

연구보고서 2023-14

# 보건복지분야 위기대응 예측모델링 연구

- 정신건강 사례를 중심으로

오미애

박성준·안수인·최호식·박민수·조현경

사람을  
생각하는  
사람들



KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



한국보건사회연구원  
KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



## ■ 연구진

연구책임자	오미애	한국보건사회연구원 연구위원
공동연구진	박성준	한국보건사회연구원 부연구위원
	안수인	한국보건사회연구원 연구원
	최호식	서울시립대학교 도시빅데이터융합학과 교수
	박민수	충남대학교 정보통계학과 교수
	조현경	한국보건사회연구원 연구원

연구보고서 2023-14

### 보건복지분야 위기대응 예측모델링 연구

- 정신건강 사례를 중심으로

발행일 2023년 12월  
발행인 이태수  
발행처 한국보건사회연구원  
주소 [30147]세종특별자치시 시청대로 370  
세종국책연구단지 사회정책동(1~5층)  
전화 대표전화: 044)287-8000  
홈페이지 <http://www.kihasa.re.kr>  
등록 1999년 4월 27일(제2015-000007호)  
인쇄처 고려씨엔피

© 한국보건사회연구원 2023  
ISBN 978-89-6827-969-0 [93510]  
<https://doi.or.kr/10.23060/kihasa.a.2023.14>

## 발|간|사

급격한 사회경제적 변화, 코로나19 상황으로 인해 정신건강 영역은 점점 중요해지고 있다. 정신건강은 단순히 정신병이나 심리적 장애가 없는 상태가 아니라, 개인의 삶에 대한 만족도와 적응력, 사회적 관계와 역할 수행능력 등을 포함하는 포괄적인 개념이다.

정신건강 문제는 성별에 따라, 연령대에 따라 주원인이 다르기 때문에 정책 대상에 초점을 맞추어 조기 발견 가능성 검토에 대한 심층연구가 필요하다. 특히 자살 고위험군의 특성의 경우, 성별, 연령대를 나누어 접근할 필요가 있다.

본 연구는 정책 대상별 특성을 파악하고 다양한 공공데이터를 활용하여 정책 대상별 예측 모델을 개발하는 것을 목적으로 한다. 공공데이터를 활용하여 성별을 고려한 정책 대상별(아동·청소년/청년/중장년/노인) 정신건강 예측 모델을 개발하고, 국정과제(예방적 건강관리 강화: 정신건강 문제 조기 발견 및 개입) 지원 기반을 마련하고자 한다.

본 보고서의 결과는 우리 연구원의 공식적인 견해가 아니라 연구진의 의견임을 밝혀 둔다.

2023년 12월

한국보건사회연구원 원장

이 태 수





Abstract .....	1
요 약 .....	3
<b>제1장 서론 .....</b>	<b>11</b>
제1절 연구의 배경 및 목적 .....	13
제2절 연구의 내용 및 방법 .....	20
<b>제2장 정신건강 및 자살 고위험군에 대한 이론적 검토 .....</b>	<b>23</b>
제1절 정신건강 고위험군의 특성 .....	25
제2절 자살 고위험군의 특성 .....	33
제3절 소결 .....	47
<b>제3장 예측 모델 적용을 위한 공공데이터 검토 .....</b>	<b>51</b>
제1절 검토 대상 공공데이터 개요 .....	53
제2절 자살 고위험군 관련 선행연구 실증 분석 변수 .....	84
제3절 소결 .....	113
<b>제4장 공공데이터를 활용한 정책 대상별 예측 모델 .....</b>	<b>115</b>
제1절 청소년건강행태조사를 활용한 예측 모델 .....	117
제2절 정신건강실태조사를 활용한 예측 모델 .....	139
제3절 소결 .....	173

<b>제5장 심층 분석 I : 자살 고위험군에 대한 위험 인자 탐색을 위한 패널데이터 생존 분석 모형 적용</b> .....	<b>179</b>
제1절 한국복지패널의 분석 개요 .....	182
제2절 통계량 산출 및 집단 간 비교 .....	193
제3절 생존 분석 결과 .....	205
제4절 소결 .....	255
<b>제6장 심층 분석 II : 텍스트 데이터 기반 자살 고위험군 탐지를 위한 딥러닝 적용</b> .....	<b>259</b>
제1절 텍스트마이닝 및 자연어처리 기반 분석 방법 .....	261
제2절 텍스트의 그룹화 분석 결과 .....	272
제3절 문장 임베딩을 활용한 분석 결과 .....	282
제4절 소결 .....	300
<b>제7장 결론 및 시사점</b> .....	<b>303</b>
제1절 결론 .....	305
제2절 시사점: 정책적 활용 측면 .....	308
제3절 시사점: 예측 모형 활용 측면 .....	314
<b>참고문헌</b> .....	<b>321</b>

# 표 목차

KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



〈표 2-1〉 자살의 위험요인과 보호요인 .....	39
〈표 2-2〉 아동·청소년 자살 행동의 위험요인과 보호요인 .....	44
〈표 3-1〉 국민건강영양조사 개요 .....	54
〈표 3-2〉 국민건강영양조사의 건강설문조사 항목 상세 .....	55
〈표 3-3〉 국민건강영양조사의 자살 관련 문항 연도별 현황 .....	57
〈표 3-4〉 국민건강영양조사의 자살 관련 문항 .....	58
〈표 3-5〉 국민건강영양조사의 자살 관련 문항 응답 빈도(표본 수) .....	59
〈표 3-6〉 정신건강실태조사(National Mental Health Survey) 개요 .....	60
〈표 3-7〉 정신건강실태조사의 조사 항목 상세 .....	62
〈표 3-8〉 정신건강실태조사의 자살 관련 문항 연도별 현황 .....	64
〈표 3-9〉 정신건강실태조사의 자살 관련 문항 .....	65
〈표 3-10〉 정신건강실태조사의 자살 관련 행동 비율(%) .....	67
〈표 3-11〉 청소년건강행태조사 개요 .....	68
〈표 3-12〉 청소년건강행태조사 조사 항목 상세(총 114개 문항) .....	69
〈표 3-13〉 청소년건강행태조사 자살 관련 문항 연도별 현황 .....	71
〈표 3-14〉 청소년건강행태조사의 자살 관련 문항 .....	72
〈표 3-15〉 청소년건강행태조사의 자살 관련 행동 비율(%) .....	73
〈표 3-16〉 한국복지패널조사 개요 .....	74
〈표 3-17〉 한국복지패널조사 조사 항목(조사 주제별) .....	76
〈표 3-18〉 한국복지패널조사의 자살 관련 문항 연도별 현황 .....	78
〈표 3-19〉 한국복지패널조사 자살 관련 문항 .....	79
〈표 3-20〉 한국복지패널조사의 자살 관련 행동 빈도 수 .....	80
〈표 3-21〉 한국의료패널조사 개요 .....	81
〈표 3-22〉 한국의료패널조사의 자살 관련 문항 연도별 현황 .....	83
〈표 3-23〉 한국의료패널조사의 자살 관련 문항 .....	83
〈표 3-24〉 한국의료패널조사의 자살 관련 행동(자살 사고) 빈도 수 .....	83
〈표 3-25〉 선행연구에서의 분석 변수 정리 요약 .....	85

〈표 3-26〉 선행연구에서의 분석 변수 정리 .....	87
〈표 3-27〉 선행연구 실증 분석 결과: 독립 변수의 통계적 유의성 .....	97
〈표 4-1〉 청소년건강행태조사 2021의 데이터 정책 대상별 기초통계 .....	118
〈표 4-2〉 청소년건강행태조사 2021의 예측 모형 적용 변수 .....	119
〈표 4-3〉 남자 중학생 집단 Logistic regression & Elastic net 결과 .....	122
〈표 4-4〉 남자 중학생 집단 Boosting Relative Importance Plot .....	124
〈표 4-5〉 여자 중학생 집단 Logistic regression & Elastic net 결과 .....	127
〈표 4-6〉 여자 중학생 집단 Boosting Relative Importance Plot .....	129
〈표 4-7〉 남자 고등학생 집단 Logistic regression & Elastic net 결과 .....	130
〈표 4-8〉 남자 고등학생 집단 Boosting Relative Importance Plot .....	133
〈표 4-9〉 여자 고등학생 집단 Logistic regression & Elastic net 결과 .....	135
〈표 4-10〉 여자 고등학생 집단 Boosting Relative Importance Plot .....	138
〈표 4-11〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 빈도표 .....	139
〈표 4-12〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 평생 자살 사고 기초 통계 .....	139
〈표 4-13〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 평생 자살 계획 기초 통계 .....	140
〈표 4-14〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 평생 자살 시도 기초 통계 .....	141
〈표 4-15〉 정신건강실태조사 2021의 데이터 자살 계획 또는 자살 시도 여부 기초통계 .....	141
〈표 4-16〉 정신건강실태조사 2021의 데이터 정책집단별 자살 계획 또는 자살 시도 여부 기초통계 .....	142
〈표 4-17〉 정신건강실태조사 2021의 예측 모형 적용 변수 .....	142
〈표 4-18〉 남자 18~39세 집단 Logistic regression & Elastic net 결과 .....	145
〈표 4-19〉 남자 18~39세 집단 Boosting Relative Importance Plot .....	148
〈표 4-20〉 여자 18~39세 집단 Logistic regression & Elastic net 결과 .....	150





〈표 4-21〉 여자 18~39세 집단 Boosting Relative Importance Plot .....	152
〈표 4-22〉 남자 40~64세 집단 Logistic regression & Elastic net 결과 .....	154
〈표 4-23〉 남자 40~64세 집단 Boosting Relative Importance Plot .....	157
〈표 4-24〉 여자 40~64세 집단 Logistic regression & Elastic net 결과 .....	159
〈표 4-25〉 여자 40~64세 집단 Boosting Relative Importance Plot .....	162
〈표 4-26〉 남자 65~79세 집단 Logistic regression & Elastic net 결과 .....	164
〈표 4-27〉 남자 65~79세 집단 Boosting Relative Importance Plot .....	167
〈표 4-28〉 여자 65~79세 집단 Logistic regression & Elastic net 결과 .....	169
〈표 4-29〉 여자 65~79세 집단 Boosting Relative Importance Plot .....	172
〈표 5-1〉 변수 설명 및 결측 비율 .....	183
〈표 5-2〉 자살 생각 결과변수의 빈도 및 비율 .....	194
〈표 5-3〉 자살 생각 1회 이상을 결과변수의 사건으로 정의 시, 기초 통계량 및 집단 간 비교 결과(자살 생각 없음 vs. 자살 생각 1회 이상) .....	196
〈표 5-4〉 자살 생각 2회 이상을 결과변수의 사건으로 정의 시, 기초 통계량 및 집단 간 비교 결과(자살 생각 1회 이하 vs. 자살 생각 2회 이상) .....	201
〈표 5-5〉 1회 이상의 자살 생각에 대한 단변수 생존 분석 결과 .....	207
〈표 5-6〉 1회 이상의 자살 생각에 대한 다변수 생존 분석 결과(모든 변수 활용) .....	210
〈표 5-7〉 1회 이상의 자살 생각에 대한 다변수 생존 분석 결과(변수선택법 적용) .....	213
〈표 5-8〉 1회 이상의 자살 생각에 대한 다변수 생존 분석 결과(모든 변수 활용, 결측 비율(50%) .....	215
〈표 5-9〉 2회 이상의 자살 생각에 대한 단변수 생존 분석 결과 .....	218
〈표 5-10〉 2회 이상의 자살 생각에 대한 다변수 생존 분석 결과(모든 변수 활용) .....	222
〈표 5-11〉 2회 이상의 자살 생각에 대한 다변수 생존 분석 결과(변수선택법 적용) .....	225
〈표 5-12〉 자살 생각 2회 이상에 관한 생존 모형의 예측된 baseline hazard 함수 .....	227
〈표 5-13〉 자살 생각 2회 이상에 관한 생존 모형의 예측된 baseline hazard 함수 .....	227
〈표 5-14〉 2회 이상의 자살 생각에 대한 다변수 생존 분석 결과(모든 변수 활용, 결측 비율(50%) .....	229

〈표 5-15〉 생존 모형 간 유의성을 지닌 변수들의 비교 .....	255
〈표 6-1〉 정신장애진단도구(K-CIDI) 항목 중 사용 변수의 개수 .....	263
〈표 6-2〉 질문별 문장 단일화 예시 .....	263
〈표 6-3〉 질문별 문맥 조건 추가 예시 .....	264
〈표 6-4〉 sentence-BERT 모델별 configuration .....	267
〈표 6-5〉 (비교 1, 비교 2) 응답에 따른 문장 임베딩 코사인 유사도 .....	268
〈표 6-6〉 (비교 3) 질문-응답 쌍에 따른 문장 임베딩 코사인 유사도 .....	269
〈표 6-7〉 응답 조합별 동일 변수에 대한 코사인 유사도 .....	272
〈표 6-8〉 반대 응답과 가까운 거리의 변수 .....	276
〈표 6-9〉 수정 전후 질문 .....	277
〈표 6-10〉 S(자살) 설문 문항군과 근접한 임베딩 공간의 질문 변수 .....	280
〈표 6-11〉 XGBOOST 정확도(accuracy) 결과 .....	288
〈표 6-12〉 t-SNE 2차원 축소 자료(perplexity=20)의 3회차 샘플에서 혼동 행렬 .....	288
〈표 6-13〉 t-SNE 2차원 자료(perplexity=20)의 3회차 샘플에서 top-10 예측기여도 (shapley value, SHAP) 상위 설문 문항 .....	290
〈표 6-14〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측 기여도(shapley value, SHAP) 상위 설문 문항 .....	291
〈표 6-15〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측 기여도(shapley value, SHAP) 상위 설문 문항 .....	292
〈표 6-16〉 자살 관련 직접 질문 변수 .....	294
〈표 6-17〉 자살 고위험군 세 그룹에 대한 XGBOOST 분류 정확도 .....	294
〈표 6-18〉 혼동 행렬 .....	295
〈표 6-19〉 t-SNE 2차원 자료(perplexity=20)의 1회차 샘플에서 top-10 예측기여도 (shapley value) 상위 설문 문항 .....	296
〈표 6-20〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항 .....	297



〈표 6-21〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항 .....	297
〈표 6-22〉 정신장애진단도구(K-CIDI) 항목 중 사용 변수 .....	299
〈표 7-1〉 예측 모델링 주요 결과 .....	307
〈표 7-2〉 자살 위험 척도 및 특성 .....	310
〈표 7-3〉 정신건강 상황에 따른 대응 모델 .....	312

# 그림 목차

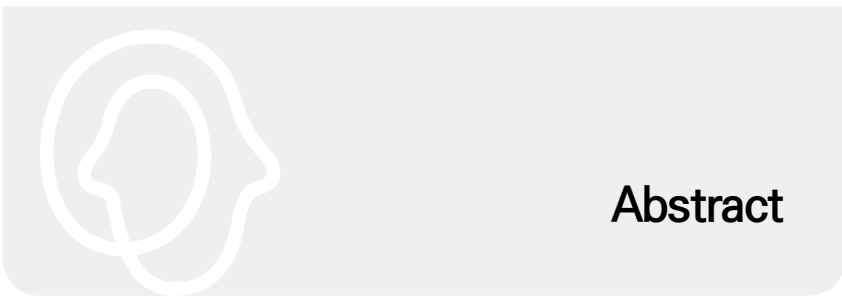
---

[그림 1-1] 자살자 수 및 자살률(2011-2021) .....	13
[그림 1-2] 월별 자살자 수(2019-2021) .....	14
[그림 1-3] OECD 국가 연령표준화 자살률 비교 .....	14
[그림 1-4] 정신질환을 진단받은 대상자의 정신건강서비스 이용률 국가별 비교 .....	15
[그림 1-5] 자살 관련 행동 인포그래픽 .....	16
[그림 1-6] 연령별 자살 생각을 .....	17
[그림 1-7] 성별 자살 생각을 .....	17
[그림 1-8] 성별에 따른 연령대별 자살충동 이유 .....	18
[그림 1-9] 연구 과정 개요 .....	21
[그림 2-1] 정신건강의 사회적 결정요인에 대한 개념화 .....	29
[그림 2-2] 정신건강 위험요인의 중첩 .....	30
[그림 2-3] 정신건강 고위험군 이행 과정 .....	31
[그림 2-4] 제2차 정신건강복지 기본계획의 체계 .....	33
[그림 2-5] 자살 이행과정의 도식화 .....	36
[그림 2-6] 자살 위험 신호 .....	41
[그림 2-7] 자살 원인과 진행 과정에 따른 개입전략 .....	46
[그림 4-1] 남자 중학생 집단 Tree Plot .....	121
[그림 4-2] 여자 중학생 집단 Tree Plot .....	126
[그림 4-3] 남자 고등학생 집단 Tree Plot .....	133
[그림 4-4] 여자 고등학생 집단 Tree Plot .....	137
[그림 4-5] 남자 18~39세 집단 Random Forest Importance Plot .....	147
[그림 4-6] 남자 18~39세 집단 Tree Plot .....	149
[그림 4-7] 여자 18~39세 집단 Random Forest Importance Plot .....	152
[그림 4-8] 여자 18~39세 집단 Tree Plot .....	153
[그림 4-9] 남자 40~64세 집단 Random Forest Importance Plot .....	156
[그림 4-10] 남자 40~64세 집단 Tree Plot .....	158
[그림 4-11] 여자 40~64세 집단 Random Forest Importance Plot .....	161



[그림 4-12] 여자 40~64세 집단 Tree Plot .....	163
[그림 4-13] 남자 65~79세 집단 Random Forest Importance Plot .....	166
[그림 4-14] 남자 65~79세 집단 Tree Plot .....	168
[그림 4-15] 여자 65~79세 집단 Random Forest Importance Plot .....	171
[그림 4-16] 여자 65~79세 집단 Tree Plot .....	173
[그림 4-17] Boosting의 “슬픔·절망감 경험 여부” 상대적 영향도(%) 비교 .....	175
[그림 4-18] Boosting의 “우울 장애 여부” 상대적 영향도(%) 비교 .....	177
[그림 5-1] 자살 생각 2회 이상의 생존 모형을 통해서 얻은 예측된 생존확률로 구분된 위험 그룹 변수에 대한 KM curve .....	228
[그림 5-2] 자살 생각 2회 이상의 생존 모형 with stepwise을 통해서 얻은 예측된 생존확률로 구분된 위험 그룹 변수에 대한 KM curve .....	228
[그림 5-3] 범주형 변수에 대한 자살 생각 1회 이상에 대한 KM curve .....	232
[그림 5-4] 범주형 변수에 대한 자살 생각 2회 이상에 대한 KM curve .....	244
[그림 6-1] sentence-BERT와 BERT의 모델 아키텍처 .....	266
[그림 6-2] 정규분포와 자유도 1인 t분포의 비교 -분포 .....	271
[그림 6-3] B(니코틴 사용에 의한 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화 .....	273
[그림 6-4] D(공포 장애 및 기타 불안 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화 .....	274
[그림 6-5] E(우울 장애 및 기분부전 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화 .....	274
[그림 6-6] J(알코올 사용과 관련된 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화 .....	275
[그림 6-7] K(강박장애 및 외상 후 스트레스 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화 .....	275
[그림 6-8] D(공포 장애 및 기타 불안 장애) 수정된 질문-응답 임베딩 t-SNE 차원 축소 시각화 .....	278

[그림 6-9] J(알코올 사용과 관련된 장애) 수정된 질문-응답 임베딩 t-SNE 차원 축소 시각화 .....	278
[그림 6-10] 질문 설문군의 임베딩 t-SNE 차원 축소 시각화 .....	279
[그림 6-11] 질문 변수 임베딩 t-SNE 차원 축소 해석 .....	281
[그림 6-12] B(니코틴 사용에 의한 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화 .....	283
[그림 6-13] D(공포 장애 및 기타 불안 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화 .....	284
[그림 6-14] E(우울 장애 및 기분부전 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화 .....	284
[그림 6-15] J(알코올 사용과 관련된 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화 .....	285
[그림 6-16] K(강박장애 및 외상 후 스트레스 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화 .....	285
[그림 6-17] t-SNE 2차원 perplexity=20의 3회차 샘플에서 설문 문항의 top-10 예측기여도(shapley value) .....	289
[그림 6-18] t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항 .....	291
[그림 6-19] t-SNE 2차원 자료(perplexity=20)의 1회차 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항 .....	295
[그림 6-20] t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항 .....	298
[그림 7-1] 위기대응 프로세스 예시 .....	314
[그림 7-2] 개인정보 보호 텍스트 분류 .....	317
[그림 7-3] Transfer learning을 활용하여 client의 암호화된 정보를 server에서 연산하는 모델 구성도 .....	318



## Abstract

### **Predictive Modeling of Crisis Response in the Health and Social Welfare Sectors:** Focusing on Mental Health Cases

Project Head: Oh, Miae

Rapid socio-economic changes and the COVID-19 situation have increasingly highlighted the importance of mental health. Mental health is not merely the absence of mental illness or psychological disorders. It is a comprehensive concept that includes an individual's satisfaction with life, adaptability, social relationships, and the ability to perform social roles.

Mental health issues vary in their primary causes depending on sex and age, necessitating in-depth research into the possibilities of early detection focused on specific policy targets. Particularly in the case of high-risk groups for suicide, it is necessary to approach the issue by sex and age group.

This study aims to understand the characteristics of different policy targets and develop, by utilizing public data, predictive models for different age groups (children and adolescents, youth, the middle-aged, and the elderly) of each sex. This will assist in laying the foundation for supporting national policy objectives (enhanced preventive health care: early detection and intervention of mental health issues).

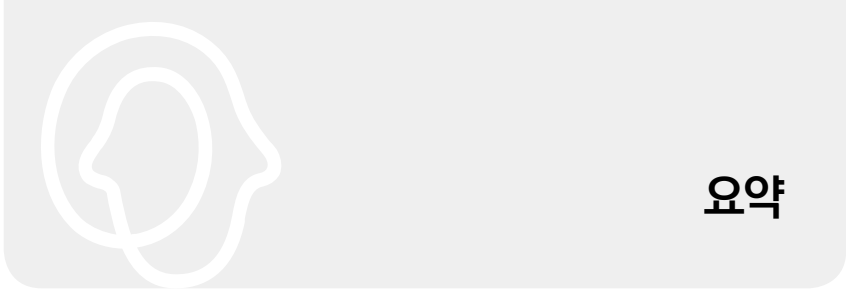
**Keyword :** Crisis Response, High-Risk Suicide Group,  
Predictive Modeling, Survival Analysis, Deep Learning

---

Co-Researchers: Park, Seongjun·An, Suin·Choi, Hosik·Park, Minsu·Jo, Hyun-kyung







## 1. 연구의 배경 및 목적

급격한 사회경제적 변화, 코로나19 상황으로 인해 정신건강 영역은 점점 중요해지고 있다. 정신건강은 단순히 정신병이나 심리적 장애가 없는 상태가 아니라, 개인의 삶에 대한 만족도와 적응력, 사회적 관계와 역할 수행능력 등을 포함하는 포괄적인 개념이다.

정신건강 문제는 성별에 따라, 연령대에 따라 주원인이 다르기 때문에 정책 대상에 초점을 맞추어 조기 발견 가능성 검토에 대한 심층연구가 필요하다. 특히, 자살 고위험군의 특성의 경우, 성별, 연령대를 나누어 접근할 필요가 있다.

본 연구에서는 정책 대상별 특성을 파악하고 다양한 공공데이터를 활용하여 정책 대상별 예측 모델을 개발하는 것을 목적으로 한다. 공공데이터를 활용하여 성별을 고려한 정책 대상별(아동·청소년/청년/중장년/노인) 정신건강 예측 모델을 개발하여 국정과제(예방적 건강관리 강화: 정신건강 문제 조기 발견 및 개입) 지원 기반을 마련하고자 한다.

