



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

# 유전성 유방암-난소암 증후군의 가족 알림 의도

연세대학교 대학원

간 호 학 과

김 예 은

# 유전성 유방암-난소암 증후군의 가족 알림 의도

지도교수 김 수

이 논문을 석사 학위논문으로 제출함

2023년 12월 일

연세대학교 대학원

간 호 학 과

김 예 은

## 김예은의 석사 학위논문을 인준함

심사위원 \_\_\_\_\_ 김 수 인

심사위원 \_\_\_\_\_ 이 향 규 인

심사위원 \_\_\_\_\_ 박 창 기 인

연세대학교 대학원

2023년 12월 일

## 차 례

표 차례 .....	iv
그림 차례 .....	iv
부록 차례 .....	iv
국 문 요 약 .....	v
I. 서론 .....	1
1.1. 연구의 필요성 .....	1
1.2. 연구의 목적 .....	4
1.3. 용어의 정의 .....	4
1) 가족 .....	4
2) 가족 알림 의도(Intention to inform family) .....	4
II. 문헌고찰 .....	6
2.1. HBOC 의 가족 알림 및 영향요인 .....	6
2.2. HBOC 의 가족 알림 의도 및 영향 요인 .....	11

III. 연구방법 .....	14
3.1. 연구 설계 .....	14
3.2. K-CASCADE 연구 .....	15
3.3. 연구 대상자 .....	16
3.4. 연구 변수 .....	17
3.5. 자료 전처리 .....	27
3.6. 자료 변환 .....	27
3.7. 자료 분석 방법 .....	27
IV. 결과 .....	30
4.1. 일반적 특성 .....	30
4.2. 가족 알림 의도 집단 별 비교 .....	34
4.3. 예측모델 구축과 성능 .....	37
4.4. 예측모델의 변수 중요도 .....	42
V. 논의 .....	47
5.1. HBOC 의 가족 알림 의도 예측 인자 .....	48
5.2. 연구의 제한점 .....	54

5.3. 연구의 의의 .....	55
1) 간호 이론 측면 .....	55
2) 간호 연구 측면 .....	56
3) 간호 실무 측면 .....	56
VI. 결론 .....	58
참고문헌 .....	61
Appendix .....	74
Abstract .....	77

## 표 차례

Table 1. Independent Variables of the Study.....	24
Table 2. Comparison of non-intention and intention groups (N=219)....	31
Table 3 Performance of the Prediction Models.....	39
Table 4. Meta Learner Summary.....	42
Table 5 Variable Importance in Random Forest.....	44

## 그림 차례

Figure 1. Knowledge Discovery in Databases process.....	15
Figure 2. ROC curve of Logistic regression.....	39
Figure 3. ROC curve of Logistic Regression with Lasso/Ridge.....	40
Figure 4. ROC curve of Random Forest.....	40
Figure 5. ROC curve of Ensemble Learning.....	41
Figure 6. Variable Importance Plot.....	46

## 부록 차례

Appendix 1. Comparison of K-CASCADE whole sample and this study' s sample.....	74
---	----

## 국 문 요 약

### 유전성 유방암-난소암 증후군의 가족 알림 의도

유전성 유방암-난소암 증후군(Hereditary Breast and Ovarian Cancer syndrome, HBOC)은 가족 구성원에 대한 암 유전자 검사가 권장된다. 가족 내에서 HBOC 초기 사례(index case)가 가족에게 유전자 검사를 권장하는 첫 단계는 가족에게 본인의 유전자 검사 결과를 알리기 위한 결정을 하는 것이다.

본 연구의 목적은 한국 HBOC 보인자의 가족 알림 의도 예측인자를 규명하는 것이다. 이를 위해 HBOC 중 가족 알림 의도가 있는 군과 알림 의도가 없는 군 간의 차이를 비교하고, 둘째, HBOC의 가족 알림 의도 예측모델을 구축하여 평가하고, 셋째, HBOC의 가족 알림 의도 예측에 가장 적절한 모델의 변수 중요도를 확인하였다.

K-CASCADE 코호트의 2021년 3월부터 2023년 9월까지 등록된 219명의 HBOC 자료를 사용하였으며, K-CASCADE 기초설문의 일반적 특성, 암 병력, 유전자 서비스, 유전자 검사를 시행한 이유, 자기 관리, 가족 환경, 의학적 환경, 신념과 태도 범주의 변수를 독립변수로 채택하였다. 종속 변수는 가족 알림 의도로, 혈연 관계에 있는 가족 구성원을 현재 진행 중인 연구에 초대할

의도로 측정하였다. 분석 시, 가족 알림 의도 여부에 따라 독립 T 검정과 카이제곱 검정을 시행하였다. 예측모델은 전통적인 변수 선택법을 활용한 로지스틱 회귀분석과 LASSO/Ridge 방법을 활용한 로지스틱 회귀분석, 랜덤 포레스트 방법을 활용하였다. 예측모델은 R 프로그램의 SuperLearner 패키지를 활용하였다. 예측모델은 민감도, 특이도, 정확도, F1 값, ROC curve 로 평가하였다. 랜덤 포레스트 모델의 변수 중요도는 지니 평균 감소량으로 측정했다.

가족 알림 의도 여부에 따라 비교했을 때, 나이가 많을수록( $Z=-2.03$ ,  $p=.044$ ), 임신 경험이 있을 때( $\chi^2=4.94$ ,  $p=.026$ ), 난소암 진단 받은 경우( $\chi^2=6.34$ ,  $p=.012$ ), 유전자 검사 후회가 적을수록( $Z=-2.63$ ,  $p=.009$ ), 인지된 유전자 검사 유용성이 높을수록( $Z=-3.26$ ,  $p=.001$ ), 그리고 19 세 이상의 자녀와 동거하는 경우( $\chi^2=4.29$ ,  $p=.038$ ) 가족 알림 의도가 높았다.

예측모델 중에서는 랜덤 포레스트 모델이 민감도 0.9463, 특이도 0.0857, 정확도 0.6712, F1 값 0.7966, AUC 0.516 으로 가장 성능이 좋은 모델이었다.

랜덤 포레스트 모델의 변수 중요도가 높은 상위 10 개 변수는 암 관련한 자기 효능감, 나이, 가족 결속력, 처음 암을 진단 받은 나이, 문제 중심 대처, 질병에 대한 가족의 지지, 의학적 의사결정 참여도, 감정 중심 대처, 암 유전 관련 이해도, 유전자 검사에 대한 후회이다.

K-CASCADE 코호트의 가족 알림 의도는 32%로 스위스 CASCADE 코호트 65%에 비해 낮은 비율을 보였다. 가족 알림 의도 여부에 따른 비교에서 밝혀지지 않았던 암 관련한 자기 효능감, 가족 결속력, 처음 암을 진단 받은 나이 등 과 같은 변수들이 예측모델을 통해 예측 변수임을 확인하였다. 다양한 요인이 HBOC 가족 알림 의도에 영향을 미칠 수 있기에 초기에 대상자의 다양한 요인을 파악하고, 적절한 유전자 서비스 및 교육을 제공하는 것이 필요하다. 추후 EMR 과 같은 객관적 데이터를 추가하여 모델의 성능 향상을 도모해 볼 수 있다.

본 연구는 한국 HBOC 의 가족 알림 의도를 이해하는데 기초 자료가 되며, 가족 알림 의도를 향상시키기 위해 다양한 사회 심리적 요인을 고려한 단계별 유전자 서비스 개발에 활용할 수 있다.

---

핵심이 되는 말: 가족 알림, 의도, HBOC, K-CASCADE

# I. 서론

## 1.1. 연구의 필요성

유전성 유방암-난소암 증후군 (Hereditary Breast and Ovarian Cancer syndrome: HBOC)은 상염색체 우성 유전되는 유전성 암의 한 종류로 *BRCA1* 또는 *BRCA2* 병인성 변이(pathogenic variant)가 대표적이다. 일반인에서는 유방암 발병률이 12%, 난소암 1.5%인데 비해 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 보인자의 경우, 80 세까지 유방암 위험률은 각각 72%와 44%, 난소암 위험률은 각각 44%와 17%이다(한국유방암학회 & 한국인유전성유방암연구회, 2021). 또한 일반 남성의 유방암 발병율은 1.5%에 비해 변이 보인자의 경우 1.2-6.5%의 위험율을 보이며, 전립선 암 위험은 15%까지 올라가며, 남녀 모두 췌장암, 대장암, 위암, 흑색종(melanoma), 담도암, 자궁내막암 등 기타 암의 위험률은 병인성 변이가 없는 사람들에 비해 2-6 배 높다(한국유방암학회 & 한국인유전성유방암연구회, 2021). HBOC 는 일반 인구집단에 비해 이른 시기에 암이 발병할 가능성이 있으므로 조기에 병인성 변이를 발견하는 것이 암 예방 활동에 도움이 된다.

HBOC 고위험군은 통상적으로 HBOC 병인성 변이 보인자일 확률이 10% 이상인 사람을 의미한다. HBOC 환자를 기준으로 1 등친(first-degree relative)에 속하는 부모 혹은 자녀, 형제에게 환자와 동일한 병인성

변이가 있을 확률은 50%이며, 2 등친(second-degree relative)에 속하는 조부모 혹은 조카, 이모, 삼촌 등은 25%, 3 등친(third-degree relative)에 속하는 사촌은 12.5%으로 고위험군에 속한다. 이들이 HBOC 를 통해 연쇄적 유전자 검사(cascade genetic testing, CGT)를 조기에 시행한다면, 조기에 병인성 변이를 확인하여 자신의 건강에 대한 의사결정을 할 때 중요한 정보를 얻을 수 있으며, 결과에 따라 예방적 수술이나 화학요법 등 암 예방활동을 선택할 수 있다(Hanley et al., 2019; Mansfield et al., 2020). HBOC 가족의 유전자 검사는 암을 사전에 예방할 수 있도록 도움이 되므로, 사회적으로도 경제적인 이익이 될 수 있다(Ficarazzi et al., 2021). CGT 의 활성화는 제한된 의료 재원을 효과적으로 재분배할 수 있는 방법 중 하나이다. 이러한 CGT 활성화에 중요한 첫 단추는 HBOC 의 가족 알림(family disclosure)이다.

현재 대부분의 국가는 개인정보 보호 관련 법으로 한 개인이 병인성 변이가 발견되어 HBOC 로 진단받았을 때, 검사를 진행한 병원 혹은 유전자 검사센터에서 직접적으로 환자의 가족에게 연락할 수 없다(Surbone, 2011; Young et al., 2020). HBOC 로 진단받은 자가 직접 가족에게 본인이 HBOC 진단받았음을 알려야 가족 구성원이 유전자 검사 관련 서비스를 선택할 수 있게 된다. 한 연구에서는 유전자 검사 서비스를 받은 HBOC 가 유전성 변이 결과에 대해 의사가 가족에게 신뢰할만한 웹사이트를 알려주기를 바라거나

유전자 검사 결과에 대한 브로슈어나 책 혹은 동영상 자료에 대한 요구도가 높았음을 확인하였다(Sarki et al., 2022). HBOC의 연쇄적 유전자 검사 비율을 높이기 위해서는 현재 행해지는 유전자 서비스에 변화가 필요하나, 가족 알림 의도에 영향을 미치는 요인에 대해 알려진 바가 적어 가족 알림 의도를 향상시키기 위한 통용되는 유전자 검사 서비스가 없다.

현재까지 연구는 HBOC가 가족에게 유전자 검사 결과를 알리는데 미쳤던 요인을 탐색한 경우가 많다. 그러나 가족 알림을 확대시키기 위해서는 특정 행동에 앞선 행동의 의도에 미치는 영향요인을 확인하는 것이 필요하다. 행동의 변화를 위해서는 그에 앞선 의도 형성을 촉진하여 의도를 행동으로 변화시키는 것을 목표로 해야 하기 때문이다(Webb & Sheeran, 2006). 따라서 본 연구는 가족에게 알리는 행동에 앞서, 가족에게 알림 의도가 있는 자와 없는 자를 비교하고 예측모델을 탐색하는 것을 목표로 한다. 또한 연구자가 변수를 선택하는 기존 접근에서 한 걸음 나아가 다양한 예측모델을 고려해봄으로써 새로운 변수 여부 및 중요도를 확인해보고자 한다.

## 1.2. 연구의 목적

본 연구는 한국 HBOC 보인자의 가족 알림 의도 예측인자를 규명하고자 하며 세부 목표는 아래와 같다.

첫째, HBOC 중 가족 알림 의도가 있는 군과 알림 의도가 없는 군 간의 차이를 비교한다.

둘째, HBOC 의 가족 알림 의도 예측모델을 구축하여 평가한다.

셋째, HBOC 의 가족 알림 의도 예측에 가장 적절한 모델의 변수 중요도를 확인한다.

## 1.3. 용어의 정의

### 1) 가족

가족은 혈연과 혼인으로 관계를 맺은 사람들의 집단이다. 본 연구에서 가족이란 HBOC 와 혈연관계로 병인성 변이 보인자일 확률이 있는 자를 의미한다.

### 2) 가족 알림 의도(Intention to inform family)

의도란 행동의 유일하고 즉각적인 선행 인자로, 주어진 행동을 수행하려는 시도이다(Ajzen, 1991). 가족 알림 의도란 HBOC 가 본인이

HBOC 이 있음을 본인 가족에게 알리는 행동의 유일하고 즉각적인 선행 인자로, 가족에게 알리는 행동을 수행하려는 시도이다.

HBOC 보인자와 그 가족을 대상으로 한 연구에 본인의 가족을 초청하려는 의도를 HBOC 알림 의도로 조작화한 선행연구(M. C. Katapodi et al., 2017)에 근거해, 본 연구에서는 K-CASCADE 기초설문 중 ‘본 연구에 (가족 구성원을) 초청할 의향이 있으십니까?’ 라는 문항에 1명 이상 초청하겠다고 응답했을 경우 가족 알림 의도가 있다고 측정하였다.

## II. 문헌고찰

국내외에서는 HBOC의 가족 알림에 대한 연구가 제한적으로 이루어져 왔다. 특히, 알림 의도를 측정하는 것은 행동보다 더 어려워, 이에 대한 연구보다는 가족 알림 자체에 주로 초점이 맞춰졌다. 그러나 행동의 의도는 실제 행동에 선행되는 의미 있는 요소로 간주되며(Chatzisarantis et al., 2019; Sohn & Lee, 2012), 이에 따라 의도와 심리사회적 특성 등의 변수의 탐색이 중요시되고 있다(Chuenphitthayavut et al., 2020).

이에 본 문헌고찰은 가족 알림과 가족 알림 의도에 대한 고찰을 병행함으로써 가족 알림의 심리사회적 특성과 의도 형성에 영향을 미치는 다양한 측면을 보다 상세하게 이해하고자 한다.

### 2.1. HBOC의 가족 알림 및 영향요인

선행연구에서 국내 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 사람 106 명을 대상으로 가족에게 병인성 변이가 있음을 알린 사람은 99 명(93%)이었고 알리지 않은 사람과 비교 시 유전성 유방암 지식도 차는 유의하지 않았다(Kang et al., 2011). 국내 다른 연구에서는 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 난소암 환자 131 명 대상의 연구에서 130 명(99%)가 가족과 유전자검사 결과를 알렸다(정고운, 2022). 터키의

연구에서는 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 사람 92 명 중 87 명 (94.5%) 이 가족에게 알림을 했다(Seven et al., 2022). 연구별 대상자가 일치하지는 않으나, 국내외에서 보인자 대부분은 가족에게 유전성 변이 결과를 알렸음을 알 수 있다. 그러나 이들 연구에서 가족 알림 비율이 높았던 것은 Kang et al. (2011)과 Seven et al. (2022)의 연구에서는 유전성 유방암 대상자를 만나는 과정과 연구 동의 과정에서 선택 편향 가능성이 있다. 정고운 (2022) 연구의 유전성 난소암 대상자는 Olaparib 치료 여부가 중요하고, 이 치료법을 사용하기 위해서는 유전자 검사 상 *BRCA* 변이 확인이 필요하다(Kim, 2016). 이에 대부분의 대상자가 유전자 검사를 시행하여 가족 알림이 잘 이루어지는 환경이 조성되었을 수 있어 해석 및 HBOC 연구결과와 비교 시 신중을 기해야 할 것이다.

