korea University Anam Hospital, Seoul 02841, Korea: kcherry@korea.ac.kr (C.-W.K.)
[3] School of Environmental Health Sciences, Korea University Anam Hospital, Seoul 02841, Korea
[4] Department of Statistics, Korea University College of Political Science and Economics, Seoul 02841, Korea
* Correspondence: cyberpelvis@gmail.com (H.P.); ecophy@hanmail.net (K.-S.L.)
† Hye Gyeong Jeong and Chae-Won Kim equally contributed for this study.

Abstract: This study uses explainable artificial intelligence to analyze major predictors of poor self-rated health, dementia and their comorbidity, which include various aspects of social support. Data came from the Korean Longitudinal Study of Ageing (2016-2018), with 5527 participants aged 56 or more. The dependent variables were: (1) poor self-rated health (SRH) with 5 categories (very good, good, middle, poor, very poor); (2) dementia (no, yes); and (3) their comorbidity with 4 categories (SRH very good, good or middle, dementia no; SRH very good, good or middle, dementia yes; SRH poor or very poor, dementia no; SRH poor or very poor, dementia yes). The 49 demographic, socioeconomic and health-related predictors were included. The accuracy and area under the receiver operating characteristic curve (AUC) were compared for logistic regression and the random forest. Logistic regression and the random forest registered similar AUCs with the range of 0.80-0.87. Based on random forest variable importance, 18 out of top-20 predictors were identical for SRH, dementia and their comorbidity: previous SRH, life satisfaction - health, age, body mass index, income, chronic diseases, life satisfaction - economic, life satisfaction - overall, grandchildren alive, friendship meeting, brothers/sisters alive, class identity, children alive, social activity - friendship, education, children in weekly contact, religion and drinker. However, some variations were found among the three dependent variables, i.e., previous SRH for comorbidity (1st) and SRH (1st) compared to dementia (11th), grandchildren alive for SRH (6th) and dementia (4th) compared to comorbidity (9th), chronic diseases for comorbidity (6th) compared to SRH (10th) and dementia (10th). In terms of Shapley Additive Explanation (SHAP) values, the probability of dementia is expected to decrease by 21% in case a predictor grandchildren alive is included to the model. This study highlights the importance of social support for the prediction of poor self-rated health, dementia and their comorbidity. Specifically, the promotion of family support and friendship activity for elders would be vital for the prevention and management of their self-rated health and dementia.

Keywords: self-rated health; dementia; family support; friendship activity; artificial intelligence

1. Introduction

Dementia can be defined as "deterioration in cognitive function beyond usual consequences of biological ageing" [1]. It comes from various conditions having major or minor effects on the brain. Its prevalence is more than 55 million in the world now, whereas its new incidences are almost 10 million on the globe every year [1]. Its global prevalence is expected to reach 153 million for the year 2050, while its global female-to-male ratio is projected to stay stable between 1.69 in 2019 and 1.67 in 2050 [2]. This global trend agrees with its local counterpart in Korea. The prevalence of elders' dementia in the nation registered a growth from 5.9% to 7.3% (588,000) during 2015-2019 [3]. The direct, indirect and total cost of dementia in the nation expanded by 56%, 36% and 46% during 2015-2019, i.e., 1400 to 2191, 1485 to 2026, and 2885 to 4218 USD million, respectively [3].

On the other hand, self-rated health can be conceptualized as "the overall perception of one's own health" [4–9]. It is usually assessed in a single item, where participants rate their current health.

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

It is considered to be an effective indicator of health status on its own right but it is expected to affect other health outcomes such as mortality based on a variety of indirect channels, e.g., socioeconomic conditions, biological, physical and mental functions, as well as health-related behaviors [4–9]. For example, a previous study on self-rated health in Japan and Korea [9] reported positive effects from socioeconomic status, daily activity and physical exercise, as well as negative impacts from chronic disease, overweight/obesity and smoking. In a similar context, existing literature confirms a strong association between poor self-rated health and the risk of dementia [10–14].

