

After submitting the paper,
we further analyzed and add new
result and discussion.

27

Feature Selection by Random Forest

Factor	Mean Decrease Accuracy	Mean Decrease Gini
Hospitalization (0,1-14,15-30, >30 days)	264.38	565.38
Care Needs Level (1/2/3/4/5)	49.76	75.16
Alzheimer's Disease	36.6	22.57
New Diagnosis of Dementia	22.49	15.31
Facility Care Service Use	13.09	26.38
Sex	4.26	13.15
DM	2.91	10.89
Age (65-74,75-84,85-94,>94)	2.49	26.77
Medical Area (Tango, CyuTan, NanTan, Kyoto.Otokuni, YamashiroKita, YamashiroMinami)	-7.88	36.07

Number of tree = 500

Number of variables = 3

28

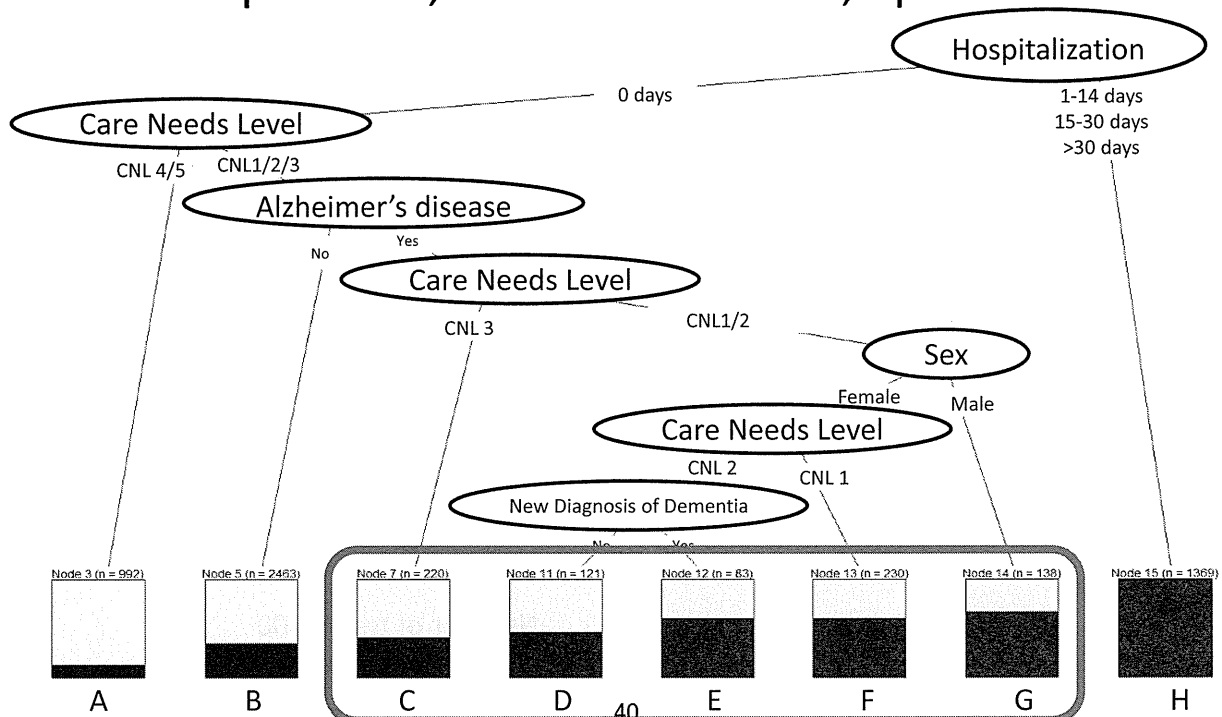
Tree Rules

n=5,616

Node Number	Rules	Prob	Occurrence
15 (H)	IF Hospitalization Period = "30,1~14,15~30 days then LTCI expenditure increase equal to 50% or above in one year	0.99	1,369
14 (G)	IF Hospitalization Period = "0" AND Alzheimer's disease = "Y" AND CNL = "1, 2" AND Sex = "Male" then LTCI expenditure increase equal to 50% or above in one year	0.67	138
13 (F)	IF Hospitalization Period = "0" AND Alzheimer's disease = "Y" AND CNL = "1/2" AND Sex = "Female" then LTCI expenditure increase equal to 50% or above in one year	0.61	230
12 (E)	IF Hospitalization Period = "0" AND CNL = "2" AND Alzheimer's disease = "Y" AND Sex = "Female" AND New Dementia = "Y" then LTCI expenditure increase equal to 50% or above in one year	0.60	83
11 (D)	IF Hospitalization Period = "0" AND CNL = "2" AND Alzheimer's disease = "Y" AND Sex = "Female" AND New Dementia = "N" then LTCI expenditure increase equal to 50% or above in one year	0.46	121
7 (C)	IF Hospitalization Period = "0" AND Alzheimer's disease = "Y" AND CNL = "3" then LTCI expenditure increase equal to 50% or above in one year	0.41	220
5 (B)	IF Hospitalization Period = "0" AND Alzheimer's disease = "N" AND CNL = "1/2/3" then LTCI expenditure increase equal to 50% or above in one year	0.35	2,463
3 (A)	IF Hospitalization Period = "0" AND CNL = "4/5" then LTCI expenditure increase equal to 50% or above in one year	0.13	992