또 Kang et al. (2011) 연구에서 자녀의 연령에 관계없이 자녀에게 알린 비율은 36%(n=36)이었다. 이는 부모가 4-24 세 자녀에게 병인성 변이를 알림에 대한 미국에서의 선행연구 55%(n=23) (Bradbury et al., 2007)보다 낮았다. 두 연구의 직접적 비교에 신중을 기해야 하나 국내 보인자의 경우 자녀에게 알리는 비율이 미국에 비해 낮은 편이라고 추정할 수 있다.

미국의 경우 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 사람 혹은 암 병력이 있는 *BRCA1* 또는 *BRCA2* 보인자를 가족으로 둔 45 명을 대상으로

수행된 인터뷰에서 가족 중 이미 HBOC가 있어 HBOC에 대해 인지하고 있는 경우 가족 환경에서 HBOC에 노출되었기 때문에 가족 알림으로 쉽게 이어지는 경향이 있으며, 본인이 인지하는 가족 알림과 관련된 위험성과 HBOC가 가족과의 관련성이 있다고 인식하는 것이 가족 알림에 중요하다고 하였다(Dean & Rauscher, 2018). 이는 *BRCA* 또는 *PALB2*, *CHEK2*, *ATM* 유전자 변이를 가지고 있는 미국인 여성 62명을 대상으로 한 질적 연구에서는 HBOC가 가족과 관련성이 없다고 생각하거나 가족에게 알렸을 경우 가족 반응에 대한 걱정을 할 때 가족 알림이 저하된다는 보고와 맥을 같이 한다(Dean et al., 2021). 반면 터키에서 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 사람을 대상으로 가족에게 위험성이 있다고 인지하는 것이 가족 알림에 미치는 영향을 검증했을 때 통계적으로 유의한 영향은 확인되지 않아(Odd ratio 3.61,  $p=.057$ )(Seven et al., 2022) 더 연구가 필요한 상황이다.

가족 친밀감 또한 가족 알림을 촉진시키는 요인으로 미국(Dean et al., 2021)과 터키에서 수행된 연구(Seven et al., 2022)를 통해 보고되었다. 즉 가족과 정서적으로 친밀할수록 가족 알림 수행이 유의했다.

한편 부모-자녀 관계인 경우의 가족 알림에 대해서는 미국에 거주하는 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 32가족(부모 42명과 그의 4-24세 사이 자녀 86명)을 대상으로 한 연구에서 자녀의 연령이 높을수록

가족 알림의 경향이 높았다(Bradbury et al., 2007). 이는 터키에서 수행된 연구에서 가족 구성원의 나이가 많을수록 가족 알림의 경향이 더 높음으로 확인된 결과 방향과 일치한다(Seven et al., 2022).

호주의 유전 관련 전문가 인터뷰 연구에서는 HBOC 부모가 자녀에게 알릴 때, 부모의 암 진단과 같은 스트레스 상황을 일부 전문가는 자녀 알림을 촉진시킬 수 있는 요인으로 보았고, 일부 전문가는 자녀에게 스트레스 사건을 투사할 수 있어 저해할 수 있다고 하였다(Young et al., 2020). 반면 유전자 변이 결과가 스트레스 상황이라는 하나 가족성 암에 대해 이해할 수 있는 계기가 되어 가족 알림을 촉진하는 요인으로 본 연구도 있다(Dean et al., 2021).

유전자 검사에 대한 만족도도 가족 알림에 유의한 것으로 보고된다. *BRCA1* 또는 *BRCA2* 병인성 변이 유전자 검사를 시행한 유대인들을 대상으로 시행한 설문연구에서 유전자 검사에 대한 만족도가 높은 사람들이 더 가족에게 알림을 확인하였다(Lieberman et al., 2018). 이는 앞선 Young et al. (2020)에서 부모가 본인의 유전자 검사에 대한 만족도가 자녀에게 병인성 변이를 알릴 때 중요하다는 진술과 일치한다.

그밖에 HBOC 보인자가 유전자 검사 결과를 복잡하다고 생각하는 경우 가족 알림에 저해요소로 보고되며(Dean et al., 2021), 교육을 많이 받은 경우 가족 알림이 더 활발하게 이루어진다고 보고하였다(Seven et al.,

2022). 그러나 Bradbury et al. (2007)의 연구에서는 대학교 이상의 학력을 가진 자보다는 고등학교 졸업까지 한 자가 가족 알림의 경향이 높다고 밝혀 유전자 검사 결과를 이해하여 가족에게 알리는데 있어서 정규 교육과정의 영향 여부를 규명할 필요가 있다.

*BRCA1* 또는 *BRCA2* 병인성 변이가 있는 부모 중 예방적 유방절제술 또는 예방적 난소난관절제술을 시행한 경우 그렇지 않은 부모에 비해 자녀에게 가족 알림의 경향이 높았으며, 부모의 암 병력은 관계가 없었다(Bradbury et al., 2007).

*BRCA1* 또는 *BRCA2* 검사를 받은 사람 대상 Lafrenière et al. (2013)의 질적 연구에서 본인이 병인성 변이를 가족에게 알리기 위한 의사결정을 할 때 환자 개인의 특성과 가족 구성원의 특성, 가족과의 접근성과 관계의 질, 가족 문화, 과거 본인 혹은 가족의 암 경험 외에도 사회-의학적 맥락이 중요하다고 진술하였다. 그러나 현재까지 사회-의학적 맥락을 양적으로 검토한 연구가 제한적이며, 이에 대한 추가적인 연구와 검토가 필요하다.

종합적으로, HBOC의 가족 알림에 영향을 미치는 다양한 요인들이 존재하나 국내 연구에서는 가족 알림에 대한 심층적이고 체계적인 탐구가 충분하지 못하다. 특히, 국내에서의 다양한 변수, 환경과 문화의 영향을

고려한 예측모델을 구현하는 것이 필요하다. 이는 HBOC의 가족 알림을 더욱 정확하게 이해하는데 기여할 것으로 사려된다.

## 2.2. HBOC의 가족 알림 의도 및 영향 요인

20-45 세에 유방암을 진단 받은 30-63 세의 미국 거주 젊은 유방암 생존자를 대상으로 한 선행연구에서(M. C. Katapodi et al., 2017). 젊은 유방암 생존자 859 명 중 28.9%만이 유전자 검사를 받았기에 HBOC 대상자와는 다소 다르다고 볼 수 있으나, 젊은 유방암 생존자의 가족도 가족성 유방암의 고위험 군으로 유전자 검사 권고 대상이 된다. 따라서 이들의 가족 알림 의도는 HBOC의 가족 알림 의도와 다소 유사한 특성을 띠 수 있다. M. C. Katapodi et al. (2017)의 연구에서도 연구에 가족을 초청할 의도를 가족 알림 의도로 조작화 하였고 의도를 LASSO 방법으로 변수를 선택하여 포아송 회귀분석(poisson regression)을 시행한 결과 가족의 크기가 클수록, 의료진보다 가족 구성원의 제안을 따르고자 하는 동기가 있을 때 가족 알림 의도가 더 높았다. 또 회귀분석 상에서 가족 알림 의도는 인종간 차이는 없었고 질병에 대한 가족의 지지와 가족 결속력도 차이를 보이지 않았으며, 유방암 위험요소에 대한 지식 수준, 보험 여부, 의료비 부담, 교육수준, 소득수준, 결혼 여부에서도 차이가 없는 것으로 확인되었다. 그러나 이 연구는 HBOC만을 대상으로 하지

않았고, 유전자 검사를 시행하지 않은 젊은 유방암 생존자도 포함되어 있었으며, 알림 의도 예측이 아닌 가족 알림 의도 영향요인을 확인하는데 그쳤다.

HBOC의 가족 알림 의도에 대한 연구는 선행연구가 적으나 최근 한 연구에서 LASSO를 활용한 로지스틱 회귀분석으로 가족 알림 의도 영향요인을 분석하였다(Sarki et al., 2022). 스위스 거주 HBOC 보인자와 가족을 대상으로 한 이 연구는 연구에 가족을 초청할 의도로 가족 알림 의도를 조작화하였고, 가족원의 나이와 교육수준, 유전자 변이 여부, 암 유무, 암 진단 여부, 암 유전 관련 이해도(genetic affinity), 결혼 유무, 질병에 대한 가족의 지지가 LASSO 방법으로 변수 선택되었다. 그 결과 나이가 많을수록, 교육 수준이 높을수록, 암을 진단받았던 적 있을수록, 암 유전 관련 이해도가 높을수록 가족 알림 의도가 높다고 하였으나 통계적으로 유의하지는 않았다. 반면 유전자 변이가 있을수록, 질병에 대한 가족의 지지가 높을수록 가족 알림 의도가 높다고 하였고 이는 통계적으로 유의한 차이가 있었다. 그러나 연구 대상자가 HBOC 보인자와 그 혈연관계 가족들로, 전체 339명 중 HBOC 보인자는 238명, 나머지 82명은 HBOC의 가족이었기에, 유전자 검사를 시행하지 않았거나 병인성 변이가 없는 대상자가 포함되었다. LASSO를 활용한 로지스틱 회귀분석을 진행하기는 하였으나 가족 알림 의도 예측이 목표가 아니었기 때문에 다양한 분석

방법을 활용하지 않았으며, 회귀분석의 예측모델 관련 평가 지표가 제시되지 않은 제한점이 있다.

이상의 문헌고찰을 통해 HBOC의 가족 알림 의도에 포괄적으로 개인의 교육수준과 같은 일반적 특성과 예방적 유방절제술과 예방적 난소난관절제술 시행과 같은 암 관련 병력, 유전자 검사 관련 정보의 복잡성을 포함한 유전자 서비스에 대한 이해, 유전자 검사 시행 이유, 가족 결속력이나 질병에 대한 가족의 지지, 가족의 형태 등을 포함한 가족 환경, 개인의 암 관련 지식 수준이나 스트레스 상황 속 대처하는 개인의 태도 등을 포함한 신념과 태도, 자기관리와 같은 변수가 영향을 미칠 수 있음이 보고되었다. 가족 알림 의도를 양적으로 규명하고자 한 연구가 있었으나, HBOC만을 대상으로 하는 것이 아니라 유방암 환자를 대상으로 할 때 유전자 변이가 없는 유방암 환자를 포함하거나, HBOC와 그의 가족을 대상으로 진행했었다. 더불어, 다양한 통계 분석을 사용하지 않고 한가지 통계 기법을 사용해 분석했다는 한계점이 있다. 따라서 가족 알림 의도에 영향을 미치는 요인을 탐색하고, 가족 알림 의도를 예측할 수 있는 모델들을 구현하고, 구축된 최적의 모델에서 각각의 변수가 얼마나 설명력을 가지는지 탐색해 볼 필요가 있다.

## III. 연구방법

### 3.1. 연구 설계

본 연구는 HBOC 가족의 알림 의도 예측인자를 규명하기 위한 예측모형 설계(Predictive modeling design)이다. 가족 알림 의도가 있는 군의 특성을 파악하고 머신러닝(machine learning)을 기반으로 가족 알림 의도 예측모형을 개발하였다.

예측모형 개발을 위하여 Knowledge Discovery in Databases (KDD, Glymour et al., 1996) 방법을 활용하였다. 보편적인 데이터 마이닝(Data mining) 방법 중 하나인 KDD 방식을 의학 관련 연구에서도 사용할 수 있다(Mavrogiorgou et al., 2021; McClelland, 2010). KDD 방식은 데이터로부터 적절한 지식을 발견하기 위한 방법으로, 적절한 데이터 선택(Data selection), 데이터 정제하는 전처리(Data pre-processing), 데이터 사용할 기법에 맞게 변환(Data transformation), 적절한 머신러닝 방법을 통한 데이터 마이닝, 만들어진 모형에 대한 해석과 평가(Model interpretation and evaluation)의 과정을 상호작용하며 반복적으로 수행된다(Glymour et al., 1996) (Fig 1).

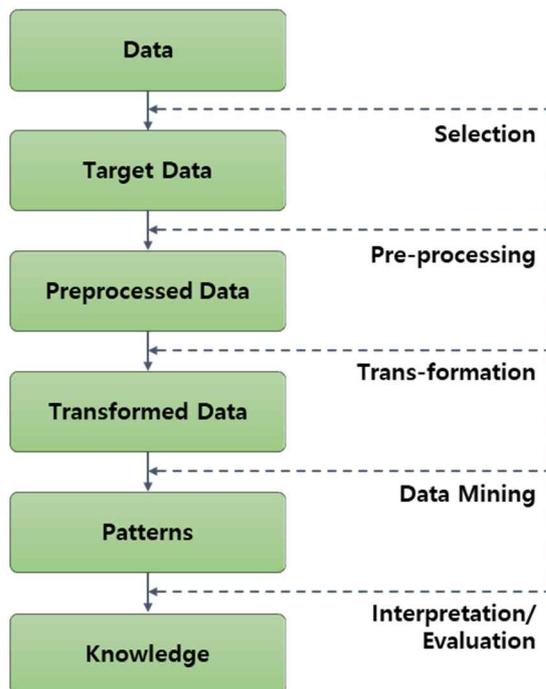


Figure 1. Knowledge Discovery in Databases process

### 3.2. K-CASCADE 연구

본 연구는 HBOC 가족을 대상으로 한 한국 유전성 암  
 케이스케이드(가족 연계) 유전자 조기검진 연구(K-CASCADE) 전향적 코호트  
 (<https://kcascade.kr/>; Kim et al., 2021) 자료를 활용하였다. K-  
 CASCADE는 한-스위스 국제협력 연구로 HBOC 중 *BRCA 1* 또는 *BRCA 2* 병인성  
 변이가 있는 자(index case)와 해당 보인자가 초청하는 혈연관계 가족이  
 참여하는 코호트 설문 연구이다. 서울특별시 소재의 4개

상급종합병원(세브란스병원, 강남세브란스병원, 삼성서울병원, 서울아산병원)과 경기도 소재의 1 개 공공병원(국립암센터)의 유방외과 혹은 부인암 외래를 통해 2021 년 3 월부터 대상자를 상시 등록(ongoing enrollment)하고 있다.

K-CASCADE 연구는 참여자 선호에 따라 종이 설문지 혹은 온라인 설문 링크와 보안코드가 제공되었으며, 스위스 CASCADE 코호트 연구 (<https://swisscascade.ch/en/>; Maria. C. Katoapodi et al., 2017)와 동일한 248 문항의 기초설문(baseline survey)을 사용하였다. 온라인 설문은 Lime survey 를 통해 스위스 연구팀과 동일한 시스템으로 구축되었으며, 참여자들의 응답은 보안코드를 사용하여 익명화 되었다. 기초설문 작성에 약 40 분 소요되었으며 참여자에게는 모바일 쿠폰(1 만원 상당 음료권)이 지급되었다.

### 3.3. 연구 대상자

K-CASCADE 의 선정기준은 *BRCA1* 또는 *BRCA2* 병인성 변이가 있고, 만 19 세 이상으로 국내에 거주하는 자이며, 혈연 관계인 가족이 있는 자이다. 제외 기준은 자가 보고로 정신질환이 있거나 컴퓨터나 모바일 접근이 어려운 경우, 한글을 읽고 쓰지 못하는 경우이다. 이 중 본 연구에 포함된

대상자는 *BRCA1* 또는 *BRCA2* 병인성 변이가 있는 보인자로 암이 발병한 적 있는 자를 대상으로 하였다.

2021년 3월부터 2023년 9월 30일까지 K-CASCADE 연구에서 확인된 잠재적 대상자 648명을 초청하였고, 이중 연구 참여에 동의한 615명에게 기초설문을 배부하여 298명의 설문이 회수되었다. 이 중 결측치가 있는 자료를 제외하여 219명의 자료를 본 연구에 활용하였다.

