

CRT 알고리즘을 이용한 우리나라 노인의 사회활동 영향요인 예측 모형 개발

변해원

호남대학교 보건과학대학 언어치료학과 교수

Development of Predictive Model of Social Activity for the Elderly in Korea using CRT Algorithm

Haewon Byeon

Professor, Department of Speech Language Pathology, Honam University

요 약 노년기의 사회참여는 사회적 상호작용의 기회를 제공하여 삶의 만족감을 고취시키기 때문에 성공적인 노화를 달성하기 위해서 중요하다. 이 연구는 우리나라 지역사회 노인을 대상으로 노년기 사회 활동의 관련요인과 사회 참여를 예측하는 통계적 분류 모형을 구축하였다. 분석 대상은 2015년도 지역사회 건강조사를 완료한 60세 이상 노인 1,864명(남 829명, 여 1,035명)이었다. 결과 변수는 지난 1달 간 사회 활동 경험(있음, 없음)으로 정의하였다. 예측모형은 Classification and Regression Trees(CRT) 알고리즘 기반 의사결정나무모형을 이용하여 구축하였다. 연구결과, 사회참여의 유의미한 분류 변수는 주관적 건강, 이웃과의 만남빈도, 친척과의 만남빈도, 배우자 동거여부이었고, 그 중에서도 가장 우선적으로 관여하는 예측 요인은 주관적 건강수준이었다. 본 연구의 결과를 기초로 도래하는 초고령사회의 성공적인 노화를 대비하기 위해서 노인의 사회 활동에 대한 사회적 관심과 지원이 요구된다.

주제어 : 성공적 노화, 사회 참여, 노인, 주관적 건강, 예측모형

Abstract The social activities of the elderly are important in successfully achieving aging by providing opportunities for social interaction to enhance life satisfaction. The purpose of this study is to identify the related factors of the elderly social activities and build a statistical classification model to predict social activities. Subjects were 1,864 elderly people (829 males, 1,035 females) who completed the community health survey in 2015. Outcome variables were defined as the experience of social activity during the past month(yes, no). The prediction model was constructed using decision tree model based on Classification and Regression Trees (CRT) algorithm. The results of this study were subjective health, frequency of meeting with neighbors, frequency of meeting with relatives, and living with spouse were significant variables of social participation. The most prevalent predictor was the subjective health level. In order to prepare for the successful aging of the super aged society based on the results of this study, social attention and support for the social activities of the elderly are required.

Key Words : Successful aging, Social activities, Elderly, Subjective health, Prediction model

1. 서론

저출산과 함께 평균수명이 증가하면서 우리나라의 노

인인구는 급증하는 추세이다. 우리나라는 2017년을 기준으로 65세 이상 노인 인구가 전체의 14% 이상을 상회하는 고령사회(aging society)로 이미 진입하였고, 2026년

*Corresponding Author : Haewon Byeon (bhwpuma@naver.com)

Received July 11, 2018

Accepted October 20, 2018

Revised September 11, 2018

Published October 28, 2018

에는 노인인구가 20%를 초과하는 초고령 사회(super-aged society)에 진입할 뿐만 아니라 2040년이 되면 우리나라 국민 3명 중 1명이 노인인 사회가 될 것으로 전망되고 있어 미래 사회에 대한 장기적인 대비가 필요하다[1]. 특히, 최근에 노인세대에 진입한 1955년에서 1963년 사이에 태어난 베이비부머 세대의 경우 건강상태, 근로능력, 의욕이 충분함에도 불구하고 은퇴 후 노년기 사회활동이 저조함으로써 경제적, 정서적 측면에서 노인의 삶이 취약성을 드러내고 있다[1].

노년기는 인생의 마지막 주기로서 지속적인 사회참여를 통해 발전을 달성한다면, 만족스럽고 성공적인 삶을 영위할 수 있다[2]. 이 같은 노년기의 사회참여는 노년기 이전에 미리 계획해서 지금까지 해오던 활동을 더 새롭게 발전시키는 것, 새로운 일이나 활동을 시작하는 것, 자기계발을 계속하는 것, 또는 취미나 특기를 새로 개발하여 즐기는 것 등이 될 수 있다. 그럼에도 불구하고, 현재까지 노인보건정책의 주요목표는 건강행위나 노년기 적용에 국한되어있으며, 노년기의 사회참여와 발달의 축진은 여전히 미흡한 실정이다[3].