## 2. 주요 연구 결과

본 보고서 「보건복지분야 위기대응 예측 모델링 연구-정신건강 사례를 중심으로」의 주요 연구 내용은 다음과 같다. 제2장에서는 연구범위 설정을 위한 선행연구로 정신건강 고위험군의 특성과 자살 고위험군의 특성을 살펴보았다. 제2차 정신건강복지 기본계획(보건복지부, 2021)에서는 정신건강 고위험군을 자살, 약물중독, 정신질환 등 광범위하게 설정하였

다. 정신건강 고위험군과 관련된 대표적인 영역은 자살이다. 자살의 경우 자살 시도자에 대한 사후관리를 포함하여 자살 생각·계획, 심각한 우울 및 불안, 자살유가족 지원 등 고위험군의 범위가 상대적으로 넓게 설정되어 있다. 2023년 정신건강 사업안내에서는 정신건강 고위험군을 중증정신질환, 사회적 재난 상황, 중독 문제, 자살 관련 위험 상황에 처한 사람 외에도 청년·여성·노인·장애인 등을 정신건강 취약계층으로 설정하여 고위험군의 조기 발견 및 개입을 강조하였다. 본 연구에서는 자살 고위험군의 특성을 이해하기 위해 자살 행동과 관련된 개념을 정리해보고 자살의 이행과정을 단계별로 살펴보았다. 자살 관련 요인은 생물학적 요인, 인구사회학적 요인, 정신증적 요인, 심리적 요인, 사회적 요인, 지역사회 요인, 대인관계 요인으로 나누어 검토하였다. 인구 특성별 자살 및 정신건강 고위험군의 특성을 파악하기 위해 선행연구를 분석하여 정리하고 자살 고위험군의 정책적 범위도 살펴보았다. 정신건강 관련 정책 중 자살 고위험군에 대한 선별과 자살 위험에 대한 예측은 예방적 정신건강 정책의 핵심적 내용이라 볼 수 있다. 따라서 본 연구에서는 정신건강 위기에 대한 사전적 대응 방안을 모색하기 위하여 자살 행동 관련 요인에 대한 검토와 자살 위험에 대한 예측 모형 탐색을 중심으로 논의하였다.

제3장에서는 예측 모델을 적용하기 위해 다양한 공공데이터를 검토해보았다. 국민건강영양조사, 정신건강실태조사, 청소년건강행태조사, 한국복지패널조사, 한국의료패널조사를 중심으로 검토하였고, 각 공공데이터의 특징과 자살 관련 문항을 상세히 설명하였다. 자살 고위험군 관련 분석 변수를 도출해보고자, '자살'과 공공데이터 키워드를 기준으로 KCI 등재 학술지를 검토하였다. 공통적인 일반적 특성 변수는 연령, 결혼상태, 교육 수준, 경제활동, 소득 및 지역 변수가 있다. 아동·청소년은 학업 관련 변수, 청년은 1인 가구 여부 및 경제 상태 관련 변수, 그리고 여성은

가족 구성이나 자녀 및 결혼 관련 변수 등이 특징적인 ‘일반적 특성 변수’에 해당한다. 일반 성인은 의료비나 만성질환 등 보건·의료 관련 변수, 노인은 보조금이나 공적 서비스 수급 관련 변수 등 제도 관련 변수, 그리고 저소득층의 경우 종교 관련 변수 등이 특징적인 ‘일반적 특성 변수’에 해당한다. 자살 고위험군 관련 선행연구의 실증 분석 결과를 살펴보면 통계적으로 유의한 변수가 다양하게 나타나고, 동일한 데이터를 활용하더라도 모형 설정에 따라 통계적으로 유의한 변수가 다르게 나타나거나 회귀계수의 부호가 다르게 나타나기도 한다. 본 연구에서 검토한 5개의 공공데이터 중에서 문항 수, 응답 빈도 수(표본 수) 및 시계열적 연속성 등을 고려하여 예측 모델 적용 데이터 및 심층 분석 데이터는 ‘정신건강실태조사’, ‘청소년건강행태조사’ 그리고 ‘한국복지패널’로 선정하였다. 제3장에서 검토한 내용을 바탕으로 제4장과 제5장, 제6장에서 실제 예측 모델을 적용하여 그 결과를 제시하였다.

제4장에서는 제3장에서 검토한 공공데이터 중 청소년건강행태조사와 정신건강실태조사를 활용하여 정책집단별 여러 예측 모형을 적용해보았다. 예측 모형은 각 모형마다의 장단점이 있기 때문에 그 집단의 특성을 다각도로 이해하기 위해서는 여러 예측 모형으로 유의미한 요인을 살펴보는 것도 중요한 부분이다. 청소년건강행태조사 분석의 종속변수는 최근 1년 동안 자살 생각을 한 적이 있느냐는 문항을 사용하였다. 정신건강실태조사에서 최근 1년 동안 자살 생각을 한 대상자는 전체의 1.3%로 예측 모형을 적용하기에는 자료의 불균형이 심하기 때문에 정신건강실태조사의 경우 종속변수로 평생 동안 자살 생각을 한 적이 있느냐는 문항을 사용하였다.

분석 결과, 아동·청소년 집단인 청소년건강행태조사 대상자들에서는 전체적으로 주관적 건강 인지, 슬픔·절망감 경험 여부, 외로움 경험 정도,

평상시 스트레스 인지 정도 요인이 자살 생각에 영향을 주는 중요한 요인으로 도출되었다. 각 집단별 특징으로 남자 중학생 집단에서는 최근 폭력으로 인한 치료 경험이 있을수록, 흡연 경험이 있을수록, 타인의 음주로 인한 폐해가 있을수록, 가정의 경제적 상태가 어려울수록, 현재 거주형태가 가족과 함께 거주하지 않을수록, 학년이 낮을수록 자살 생각을 할 가능성이 높았다. 여자 중학생 집단에서는 학습할 목적으로 주말에 앉아서 보낸 평균 시간이 많을수록, 평상시 스트레스 인지가 높을수록, 슬픔-절망 경험이 있을수록, 외로움을 많이 느낄수록, 학교에서 손상 경험이 있을수록, 최근 폭력으로 인한 치료 경험이 있을수록, 음주 경험이 있을수록, 흡연 경험이 있을수록, 치료 목적을 제외하고 습관적 약물 경험이 있을수록, 가정의 경제적 상태가 어려울수록, 현재 거주형태가 가족과 함께 거주하지 않을수록, 학년이 낮을수록 자살 생각을 할 가능성이 높았다. 남자 고등학생 집단에서는 일반계 고등학교일 때, 최근 폭력으로 인한 치료 경험이 있는 경우, 성관계 경험이 있는 경우 자살 생각을 할 가능성이 높았다. 여자 고등학생 집단 모형에서는 최근 폭력으로 인한 치료 경험이 있는 경우, 음주 경험이 있는 경우, 흡연 경험이 있는 경우, 성관계 경험이 있는 경우, 치료 목적을 제외하고 습관적 약물 경험이 있는 경우, 스마트폰 주말 평균 사용 시간이 많을수록, 가족과 함께 거주하지 않는 경우, 가정형편으로 인한 경제적 도움을 받았을 때 자살 생각을 할 가능성이 높았다. 부스팅의 상대적 영향도 분석 결과, 여자 고등학생 집단은 다른 집단에 비해 슬픔-절망감 경험 여부의 비중이 상대적으로 더 높았다.

정신건강실태조사에서는 연령과 성별로 6개의 집단을 나누어 각 정책 집단별 예측 모형을 구축하였다. 전체적으로는 우울장애가 있는 경우, 삶의 질 만족도, 현재 삶에 대한 만족도가 자살 생각에 유의미한 영향을 미치는 요인으로 도출되었다. 남자 18~39세 집단에서는 만연령이 높을수

록, 정신건강 문제로 전문가와 상의한 경험이 있는 경우 자살 생각을 할 가능성이 높았다. 여자 18~39세 집단에서는 정신건강 문제로 전문가와 상의한 경험, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 적이 있는 경우, 연령, 가구 총소득이 자살 생각에 영향을 미치는 중요한 요인이었다. 남자 40~64세의 집단 예측 모형에서는 연령이 높을수록, 자녀가 적을수록, 직업이 없는 경우, 의료급여 수급권자일 경우, 의료급여 수급권자가 아닐 때 가구 총소득이 높을수록, 알코올 사용장애가 있는 경우, 불안장애, 외상 후 스트레스 장애가 자살 생각에 영향을 주는 중요한 요인으로 도출되었다. 여자 40~64세 집단의 예측 모형에서는 부스텝의 상대적 영향도에서 우울장애 유무가 전체의 60%를 차지할 정도로 중요한 요인으로 도출되었다. 알코올 사용장애, 니코틴 사용장애, 외상 후 스트레스 장애가 자살 생각에 영향을 미치는 중요한 요인이라고 볼 수 있다. 남자 60~79세 집단의 예측 모형에서는 가구 총소득, 불안장애, 정신건강 문제로 전문가와 상담한 경험이 있는 경우, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 경험이 있는 경우에 자살 생각을 할 가능성이 높았다. 여자 65~79세 집단 예측 모형에서는 자녀의 수, 가구 총소득, 만연령, 정신건강 문제로 전문가와 상의한 경험이 있는 경우, 알코올 사용장애가 있는 경우 자살 생각을 할 가능성이 높았다. 이러한 집단별 특성 차이는 정책 대상별로 맞춤형 서비스를 제공한다고 할 경우 대상자의 우선순위를 정할 때 유용하게 활용될 수 있을 것이다.

제5장에서는 패널 데이터의 정신건강 관련 문항 분석을 통해 자살 생각을 event로 고려하여 이와 관련된 요인들을 심층적으로 살펴보고자, 한국복지패널 데이터를 활용하여 생존 분석 모형을 적용해보았다. 2012년 7차부터 2021년 16차까지의 한국복지패널조사 데이터를 활용하였고, 생존 모형은 단변수 분석과 다변수 분석으로 나누어 진행하였다. 결

과변수는 총동적 선택을 고려하여 '1회 이상 자살 생각'을 한 경우와 '2회 이상 자살 생각'을 한 경우로 사건을 나누어서 분석하였다. 활용된 모든 다변수 분석에서 유의했던 변수는 우울과 좌절감이었다. 또한, 불면, 무기력, 건강상태, 가부장적 관념 등의 변수들이 그 뒤를 이었다. 자살 생각을 1회와 2회로 각각 정의했을 시 두 모형 간의 차이는 크게 없었으나 2회로 정의할 때 가족관계 만족도, 가족생활 만족도, 자녀들의 형제자매 관계 만족도 등의 변수들이 새롭게 유의성을 갖게 되었다. 자살 생각을 1회로 정의할 때에는 집 점유, 성별, 배우자 관계 만족도 등의 변수들이 유의성을 가졌고, 이는 성별의 차이나 경제적 여유 등의 이유로 인해 발생되는 변수들이 함께 모형에서 유의하였다. 결측 비율을 50% 미만으로 확장할 때 '근로'와 관련된 4가지 변수들이 새롭게 추가됨에 따라 이와 관련된 근로계약, 근로 지속 가능성 등의 변수들이 유의성을 갖게 됨을 확인할 수 있었으며, 4가지 변수 추가 시 직업 만족도, 사회적 친분 관계 만족도, 여가생활 만족도 등의 사회생활과 밀접한 변수들이 새롭게 유의성을 가짐을 확인하였다.

제6장에서는 개별 문항의 집계점수가 가진 정보의 축약성을 극복하고자, 거대언어 모형을 활용한 설문 문항의 문맥을 파악하고 변수화하는 방법을 제안하였다. 정신건강실태조사 데이터를 활용하여 자살 사고, 자살 계획, 자살 시도 대상자 응답 같은 스몰 데이터에서의 정보량을 최대화하기 위한 방안으로 기계학습과 딥러닝 방법을 수행하여 자살 고위험군의 집단적 특성을 파악하고자 하였다. 이 장에서는 실험적으로 범주화된 응답 변수가 아닌 자연어 형태의 질문과 응답 데이터를 활용하여 자살 고위험군과 저위험군 분류 성능, 자살 고위험군 내에서 자살 사고, 자살 계획, 자살 시도 그룹의 분류 성능을 확인해 보았다는 데 의미가 있다.

### 3. 결론 및 시사점

결론 및 시사점은 이 연구의 요약인 결론 파트와 시사점 파트를 정책적 활용 측면과 예측 모형 활용 측면으로 나누어 제시하였다. 정책적 활용 측면은 자살 고위험군에 대한 선별 방안, 정신건강 문제 고위험군에 대한 대응 방안에서 예측 모형이 어떠한 역할을 할 수 있는지를 검토하였고, 자살 고위험군 관련 다차원적 인프라 구축의 필요성을 언급하였다. 예측 모형 활용 측면은 세분화 모형의 장점 및 고려사항, 희소 데이터(sparse data), 개인정보를 보호할 수 있는 예측 모델링으로 나누어 검토하였다.

본 연구는 정신건강 사례에서 맞춤형 복지를 실현하기 위한 근거 기반 기초연구로, 이 연구의 결과가 정책적으로 많이 활용될 수 있기를 기대한다.

주요 용어 : 위기대응, 자살 고위험군, 예측 모델링, 생존 분석, 딥러닝





사람을  
생각하는  
사람들



KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



# 제 1 장

## 서론

제1절 연구의 배경 및 목적

제2절 연구의 내용 및 방법



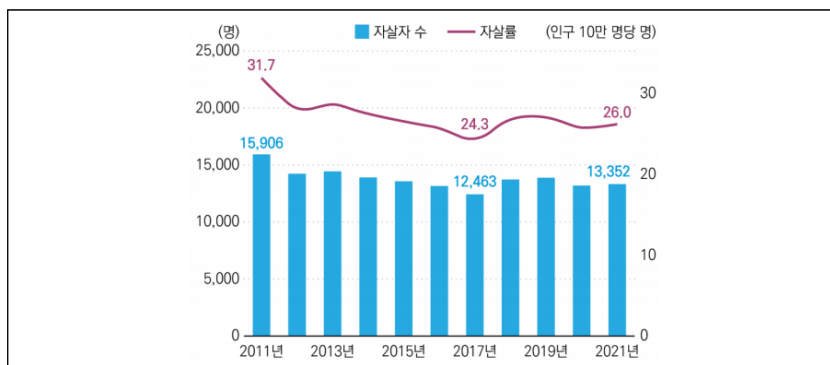
# 제 1 장 서론

## 제1절 연구의 배경 및 목적

### 1. 연구의 배경

급격한 사회경제적 변화, 코로나19 상황으로 인해 정신건강 영역은 점점 중요해지고 있다. 고의적 자해로 인한 사망자 수인 자살률은 정신건강 문제의 심각성과 사회적 비용을 반영하는 지표인데, [그림 1-3]을 살펴보면 우리나라 인구 10만 명당 자살률은 23.6명으로, OECD 국가 중 가장 높은 수준이다. 이렇듯 정신건강 문제가 심각함에도 불구하고 정신질환을 경험한 사람들 가운데 정신건강 서비스를 이용한 비율은 12.1%로 OECD 국가 등 선진국의 절반에 채 미치지 못하는 수준이다.

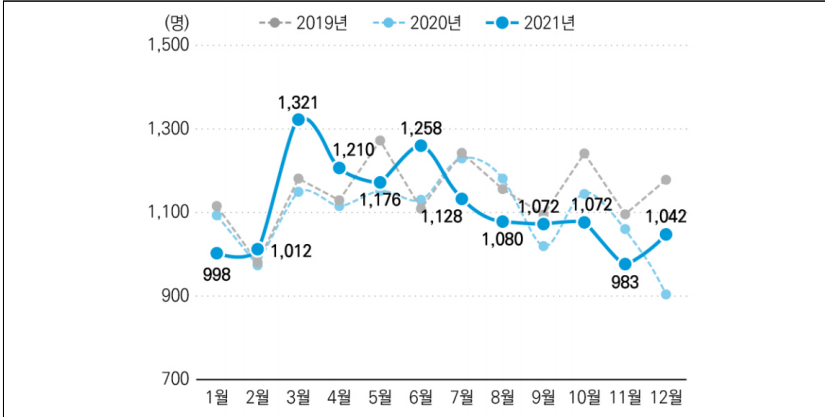
[그림 1-1] 자살자 수 및 자살률(2011-2021)



자료: 국립정신건강센터, 건강보험심사평가원, 한국보건사회연구원. (2022). 국가 정신건강현황 보고서 2021. 국립정신건강센터. p.26.

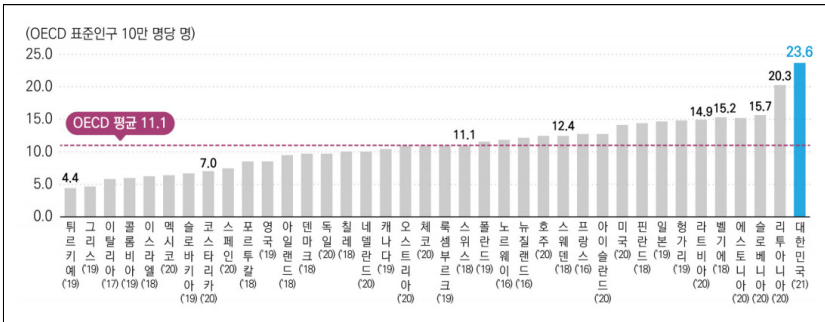
14 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

[그림 1-2] 월별 자살자 수(2019-2021)



자료: 국립정신건강센터, 건강보험심사평가원, 한국보건사회연구원. (2022). 국가 정신건강현황 보고서 2021. 국립정신건강센터. p.26.

[그림 1-3] OECD 국가 연령표준화 자살률 비교

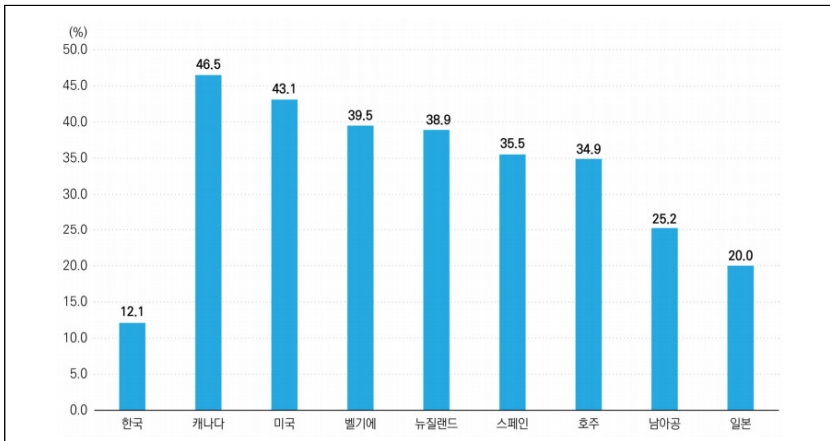


주: OECD 평균은 자료 이용이 가능한 38개 국가의 가장 최근 자료를 이용하여 계산함.  
 자료: 국립정신건강센터, 건강보험심사평가원, 한국보건사회연구원. (2022). 국가 정신건강현황 보고서 2021. 국립정신건강센터. p.26.

정신장애의 유병률 및 정신건강서비스 이용 현황을 파악하기 위해 실시한 ‘정신건강실태조사(2021)’ 결과, 우리나라 국민 3~4명 중 1명은 평생에 한 번 이상 우울이나 불안 등의 정신건강 문제를 경험하고 있으며, 여성이 남성보다 1.5배 이상 높았으며 정신장애 중 가장 흔한 것으로 알

코울 사용장애와 니코틴 사용장애를 꼽았다(국립정신건강센터, 2021). 주요 결과를 살펴보면 정신장애 1년 유병률은 8.5%였으며, 니코틴 사용장애를 제외한 1년 유병률은 6.4%로 여자가 남자보다 1.5배 높았다. 지난 1년간 정신건강 문제를 경험한 사람은 약 355만 명으로 추산되고, 성인 4명 중 1명이 평생 한 번 이상 정신건강 문제를 경험하고 있는 것으로 나타났다. 자살 사고, 자살 계획 및 자살 시도를 살펴보면, 성인의 10.7%는 평생 한 번 이상 심각하게 자살을 생각하며, 2.5%는 자살을 계획하고, 1.7%는 자살을 시도하였다. 지난 1년 간 성인의 1.3%가 한 번 이상 심각하게 자살을 생각하고, 0.5%가 자살을 계획하며, 0.1%가 자살을 시도한 것으로 나타났다. 또한, 자살 생각자의 56.8%, 자살 계획자의 83.3%, 자살 시도자의 71.3%가 평생 한 번 이상 정신장애를 경험한 것으로 나타났다.

[그림 1-4] 정신질환을 진단받은 대상자의 정신건강서비스 이용률 국가별 비교

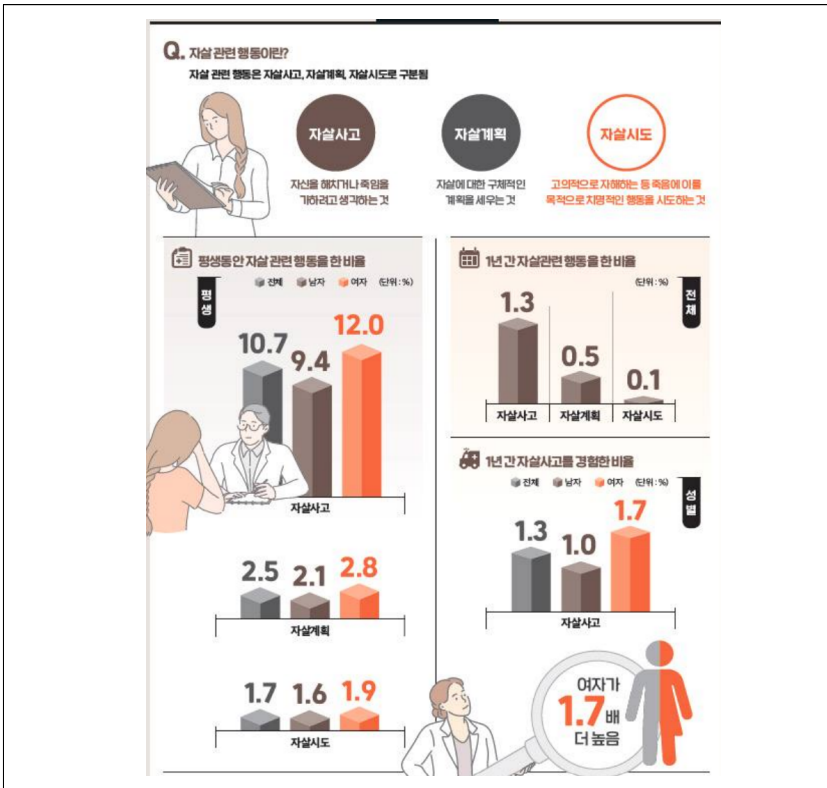


주: 한국은 평생, 그 외 국가는 12개월임.

자료: 국립정신건강센터, 건강보험심사평가원, & 한국보건사회연구원. (2022). 국가 정신건강현황 보고서 2021. 국립정신건강센터. p.29.

보건복지부는 2022년 2분기 코로나19 국민정신건강 실태조사 결과를 코로나19 중앙재난안전대책본부 회의에서 발표하였는데, 이는 코로나19로 인한 국민 정신건강 실태 및 현황을 파악하여 국민에게 필요한 정신건강 서비스를 제공하기 위한 것으로, 그간의 국민 정신건강 현황 및 변화 추이를 파악하는 데 유의미하다. 발표에 따르면 이는 사회적 거리두기 해제 후 첫 조사로, 우울·불안 등 정신건강 지표가 개선되었으며, 소득 감소 및 고립 등 현실적인 문제로 인한 자살 생각률은 증가하였음을 알 수 있다.