### 3.4. 연구 변수

본 연구는 K-CASCADE 기초설문 248 문항 중 병인성 변이 보인자가 아닌 경우 응답해야하는 문항과 대상자 개별 치료 상황에 따라 응답 여부가 달라지는 문항은 제외하고, 148 문항을 활용하였다. 148개 문항은 <Table 1>과 같이 총 50개의 변수로 변환하여 분석하였다.

가족 알림 의도는 K-CASCADE 설문 중 ‘본 연구에 (가족 구성원을) 초청할 의향이 있으십니까?’ 라는 문항에 1명 이상 초청 의향이 있다고 응답한 경우 가족 알림 의도가 있음으로, 아무도 초청하지 않겠다고 응답한 경우 가족 알림 의도가 없음으로 측정했다.

독립변수는 일반적 특성과 암 병력, 유전자 서비스, 유전자 검사를 시행한 이유, 자기관리, 가족 환경, 의학적 환경, 신념과 태도 범주로 분류하였다(Table 1).

일반적 특성으로 12 개 변수로 구성하였다. 나이는 데이터를 수집한 시기의 나이를 숫자형으로 측정하였으며, 결혼 유무는 데이터를 수집한 시기에 결혼 상태 여부를 이분형으로(yes/no) 측정하였다. BMI 는 한국 표준체중 기준으로  $23\text{kg}/\text{m}^2$  이하와 초과, 교육수준은 대학교 졸업 이상과 미만, 직업은 전일제 근무 여부로 측정하였다. 흡연은 흡연해 본적 없는 자와 과거나 현재 흡연을 한 자로 구분하여 측정하였다. 음주는 지난 4주간 음주를 해본 적 없는 자와 음주를 해본 자로 측정하였다. 신체활동은 지난 4주간 운동을 하지 않은 자와 해본 자로 측정하였다. 임신 경험은 임신을 안해본 자와 1회 이상 해본 자로 측정하였다. 유방 조직검사는 유방 조직검사를 받아본 자와 받아본 적 없는 자로 측정하였다. 호르몬요법은 폐경기 호르몬 요법 치료를 받은 적 있는 자와 치료 받은 적 없는 자로 측정하였다. 경구 피임약은 경구 피임약을 복용해본 적 있는 자와 복용해본 적 없는 자로 측정하였다.

암 병력 관련 5 가지 변수로 구성하였다. 암 첫 진단 나이는 첫 번째로 암을 진단 받았을 때의 나이를 숫자형으로 측정하였다. 유방암 진단은 유방암을 진단 받은 적 있는 자와 진단 받은 적 없는 자로, 난소암 진단은 난소암을 진단 받은 적 있는 자와 진단 받은 적 없는 자로 측정하였다. 예방적 유방절제술은 유방암 발생 전 유방절제술을 시행한

자와 시행하지 않은 자로, 예방적 난소난관절제술은 난소암 발생 전 난소난관절제술을 시행한 자와 시행하지 않은 자로 측정하였다.

유전자 서비스(genetic services)는 유전자 검사 후회(decisional regret)와 유전자 검사에 대한 유용성으로 구성하였다. 유전자 검사 후회는 5개의 문항으로, 각 문항은 7점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 매우 그렇다)로 측정하여 평균점수를 활용하였으며, 점수가 높을수록 유전자 검사에 대한 후회 정도가 낮음을 의미한다. 인지된 유전자 검사 유용성은 혈연관계 가족이 유전자 검사를 하는 것이 그들에게 도움이 된다고 생각하는지를 7점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 매우 그렇다)로 측정하였으며, 점수가 높을수록 가족이 유전자 검사를 하는 것이 가족에게 도움이 될 거라고 생각함을 의미한다.

유전자 검사를 시행한 이유는 총 11개 문항으로 각 문항은 7점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 매우 그렇다)로 측정하였다. 암 위험성은 점수가 높을수록 본인의 미래 암 위험성에 대해 알고 싶어서 유전자 검사를 시행했음을 의미한다. 유전자 변이 확인은 점수가 높을수록 본인이 암과 관련된 유전자 변이를 가지고 있는지 알고 싶어서 유전자 검사를 시행했음을 의미한다. 가족의 이익은 점수가 높을수록 본인의 가족에게 도움이 될 수 있다고 생각해서 유전자 검사를 시행했음을 의미한다. 의사의 권유는 점수가 높을수록 의사가 검사를 해야한다고

권해서 유전자 검사를 시행했음을 의미한다. 암 치료 계획은 점수가 높을수록 유전자 검사 결과에 따라 본인의 암 치료 계획이 달라질 수 있기 때문에 유전자 검사를 시행했음을 의미한다. 자녀의 위험성은 점수가 높을수록 본인 자녀들에게 미치는 위험을 알고 싶어 유전자 검사를 시행했음을 의미한다. 가족의 바람은 점수가 높을수록 가족들이 유전자 검사하는 것을 원하여 유전자 검사를 시행했음을 의미한다. 의료보험 적용은 점수가 높을수록 의료보험으로 검사 비용을 낼 수 있어서 유전자 검사를 시행했음을 의미한다. 가족 계획은 점수가 높을수록 유전자 검사 결과에 따라서 가족 계획을 바꿀수도 있어 유전자 검사를 시행했음을 의미한다. 병인성 변이 보인자 가족은 점수가 높을수록 가족 중 암과 관련된 유전자 변이를 가지고 있는 사람이 이미 있어서 유전자 검사를 시행했음을 의미한다. 가까운 거리는 점수가 높을수록 유전자 검사 센터가 집과 가까워 유전자 검사를 시행했음을 의미한다.

자기관리(self-management)는 암 관련 자기 효능감과 재발 관련 걱정으로 구성하였다. 암 관련 자기 효능감은 총 14 개 문항, 각 문항은 7 점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 매우 그렇다)로 측정하여 평균평점을 활용하였으며, 점수가 높을수록 암 진단에 대해 자신 있게 대처하여 자기 효능감이 높음을 의미한다. 암 재발 관련 걱정은 총 4 개 문항, 각 문항은 7 점 리커트 척도(1 점: 전혀 - 7 점: 항상; 1 점: 전혀 -

7 점: 매우 많이)로 측정하여 평균평점을 활용하였으며, 점수가 높을수록 암 재발에 대해 걱정이 많음을 의미한다.

가족 환경(family environment)은 7 가지 변수로 구성하였다. 가족 결속력(family cohesion)은 총 15 개 문항을 7 점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 항상 그렇다)로 측정하여 평균평점을 활용하였으며, 점수가 높을수록 가족 결속력이 높음을 의미한다. 질병에 대한 가족의 지지는 총 15 개 문항, 각 문항은 7 점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 항상 그렇다)로 측정하여 평균평점을 활용하였으며, 점수가 높을수록 본인의 건강에 문제가 생겼을 때 본인의 가족이 더 많은 지지를 해줄 것이라고 생각함을 의미한다. 동거 가족에 대한 변수는 동거 가족 없이 혼자 사는지, 배우자나 파트너와 함께 동거하는지, 19 세 미만의 가족과 동거하는지, 19 세 이상의 자녀와 동거하는지, 부모와 함께 동거하는지, 각각 이분형으로(yes/no) 측정하였다.

의학적 환경(medical environment)은 3 가지 변수로 구성하였다. 의학적 의사결정 참여도(medical decision participation)는 총 4 개의 문항, 7 점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 항상 그렇다)로 측정하여 평균평점을 활용하였으며, 점수가 높을수록 최근 12 개월 동안 가장 자주 본 의료진으로부터 본인이 받은 치료에 대한 의사결정을 할 때 참여를 많이 했음을 의미한다. 의료비 지불 어려움은 7 점 리커트

척도(1 점: 전혀 없다 - 7 점: 항상 있다)로 측정하였고, 점수가 높을수록 최근 12 개월 동안 비용 때문에 의사에게 진찰이나 진료를 받지 못하였음을 의미한다. 매년 같은 의료진에게 건강검진 시행 여부는 한두 명의 의료진이 본인의 건강관리를 책임지는지 여부로 측정하였다.

신념과 태도(belief and attitude)는 7 가지 변수로 구성하였다. 암 관련 지식은 총 4 개의 문항으로, 각 문항에 정답을 맞출 경우 1 점, 틀리거나 모르겠다고 응답할 경우 0 점으로 측정하고 합산점수를 활용한다. 점수가(가능 점수 범위: 0-4) 높을수록 암 관련하여 지식 수준이 높음을 의미한다. 유전성 암 관련 지식은 1 개의 단일 문항으로 정답을 맞출 경우 1 점, 틀리거나 모르겠다고 응답할 경우 0 점으로 측정한다. 점수가 높을수록 유전성 암 관련 지식수준이 높음을 의미한다. 암 유전 관련 이해도(genetic affinity)은 암에 대한 암 유전학과 암 위험에 대해 알고 있다는 인식 정도를 의미하며(Pedrazzani et al., 2021) 총 2 문항으로, “암에 걸릴 가능성에 대해 얼마나 잘 알고 있다고 생각하십니까?” 와 “암의 유전학에 대해 얼마나 알고 있습니까?” 라는 질문에 7 점 리커트 척도(1 점: 전혀 모른다 - 7 점: 매우 잘 안다)로 평균평점을 측정하였다. 점수가 높을수록 유전친화력이 높음, 즉 암 관련 유전정보에 대한 이해도가 높음을 의미한다(Dar-Nimrod et al., 2019). 암 재발에 대한 신념은 10 점 리커트 척도(1 점: 절대 그렇지 않을 것이다 - 10 점: 반드시 그럴

것이다)로 측정하며, 점수가 높을수록 본인이 암에 걸릴 가능성이 있다고 더 많이 믿음을 의미한다. 스트레스 상황에 대처하는 방식은 문제 중심적 대처(coping: problem-focused, 8 문항) 감정 중심적 대처(coping: emotion-focused, 10 문항), 회피적 대처(coping: avoidant, 7 문항) 3 가지로 구분되며, 각 문항은 7 점 리커트 척도(1 점: 전혀 그렇지 않다 - 7 점: 항상 그렇다)로 측정하여 평균평점을 활용하였다. 문제 중심적 점수가 높을수록 스트레스 받는 상황에 대처할 때 문제를 해결하기 위한 방식을 취함을 의미하며, 감정 중심적 점수는 높을수록 스트레스 받는 상황에서 자신의 감정에 집중하여 대처하는 것을 의미한다. 회피적 점수는 높을수록 스트레스 받는 상황에서 스트레스 상황을 회피하기 위한 대처방식을 택함을 의미한다.

**Table 1. Independent Variables of the Study**

Category	Variables	Type /Level	Possible Range	Description
<b>General Characteristics</b>				
	Age	Numeric		
	Married	Y/N		
	BMI (ref=under 23)	Y/N		
	Education level (ref=Bachelor's degree or higher)	Y/N		
	Work (ref=full-time)	Y/N		
	Smoking (ref=never smoke)	Y/N		
	Drinking (ref=nondrinker)	Y/N		
	Physical activity (ref=Yes)	Y/N		
	Pregnancy experience (ref=Yes)	Y/N		
	Breast biopsy (ref=Yes)	Y/N		
	Menopausal hormone therapy (ref=Yes)	Y/N		
	Oral contraceptives (ref=Yes)	Y/N		
<b>Cancer History</b>				
	Age at first cancer diagnosis	Numeric		
	Breast cancer (ref=Yes)	Y/N		
	Ovarian cancer (ref=Yes)	Y/N		
	Prophylactic mastectomy (ref=Yes)	Y/N		
	Prophylactic salpingo-oophorectomy (ref=Yes)	Y/N		
<b>Genetic Services</b>				
	Decisional Regret	Numeric / 7-point Likert scale, point average.	1-7	5 items. A higher score means a lower decisional regret associated with genetic testing.
	Perceived GT usefulness	Numeric / 7-point Likert scale.	1-7	1 item. A higher score means a higher perceived usefulness for blood relatives to have genetic testing.

Category	Variables	Type /Level	Possible Range	Description
<b>Reason for Genetic test</b>				
	Risk of cancer	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because I wanted to know more about my future cancer risk"
	Check gene mutation	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because I wanted to know if I have a mutation connected to cancer"
	Family benefit	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because it may benefit my family"
	Physician advice	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because my physician advised that I should do it"
	Cancer treatment	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because results may change my cancer treatment"
	Risk of children	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because I wanted to know the risk for my children"
	Family's wish	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because my family members want me to"
	Insurance covered	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because my medical insurance covered the cost"
	Family planning	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because results may change my family planning"
	Carrier family	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because I already know someone in my family who has a mutation connected to cancer"
	Close distance	Numeric / 7-point Likert scale.	1-7	"I had genetic testing because the genetic testing center is close to home"
<b>Self-management</b>				
	Self-efficacy dealing with cancer	Numeric / 7-point Likert scale, point average.	1-7	A higher score means a higher self-efficacy in dealing with cancer.
	Recurrence concerns	Numeric / 7-point Likert scale, point average.	1-7	A higher score means more concerns about cancer recurrence.
<b>Family environment</b>				
	Family cohesion	Numeric / 7-point Likert scale, point average.	1-7	A higher score means a higher family coherence.
	Family support in illness	Numeric / 7-point Likert scale, point average.	1-7	A higher score means a higher family support in illness.

Category	Variables	Type /Level	Possible Range	Description
	Live alone (ref=Yes)	Y/N		
	Live with partner (ref=Yes)	Y/N		
	Live with child under 19-year-old (ref=Yes)	Y/N		
	Live with child over 19-year-old (ref=Yes)	Y/N		
	Live with parents (ref=Yes)	Y/N		
<b>Medical environment</b>				
	Medical decision participation	Numeric / 7-point Likert scale, point average.	1-7	A higher score means more involvement in the medical decisions of medical staff.
	Difficulty affording medical costs	Numeric / 7-point Likert scale.	1-7	A higher score means having more difficulty in paying for medical costs.
	Annual check-up with same staff (ref=Yes)	Y/N		Receive annual checkups and/or preventive screenings from the same health professionals or not.
<b>Belief and Attitude</b>				
	Knowledge of cancer	Numeric / 1 point for each item, sum-up score.	0-4	A higher score means higher knowledge of breast cancer risk factors.
	Knowledge of genetic cancer	Numeric / 1 item, 1-point.	0-1	Score 1 is known about genetic cancer.
	Genetic affinity	Numeric / 7-point Likert scale, point average.	1-7	A higher score means a higher genetic affinity.
	Belief of cancer recurrence	Numeric / 10-point Likert scale.	1-10	A higher score means higher belief of chances for getting other cancer.
	Coping: problem-focused	Numeric / 7-point Likert scale, point average.	1-7	A higher score means more problem-focused coping in stressful events.
	Coping: emotion-focused	Numeric / 7-point Likert scale, point average.	1-7	A higher score means more emotion-focused coping in stressful events.
	Coping: avoidant	Numeric / 7-point Likert scale, point average.	1-7	A higher score means more avoidant in stressful events.

### 3.5. 자료 전처리

결측치가 없는 219 명의 자료를 활용하였고 전체 298 명의 자료와 219 명의 자료의 변수 간의 차이가 있는지 독립 표본 T 검정과 카이제곱 검정을 시행하였고, 통계적으로 유의한 차이가 없음을 확인하였다(부록 1).

### 3.6. 자료 변환

본 연구의 예측모델은 가족 알림 의도 여부를 분류하기 위한 모델로, 분류를 원활히 하기 위하여 범주형 자료를 원-핫 인코딩(one-hot encoding)하여, 0 과 1 로 변환하였다(Bruce et al., 2021).

### 3.7. 자료 분석 방법

기술통계는 Stata(16.1 version)을, 예측모델은 R studio(4.1.2 version) 프로그램을 활용하여 분석했으며, 통계적 유의성을 평가할 경우 유의확률 0.05(양측검정)을 활용하였다. 자세한 자료 분석 방법은 아래와 같다.

첫째, HBOC 중 가족 알림 의도가 있는 군과 알림 의도가 없는 군의 특성을 변수에 따라 평균과 표준편차 또는 빈도와 백분율로 기술했으며,

가족 알림 의도 유무에 따른 차이를 확인하기 위해 독립 표본 T 검정과 카이제곱 검정을 시행했다.