However, little study has been done on social support, dementia and self-rated health. A previous study [15] made a rare contribution for this line of research to use a machine learning model for testing: whether social determinants are major determinants of the association among cerebrovascular disease, hearing loss and cognitive impairment in a middle-aged or older population (hypothesis 1); and whether the association among the three diseases is very strong in the middle-aged or older population (hypothesis 2). Data came from the Korean Longitudinal Study of Aging (2014-2016), with 6060 participants aged ≥ 53 years. The findings of this study supported the two hypotheses, highlighting the importance of preventive measures, family support, socioeconomic status and friendship activity for managing the three diseases. But this study did not examine whether there exists a variation among different diseases regarding the importance of social support. In this vein, this study evaluates the following hypotheses based on machine learning models:

Hypothesis 1. *The predictors of social support are major predictors of dementia, poor self-rated health and their comorbidity in the middle-aged or old*
Hypothesis 2. *There exist some variations among dementia, poor self-rated health and their comorbidity in the middle-aged or old in terms of the importance of social support*

2. Methods

2.1. Participants and Variables

The data source of this study was the Korean Longitudinal Study of Ageing (KLoSA) (2016-2018). This study did not require either the approval of the ethics committee or the informed consent of human subjects given that (1) data were publicly available (https://survey.keis.or.kr/eng/klosa/klosa01.jsp) and (2) data were de-identified (patient anonymity was preserved). This study used the recoded version of data for immediate analysis ("light version"). But there were still missing values in some independent variables hence they were replaced with their median values (median imputation). The final sample of this study consisted of 5961 subjects aged 56 or more. The dependent variables were: (1) poor subjective or self-rated health (SRH) with 5 categories (very good, good, middle, poor, very poor); (2) dementia (no, yes); and (3) their comorbidity with 4 categories (SRH very good, good or middle, dementia no; SRH very good, good or middle, dementia yes; SRH poor or very poor, dementia no; SRH poor or very poor, dementia yes). The KLoSA questions on dementia and poor self-rated health in 2016 and 2018 were "Since the last survey, have you ever been diagnosed by a doctor dementia? 1. Yes. 5. No." and "How do you evaluate your health condition? 1. Very Good. 2. Good. 3. Middle. 4. Poor. 5. Very Poor.".

The independent variables were the following 49 predictors in 2016: (1) 5-category SRH (very good, good, middle, poor, very poor); (2) demographic information including age, gender, marital status (married, separated, divorced, widowed, unmarried); (3) socioeconomic conditions such as educational level (elementary school or below, junior high school, senior high school, college or above), income (monthly, normalized between 0 and 1), health insurance (Medicare, Medicaid), economic activity (employed, unemployed); (4) social support, that is, friendship meeting, social activities (religious, friendship, leisure, association), grandchildren (alive, raised, taken care), children (alive, cohabiting, meeting, contact, proximity), brothers/sisters (alive, cohabiting), parents (alive, cohabiting, meeting, contact, proximity); (5) health-related information, i.e., body mass index, smoker (non, former, current), drinker (non, former, current); and (6) other determinants including region (big urban, small urban, rural), religion (non, Protestant, Catholic, Buddhist, Won-Buddhist,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

other), residential type (apartment, other), subjective class (high-A, high-B, middle-A, middle-B, low-A, low-B), life satisfaction - health (0-100), life satisfaction - economic (0-100) and life satisfaction - overall (0-100).

2.2. Analysis

Logistic regression and the random forest were compared for the prediction of dementia, SRH and their comorbidity. For the random forest, the number of trees was 100, the criterion of split was GINI and the max of depth was not determined. Data on 5961 participants were divided into training and validation sets with a 75:25 ratio (4471 vs. 1490 observations). Criteria for validating the models trained were (1) accuracy, a ratio of correct predictions among 1490 observations, and (2) the area under the receiver-operating-characteristic curve (AUC), the plot of sensitivity over 1 - specificity. Here, the AUC measures how much sensitivity can be secured in case the threshold of sensitivity increases from 0 to 1 (that is, specificity increases from 0 to 1). Variable importance from the random forest, a GINI gap between a complete model and a model excluding a certain variable, was used for testing the two hypotheses of this study. The evaluation of Hypothesis 1 was based on whether the predictors of social support are top-20 predictors of dementia, SRH and their comorbidity across board. The evaluation of Hypothesis 2 was based on whether some variations are found among dementia, SRH and their comorbidity in terms of the variable importance ranking of social support. Finally, the Shapley Additive Explanation (SHAP) values were calculated to analyze the direction of association between a major predictor and its dependent variable in the model (random forest). The SHAP value of a particular predictor for a particular observation measures a difference between what the model (the random forest) predicts for the probability of the dependent variable for the observation with and without the predictor (https://github.com/shap/shap). Python 3.52 (Centrum voor Wiskunde en Informatica, Amsterdam, Netherlands) was employed for the analysis on January 2023.

3. Results

Descriptive statistics for participants' categorical and continuous variables are shown in Tables 1 and 2, respectively. The proportions of those with SRH poor/very poor and dementia in 2018 were 28.7% (1712) and 1.5% (89), respectively. The proportion of those with either or both disease in 2018 was 28.8% (1717). The proportion of those with SRH poor/very poor in 2016 was 26.8% (1598). The average numbers of family members or the average frequencies of social activity in a month were 4 (grandchildren alive), 3 (friendship meeting), 3 (brothers/sisters alive), 3 (children alive), 2 (social activity - friendship) and 2 (children in weekly contact). In the case of median imputation for missing values, logistic regression and the random forest registered similar AUCs with the range of 0.80-0.87 (Table 3). These performance measures were higher than those in the case of piece-wise deletion for missing values, i.e., 0.27-0.84.