노년기의 사회활동은 삶의 질적 측면에서 중요하다. 사회활동은 주변사람과 관계 형성을 통해서 사회와의 접촉을 지속적으로 유지시킨다. 또한, 개인적인 만족감을 고취시켜서 성공적 노화에 긍정적 영향을 주고 궁극적으로 개인의 삶의 질을 증진시킨다[4]. 노인의 여가활동은 신체적, 정서적으로 만족스러운 삶을 영위하는 데 긍정적인 영향을 미치며[5], 지역사회의 다양한 경험의 기회를 제공하여 타인과 상호작용하는 수단이 되고, 궁극적으로 노인의 건강을 증진시킬뿐만 아니라[6], 활동적인 여가활동을 통해 심리적 안정과 삶의 만족도를 제공한다[7]. 종합하면, 노년기의 사회참여는 신체적 정신적 건강을 증진시키고 사회적 상호작용의 기회를 제공하여 삶의 만족감을 고취시키고 이를 통해 성공적인 노화를 달성할 수 있다.

성공적 노화는 노년기에 형성되는 신체적, 심리적, 사회적 적응과 역할들에 초점을 두고 있으며, 특히, 다양한 사회활동과 관련이 있다[8]. Tajvar & Montazeri [6]에 따르면, 사회참여의 빈도가 높은 노인일수록 삶의 만족도가 높았다. 이처럼 노년기의 삶의 질을 향상시키고, 성공적 노화를 달성하기 위해서는 노년기 사회참여의 관련 요인의 규명이 필요하다. 특히, 사회참여는 환경, 개인요인, 경제적 요소 등 다차원적 요인에 기인하기 때문에 노

년기 사회참여의 특성을 보다 정확하게 파악하기 위해서는 위험 요인 간의 경로나 복합적 요인을 설명할 수 있는 예측모델의 구축이 필요하다. 그럼에도 불구하고, 우리나라 노인의 사회활동에 관한 관심은 대부분 고용 등 경제활동에 국한되어 있으며, 건강행위, 인구사회학적, 환경적 요인을 포괄하여 사회적 활동 참여의 영향요인을 탐색 연구는 미흡한 실정이다.

최근에는 요인을 탐색하는 방법으로 의사결정나무분석이 활용되고 있다. 데이터마이닝의 대표적인 통계적 분류 기법인 의사결정나무는 단일 종속변수에 대해 여러 복합적인 요인들의 계층적 영향을 나무 형태의 시각적인 모형으로 경로를 표현하기 때문에 도출된 결과를 해석하기 수월할 뿐만 아니라 의사결정에 유용한 규칙들을 분석하는데 용이하므로 관심변수에 영향을 미치는 요인들을 발견하고 의사결정을 선택하기 위한 분석방법으로 다양한 분야에서 이용되고 있다[9]. 본 연구는 우리나라 지역사회 전국조사의 데이터를 대상으로 Classification and Regression Trees(CRT) 알고리즘을 이용하여 노년기 사회 활동의 관련요인과 사회 참여를 예측하는 통계적 분류 모형을 구축하였다.

2. 연구방법

2.1 연구대상

질병관리본부가 전국의 254개 보건소 및 35개 대학과 협력하여 2015년 8월 31일부터 11월 8일까지 조사를 수행한 2015년도 지역사회건강조사의 원시자료를 분석하였다. 지역사회건강조사는 근거중심의 보건사업 수행의 기반을 마련하고, 지역건강통계를 생산하기 위한 목적으로 수행되었으며, 조사항목은 건강행태, 의료이용, 삶의질, 경제활동 등으로 구성되어있다. 조사방법은 훈련된 조사원이 대상자의 집을 직접 방문하여 노트북을 사용하여 1:1면접조사를 수행하는 방식을 이용하였다[10].

본 연구에서는 건강설문과 가구조사를 완료한 60세 이상 노인 1,864명(남 829명, 여 1,035명)을 분석하였다.