Decision Tree (CART)

- Min Split = 70, Mini Bucket = 50, cp = 0.001



Tree Rules Table

Rule	Hospitalization	CNL	Alzheimer's disease	Sex	New Dementia	Probability	N
H	O					0.99	1,369
G	X	1,2	O	M		0.67	138
F	X	1	O	F		0.61	230
E	X	2	O	F	O	0.6	83
D	X	2	O	F	X	0.46	121
C	X	3	O			0.41	220
B	X	1,2,3	X			0.35	2,463
A	X	4,5				0.13	992

31

The evaluation results

Evaluation	CART	Random Forest
Prediction Accuracy	0.7421	0.7409
Sensitivity	0.5891	0.5547
Specificity	0.9003	0.9333
AUC	0.8087	0.8097

(validation set, n=2,408)

We decreased the number of variables but got an similar level of prediction accuracy and AUC of model.

Limitations

- In order to avoid over-fitting, we resample our subjects ratio of case to control as 1:1. Although we resample our subjects randomly, the bias that result from sampling could occur.
- The individual's economic status, and care resource of individual and communities may affect LTC expenditure increase. However, these factors are not available in this study.

33

Discussion (1/6)

- The weight ranking of factors resulted from Random Forest and CART are similar.
 - Days of hospitalization more than 30 Days.
 - Care Needs Level
 - Alzheimer's disease
 - Etc
- The decision tree with 8 terminal nodes, based on CART, resulted in

high prediction accuracy 0.742

high AUC 0.809

34

Discussion (2/6)

- The relevant factors of LTCI expenditure increase
 1. Hospitalization
 - The risk of hospitalization-associated disability at 30 days after hospital discharge increases with advanced age. (Chodos et al.(2015))

35

Discussion (3/6)

2. Care Needs Level
 - Care Needs Level is a factor of LTCI expenditure(Olivares-Tirado et al.(2011)) . Lower Care Needs Level may show unstable conditions that may require the individuals to adjust their care plan.
 - The limitation of reimbursement increased as care needs level getting higher, therefore, the amount of same proportion by different CNL is different.
 - e.g. 1% of CNL 1= 16,692 yen: 1% of CNL 5= 36,065 yen
- Therefore, comparing with higher Care Needs Level (4,5), it is easier to exceed the 50% increase for insured with lower Care Needs Level (1,2,3).

36

Discussion (4/6)

3. Alzheimer's disease (Rule B/C)

- Alzheimer's disease predict LTCI expenditure increase especially among elderly with lower care needs level and without hospitalization.
- Alzheimer's disease is a risk factor of higher healthcare utilization (Sauvaget et al.(2002); Delavande et al.(2013); Chung et al.(2014);), and associated to medical comorbidity (Doraiswamy et al.(2002)), individuals with Alzheimer's disease may require more care service.

37

Discussion (5/6)

4. New Diagnosis of Dementia (Rule D/E)

- Even the same service provided, if the elderly were dementia patient, the reimbursement would be higher than individuals without dementia.
- The type/hours of service may increase as the dependant level of dementia gets higher.

38

Discussion (6/6)

5. Sex (Rule F/G)

- The medical costs of male increase more rapidly than female in elderly. (Kim et al. (2013)) LTCI costs may have similar trends.

39

Conclusion

According to our analysis of LTC service, the algorithm (8 terminal nodes) based on CART predict LTCI expenditure increase with high prediction accuracy of 0.74 and AUC of 0.81.

The result implies that hospitalization, lower care needs level, Alzheimer's disease, new diagnosis of dementia and male gender would be predictors of LTCI expenditure increase.

In the elderly with lower care needs level and without hospitalization, Alzheimer' disease would be a powerful indicator to predict LTCI expenditure increase.

40

THANK YOU FOR YOUR LISTENING.

41

Case-to-Control Ratio

- Most of data sets in the real world have unbalance class distribution and may result in unrealistically high accuracy.
 - Most learning algorithm tend to omit the small class because of it is not supported statistically.

The Reimbursements Claims by CNL

Care Needs Level	Reimbursements Claims(yen)/monthly
Support level 1	50,030
Support level 2	104,730
Care Needs Level 1	166,920
Care Needs Level 2	196,160
Care Needs Level 3	269,310
Care Needs Level 4	308,060
Care Needs Level 5	360,650

43

Alzheimer's Disease * New Dementia

		New Dementia			
		N	Y	Total	
Alzheimer's Disease	N	Count	6114	377	6491
		% of Total	76.2%	4.7%	80.9%
	Y	Count	985	548	1533
		% of Total	12.3%	6.8%	19.1%
Total		Count	7099	925	8024
		% of Total	88.5%	11.5%	100.0%