Based on random forest variable importance (Table 4), 18 out of top-20 predictors were identical for SRH, dementia and their comorbidity: previous SRH, life satisfaction - health, age, body mass index, income, chronic diseases, life satisfaction - economic, life satisfaction - overall, grandchildren alive, friendship meeting, brothers/sisters alive, class identity, children alive, social activity - friendship, education, children in weekly contact, religion and drinker. This result supports Hypothesis 1: The predictors of social support are major predictors of dementia, poor self-rated health and their comorbidity in the middle-aged or old. However, some variations were found among the three dependent variables, i.e., previous SRH for comorbidity (1st) and SRH (1st) compared to dementia (11th), grandchildren alive for SRH (6th) and dementia (4th) compared to comorbidity (9th), chronic diseases for comorbidity (6th) compared to SRH (10th) and dementia (10th). This finding supports Hypothesis 2: There exist some variations among dementia, poor self-rated health and their comorbidity in the middle-aged or old in terms of the importance of social support. In terms of SHAP, the probability of dementia is expected to decrease by 21% in case a predictor grandchildren alive is included to the model (Table 5 and Figure 1). There exists a strong negative association between the predictor grandchildren alive and the dependent variable dementia.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

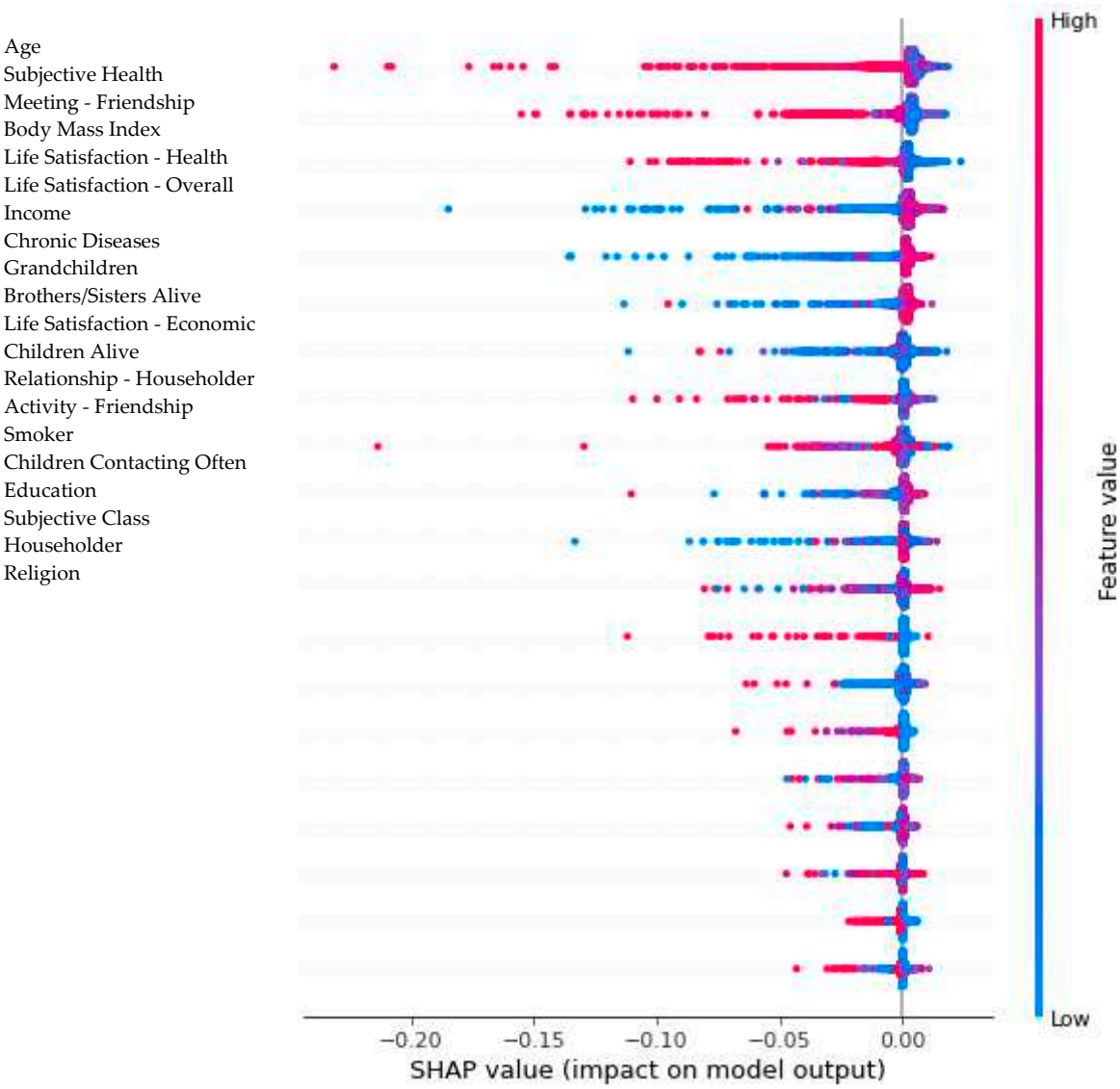

Figure 1. Random Forest SHAP Dependence Plots for Dementia. Legend: The SHAP value of a particular predictor for a particular observation measures a difference between what the model (the random forest) predicts for the probability of the dependent variable for the observation with and without the predictor. In terms of SHAP, for example, the probability of dementia is expected to decrease by 21% in case a predictor grandchildren alive is included to the model.

Table 1. Descriptive Statistics for Participants' Categorical Variables for Year 2018/2016.

Variable	***Count***	***Percentage (%)***
Dependent Variables (in 2018)		
Comorbidity		
NN[a]	4244	71.2