[그림 1-5] 자살 관련 행동 인포그래픽

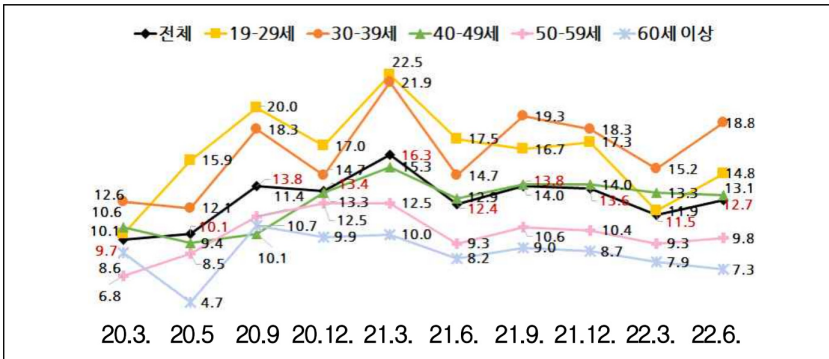


자료: 2021년 정신건강실태조사 통계인포그래픽

실태조사 연구진은 두려움, 불안은 시간이 경과하면서 적절히 감소하고 있지만 우울의 감소 정도는 상대적으로 낮게 나타나고 있음을 확인하였다. 또한, 코로나19 기간 누적된 소득 감소, 고립 등 현실적인 문제로 인해 정신건강이 더 악화되거나 자살이 증가할 우려에 대비하여, 경제적·사회적으로 취약한 계층에 대한 적극적인 지원이 필요함을 강조하였다.

[그림 1-6] 연령별 자살 생각을

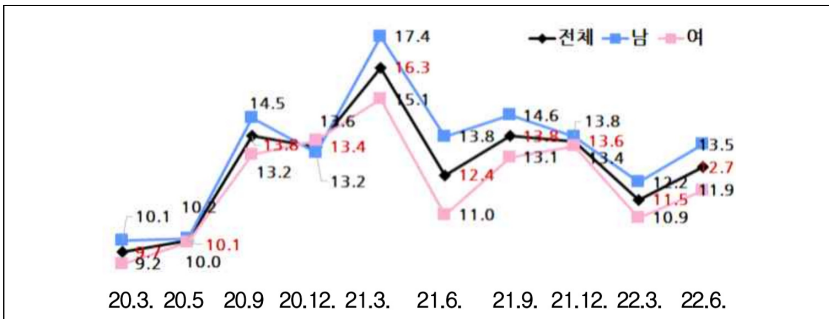
(단위: %)



자료: 보건복지부. (2022. 8. 10). 2022년 2분기 「코로나19 국민 정신건강 실태조사」 결과 발표 보도참고자료. p.3.

[그림 1-7] 성별 자살 생각을

(단위: %)



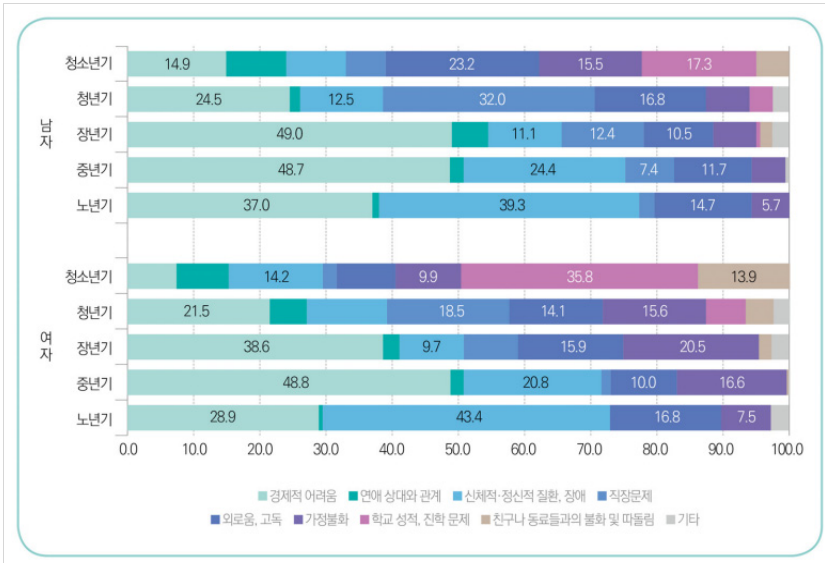
자료: 보건복지부. (2022. 8. 10). 2022년 2분기 「코로나19 국민 정신건강 실태조사」 결과 발표 보도참고자료. p.3.

정신건강은 단순히 정신병이나 심리적 장애가 없는 상태가 아니라, 개인의 삶에 대한 만족도와 적응력, 사회적 관계와 역할 수행능력 등을 포함하는 포괄적인 개념이다. 정신건강은 개인의 특성과 환경 요인이 상호 작용하는 과정에서 형성되며, 변화하고 발전할 수 있는 동적인 상태이며, 긍정적인 측면과 부정적인 측면으로 구분할 수 있는데, 긍정적인 측면은 자기효능감, 자아존중감, 자기결정력, 목표지향성, 삶의 의미와 목적 등을 포함한다(김동배, 안인경, 2004).

정신건강 문제는 성별에 따라, 연령대에 따라 주원인이 다르기 때문에 정책 대상에 초점을 맞추어 조기 발견 가능성 검토에 대한 심층연구가 필요하다. 자살 고위험군의 특성의 경우, 성별, 연령대를 나누어 접근할 필요가 있다.

[그림 1-8] 성별에 따른 연령대별 자살충동 이유

(단위: %)



자료: 보건복지부, & 한국생명존중희망재단. (2021). 2021 자살예방백서, p.204.



남자 청소년기의 자살충동 이유 중 가장 큰 비중을 차지하는 것은 ‘외로움, 고독’이고, 여자 청소년기의 자살충동 이유 중 가장 큰 비중을 차지하는 것은 ‘학교 성적, 진학 문제’로, 선제적 정신건강예방 관리를 위해서는 맞춤형 접근 필요하다.

국정과제에서는 ‘예방적 건강관리 강화(67)’로 신체건강뿐 아니라 마음건강에 대한 투자를 확대하여 정신질환자·자살 고위험군 지원 강화 및 정신건강 문제 대응체계를 확립하고자 한다.

## 2. 연구의 목적

빅데이터를 활용하여 위기 요인이 있는 대상자를 선제적으로 예측·발굴하는 예측 모델은 복지 사각지대 발굴시스템, 위기아동 발굴시스템에서 정책적으로 활용되고 있다.

복지 사각지대 발굴 시스템은 단전·단수·사회보험료 체납 등의 정보를 수집·분석하여 복지 사각지대 대상자를 선제적으로 발굴·지원하는 시스템이다. 위기아동 발굴시스템은 학대위험 아동 및 가구를 조기에 발견할 수 있는 체계를 마련함으로써 선제적으로 아동학대 발생을 예방하고 아동 및 가족의 상황에 부합한 적절한 아동복지서비스를 제공할 수 있는 지원 시스템이다. 두 시스템 모두 정신건강 고위험군을 예측·발굴하기에는 정신건강 관련 정보가 부족한 상황이다. 정신건강 관련 시스템으로 ‘정신건강 사례관리 시스템’이 별도로 있으나 연구용으로 활용하기에는 물리적, 데이터 품질 측면에서 한계가 있다.

정신건강과 관련하여 자살 고위험군의 특성을 분석하여 고위험군 발굴을 위한 초기 예측 모델을 오미애, 전진아, 김은하, & 진재현 (2020)에서 개발하였으나, 다양한 데이터, 최신 방법론을 활용하여 더 심층적이고 다

각도로 분석할 필요성이 있다.

선행연구에서는 건강보험 자료를 활용하여 모형을 구축하였으나, 다양한 실태조사를 검토하여 정신건강 고위험군의 특성을 파악할 필요가 있다. 예측 관련 최신 방법론을 적용하여 예측 모델링의 성능을 검토하고 AI 기반 예측 모델링 정책 적용 가능성도 함께 모색하고자 한다.

본 연구는 정책 대상별 특성을 파악하고 다양한 공공데이터를 활용하여 정책 대상별 예측 모델을 개발하는 것을 목적으로 한다. 공공데이터를 활용하여 성별을 고려한 정책 대상별(아동·청소년/청년/중장년/노인) 정신건강 예측 모델을 개발하여 국정과제(예방적 건강관리 강화: 정신건강 문제 조기 발견 및 개입) 지원 기반을 마련하고자 한다.

## 제2절 연구의 내용 및 방법

본 보고서 「보건복지분야 위기대응 예측모델링 연구\_정신건강 사례를 중심으로」의 주요 연구 내용은 다음과 같다. 제2장에서는 연구범위 설정을 위한 선행연구로 정신건강 고위험군의 특성과 자살 고위험군의 특성을 살펴보았다. 제3장에서는 예측 모델 적용을 위해 활용 가능한 다양한 공공데이터를 검토해보았다. 국민건강영양조사, 정신건강실태조사, 청소년건강행태조사, 한국복지패널조사, 한국의료패널조사를 중심으로 검토하였고, 각 공공데이터의 특징과 자살 관련 문항을 상세히 설명하였다. 자살 고위험군 관련 분석 변수를 도출해보고자, ‘자살’과 공공데이터 키워드를 기준으로 KCI 등재 학술지를 검토하였다. 제4장에서는 제3장에서 검토한 공공데이터 중에서 청소년건강행태조사와 정신건강실태조사를 활용하여 정책집단별 여러 예측 모형을 적용해보았다. 제5장에서는

패널 데이터에서의 정신건강 관련 문항 분석을 통해 자살 생각을 event 로 고려하여 이와 관련된 요인들을 심층적으로 살펴보고자, 한국복지패 널 데이터를 활용하여 생존 분석 모형을 적용해보았다. 제6장에서는 개 별 문항의 집계점수가 가진 정보의 축약성을 극복하고자, 거대언어 모형

[그림 1-9] 연구 과정 개요

연구 단계	내용		연구 방법 및 분석데이터	
	1장 서론	연구의 배경 및 목적	연구의 내용 및 방법	전문가 자문회의
2장 정신건강 및 자살 고위험군에 대한 이론적 검토	정신건강 고위험군의 특성		문헌연구	
	자살 고위험군의 특성			
	소결			
3장 예측 모델 적용을 위한 공공데이터 검토	검토 대상 공공데이터 개요		연구진	
	자살 고위험군 관련 선행연구 실증 분석 변수			
	소결			
4장 공공데이터를 활용한 정책 대상별 예측 모델	청소년건강행태조사를 활용한 예측 모델		연구진  청소년건강행태조사 정신건강실태조사	
	정신건강실태조사를 활용한 예측 모델			
	소결			
5장 심층분석 I : 자살고위험군에 대한 위험 인자 탐색을 위한 패널데이터 생존 분석 모형 적용	한국복지패널 분석 개요		연구진  한국복지패널조사	
	통계량 산출 및 집단 간 비교			
	생존 분석 결과			
	소결			
6장 심층분석 II : 텍스트 데이터 기반 자살 고위험군 탐지 를 위한 딥러닝 적용	텍스트마이닝 및 자연어처리 기반 분석 방법		연구진  정신건강실태조사	
	텍스트 데이터의 그룹화 분석 결과			
	문장 임베딩을 활용한 분석 결과			
	소결			
7장 결론 및 시사점	결론	시사점 : 정책적 활용 & 예측모형 활용	전문가 자문회의	연구진 논의

을 활용한 설문 문항의 문맥을 파악하고 변수화하는 방법을 제안하여 정신건강실태조사 데이터를 분석하였다. 마지막으로 제7장 결론에서는 이 연구의 시사점을 정책적 활용 측면과 예측 모형 활용 측면으로 나누어 제시하였다. 정책적 활용 측면은 자살 고위험군에 대한 선별 방안, 정신건강 문제 고위험군에 대한 대응 방안에서 예측 모형이 어떠한 역할을 할 수 있는지를 검토하였고, 자살 고위험군 관련 다차원적 인프라 구축의 필요성을 언급하였다. 예측 모형 활용 측면은 세분화 모형의 장점 및 고려사항, 희소 데이터(sparse data), 개인정보를 보호할 수 있는 예측 모델링으로 나누어 검토하였다.

이 보고서 작성을 위해 국내외 문헌 연구, 전문가 자문회의, 데이터 분석(R/Python) 등 다양한 방법을 활용하였다.



## 제2장

### 정신건강 및 자살 고위험군에 대한 이론적 검토

제1절 정신건강 고위험군의 특성

제2절 자살 고위험군의 특성

제3절 소결



## 제 2 장

# 정신건강 및 자살 고위험군에 대한 이론적 검토

### 제1절 정신건강 고위험군의 특성

#### 1. 정신건강의 개념

정신건강 관련 위기대응 시 정신건강 고위험군에 대한 범위를 설정할 필요가 있으며, 이에 따라 정신건강 및 정신건강 고위험군에 대한 개념을 살펴보고자 하였다. 먼저, 정신건강이란 정신적 안녕을 의미하며, 이는 한 개인이 자신의 능력을 실현하고, 일상적인 삶의 스트레스를 대처하며 생산적 활동이 가능하고, 자신이 속한 지역사회에 기여할 수 있는 건강한 상태를 의미한다. 여기서 건강이란 병이나 질환이 없는 상태를 넘어 신체적·정신적·사회적 안녕이 이루어진 상태를 의미한다(WHO, 2020).

정신적으로 건강한 상태는 행복감, 자기존중감 등과 같은 긍정적 정서 상태와 관련성을 가진다(Bunton, Macdonald, & Macdonald, 2003). 정신적으로 건강한 상태를 3가지 측면으로 설명하면, 첫째, 개인이 자신이 가진 가능성을 충분히 발휘하여 자아를 실현할 수 있는 상태이다. 둘째, 개인 주변을 둘러싸고 있는 환경을 자신의 힘으로 통제할 수 있는 상태이다. 셋째, 개인이 경험하는 문제 상황을 인식하고 직면하며 이를 해결할 수 있는 자율성을 가진 상태를 일컫는다. 반면 정신건강은 신체적·사회적 측면에서 건강하지 않은 상태와도 관련성을 가진다(Walker, Moodie, & Herrman, 2004). 또한, 정신건강은 신체적·사회적 건강 등 다양한 건강 영역과도 관련성을 가진다(Walker, Moodie, & Herrman,

2004). 대표적으로 우울증 및 불안증은 각종 신체적 질병과도 관련성을 가지는 것으로 알려져 있다. 또한 건강행태와 관련하여, 각종 물질의 오남용이나 폭력, 사회적 기능을 수행하지 못하는 상태 등이 개인의 정신건강과 관련된다.

한편, 정신건강의 개념을 6가지 모델로 분류하여 설명하면 아래와 같다(Vaillant GE, 2003). 첫째, 정신건강을 정상 이상의 상태로 보는 관점(Mental Health as Above Normal)이다. 이는 정신건강을 이상(異常) 상태로 정의하여, 특정 지수를 활용하여 정량화가 가능하며, 이를 바탕으로 정신건강의 수준을 측정할 수 있다고 보는 관점이다. 이 관점에서 치료적 개입은 특정 지표의 변화에 초점이 맞추어진다.

둘째, 긍정적 심리학으로 보는 관점(Mental Health as Positive Psychology)이다. 여기서 정신건강은 개인이 사회적 측면에서 충분히 기능할 수 있는 상태를 의미한다. 다시 말해 자신의 삶 속에서 나타나는 상황에 대응하고, 스트레스를 대처하며, 자신의 능력을 발휘할 수 있는 상태를 일컫는다. 이러한 접근은 인본주의 심리학 등 관련 학문의 근간이 되지만 치료적·정책적 개입의 명확성이 다소 떨어진다는 의견이 있다.

셋째, 성숙으로 보는 관점(Mental Health as Maturity)이다. 이는 정신건강 상태를 하나의 인간 발달과정으로 이해하는 관점이며, 대표적으로 에릭슨(Erikson)의 발달단계 등이 이에 속한다. 해당 관점에서 말하는 성숙은 뇌의 발달 외에도 시간의 흐름에 따른 정서적·사회적 기능의 발달을 의미하기 때문에 다면적인 평가가 요구된다.

넷째, 사회-정서 지능으로 보는 관점(Mental Health as Social-Emotional Intelligence)이다. 여기서 정신건강이란 자신의 정서 수준을 적절하게 평가하고 사회적으로 용인될 수 있는 방식으로 표현하는 기능을 의미한다. 이는 1) 자신의 감정을 정확하게 인식하는 능력,



2) 불안 및 우울에 적절히 대처하는 능력, 3) 타인의 감정을 적절하게 인식하고 반응하는 능력, 4) 타인과의 친밀한 관계를 유지하는 능력, 5) 자신의 충동에 대처하고 만족감을 지연시킬 수 있는 능력 등을 통하여 확인할 수 있다.

다섯째, 주관적 안녕으로 보는 관점(Mental Health as Subjective Well-Being)이다. 개인의 주관적 행복감 수준이 높은 경우를 정신적으로 건강하다고 본다. 주관적 안녕 수준은 사회환경의 요인에 크게 영향을 받으며, 사회적 분위기나 종교 등의 내재화를 통하여 변화될 수 있다. 또한 자아존중감, 자기효능감 등 긍정적 심리 상태와 상당한 관련성을 가지고 있는 것으로 알려져 있다.

여섯째, 회복탄력성으로 보는 관점(Mental Health as Resilience)이다. 이는 내·외부적 스트레스 발생에 적극적으로 대처하고 극복할 수 있는 상태를 의미하며, 특히 외부에서의 부정적 자극을 적절한 방식으로 반응하여 사회관계를 유지하는 적극적 방식을 강조한다. 이 관점에서는 심리학 영역에서 긍정적 기능의 방어기제가 중요하게 평가된다.

## 2. 정신건강 고위험군

정신건강 고위험군은 정신건강에 대한 다양한 해석에 따라 다르게 정의될 수 있다. 정신건강의 개념 정의에 근거하여 정신건강 고위험군을 설명하면, 자신의 능력을 실현하고 일상적인 삶의 스트레스를 대처하고 생산적 활동을 하는 데 어려움을 겪으며, 자신이 속한 지역사회에 기여할 수 있는 건강한 상태를 유지하는 데 위험요인을 가진 집단으로 설정할 수 있다(정진욱 외, 2012).

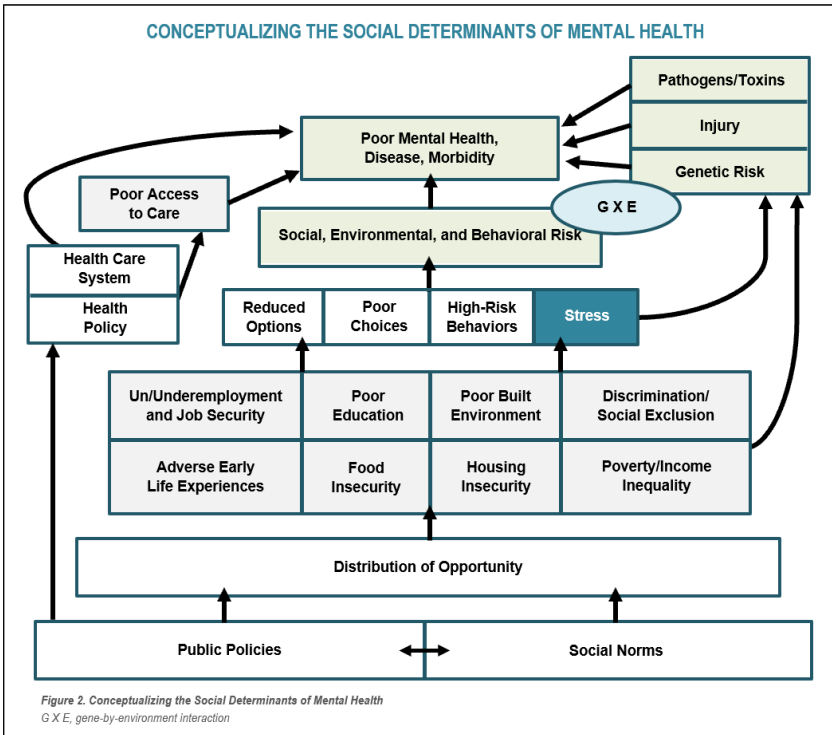
앞에서 제시한 정신건강을 설명하는 모델을 기준으로 고위험군을 설정

하면, 첫째, 정신건강 관련 측정도구를 활용하여, 이상 상태가 확인된 집단, 즉 정신건강의 상태를 확인할 수 있는 각종 측정도구를 활용하여 이상 상태로 볼 수 있는 결과나 정신질환의 진단 기준에 근거하여 정신건강 상태가 좋지 않다고 판단된 집단을 의미한다. 둘째, 사회적 측면에서의 기능을 충분히 이행하는 데 어려움을 겪는 경우이다. 다시 말해 자신의 삶 속에서 경험하는 상황에 적절하게 대응하고 사회적 기능을 수행하는 데 어려움을 겪는 집단을 의미한다. 셋째, 정서적·사회적 기능이 개인의 발달단계에 맞게 발휘되지 못하는 상태에 이른 집단을 의미한다. 넷째, 사회적으로 용인될 수 있는 방식으로 사회·정서적 반응을 하지 못하는 집단을 말한다. 즉, 자신의 감정을 정확하게 인식하지 못하면서 부정적 상태에 대해 적절하게 대처하지 못하는 경우, 타인의 감정을 인식하고 친밀한 관계를 유지하는 데 어려움을 경험하는 경우 등을 말한다. 다섯째, 주관적 안녕감이 낮은 집단, 즉 주관적으로 자신이 행복하다고 느끼지 못하는 경우나 자아존중감·자기효능감, 삶에 대한 주체성 등이 낮은 집단을 의미한다. 여섯째, 회복탄력성이 낮은 집단을 말한다. 다시 말해 내·외부적 스트레스 상황에 적절하게 대처하지 못하거나 스트레스 상황 이후 이를 극복하는 데 어려움을 겪는 경우나 사회적 관계를 유지하는 방식으로 자신에 대한 심리적 방어 기전을 활용하지 못하는 경우를 말한다.

이와 다르게, 정신건강 고위험군은 정신건강 관련 위험요인을 중심으로 설명할 수 있다. 먼저, 정신건강 문제를 발생시킬 수 있는 유전적 위험이나 관련 질병, 건강행태 문제 등이 개인의 정신건강에 영향을 미칠 수 있으며, 사회적·환경적·행동적 위험요인 등이 이러한 요인과 상호작용하여 개인의 정신건강에 영향을 미칠 수 있다. 각종 스트레스 상황이나 사회적·경제적으로 취약한 상태 등은 정신건강의 사회적 결정요인이라 볼 수 있다. 좀 더 거시적 관점에서는 관련 정책이나 사회적 규범 등이 이에

영향을 미칠 수 있으며, 개인의 정신건강을 증진시키기 위한 각종 보건복지 서비스 등이 보호요인으로 작용할 수 있다.