44

認知症高齢者の日常生活自立度

レベル	判断基準
I)	「何らかの認知症を有するが、日常生活は家庭内および社会的にほぼ自立している状態」 基本的には在宅で自立した生活が可能なレベルです。
II a)	「日常生活に支障をきたすような症状・行動や意思疎通の困難さが家庭外で多少見られても、誰かが注意していれば自立できる状態」
II b)	「日常生活に支障をきたすような症状・行動や意思疎通の困難さが家庭内で見られるようになるが、誰かが注意していれば自立できる状態」
III a)	「日常生活に支障をきたすような症状・行動や意思疎通の困難さが主に日中を中心に見られ、介護を必要とする状態」
III b)	判断基準「日常生活に支障をきたすような症状・行動や意思疎通の困難さが夜間にも見られるようになり、介護を必要とする状態」
IV)	「日常生活に支障をきたすような症状・行動や意思疎通の困難さが頻繁に見られ、常に介護を必要とする状態」
M)	「著しい精神症状や周辺症状あるいは重篤な身体疾患が見られ、専門医療を必要とする状態」

45

Application of Machine Learning in Predicting Risk Factors of Care Needs Level Deterioration among Elderly with Dementia

Huei-Ru Lin

And

Chia-Lun Lo, Tetsuya Otsubo, Yuichi Imanaka

NCWIA June 2015



國立中正大學 College of Management National Chung Cheng University
醫療資訊管理中心 醫療資料探勘實驗室

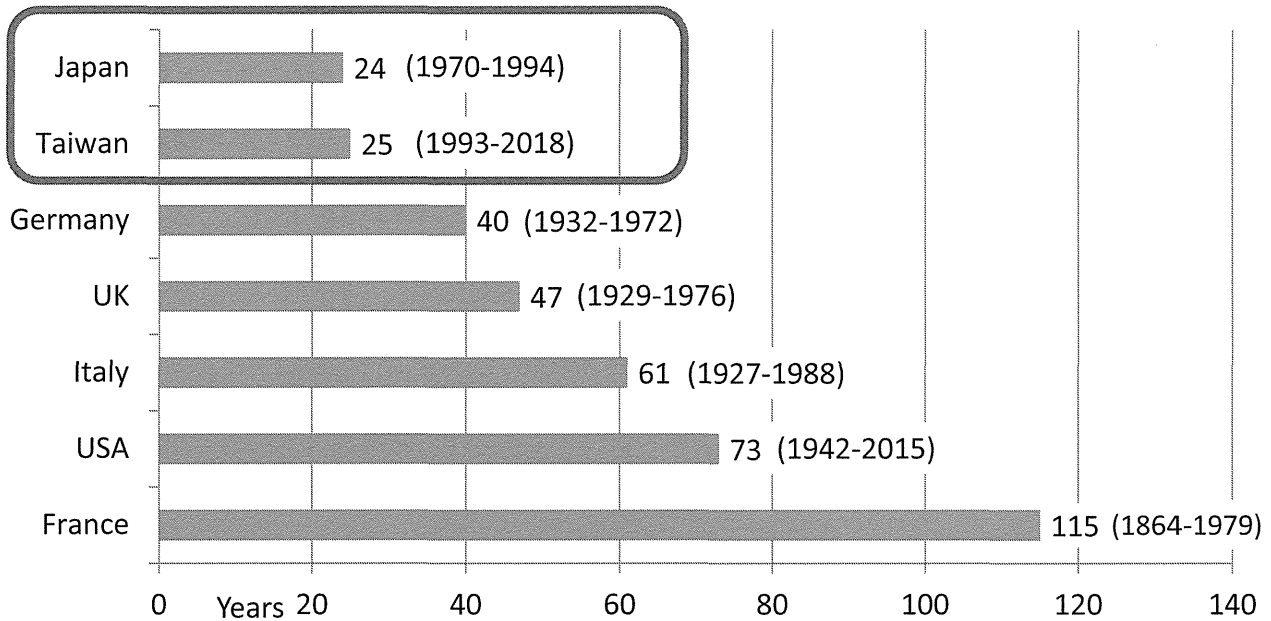


KYOTO UNIVERSITY
FACULTY / GRADUATE SCHOOL OF MEDICINE

BACKGROUND



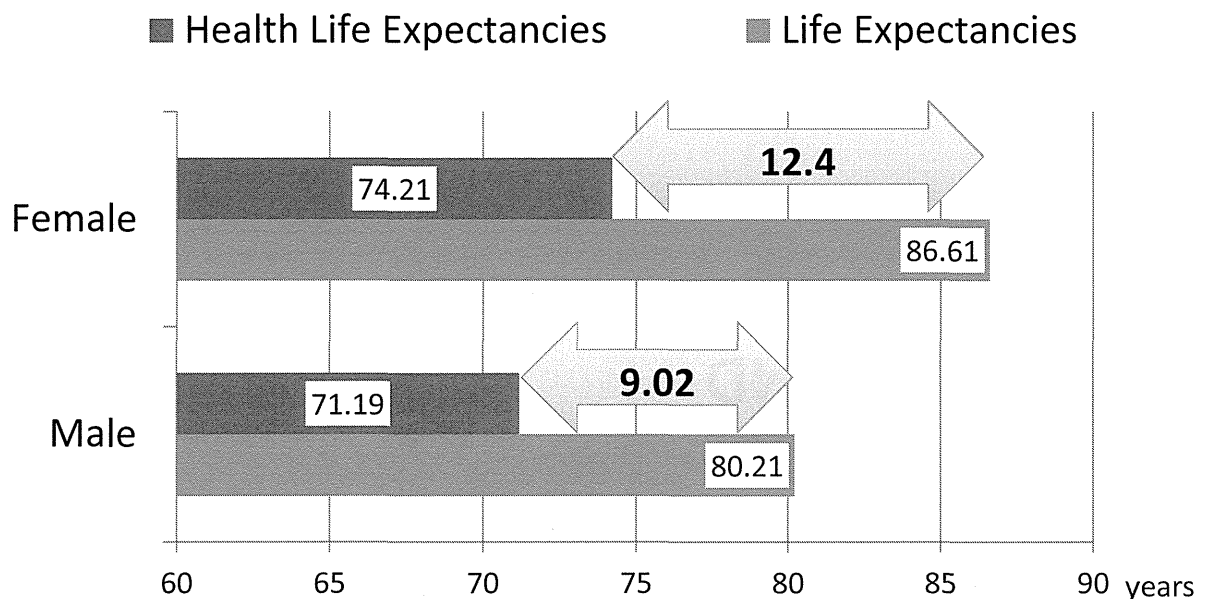
Number of Years for percent of Population Age 65 & Over to Rise from 7% to 14 %