YN[b]	84	27.3
NY[c]	5	0.1
YY[d]	1628	1.4
Subjective/Self-Rated Health		
Very Good	81	1.4
Good	1471	24.7
Middle (Neither Good nor Poor)	2697	45.2
Poor	1332	22.3

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

Very Poor	380	6.4
Dementia		
Yes	89	1.5
No	5872	98.5
Subjective/Self-Rated Health (in 2016 Hereafter)		
Very Good	56 1587	0.9
Good	1587	26.2
Middle (Neither Good nor Poor)	2720	45.6
Poor	1307	21.9
Very Poor	291	4.9
Householder		
Yes	3378	56.7
No	2583	43.3
Relationship - Householder		
Spouse	5553	93.2
Parents	32	0.5
Children not Married	75	1.3
Children Married	276	4.6
Brothers/Sisters	9	0.2
Grandchildren	1	0
Grandparents	0	0
Other	15	0.3
Gender		
Male	2489	41.8
Female	3472	58.2
Marriage		
Married	4536	76.1
Separated	35	0.6
Divorced	121	2
Widowed	1228	20.6
Unmarried	41	0.7
Grandchildren under Care, Aged Less than 10		
Yes	242	4.1
No	5719	95.9
Grandchildren under Care, Aged Less than 10 (Last Year)		
Yes	66	1.1
No	5895	98.9
Parents Alive		
Father & Mother	230	3.9
Father	93	1.6
Mother	886	14.9
None	4752	79.7

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

Parents Cohabiting		
Yes	170	2.9
No	5791	97.1
Father Not Cohabiting		
Not Cohabiting with Other Children	5921	99.3
Cohabiting with Other Children	38	0.6
Other	2	0
Mother Not Cohabiting		
Not Cohabiting with Other Children	5683	95.3
Cohabiting with Other Children	227	3.8
Other	51	0.9
Health Insurance		
Health Insurance	5673	95.2
Medicare	288	4.8
Economic Activity		
Employed	2210	37.1
Unemployed	3751	63
Religion		
Non	3412	57.2
Protestant	1011	17
Catholic	414	6.9
Buddhist	1082	18.2
Won-Buddhist	13	0.2
Other	29	0.5
Drinker		
Current	1931	32.4
Former	984	16.5
Non	3046	51.1
Smoker		
Non	4132	69.3
Former	1210	20.3
Current	619	10.4
Education		
Elementary or Below	2559	42.9
Junior High	1059	17.8
Senior High	1737	29.1
College or Above	606	10.2

[a] NN for Subjective Health very good, good or middle Dementia No
[b] YN for Subjective Health poor or very poor Dementia No
[c] NY for Subjective Health very good, good or middle Dementia Yes
[d] YY for Subjective Health poor or very poor Dementia Yes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

Table 2. Descriptive Statistics for Participants' Continuous Variables for Year 2016.