[그림 2-1] 정신건강의 사회적 결정요인에 대한 개념화



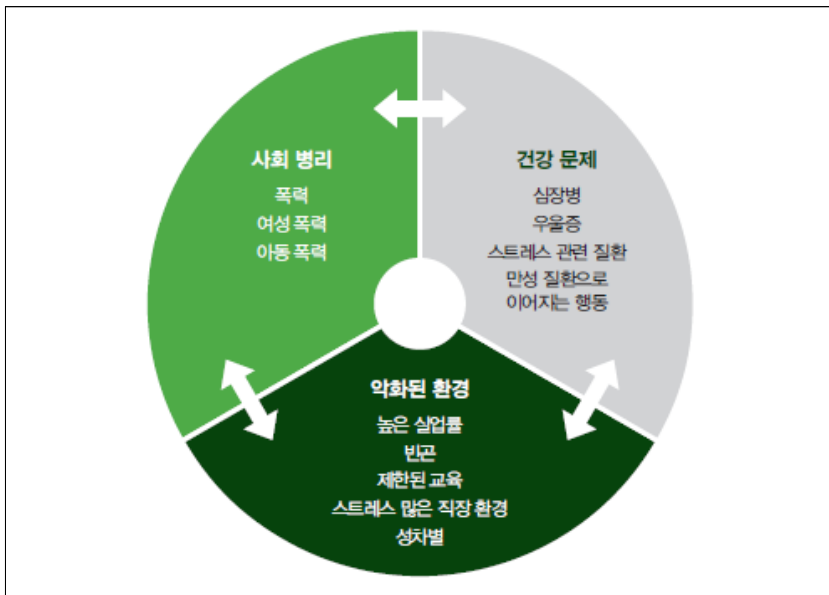
자료: Compton, M. T., & Shim, R. S. (2015). The social determinants of mental health. *Focus*, 13(4), 419-425. p.420.

### 3. 정신건강 고위험군의 이행

신체적·정서적 요인 외에도 사회적·경제적 요인 등 다양한 정신건강 위험요인은 상호 영향을 미치며, 정신건강에 부정적인 영향을 미치는 상황은 중첩되어 나타나는 경향이 있다. 이 때문에 정신건강의 위험 수준을

진단하거나 정신건강 문제에 예방적 접근을 할 시에는 다각적 접근이 요구된다. 예를 들어 개인의 사회·경제적 취약성은 지역 내 사회문제와 높은 관련성을 가진다. 이러한 상황 속에서 개인의 신체적·정신적 건강 문제가 발생하며, 이에 대응하기 위한 사회적·제도적 방안이 마련되지 못한 경우에 정신건강의 문제가 심화될 수 있다. 이러한 관점에서 보았을 때, 정신건강 고위험군의 특성을 이해하기 위해서는 다차원적으로 접근할 필요가 있으며, 고위험군에 대해 진단하거나 개입 등을 할 경우에 정신건강 위험요인의 복합성이 고려되어야 한다.

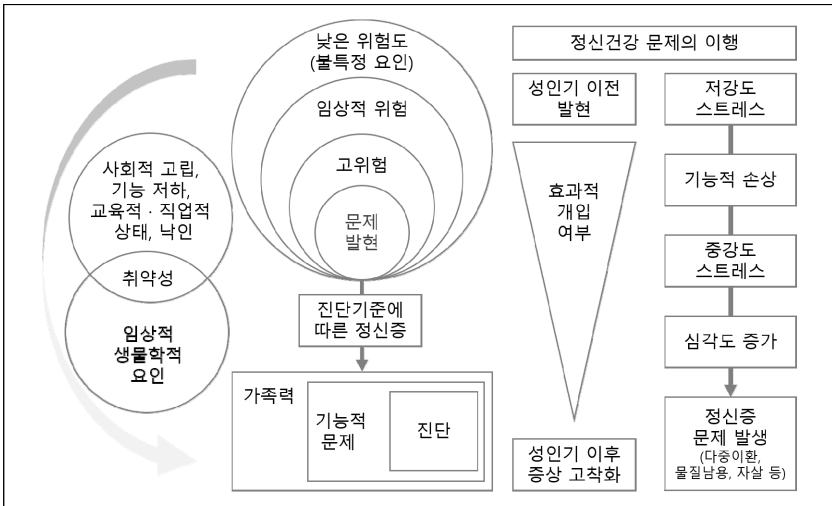
[그림 2-2] 정신건강 위험요인의 중첩



자료: Saxena, S., Izutsu, T., & Tsutsumi, A. (2008). 정신건강증진: 개념·증거·실천(김민석 외 역). 서울: 포널스출판사. p.240.

정신건강과 관련된 여러 위험요인들은 전 생애에 걸쳐 영향을 미치는데, 생애 초기에 나타난 정신건강 문제는 점진적으로 심화되어 정신건강 고위험군으로 이행하는 데 영향을 미친다. 또한 정신건강 관련 위험요인은 여러 상황 속에서 질병으로 이행하는 데 영향을 미치기 때문에 정신건강 관련 문제에 대한 조기 발견과 개입이 중요하다. 특히 정신건강 관련 문제들은 여러 증상이 동시에 나타나는 경향이 있으며, 물질남용이나 자살 등과 같은 문제를 동반할 수 있다. 이뿐만 아니라 스트레스 상황에서 적절한 개입이 이루어지지 않는 경우에는 정신건강 문제가 만성화되거나 증상의 심각도 수준이 높아질 수 있다.

[그림 2-3] 정신건강 고위험군 이행 과정



주: 원 자료의 도표를 번역, 편집 등 수정함.

자료: Colizzi, M., Lasalvia, A., & Ruggeri, M. (2020). Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care?. *International journal of mental health systems*, 14(1), 1-14. p.4.

#### 4. 정신건강 고위험군의 범위

정신건강 고위험군의 범위를 설정하는 것은 이상에서 논의한 바와 같이 불분명하다. 따라서 정신건강증진 또는 정신건강 고위험군에 대한 정책 범위 설정에서는 일정 수준의 정의를 채택하는 것이 타당하다고 볼 수 있다(정진욱 외, 2012). 이에 따라 우리나라의 정책 특성을 바탕으로 정신건강 고위험군의 범위를 검토하였다.

먼저, 제2차 정신건강복지 기본계획에서는 정신건강 고위험군을 자살, 약물중독, 정신질환이 있는 사람 등으로 광범위하게 설정하고 있다. 정신건강 고위험군과 관련된 대표적인 영역은 자살이다. 자살은 자살 시도자에 대한 사후관리를 포함하여 자살 생각·계획, 심각한 우울 및 불안, 자살 유가족 지원 등 제2차 기본계획에서 고위험군의 범위가 상대적으로 넓게 설정되어 있다. 한편, 약물중독 영역에서는 각종 약물중독 사례에 대한 지속적 관리 대상을 고위험군으로 설정하였다. 정신질환 영역에서는 진단을 통하여 고위험군을 설정하며, 그 외에도 고위험군에 속할 가능성이 있는 경우를 조기에 발견하기 위하여 정신건강 위험요인에 대한 관리 범위를 폭넓게 설정하고 있다. 제2차 기본계획에는 코로나19 등 사회적 재난 상황에서의 심각한 우울 및 불안 발생 등도 정신건강 고위험군 범위에 포함되어 있다.

그 밖에 2023년 정신건강 사업안내에서는 정신건강 고위험군을 중증 정신질환, 사회적 재난 상황, 중독 문제, 자살 관련 위험 상황 외에도 청년·여성·노인·장애인 등을 정신건강 취약계층으로 설정하여 고위험군의 조기발견 및 개입을 강조하고 있다. 특히 관련 사업내용에는 정신건강 고위험군에 대한 모니터링과 사후관리, 스크리닝을 통한 조기 대응 방안 등이 포함되어 있다.

[그림 2-4] 제2차 정신건강복지 기본계획의 체계



자료: 보건복지부 보도자료. (2021.1.14.). 대한민국 정책 브리핑(온국민 마음건강 종합대책 브리핑). p.3. <https://www.korea.kr/briefing/policyBriefingView.do?newsId=156431976>에서 2023.12.27. 인출

## 제2절 자살 고위험군의 특성

### 1. 자살 행동 관련 개념

자살 행동을 설명하기 위한 다양한 개념으로 주로 자살, 자살 생각, 자살 시도, 자살 계획, 자해 등이 활용되고 있다(Jakobsen. et al, 2023). 먼저 자살(suicide)은 죽음에 대한 의지를 가지고 자신을 해쳐 죽음에 이

르게 한 행위를 의미한다. 자살 생각(suicide ideation)은 자신의 목숨을 끊기 위한 직접적 행위를 하지는 않았으나 자살을 원하거나 이를 고려한 경험이 있음을 의미한다. 자살 계획(suicide plan)은 자살 시도와 관련된 구체적인 절차를 구상한 상황을 말한다. 자살 시도(suicide attempt)는 죽음에 이를 정도로 치명적이지는 않지만 죽음에 이르고자 의도한 행위를 의미한다. 자살은 자기 보고식 정보수집이 어렵기 때문에 자살 행동을 설명하는 용어들이 활용되고 있다. 예를 들어 자살 위험군을 설정할 때나 자살의 영향요인을 규명할 시에 자살 생각, 자살 계획, 자살 시도 등의 개념이 주요 변인으로 활용된다.

한편, 자살 시도와 자해는 유사한 의미를 가지지만 구분될 필요가 있다. 자해(Self-harm)는 치사율이 낮은 방식으로 자신의 신체에 위해를 가하는 경우가 많으며, 개인의 심리적 고통을 표현하는 방식으로 활용되는 경우가 많다. 자살을 목적으로 자신에게 위해를 가하는 것은 아니지만 위해 정도가 심할 경우에는 죽음에 이를 수 있다. 이러한 경우에는 자해와 자살 시도를 구분하기 어렵다. 이 때문에 자살 시도와 자해를 구분하기 위하여 비자살적 자해('non-suicidal' self-harm), 준자살 행동(parasuicide) 등의 개념이 활용되고 있으며, 실제 자살의 의도를 가지고 행동을 하였으나 자살에 이르지 못한 경우를 중단된 자살(aborted suicide)로 표현하는 경우도 있다. 이 외에도 자살 위험과 관련된 용어로는 습관성 비자살적 자해(habitual non-suicidal self harm), 자기 신체절단(self-mutilation), 고의적 자해 증후군(deliberate self-harm syndrome), 자기 학대(self-abuse), 자기 부상(self-wounding) 등이 있다 (김수진, & 김봉환, 2015).

요약하면, 자살 위험을 설명하기 위하여 자살 행동과 관련된 개념을 활용할 경우에 각 개념의 차이가 고려될 필요가 있다. 자살은 심리부검 등



과 같은 사후 보고형식 외에 그 특성을 파악하는 데 한계가 있기 때문에 자기 보고식 조사가 가능한 자살 행동 관련 개념이 주로 활용되고 있는 실정이다. 자살 행동은 자살 위험성을 설명하는 데 활용될 수 있으나 자살 이행 여부와는 다소 차이가 있으며, 특히 비습관적 자살 시도로 인한 자살 이행 사례는 자살 위험 징후를 파악하기 어렵다. 다만 자살 행동 관련 개념은 자살 이행의 가능성을 증대시키고, 자살 위험요인 간의 관련성이 크기 때문에 간접적으로 자살 위험 수준을 측정할 수 있다는 점에서 활용 가능성이 높다고 볼 수 있다. 따라서 자살 위험 수준의 측정 또는 자살 위험군의 구분 등에 자살 행동을 설명하는 개념들이 대안으로 활용될 수 있다.

## 2. 자살의 이행과정

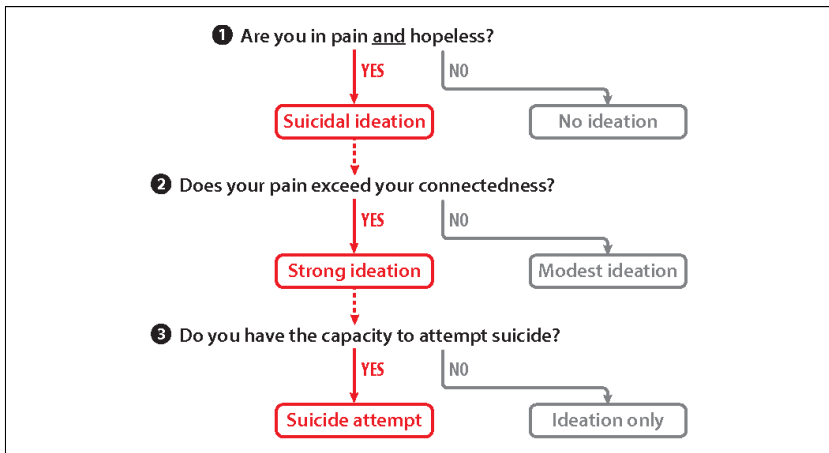
자살의 이행과정을 설명하는 이론 가운데 ‘3단계 이론(3-Step-Theory, 3ST)’에서는 자살 시도로 전이하는 과정은 자살 생각의 촉발, 자살 생각의 심화, 자살 시도로의 전이로 구분하고 있다(Klonsky, May, & Saffer, 2016). 첫 번째 단계는 고통과 무기력감에 의하여 자살 생각이 촉발됨을 말한다. 이는 일상생활에서 경험하는 다양한 스트레스와 사회적 배제, 타인과의 관계에서 발생하는 갈등 등에 기인할 수 있다.

두 번째 단계는 심리적 고통으로 인하여 유대감이 잠식되는 상황을 말한다. 가족, 지인 등 타인과의 관계나 가치 있는 사회적 역할 등에 대한 수용 정도가 자살 위험을 감소시키는 보호요인(유대감)으로 작용한다. 그러나 첫 번째 단계에서 겪는 심리적 고통이 보호요인으로서의 유대감이 가지는 효과를 넘어설 경우 자살 생각이 심화될 수 있다.

세 번째 단계에서는 자살 생각이 자살 시도로 전이되는 단계를 말한다.

이때 기질적 요인(dispositional factors), 실행적 요인(practical factors), 습득적 요인(acquired factors)이 영향을 미치게 된다. 기질적 요인은 고통에 대한 인내 수준이 낮거나 죽음에 대한 두려움이 높은 경우 등 태생적·유전적 특성에 기인한다. 실행적 요인은 자살 행동과 관련된 정보에 대한 접근성, 총기·약물 등 치명적인 자살 수단에 대한 접근성 등을 일컫는다. 습득적 요인은 생애 과정 중 고통이나 주변 사람들의 죽음 등에 대한 경험을 통하여 학습하는 능력을 의미한다.

[그림 2-5] 자살 이행과정의 도식화



자료: Klonsky, E. D., May, A. M., & Saffer, B. Y. (2016). Suicide, suicide attempts, and suicidal ideation. Annual review of clinical psychology, 12, 307-330. p.320.

### 3. 자살 관련 요인에 대한 검토

자살은 정신건강 관련 요인과 마찬가지로 개인 내적 요인 외에도 사회·경제적 환경 요인의 영향을 받는다. 이에 따라 자살과 관련된 주요 요인들은 생태학적 관점에서 설명되고 있으며, 각 영역 내 요인들이 가지는

작용에 따라 위험요인과 보호요인으로 구분한다.

생태학적 관점에서 자살에 영향을 미치는 요인은 크게 생물학적 요인, 인구사회학적 요인, 정신증적 요인, 심리적 요인, 사회적 요인, 지역사회 요인, 대인관계 요인으로 구분될 수 있다(Cramer, & Kapusta, 2017). 생물학적 요인 가운데 성별은 자살 행동과 관련성을 가진다. 남성은 자살 성공률이 높은 반면, 여성은 자살 시도율이 상대적으로 높다. 또한 신경 전달물질인 세로토닌의 기능상 문제로 인하여 우울·불안·초조 등이 발생하고 이에 따라 자살 행동에 이를 수 있다. 또한 자살 행동의 가족력도 주요 위험요인으로 구분된다. 반면에 항우울제 등과 같은 약물치료나 기분 안정을 위한 개입 등은 보호요인으로 작용한다.

인구사회학적 요인 가운데 성 소수자로서의 정체성이나 자살 행동에 허용적인 종교관, 사회경제적으로 취약한 인종이나 고령의 경우에 자살 위험 수준이 상대적으로 높다고 볼 수 있다. 또한 고위험 직업에 종사하는 경우, 총기 등과 같은 자살 수단의 확보, 실업이나 실직, 경제적 문제, 차별경험 등도 주요 위험요인으로 구분된다. 반면에 위에서 설명된 상황에 해당되지 않은 경우나 생명을 중시하는 종교관 및 가치관을 가진 경우에는 보호요인으로 작용할 수 있다.

정신증적 요인으로는 우울증, 조현병, 성격장애 등 정신질환이 주된 위험요인으로 구분된다. 또한 각종 물질 남용 등과 같은 행동적 문제도 이에 포함된다. 반면에 정신증적·행동적 문제에 대한 관리나 치료는 보호요인으로 작용한다.

심리적 요인 가운데 이전 자살 시도 경험이나 자살 생각·자살 계획, 자해 경험 등이 대표적인 위험요인으로 구분된다. 또한 위험요인으로 감정 조절의 어려움, 심각한 기분 변화, 충동, 과도한 부담이나 완벽주의, 수면 장애, 피로 등이 있다. 이뿐만 아니라 건강상 문제나 낮은 자존감, 내향적

성격, 부적절 사고나 죽음에 대한 허용적 태도, 사회적 낙인 등도 위험요인으로 작용한다. 한편, 심리적 위험요인에 대한 대처나 해결 능력, 자살에 대한 비허용적 태도나 긍정적 성격, 삶에 대한 만족과 현실적 사회인식, 높은 자존감·효능감, 자아탄력성 등은 심리적 보호요인으로 구분된다.

사회적 요인 가운데 자살 위험요인으로는 경기침체나 빈곤 등과 같은 사회경제적 위기, 총기 등 자살 수단에 허용적인 지역, 정신건강에 대한 부정적 인식과 낙인 등이 있다. 이 외에도 사회적으로 고립된 지역이거나 계절의 변화, 바이러스나 기생충 등에 위험성이 높은 지역 등도 위험요인에 해당된다. 반면, 사회경제적 문제가 적은 경우나 자살 수단에 비허용적인 지역, 정신건강 증진을 위한 제도적 기반이 마련된 지역은 보호요인으로 구분될 수 있다.

지역사회 요인을 살펴보면, 각종 폭력에 노출된 지역사회, 자살률이 높은 지역, 의료 접근성이 높은 지역 등이 자살 위험도가 높은 곳으로 구분된다. 반면에 지역사회 위기 지원을 위한 연계가 활발한 경우나 정신건강 관리를 위한 각종 사업이 이루어지고 있는 경우, 지역 내 공동체가 활성화되어 있는 경우, 학교 기반의 정신건강 증진 프로그램이 이루어지고 있는 경우에는 자살에 대한 위험 수준이 낮은 것으로 알려져 있다.

끝으로 대인관계 요인 가운데 주변에 자살 수단을 소유하고 있는 경우, 가족 및 지인의 자살, 가정폭력이나 가족 내 갈등, 정신질환에 대한 가족력, 대인관계에 대한 어려움, 사회적 고립 등은 자살 위험요인으로 구분된다. 반면에 가족 및 주변으로부터의 사회적 지지, 가족에 대한 책임감, 타인과의 유대, 대인관계 문제에 대한 대응 능력 등은 보호요인으로 구분된다.

〈표 2-1〉 자살의 위험요인과 보호요인

	위험요인	보호요인
생물학적 요인	남성(자살)/여성(자살 시도) 세로토닌 기능 장애 자살 행동의 가족력	항우울제 사용 기분 안정 치료 항정신증 약물 사용
인구 사회학적 요인	성 소수자 정체성 종교/영성(속명론적 자살) 인종적 특성 고령자/중년 고위험 직업 충기 소유 감금 높은 지각/주관적 스트레스 실직/실업 재정적 부담 최근 정신병원에서 퇴원 괴롭힘/차별 범죄 피해	이성에 성적 취향 종교/영성(결속)
정신증적 요인	우울증, 조울증, 외상 후 스트레스 장애, 불안 및 정신분 열증 경계선 성격과 같은 성격 장애 물질 사용/남용 알코올 사용/남용	치료 동기
심리적 요인	이전 자살 시도 현재 자살 생각 자살 의도의 존재 자살 계획의 존재 치명적인 수단에 대한 접근/존재 자살 준비 행동 이전 또는 현재의 자살하지 않은 자해 기타 자살 이력(예: 생각) 가망 없음 낮은 자제력/높은 충동성 침략 동요 감정 조절 장애 심한 기분 변화 어린 시절 학대 부담감 거부/방해 소속 만성 질환 급성 건강 증상 피로 수면 장애/장애 신경증	대처 기술 문제 해결 능력 자살에 대한 도덕적 반대 생존 신념/살고자 하는 욕구 자살/죽음에 대한 두려움 사회적 반감에 대한 두려움 낙천주의 희망/긍정적인 미래지향성 삶의 만족 온전한 현실적/객관적 평가 높은 자존감/자기효능감 탄력성 외향성 삶에 대한 부가적 사유

40 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

	위험요인	보호요인
	내향성 경험에 대한 제한된 개방성 완벽주의 노숙자 낮은 자존감 창피함 신체적 통증 내성 자살/죽음에 대한 두려움 없음 사고 오류/부정적 사고 정신병/정신적 고통 내면화된 낙인	
사회적 요인	경기 침체/공황 총기법 제한이 낮은 지역 계절 변화 정신건강 및 치료에 대한 낙인 대기 오염 물질 바이러스/기생충 빈곤 산악 지역	건강한 경제상태 총기법 제한이 높은 지역 정신건강 기금
지역사회 요인	지역사회 폭력에 노출 지역 자살 전염 특성 의료 접근성의 장벽	위기 지원 라인/핫라인 건강관리/정신건강 관리 접근성 효과적인 정신건강 관리 훈련된 정신건강 게이트키퍼 공동체 참여 학교 기반 지원 및 증재 프로그램
대인관계 요인	총기 소지 자살/전염에 노출 가정폭력 가족 갈등 정신질환의 가족력 자살/시도의 가족력 관계 불안정 사랑하는 사람의 죽음 연애의 단절 사회적 고립/금단 폭력에 노출	사회적 지원의 존재 사회적 지원 이용 인지된 사회적 지원 자살이 어린이/가족에게 해롭다는 우려 가족에 대한 책임감 건강하고 장기적인 헌신적 관계/결혼 도움을 구하는 행동 집에 있는 아이들 임신 협력 배려 사회적 연결성 돌봄 제공자와의 연락 의료인과의 연계 지원 인지 행동 치료 변증법적 행동 요법

주: 미국의 지역적 특성 관련 요인은 생략

자료: Cramer, R. J., & Kapusta, N. D. (2017). A social-ecological framework of theory, assessment, and prevention of suicide. *Frontiers in psychology*, 8, 1756. p.4-5.

한편, 자살 위험에 놓인 개인에게서 자살 위험 신호를 관찰할 수 있다. 이는 자살을 예방하기 위한 실천적 접근방법에 주로 활용된다. 특히 주변 사람들이 자살의 위험에 노출된 경우를 감지할 수 있도록 아래의 위험 신호들을 자살예방을 위한 지침으로 활용한다.













자살 관련 위험요인과 위험 신호는 자살 행동의 이행 중에 중복되어 나타날 수 있다. 예를 들어 자살 계획의 수립이나 자살 수단의 확보 이전에 이러한 행동이 확인될 수 있으며, 자살 실행 이전에도 위험 신호 행동이 발견될 수 있다. 다만 자살 위험요인과 자살 위험 신호가 자살 이행에 직접적인 영향을 미친다고 보기에는 어렵기 때문에 자살 위험 예측을 위한 연구에서는 해석에 주의가 필요하다.