Directorate General of Budget, Accounting and Statistics, Executive Yuan (2006): 2006 Social Indicators Annual Report in Taiwan.

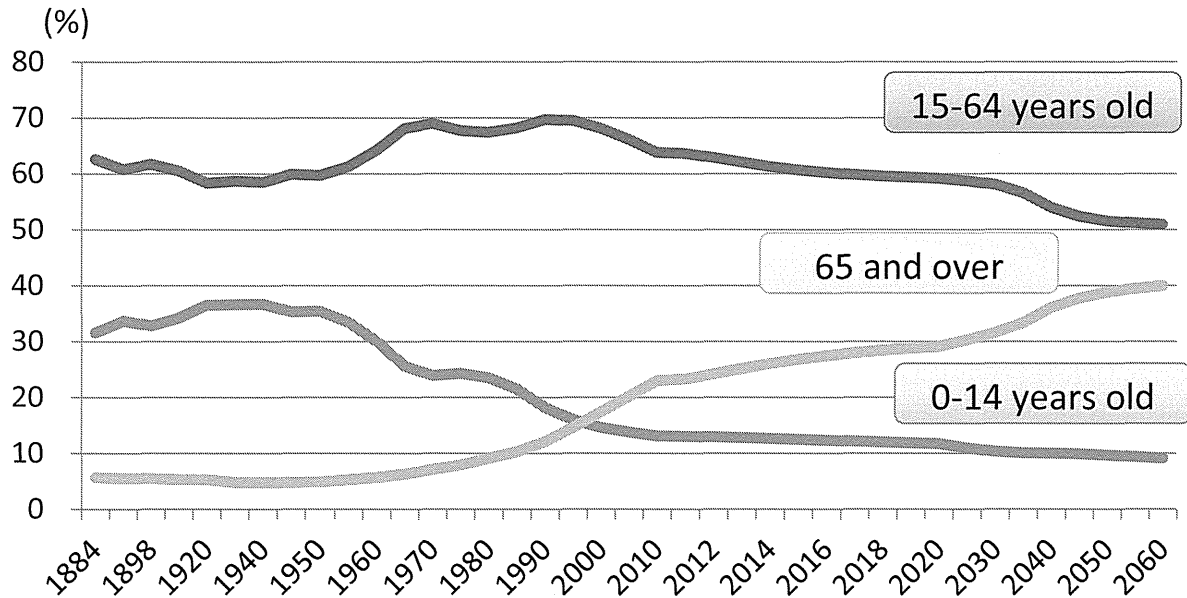
3

Life expectancies of Japan.