Variable	*Mean*	*SD*	*Min*	*25%*	*50%*	*75%*	*Max*
Age	69	9	55	61	68	76	100
Meeting - Friendship	3	3	0	1	3	4	10
Activity - Religious	1	3	0	0	0	0	16
Activity - Friendship	2	4	0	0	1	2	16
Activity - Leisure	0	2	0	0	0	0	16
Activity - Association	0	1	0	0	0	0	16
Body Mass Index	23	3	12	22	23	24	42
# Chronic Diseases	1	1	0	0	1	2	7
# Children Alive	3	1	0	2	3	4	9
Income (Last Year)	0	0	0	0	0	0	1
# Children Cohabiting	1	0	1	1	1	1	5
# Children Cohabiting, Single	1	0	1	1	1	1	4
# Children Living Nearby*	1	0	1	1	1	1	6
# Children Meeting Often**	1	0	1	1	1	1	6
# Children Contacting Often**	2	1	1	2	2	2	9
# Grandchildren	4	4	0	1	3	6	34
# Grandchildren Under Care	1	0	1	1	1	1	8
# Grandchildren Under Care (Last Year)	1	0	1	1	1	1	3
# Brothers/Sisters Alive	3	2	1	2	3	5	11
# Brothers/Sisters Cohabiting	1	0	1	1	1	1	5
# Cohabiting Months with Father	1	1	0	1	1	1	12
# Cohabiting Months with Mother	2	2	0	2	2	2	12
Distance to Father***	3	0	1	3	3	3	4
Monthly Frequency - Meeting Father	7	0	1	7	7	7	9
Monthly Frequency - Contacting Father	4	0	1	4	4	4	8
Distance to Mother***	3	0	1	3	3	3	4
Monthly Frequency - Meeting Mother	7	1	1	7	7	7	10
Monthly Frequency - Contacting Mother	4	1	1	4	4	4	10
Life Satisfaction - Health	59	19	0	50	60	70	100
Life Satisfaction - Economic	56	19	0	50	70	70	100
Life Satisfaction - Overall	63	16	0	50	70	70	100
Subjective Class (1-6 Scale)	4	1	1	4	4	5	6

Note: * Within 30-minute Distance by Public Transportation; ** Meeting Once or More Often a Week; *** Unit 30 Minutes: 1 30 Minutes, 2 60 Minutes, 3 90 Minutes, 4 120 Minutes.

Table 3. Model Performance.

Variable	Imputation				Deletion			
	Accuracy		*AUC*		*Accuracy*		*AUC*	
	LR	**RF**	**LR**	**RF**	**LR**	**RF**	**LR**	**RF**
Subjective Health	0.55	0.58	0.82	0.80	0.56	0.58	0.82	0.80
Dementia	0.98	0.98	0.87	0.83	0.99	0.99	0.30	0.27
Comorbidity	0.79	0.80	0.87	0.87	0.79	0.8	0.83	0.84

Note: LR Logistic Regression, RF Random Forest, AUC Area under the Receiver-Operating-Characteristic Curve.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

Table 4. Random Forest Variable Importance.

Comorbidity		Subjective Health		Dementia	
Subjective Health - Previous Period	0.1185	Subjective Health - Previous Period	0.0843	Body Mass Index	0.1013
Life Satisfaction - Health	0.0820	Body Mass Index	0.0738	Age	0.0903
Age	0.0746	Income	0.0732	Income	0.0847
Body Mass Index	0.0661	Age	0.0720	# Grandchildren	0.0646
Income	0.0645	Life Satisfaction - Health	0.0577	Life Satisfaction - Health	0.0490
# Chronic Diseases	0.0532	# Grandchildren	0.0456	Life Satisfaction - Economic	0.0484
Life Satisfaction - Economic	0.0497	Life Satisfaction - Economic	0.0454	Meeting - Friendship	0.0481
Life Satisfaction - Overall	0.0446	Meeting - Friendship	0.0416	Life Satisfaction - Overall	0.0469
# Grandchildren	0.0440	# Brothers/Sisters Alive	0.0409	# Children Alive	0.0461
Meeting - Friendship	0.0402	# Chronic Diseases	0.0403	# Chronic Diseases	0.0450
# Brothers/Sisters Alive	0.0334	Life Satisfaction - Overall	0.0388	Subjective Health - Previous Period	0.0433
Subjective Class	0.0316	Subjective Class	0.0327	# Brothers/Sisters Alive	0.0382
# Children Alive	0.0309	# Children Alive	0.0322	# Children Contacting	0.0299
Activity - Friendship	0.0282	Activity - Friendship	0.0315	Subjective Class	0.0280
Education	0.0264	Education	0.0265	Religion	0.0237
# Children Contacting	0.0233	# Children Contacting	0.0262	Activity - Friendship	0.0204
Religion	0.0206	Religion	0.0252	Drinker	0.0193
Drinker	0.0168	Drinker	0.0199	Relationship - Householder	0.0190
Economic Activity	0.0143	Smoker	0.0164	Smoker	0.0179
Marriage	0.0138	Marriage	0.0134	Education	0.0177

Table 5. Random Forest Shapley Additive Explanation (SHAP) Values.