둘째, HBOC의 가족 알림 의도 예측모델은 R studio의 SuperLearner 패키지(Polley et al., 2023)를 활용했다. SuperLearner는 라이브러리에 내장된 머신러닝(machine learning) 기법들을 조합하여 최적의 예측모델을 구현하는 패키지이다. Super Learner의 라이브러리 중 이분형 종속변수에 활용할 수 있는 방법 중 전통적인 변수 선택법을 통한 로지스틱 회귀분석(Logistic Regression)과 Lasso와 Ridge 방법을 통한 로지스틱 회귀분석(Logistic regression with Lasso/Ridge), 분류 모델을 위해 가장 널리 사용되는 트리(tree) 방법 중 하나인 랜덤 포레스트(Random Forest), 그리고 이 세 가지 머신러닝을 조합하는 앙상블 학습 기법(Ensemble learning)을 시행하였다. 교차 검증은 10-folds 교차 검증을 활용하였다.

HBOC의 가족 알림 의도 예측모델은 민감도(sensitivity)와 특이도(specificity), 정확도(accuracy), F1 점수(F1-score), 수신자 판단(Receiver Operating Characteristic, ROC) 곡선을 통해 성능을 평가하였다. F1 점수는 모델의 정확성을 측정하는 기계 학습 평가 지표 중 하나로, 데이터세트의 불균형에 민감한 민감도와 특이도를 보완하기 위하여 사용하였다(Sokolova et al., 2006). 민감도와 특이도, 정확도, F1 점수는 다음과 같이 계산하였다.

$$\begin{aligned}
 \text{Sensitivity} &= \frac{TP}{TP + FN} \\
 \text{Specificity} &= \frac{TN}{TN + FP} \\
 \text{PPV} &= \frac{TP}{TP + FP} \\
 \text{NPV} &= \frac{TN}{TN + FN} \\
 \text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN} \\
 \text{F}_1 \text{ score} &= 2 \times \frac{\text{Sensitivity} \times \text{PPV}}{\text{Sensitivity} + \text{PPV}}
 \end{aligned}$$

\* TP: True Positives, TN: True Negatives, FP: False Positives, FN: false negatives

양상블 학습 기법에 대한 추가분석으로 양상블 학습 기법에 활용된 3 가지 하위 모델과 양상블 학습 기법 모델의 MSE(Mean Standard Error) 값을 확인하였으며, 해당 값은 작을수록 모델의 성능이 좋음을 의미한다.

셋째, HBOC 의 가족 알림 의도 예측모델 중 가장 성능이 우수한 모델의 변수 중요도는 로지스틱 회귀분석의 경우 Z 점수의 절댓값이 클수록, 랜덤 포레스트의 경우에는 지니 평균 감소량(mean decrease of Gini)이 클수록 변수가 중요도가 높고 평가하였다.

## IV. 결과

### 4.1. 일반적 특성

연구에 활용한 219 명의 특성은 <Table 2>와 같다. 대상자는 모두 여성이고 한국인이었으며, 평균 연령은  $47.22 \pm 9.78$  세(range: 26-75), 기혼자는 164 명(74.89%)이다. BMI 표준체중 이하(23 이하)는 142 명(64.84%)이며, 대졸 이상의 교육수준을 가진 자는 162 명(73.97%)이다. 임신 경험이 있는 사람은 171 명(78.08%)이었으며, 유방 조직검사를 해본 사람은 171 명(78.08%)이다.

암 관련 병력으로 암 첫 진단 나이는  $41.65 \pm 9.49$  세였으며, 유방암 진단자는 159 명(72.60%), 난소암 진단자는 63 명(28.77%)이었으며, 유방암과 난소암 모두 진단받은 자는 6 명(2.74%)이었다. 예방적 유방절제술을 받은 사람은 41 명(18.72%)이고, 예방적 난소난관절제술을 받은 사람은 144 명(65.75%)이다.

유전자 서비스 관련하여 유전자 검사 후회 점수는  $6.06 \pm 1.10$  점(range: 1-7)으로 유전자 검사에 대한 후회 정도가 낮았다. 인지된 유전자 검사 유용성  $5.96 \pm 1.22$  점(range: 1-7)으로 높은 수준으로 대상자들은 유용하게 인지함을 확인하였다.

Table 2. Comparison of non-intention and intention groups (N=219)

	Total (N=219)	Non- intention (n=149) M±SD or n(%)	Intention (≥1) (n=70) M±SD or n(%)	Z or $\chi^2$	<i>p</i>
<b>General Characteristics</b>					
Age	47.22±9.78	46.31±9.63	49.16±9.89	-2.03	<b>.044*</b>
Married	164(74.89)	109(73.15)	55(78.57)	0.74	.389
BMI (ref=under 23)	142(64.84)	104(69.80)	38(54.29)	5.03	<b>.025*</b>
Education (ref=Bachelor's degree or higher)	162(73.97)	115(77.18)	47(67.14)	2.49	.114
Work (ref=full-time)	54(24.66)	40(26.85)	14(20.00)	1.20	.273
Smoking (ref=never smoke)	191(87.21)	134(89.93)	57(81.43)	3.09	.079
Drinking (ref=nondrinker)	199(90.87)	138(92.62)	61(87.14)	1.72	.190
Physical activity (ref=Yes)	201(91.78)	137(91.95)	64(91.43)	0.02	.896
Pregnancy experience (ref=Yes)	171(78.08)	110(73.83)	61(87.14)	4.94	<b>.026*</b>
Breast biopsy (ref=Yes)	171(78.08)	122(81.88)	49(70.00)	3.93	<b>.048*</b>
Menopausal hormone therapy (ref=Yes)	16(7.31)	11(7.38)	5(7.14)	0.01	.949
Oral contraceptives (ref=Yes)	46(21.00)	28(18.79)	18(25.71)	1.38	.241
<b>Cancer History</b>					
Age at first cancer diagnosis	41.65±9.49	40.95±9.65	43.14±9.04	-1.60	.111
Breast cancer (ref=Yes)	159(72.60)	113(75.84)	46(65.71)	2.45	.117
Ovarian cancer (ref=Yes)	63(28.77)	35(23.49)	28(40.00)	6.34	<b>.012*</b>
Prophylactic mastectomy (ref=Yes)	41(18.72)	27(18.12)	14(20.00)	0.11	.740
Prophylactic salpingo- oophorectomy (ref=Yes)	144(65.75)	96(64.43)	48(68.57)	0.36	.547
<b>Genetic Services</b>					
Decisional Regret	6.06±1.10	5.93±1.16	6.34±0.91	-2.63	<b>.009**</b>
Perceived GT usefulness	5.96±1.22	5.78±1.33	6.34±0.85	-3.26	<b>.001**</b>
<b>Reason for Genetic test</b>					
Risk of cancer	6.03±1.30	5.97±1.30	6.16±1.32	-0.97	.331
Check gene mutation	5.99±1.39	5.93±1.42	6.11±1.30	-0.90	.367
Family benefit	5.96±1.27	5.89±1.24	6.11±1.30	-1.25	.213
Physician advice	5.90±1.35	5.84±1.38	6.03±1.30	-0.97	.335
Cancer treatment	5.88±1.51	5.80±1.52	6.04±1.50	-1.12	.266
Risk of children	5.51±1.80	5.38±1.84	5.79±1.69	-1.58	.116

	Total (N=219)	Non- intention (n=149) M±SD or n(%)	Intention (≥1) (n=70) M±SD or n(%)	Z or $\chi^2$	<i>p</i>
Family's wish	4.82±2.01	4.61±2.02	5.26±1.92	-2.24	<b>.026*</b>
Insurance covered	4.45±2.29	4.34±2.27	4.67±2.32	-0.99	.322
Family planning	3.74±2.34	3.68±2.33	3.87±2.38	-0.57	.570
Carrier family	3.29±2.40	3.34±2.41	3.17±2.38	0.49	.624
Close distance	2.58±1.94	2.57±1.90	2.60±2.03	-0.11	.917
<b>Self-management</b>					
Self-efficacy dealing with cancer	4.99±1.12	4.89±1.11	5.19±1.12	-1.83	.069
Recurrence concerns	4.80±1.33	4.72±1.39	4.98±1.19	-1.32	.189
<b>Family environment</b>					
Family cohesion	5.52±1.04	5.45±1.08	5.67±0.93	-1.43	.153
Family support in illness	5.83±1.03	5.76±1.08	5.98±0.91	-1.46	.145
Live alone	18(8.22)	14(9.40)	4(5.71)	0.86	.355
Live with partner	160(73.06)	107(71.81)	53(75.71)	0.37	.544
Live with child under 19-year-old	83(37.90)	59(39.60)	24(34.29)	0.57	.450
Live with child over 19-year-old	67(30.59)	39(26.17)	28(40.00)	4.29	<b>.038*</b>
Live with parents	35(15.98)	26(17.45)	9(12.86)	0.75	.387
<b>Medical environment</b>					
Medical decision participation	5.62±1.18	5.52±1.16	5.84±1.21	-1.90	.059
Difficulty affording medical costs	1.58±1.13	1.65±1.28	1.41±0.71	1.45	.150
Annual check-up with same staff (ref=Yes)	166(75.80)	109(73.15)	57(81.43)	1.78	.182
<b>Belief and Attitude</b>					
Knowledge of cancer	2.79±1.35	2.77±1.35	2.83±1.36	-0.32	.747
Knowledge of genetic cancer	0.68±0.47	0.66±0.47	0.70±0.46	-0.52	.602
Genetic affinity	3.51±1.28	3.43±1.22	3.66±1.39	-1.27	.206
Belief of cancer recurrence	6.33±2.24	6.24±2.25	6.53±2.21	-0.91	.366
Coping: problem-focused	5.07±0.92	5.01±0.91	5.19±0.92	-1.36	.176
Coping: emotion-focused	3.78±0.66	3.74±0.67	3.85±0.64	-1.09	.277
Coping: avoidant	2.74±0.70	2.75±0.73	2.71±0.65	0.38	.707

\* $p < 0.05$ , \*\* $p < 0.01$ .

유전자 검사 시행한 이유로 높은 이유가 되었던 순서대로, 미래 암 위험성을 확인하고 싶어서  $6.03 \pm 1.30$  점(range: 1-7), 암과 관련된 유전자 변이 확인을 위해서가  $5.99 \pm 1.39$  점(range: 1-7), 가족에게 도움이 될 것 같아서  $5.96 \pm 1.27$  점(range: 1-7), 의사의 권유는  $5.90 \pm 1.35$  점(range: 1-7), 미래 암 치료 계획 때문에는  $5.88 \pm 1.51$ (range: 1-7), 자녀에게 미치는 위험성 확인을 위해서는  $5.51 \pm 1.80$  점(range: 1-7), 가족의 바람 때문에는  $4.82 \pm 2.01$  점(range: 1-7), 검사 비용의 의료보험 적용이 되어서는  $4.45 \pm 2.29$  점(range: 1-7)이었다. 유전자 검사를 시행한 이유 중 낮은 이유로는 출산 계획 때문에는  $3.74 \pm 2.34$  점(range: 1-7), 유전자 변이 있는 가족이 있기 때문에는  $3.29 \pm 2.40$  점(range: 1-7), 유전자 검사 센터와의 거리가 가까워서  $2.58 \pm 1.94$  점(range: 1-7)이었다.

자기관리 관련 항목으로 암 관련한 자기 효능감은  $4.99 \pm 1.12$  점(range: 1-7)이고, 암 재발에 대한 걱정은  $4.80 \pm 1.33$  점(range: 1-7)으로 둘다 보통 정도의 수준을 보인다.

가족 환경으로 가족 결속력은  $5.52 \pm 1.04$  점(range: 1-7)이고, 질병에 대한 가족의 지지는  $5.83 \pm 1.03$  점(range: 1-7)으로 높은 편이었다. 혼자 사는 사람은 18 명(8.22%)으로 낮은 비율을 보였다. 동거 가족이 있는 사람들 중에 파트너와 함께 사는 사람은 160 명(73.06%), 19 세 미만 자녀와

사는 사람은 83 명(37.90%), 19 세 이상 자녀와 사는 사람은 67 명(30.59%), 부모와 같이 사는 사람은 35 명(15.98%)이었다.

의학적 환경에서 의학적 의사결정 참여도(Medical decision participation)는  $5.62 \pm 1.18$  점(range: 1-7)으로 높은 편이며, 의료비 지불 어려움은  $1.58 \pm 1.13$  점(range: 1-7)으로 매우 낮은 편이었다. 매년 같은 의료진에게 검진을 받는 사람은 166 명(75.80%)이다.

태도와 신념에서 암 관련 지식은  $2.79 \pm 1.35$  점(range: 0-4), 유전성 암 관련 지식은  $0.68 \pm 0.47$  점(range: 0-1)으로 높은 편이다. 암 유전 관련 이해도는  $3.51 \pm 1.28$  점(range: 1-7)이고, 암 재발 가능성에 대한 믿음은  $6.33 \pm 2.24$  점(range: 1-10)으로 둘 다 보통 수준을 보인다. 스트레스 상황에 대처하는 기전으로 문제 해결 중심적 대처는  $5.07 \pm 0.92$  점(range: 1-7)으로 높은 수준을 보였고, 감정 중심적 대처는  $3.78 \pm 0.66$  점(range: 1-7)으로 보통 수준이었으며, 회피적 대처는  $2.74 \pm 0.70$  점(range: 1-7)으로 낮은 수준이었다.

## 4.2. 가족 알림 의도 집단 별 비교

가족 알림 의도가 있는, 즉 K-CASCADE 연구에 가족을 초청할 의도가 있는 군은 총 70 명(31.96%)이다(Table 2). 가족 알림 의도가 있는 군과

없는 군은 나이, BMI, 임신 여부, 유방 조직검사 여부에서 통계적으로 유의한 차이가 있었다.

가족 알림 의도가 있는 군은  $49.16 \pm 9.89$  세, 가족 알림 의도가 없는 군은  $46.31 \pm 9.63$  세로, 가족 알림 의도가 있는 군이 통계적으로 유의하게 나이가 많았다( $p=.044$ ). BMI 표준체중 이하인 비율은 가족 알림 의도가 있는 군( $n=38$ , 54.29%)이 가족 알림 의도가 없는 군( $n=104$ , 69.80%) 보다 높았다( $p=.025$ ). 임신 여부도 가족 알림 의도가 있는 군이( $n=61$ , 87.14%) 가족 알림 의도가 없는 군 ( $n=110$ , 78.08%) 보다 높았다( $p=.026$ ). 유방 조직검사 경험자는 가족 알림 의도가 없는 군( $n=122$ , 81.88%)이 알림 의도가 있는 군( $n=49$ , 70.00%) 보다 높았다( $p=.048$ ). 이외 결혼 여부, 교육수준, 정규직 직업 여부, 흡연 여부, 음주 여부, 신체적 활동 여부, 호르몬 치료 여부, 경구 피임약 복용 여부는 두 군 간의 통계적으로 유의한 차이가 없었다.

암 관련 병력으로 가족 알림 의도가 있는 군의 난소암 진단자는 28명(40.00%), 가족 알림 의도가 없는 군은 35명(23.49%)으로 가족 알림 의도가 있는 군의 난소암 비율이 높았다( $p=.012$ ). 그러나 처음 암을 진단받은 나이나 유방암 진단 비율, 예방적 유방절제술, 예방적 난소난관절제술 시행 비율은 두 군 간의 유의한 차이는 없었다.