[그림 2-6] 자살 위험 신호

“자살 경고 신호”

- 각종 부담되는 상황에 대한 이야기
- 고립된 상태
- 불안감의 증가
- 갇힌 느낌, 참을 수 없는 고통에 대한 이야기
- 각종 물질 오남용
- 치명적인 수단에 대한 탐색
- 분노 또는 짜증의 증가
- 극심한 기분 변화
- 절망스러움에 대한 표현
- 과도한 또는 부족한 수면
- 죽고싶다는 이야기 또는 글 게시
- 자살 계획 수립

Know the Suicide  
**WARNING SIGNS**

- Feeling like a burden 
- Being isolated 
- Increased anxiety 
- Feeling trapped or in unbearable pain 
- Increased substance use 
- Looking for a way to access lethal means 
- Increased anger or rage 
- Extreme mood swings 
- Expressing hopelessness 
- Sleeping too little or too much 
- Talking or posting about wanting to die 
- Making plans for suicide 

자료: CDC. (2018). Centers for Disease Control and Prevention 홈페이지 CDC's VitalSigns pdf 파일. p.3. <https://www.cdc.gov/suicide/factors>, 2023.12.27. 인출

#### 4. 인구 특성별 자살 및 정신건강 고위험군 특성

인구 특성은 자살 위험이나 정신건강 상태에 영향을 미치는 것으로 알려져 있다. 먼저, 성별에 따른 특성을 살펴보면 다음과 같다(Kiely, Brady, & Byles, 2019). 여성은 남성에 비하여 우울 및 불안 장애 수준이 상대적으로 높은 것으로 알려져 있다. 또한, 연령의 증가에 따라 우울증을 포함한 각종 인지장애(치매 등)가 발생하며 이러한 정신건강 문제는 성별에 따라 다르게 나타나는 경향이 있다. 여성은 정신건강 관련 질환의 유병률이 높지만, 그로 인한 사망 연령도 높다. 다시 말해 남성에 비하여 상대적으로 정신건강 문제가 지속되는 경향을 보인다는 것이다. 반면에 남성은 여성에 비하여 자살 사망률이 높은 특성을 보인다. 여성은 우울이나 불안 등으로 인하여 자해, 자살 관련 행동의 수준이 남성에 비해 빈번하게 나타나지만 자살 성공으로 이어지는 경우가 남성에 비하여 적다고 보고되고 있다. 그 외에도 여성은 각종 폭력 상황이나 자살과 관련된 직·간접적 경험이 정신건강 문제로 이행될 가능성이 상대적으로 높다. 한편 연령이 높은 남성은 정신건강 문제에 대한 지원이나 치료 등에 대하여 상대적으로 소극적인 태도를 보이는 것으로 알려져 있다.

아동 및 청소년의 자살 행동과 정신건강 특성을 살펴보면 다음과 같다(Colizzi, Lasalvia, & Ruggeri, 2020). 아동 및 청소년기 정신건강 문제에 대한 대응은 성인기 정신건강 문제에 대한 예방과 만성질환으로의 이행 가능성 감소, 정신건강 치료 등과 관련된 비용의 감축 등에 영향을 미친다. 즉, 아동·청소년기에 정신건강에 대한 조기 개입은 정신건강 문제가 심화될 수 있는 위험성을 줄이고 생애 초기 심리적·정서적 발달에 긍정적인 영향을 미칠 수 있다. 정신건강 서비스는 적절한 개입 시점에 맞게 제공되어야 하는데, 일반적으로 생애 초기에 이루어지는 개입이 효



과적인 것으로 보고되고 있다. 정리하면, 전 생애 정신건강 증진을 위하여 아동 및 청소년기에 나타날 수 있는 정신건강 문제를 조기에 발견하고 개입하는 것이 1차 예방적 접근의 측면에서 중요하다.

청소년의 자살 행동은 다른 생애주기의 자살 행동과 다음의 차이를 가진다(박성준, 2019). 첫째, 청소년의 자살은 자신의 삶 속에서 경험하는 어려움을 표현하는 극단적 방식 중 하나이기 때문에 자살 시도 또는 자해가 반복적으로 나타날 가능성이 높다. 둘째, 아동 및 청소년은 자살 시도에 대한 사전계획이 부재하거나 불명확하여 계획과 실행 간의 기간이 짧은 특성을 보이며, 이에 따라 자살 시도에 따른 자살 성공률이 성인에 비하여 낮다. 셋째, 청소년은 성인에 비하여 정신질환 문제보다는 일시적인 스트레스나 우울·불안감에 의하여 자살이 시도될 가능성이 높다. 특히 학교 등 사회에서의 부적응이나 타인과의 관계 문제, 학업성취에 대한 부담 등이 자살 행동에 영향을 미친다. 넷째, 청소년은 자신의 자살 경험이나 매체를 통한 자살 정보에 대하여 성인에 비해 민감하게 반응하는 특성을 보인다. 반면에 성인은 경제적 어려움이나 사회적 고립 등 본인에게 기인한 문제에 더 많은 영향을 받는다. 다섯째, 자살 관련 정보 수집, 자살 수단의 습득 및 이용, 자살 계획의 실행 등의 자살 행동은 충분한 인지적 발달이 선행되어야 하기 때문에 연령이 증가됨에 따라 이러한 위험성이 증대되는 특성을 보인다.

아동·청소년 자살의 보호요인과 위험요인은 일반적인 자살 관련 요인과 유사하지만 가족이나 지인의 특성, 학교나 매체를 통한 환경적 특성에 상대적으로 민감하게 반응하는 경향을 보인다. 대표적으로 가족과의 관계나 가족 구성원의 정신건강 문제가 개인의 자살 행동에 영향을 미치며, 학교 내에서의 역할이나 구성원 간의 관계는 자살의 위험·또는 보호요인으로 적용하고 있다.

〈표 2-2〉 아동·청소년 자살 행동의 위험요인과 보호요인

	위험요인	보호요인
개인적 특성	성별(남성) 낮은 자아존중감 성적 성향 우울, 불안 주의력 결핍 낮은 문제해결 능력 물질 오남용 무망감 완벽주의 충동적·공격적 성향 이전 자살 시도 경험	높은 자아존중감 높은 회복탄력성 정서적 심리적 안정감 행복감
가족 특성	가족의 자살 이력 부모의 정신건강 문제 부모의 분리(이혼 등) 부모의 죽음 가족 내 관계 문제	가족 구성원 간 관계 가족 간 결속력
환경적 요인	낮은 사회경제적 지위 낮은 교육 수준(본인, 부모) 자살 수단에 대한 접근성 학교 내 문제 타인의 자살 행동에 대한 노출	학교에 대한 소속감 동료와의 애착관계 교사의 지지 종교 자살 수단에 대한 접근성 제한
부정적 생활사건	아동학대 등 아동기의 경험한 어려움 대인관계에 대한 어려움 스트레스 사건	-

자료: 정승. (2019). 서울 학생 자살 고위험군 지원 방안 연구. 서울특별시교육청교육연구정보원. p.15, 17.

노인의 자살 사망률은 다른 연령층에 비하여 높으며, 이에 따라 자살 위험요인을 설정할 시에 ‘고령 여부’가 필수적으로 고려된다. 노인은 다른 연령 집단에 비하여 자살 관련 위험요인에 민감하게 반응하는 것으로 알려져 있다(송영달, 손지아, & 박순미, 2010). 노인은 다른 연령 집단에 비하여 독거 비율이 높고 교육 수준이 낮으며 불안정한 경제 상태인 경우가 많다. 노화에 따른 신체적·정신적 건강의 취약성이 증대되는 가운데 우울·불안 등의 정서적 어려움을 가지는 경우가 많다. 또한, 심리적 측면

에서는 자아존중감과 자기효능감 수준이 낮은 특성을 보인다. 반면에 비동거 가족의 관계나 지역사회 인근 주민과의 관계 등이 노인의 정신건강에 영향을 미친다. 특히 주변 사람들과의 교류나 사회활동이 활발한 경우에는 자살의 위험 수준이 낮아질 수 있다.

사회환경적 측면에서 보았을 때, 노인의 자살 사망률은 노인에 대한 사회적 인식과도 관련성을 가진다(Yuryev. et al, 2010). 노인에 대한 존중이나 우호감 등의 인식 수준이 높은 사회 또는 지역에서는 노인의 자살 사망률이 상대적으로 낮게 나타난다. 또한 노인과 그 가족이 함께 사는 문화를 가진 경우, 노인의 사회경제적 활동에 우호적인 태도를 보이는 사회의 경우에 노인의 자살 사망률이 낮게 나타난다.

## 5. 자살 고위험군의 정책적 범위

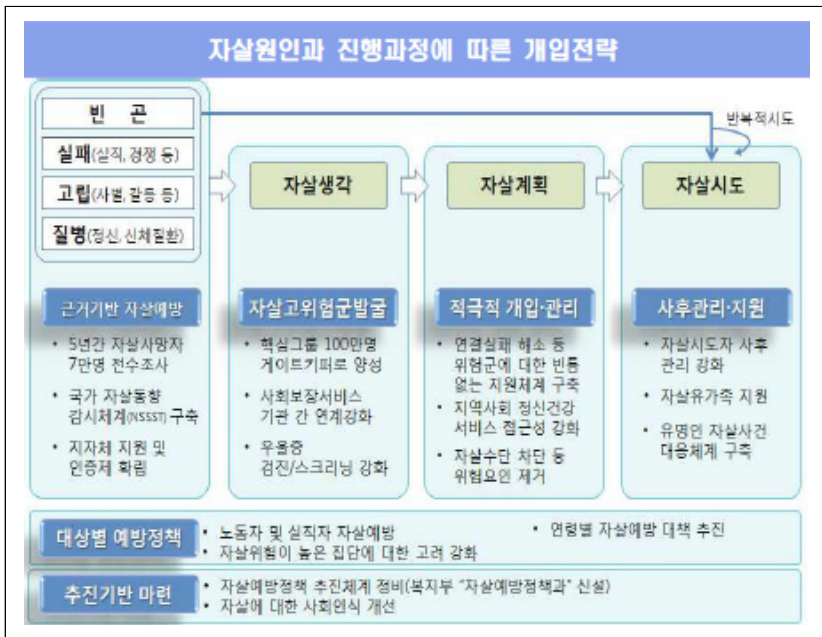
「자살예방 및 생명존중문화 조성에 관한 법률(이하 ‘자살예방법’)」에서는 별도로 자살 고위험군을 정의하고 있지 않으나 자살예방에 대한 주요 정책 대상을 설정하고 있다. 먼저, 동법 제12조의 2(자살 시도자 등의 사후관리), 제12조의 3(형사사법정보의 제공 등)에 따라 자살 시도 여부, 자살예방센터로부터 자살 위험 관련 정보 등이 확인된 경우에는 해당 법률에 근거한 각종 정책의 주요 대상으로 선정된다.

한편 2018년에 수립된 제4차 자살예방기본계획(자살예방 국가행동계획)에서는 자살 고위험군 발굴을 위한 자살 위험 개입 전략을 수립하였다. 해당 계획에서는 자살과 관련된 위험요인이나 자살 행동 여부(자살 생각, 계획, 시도) 등을 바탕으로 자살 고위험군을 발굴·관리할 것을 강조하였다. 자살 고위험군의 발굴을 위하여 인력양성, 관계기관 연계, 스크리닝 강화 외에도 자살 사망자에 대한 전수조사나 자살 동향 감시체계 구

축 등이 주요 전략으로 제시되었다.

이후 제5차 자살예방기본계획에서는 자살 고위험군의 대응 전략 강화를 위하여 자살 고위험군의 범위가 확대되었다(전진아 외, 2022). 지역사회에서의 보건·복지 사례관리 대상 가운데 위기가구(저소득, 사회적 고립, 기타 사회적 위험성에 노출된 경우)의 경우, 자살 위험 수준이 상대적으로 높다고 판단하고 있으며, 특히 자살 고위험군에 대한 조기 개입의 근거로 지역사회 및 해당 가구의 취약성 등이 사전에 검토되고 있다.

[그림 2-7] 자살 원인과 진행 과정에 따른 개입전략



자료: 관계부처 합동. (2018). 자살예방 국가 행동계획. p.19

대상별로 보건·복지 기관에서 자살 고위험군을 조기 발굴 것이 강조되고 있는데, 아동·청소년에 대해서는 지역아동센터, 드림스타트, 청소년쉼

터, 상담복지센터, 학교밖 청소년 지원센터, 위클래스 등의 기관에서 조기 발굴 및 초기 자원 연계를 강조하였다. 여성에 대해서는 새일센터, 건강가정지원센터, 모자보호 기관 등을 통하여 사회적 고립을 예방하고 정신건강 증진을 위한 지원을 추진하고 있으며, 청년에 대해서는 청년마음 건강지원사업을 통하여 자살예방센터 및 정신건강 의료기관과의 연계를 강화하였다. 노인의 경우, 독거노인 및 치매환자 가족에 대한 지원 기관을 통하여 자살 위험 상황에 대응할 수 있도록 지원하고 있으며, 그 외에 각 대상별 정신건강 관련 사업을 추진하는 주체와 보건·복지 서비스를 제공하는 주체를 중심으로 자살 고위험군에 대한 조기 개입 지원을 하고 있다. 또한 자살 위험요인 가운데 지역적 특성을 고려해, 자살 고위험 장소를 설정하여 자살 수단에 대한 접근성을 관리하는 방안이 검토되고 있다.

이상의 자살 고위험군 발굴 및 조기 개입을 위한 전략 추진 시, 위기 대응을 위한 각종 데이터 활용이 강조되고 있다. 예를 들어 사회보장정보시스템의 활용을 통한 고위험군 발굴, 지역단위 사회보장 전달체계를 활용한 고위험군 발굴 등이 있으며, 관련 법률에 근거하여 자살 시도, 사회적 재난 등 관련 정보는 본인의 동의 없이 정보 연계가 가능하게 되었다. 향후 자살 고위험 관련 정보를 활용한 데이터베이스 및 통합분석시스템 구축을 추진할 예정이다.

### 제3절 소결

이 장에서는 정신건강 및 자살과 관련된 주요 이론과 정책들을 바탕으로 정신건강 및 자살 고위험군을 정의하고, 관련 위험 발생의 원인과 과정 등에 관한 내용을 검토하였다. 이를 통해 도출될 수 있는 내용을 정리하면 아래와 같다.

첫째, 정신건강 관련 위기 대응을 위한 고위험군의 범위 설정이 필요하다. 특히 정신건강 및 자살 관련 위기에 대응하기 위해서는 위기 상황에 대한 진단 또는 위기로 이행되는 과정에 대한 검토가 요구된다. 즉, 정신건강과 관련된 주요 위험요인들에 대한 검토와 정신건강 문제가 발생하는 진행 과정의 점검 등이 정신건강 관련 위기대응을 위한 전략 수립에 필요하다. 또한 정신건강과 관련된 대응의 우선순위를 설정하기 위해서는 고위험군에 대한 기준이 설정될 수 있다. 이를 통해 자살을 포함하여 정신건강 증상의 악화나 만성화로 인한 각종 문제에 대응하기 위한 자원의 배분과 개입의 강도 등을 결정할 수 있다.

둘째, 자살 고위험군에 대한 대응 등과 같이 정신건강 관련 위기 대응 방안에 대한 세부적인 사항들이 검토되어야 한다. 앞에서 논의된 바와 같이 정신건강 문제의 범위는 상당히 넓기 때문에 고위험군에 대한 적절한 범위 및 개념 설정이 필요하다. 다양한 정신건강 문제 가운데 자살 행동은 상대적으로 위험 수준이 높다고 볼 수 있기 때문에 정책 대상 설정과 위기대응의 우선순위가 높다고 볼 수 있다. 한편, 자살 관련 요인들은 정신건강 문제와 관련된 각종 위험·보호요인과 중첩되는 경향을 보인다. 또한 정신건강 문제가 심화되었을 때 자살 행동을 보일 가능성이 증대된다. 이뿐만 아니라 자살 행동은 정신건강 문제의 결과임과 동시에 자살 완료(completed suicide)를 비롯한 각종 위험 상황의 원인이 되기도 한다. 따라서 정신건강 위기대응을 위한 구체적이고 효과적인 접근방법이 제시되기 위해서는 자살 고위험군에 대한 대응 전략 등과 같이 세부적인 문제 상황에 대한 논의가 필요하며, 이에 따라 본 연구에서는 자살 위험에 대한 예측 모형을 중심으로 정신건강 위기대응 방안을 모색하고자 하였다.

셋째, 정신건강 관련 정책에서 자살 고위험군을 주요 정책 대상으로 설정하고 있다. 정신건강과 관련된 정책 및 각종 사업에서 자살 고위험군에

대한 대응 또는 지원을 적극 추진하고 있다. 이는 자살 위험이 다른 정신 건강 문제보다 그 위험 수준이 높으며, 여러 정신건강 문제에 대한 결과란 인식이 반영된 결정으로 볼 수 있다. 또한, 자살은 사후적 대처가 불가능한 정신건강 영역이기 때문에 위기에 대한 예측과 신속한 사전 대응이 상당히 중요하다. 이것이 자살예방을 위한 별도의 기본계획이 수립되는 배경이 되기도 한다. 정리하면, 정신건강 관련 정책 중 자살 고위험군에 대한 선별과 자살 위험에 대한 예측은 예방적 정신건강 정책의 핵심적 내용이라 볼 수 있다. 이를 바탕으로 본 연구에서는 정신건강 위기에 대한 사전적 대응 방안을 모색하기 위하여 자살 행동 관련 요인에 대한 검토와 자살 위험에 대한 예측 모형 탐색을 중심으로 논의하고자 한다.







## 제3장

### 예측 모델 적용을 위한 공공데이터 검토

제1절 검토 대상 공공데이터 개요

제2절 자살 고위험군 관련 선행연구 실증 분석 변수

제3절 소결



## 제 3 장

# 예측 모델 적용을 위한 공공데이터 검토

### 제1절 검토 대상 공공데이터 개요

#### 1. 국민건강영양조사

국민건강영양조사는 국민의 건강 수준, 건강 관련 의식 및 행태, 식품과 영양 섭취 실태를 조사하고자 실시된다. 국민건강증진법에 근거하여 질병관리청에서 만 1세 이상 가구원 약 1만 명을 매년 조사하고 있으며, 해당 조사를 통해 국가통계를 생산하고 국제기구에 제출하기도 한다. 또한, 우리나라 국민의 건강 수준을 평가하고 건강정책의 수립, 수행 및 평가 등에도 활용하고 있다.

국민건강영양조사는 제1기(1998년) 조사를 시작으로 2023년 2월 현재 제9기(2022~2024년) 조사를 수행 중이다. 과거 제3기(2005년) 조사까지는 3년 주기의 2~3개월의 단기 조사체계로 실시하였으나, 제4기(2007~2009년) 조사부터는 연중 조사체계로 개편되었다. 1개 기수 내의 3개 연도가 각기 독립적인 3개의 순환표본으로, 전국을 대표하는 확률 표본이 될 수 있도록 하는 순환표본조사 방식을 도입하였다.

〈표 3-1〉 국민건강영양조사 개요

구분	설명
조사 주기	1년
조사 대상	매년 만 1세 이상 가구원 약 1만 명
조사 방법	- 전문 조사수행팀이 이동검진 차량과 함께 방문하여 검진, 건강설문 및 영양조사 - 순환표본조사(Rolling Sampling Survey)
조사 기간	- 1~12월(12개월), 연중 조사(연 48주) - 매주 4개 지역(연 192개 지역) 조사, 1개 지역(조사구)을 3일간 조사
목표 표본 크기	- 576개 조사구, 14,400가구 - 연간 192개 조사구, 4,800가구(1개 조사구 당 25가구)
표본 추출 방법	2단계 층화집락 표본 추출
표본 추출 틀	- 표본설계 시점에서 가용한 가장 최근 시점의 인구주택총조사 자료 - 인구주택총조사(2016)
표본 단위	(병합) 조사구 → 가구
층	시도, 동읍면, 주택유형(내재적층: 성별, 연령, 가구주 나이, 주거면적 등)
주요 내용	- 조사 대상자의 생애주기별 특성에 따라 소아(1~11세), 청소년(12~18세), 성인(19세 이상)으로 구분하고 각 특성에 맞는 조사 항목을 적용 - 조사 분야 1) 검진 조사: 비만, 고혈압, 당뇨병, 이상지질혈증, 간질환, 신장질환, 빈혈, 구강질환, 악력, 체성분, 시력검사, 이비인후질환 2) 건강설문조사: 가구조사, 이환, 예방접종 및 건강검진, 활동제한 및 삶의 질, 손상, 의료이용, 신체활동, 여성건강, 교육 및 경제활동, 비만 및 체중조절, 음주, 안전의식, 수면건강 및 정신건강, 흡연, 구강건강 3) 영양조사: 식품 및 영양소 섭취 현황, 식생활 행태, 식이보충제, 영양지식, 식품 안정성, 수유 현황, 이유보충식

주: 2023년 2월 시점에서 원시자료 구득이 가능한 제8기(2019~2021년) 국민건강영양조사 기준 자료: 질병관리청. 국민건강영양조사 홍보홈페이지. 국민건강영양조사 개요. [https://knhanes.kdca.go.kr/knhanes/sub01/sub01\\_02.do](https://knhanes.kdca.go.kr/knhanes/sub01/sub01_02.do)에서 2023.2.19. 인출.  
질병관리청. (2023). 국민건강영양조사 원시자료 이용지침서 제8기(2019-2021). pp.3-4.

국민건강영양조사의 건강설문조사에는 정신건강에 관한 문항들이 포함되어 있는데, 자살 외에 스트레스, 우울 및 정신 문제 상담 등에 관한 문항도 있다. 또 정신건강 영역 외에 삶의 질 영역에서 불안이나 우울 그리고 행복에 관한 문항들을 포함하고 있다. 특히, 자살에 관한 문항을 자살 생각·계획·시도로 자세히 구분하여 제시한 것이 특징적이다.