Predictor/Dependent Variable	Comorbidity		Self-Rated Health		Dementia	
	Min	*Max*	*Min*	*Max*	*Min*	*Max*
# Chronic Diseases	-0.0062	0.0024	-0.1060	0.0811	-0.0964	0.0107
Subjective Health	-0.0079	0.0026	-0.1788	0.2099	-0.1336	0.0155
Religion	-0.0008	0.0008	-0.0172	0.0633	-0.0297	0.0063
Meeting - Friendship	-0.0017	0.0010	-0.0646	0.0441	-0.0966	0.0169
Activity - Religious	-0.0005	0.0001	-0.0320	0.0317	-0.0468	0.0042
Activity - Friendship	-0.0041	0.0017	-0.0472	0.0910	-0.0244	0.0077
Activity - Leisure	-0.0003	0.0000	-0.0278	0.0440	-0.0105	0.0019
Activity - Association	-0.0015	0.0031	-0.0363	0.0961	-0.1373	0.0014
Householder	-0.0008	0.0011	-0.0150	0.0148	-0.0219	0.0063
Relationship - Householder	-0.0006	0.0001	-0.0239	0.0233	-0.0509	0.0123
# Children Alive	-0.0016	0.0039	-0.0331	0.0378	-0.0690	0.0100
# Children Cohabiting	-0.0006	0.0001	-0.0136	0.0254	-0.0752	0.0031
# Children Cohabiting, Single	0.0000	0.0000	-0.0071	0.0282	-0.0094	0.0009
# Children Living Nearby	-0.0014	0.0017	-0.0214	0.0615	-0.0295	0.0039
# Children Meeting Often	-0.0008	0.0001	-0.0103	0.0608	-0.0440	0.0016
# Children Contacting Often	-0.0008	0.0011	-0.0435	0.0399	-0.0382	0.0048
# Grandchildren	-0.0011	0.0036	-0.0478	0.0424	-0.2100	0.0167
Grandchildren Under Care	-0.0003	0.0000	-0.0183	0.0180	-0.0109	0.0013
# Grandchildren Under Care	-0.0005	0.0000	-0.0099	0.0056	-0.0171	0.0020
Grandchildren Under Care (Last Year)	0.0000	0.0000	-0.0099	0.0239	-0.0003	0.0005
# Grandchildren Under Care (Last Year)	0.0000	0.0000	-0.0006	0.0052	-0.0002	0.0010
# Brothers/Sisters Alive	-0.0008	0.0005	-0.0271	0.0652	-0.1114	0.0077
# Brothers/Sisters Cohabiting	0.0000	0.0000	-0.0009	0.0031	0.0000	0.0000
Parents Alive	-0.0002	0.0002	-0.0248	0.0384	-0.0024	0.0023
# Cohabiting Months with Father	0.0000	0.0000	-0.0335	0.0129	-0.0001	0.0001
# Cohabiting Months with Mother	-0.0002	0.0001	-0.0124	0.0284	-0.0011	0.0010
Parents Cohabiting	-0.0002	0.0000	-0.0227	0.0183	-0.0003	0.0003
Father Cohabiting with Other Children	0.0000	0.0000	-0.0005	0.0137	0.0000	0.0002
Distance to Father	0.0000	0.0000	-0.0180	0.0362	0.0000	0.0000
Monthly Frequency - Meeting Father	0.0000	0.0000	-0.0257	0.0235	0.0000	0.0000
Monthly Frequency - Contacting Father	0.0000	0.0000	-0.0254	0.0233	-0.0001	0.0003
Mother Cohabiting with Other Children	0.0000	0.0000	-0.0230	0.0204	-0.0002	0.0008
Distance to Mother	-0.0001	0.0001	-0.0146	0.0330	-0.0002	0.0002
Monthly Frequency - Meeting Mother	-0.0002	0.0000	-0.0151	0.0694	-0.0006	0.0001
Monthly Frequency - Contacting Mother	-0.0004	0.0001	-0.0336	0.0501	-0.0003	0.0003
Education	-0.0022	0.0017	-0.0416	0.0753	-0.0463	0.0074