유전자 서비스 관련하여 유전자 검사 후회 척도에서 가족 알림 의도가 있는 군은( $6.34 \pm 0.91$  점) 가족 알림 의도가 없는 군( $5.93 \pm 1.16$  점)에 비해 통계적으로 유의하게 높아( $p=.009$ ), 유전자 검사에 대한 후회가 적은 것으로 확인하였다. 인지된 유전자 검사 유용성은 가족 알림 의도가 있는 군이( $6.34 \pm 0.85$  점) 가족 알림 의도가 없는 군( $5.78 \pm 1.33$  점)에 비해 통계적으로 유의하게 높아( $p=.001$ ) 유전자 검사를 더 유용하게 인식하고 있었다

유전자 검사를 시행한 이유 중 가족의 바람으로 유전자 검사를 시행한 경우만 가족 알림 의도가 있는 군의 점수( $5.26 \pm 1.92$  점)가 가족 알림 의도가 없는 군( $4.61 \pm 2.02$  점)에 비해 높았다( $p=.026$ ). 이외 자녀에게 미치는 위험성을 확인하거나 암과 관련된 유전자 변이 확인하고 싶어서, 미래에 암 위험성 확인하기 위해, 가족 중 유전자 변이 보인자가 있거나, 출산계획을 위해, 가족에게 유익할 것 같아서, 의사가 권유해서, 유전자 검사 센터와의 거리가 가까워서, 검사 비용의 의료보험 적용되기 때문에, 앞으로 암 치료 계획에 활용하기 위해서 유전자 검사를 시행하였다고 응답한 점수는 두 군 간에 유의한 차이는 없었다.

자기관리와 관련된 변수로, 암 관련한 자기 효능감과 암 재발에 대한 걱정은 가족 알림 의도에 따라 유의한 점수 차이가 없었다.

가족 환경에서 가족 알림 의도가 있는 군은 19 세 이상 자녀와 함께 거주하는 경우가 40%로, 가족 알림 의도가 없는 군의 26.17%에 비해 더 높았다( $p=.038$ ). 두 군 간의 가족 결속력과 질병에 대한 가족의 지지에 대한 유의한 점수 차이는 없었으며, 혼자 살거나, 파트너 또는 19 세 미만의 자녀, 부모와 같이 사는 사람의 비율 차도 나지 않았다.

의학적 환경에 대한 변수, 의학적 의사결정 참여도, 의료비 지불 어려움 정도, 매년 같은 의료진에게 검진을 받는지 여부는 가족 알림 의도에 따른 차이가 없었다.

태도와 신념 관련된 변수로 암 관련 지식과 유전성 암 관련 지식, 암 유전 관련 이해도, 암 재발 가능성에 대한 믿음에 대한 두 군의 차이는 없었다. 또 스트레스 상황에 대처하는 방식 중 문제 해결에 치중하여 대처하는 정도를 보는 문제 중심적 대처 점수와 스트레스 상황에 감정에 치중하여 대처하는 정도를 보는 감정 중심적 대처 점수, 스트레스 상황을 회피하고자 하는 정도를 보는 회피적 대처 점수도 유의한 차이가 없었다.

### 4.3. 예측모델 구축과 성능

전통적인 변수 선택법을 통한 로지스틱 회귀분석 모델의 민감도는 0.7181, 특이도는 0.2286, 정확도는 0.5616 (95% CI 0.4932, 0.6284),

F1 값은 0.6903 이며, AUC 는 0.473 으로 전반적인 예측모델로서의 성능이 좋지 않은 편이다(Figure 2, Table 3).

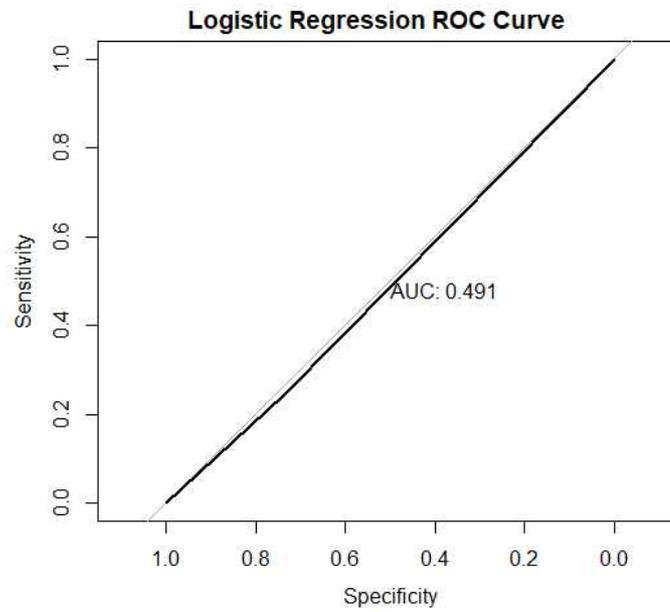
LASSO/Ridge 방법을 통한 로지스틱 회귀분석의 민감도는 1.0000, 특이도는 0.0000, 정확도는 0.6804 (95% CI 0.6142, 0.7416), F1 값은 0.8098 이며, AUC 는 0.500 이다(Figure 3, Table 3). 민감도와 특이도를 미루어 보아 해당 모델은 모두 가족 알림 의도가 있다고 예측하는 모델로, 정확도가 높게 측정되기는 하였으나 추후 대상자를 더 수집하여 예측모델의 성능을 다시 검토할 필요성이 있다.

랜덤 포레스트의 민감도는 0.9463, 특이도는 0.0857, 정확도는 0.6712 (95% CI 0.6047, 0.7330), F1 값은 0.7966 이며, AUC 는 0.516 로 LASSO/Ridge 방법을 통한 로지스틱 회귀분석과 F1 값은 비슷하고, AUC 값은 더 높았다(Figure 4, Table 3).

양상불 학습기법의 민감도는 0.9866, 특이도는 0.0571, 정확도는 0.6895 (95% CI 0.6237, 0.7501), F1 값은 0.8122 이며, AUC 는 0.522 으로, 전반적인 모델 성능이 앞선 모델들에 비해 우수한 편이다(Figure 5, Table 3).

*Table 3 Performance of the Prediction Models*

	Sensitivity	Specificity	Accuracy (95% CI)	F1-score
Logistic Regression	0.7383	0.2429	0.5799 (0.5155, 0.6461)	0.7051
Logistic Regression with Lasso/Ridge	1.0000	0.0000	0.8404 (0.6142, 0.7416)	0.8098
Random Forest	0.9463	0.0857	0.6712 (0.6047, 0.7330)	0.7966
Ensemble Learning	0.9866	0.0571	0.6895 (0.6237, 0.7501)	0.8122



*Figure 2. ROC curve of Logistic regression*

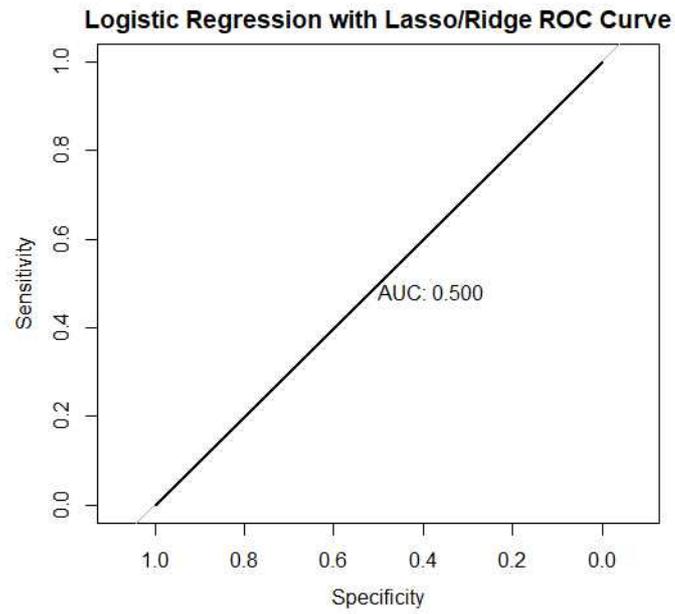


Figure 3. ROC curve of Logistic Regression with Lasso/Ridge

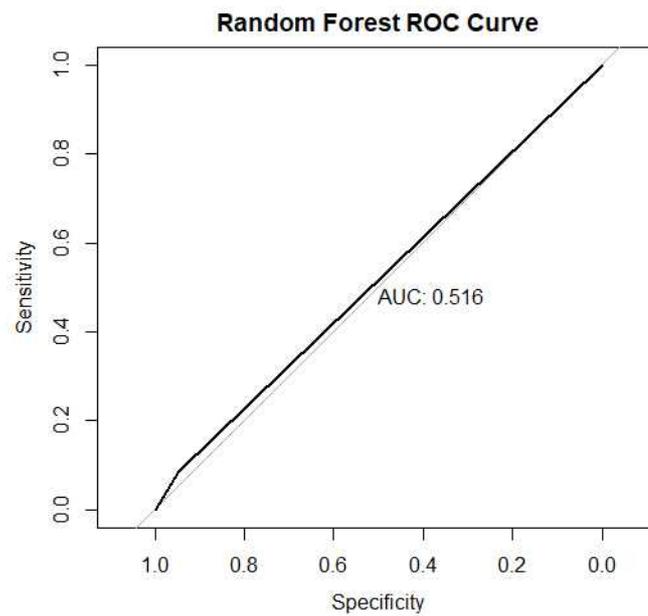
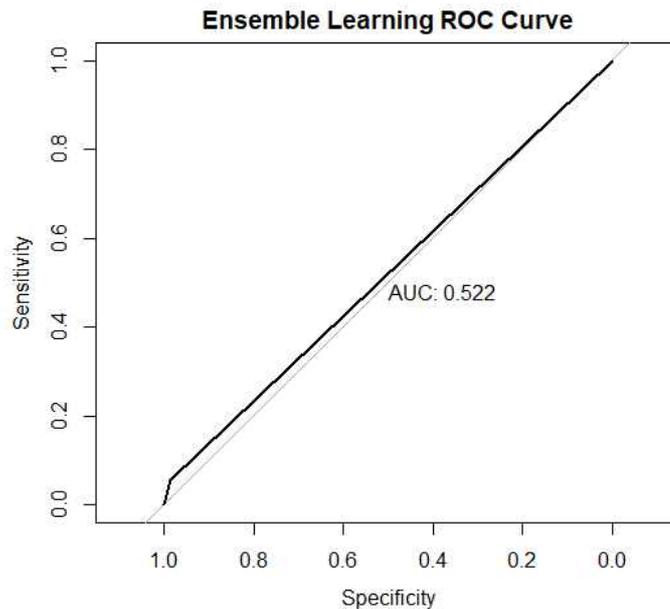


Figure 4. ROC curve of Random Forest



*Figure 5. ROC curve of Ensemble Learning*

예측모델 성능에 대한 평가를 종합하였을 때, 로지스틱 회귀분석 모델은 정확도와 F1 값, AUC 값이 다른 모델에 비해 현저히 낮아 가장 성능이 낮은 모델이다. 그 다음으로 LASSO/Ridge 방법을 통한 로지스틱 회귀분석 모델은 정확도와 F1 값이 랜덤 포레스트 모델에 비해 소폭 높기는 하나, 특이도가 0 이고, AUC 가 랜덤 포레스트 모델 값에 비해 낮은 것을 근거로 그 다음으로 성능이 낮은 모델로 평가하였다. 랜덤 포레스트 모델의 경우 단일 기법으로 만든 예측모델 중 가장 좋은 성능을 보이는 모델로 평가하였다.

양상블 학습기법으로 만든 모델의 경우, 정확도는 LASSO/Ridge 방법을 통한 로지스틱 회귀분석 모델에 비해 낮지만, 그 외 다른 모델 중에서는 정확도가 높으며, F1 값과 AUC 값이 제일 높았다. 이를 통해 단일 기법 예측모델에 비해 양상블 학습기법이 예측 성능은 더 우수함을 확인할 수 있었다.

양상블 학습기법에 대한 추가 분석으로 각각 모델과 양상블 모델의 Mean Square Error (MSE)를 바탕으로 로지스틱 회귀분석 모델의 MSE가 가장 커서 가장 성능이 좋지 않음을 확인할 수 있었고, 그 외 3가지 방법은 비슷한 수준의 MSE를 보임을 확인하였다(Table 4).

*Table 4. Meta Learner Summary*

<b>Algorithm</b>	<b>MSE</b>	<b>Standard Error</b>	<b>Min</b>	<b>Max</b>
Logistic Regression	0.2727	0.0206	0.1799	0.3436
Logistic Regression with Lasso/Ridge	0.2172	0.0114	0.2114	0.2257
Random Forest	0.2174	0.0121	0.1895	0.2375
Ensemble Learning	0.2173	0.0117	0.1998	0.2310

*\*MSE: Mean Square Error*

#### 4.4. 예측모델의 변수 중요도

예측모델의 변수 중요도는 단일 예측모델 중 가장 성능이 우수했던 랜덤 포레스트 모델을 기준으로 분석하였다(Table 5, Figure 6).

평균 지니 감소량(Mean Decrease Gini)이 높은 항목은 순서대로 암  
관련한 자기 효능감, 나이, 가족 결속력, 처음 암을 진단받은 나이,  
문제중심 대처, 질병에 대한 가족의 지지, 의학적 의사결정 참여도,  
감정중심 대처, 암 유전 관련 이해도, 유전자 검사에 대한 후회, 회피적  
대처였다. 이 항목들 중 나이와 유전자 검사에 대한 후회를 제외하고는  
가족 알림 의도가 있는 군과 없는 군을 비교할 때는 유의한 차이를 보이지  
않았던 변수였다.

*Table 5 Variable Importance in Random Forest*

<b>Category</b>	<b>Mean Decrease Gini</b>
Self-efficacy dealing with cancer	5.7385
Age	4.9610
Family cohesion	4.3526
Age at first cancer diagnosis	4.2446
Coping: problem-focused	4.2186
Family support in illness	4.2019
Medical decision participation	4.1677
Coping: emotion-focused	3.9734
Genetic affinity	3.8612
Decisional Regret	3.3777
Recurrence concerns	3.3251
Coping: avoidant	3.2634
Belief of cancer recurrence	2.8585
Family planning	2.4919
Family's wish	2.4330
Perceived GT usefulness	2.3364
Insurance covered	2.3202
Close distance	2.0151
Cancer treatment	2.0106
Physician advice	1.8659
Knowledge of cancer	1.8560
Risk of children	1.7222
Carrier family	1.6700
Risk of cancer	1.5184
Family benefit	1.4743
Check gene mutation	1.3313
BMI	1.3224
Smoking	1.1069
Live with child over 19-year-old	1.0915
Ovarian cancer	1.0896

<b>Category</b>	<b>Mean Decrease Gini</b>
Difficulty affording medical costs	0.9433
Oral contraceptives	0.9184
Drinking	0.8594
Pregnancy experience	0.8191
Education	0.8094
Breast biopsy	0.8066
Annual check-up with same staff	0.7879
Knowledge of genetic cancer	0.7150
Prophylactic mastectomy	0.6646
Prophylactic salpingo-oophorectomy	0.6604
Work	0.5890
Breast cancer	0.5567
Live with child under 19-year-old	0.5466
Live with parents	0.5181
Live with partner	0.4798
Married	0.3876
Physical activity	0.3578
Menopausal hormone therapy	0.3003
Live alone	0.2427