4

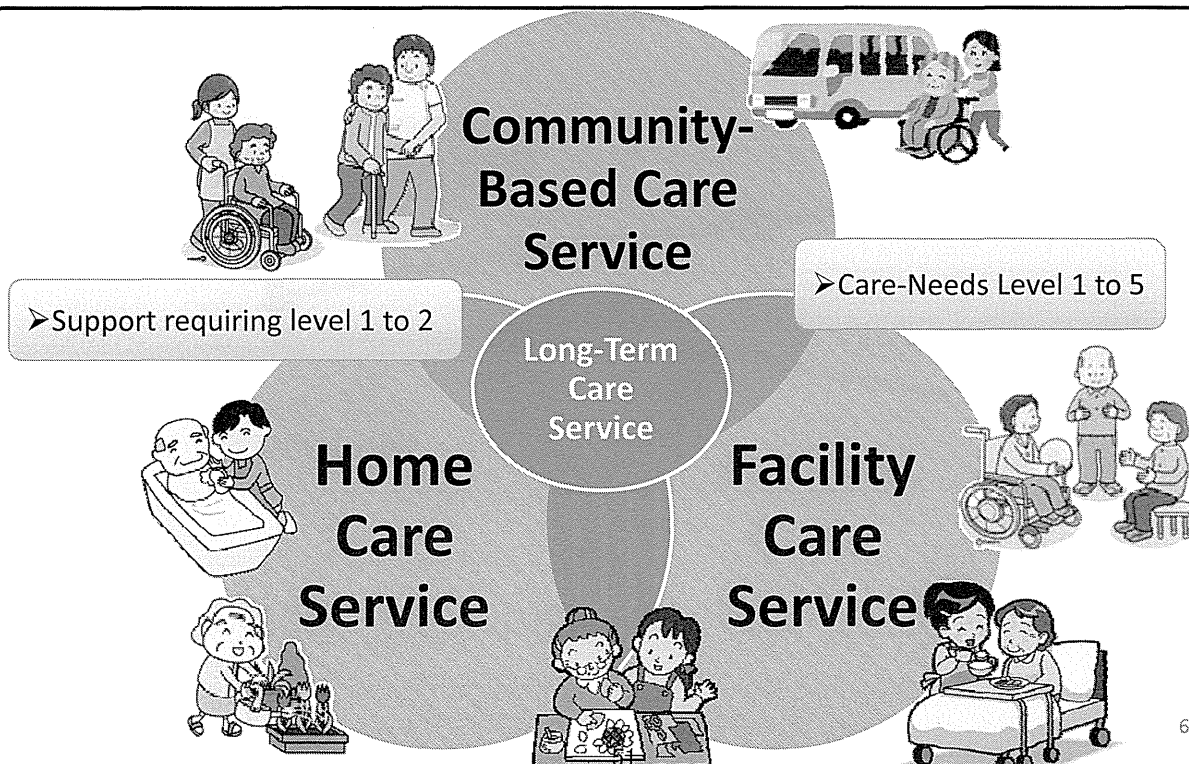
Age Structure of Population



National Institute of Population and Society Research. (2012) Population Statistics of Japan 2012. from <http://www.ipss.go.jp/p-info/e/psj2012/PSJ2012.asp>; Access 15 May 2015.