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

Gender	-0.0011	0.0007	-0.0130	0.0196	-0.0108	0.0058
Age	-0.0016	0.0028	-0.1215	0.1101	-0.2239	0.0176
Marriage	-0.0034	0.0011	-0.0434	0.0252	-0.0166	0.0092
Body Mass Index	-0.0015	0.0065	-0.0510	0.0817	-0.1325	0.0136
Smoker	-0.0014	0.0014	-0.0217	0.0376	-0.0318	0.0042
Drinker	-0.0010	0.0011	-0.0146	0.0353	-0.0321	0.0064
Health Insurance	-0.0012	0.0000	-0.0308	0.0224	-0.0264	0.0022
Economic Activity	-0.0018	0.0014	-0.0159	0.0375	-0.0101	0.0029
Income	-0.0011	0.0018	-0.0467	0.0961	-0.0827	0.0162
Life Satisfaction - Health	-0.0040	0.0017	-0.0948	0.0935	-0.0769	0.0101
Life Satisfaction - Economic	-0.0029	0.0009	-0.0699	0.0635	-0.1263	0.0125
Life Satisfaction - Overall	-0.0028	0.0042	-0.0345	0.0558	-0.0428	0.0074
Subjective Class	-0.0019	0.0037	-0.0636	0.0720	-0.0389	0.0087

4. Discussion

Existing literature reports inconsistent results on social support and dementia [16–18]. A retrospective cohort study adopted 355 elderly Americans and linear regression to find that a positive association exists between perceived emotional support and the Repeatable Battery for the Assessment of Neuropsychological Status score (cognitive function) in women only [16]. A retrospective cohort study, which employed 5852 elderly Koreans and Cox proportional hazard regression, reached a similar conclusion with its U.S. counterpart, i.e., emotional support has its protective effect on incident dementia among women only [17]. On the contrary, a prospective cohort study employed 14,088 elderly Japanese and Cox proportional hazard regression to discover that the protective effect of family support on incident dementia exists among men only [18]. The SHAP results in Table 5 of this study show that there exists, in general, a strong negative association between the predictor social support and the dependent variable dementia among all participants. In terms of SHAP, for example, the probability of dementia is expected to decrease by 21% in case a predictor grandchildren alive is included to the model. Artificial intelligence is a data-driven approach hence more examination is needed for more conclusive findings on social support and dementia.

On the other hand, existing literature reports that poor self-rated health results in the increased risk of mortality [19,20]. In a retrospective cross-sectional study with 2819 breast cancer patients in the United States, poor self-rated health was found to have a positive association with cancer mortality with the hazard ratio of 3.05 [19]. In a retrospective cohort study with 710 elderly men in Europe, poor self-rated health was also discovered to have a positive relationship with cancer mortality with the hazard ratio of 2.41 [20]. In other words, poor self-rated health requests due attention to its fatal role in the risks of mortality. Moreover, poor self-rated health in the previous period was the most important predictor for the comorbidity of dementia and poor self-rated health in the current period based on the findings of this study. Based on random forest variable importance (Table 4) in this study, 18 out of top-20 predictors were identical for SRH, dementia and their comorbidity, and six of these top predictors represented social support i.e., grandchildren alive, friendship meeting, brothers/sisters alive, children alive, social activity - friendship and children in weekly contact. Little literature is available and more investigation is needed on dynamic interactions among social support, poor self-rated health and dementia in a variety of study populations in different environments. The SHAP value can be considered to be the weighted average of all possible real-world scenarios [21] hence explainable artificial intelligence based on the SHAP including this study can be a very attractive option for this line research.

This study had some limitations. Firstly, this study had limited memory capacity and used a weak version of the longitudinal design with data in two waves. Poor self-rated health, dementia and their comorbidity in 2018 (Wave 7) served as the dependent variable of the models, while, the three diseases in 2016 (Wave 6) and the demographic, socioeconomic and health-related factors in 2016 (Wave 6) served as the independent variables of the models. Expanding memory capacity and employing a strong version of the longitudinal design with data in all seven waves would strengthen the performance of explainable artificial intelligence. Secondly, this study kept outliers to protect the sample size. Comparing different outlier management strategies would further the horizon of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

research on this topic. Finally, sub-group analysis would make a good contribution for this line of research, e.g., 55-64, 65-74 and 75+ in age.

In conclusion, this study highlights the importance of social support for the prediction of poor self-rated health, dementia and their comorbidity. Specifically, the promotion of family support and friendship activity for elders would be vital for the prevention and management of their self-rated health and dementia.

Author Contributions: H.G.J., C.-W.K., H.P. and K.-S.L. designed the study. H.G.J., C.-W.K., H.P. and K.-S.L. collected, analyzed, and interpreted the data. H.G.J., C.-W.K., H.P. and K.-S.L. wrote and reviewed the manuscript. All authors approved the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Health and Welfare of South Korea under Korea Health Industry Development Institute grants (No. HI21C1560; HI22C1302 (Korea Health Technology R&D Project)). The funder had no role in the design of the study, the collection, analysis and interpretation of the data and the writing of the manuscript.