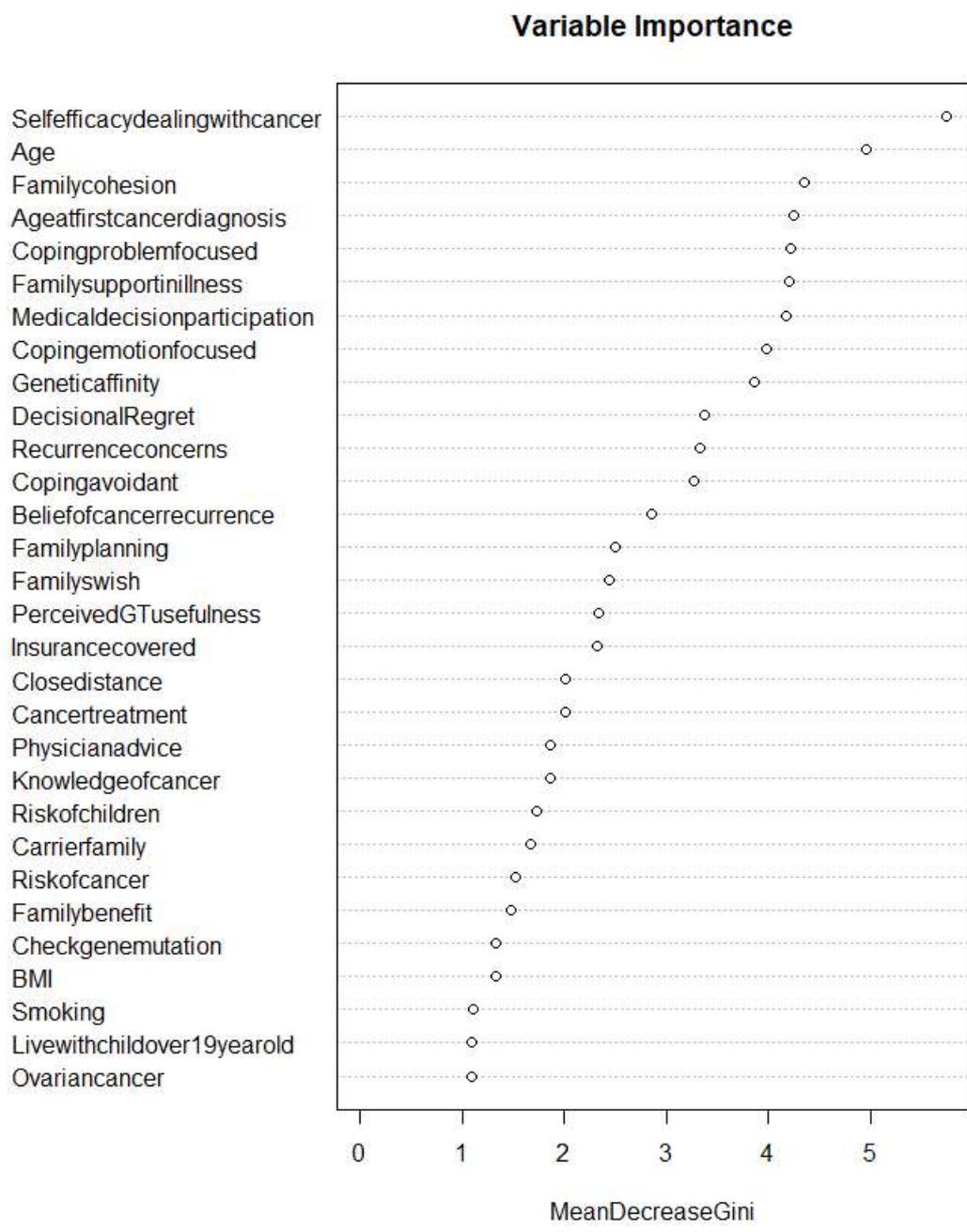


Figure 6. Variable Importance Plot

## V. 논의

가족 알림 의도는 K-CASCADE 코호트에 속한 HBOC 에서 31.96%로 확인되어 연구에 가족 초청 의도로 동일하게 조작화한 스위스 CASCADE 코호트에서 보고된 가족 알림 의도 65% (Sarki et al., 2022) 보다 현저히 낮았다. 또한 혈연관계 가족에게 유전자 검사를 알리는 비율을 보고한 터키에서 시행한 연구에서의 94.5% (Seven et al., 2022), 미국에서 부모가 자녀에게 알린 것을 조사한 55%(Bradbury et al., 2007)를 감안하면 직접적 비교가 어려울지라도 한국 코호트 참여자의 가족 알림 의도 비율은 낮은 것으로 사료된다. 이는 아시아인이 다른 인종에 비해 유전자 검사 결과를 알리는 비율이 낮았다는 선행 문헌과 일치한다(Cheung et al., 2010). 동양과 서양인들의 사회문화적 맥락의 차이가 가족 알림에 반영되었을 수 있으며(Barlow-Stewart et al., 2006; Suh, 2008), 추후 한국 코호트와 다른 코호트를 비교해 볼 필요가 있다.

예측모델에서는 전통적인 변수 선택법을 통한 로지스틱 회귀분석, Lasso/Ridge 방법을 통한 로지스틱 회귀분석, 랜덤 포레스트, 3 가지 방법을 혼합한 모델 순으로 성능이 좋았다. 모든 예측모델이 최신의 통계기법을 활용했을 때 성능이 좋을 것이라고 말할 수는 없지만,

예측모델을 만들 때 전통 방법 이외에도 혼합 모델을 포함한 여러 방법을 사용해 보는 것이 더 높은 성능을 가진 예측모델을 만들 때 필요하다.

예측모델을 만들어 변수 중요도를 확인함으로써 암 관련한 자기 효능감, 가족 결속력, 암 첫 진단 나이와 같이, 가족 알림 의도 유무에 따른 비교에서는 유의하다고 밝혀지지는 않았으나 예측모델에서는 중요한 변수로 영향을 미침을 확인할 수 있었다.

예측모델은 모두 0.6 미만의 AUC 를 보여 다소 미흡한 정확도를 보인다고 볼 수 있으나(Muller et al., 2005; Šimundić, 2009), 이 모델은 심리사회적 요인으로 구성된 모델이며 표본 수가 충분히 크지 않아 다소 낮게 측정되었을 수 있다. 이에 대상자를 더 모집하여 다시 분석해 볼 필요가 있다. 또한 심리사회적 요인뿐만 아니라 객관적인 전자의무기록(Electronic Medical Record)을 포함한다면 모델의 성능이 더 높아질 수 있을 것이다.

### 5.1. HBOC 의 가족 알림 의도 예측 인자

예측모델 적용 결과 가족 알림 의도에 영향을 미치는 요인으로 *BRCA1* 또는 *BRCA2* 보인자의 나이가 많을수록 높은 알림 의도를 유의하게 보였으며, 이는 스위스에서의 선행 연구 결과와 일치하였다(M. C. Katapodi et al., 2017; Sarki et al., 2022).

또한 본 연구에서 암 첫 진단 나이가 가족 알림 의도의 높은 예측력 변수로 확인되었으나, 선행 연구에서 이에 대해 밝혀진 바는 적다. 알림 의도가 있는 군의 평균 나이가 49.16 세이고, 암 첫 진단 나이가 43.14 세, 알림 의도가 없는 군의 평균 나이가 46.31 세, 암 첫 진단 나이가 40.95 세로, 평균 연령으로 보았을 때 두 군 모두 암 진단 후 약 3년 된 상태로 암에 대한 노출 기간의 차이가 아닌 나이와 유사한 성격을 가진 변수로 암 첫 진단 나이가 가족 알림 의도에 영향을 미친 것으로 추정할 수 있다. 향후 연구에서는 암 첫 진단 나이보다는 암 진단 후 몇 년이 지났는지, 유전자 검사를 시행하고 몇 년이 지났는지, 암 진단 시기와 유전자 검사를 시행한 시기는 어떤 차이가 있는지와 같이 일반적 나이 변수와는 구분되는 다른 변수들을 탐색해 볼 필요가 있다.

난소암이 있으면 가족 알림 의도가 더 높았으나, 유방암 진단 여부는 가족 알림 의도에 영향을 미치지 않았다. 선행연구(M. C. Katapodi et al., 2017; Sarki et al., 2022)에서는 난소암 대상자가 거의 없었기에 직접적 비교가 불가하나 난소암의 경우 Olaparib 등 PARP 저해제(Poly ADP-ribose polymerase inhibitor)를 활용한 치료계획을 위해 유전자 검사의 실시와 가족 알림 양상이 유방암과 다른 맥락임을 고려할 때 난소암 HBOC 에 대한 지속적 연구가 요구된다.

또한 본 연구결과에서 유전자 검사에 대한 후회 정도가 적고, 가족의 유전자 검사 유용성을 높게 평가할수록, 가족 구성원이 원해서 유전자 검사를 시행한 경우 가족 알림 의도가 더 높았다. 이는 선행 연구에서 보고된 HBOC 병인성 변이 결과에 대한 개인의 반응과 그 가족에게 미치는 의미가 중요한 영향을 미친다는 결과와 부합한다(Seven et al., 2021). 유전 상담 및 관련 진료 등의 유전자 서비스를 받는 것은 HBOC의 가족 알림에 첫 단추가 되는 단계로(Pedrazzani et al., 2022), 초기에 유전자 검사 시행 이유를 파악하고 검사의 유용성에 대한 교육을 제공하는 것이 가족 알림 의도 향상에 도움이 될 수 있을 것이다.

암에 관련한 자기 효능감과 암 재발에 대한 걱정은 가족 알림 의도 여부에 따라 통계적으로 유의한 변수는 아니었지만 가족 알림 의도를 예측하는데 좋은 변수였다. 즉, 암에 관련한 자기 효능감이 높을수록, 그리고 암 재발에 대한 걱정이 많을수록 가족 알림 의도가 있었다. 이는 자기 효능감이 높을수록 유전자 서비스 제공자에 대한 의존도가 낮고 유전자 검사와 예방활동에 대한 의도가 높았기에 가족 알림 의도에도 영향을 준 것으로 보인다(Katapodi et al., 2020). 암에 대한 재발에 대한 걱정은 주로 과거 가족 내에서 나타나며, 특히 가족의 암 경험이나 사망 여부가 영향을 미친다(Silva et al., 2022). HBOC 뿐만 아니라 가족 전체가 이러한 가족의 암 경험과 손실을 함께 겪었을 가능성이 높기 때문에,

HBOC가 본인의 암 재발에 대한 걱정을 갖고 있다면, 앞으로 자신이 암에 걸릴 경우에 가족들이 다시 한 번 비슷한 어려움을 경험하게 될 것에 대한 우려가 가족 알림 의도에 영향을 미치는 것으로 생각된다.

19세 이상의 자녀와 동거하는 사람이 가족 알림 의도가 유의하게 있었는데, 이는 선행 연구와 결과가 일치한다(Bradbury et al., 2007). NCCN 임상지침(NCCN, 2023)에서 성인 자녀의 유전자 검사를 권고하고 있으며 한국유방암학회의 임상지침(한국유방암학회 & 한국인유전성유방암연구회, 2021)에서도 19세 이상의 자녀에게 유전자 검사를 권장하며 검사 결과에 따라 예방적 중재를 시작할 것을 권고하는 방향과 일치하는 결과이다. 한국 코호트 예측모델 상에서도 가족 결속력과 질병 상황 속 가족의 지지는 가족 알림 의도를 예측하는 효과가 있는 변수로 확인되었으며, 이는 선행문헌에서 가족 환경(Lafrenière et al., 2013)과 질병 상황 속 가족의 지지(Sarki et al., 2022)가 의도에 영향을 미친다고 알려진 것과 일치했다.

의학적 의사결정 참여도가 가족 알림 의도를 예측하는데 좋은 변수로 확인된 본 연구의 결과는, 앞서 자기 관리의 암 관련 자기 효능감이 가족 알림 의도를 예측하는데 좋은 변수였고 HBOC의 암 관련 자기 효능감의 일종으로 의학적 의사결정 참여도가 가족 알림 의도의 예측에 영향을 미치는 것으로 보인다.

암 유전학에 대해 이해하고 있는 정도를 의미하는 암 유전 관련 이해도와 암 재발 가능성에 대한 믿음은 가족 알림 의도를 예측하는데 좋은 변수로 확인되었다. 이는 암 진단을 이미 받았고 또 유전자 서비스를 받아 암 유전 관련 이해도가 높을 수 있는(Pedrazzani et al., 2021) HBOC 단일 집단을 대상으로 했기 때문인 것으로 보인다. 본 연구에서는 스트레스 상황에 문제해결 중심적, 감정 중심으로 대처하는 것이 알림 의도에 영향을 미쳤는데, 이는 HBOC 결과를 듣고 가족에게 알려야 하는 스트레스 상황에서 해결하기 위한 대처 방법으로 문제해결 혹은 감정을 공유하는 대처 방법을 선택하는 사람들일수록 알림 의도가 높다고 생각할 수 있다. 이 결과는 선행연구에서 개인의 성격이 가족 알림에 영향이 있었다고 밝힌 것(Dick et al., 2022)보다 구체적으로 성격 중 문제 상황에 대처하는 양상에 대해 확인할 수 있었다.

한편 교육 수준과 결혼 유무에 대해서는 일관된 선행 연구결과가 없었으며, 본 연구에서는 교육 수준과 결혼 여부가 가족 알림 의도에 영향을 미치지 않음을 확인하였다. 한국 코호트 참여자들의 경우 결혼 유무보다는 임신 경험 유무가 가족 알림 의도에 영향을 미쳤으나 통계적으로 유의하지는 않았다. 그러나 가족 알림 의도가 있는 사람들이 유전자 검사를 시행한 이유로 자녀의 유전자 변이 위험성을 알고 싶다는 것과 가족 계획을 위해서, 가족에게 유익하기 때문에 시행했다고

응답했으며, 동거 가족 중 19 세 이상의 자녀 유무에 따라 가족 알림 의도에 차이가 있는 것으로 미루어 보아 자녀가 있고 없음이 영향을 미칠 수 있다. 유방 조직검사 경험이 가족 알림 의도가 있는 군이 70.00%, 가족 알림 의도가 없는 군이 81.88%로 가족 알림 의도가 있는 군이 11% 정도 적었으나 이는 유방암 진단받은 자의 비율이 가족 알림 의도가 있는 군이 65.71%, 가족 알림 의도가 없는 군에서는 75.84%로 유방암 진단 비율이 10% 정도 더 적은 것과 관련이 있는 것으로 보인다. 또한 예측모델 내에서 유방 조직검사 경험은 낮은 변수 중요도를 보였기 때문에 가족 알림 의도를 예측하는 유의한 변수로 보기에 다소 어려움이 있다.

HBOC 의 예방적 유방절제술이나 난소난관절제술은 가족 알림 의도에 영향을 미치지 않았으나 선행연구에서 터키(Seven et al., 2022)와 미국(Bradbury et al., 2007)의 경우 예방적 수술을 시행한 자들이 더 많은 가족 알림을 했다는 것과 다르게 확인되었다. 본 연구에서의 유방절제술은 약 20%, 예방적 난소난관절제술은 약 65% 로, 터키에서 각각 25%, 23% (Seven et al., 2022), 미국에서 각각 50%, 85% (Bradbury et al., 2007) 비율을 보인 것과 대조된다. 이러한 차이는 국가별로 사회문화적 요인과 국가 건강보험 체계가 영향을 미치는 것으로 볼 수 있다(Park et al., 2020).

일반적 의료비 지불 관련은 한국 코호트 참여자의 경우 의료비 지불  
관련한 어려움을 거의 호소하지 않았고, 이는 한국의 건강보험체계가  
영향을 미쳤을 것으로 보인다. *BRCA1* 또는 *BRCA2* 검사 비용이 한국의 경우  
170-190 만원 정도이며, 건강보험 기준에 해당할 시 8-9 만원의  
자기부담금이 발생하는 것에 반해(한국유방암학회 &  
한국인유전성유방암연구회, 2021), 미국의 경우 \$2000-4000 로 보험 혜택을  
누릴 수 없는 사람은 경제적으로 부담을 느낄 수 있다(Guo et al., 2020).

개인의 암과 관련된 지식 수준은 선행연구(Kang et al., 2011; M.  
C. Katapodi et al., 2017)와 동일하게 가족 알림 의도에 영향을 미치지  
않았다. 이는 HBOC 가 전반적으로 높은 수준의 HBOC 에 대한 지식을  
보유하고 있는 경향이 있어서(Claes et al., 2003), HBOC 안에서 그룹을  
나누어 비교했을 때 유의한 차이가 나타나지 않았을 수 있다.

## 5.2. 연구의 제한점

통상적으로 가족 알림 의도는 HBOC 가 본인의 유전자 검사 결과를  
최초로 알리고자 하는 의도를 의미하나 본 연구에서는 HBOC 가 본인의  
유전자 검사를 알리려는 여부가 최초인지는 확인할 수가 없었다. 그러나  
연구에의 초청 의도로 조작화한 선행연구(Sarki et al., 2022)와 맥을 같이  
하여 분석함으로써 HBOC 관련한 사회문화적 맥락에 대한 이해에 기여하고자

하였다. 또 K-CASCADE 는 현재 계속 대상자를 모집하고 있는 상태(ongoing enrollment)로 앞으로 추가될 대상자들의 특성에 따라 현재 밝힌 내용과 다른 요인들이 가족 알림 의도에 주요한 요인으로 밝혀질 수 있음을 고려해야 한다. 추후에는 유전자 변이에 대해 처음으로 가족에게 알리는 것인지 묻는 설문 문항의 추가 혹은 전자의무기록과 대응시켜 검사 결과 보고일을 기준으로 대상자를 제한해 분석하는 방법으로 보완할 수 있을 것이다.