2.2 변수의 정의

결과 변수는 지난 1달 간 사회활동참여 경험(있음, 없음)으로 정의하였다. 설명변수는 연령(65-75세, 75세 이상), 성, 최종학력(초등학교 이하, 중학교, 고등학교, 대학

졸업 이상), 현재 취업 상태(취업, 미취업), 가구 월 평균 총 소득(200만원 미만, 200-400만원, 400만원 이상), 배우자 유무(배우자가 있고 함께 살고 있음, 배우자가 있으나 함께 살고 있지 않음, 배우자 없음), 음주빈도(주1회 이하, 주1회 이상), 흡연(현재흡연자, 비흡연자), 주관적 건강상태(좋음, 보통, 나쁨), 주당 걷기 실천 일수(2일 이상, 1일 미만), 최근 1년간 우울감 경험(없음, 있음), 최근 2주간 질환 및 사고경험(없음, 있음), 주관적 스트레스 수준(있음, 없음), 이웃과의 만남빈도(1달에 한번 미만, 1달에 두 번 이상), 친척과의 만남빈도(1달에 한번 미만, 1달에 두 번 이상)를 포함하였다.

2.3 CRT 알고리즘

의사결정나무는 의사결정이 수행되기까지 규칙을 도식화하여 관심변수와 관련이 있는 그룹을 몇 개의 하위 집단으로 분류하거나 예측하는 데이터마이닝 분석방법이다. 결과가 나무구조에 의한 분석방법으로 표현되므로 연구자가 분석과정을 해석하기 쉽다는 장점을 가지고 있으며, 분석의 정확도보다 분석과정의 설명과 해석이 필요한 경우에 유용하게 사용된다. 의사결정나무분석은 분류규칙에 따라 Classification and Regression Trees(CRT) 등 여러 분석방법이 있으며, 이진분리(binary split)를 수행하기 위한 Regression Trees의 기본 알고리즘은 Least-Squared Deviation (LSD)가 사용된다(Fig. 1).

$$i(t) = \frac{\sum_{n \in h(t)} w_n f_n (y_n - \bar{y}(t))^2}{\sum_{n \in h(t)} w_n f_n}$$

where

$$p_L = N_w(t_L) / N_w(t), \quad p_R = N_w(t_R) / N_w(t), \quad N_w(t) = \sum_{n \in h(t)} w_n f_n$$

$$\bar{y}(t) = \frac{\sum_{n \in h(t)} w_n f_n y_n}{N_w(t)}$$

Fig. 1. Least Squares Deviation impurity measures

CRT는 목표변수가 범주형인 경우에는 Gini Index를 연속형인 경우에는 분산의 감소량을 이용하여 이진분리(binary split)를 수행하는 알고리즘이다. Gini Index는 불순도를 산출하는 지수로서 특정 개체가 목표변수의 i번

째 범주로부터 추출되었고, 그 개체가 목표변수의 j번째 범주에 속한다고 오분류할 확률인 P(i|P(j))를 모두 더하여 오분류 확률의 추정치를 계산한 것이다[11].

이 연구의 모형에서 CRT 알고리즘에 대한 의사결정 규칙의 기준값은 0.05로 설정하였고, 부모마디 수는 250명, 자식마디 수는 150명, 분지가지 개수는 각각 4개로 제한하였다. 모형의 타당성 평가는 10-fold 교차타당성 검정을 이용하였다.

3. 결과

3.1 대상자의 일반적 특성

전체 대상자 1,864명 중에서 지난 1달간 사회활동 경험이 없는 노인은 1,035명(55.5%)이었다(Table 1). 카이제곱검정 결과, 지난 1달간 사회활동 경험이 있는 노인과 없는 노인은 연령, 성별, 최종학력, 경제활동여부, 가구소득, 배우자 유무, 음주빈도, 주관적 건강상태, 주당 걷기 실천일수, 최근 1달간 우울감 경험, 최근 2주간 질환 및 사고경험, 이웃과의 만남빈도, 친척과의 만남빈도에서 유의한 차이가 있었다($p < 0.05$). 75세 이상(60.1%), 여성(48.5%), 초등학교 졸업이하(55.9%), 미취업자(48.1%), 가구소득 200만원 미만(51.8%), 배우자 사별(68.8%), 주1회 이하 음주(47.2%), 주관적 건강상태 나쁨(60.6%), 주당 1일미만 걷기 실천(70.6%), 최근 1달간 우울감 경험자(60.1%), 최근 2주간 질환 및 사고경험자(51.8%), 이웃과의 만남빈도가 월1회 미만(56.8%), 친척과의 만남빈도 월1회 미만(59.2%)에서 사회활동 무경험 비율이 높았다.