Data Availability: The data used for this study are available from the Korean Longitudinal Study of Ageing (KLoSA) (https://survey.keis.or.kr/eng/klosa/klosa01.jsp).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Health Organization. Dementia. https://www.who.int/news-room/fact-sheets/detail/dementia (Accessed 27 December 2022; Updated 20 September 2022).
2. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022;7:e105-e125.
3. Shon C, Yoon H. Health-economic burden of dementia in South Korea. BMC Geriatr 2021;21:549-558.
4. Jylh¨a M. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. Soc Sci Med 2009;69;307-316.
5. Mavaddat N, Valderas JM, van der Linde R, Khaw KT, Kinmonth AL. Association of self-rated health with multimorbidity, chronic disease and psychosocial factors in a large middle-aged and older cohort from general practice: a cross-sectional study. BMC Fam Pract 2014;15:185.
6. Sargent-Cox K, Cherbuin N, Morris L, Butterworth P, Anstey KJ. The effect of health behavior change on self-rated health across the adult life course: a longitudinal cohort study. Prev Med 2014;58:75-80.
7. Harris SE, Hagenaars SP, Davies G, David Hill W, Liewald DCM, Ritchie SJ, Marioni RE; METASTROKE Consortium, International Consortium for Blood Pressure Genome-Wide Association Studies; International Consortium for Blood Pressure Genome-Wide Association Studies; CHARGE Consortium Aging and Longevity Group; CHARGE Consortium Cognitive Group; Sudlow CLM, Wardlaw JM, McIntosh AM, Gale CR, Deary IJ. Molecular genetic contributions to self-rated health. Int J Epidemiol 2017;46:994-1009.
8. Stephan Y, Sutin AR, Bayard S, Kriˇzan Z, Terracciano A. Personality and sleep quality: evidence from four prospective studies. Health Psychol 2018;37:271-281.
9. Park JH, Lee KS. Self-rated health and its determinants in Japan and South Korea. Public Health 2013;127:834-843.
10. Weisen SF, Frishman WH, Aronson MK, Wassertheil-Smoller S. Self-rated health assessment and development of both cardiovascular and dementing illnesses in an ambulatory elderly population: a report from the Bronx Longitudinal Aging Study. Heart Dis 1999;1:201-205.
11. Yip AG, Brayne C, Matthews FE; MRC Cognitive Function and Ageing Study. Risk factors for incident dementia in England and Wales: The Medical Research Council Cognitive Function and Ageing Study. A population-based nested case-control study. Age Ageing 2006;35:154-160.
12. Montlahuc C, Soumar'e A, Dufouil C, Berr C, Dartigues JF, Poncet M, Tzourio C, Alp´erovitch A. Self-rated health and risk of incident dementia: a community based elderly cohort, the 3C study. Neurology 2011;77:1457-1464.
13. St John P, Montgomery P. Does self-rated health predict dementia? J Geriatr Psychiatr Neurol 2013;26:41-50.
14. Aschwanden D, Aichele S, Ghisletta P, Terracciano A, Kliegel M, Sutin AR, Brown J, Allemand M. Predicting cognitive impairment and dementia: a machine learning approach. J Alzheimers Dis 2020;75:717-728.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2023 doi:10.20944/preprints202309.2007.v1

15. 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.
16. Pillemer SC, Holtzer R. The differential relationships of dimensions of perceived social support with cognitive function among older adults. Aging Ment Health 2016;20:727-735.
17. Oh DJ, Yang HW, Kim TH, Kwak KP, Kim BJ, Kim SG, Kim JL, Moon SW, Park JH, Ryu SH, Youn JC, Lee DY, Lee DW, Lee SB, Lee JJ, Jhoo JH, Bae JB, Han JW, Kim KW. Association of low emotional and tangible support with risk of dementia among adults 60 years and older in South Korea. JAMA Netw Open 2022;5:e2226260.
18. Murata C, Saito T, Saito M, Kondo K. The Association between social support and incident dementia: a 10-year follow-up study in Japan. Int J Environ Res Public Health 2019;16:239.
19. Adeyemi OJ, Gill TL, Paul R, Huber LB. Evaluating the association of self-reported psychological distress and self-rated health on survival times among women with breast cancer in the U.S. PLoS One 2021;16:e0260481.
20. Giltay EJ, Vollaard AM, Kromhout D. Self-rated health and physician-rated health as independent predictors of mortality in elderly men. Age Ageing 2012;41:165-171.
21. Panda C, Mishra AK, Dash AK, Nawab H. Predicting and explaining severity of road incident using artificial intelligence, SHAP and feature analysis. International Journal of Crashworthiness 2022; https://doi.org/10.1080/13588265.2022.2074643.