### 5.3. 연구의 의의

#### 1) 간호 이론 측면

본 연구는 HBOC 의 가족 알림 의도에 미치는 영향요인을 탐색하는 연구로, 일반적 특성과 암 병력, 유전자 서비스, 유전자 검사를 시행한 이유, 자기 관리, 가족 환경, 의학적 환경, 신념과 태도의 다양한 범주의 변수들이 가족 알림 의도에 미치는 영향을 탐색하였다. 가족 알림 의도에 영향을 미치는 요인으로 나이, 가족 결속력, 질병 상황 속 가족의 지지, 유전자 검사 후회, 가족에게 유전자 검사가 유용하다는 인식, 19 세 이상의 자녀와 동거 여부, 스트레스 상황에 문제해결 중심적, 감정 중심적 대처하는 것은 선행연구와 동일함을 확인하였고 암에 관련한 자기 효능감, 암 첫 진단 나이, 의학적 의사결정 참여도, 암 유전 관련 이해도, 암

재발에 대한 걱정, 암 재발 가능성에 대한 믿음, 난소암 진단 여부, 유전자 검사 시행 이유는 본 연구에서의 독특한 요인임을 밝힘으로써 HBOC의 가족 알림 의도의 영향요인을 규명한 간호 지식체를 제공할 수 있다.

## 2) 간호 연구 측면

HBOC의 가족 알림 의도에 미치는 영향요인을 탐색하기 위해 K-CASCADE 코호트 자료를 바탕으로 KDD 방식을 통해 자료를 처리하고 HBOC의 가족 알림 의도 유무에 따라 변수별로 비교하고, 나아가 예측모델을 구축하고 평가하여 최적의 예측모델을 선정하였고, 선정한 예측모델의 변수 중요도를 탐색해보았다. 예측모델을 구축할 때 LASSO 변수 선택법을 활용한 로지스틱 회귀분석만을 활용하였던 Sarki et al. (2022)의 연구에서 나아가 전통적인 변수 선택법을 활용한 로지스틱 회귀분석과 LASSO/Ridge 변수 선택법을 활용한 회귀분석과 랜덤 포레스트 방법을 활용하여 다양한 통계분석 방법을 제시하였으므로 간호 연구 방법론 발전에 기여할 수 있다.

## 3) 간호 실무 측면

본 연구는 간호 실무 측면에서 HBOC의 가족 알림 의도 향상을 위한 중재 개발과 중재 제공에 근거자료가 될 수 있다. 한국 HBOC 코호트를 대상으로 가족 알림 의도에 대해 다양한 분류의 변수를 탐색해봄으로써,

한국 코호트의 가족 알림 의도에 따른 특징을 확인해 볼 수 있었다. 가족 알림 의도에 영향을 미치는 요인으로 확인된 암에 관련한 자기 효능감, 스트레스 상황에 대처 방식, 의학적 의사결정 참여도, 암 유전 관련 이해도는 HBOC 보인자 및 가족을 상담, 지지하는 간호사 및 유전상담사들이 활용할 수 있다. 또한 선행연구에서 지적했듯이(Baroutsou et al., 2021; Katapodi et al., 2018) 가족 알림 의도 향상을 위한 단계별 중재를 개발하고 적용하는 것이 도움이 될 것이다.

## VI. 결론

본 연구는 한국인 *BRCA1* 또는 *BRCA2* 병인성 변이를 가진 자와 혈연관계 가족을 대상으로 한 K-CASCADE 코호트를 통해 한국인에서의 가족 알림 의도 예측모델을 탐색한 첫 연구로서 의의가 있다. K-CASCADE 코호트에서 HBOC의 병인성 변이를 가족에게 알리려는 의도는 31.96%로 다소 낮았다.

유전자 검사 결과를 가족에게 알리는 의도를 예측하는데 다음 10 가지 요인이 가장 영향을 미침을 확인하였고 모델의 정확도는 67.12%였다: HBOC의 유전자 서비스와 관련해 유전자 검사에 대한 후회 정도; 자기 관리에서 암관련 자기 효능감; 가족 환경에서 가족 결속력, 질병에 대한 가족의 지지; 의학적 환경에서 본인이 의학적 의사결정에 참여하는 정도; 태도와 신념 관련하여 암 유전 관련 이해도와 스트레스 상황에 대처하는 문제 해결 중심적 방식과 감정 중심적 방식, 일반적 특성에서 나이, 그리고 암 진단 나이가 유의하였다.

HBOC라는 유전자 검사 결과를 듣게 되는 순간부터 적절한 유전자 서비스 제공이 필요하며 초기에 대상자에 대한 다양한 요인을 파악하여 앞으로 HBOC의 가족 알림 향상을 위해 필요한 것을 적절히 제공할 필요가 있다. 초기에는 대상자가 스트레스 상황에 대처하는 방식이나 암 유전 관련

이해도 수준을 확인하고, 대상자가 스스로 본인의 의학적 상황에 대해 이해하고 의사결정을 하며 자기 효능감이 향상될 수 있는 환경 조성  
과 유전자 가족의 유전자 검사에 대한 유용성에 대한 교육, 그리고 점차 가족을 참여시킬 수 있도록 적절한 유전자 서비스 개발 및 적용이 필요하다. 또 코호트를 계속 모집하여 현재 예측모델의 성능의 향상 여부와 알림 의도 예측에 도움이 된다고 밝혀진 변수의 설명력이 계속 일치하는지 확인해 볼 필요가 있다.

본 연구 결과를 바탕으로 다음을 제안한다.

첫째, HBOC의 가족 알림 향상을 위하여 초기 HBOC의 암 병력과 가족 환경을 포함한 다양한 사회 심리적 요인을 파악하고, 이에 맞춘 적절한 단계별 유전자 서비스를 개발하는 것이 필요하다. 가족들 간의 소통 및 지원체계를 강화하고, 유전자 서비스를 개선하여 가족 구성원들이 유전자 정보를 더 쉽게 이해하고 받아들일 수 있도록 지원하는 방안을 고려해야 한다.

둘째, HBOC의 가족 알림 향상을 위하여 HBOC에게 CGT 중요성을 인식시키는 노력이 필요하다. HBOC와 관련된 유전자 서비스에 대한 인식이 높아져야 하며, 이를 위해 국가 차원에서 정책을 마련하여 유전자 서비스 관리를 강화하는 것이 필요하다. 일관되고 효율적인 유전자 서비스를 제공하기 위해 국가적 차원에서 유전자 서비스 관리를 위한 표준화된

정책을 시행할 필요가 있다. 또한, 전문가들 간의 협력을 강화하여 고도의 유전자 서비스 품질을 제공하고, 교육 및 홍보를 통해 대중에게 유전자 서비스의 중요성을 알리는 것이 중요하다.

## 참고문헌

- 정고운. (2022). *BRCA1/2 변이를 가진 난소암 환자들의 가족 의사소통 과정과 영향요인에 대한 설문연구* [연세대학교 간호대학원]. 서울.  
<https://www.riss.kr/link?id=T16069773>
- 한국유방암학회, & 한국인유전성유방암연구회. (2021). *유전성 유방암 유전상담 매뉴얼*. 군자출판사.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.  
[https://doi.org/https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/https://doi.org/10.1016/0749-5978(91)90020-T)
- Barlow-Stewart, K., Yeo, S. S., Meiser, B., Goldstein, D., Tucker, K., & Eisenbruch, M. (2006). Toward cultural competence in cancer genetic counseling and genetics education: Lessons learned from Chinese-Australians. *Genetics in Medicine*, 8(1), 24-32.  
<https://doi.org/https://doi.org/10.1097/01.gim.0000195884.86201.a0>
- Baroutsou, V., Underhill-Blazey, M. L., Appenzeller-Herzog, C., & Katapodi, M. C. (2021). Interventions Facilitating Family Communication of Genetic Testing Results and Cascade Screening in

Hereditary Breast/Ovarian Cancer or Lynch Syndrome: A Systematic Review and Meta-Analysis. *Cancers*, 13(4), 925.

<https://www.mdpi.com/2072-6694/13/4/925>

Bradbury, A. R., Dignam, J. J., Ibe, C. N., Auh, S. L., Hlubocky, F. J., Cummings, S. A., White, M., Olopade, O. I., & Daugherty, C. K. (2007). How often do BRCA mutation carriers tell their young children of the family's risk for cancer? A study of parental disclosure of BRCA mutations to minors and young adults. *Journal of Clinical Oncology*, 25(24), 3705-3711.

<https://doi.org/10.1200/jco.2006.09.1900>

Bruce, P., Bruce, A., & Gedeck, P. (2021). *Practical Statistics for Data Scientists* [데이터 과학을 위한 통계 ] (이준용, Trans.; 2nd ed.). Hanbit Media. (O'Reilly Media)

Chatzisarantis, N. L. D., Yli-Piipari, S., Schriefer, L. S., Wang, D., Barkoukis, V., & Hagger, M. S. (2019). Is the relationship between physical activity intentions and behaviour convex? A test across 13 studies. *Psychology of Sport and Exercise*, 43, 114-122.

<https://doi.org/https://doi.org/10.1016/j.psychsport.2019.01.013>

- Cheung, E. L., Olson, A. D., Yu, T. M., Han, P. Z., & Beattie, M. S. (2010). Communication of BRCA results and family testing in 1,103 high-risk women. *Cancer Epidemiol Biomarkers Prev*, *19*(9), 2211-2219. <https://doi.org/10.1158/1055-9965.Epi-10-0325>
- Chuenphitthayavut, K., Zihuang, T., & Zhu, T. (2020). The prediction of behavioral intention to use online mental health interventions. *PsyCh Journal*, *9*(3), 370-382. <https://doi.org/https://doi.org/10.1002/pchj.333>
- Claes, E., Evers-Kiebooms, G., Boogaerts, A., Decruyenaere, M., Denayer, L., & Legius, E. (2003). Communication with close and distant relatives in the context of genetic testing for hereditary breast and ovarian cancer in cancer patients. *American Journal of Medical Genetics Part A*, *116A*(1), 11-19. <https://doi.org/https://doi.org/10.1002/ajmg.a.10868>
- Dar-Nimrod, I., MacNevin, G., Godwin, A., Lynch, K., Magory Cohen, T., Ganesan, A., & Morandini, J. (2019). Genetic Knowledge within a National Australian Sample: Comparisons with Other Diverse Populations. *Public Health Genomics*, *21*(3-4), 133-143. <https://doi.org/10.1159/000496381>

- Dean, M., & Rauscher, E. A. (2018). Men's and Women's Approaches to Disclosure About BRCA-Related Cancer Risks and Family Planning Decision-Making. *Qualitative Health Research, 28*(14), 2155–2168.  
<https://doi.org/10.1177/1049732318788377>
- Dean, M., Tezak, A. L., Johnson, S., Pierce, J. K., Weidner, A., Clouse, K., Pal, T., & Cragun, D. (2021). Sharing genetic test results with family members of BRCA, PALB2, CHEK2, and ATM carriers. *Patient Education and Counseling, 104*(4), 720–725.  
<https://doi.org/10.1016/j.pec.2020.12.019>
- Dick, J., Tüchler, A., Brédart, A., Vitinius, F., Wassermann, K., Rhiem, K., & Schmutzler, R. K. (2022). Psychological factors and the uptake of preventative measures in BRCA1/2 pathogenic variant carriers: results of a prospective cohort study. *Hereditary Cancer in Clinical Practice, 20*(1), 38.  
<https://doi.org/10.1186/s13053-022-00244-y>
- Ficarazzi, F., Vecchi, M., Ferrari, M., & Pierotti, M. A. (2021). Towards population-based genetic screenings for breast and ovarian cancer: A comprehensive review from economic evaluations

to patient perspectives. *Breast*, 58, 121-129.

<https://doi.org/10.1016/j.breast.2021.04.011>

Glymour, C., Madigan, D., Pregibon, D., & Smyth, P. (1996). Statistical inference and data mining. *Communications of the ACM*, 39(11), 35-41. <https://doi.org/10.1145/240455.240466>

Guo, F., Scholl, M., Fuchs, E. L., Berenson, A. B., & Kuo, Y.-F. (2020). BRCA testing in unaffected young women in the United States, 2006-2017. *Cancer*, 126(2), 337-343.

<https://doi.org/https://doi.org/10.1002/cncr.32536>

Hanley, G. E., McAlpine, J. N., Cheifetz, R., Schrader, K. A., McCullum, M., & Huntsman, D. (2019). Selected medical interventions in women with a deleterious BRCA mutation: a population-based study in British Columbia. *Current Oncology*, 26(1), e17-e23.

<https://doi.org/10.3747/co.26.4068>

Kang, E., Park, S. K., Kim, K. S., Choi, D. H., Nam, S.-J., Paik, N. S., Lee, J. W., Lee, M. H., & Kim, S.-W. (2011). Communication with Family Members about Positive BRCA1/2 Genetic Test Results in Korean Hereditary Breast Cancer Families. *Journal of genetic medicine*, 8(2), 105-112.

Katapodi, M. C., Duquette, D., Yang, J. J., Mendelsohn-Victor, K., Anderson, B., Nikolaidis, C., Mancewicz, E., Northouse, L. L., Duffy, S., Ronis, D., Milliron, K. J., Probst-Herbst, N., Merajver, S. D., Janz, N. K., Copeland, G., & Roberts, S. (2017). Recruiting families at risk for hereditary breast and ovarian cancer from a statewide cancer registry: a methodological study. *Cancer Causes Control, 28*(3), 191-201.  
<https://doi.org/10.1007/s10552-017-0858-2>

Katapodi, M. C., Jung, M., Schafenacker, A. M., Milliron, K. J., Mendelsohn-Victor, K. E., Merajver, S. D., & Northouse, L. L. (2018). Development of a Web-based Family Intervention for BRCA Carriers and Their Biological Relatives: Acceptability, Feasibility, and Usability Study. *JMIR Cancer, 4*(1), e7.  
<https://doi.org/10.2196/cancer.9210>

Katapodi, M. C., Ming, C., Northouse, L. L., Duffy, S. A., Duquette, D., Mendelsohn-Victor, K. E., Milliron, K. J., Merajver, S. D., Dinov, I. D., & Janz, N. K. (2020). Genetic Testing and Surveillance of Young Breast Cancer Survivors and Blood

Relatives: A Cluster Randomized Trial. *Cancers*, 12(9), 2526.

<https://www.mdpi.com/2072-6694/12/9/2526>

Katapodi, M. C., Viassolo, V., Caiata-Zufferey, M., Nikolaidis, C.,  
Bührer-Landolt, R., Buerki, N., Graffeo, R., Horváth, H. C.,  
Kurzeder, C., Rabaglio, M., Scharfe, M., Urech, C., Erlanger, T.  
E., Probst-Hensch, N., Heinimann, K., Heinzelmann-Schwarz, V.,  
Pagani, O., & Chappuis, P. O. (2017). Cancer Predisposition  
Cascade Screening for Hereditary Breast/Ovarian Cancer and Lynch  
Syndromes in Switzerland: Study Protocol [Proposal]. *JMIR  
Research Protocols*, 6(9), e184.

<https://doi.org/10.2196/resprot.8138>

Kim, J. W. (2016). Targeted therapy and immunotherapy in ovarian cancer.  
*Journal of the Korean Medical Association*, 59(3), 180-188.