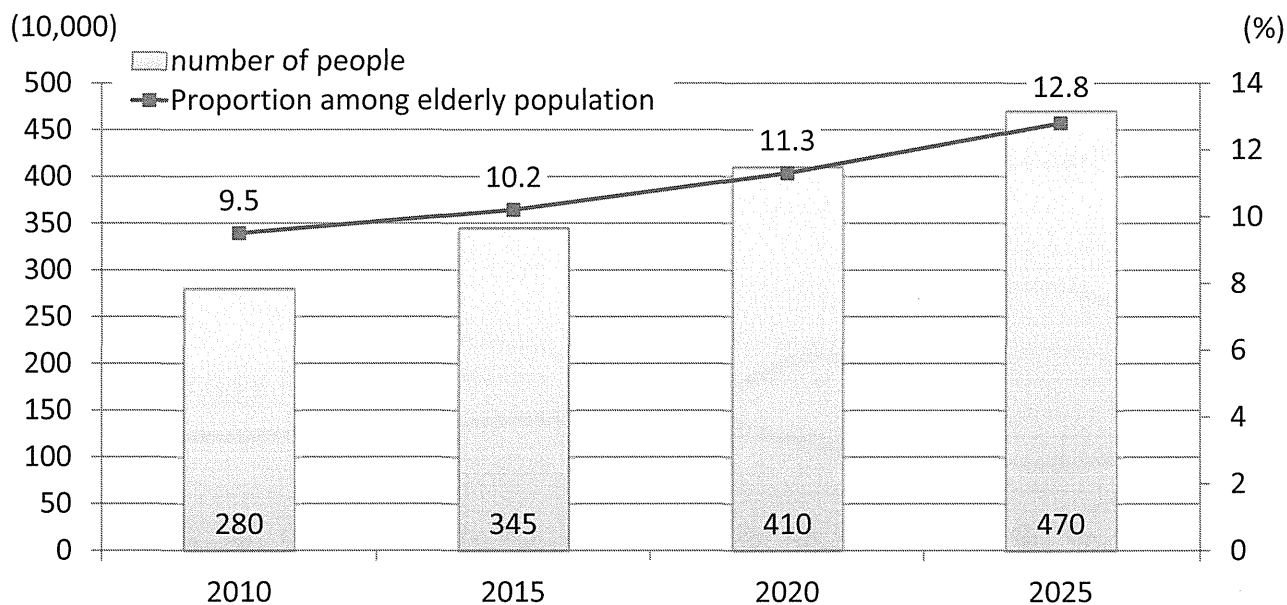
5

Long-Term Care Insurance Service - Implemented since 2000



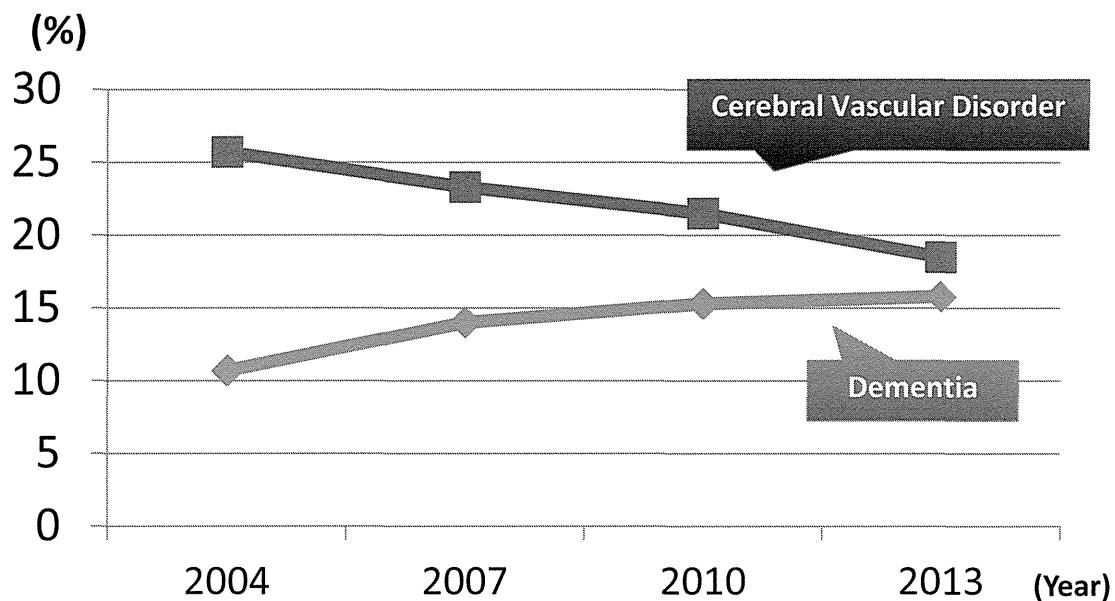
6

Estimated Number of The Elderly with Dementia based on the LTCI statistics



Ministry of Health Labour and Welfare, Health and Welfare Bureau for Elderly, 2013

The main reason of needing long-term care service



8

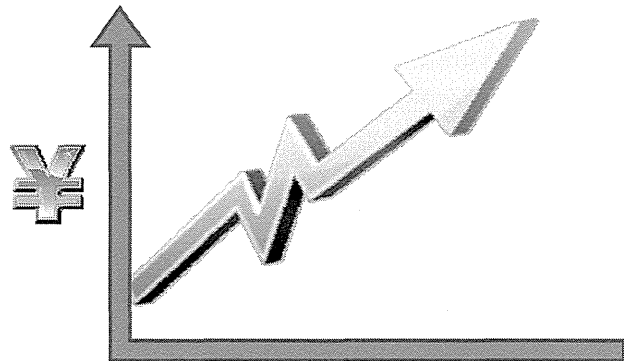
The influence on LTC by dementia

Dementia



Expenditure

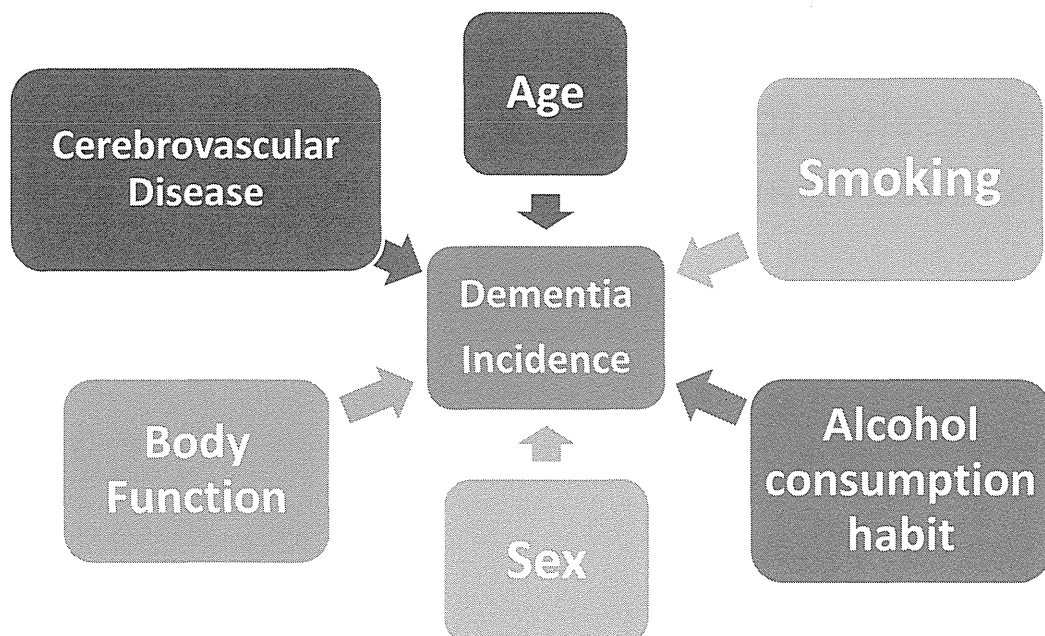
Utilization



Arling, G. et al., *Med Care*, 2013, 51(7), 575-581.
Langa, K. M. et al., *Alzheimer Dis Assoc Disord*, 2004, 18(2), 90-98.