3.2 CRT 알고리즘에 의한 노년기 사회참여 모형

CRT 알고리즘을 이용한 우리나라 노인의 사회참여 경험 예측모형은 Fig. 2에 제시하였다. CRT 알고리즘 분류모형 구축 결과, 사회참여의 유의미한 분류 변수는 주관적 건강, 이웃과의 만남빈도, 친척과의 만남빈도, 배우자 동거여부였고, 그 중에서도 가장 우선적으로 관여하는 예측 요인은 주관적 건강수준이었다.

Table 2는 CRT 알고리즘을 이용한 노년기 사회참여 예측모형의 이익 도표이다. 사회참여 경험이 없는 노인을 효과적으로 예측하는 유의미한 경로로 총 2개의 노드가 확인되었다. 먼저, 이익지표 값이 가장 큰 제1경로는 이웃과의 만남빈도가 월 1회 미만이고, 주관적 건강수준

Table 1. The general characteristics of the subjects

Variables	Experience of social activity during the past month		p
	Yes (n=1,020)	No (n=844)	
Gender			0.002
Male	487 (58.7)	342 (41.3)	
Female	533 (51.5)	502 (48.5)	
Age			<0.001
65-74	784 (61.6)	489 (38.4)	
75+	236 (39.9)	355 (60.1)	
Education			<0.001
Elementary school	335 (44.1)	424 (55.9)	
Middle school	191 (55.8)	151 (44.2)	
High school	290 (61.8)	179 (38.2)	
University	199 (69.6)	87 (30.4)	
Monthly mean total household income			<0.001
<2 million KRW	526 (48.2)	566 (51.8)	
2-4 million KRW	289 (63.2)	168 (36.8)	
4 million KRW ≤	184 (65.2)	98 (34.8)	
Current employment status			<0.001
Employment	284 (63.5)	163 (36.5)	
Unemployment	736 (51.9)	681 (48.1)	
Marital status			<0.001
Living with a spouse	762 (60.2)	504 (39.8)	
Married but not living with a spouse	252 (43.4)	329 (56.6)	
Not married	5 (31.3)	11 (68.8)	
Walking per week			<0.001
Equal to or more than 2 days	955 (58.1)	688 (41.9)	
Less than 1 day	65 (29.4)	156 (70.6)	
Drinking frequency			<0.001
Less than once a week	830 (52.8)	742 (47.2)	
More than once	189 (65.2)	101 (34.8)	
Smoking			0.320
Current smoker	89 (51.1)	85 (48.9)	
Nonsmoker	931 (55.1)	759 (44.9)	
Subjective health status			<0.001
Good	334 (63.7)	190 (36.3)	
Normal,	432 (62.2)	263 (37.8)	
Bad	254 (39.4)	391 (60.6)	
Subjective stress level			0.196
Yes	661 (53.7)	571 (46.3)	
No	359 (56.8)	273 (43.2)	
Depression experience in the past 1 year			<0.001
Yes	65 (39.9)	98 (60.1)	
No	954 (56.1)	746 (43.9)	
Disease or accident experience in the past 2 weeks			0.012
Yes	149 (48.2)	160 (51.8)	
No	870 (56.0)	683 (44.0)	
Frequency of meeting a neighbor			<0.001
Less than once a month	266 (43.2)	350 (56.8)	
More than twice a month	747 (60.9)	480 (39.1)	
Frequency meeting a relative			<0.001
Less than once a month	151 (40.8)	219 (59.2)	
More than twice a month	868 (58.2)	623 (41.8)	

이 나쁜 노인으로 72.1%가 지난1달간 사회참여 경험이 없다고 응답하였고, 이익지표는 173.4%이었다. 다음으로 이익지표 값이 큰 제2경로는 이웃과의 만남빈도가 월 2

회 이상이지만, 배우자가 사별하였거나 배우자가 있더라도 별거중인 주관적 건강 수준이 나쁜 노인으로 62.7%가 지난1달간 사회참여 경험이 없다고 응답하였고, 이익지

표는 150.7%이었다.

10-fold 교차타당성 검정의 수행 결과, 모형의 예측정확도는 65.1% 이었고, 위험지수는 크로스 분류모형의 위험지수는 0.349, 오분류율은 35%, 표준오차는 0.009로 도출되어, 예측모형의 위험지수 0.369 및 오분류율 37%, 표준오차 0.009와 유사하였다.