(표 3-2) 국민건강영양조사의 건강설문조사 항목 상세

조사영역	대상 연령	조사 항목
가구조사	만 19세 이상	가구원 수, 세대유형, 기초생활 수급자 여부, 주택소유 여부, 거주 주택 유형, 가구 총소득, 성, 연령, 결혼상태, 가구주와의 관계, 건강보험 가입, 민간보험 가입
교육	만 1세 이상	학력, 졸업 여부
	만 19-64세	부모 학력
경제활동	만 15세 이상	경제활동 여부, 취업형태, 종사상지위, 평균 근로시간, 근로시간형태, 현재 직업
	만 19세 이상	정규직 여부, 최장 직업
이환	만 1세 이상	최근 2주간 이환, 만성질환별(성인 30개, 소아청소년 7개) 이환
의료이용	만 1세 이상	(병의원)미충족의료·이유, 외래이용, 입원이용
건강검진	만 19세 이상	건강검진/암검진 수진 여부, 건강검진/암검진 방법 <sup>1)</sup>
예방접종	만 1세 이상	인플루엔자 예방접종 여부
활동제한	만 12세 이상	활동제한 여부, 활동제한 이유, (월간)와병경험/일수, (월간)결근결석 경험/일수
삶의 질	만 1세 이상	주관적 건강 인지
	만 19세 이상	건강 관련 삶의 질 지수 (1) EQ-5D <sup>12)</sup> , EuroQol-5 Dimension: 운동능력, 자기관리, 일상활동, 통증·불편, 불안·우울, (2) HINT-8 <sup>13)</sup> , Health-related quality of life Instrument with 8 items: 계단오르기, 통증, 기운, 일하기, 우울, 기억, 잡자기, 행복
손상	만 1세 이상	손상 경험·치료처
흡연	만 12-18세	일반담배(궐련) 평생 흡연/현재 흡연/처음 흡연 시작 연령/흡연량
	만 19세 이상	일반 담배(궐련) 평생 흡연/현재 흡연/과거 흡연/처음 흡연 시작 연령/매일 흡연 시작 연령 <sup>3)</sup> /흡연량, 궐련형 전자담배 평생 흡연/현재 흡연/흡연량/사용 이유 <sup>2)</sup> , 액상형 전자담배·담배 종류별 평생/월간 사용, 금연 기간, 금연 시도, 금연 계획, 니코틴 의존 <sup>1)</sup> , 금연 방법 <sup>1)</sup> , 가정·직장·공공장소 실내 간접흡연
음주	만 12세 이상	평생 음주, 음주 시작 연령, 음주빈도, 음주량, 폭음빈도, 간접폐해
	만 19세 이상	절주 권고경험 여부, 음주문제 상담경험 여부, 과거/현재 알코올 의존
신체활동	만 12-18세	하루 60분 신체활동 실천, 앉아서 보내는 시간, 근력운동
	만 19세 이상	국제신체활동설문(GPAQ, Global Physical Activity Questionnaire: 일/여가 고강도·중강도

56 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

조사영역	대상 연령	조사 항목
		신체활동, 이동 시 활동, 걷기, 앉아서 보내는 시간, 근력운동
수면건강	만 12세 이상	주중/주말 하루 평균 수면시간, 주중/주말 잠든 시각, 일어난 시각 <sup>3)</sup>
	만 40세 이상	폐쇄성수면무호흡증 선별도구(STOP-Bang: 코골이, 피곤함, 수면무호흡 목격자)
정신건강	만 12세 이상	스트레스 인지, 우울감 경험, 자살 생각·계획·시도, 정신문제 상담 경험
	만 19세 이상	우울증 선별도구 <sup>2)</sup> (PHQ-9, Patient Health Questionnaire-9), 범불안장애 선별도구 <sup>3)</sup> (GAD-7, Generalized Anxiety Disorder-7)
안전의식	만 1-5세	자동차 보호장구 착용
	만 1-11세	자동차 앞좌석 이용, 자전거 헬멧 착용
	만 12세 이상	동승 차량 자동차 안전벨트 착용
	만 19세 이상	운전 시 안전벨트 착용, 앞/뒷좌석 안전벨트 착용, 자동차·오토바이·자전거 음주운전 경험, 음주운전 차량 동승
비만 및 체중조절	만 6세 이상	주관적 체형 인지, 체중조절 시도, 체중조절 방법 <sup>1)</sup>
	만 19세 이상	체중 변화 여부, 체중 감소 정도, 체중 증가 정도
여성건강	만 10세 이상	현재 월경 여부, 초경 연령
	만 15세 이상	임신 경험, 출산 경험, 출산연령
	만 19세 이상	모유수유 경험·자녀 수·기간, 폐경 연령, 경구피임약 복용 경험
구강건강	만 1세 이상	칫솔질 여부, 치아 손상, 구강검진, 치과 이용, (치과)미충족의료
	만 12세 이상	구강용품 사용
	만 19세 이상	저작 불편, 발음 불편

주1: 2023년 2월 시점에서 원시자료 구득이 가능한 제8기(2019~2021년) 국민건강영양조사 기준.  
 2: 1) 순환조사 문항으로 제8기 1차년도(2019) 조사, 2) 순환조사 문항으로 제8기 2차년도(2020) 조사, 3) 순환조사 문항으로 제8기 3차년도(2021) 조사  
 자료: 질병관리청. (2023). 국민건강영양조사 원시자료 이용지침서 제8기(2019-2021). p.6.

자살에 대한 생각·계획·시도 문항은 순환조사 문항이 아닌 매년 조사하는 문항으로, 자살 생각과 자살 시도에 대한 조사 항목은 제1기(1998년) 조사부터 계속해서 조사하고 있다. 이에 반해 자살 계획에 대한 문항

은 제6기(2013~2015년) 조사부터 계속해서 포함하고 있으며 그 이전에는 2005년도 조사를 제외하고는 모두 조사하지 않았다. 그리고 자살 관련 문항을 매년 모든 응답자를 대상으로 조사하였던 것은 아니다. 2014년, 2016년, 2018년 및 2020년은 청소년(만 12~18세)만을 대상으로 조사하였다.

국민건강영양조사의 자살 관련 문항은 모두 지난 1년간의 생각·계획·시도에 대해 묻고 있다. 제6기(2013~2015년) 조사부터는 자살 생각 문항의 질문에 ‘진지하게’라는 문구가 추가되었으며, 이후로는 응답 빈도가 줄어든 것을 알 수 있다. 자살 관련 문항은 조건부 로직이 설정되어 있어 맨 첫 질문인 자살 생각에 대한 응답 빈도에 나머지 문항들의 응답 빈도도 의존함을 알 수 있다.

〈표 3-3〉 국민건강영양조사의 자살 관련 문항 연도별 현황

조사 항목	제1기 (1998)	제2기 (2001)	제3기 (2005)	제4기 (2007~ 2009)	제5기 (2010~ 2012)	제6기 (2013~ 2015)	제7기 (2016~ 2018)	제8기 (2019~ 2021)
1년간 자살 생각	●	●	●	●●●●	●●●●	●○●○	●○●○	●○●○
자살 생각 원인			●					
1년간 자살 계획			●			●○●●●●●●		
1년간 자살 시도	●	●	●	●●●●	●●●●	●○●●●●●●		
1년간 자살 시도 횟수	●	●	●					
1년간 자살 시도로 인한 치료	●	●	●					

주: ●은 조사 대상이 성인(만 19세 이상)과 청소년(만 12~18세), ○은 조사 대상이 청소년(만 12~18세)임.  
 자료: 질병관리청. (2023). 국민건강영양조사 원시자료 이용지침서 제8기(2019-2021). p.63.

〈표 3-4〉 국민건강영양조사의 자살 관련 문항

조사 항목	설문 문항
1년간 자살 생각	(01~'05) <b>지난 1년 동안</b> 죽고 싶다는 생각을 해 본 적이 있습니까? 예, 아니오 (07~'12) <b>최근 1년 동안</b> 죽고 싶다는 생각을 해 본 적이 있습니까? 예, 아니오 (13~) <b>최근 1년 동안 진지(심각)하게</b> 자살을 생각한 적이 있습니까? 예, 아니오
자살 생각 원인	(05) 죽고 싶다는 생각을 했던 가장 큰 이유는 무엇이었습니까? (1) 경제문제 (2) 직장생활(직장 내 대인관계 포함) (3) 부모자녀 문제(형제, 친척, 고부간 문제) (4) 배우자 문제 (5) 생활환경 문제(주거 등) (6) 취업 문제(진로) (7) 본인의 건강 문제 (8) 가족의 건강 문제 (9) 기타(무엇 )
1년간 자살 계획	(05) <b>지난 1년 동안</b> 자살 계획을 구체적으로 세워 본 적이 있습니까? 예, 아니오 (13~) <b>최근 1년 동안</b> 자살하기 위해 구체적인 계획을 세운 적이 있습니까? 예, 아니오
1년간 자살 시도	(07~) <b>최근 1년 동안</b> 실제로 자살 시도를 해 본 적이 있습니까? 예, 아니오 (01~'05) <b>지난 1년 동안</b> 실제로 자살 시도를 해 본 적이 있습니까? 예, 아니오
1년간 자살 시도 횟수	(01) 자살 시도를 해 본 적이 있다면 몇 번이나 시도를 하셨습니까? (1) 1번 (2) 2~3번 (3) 4~5번 (4) 6번 이상 (05) <b>지난 1년 동안</b> 자살을 몇 번이나 시도하셨습니까? (1) 1번 (2) 2~3번 (3) 4~5번 (4) 6번 이상
1년간 자살 시도로 인한 치료	(01) 자살 시도 하신 당시 치료를 위해서 병원을 방문하셨습니까? 예, 아니오 (05) <b>지난 1년 동안</b> 자살 시도를 하신 당시 이로 인한 손상 때문에 의료 기관에서 치료를 받은 적이 있습니까? 예, 아니오

자료: 남정자, 김혜련, 이상호, 최은영, 윤강재, 박인화. (1999.12). 1998년도 국민건강·영양조사: 총괄보고서. 보건복지부·한국보건사회연구원. 조사표 p. 21.  
최정수, 남정자, 김정희, 고민정, 조재국, 이연희, ... 서희재. (2002). 2001년도 국민건강·영양조사: 총괄편. 보건복지부·한국보건사회연구원. p. 350.  
강은정, 최은진, 송현중, 유근춘, 김나연, 남정자, ... 이연경. (2006). 국민건강 영양조사 제3기 (2005): 총괄. 보건복지부·한국보건사회연구원. p. 453.  
질병관리청. 국민건강영양조사. 각 연도별 조사표.



(표 3-5) 국민건강영양조사의 자살 관련 문항 응답 빈도(표본 수)

	자살 생각						자살 계획						자살 시도					
	예	아니오	비해당	모름	결측	계	예	아니오	비해당	모름	결측	계	예	아니오	비해당	모름	결측	계
	1998년	2756	8052		28252	39060								112	2644	8052		28252
2001년	1756	7405	9	28708	37878								78	1669	7405	18	28708	37878
2005년	1652	7183		25310	34145	157	1495	7183		25310	34145		53	1590	7183		25319	34145
2007년	590	2756		10	1238	4594							37	530	2766	23	1238	4594
2008년	1444	6186		10	2104	9744							86	1347	6196	11	2104	9744
2009년	1488	6969	1558	36	482	10533							91	1392	8563	5	482	10533
2010년	1038	5943	1384	23	570	8958							58	977	7327	26	570	8958
2011년	961	5720	1170	44	623	8518							50	911	6890	44	623	8518
2012년	880	5319	1001	137	721	8058							42	837	6320	138	721	8058
2013년	299	5678	1107	193	741	8018	86	5891	1107	193	741	8018	47	5929	1107	194	741	8018
2014년	25	464	6400	2	659	7550	6	483	6400	2	659	7550	5	484	6400	2	659	7550
2015년	324	5611	804	165	476	7380	97	5839	804	164	476	7380	49	5887	804	164	476	7380
2016년	24	528	7243	2	353	8150	117	6450	1115	115	353	8150	43	6524	1115	115	353	8150
2017년	323	6310	990	76	428	8127	84	6548	990	77	428	8127	43	6590	990	76	428	8127
2018년	19	457	7164		352	7992	96	6562	929	53	352	7992	43	6617	929	51	352	7992
2019년	329	6349	925	105	402	8110	87	6592	925	104	402	8110	28	6652	925	103	402	8110
2020년	11	420	6663		265	7359	116	6172	743	63	265	7359	31	6257	743	63	265	7359
2021년	260	5771	646	51	362	7090	81	5951	646	50	362	7090	30	6003	646	49	362	7090

자료: 질병관리청. 국민건강영양조사. 각 연도별 원시자료에서 추출.

## 2. 정신건강실태조사

정신건강실태조사는 정신건강증진 정책 수립을 지원하기 위해 정신건강복지법 제10조에 근거하여 5년 주기로 실시하는 실태조사이다. 정신질환의 인구학적 분포, 유병률 및 유병요인, 성별·연령 등 인구학적 특성에 따른 정신질환의 치료 이력, 정신건강증진시설 이용 현황, 정신질환으로 인한 사회적·경제적 손실, 정신질환자의 취업·직업훈련·소득·주거·경제상태, 정신질환자에 대한 복지서비스 및 정신질환자 가족의 사회·경제적 상황 등을 조사한다.

조사 대상은 만 18세 이상 만 79세 이하 대한민국 국민 5,500명(가구당 1명)이며 조사원에 의한 가구방문 면접조사 형태로 실시된다. 정신건강실태조사는 2001년 ‘정신질환자 실태조사(최초 작성 승인 제 117050호)’를 시작으로 5년 주기로 작성되었으며, 2021년 5차 조사부터는 조사명칭이 ‘정신건강실태조사’로 변경되었다.

〈표 3-6〉 정신건강실태조사(National Mental Health Survey) 개요

구분	설명
조사 주기	5년
조사 대상	만 18세 이상 만 79세 이하 대한민국 국민 5,500명(가구당 1명) (최종 응답 만 18세 이상 만 79세 이하 5,511명, 가구당 1명)
조사 방법	- 조사원에 의한 가구방문 면접조사(태블릿 PC 사용)
조사 기간	- 조사 실시 기간: 2021년 6월 ~ 2021년 8월 - 조사 기준 시점: 2021년 1월 1일 0시 기준 - 조사 대상 기간: 2020년 1월 1일 ~ 2020년 12월 31일(1년) * 유병률 관련 문항의 “지난 1년” 기준은 조사 시점으로부터 이전 1년
목표 모집단	조사기준 시점에 대한민국에 거주하는 만 18세 이상 만 79세 이하 일반 국민
표본 추출틀	통계청에서 제공하는 2019년 기준 인구총조사의 조사구 및 가구 명부
조사 모집단	인구총조사의 조사구 중에서 섬, 기숙 시설, 특수사회시설, 관광호텔 및 외국인 조사구를 제외한 아파트 조사구 및 일반 조사구에 속한 가구에 거주하는 만 18세 이상 만 79세 이하의 국민

구분	설명	
층화	- 1차 층화: 시·도 - 2차 층화: 조사구 특성을 고려하여 동부의 경우 아파트와 일반, 읍면부로 구분	
표본 추출	조사구의 가구 수를 기준으로 확률비례계통추출법으로 조사구 선정	
조사 항목	[정신장애 진단도구(K-CIDI)] A. 응답자 인적 사항 B. 니코틴 사용에 의한 장애 D. 공포 장애 및 기타 불안 장애 E. 우울장애 및 기분부전 장애 J. 알코올 사용과 관련된 장애 K. 강박장애 및 외상 후 스트레스 장애 S. 자살 SR. 정신건강서비스 이용	[부가도구] 1. 게임중독 2. 건강위험요인 및 동반질환 3. 삶의 질 및 삶의 만족도 4. 성인 ADHD 5. 애도 및 회복탄력성 6. 사회적 고립 7. 자해행동 8. 외로움 9. 정신질환에 대한 인식

주: 2021년 정신건강실태조사 기준

자료: 보건복지부 국립정신건강센터 정신건강연구소, 정신건강실태조사 누리집. <https://mhs.ncmh.go.kr/front/svyAdult.do>에서 2023.6.26. 인출.

보건복지부 국립정신건강센터. (2021). 2021년 정신건강실태조사 보고서. pp.32-47.

정신건강실태조사에서는 정신장애진단도구인 K-CIDI(2.1버전)를 사용하여 조사 항목을 구성하고, 그 외에도 기타 부가도구를 사용하여 위험 및 보호요인을 파악한다. 또 자살 관련 행동의 평생, 1년, 1개월 유병률을 산출하기 위해 자살 관련 행동에 대한 문항들을 구성하는데, 2021년 정신건강실태조사에서는 과거 조사와 동일하게 자살 관련 행동(자살 사고, 자살 계획, 자살 시도)을 조사하고 각각의 평생/1년/1개월 유병률을 산출하였다.

〈표 3-7〉 정신건강실태조사의 조사 항목 상세

도구명	구분	항목
정신 장애 진단 도구 (K-CIDI)	A. 응답자 인적 사항	성별, 연령, 신장 및 체중, 지난 1년간 월평균 가구 소득 등
	B. 니코틴 사용에 의한 장애	지난 1달간 흡연 여부 및 개수, 시기·흡연욕구 증가 및 중단배 여부, 과다흡연 경험 여부 등
	D. 공포 장애 및 기타 불안 장애	공포나 불안, 두려움, 발작에 대하여 나타나는 문제 증세 경험 또는 기피 여부 등
	E. 우울장애 및 기분부전 장애	2주 이상 무기력한 상태에서 식욕/몸무게 변화, 수면 문제, 지체/안절부절못함, 무가치감/죄책감, 자신감 부족, 사고 장애, 자살에 대한 생각 경험 여부 등
	J. 알코올 사용과 관련된 장애	평생 및 지난 1년간 음주량, 음주 빈도, 폭음횟수, 음주로 인한 문제 등
	K. 강박장애 및 외상 후 스트레스 장애	2주 이상 불쾌한 생각으로 인한 괴로움, 문제 증상, 강박 증세 경험 및 일상생활 애로 사항, 심한 스트레스나 충격적인 사건으로 인한 신체적/정신적 문제 등
	S. 자살	<b>자살 사고, 계획, 시도 등</b>
	SR. 정신건강서비스 이용	정신건강 문제로 인한 의료 상담 여부 및 시기, 정신과 약물치료에 대한 경험/생각 등
부가 도구	1. 게임중독	게임중독은 인터넷 게임장애 구조적 면담검사(Structured Clinical Interview for Internet Gaming Disorder, SCI-IGD) 척도와 인터넷 게임 습관과 관련된 몰두, 내성, 갈망/급단, 통제력 상실, 다른 영역에 대한 태만, 병식 등의 항목으로 구성된 게임 생활 습관 설문 문항을 사용하여 조사
	2. 건강위험요인 및 동반질환	- 국제신체활동설문(International Physical Activity Questionnaire, IPAQ)을 사용하여, 지난 7일간 신체활동에 소모한 시간에 대해 조사 - 동반질환에 대한 문항으로 기존 당뇨 및 고혈당, 고혈압, 고지혈증, 뇌졸중 또는 중

도구명	구분	항목
		풍, 심장병, 암의 6개 분류에서 심장병을 심근경색증과 협심증으로 나누어 총 7개의 문항으로 제시하여, 진단 시기와 현재 이환 여부를 조사
	3. 삶의 질 및 삶의 만족도	- EuroQol EQ-5D 척도를 사용하여 5가지 측면에서 삶의 질에 대해 조사 - WHOQOL-BREF 중 전반적 삶의 질과 건강 상태 질문을 사용하여 조사
	4. 성인 ADHD	성인 ADHD는 DSM 진단 기준을 근거로 만들어진 WHO 성인 ADHD 자가보고척도(Adult ADHD Self-Report Scale, ASRS)를 이용하여 조사
	5. 애도 및 회복탄력성	애도는 한글판 지속비애장애-13(Korean Version of the Prolonged Grief Disorder Scale, PG-13-K) 척도를, 회복탄력성은 한국형 코너-데이비스 자아탄력성 척도(Korean Version of the Connor-Davidson Resilience Scale, K-CD-RISC) 10문항을 사용하여 조사
	6. 사회적 고립	사회적 고립은 Teo 외(2015) 논문의 은둔형 외톨이 문항 일부를 번안 및 수정하여 조사
	7. 자해행동	구체적인 자해행동 조사를 위해 자해행동척도(Deliberate Self-Harm Inventory, DSHI)의 항목을 사용하여 조사
	8. 외로움	외로움은 외로움과 사회적 고립 척도(Loneliness and Social Isolation Scale, LSIS) 6문항과 외로움의 특성(빈도, 정도, 기간) 분석을 위한 연구진 개발 문항(3문항)을 사용하여 조사
	9. 정신질환에 대한 인식	정신질환 및 정신질환자에 대한 인식을 질문하기 위해 수정판 지각된 평가절하 및 차별 척도를 사용하여, 4가지 주요 정신건강 문제(조현병, 알코올 중독, 자살 시도, 우울증)에 대한 인식 및 낙인을 조사

주: 2021년 정신건강실태조사 기준  
 자료: 보건복지부 국립정신건강센터. (2021). 2021년 정신건강실태조사 보고서. pp.50-51.

자살 문제가 우리 사회의 심각한 사회 문제 중 하나로 대두됨에 따라 2011년 정신건강실태조사부터는 자살 관련 행동(자살 생각, 자살 계획, 자살 시도)에 대한 문항이 추가 도구로 등장하기 시작하였다. WHO에서 자살예방사업의 일환으로 시행한 조사인 WHO-Suicide prevention Multisite Intervention Study on Suicidal Behaviors(SUPRE-MISS)의 설문 문항을 우리나라의 상황에 맞춰 수정하여 구성하였다. 2016년 정신질환실태 역학조사의 경우, 부가도구로 자살실태에 관한 조사를 위해 WHO World Mental Health CIDI의 설문지의 자살 생각, 자살 계획, 자살 시도에 대한 설문 문항을 사용하였다. 2021년 정신건강실태조사의 자살 관련 문항은 2016년의 문항과 거의 동일하나 자살 생각·계획·시도를 한 1년 이내의 시점을 구체적으로 묻는 보기 문항이 추가되었다.

〈표 3-8〉 정신건강실태조사의 자살 관련 문항 연도별 현황

조사 항목	2001년	2006년	2011년	2016년	2021년
자살 사고			●	●	●
자살 계획			●	●	●
자살 시도			●	●	●
내 상황				●	●

자료: 이충경, 박강규, 조맹제, 정은기, 김장규, 김선옥, ....., & 배안. (2001). 정신질환실태 역학조사 보고서. 보건복지부·국립서울정신병원. pp.281-297.  
 조맹제, 함봉잔, 홍진표, 이동우, 이준영, 정인원, ....., & 이해우. (2006). 정신질환실태 역학조사 2006. 보건복지부·서울대학교 의과대학. pp.48-50.  
 조맹제, 박종익, 배안, 배재남, 손정우, 안준호, ....., & 전승범. (2011). 2011년도 정신질환실태 역학조사. 보건복지부·서울대학교 의과대학. p.51, pp. 339-346.  
 홍진표, 이동우, 함봉진, 이소희, 성수정, 윤탁, ....., & 김선용. (2016). 2016년도 정신질환실태 조사. 보건복지부·삼성서울병원. p. 8, pp. 382-390, 부록 p.13.  
 보건복지부 국립정신건강센터. (2021). 2021년 정신건강실태조사 보고서. pp.191-194.

정신건강실태조사의 자살 관련 문항은 자살 생각·계획·시도와 그 시점을 함께 묻고 있다. 2016년 정신건강실태조사부터는 자살에 대한 의도성

을 묻는 문항으로 내 상황에 대한 문항이 추가되었다. 정신건강실태조사의 자살 관련 행동 비율을 살펴보면 여성의 자살 관련 행동의 비율이 남성보다 높게 나타나는 경향이 있음을 알 수 있다.

〈표 3-9〉 정신건강실태조사의 자살 관련 문항

조사 항목	설문 문항
자살 사고	S1) 자살하는 것에 대해 <b>진지하게 생각</b> 한 적이 한 번이라도 있습니까? 예, 아니오 S2_ONS) 자살하는 것에 대해 처음으로 생각한 때는 언제입니까? (1) 지난 2주 이내 (2) 2주에서 1개월 이내 (3) 1개월에서 6개월 이내 (4) 6개월에서 1년 이내 (5) 지난 1년 이내. 정확히는 모름 (6) 1년 초과 S2_AGO) 처음으로 자살하는 것을 생각한 때가 만으로 몇 살이었습니까? 처음 나이: 만 _____세 S2_REC) 자살하는 것에 대해 마지막으로 생각한 때는 언제입니까? (1) 지난 2주 이내 (2) 2주에서 1개월 이내 (3) 1개월에서 6개월 이내 (4) 6개월에서 1년 이내 (5) 지난 1년 이내. 정확히는 모름 (6) 1년 초과 S2_AGR) 마지막으로 자살하는 것을 생각한 때가 만으로 몇 살이었습니까? 마지막 나이: 만 _____세 S2A) <b>현재</b> 진지하게 자살에 대해서 생각하고 있습니까? 예, 아니오
자살 계획	S3) 자살하려고 <b>구체적으로 계획</b> 을 세운 적이 있습니까? 예, 아니오 S4_ONS) 자살하는 것에 대해 처음으로 계획을 세운 때는 언제입니까? (1) 지난 2주 이내 (2) 2주에서 1개월 이내 (3) 1개월에서 6개월 이내 (4) 6개월에서 1년 이내 (5) 지난 1년 이내. 정확히는 모름 (6) 1년 초과 S4_AGO) 처음으로 자살하는 계획을 세운 때가 만으로 몇 살이었습니까? 처음 나이: 만 _____세 S4_REC) 자살하는 것에 대해 마지막으로 계획을 세운 때는 언제입니까? (1) 지난 2주 이내 (2) 2주에서 1개월 이내 (3) 1개월에서 6개월 이내 (4) 6개월에서 1년 이내 (5) 지난 1년 이내. 정확히는 모름 (6) 1년 초과 S4_AGR) 마지막으로 자살하는 계획을 세운 때가 만으로 몇 살이었습니까? 마지막 나이: 만 _____세 S4A) <b>현재</b> 구체적으로 자살하려고 계획을 가지고 있습니까? 예, 아니오

조사 항목	설문 문항
자살 시도	<p>S5) 자살을 시도한 적이 있습니까? 예, 아니오</p> <p>S5A) 자살을 시도한 횟수는 몇 회입니까? ____회</p> <p>S5B) 그 중 술을 마시고 자살을 시도한 횟수는 몇 회입니까? ____회</p> <p>S6_ONS) 자살을 처음으로 시도한 때는 언제입니까? (1) 지난 2주 이내 (2) 2주에서 1개월 이내 (3) 1개월에서 6개월 이내 (4) 6개월에서 1년 이내 (5) 지난 1년 이내. 정확히는 모름 (6) 1년 초과</p> <p>S6_AGO) 처음으로 자살을 시도한 때가 만으로 몇 살이었습니까? 처음 나이: 만 ____세</p> <p>S6_REC) 자살을 마지막으로 시도한 때는 언제입니까? (1) 지난 2주 이내 (2) 2주에서 1개월 이내 (3) 1개월에서 6개월 이내 (4) 6개월에서 1년 이내 (5) 지난 1년 이내. 정확히는 모름 (6) 1년 초과</p> <p>S6_AGR) 마지막으로 자살을 시도한 때가 만으로 몇 살이었습니까? 마지막 나이: 만 ____세</p> <p>S7) 자살을 시도한 때 중에서 (가장 심각했던 때/ 1회만 시도한 경우 그 때)를 기억하십시오. 자살 시도 중 (가장 심각했던 때/ 1회만 시도한 경우 그때)에 당시 몇 살이었습니까? 만 ____세</p>
내 상황	<p>S7A) 내 상황을 가장 잘 나타내는 것을 보기에서 한 가지만 선택해 주십시오.</p> <ol style="list-style-type: none"> <li>1. 나는 진짜 죽으려고 했으며, 내가 죽지 않은 것은 단지 행운이다</li> <li>2. 나는 죽으려고 시도했지만, 확실한 방법이 아니라는 것을 알고 있었다</li> <li>3. 나는 단지 도움이 필요해서 그랬지, 죽으려고 한 것은 아니다</li> </ol>

주: 2021년 정신건강실태조사 기준  
 자료: 보건복지부 국립정신건강센터. (2021). 2021년 정신건강실태조사 보고서. pp.191-194.