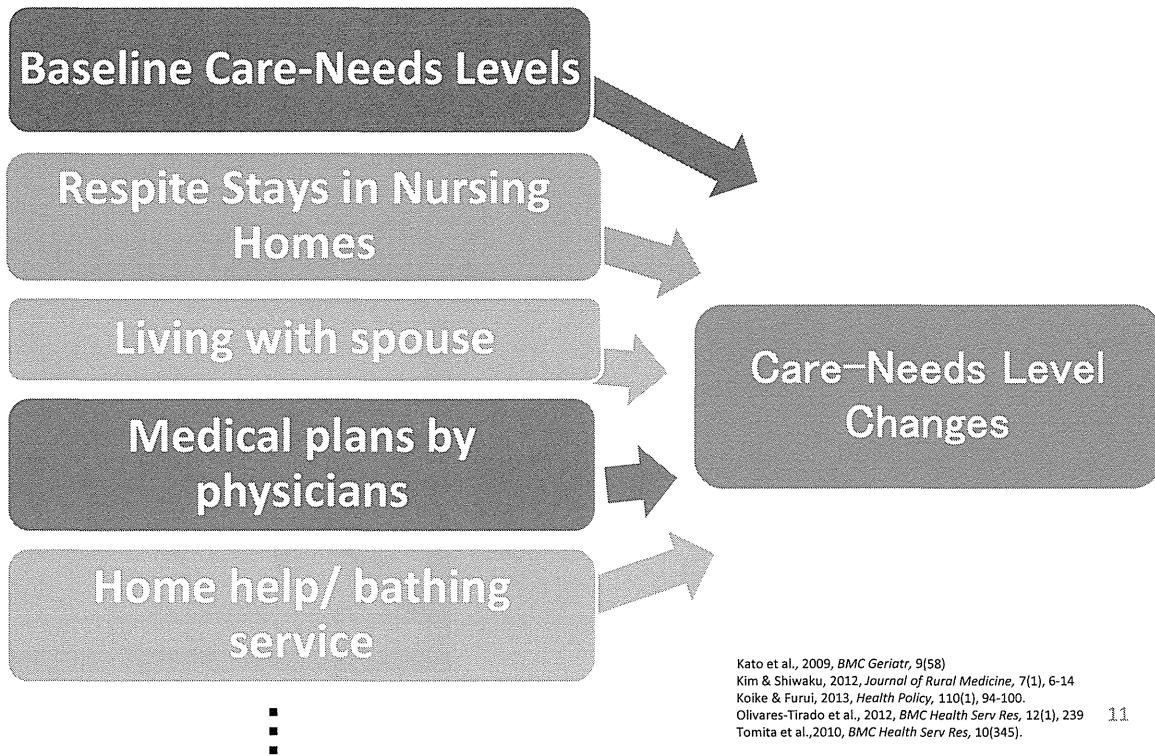


The risk factors of dementia incidence.



Anstey et al., *Am J Geriatr Psychiatry*, 2009, 17(7), 542-555.
Luck et al., *Am J Geriatr Psychiatry*, 2012, 20(11), 943-954.
Rakowski et al., *Clin J Am Soc Nephrol*, 2002, 1(5), 1000-1005. 53

Factors related to Care-Needs Level Changes



Former Methods

- Most studies conducted logistic/multiple regression to clarify the relationship between risk factors and care needs level deterioration.

Aim

- To predict the deterioration of dementia insured of long-term care insurance system by decision tree.

13

Population

- 6,876 people who utilized for long-term care service, insured of Long-Term Care Insurance with dementia and aged above 65 years in June 2010 in Kyoto prefecture.

14

Method

- Predictor
 - Care needs level deteriorated.
- Factors
 - Age, sex, baseline care-needs level, the type service use, medical area, and diagnosis of new dementia were used.
- Statistical analysis
 - Descriptive analysis
 - CART and random forest
 - Rstudio 0.98.1102

15

Descriptive statistics (1/3)

Variables	Range	Descriptive Statistics
Sex	Male/Female	Male:1,696 Female:5,180
Age Group	65-74/ 75-84/ >85	65-74:422 75-84:2,725 >85:3,729
Medical Area	Tango/ Cyutan/ Nantan/ KyotoOtokuni/ Yamashirokita/ Yamashirominami	Tango: 463 Cyutan: 617 Nantan: 448 KyotoOtokuni:4,230 Yamashirokita: 886 Yamashirominami: 232

Descriptive statistics (2/3)

Variables	Range	Descriptive Statistics
Baseline Care Needs Level	1 to 5	CNL1:949 CNL2:1,110 CNL3:1,911 CNL4:1,638 CNL5:1,268
New Dementia Diagnosis	Yes/No	Yes:719 No: 6,157