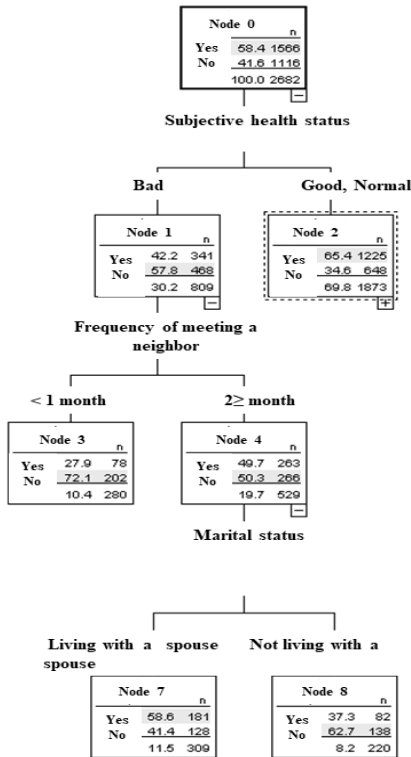


Fig. 2. Prediction model for participating in social activity for the elderly in Korea

Table 2. Gains chart of predictor variable

Node no	Response % ¹	Gain Index % ²	Characteristics
3	72.1	173.4	Subjective health = poor; Frequency of meeting with neighbors = less than once a month
8	62.7	150.7	Subjective health = poor; Frequency of meeting with neighbors = more than once a month; Whether a spouse lives together = Married but not living with a spouse or Not married

¹ Response (%): No social activity experience in subjects

² Gain index (%):=173.4 in total 8 node

4. 논의 및 결론

본 연구에서 우리나라노인의 사회참여 예측모형을 구축한 결과, 주관적 건강, 이웃과의 만남빈도, 친척과의 만남빈도, 배우자 동거여부이었고, 그 중에서도 가장 우선적으로 관여하는 예측 요인은 주관적 건강수준이었다. 이러한 결과는 주관적 건강이 사회참여와 밀접한 관계가 있다는 Takeyachi 등[12]의 결과와 유사하다.

Pinquart[13]의 메타연구에 따르면 자가 평가에 의한 주관적인 건강은 본인 스스로 건강을 정의내리고 본인의 건강을 측정하는 것이므로 노인의 건강상태를 더욱 총체적으로 반영하는 것으로 알려져 있다. 이러한 주관적인 건강상태는 개인차에 따라 상이하지만 사회활동을 평가하는 가장 핵심적인 요인이라 성공적 노화의 핵심이다 [13]. 본 연구의 결과와 유사하게 Tajvar 등[7]의 연구에서도 주관적 건강수준이 높을수록 사회참여와 삶의 질이 더 높은 것으로 보고되었다.

이웃이나 친척과의 정기적인 접촉이 사회참여의 주요 예측요인이었다. 이웃과의 사회적 접촉이 빈번할수록 사회참여의 가능성이 더 높았고, 반면 친구, 이웃과 접촉이 낮을수록 사회참여의 가능성이 낮았다. 본 연구의 결과와 마찬가지로 Seeman 등[14]의 연구에서도 친구, 친척, 이웃과의 사회적 접촉관계뿐만 아니라 타인과의 사회적 접촉도 성공적 노화의 주요 요인이었다. 흥미로운 점은 우정의 크기보다 사람들과의 정기적인 접촉이 더 중요하다는 것이다. 따라서 본 연구를 기초로 노년기 사회참여를 증진시키기 위해서 이웃이나 친척과의 접촉횟수의 확대할 수 있는 사회적 지지가 요구된다.

반면, 노인의 경제활동이나 소득수준은 사회참여의 주요예측요인으로 도출되지 않았다. 그럼에도 불구하고, 우리나라는 그동안 노년기 사회참여를 확대하는 방안으로 개인의 안정적인 소득 보장을 위한 고용이나 평생교육 등 노동력의 질을 높이기 위한 노인 복지에 대한 투자가 지속적으로 수행되어왔다[15]. 노년기 삶의 질을 확보해 줄 수 있는 기초보장은 노인 계층이 건강을 유지하면서 경제 활동을 지속 시켜 나갈 수 있도록 사회 환경을 조성한다는 측면에서는 중요하지만, 기초보장만으로는 사회참여와 활동적(active ageing)인 노인문화의 조성으로 연결되는 것은 아니다. 따라서 노인의 사회참여를 확대하기 위해서 소득보장 외에 사회연결망을 확대할 수 있는 정책적 마련이 시급하다.