<https://doi.org/10.5124/jkma.2016.59.3.180>

Kim, S., Aceti, M., Baroutsou, V., Bürki, N., Caiata-Zufferey, M.,  
Cattaneo, M., Chappuis, P. O., Ciorba, F. M., Graffeo-Galbiati,  
R., Heinzelmann-Schwarz, V., Jeong, J., Jung, M. M., Kim, S. W.,  
Kim, J., Lim, M. C., Ming, C., Monnerat, C., Park, H. S., Park,  
S. H., . . . Katapodi, M. C. (2021). Using a Tailored Digital

Health Intervention for Family Communication and Cascade Genetic Testing in Swiss and Korean Families With Hereditary Breast and Ovarian Cancer: Protocol for the DIALOGUE Study. *JMIR Research Protocols*, *10*(6), e26264. <https://doi.org/10.2196/26264>

Lafrenière, D., Bouchard, K., Godard, B., Simard, J., & Dorval, M. (2013). Family Communication Following BRCA1/2 Genetic Testing: A Close Look at the Process. *Journal of Genetic Counseling*, *22*(3), 323-335. <https://doi.org/https://doi.org/10.1007/s10897-012-9559-x>

Lieberman, S., Lahad, A., Tomer, A., Koka, S., BenUziyahu, M., Raz, A., & Levy-Lahad, E. (2018). Familial communication and cascade testing among relatives of BRCA population screening participants. *Genetics in Medicine*, *20*(11), 1446-1454. <https://doi.org/https://doi.org/10.1038/gim.2018.26>

Mansfield, C. A., Metcalfe, K. A., Snyder, C., Lindeman, G. J., Posner, J., Friedman, S., Lynch, H. T., Narod, S. A., Evans, D. G., & Liede, A. (2020). Preferences for breast cancer prevention among women with a BRCA1 or BRCA2 mutation. *Hereditary Cancer in*

*Clinical Practice*, 18, 20. <https://doi.org/10.1186/s13053-020-00152-z>

- Mavrogiorgou, A., Kiourtis, A., Manias, G., & Kyriazis, D. (2021, 24-26 May 2021). An Optimized KDD Process for Collecting and Processing Ingested and Streaming Healthcare Data. 2021 12th International Conference on Information and Communication Systems (ICICS),
- McClelland, M. K. (2010). Benefits of Knowledge Discovery Process for Biomedical Population Study. *Journal of Information Systems Applied Research*, 3(2), 1-21.
- Muller, M. P., Tomlinson, G., Marrie, T. J., Tang, P., McGeer, A., Low, D. E., Detsky, A. S., & Gold, W. L. (2005). Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia? *Clinical Infectious Diseases*, 40(8), 1079-1086. <https://doi.org/10.1086/428577>
- NCCN. (2023). *Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic*. Retrieved Sep 27 from NCCN.org
- Park, S., Kim, Y., & Kim, S. (2020). Factors associated with the decision to undergo risk-reducing salpingo-oophorectomy among

women at high risk for hereditary breast and ovarian cancer: a systematic review. *Korean Journal of Women Health Nursing* 26(4), 285-299. <https://doi.org/10.4069/kjwhn.2020.11.19>

Pedrazzani, C., Aceti, M., Schweighoffer, R., Kaiser-Grolimund, A., Bürki, N., Chappuis, P. O., Graffeo, R., Monnerat, C., Pagani, O., Rabaglio, M., Katapodi, M. C., & Caiata-Zufferey, M. (2022). The Communication Chain of Genetic Risk: Analyses of Narrative Data Exploring Proband-Provider and Proband-Family Communication in Hereditary Breast and Ovarian Cancer. *Journal of Personalized Medicine*, 12(8). <https://doi.org/10.3390/jpm12081249>

Pedrazzani, C., Ming, C., Bürki, N., Caiata-Zufferey, M., Chappuis, P. O., Duquette, D., Heinimann, K., Heinzelmann-Schwarz, V., Graffeo-Galbiati, R., Merajver, S. D., Milliron, K. J., Monnerat, C., Pagani, O., Rabaglio, M., & Katapodi, M. C. (2021). Genetic Literacy and Communication of Genetic Information in Families Concerned with Hereditary Breast and Ovarian Cancer: A Cross-Study Comparison in Two Countries and within a Timeframe of More Than 10 Years. *Cancers*, 13(24), 6254. <https://www.mdpi.com/2072-6694/13/24/6254>

Polley, E., LeDell, E., Kennedy, C., Lendle, S., & Laan, M. v. d.

(2023). *Package 'SuperLearner'*. In (Version 2.0-28.1)

<https://github.com/ecpolley/SuperLearner>

Sarki, M., Ming, C., Aissaoui, S., Bürki, N., Caiata-Zufferey, M.,

Erlanger, T. E., Graffeo-Galbiati, R., Heinimann, K.,

Heinzelmann-Schwarz, V., Monnerat, C., Probst-Hensch, N.,

Rabaglio, M., Zürrer-Härdi, U., Chappuis, P. O., Katapodi, M. C.,

& On Behalf Of The Cascade, C. (2022). Intention to Inform

Relatives, Rates of Cascade Testing, and Preference for Patient-

Mediated Communication in Families Concerned with Hereditary

Breast and Ovarian Cancer and Lynch Syndrome: The Swiss CASCADE

Cohort. *Cancers (Basel)*, *14*(7).

<https://doi.org/10.3390/cancers14071636>

Seven, M., Shah, L. L., Daack-Hirsch, S., & Yazici, H. (2021).

Experiences of BRCA1/2 Gene Mutation-Positive Women With Cancer

in Communicating Genetic Risk to Their Relatives. *Cancer Nursing*,

*44*(3), E142-E150. <https://doi.org/10.1097/ncc.0000000000000796>

Seven, M., Shah, L. L., Yazici, H., & Daack-Hirsch, S. (2022). From

Probands to Relatives: Communication of Genetic Risk for

Hereditary Breast-Ovarian Cancer and Its Influence on Subsequent Testing. *Cancer Nursing*, 45(1), E91-E98.

<https://doi.org/10.1097/ncc.0000000000000876>

Silva, E., Gomes, P., Matos, P. M., Silva, E. R., Silva, J., Brandão, C., Castro, F., Neves, M. C., & Sales, C. M. D. (2022). “I have always lived with the disease in the family” : family adaptation to hereditary cancer-risk. *BMC Primary Care*, 23(1), 93.

<https://doi.org/10.1186/s12875-022-01704-z>

Šimundić, A. M. (2009). Measures of Diagnostic Accuracy: Basic Definitions. *Ejifcc*, 19(4), 203-211.

Sohn, Y., & Lee, B. (2012). An Efficacy of Social Cognitive Behavior Model based on the Theory of Planned Behavior : A Meta-Analytic Review [An Efficacy of Social Cognitive Behavior Model based on the Theory of Planned Behavior : A Meta-Analytic Review]. *Korean Journal of Journalism & Communication Studies*, 56(6), 127-161.

<http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02060403>

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for

- Performance Evaluation. In A. Sattar & B.-h. Kang, *AI 2006: Advances in Artificial Intelligence* Berlin, Heidelberg.
- Suh, E. E. (2008). The sociocultural context of breast cancer screening among Korean immigrant women. *Cancer Nursing, 31*(4), E1-10.  
<https://doi.org/10.1097/01.NCC.0000305742.56829.fc>
- Surbone, A. (2011). Social and ethical implications of BRCA testing. *Annals of Oncology, 22*, i60-i66.  
<https://doi.org/https://doi.org/10.1093/annonc/mdq668>
- Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin, 132*(2), 249-268.  
<https://doi.org/https://doi.org/10.1037/0033-2909.132.2.249>
- Young, A. L., Butow, P. N., Tucker, K. M., Wakefield, C. E., Healey, E., & Williams, R. (2020). When to break the news and whose responsibility is it? A cross-sectional qualitative study of health professionals' views regarding disclosure of BRCA genetic cancer risk. *BMJ Open, 10*(2), e033127.  
<https://doi.org/10.1136/bmjopen-2019-033127>

## Appendix

*Appendix 1. Comparison of K-CASCADE whole sample and this study's sample*

	Whole sample (n=298)	Study sample (n=219)	Z or $\chi^2$	df	p
	M±SD or n(%)	M±SD or n(%)			
<b>Intention to Inform</b>	89(29.87)	70(31.96)	0.26	515	.610
<b>General Characteristics</b>					
Age	48.23±9.90	47.22±9.78	1.15	510	.252
Married	220(74.83)	164(74.89)	0.00	511	.988
BMI (ref=under 23)	194(65.76)	142(64.84)	0.05	512	.828
Education (ref=Bachelor's degree or higher)	193(67.48)	162(73.97)	2.50	503	.114
Work (ref=full-time)	65(22.73)	54(24.66)	0.26	503	.612
Smoking (ref=never smoke)	259(87.50)	191(87.21)	0.01	513	.923
Drinking (ref=nondrinker)	267(90.20)	199(90.87)	0.06	513	.799
Physical activity (ref=Yes)	276(92.62)	201(91.78)	0.12	515	.725
Pregnancy experience (ref=Yes)	240(80.81)	171(78.08)	0.58	514	.447
Breast biopsy (ref=Yes)	230(77.44)	171(78.08)	0.03	514	.863
Menopausal hormone therapy (ref=Yes)	21(7.12)	16(7.31)	0.01	512	.935
Oral contraceptives (ref=Yes)	59(19.87)	46(21.00)	0.10	514	.751
<b>Cancer History</b>					
Age at first cancer diagnosis	42.59±9.81	41.65±9.49	1.09	512	.278
Breast cancer (ref=Yes)	212(71.14)	159(72.60)	0.13	515	.715
Ovarian cancer (ref=Yes)	98(32.89)	63(28.77)	1.00	515	.318

	Whole sample (n=298)	Study sample (n=219)	Z or $\chi^2$	df	<i>p</i>
	M±SD or n(%)	M±SD or n(%)			
Prophylactic mastectomy (ref=Yes)	51(17.11)	41(18.72)	0.22	515	.637
Prophylactic salpingo-oophorectomy (ref=Yes)	99(33.22)	75(34.25)	0.06	515	.807
<b>Genetic Services</b>					
Decisional Regret	6.04±1.14	6.06±1.10	-0.24	505	.810
Perceived GT usefulness	5.99±1.17	5.96±1.22	0.28	501	.776
<b>Reason for Genetic test</b>					
Risk of cancer	6.03±1.31	6.03±1.30	-0.03	500	.975
Check gene mutation	6.00±1.39	5.99±1.39	0.07	501	.942
Family benefit	5.92±1.30	5.96±1.27	-0.32	499	.750
Physician advice	5.90±1.33	5.90±1.35	0.01	499	.992
Cancer treatment	5.86±1.48	5.88±1.51	-0.11	493	.915
Risk of children	5.60±1.74	5.51±1.80	0.58	499	.561
Family's wish	4.96±1.93	4.82±2.01	0.79	498	.430
Insurance covered	4.52±2.25	4.45±2.29	0.36	489	.717
Family planning	3.72±2.32	3.74±2.34	-0.09	492	.926
Carrier family	3.30±2.37	3.29±2.40	0.06	500	.953
Close distance	2.62±1.92	2.58±1.94	0.24	494	.814
<b>Self-management</b>					
Self-efficacy dealing with cancer	5.02±1.10	4.99±1.12	0.36	512	.716
Recurrence concerns	4.74±1.33	4.80±1.33	-0.54	513	.592
<b>Family environment</b>					
Family cohesion	5.56±1.02	5.52±1.04	0.38	505	.702
Family support in illness	5.85±1.01	5.83±1.03	0.25	505	.805
Live alone	22(7.46)	18(8.22)	0.10	512	.750

	Whole sample (n=298)	Study sample (n=219)	Z or $\chi^2$	df	<i>p</i>
	M±SD or n(%)	M±SD or n(%)			
Live with partner	214(72.54)	160(73.06)	0.02	512	.896
Live with child under 19-year-old	107(36.27)	83(37.90)	0.14	512	.705
Live with child over 19-year-old	100(33.90)	67(30.59)	0.63	512	.429
Live with parents	40(13.56)	35(15.98)	0.59	512	.442
<b>Medical environment</b>					
Medical decision participation	5.62±1.22	5.62±1.18	-0.01	507	.992
Difficulty affording medical costs	1.54±1.06	1.58±1.13	-0.38	505	.705
Annual check-up with same staff (ref=Yes)	207(73.40)	166(75.80)	0.37	499	.542
<b>Belief and Attitude</b>					
Knowledge of cancer	2.59±1.44	2.79±1.35	-1.58	515	.114
Knowledge of genetic cancer	0.65±0.48	0.68±0.47	-0.68	503	.497
Genetic affinity	3.45±1.27	3.51±1.28	-0.51	506	.611
Belief of cancer recurrence	6.19±2.27	6.33±2.24	-0.67	501	.506
Coping: problem- focused	5.03±0.95	5.07±0.92	-0.42	505	.677
Coping: emotion- focused	3.78±0.71	3.78±0.66	0.11	505	.914
Coping: avoidant	2.75±0.72	2.74±0.70	0.11	505	.916

## Abstract

### Intention to Inform Family in Hereditary Breast and Ovarian Cancer Syndrome

Kim, Yeeun

Dept. of Nursing

The Graduate School

Yonsei University

Cancer genetic testing (CGT) is recommended for family members of persons identified to have Hereditary Breast and Ovarian Cancer syndrome (HBOC). The first step for an index case with HBOC in family communication for CGT, is making a decision to inform family members of their genetic test result.

This study aimed to identify predictors of intention to inform family of HBOC in Korea. Specific aims were to compare the difference between groups with the intention to inform and non-intention to inform; to develop prediction models of the intention to inform; and finally, to

identify the predictor in the best prediction model of the intention to inform.

This study used data from the K-CASCADE cohort, as part of an international cohort of HBOC index cases. Data from 219 participants who were enrolled from March 2021 to September 2023, were analyzed. From the K-CASCADE baseline survey, items on general characteristics, cancer history, genetic services, reason for genetic test, self-management, family environment, medical environment, and belief and attitude were used for independent variables. The outcome variable, i.e., intention to inform, was operationalized as the intention to invite blood-related (at-risk) family members to the ongoing study. For analysis, T-test and chi-square test were used to compare the intention group and non-intention group. Prediction models were made by logistic regression using the traditional variable selection method, logistic regression with LASSO variable selection method, and random forest. Prediction models were constructed by the SuperLearner package in R and evaluated by performance indices (sensitivity, specificity, accuracy, F1 score, and ROC curve). The importance of variables in the random forest model was measured by the mean decrease of Gini.

Between the intention group and non-intention group, older age ( $Z=-2.03$ ,  $p=.044$ ), pregnancy experience ( $\chi^2=4.94$ ,  $p=.026$ ), ovarian cancer ( $\chi^2=6.34$ ,  $p=.012$ ), lower decisional regret ( $Z=-2.63$ ,  $p=.009$ ), higher perceived genetic test usefulness ( $Z=-3.26$ ,  $p=.001$ ), and living with child(ren) over 19 years-old ( $\chi^2=4.29$ ,  $p=.038$ ) were significantly related to higher intention to inform.

Among the prediction models, the random forest model was the best-performing model with a sensitivity of 0.9463, specificity of 0.0857, F1 score of 0.7966, and AUC of 0.516.

Self-efficacy dealing with cancer, age, family cohesion, age at first cancer diagnosis, problem-focused coping, family support in illness, medical decision participation, emotion-focused coping, genetic affinity, and decisional regret were the highest variables of importance, presented in order.

Intention to inform family in the K-CASCADE cohort was 32%, which was lower than the 65% reported in the Swiss cohort. Compared to simple regression findings on the intention to inform family, previously undisclosed variables such as self-efficacy dealing with cancer, family cohesion, and age at the first cancer diagnosis were identified as

predictive factors through predictive modelling. Considering the diverse factors influencing the intention to inform family in HBOC, early identification of these variables is crucial. Nurses and genetic counselors can use this study's findings when providing genetic services and considering necessary education according to client needs. Additionally, incorporating objective data such as Electronic Medical Records, could enhance the model's performance in future studies.

This study lays the foundation for understanding the intention to inform family in Korean HBOC cases. Developing family-based interventions that address the various socio-psychological factors identified across each step of the genetic service process, can significantly contribute to improving the intention to inform in HBOC.

---

Key words : HBOC, inform family, intention, K-CASCADE