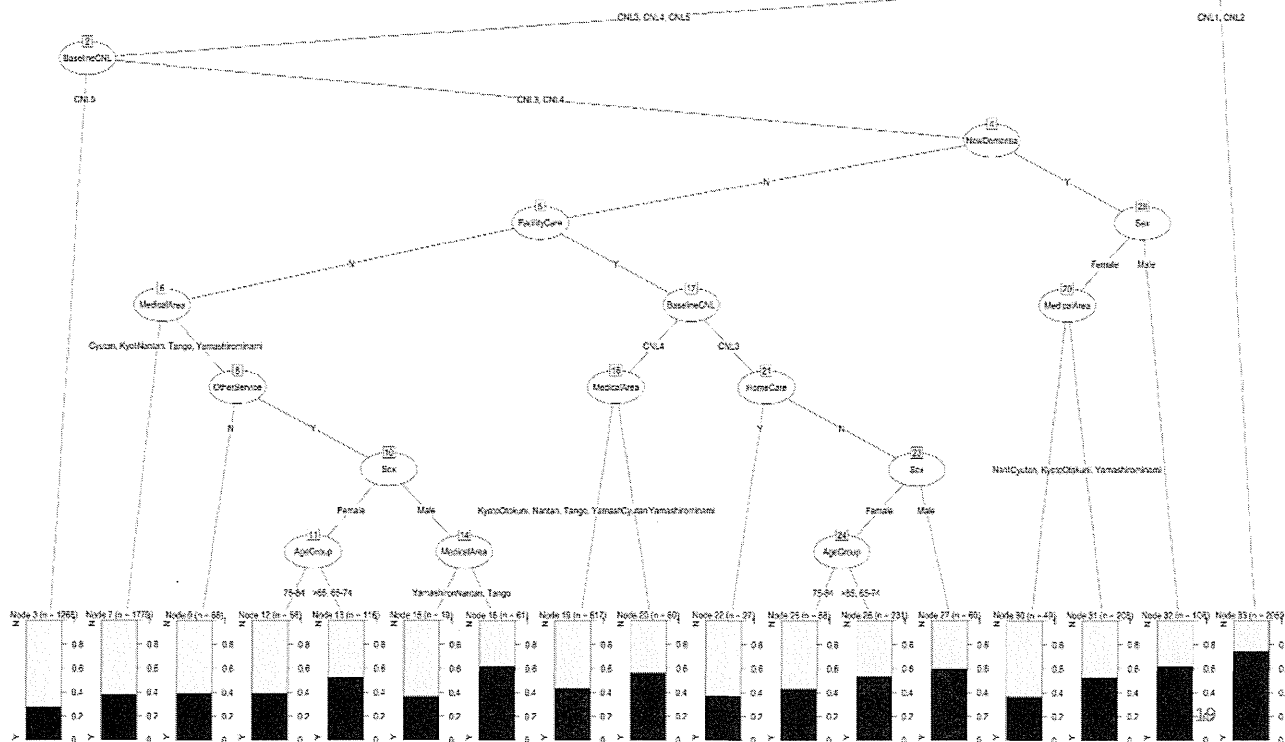
17

Descriptive statistics (3/3)

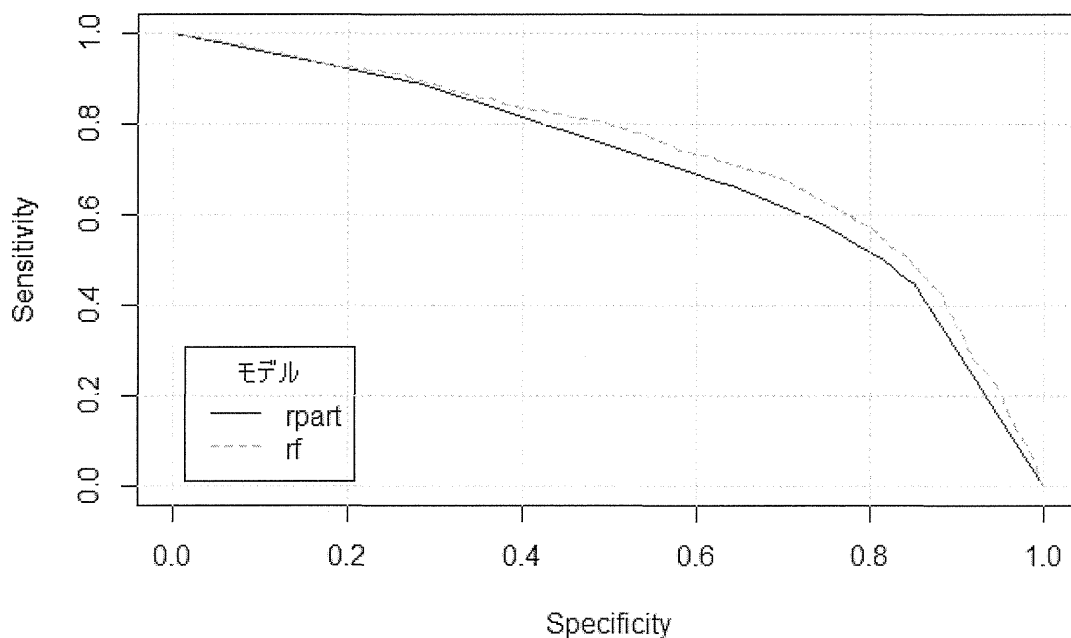
Variables	Range	Descriptive Statistics
Facility Care Service	Yes/No	Yes:1,904 No: 4,972
Home Care Service	Yes/No	Yes:4,709 No: 2,167
Community-Based Care Service	Yes/No	Yes:798 No: 6,078
Other Service	Yes/No	Yes:5,547 No: 1,329

18

Decision tree model fitted using the CART algorithm



THE SENSITIVITY/SPECIFICITY PLOT OF CART AND RANDOM FOREST MODEL.



AUC of CART: 0.6973
 AUC of random forest: 0.7279