〈표 3-10〉 정신건강실태조사의 자살 관련 행동 비율(%)

구분	2011년			2016년			2021년		
	전체 (비율, %)	남자 (비율, %)	여자 (비율, %)	전체 (비율, %)	남자 (비율, %)	여자 (비율, %)	전체 (비율, %)	남자 (비율, %)	여자 (비율, %)
평생									
자살 사고	15.6	13.4	17.9	15.4	14.0	16.7	10.7	9.4	12.0
자살 계획	3.3	2.4	4.1	3.0	3.1	3.0	2.5	2.1	2.8
자살 시도	3.2	2.4	4.1	2.4	2.3	2.5	1.7	1.6	1.9
1년									
자살 사고	3.7	3.1	4.4	2.9	3.1	2.8	1.3	1.0	1.7
자살 계획	0.7	0.4	1.0	0.4	0.3	0.5	0.5	0.3	0.7
자살 시도	0.3	0.1	0.5	0.1	0.0	0.2	0.1	0.1	0.1
1개월									
자살 사고	2.6	2.2	3.0	0.3	0.3	0.3	0.5	0.2	0.9
자살 계획	0.3	0.2	0.5	0.1	0.1	0.1	0.3	0.1	0.5
자살 시도	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-

자료: 조맹제, 박종익, 배안, 배재남, 손경우, 안준호, .... & 전승범. (2011). 2011년도 정신질환실태 역학조사. 보건복지부·서울대학교 의과대학. p.341.  
 홍진표, 이동우, 함봉진, 이소희, 성수정, 윤탁, .... & 김선웅. (2016). 2016년도 정신질환실태 조사. 보건복지부·삼성서울병원. p. 384.  
 보건복지부 국립정신건강센터. (2021). 2021년 정신건강실태조사 보고서. p.100.

### 3. 청소년건강행태조사

청소년건강행태조사는 청소년의 건강행태 현황을 파악하고 청소년 건강증진사업 수행을 위한 보건지표를 산출하고 국가 간 비교 가능한 청소년 보건지표를 산출하고자 실시하는 조사이다. 2004년 국가 만성병 감시체계 구축 계획의 일환으로 수립된 조사로 제1차 조사는 2005년에 실시했고, 매년 실시되고 있다.

〈표 3-11〉 청소년건강행태조사 개요

구분	설명			
조사 주기	1년			
조사 대상	전국의 중학교 및 고등학교 각각 400개교 1,200개 학급(학년별 1개 학급)의 청소년 56,213명(최종 응답 51,850명, 응답률 92.2%)			
조사 방법	익명성 자기 기입식 온라인 조사, 학교 내 컴퓨터실 일괄 참여 조사			
조사 기간	2022년 8월 29일 ~ 2022년 10월 25일			
목표 모집단	만 13세부터 만 18세까지의 우리나라 청소년			
표본 추출틀	조사 전년도 4월 기준 전국 중·고등학교			
조사 모집단	2021년 기준의 중·고등학교 재학생(전국 중학교 1학년~고등학교 3학년)			
표본 추출	2단계 층화집락추출법 - 층화 변수: 시도, 학교유형, 학급유형 - 추출 단위: 학교, 학급(학년별)			
조사 내용 (영역)	흡연 음주 신체활동 식생활	비만 및 체중조절 <b>정신건강</b> 손상 및 안전의식 구강건강	개인위생 성행태 아토피·천식 약물	인터넷중독 건강형평성 폭력 기타

주: 제18차(2022년) 청소년건강행태조사 기준  
 자료: 질병관리청. 청소년건강행태조사 누리집. 사업소개 페이지. <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.5. 인출.  
 통계청. 통계설명DB. 청소년건강행태조사 통계설명자료. <https://www.narastat.kr/meta/svc/index.do?confmNo=117058>에서 2023.7.5. 인출.  
 한국조사연구학회. (2020). 국가건강조사 2020년 표본설계. <제1세부 과제> 청소년건강행태조사 2020년 표본설계. 질병관리본부·한국조사연구학회. p.15

전국의 중학교 400개교 1,200개 학급(학년별 1개 학급)과 고등학교 400개교 1,200개 학급(학년별 1개 학급)의 56,213명 청소년을 대상으로 흡연, 음주, 신체활동, 식생활, 비만 및 체중조절, 정신건강, 손상 및 안전인식, 구강건강, 개인위생, 성행태, 아토피·천식, 약물, 인터넷 중독, 건강형평성, 폭력, 주관적 건강 인지 및 코로나19 후유증 등을 조사한다.

청소년건강행태조사는 조사 항목을 16개 영역으로 구분하여 총 114개 문항으로 구성되어 있다. 조사 항목은 청소년건강행태조사 누리집을 통해 상시로 제안할 수 있으며 흡연 분과, 음주 분과, 비만 분과, 신체활동 분과, 영양 분과, 손상 및 안전의식 분과, 정신건강 분과, 구강 분과, 이환 분과 및 사회경제위치 분과 등 분과별 자문위원회를 구성하여 도입 가능한 조사 항목을 검토하고 있다.

〈표 3-12〉 청소년건강행태조사 조사 항목 상세(총 114개 문항)

영역	문항 수	주요 조사 내용
흡연	14	평생/월간 일반담배(필련) 흡연 경험, 처음/매일 일반담배(필련) 흡연 경험 연령, 월간 일반담배(필련) 흡연량, 담배 구매 용이성, 금연 시도 경험, 평생/월간 액상형 전자담배 사용 경험, 평생/월간 필련형 전자담배 사용 경험, 월간 필련형 전자담배 사용량, 가정/공공장소 실내 간접흡연
음주	6	평생/월간 음주 경험, 처음 음주 경험 연령, 월간 음주량, 만취 경험, 주류 구매 용이성
신체활동	15	하루 60분 이상 신체활동, 고강도/중강도 신체활동 일수/시간, 근력강화운동, 주중/주말 앉아서 보내는 시간, 주간 걷기 일수/시간, 등하교나 등하원 시 신체활동 일수/시간, 주간 체육시간 운동횟수, 스포츠활동 팀 수, 신체활동 장려요인
식생활	12	아침결식/이유, 과일섭취/이유, 패스트푸드/채소/야식 섭취, 단맛나는 음료/고카페인음료/우유 섭취, 학교에서 영양식습관 교육, 물 섭취
비만 및 체중조절	7	신장, 체중, 월간 체중 조절, 신체이미지 왜곡 인지, 체중조절방법, 먹방 및 쿡방 시청빈도, 먹방 및 쿡방 시청이 식습관에 미치는 영향

70 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

영역	문항 수	주요 조사 내용
정신건강	10	스트레스 인지, 주관적 수면 충족, 주중/주말 수면시간, 우울감 경험, 자살 생각/계획/시도, 외로움, 범불안장애(GAD-7)
손상 및 안전의식	3	안전벨트 착용(승용차, 택시 앞/뒷좌석, 고속버스), 학교 손상으로 인한 병원치료, 학교에서 안전교육 경험
구강건강	10	하루 동안 칫솔질, 학교에서 점심식사 후 칫솔질 실천, 취침 전 칫솔질 실천, 구강용품 사용, 치과진료 경험, 실란트(치아홈 메우기) 경험, 스케일링 경험, 주관적 구강건강 인지, 구강증상 경험, 구강보건교육 경험
개인위생	2	비누 이용 손 씻기 실천(학교/집에서 식사 전/화장실 사용 후, 외출 후 귀가 시), 학교에서 개인 위생교육 경험
성행태	4	성관계 경험, 피임 경험/방법, 학교에서 성교육 경험
아토피·천식	6	천식/알레르기비염/아토피피부염 의사진단
약물	1	습관적 및 의도적 평생 약물 사용 경험
인터넷중독	2	주중/주말 스마트폰 사용 여부/시간
건강행태	13	성별, 태어난 연도/월, 학년, 가족구성원, 친부모/양부모와의 거주 여부, 친부모 학력, 친부모 한국 국적 여부, 친부모 태어난 나라, 주관적 학업성적, 주관적 경제적 상태, 현재 거주형태
폭력	1	폭력으로 인한 병원 치료 경험
기타	8	주관적 건강 인지, 코로나19 후유증

자료: 질병관리청. 청소년건강행태조사 누리집. 조사 내용 페이지. <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.5. 인출.

청소년건강행태조사의 경우 2005년 제1차 조사부터 자살 사고와 자살 시도에 관한 문항이 계속해서 포함되어 있다. 또한, 자살 계획에 대한 문항은 2011년 제7차 조사부터 계속해서 포함되고 있다. 2010년 제6차 조사에서 자살 시도 후 병원 치료 경험에 관한 문항이 추가된 이후 2017년 제13차 조사까지 포함되었으나 제외되었고 그 이후 2020년 제16차 조사에서 한번 포함하였으나 다시 제외되었다. 자살 시도 후 병원 치료 경험에 관한 문항은 자살 시도 경험이 있는 응답자들만을 대상으로 한다.

〈표 3-13〉 청소년건강행태조사 자살 관련 문항 연도별 현황

조사 항목	제1차 (2005)	제2차 (2006)	제3차 (2007)	제4차 (2008)	제5차 (2009)	제6차 (2010)	제7차 (2011)	제8차 (2012)	제9차 (2013)
자살 사고	●	●	●	●	●	●	●	●	●
자살 계획							●	●	●
자살 시도	●	●	●	●	●	●	●	●	●
자살 시도 후 병원 치료 경험						●	●	●	●
조사 항목	제10차 (2014)	제11차 (2015)	제12차 (2016)	제13차 (2017)	제14차 (2018)	제15차 (2019)	제16차 (2020)	제17차 (2021)	제18차 (2022)
자살 사고	●	●	●	●	●	●	●	●	●
자살 계획	●	●	●	●	●	●	●	●	●
자살 시도	●	●	●	●	●	●	●	●	●
자살 시도 후 병원 치료 경험	●	●	●	●			●		

자료: 질병관리청. 청소년건강행태조사 누리집. 설문지 페이지 <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.10. 인출.

질병관리청. 청소년건강행태조사. 각 연도별 조사표

청소년건강행태조사의 자살 관련 문항은 모두 최근 12개월 동안(지난 일년 동안)의 생각·계획·시도에 대해 묻고 있다. 2008년 제4차 조사부터는 ‘지난 일 년 동안’이라는 문구를 ‘최근 12개월 동안’으로 변경하였다. 자살 생각을, 자살 계획을 및 자살 시도율은 하락하는 추세를 보이다가 2016년 통계부터 다소 반등하는 것으로 나타난다. 자살 시도율의 하락 추세에 따라 자살 시도 후 병원치료 경험률 또한 하락 추세를 보였다.

〈표 3-14〉 청소년건강행태조사의 자살 관련 문항

조사 항목	설문 문항
자살 사고	(‘05~’07) <b>지난 일 년 동안, 심각하게</b> 자살을 생각한 적이 있습니까? (1) 지난 일 년 동안 생각한 적이 없다 (2) 지난 일 년 동안 생각한 적이 있다 (‘08~’22) <b>최근 12개월 동안, 심각하게</b> 자살을 생각한 적이 있습니까? (1) 최근 12개월 동안 생각한 적이 없다 (2) 최근 12개월 동안 생각한 적이 있다
자살 계획	(‘11~’22) <b>최근 12개월 동안</b> , 자살하기 위해 구체적인 계획을 세운 적이 있었습니까? (1) 최근 12개월 동안 없다 (2) 최근 12개월 동안 있다
자살 시도	(‘05~’07) <b>지난 일 년 동안</b> , 자살을 시도한 적이 있었습니까? (1) 지난 일 년 동안 시도한 적이 없다 (2) 지난 일 년 동안 시도한 적이 있다 (‘08~’22) <b>최근 12개월 동안</b> , 자살을 시도한 적이 있었습니까? (1) 최근 12개월 동안 시도한 적이 없다 (2) 최근 12개월 동안 시도한 적이 있다
자살 시도로 병원 치료	(‘10~’17, ’20) <b>최근 12개월 동안</b> , 자살을 시도하여 병원에서 치료를 받은 적이 있습니까? (1) 최근 12개월 동안 자살을 시도하여 병원에서 치료를 받은 적이 있다. (2) 최근 12개월 동안 자살을 시도하여 병원에서 치료를 받은 적이 없다.

자료: 질병관리청. 청소년건강행태조사 누리집. 설문지 페이지 <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.10. 인출,  
 질병관리청. 청소년건강행태조사. 각 연도별 조사표

〈표 3-15〉 청소년건강행태조사 의 자살 관련 행동 비율(%)

	자살 사고	자살 계획	자살 시도	자살 시도로 병원 치료
제1차(2005년)	22.6%		4.7%	
제2차(2006년)	23.4%		5.5%	
제3차(2007년)	23.7%		5.8%	
제4차(2008년)	18.9%		4.7%	
제5차(2009년)	19.1%		4.6%	
제6차(2010년)	19.3%		5.0%	0.6%
제7차(2011년)	19.6%	6.8%	4.3%	0.6%
제8차(2012년)	18.3%	6.3%	4.1%	0.6%
제9차(2013년)	16.6%	5.7%	4.1%	0.6%
제10차(2014년)	13.1%	4.4%	2.9%	0.5%
제11차(2015년)	11.7%	3.8%	2.4%	0.4%
제12차(2016년)	12.1%	4.0%	2.4%	0.4%
제13차(2017년)	12.1%	3.9%	2.6%	0.5%
제14차(2018년)	13.3%	4.4%	3.1%	
제15차(2019년)	13.1%	4.0%	3.0%	
제16차(2020년)	10.9%	3.6%	2.0%	0.3%
제17차(2021년)	12.7%	4.0%	2.2%	
제18차(2022년)	14.3%	4.5%	2.6%	

자료: 질병관리청, 청소년건강행태조사 누리집, 통계집 페이지. <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.10. 인출.  
질병관리청, 청소년건강행태조사, 각 연도별 통계.

#### 4. 한국복지패널조사

한국복지패널조사는 인구 집단별 생활 실태와 복지 욕구를 파악하여 정책 지원, 기획, 수행, 성과 평가 및 개선 등에 기여하고자 매년 실시되는 조사이다. 3개 패널(한국보건사회연구의 차상위 빈곤패널, 자활패널과 서울대학교의 한국복지패널)이 통합된 패널조사로 2006년 1차년도 조사가 실시되었다. 한국복지패널조사의 주요 조사 내용은 가구 조사, 가구원 조사 그리고 복지인식 조사로 구분할 수 있으며, 복지인식 조사의 경우 3년 주기 부가 조사로 각 연도마다 아동, 복지인식, 장애인 등 별도의 주제를 개발하여 일회성 혹은 경우에 따라 반복적으로 조사한다.

〈표 3-16〉 한국복지패널조사 개요

구분	설명
조사 주기	1년
조사 대상	<ul style="list-style-type: none"> <li>- 조사 대상은 표본 가구와 표본 가구에 속한 15세 이상 가구원, 그리고 부가조사 대상으로 구분</li> <li>- 조사 대상 8,169가구=최초 원 표본 가구+신규 가구-탈락 가구</li> <li>- 원 표본 가구: 2005년 인구주택총조사 자료'로부터 확률비례추출한 '2006년 국민생활실태조사' 최종 조사 완료 가구의 소득 자료를 기준으로 일반가구와 저소득층 가구를 구분하여 두 층으로부터 각각 3,500가구씩 총 7,000가구를 표본으로 선정하였으며 최종 패널가구로 구축된 표본가구는 7,072가구</li> <li>- 부가조사 대상: 14차 2019년 복지인식 부가조사 대상자 중 16차 2021년 표본에 포함된 가구원 3,130명을 조사 대상</li> </ul>
조사 방법	CAPI를 활용한 직접 면접 조사
조사 기간	<ul style="list-style-type: none"> <li>- 본 조사: 2022년 3월 28일 ~ 2022년 7월 2일</li> <li>- 신규 조사: 2022년 7월 5일 ~ 2022년 8월 17일</li> </ul>
목표 모집단	2018년 현재 전국에 거주하는 가구
표본 추출률	2005년 인구주택총조사 90% 자료에서 섬·특수시설 조사구를 제외한 전체 23만여 개 조사구
조사 모집단	2005년 인구센서스 90% 조사구(섬·특수시설 조사구 제외)
표본 추출	- 층화이중추출법을 이용하여 최종 패널 가구를 선정하였음.



구분	설명
	<p>1단계: 인구센서스 90% 자료로부터 517개 조사구를 표본으로 추출하여 가구의 소득 및 가구원의 경제활동 상태 등을 조사함.</p> <p>2단계: 1단계 조사 자료로부터 일반 가구와 저소득층 가구를 각각 3,500가구씩 총 7,000가구를 표본가구로 추출하였음.</p>
조사 항목	<ul style="list-style-type: none"> <li>- 가구 조사(가구용 조사표): 가구 일반사항, 건강 및 의료, 경제활동 상태, 사회보험 가입, 퇴직(연)금제, 개인연금 가입, 주거, 생활비(소비지출), 주관적 최저생계비, 소득, 부채, 이자, 가구 재산, 생활 여건, 국민기초생활보장제도, 근로(자녀)장려세제, 복지서비스 이용, 노인 가구의 복지서비스 이용, 아동 가구의 복지서비스 이용, 장애인 가구의 복지서비스 이용, 가족</li> <li>- 가구원 조사(가구원용 조사표): 사회보험·퇴직금(퇴직보험)·개인연금 수급, 근로, 생활실태·만족 및 의식, 사회적 환경에 대한 의식, 생활습관, 가족관계 및 정신건강, 교육, 개인사</li> <li>- 복지인식 조사(부가 조사표): 전반적인 사회적·정치적 인식과 태도, 복지 재원 및 대상범위, 정치참여와 성향</li> </ul>

주: 제17차년도 2022년 한국복지패널조사 기준  
 자료: 정은희 외. (2022). 2022년 한국복지패널 조사분석 보고서. 한국보건사회연구원, pp.3-10.  
 한국보건사회연구원. 한국복지패널 홈페이지. <https://www.koweps.re.kr:442>에서  
 2023.7.10. 인출,  
 한국보건사회연구원, & 서울대학교 사회복지연구소. (2022). 한국복지패널 17차년도 조사자  
 료 유저가이드. 한국보건사회연구원. pp.1-6.

한국복지패널조사는 표본 추출 시 중위소득 60% 미만 저소득층에 전체 표본의 약 50%를 할당하므로 저소득층 대상 정책 또는 빈곤 연구에 적합한 공공데이터이다. 또한, 전국적인 대표성을 갖는 조사로 ‘제주도’ 지역을 포함하였을 뿐 아니라 ‘농어가’ 가구 유형까지를 포함한다. 주제 별로는 가구원 배경 및 개인사, 근로활동, 경제 상황, 가구 여건(복지 욕구), 사회보장 가입 및 수급 현황, 사회적 관계망 및 정치 문제 인식 등 기타 등의 항목이 있으며, 내용에 있어서 광범위한 사회과학 영역의 내용을 포함하고 있음을 알 수 있다.

조사기준 연도는 전년도이고, 제17차년도 한국복지패널조사의 경우 2022년 조사로 조사기준 연도는 2021년이다. 유량 변수에 해당하는 조사 항목은 전년도 1월 1일부터 12월 31일까지의 기간을 기준으로 하고, 저장 변수에 해당하는 조사 항목은 전년도 12월 31일을 기준으로 한다.

76 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

〈표 3-17〉 한국복지패널조사 조사 항목(조사 주제별)

주제	세부주제	항목 포함 영역
1. 가구원 배경 및 개인사	1) 가구원 일반사항	가구용 I. 가구 일반사항 가구원용 F. 교육
	2) 부모세대	가구원용 G. 개인사
	3) 개인사	가구원용 G. 개인사
2. 근로활동	1) 직업이력	가구원용 G. 개인사
	2) 경제활동상태	가구용 III. 경제활동상태 가구원용 B. 근로
	3) 고용지원프로그램	가구원용 B. 근로가구원용
3. 경제 상황	1) 소득	가구용 VIII. 소득
	2) 지출 및 저축	가구용 VII. 생활비, X. 재산
	3) 재산 및 부채	가구용 IX. 부채 및 이자, X. 재산
	4) 기초생활	가구용 XI. 생활 여건
4. 가구 여건 (복지 욕구)	1) 보육, 교육	가구용 XV. 아동가구의 복지서비스 이용
	2) 건강 및 의료	가구용 II. 건강 및 의료A, V. 의료B
	3) 주거	가구용 VI. 주거
	4) 가족관계	가구용 XVII. 가족 가구원용 E. 생활습관, 가족관계 및 정신건강
	5) 기타 가구 여건	아동 부가조사, 장애인 부가조사
5. 사회보장 가입 및 수급 현황	1) 사회보험	가구용 IV. 사회보험, 퇴직(연금, 개인연금 가입 가구원용 A. 사회보험, 퇴직금, 개인연금 수급
	2) 공공부조	가구용 I. 가구 일반사항(기초보장 수급 형태) XIII. 국민기초생활보장 XIII-1. 근로(자녀)장려세제
	3) 사회복지서비스	가구용 XIII. 가구의 복지서비스 이용 XIV. 노인 가구의 복지서비스 이용 XV. 아동 가구의 복지서비스 이용 XVI. 장애인 가구의 복지서비스 이용 장애인 부가조사
6. 기타	1) 기타 생활 여건 및 사회적 관계망	가구원용 C. 생활실태·만족 및 의식 D. 사회적 환경에 대한 의식
	2) 생활습관 및 정신건강	가구원용 E. 생활습관, 가족관계 및 정신건강
	3) 복지·사회·정치문제 인식	복지인식 부가조사 가구원용 D. 사회적 환경에 대한 의식(성역할)

자료: 한국보건사회연구원, & 서울대학교 사회복지연구소. (2022). 한국복지패널 17차년도 조사자료 유저가이드. 한국보건사회연구원. p.21

한국복지패널조사에서는 응답 가구(원)을 원가구(원)과 신규 가구(원)으로 구분하고 있다. 원가구는 전년도 조사 당시 조사 완료된 패널가구이고, 신규 가구는 조사 기준 연도 9월 30일 이전에 분가, 결혼, 직장, 학업, 위탁, 이혼, 별거 등의 이유로 원가구(가구주 기준)로부터 분리된 가구이다. 원가구원은 전년도 가구원 조사 응답 가구원이고, 신규 가구원은 전년도 가구원 조사에 응답하지 않은 가구원이면서, 1) 조사 기준 연도에 만 15세가 된 가구원이면서 중고등학생이 아닌 자 또는 2) 조사 기준 연도 2월에 고등학교를 졸업(중퇴 포함)한 가구원 또는 3) 조사 기준 연도에 원가구 혹은 신규 가구에 새로 진입한 가구원에 해당하는 자이다.