본 연구의 모형을 기초로 고령사회의 성공적인 노화를 위해서 노년기 사회참여에 대한 사회적 관심과 제도적 마련이 요구된다.

REFERENCES

- [1] M. Hwang & S. Jung. (2011). The ageing society of Korea and the population estimate. *Korea Journal of Population Studies*, 34(2), 113-133.
UCI : <http://uci.or.kr/G704-000152.2011.34.2.004>
- [2] A. Sarvimäki & B. Stenbock Hult. (2000). Quality of life in old age described as a sense of well being, meaning and value. *Journal of Advanced Nursing*, 32(4), 1025-1033.
DOI: <https://doi.org/10.1046/j.1365-2648.2000.01568.x>
- [3] S. Kwon & M. R. Reich. (2005). The changing process and politics of health policy in Korea. *Journal of Health Politics Law*, 30(6), 1003-1026.
DOI: <https://doi.org/10.1215/03616878-30-6-1003>
- [4] V. H. Menec. (2003). The relation between everyday activities and successful aging: A 6-year longitudinal study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58(2), S74-S82.
DOI: <https://doi.org/10.1093/geronb/58.2.S74>
- [5] K. R. Fox. (1999). The influence of physical activity on mental well-being. *Public health nutrition*, 2(3a), 411-418.
DOI: <https://doi.org/10.1017/S1368890099000567>
- [6] T. W. Lee, I. S. Ko & K. J. Lee. (2006). Health promotion behaviors and quality of life among community-dwelling elderly in Korea: A cross-sectional survey. *International journal of nursing studies*, 43(3), 293-300.
DOI: <https://doi.org/10.1016/j.ijnurstu.2005.06.009>
- [7] M. Tajvar, M. Arab & A. Montazeri. (2008). Determinants of health-related quality of life in elderly in Tehran, Iran. *BMC public health*, 8(1), 323. DOI: <https://doi.org/10.1186/1471-2458-8-323>
- [8] A. Bowling & P. Dieppe. (2005). What is successful ageing and who should define it?. *Bmj*, 331(7531), 1548-1551.
DOI: <https://doi.org/10.1136/bmj.331.7531.1548>
- [9] H. Byeon. (2015). The factors of participating in a smoking cessation program using integrated method of decision tree and neural network algorithm. *Journal of the Korea Convergence Society*, 6(2), 25-30.
DOI : 10.15207/JKCS.2015.6.2.025
- [10] Korea Center for Disease Control and Prevention. (2017). National sample design for 2015 community health survey. Seoul: Korean Association for Survey Research.
- [11] J. Choi & D. Seo. (1999). Decision trees and its applications. *Journal of Korean Official Statistics*, 4(1), 61-83.
- [12] Y. Takeyachi, S. I. Konno, K. Otani, K. Yamauchi, I. Takahashi, Y. Suzukamo & S. I. Kikuchi. (2003). Correlation of low back pain with functional status, general health perception, social participation, subjective happiness, and patient satisfaction. *Spine*, 28(13), 1461-1466.
DOI: 10.1097/01.BRS.0000067091.88283.B6
- [13] M. Pinquart. (2001). Correlates of subjective health in older adults: A meta-analysis. *Psychology and aging*, 16(3), 414-426.
- [14] T. E. Seeman, T. M. Lusignolo, M. Albert & L. Berkman. (2001). Social relationships, social support, and patterns of cognitive aging in healthy, high-functioning older adults: MacArthur studies of successful aging. *Health psychology*, 20(4), 243.
DOI: <http://dx.doi.org/10.1037/0278-6133.20.4.243>
- [15] H. G. Koo. (2016). A study on proposal to develop senior related policies : convergence approach of both age norm and hoping activities. *Journal of Digital Convergence*, 14(4), 101-110.
DOI : <http://doi.org/10.14400/JDC.2016.14.4.101>

변 해 원 (Haewon Byeon)

[정회원]



- 2013년 2월 : 아주대학교 대학원 의학과 (이학박사)
- 2009년 9월 ~ 2011년 2월 : 아주대학교의료원 연구원
- 2011년 3월 ~ 2013년 2월 : 대림대학교 언어재활과 교수
- 2013년 3월 ~ 2017년 2월 : 남부대학교 언어치료학과 교수
- 2017년 3월 ~ 현재 : 호남대학교 언어치료학과 교수
- 관심분야 : 예측 모형, 데이터 마이닝, 의사소통장애
- E-Mail : bhwpuma@naver.com