한국복지패널조사의 경우 자살 관련 문항을 자살 사고, 자살 계획 그리고 자살 시도로 분류할 수 있다. 자살 관련 문항은 2011년 6차 조사부터 조사표에 포함되기 시작하였다.

2011년 6차 조사부터 포함되기 시작한 문항은 지금까지(생애) 여부, 처음 시기 그리고 마지막 시기에 대한 문항이다. 자살 사고에 대한 문항은 1) 지금까지의 자살 사고 여부, 2) 처음 자살 사고 시기 그리고 3) 마지막 자살 사고 시기로 구분된다. 자살 계획에 대한 문항은 1) 지금까지 구체적인 자살 계획 여부, 2) 처음 자살 계획 시기 그리고 3) 마지막 자살 계획 시기로 구분된다. 자살 시도에 대한 문항은 1) 지금까지 자살 시도 여부, 2) 처음 자살 시도 시기 그리고 3) 마지막 자살 시도 시기로 구분된다. 지금까지의 자살 사고 여부 문항은 2022년 17차 문항까지 포함되어 있으며 그 외 문항들은 2019년 14차 문항까지 포함되어 있다. 그리고 2011년 6차 조사에서는 신규 가구원과 원가구원 모두 응답하였으며, 2012년 7차 조사부터는 신규 가구원만 응답하였다.

2012년 7차 조사부터 포함되기 시작한 문항은 지난 한 해 동안의 자살 관련 행동이 있었는지에 대한 문항으로 지난 한 해 동안 자살에 대해 진지하게 생각한 적이 있는지 여부, 지난 한 해 동안 구체적인 계획을 세웠

는지 여부 그리고 지난 한 해 동안 자살을 시도했는지 여부 문항이다. 지난 한 해 동안의 자살 사고 여부 문항은 2022년 17차 문항까지 포함되어 있으며, 그 외 문항들은 2019년 14차 문항까지 포함되어 있으며, 2012년 7차 조사부터 응답하였다.

〈표 3-18〉 한국복지패널조사의 자살 관련 문항 연도별 현황

조사 항목		1차 (2006)	2차 (2007)	3차 (2008)	4차 (2009)	5차 (2010)	6차 (2011)	7차 (2012)	8차 (2013)	9차 (2014)
지금 까지/ 처음/ 마지막	자살 사고						●	●	●	●
	자살 계획						●	●	●	●
	자살 시도						●	●	●	●
지난 한 해 동안	자살 사고							●	●	●
	자살 계획							●	●	●
	자살 시도							●	●	●
조사 항목		10차 (2015)	11차 (2016)	12차 (2017)	13차 (2018)	14차 (2019)	15차 (2020)	16차 (2021)	17차 (2022)	X
지금 까지/ 처음/ 마지막	자살 사고	●	●	●	●	●	●	●	●	
	자살 계획	●	●	●	●	●				
	자살 시도	●	●	●	●	●				
지난 한 해 동안	자살 사고	●	●	●	●	●	●	●	●	
	자살 계획	●	●	●	●	●				
	자살 시도	●	●	●	●	●				

자료: 한국보건사회연구원. 한국복지패널 홈페이지 코딩북 페이지. 한국복지패널 결합데이터(1~17차) 코드북. <https://www.koweps.re.kr:442/data/book/list.do>에서 2023.7.17. 인출.  
 한국보건사회연구원. 한국복지패널 홈페이지 설문지 페이지. <https://www.koweps.re.kr:442/data/survey/list.do>에서 2023.7.17. 인출.  
 한국보건사회연구원, & 서울대학교 사회복지연구소. 한국복지패널조사. 각 연도별 조사표

한국복지패널의 자살 관련 문항은 모두 조사일 현재를 기준으로 자살 생각·계획·시도에 대해 묻고 있다. 생애 경험과 지난 한 해 동안의 경험 모두를 묻고 있어 2006년부터 자살 관련 경험에 대한 연도별 종단 연구가 가능하다.

〈표 3-19〉 한국복지패널조사 자살 관련 문항

조사 항목	설문 문항
지금까지 처음 마지막	자살 사고 (‘06~’19) (모든 응답자) 귀하께서는 지금까지 자살하는 것에 대해 진지하게 생각한 적이 한 번이라도 있습니까? (※ 조사시점(20**년 조사일 현재)을 기준으로 현재까지 기간에 대해 응답해 주십시오.) (1) 예, (2) 아니오 (위 문항의 (1)번 응답자만) 귀하께서 자살하는 것에 대해서 처음/마지막으로 생각한 때는 언제입니까? 처음 나이 만 _____세, 마지막 나이 만 _____세 (‘20~’22) (모든 응답자) 귀하께서는 지금까지 자살하는 것에 대해 진지하게 생각한 적이 한 번이라도 있습니까? (※ 조사시점(20**년 조사일 현재)을 기준으로 현재까지 기간에 대해 응답해 주십시오.) (1) 예, (2) 아니오
	자살 계획 (‘06~’19) (모든 응답자) 귀하께서는 지금까지 자살하려고 구체적으로 계획을 세운 적이 있습니까? (※ 조사시점(20**년 조사일 현재)을 기준으로 현재까지 기간에 대해 응답해 주십시오.) (1) 예, (2) 아니오 (위 문항의 (1)번 응답자만) 귀하께서 자살하는 것에 대해서 처음/마지막으로 계획을 세운 때는 언제입니까? 처음 나이 만 _____세, 마지막 나이 만 _____세
	자살 시도 (‘06~’19) (모든 응답자) 귀하께서는 지금까지 자살을 시도한 적이 있습니까? (※ 조사시점(20**년 조사일 현재)을 기준으로 현재까지 기간에 대해 응답해 주십시오.) (1) 예, (2) 아니오 (위 문항의 (1)번 응답자만) 귀하께서 자살을 처음/마지막으로 시도한 때는 언제입니까? 처음 나이 만 _____세, 마지막 나이 만 _____세
지난 한 해 동안	자살 사고 (‘07~’19) (모든 응답자) 귀하께서는 20**년 조사일 현재를 기준으로 지난 한 해 동안 자살하는 것에 대해 진지하게 생각한 적이 한 번이라도 있습니까? (1) 예, (2) 아니오 (‘20~’22) (모든 응답자) 귀하께서는 20**년 조사일 현재를 기준으로 지난 한 해 동안 자살하는 것에 대해 진지하게 생각한 적이 한 번이라도 있습니까? (1) 예, (2) 아니오
	자살 계획 (‘07~’19) (모든 응답자) 귀하께서는 20**년 조사일 현재를 기준으로 지난 한 해 동안 자살하려고 구체적으로 계획을 세운 적이 있습니까? (1) 예, (2) 아니오
지난 한 해 동안	자살 시도 (‘07~’19) (모든 응답자) 귀하께서는 20**년 조사일 현재를 기준으로 지난 한 해 동안 자살을 시도한 적이 있습니까? (1) 예, (2) 아니오

자료: 한국보건사회연구원. 한국복지패널 홈페이지 설문지 페이지. <https://www.koweps.re.kr:442/data/survey/list.do>에서 2023.7.17. 인출.

한국보건사회연구원, & 서울대학교 사회복지연구소. 한국복지패널조사. 각 연도별 조사표

〈표 3-20〉 한국복지패널조사의 자살 관련 행동 빈도 수

	자살 사고*	자살 계획*	자살 시도*	지난 한 해 동안 자살 사고	지난 한 해 동안 자살 계획	지난 한 해 동안 자살 시도
1차(2006년)						
2차(2007년)						
3차(2008년)						
4차(2009년)						
5차(2010년)						
6차(2011년)	1334	243	160			
7차(2012년)	366	86	60	392	55	21
8차(2013년)	18	2	1	565	75	18
9차(2014년)	9	3	3	602	74	26
10차(2015년)	21	10	4	366	57	9
11차(2016년)	22	4	3	321	38	11
12차(2017년)	13	2	0	255	48	12
13차(2018년)	10	3	1	281	36	10
14차(2019년)	7	0	0	232	32	11
15차(2020년)	6			189		
16차(2021년)	8			184		
17차(2022년)	129			164		

주: \*는 2011년 6차 조사는 진 가구원 대상이며, 2011년 7차 조사부터 신규 가구원만 응답하였고, 응답 사례 수가 적으니 해석에 주의를 요함.

자료: 한국보건사회연구원, 한국복지패널 홈페이지 데이터 페이지, 한국복지패널 결합 데이터(1~17차), <https://www.koweps.re.kr:442/data/data/list.do>에서 2023.7.18. 인출.

## 5. 한국의료패널조사

한국의료패널조사는 질환, 의료이용 행태와 의료비 지출 규모 및 건강 관련 인식과 행태 등에 대한 개인 및 가구 단위의 통계를 생산하고 관련 요인들을 분석할 수 있는 패널 데이터를 구축하고 보건·의료 정책 수행의 기초 자료를 제공하자 매년 실시되는 조사이다. 제1기 한국의료패널은 2008년 1차 조사를 시작으로 2019년 총 14차 조사까지 진행되었으며, 이를 개편하여 2020년부터는 제2기 한국의료패널조사가 시작되었다. 제2기 한국의료패널의 조사 영역은 제1기 한국의료패널과 유사하나 사회변화, 의료이용 현황 등을 반영하여 일부 조사 내용을 개선하였다. 개인 및 가구의 의료이용 및 의료비 지출 현황을 살펴볼 수 있게 의과, 한방, 치과, 건강검진으로 나누어 의료이용의 세부 정보를 조사한다. 그리고 고령화 등 사회 변화를 살펴볼 수 있도록 돌봄비용, 건강증진비용, 보건의료 제품 구입비용 등도 조사한다.

〈표 3-21〉 한국의료패널조사 개요

구분	설명
조사 주기	1년
조사 대상	2기 원표본: 전국 17개 시도에 거주하고 있는 가구와 가구원 (전체 조사구 수는 약 700개 조사구로서 전국 약 8,500가구와 그 가구에 속해 있는 가구원을 대상으로 조사)
조사 방법	CAPI를 활용한 직접 면접 조사
조사 기간	- 2019년 사전 조사: 2019년 8월 ~ 2019년 11월 - 2020년 1차 조사: 2020년 6월 ~ 2020년 10월 - 2021년 2차 조사: 2021년 3월 ~ 2021년 7월
표본 추출틀	2016년 등록센서스
추출 방법	확률비례 2단 층화집락추출
층화 변수	17개 시·도, 동부/읍·면부

82 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

구분	설명
표본 조사구 배분	지역별로 제공근 비례 배분
표본 조사구	708개
가구 추출 방법	계통 추출
가구 수	8,500가구
조사 항목	<ul style="list-style-type: none"> <li>- 가구 조사:(가구용 조사표): 가구 및 가구원 일반 사항, 경제활동상태, 소득(가구원, 가구), 부채 및 자산, 지출, 만성질환 이환, 일반의약품 및 의료 관련 지출, 의료서비스 이용(응급(환자 경험 포함), 입원(환자 경험 포함), 외래, 건강검진), 장기요양 및 기타 돌봄 서비스(출산 포함), 민간의료보험1</li> <li>- 가구원 조사(가구원용 조사표): 건강수준(정신건강, 삶의 질), 건강생활 습관, 의료 접근성, 상용치료원, 민간의료보험2</li> <li>- 2021년에는 부가조사(HLIT)를 실시하였음.</li> </ul>

주: 제2기 한국의료패널조사 기준  
 자료: 한국보건사회연구원, & 국민건강보험공단. 한국의료패널 홈페이지. 조사개요 페이지.  
<https://www.khp.re.kr:444/>에서 2023.7.31. 인출.

제2기 한국의료패널조사 2020년 연간데이터를 예를 들어 살펴보면, 조사 기준 시점은 조사 항목별로 상이하다. 저량 변수에 해당하는 조사 항목은 조사일 또는 2020년 12월 31일을 기준 시점으로 하고, 유량 변수에 해당하는 조사 항목의 경우 2020년 1월 1일부터 2020년 12월까지를 기준 시점으로 한다. 건강생활 습관 및 건강수준 등에 관한 조사 항목들은 현재, 평소, 최근 1개월간 및 최근 1년 동안 등 문항별 시점이 상이하다.

한국의료패널에서 자살 관련 문항은 조사 영역 중 건강수준 영역에 해당하며 모두 최근 1년 동안의 자살 사고 여부에 대해서만 묻고 있다. 제1기 한국의료패널의 경우 2008년, 2014년, 2015년 및 2016년 조사에서는 자살 관련 문항이 제외되어 있어 시계열적 단절이 발생한다. 또한, 한국의료패널은 제1기에서 제2기로 개편되면서 패널 가구들이 교체되었기에 2019년부터 시계열 단절이 발생함에 유의할 필요가 있다.



〈표 3-22〉 한국의료패널조사의 자살 관련 문항 연도별 현황

조사 항목	2008년	2009년	2010년	2011년	2012년	2013년	2014년
자살 사고		●	●	●	●	●	
	2015년	2016년	2017년	2018년	2019년*	2020년*	2021년*
			●	●	●	●	●

주: \*는 2기 한국의료패널조사에 해당함.  
 자료: 한국보건사회연구원, & 국민건강보험공단. 한국의료패널 홈페이지. 설문지 페이지.  
<https://www.khp.re.kr:444/web/data/board/list.do?bbsid=63>에서 2023.7.17. 인출.  
 한국보건사회연구원, & 국민건강보험공단. 한국의료패널조사. 각 연도별 조사표.

〈표 3-23〉 한국의료패널조사의 자살 관련 문항

조사 항목	설문 문항
자살 사고	(’09) 지난 1년 동안 죽고 싶다는 생각을 해 본 적이 있습니까? (1) 예, (2) 아니오(없었음) (’10~’13, ’17~’18) 최근 1년간 죽고 싶다는 생각을 해 본 적이 있습니까? (1) 있음, (2) 없음 (’19~’21) 최근 1년 동안 죽고 싶다는 생각을 하신 적이 있습니까? (1) 있음, (2) 없음

자료: 한국보건사회연구원, & 국민건강보험공단. 한국의료패널 홈페이지. 설문지 페이지.  
<https://www.khp.re.kr:444/web/data/board/list.do?bbsid=63>에서 2023.7.17. 인출.  
 한국보건사회연구원, & 국민건강보험공단. 한국의료패널조사. 각 연도별 조사표.

〈표 3-24〉 한국의료패널조사의 자살 관련 행동(자살 사고) 빈도 수

		전체	남자	여자			전체	남자	여자
제1기	2008년				제1기	2015년			
	2009년	1449	485	964		2016년			
	2010년	877	303	574		2017년	409	161	248
	2011년	729	262	467		2018년	353	146	207
	2012년	667	262	405	제2기	2019년	862	320	542
	2013년	636	264	372		2020년	563	218	345
	2014년					2021년			

자료: 한국보건사회연구원, & 국민건강보험공단. 한국의료패널조사. 각 연도별 데이터 원시자료.

## 제2절 자살 고위험군 관련 선행연구 실증 분석 변수

### 1. 자살 고위험군 관련 선행연구 실증 분석 변수

자살 고위험군 관련 분석 변수를 도출하고자 ‘자살’과 공공데이터 키워드를 기준으로 KCI 등재 학술지를 검토하였다. 본 연구에서 활용하고자 하는 공공데이터는 ‘한국의료패널’, ‘한국복지패널’, ‘국민건강영양조사’, ‘정신건강실태조사(구 정신질환실태 역학조사)’, 및 ‘소아·청소년 정신건강실태조사’로 해당 공공데이터명과 ‘자살’ 키워드가 초록에 등장하는 학술지를 검토 대상으로 하였다.

선행연구에서 자살 고위험군 관련 실증 분석을 위한 모형 설정 시 종속 변수를 살펴보면 크게 자살 관련 변수로는 1) 자살 생각, 2) 자살 시도 및 3) 자살 계획이 있으며, 그 외로 4) 스트레스, 5) 우울 및 6) 불안이 있다. 선행연구에서 설정한 자살 관련 종속변수는 범주형 변수로 대부분의 연구에서는 카이제곱 검정통계량을 사용한 독립성 검정 또는 다중 회귀 분석 결과 유의한 변수만을 대상으로 다변량 로지스틱 회귀 분석을 실시하였다.

선행연구에서 설정한 자살 관련 독립변수는 범주형 변수와 연속형 변수를 다양하게 사용하였으며 대부분의 연구에서는 특성별 분류에 따라 독립변수를 그룹화하였다. 먼저, 특성별 분류는 1) (인구·사회·경제적) 일반적 특성, 2) 건강행태 특성, 3) 신체건강 특성, 4) 정신건강 특성, 5) 심리·사회적 특성, 6) 가족 특성 그리고 7) 기타 미분류로 정리할 수 있다.

일반적 특성 변수는 모든 분석 대상에서 1) 공통적으로 활용하는 변수와 해당 연구의 분석 대상(아동·청소년/청년/여성/성인/노인/저소득층)

에 따라 일반적 특성 변수로 분류하기도 하는 2) 특징적인 일반적 특성 변수로 구분할 수 있다. 공통적인 일반적 특성 변수는 연령, 결혼상태, 교육 수준, 경제활동, 소득 및 지역 변수가 있다. 아동·청소년은 학업 관련 변수, 청년은 1인 가구 여부 및 경제 상태 관련 변수, 그리고 여성은 가족 구성이나 자녀 및 결혼 관련 변수 등이 특징적인 일반적 특성 변수에 해당한다. 일반 성인은 의료비나 만성질환 등 보건·의료 관련 변수, 노인은 보조금이나 공적 서비스 수급 관련 변수 등 제도 관련 변수 그리고 저소득층의 경우 종교 관련 변수 등이 특징적인 일반적 특성 변수에 해당한다.

〈표 3-25〉 선행연구의 분석 변수 정리 요약

종속변수	독립변수		
	특성별 분류	변수명	
<ul style="list-style-type: none"> <li>- 자살 생각 (자살 사고, 자살충동)</li> <li>- 자살 시도</li> <li>- 자살 계획</li> <li>- 스트레스 (스트레스 인지)</li> <li>- 우울 (우울감 경험)</li> <li>- 불안</li> </ul>	일반적 특성 (인구·사회·경제적)	공통적	연령, 결혼 상태(배우자 유무, 동거상태), 교육 수준(최종 학력), 경제활동 참여 상태(경제활동 여부, 고용 형태), 월평균 소득(연간 총가구소득, 가구당 수입, 가구 총소득, 개인별 연간 근로소득, 월수입), 거주 지역
		아동·청소년	학교 유형, 학업 수준, 학년, 학업 성적, 경제적 수준
		청년	1인 가구 여부, 실직 경험, 주관적 경제 상태
		여성	대가족 여부, 2명 이상의 자녀 수 여부, 혼인관계 유지 여부, 직업 형태
		성인	개인지출 의료비, 만성질환 상태
		노인	정부보조금 수급 여부, 공적 복지서비스 수급 여부(국민기초생활보장 수급 여부와 의료급여 수급 여부 및 1, 2종 여부, 노인장기요양보험 수급 여부), 만성질환 여부, 의료보장형태
		저소득층	종교

종속변수	독립변수	
	특성별 분류	변수명
	건강행태 특성	주관적 건강 상태, 수면의 질, 규칙적인 신체활동 유무, 현재 흡연 유무, 현재 음주 유무, 체질량지수, 병·의원 미충족 의료 경험 유무, 수면시간, 음주 빈도, 평균 음주량, 과음빈도, 운동 여부, 규칙적인 식사 여부, 체중조절 노력, 지각된 스트레스
	신체건강 특성	알레르기 질환 유무, 만성질환 유무, 시력문제 유무, 청력문제 유무, 섭식문제 유무, 활동제한 유무, 장애 여부, 미충족 의료 경험 여부, 신체적 일상생활활동(ADL), 도구적 일상생활활동(IADL), 신체 증상 특성(통증 증상군, 소화기 증상군, 가성신경증상군), 동반 상병 지수(CCI 지수, Charlson comorbidity Index), 질환 및 사고중독으로 인해 몸이 아프거나 불편했던 경험, 질병이환
	정신건강 특성	우울감, 스트레스, 자살 생각 유무, 스트레스 인지 정도, 불안감, 정신건강 문제로 전문가 상담 유무, 약물복용 유무, 좌절감, 문제음주 경험 여부, 주말 수면시간, 삶의 질, 자살충동, 진단군(우울장애 진단, 불안장애 진단, 중복진단군), 스트레스(경제적/질병/가족 간 갈등/사고), 스트레스 정도(부모/형제자매/외모/신체 건강/심리 건강/경제적 문제/친구/이성/선·후배/선생님/진로/학업 등으로 인한 스트레스)
	심리·사회적 특성	사회 안전 인식, 사회 신뢰도, 차별 경험, 대인관계 욕구, 자아존중감, 물질적 박탈(음식/주거/생활필수재/의료/교육/신용박탈), 박탈 경험(식생활/주거/교육/사회보장/직업/사회적/건강), 주관적 계층인식, 부모지원, 친구관계, 교사관계
	가족 특성	가족건강성(가족탄력성, 상호 존중과 수용, 질적 유대감, 경제적 안정과 협력, 가족 문화와 사회 참여), 어린 시절 가족관계, 어린 시절 학대 경험
	기타 미분류	아토피 피부염 환아(아토피 피부염 유병 소아청소년)의 주양육자(어머니) 여부

(표 3-26) 선행연구에서의 분석 변수 정리

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	종속변수	독립변수
1	설문조사 전문 업체의 온라인 마스터 샘플 데이터	같이만 다른 그들, 청년: 성별, 자살 생 각과 자살 시 도 영향요인의 탐색 연구	강지원 외	2023	보건사회 연구	청년 (만 20~ 39세)	<b>자살 생각</b> (지난 1년간 자살 생각 여부) <b>자살 시도</b> (지난 1년간 자살을 시도한 적이 있는지)	<b>사회경제적 특성</b> (연령, 결혼 상태, 교육 수준, 1인 가구 여부, 고용 형태, 실직 경험, 월평균 소득, 주관적 경제 상 태), <b>건강 특성</b> (주관적 건강 상태, 수면의 질, 우울감, 스 트레스), <b>심리사회적 특성</b> (사 회 안전 인식, 사회 신뢰도, 차별 경험, 대인관계 욕구), <b>가족 특성</b> (가족건강성(가족탄 력성, 상호 존중과 수용, 질 적 유대감, 경제적 안정과 힘 력, 가족 문화와 사회 참여), 어린 시절 가족관계, 어린 시 절 학대 경험)
2	청소년 건강행태 조사	알레르기 질환 청소년의 정신 건강: 2020 청소년 건강 행태 온라인 조사 활용	박초희	2022	산업융합 연구	청소년 (중학교 1학년~ 고등학교 3학년)	<b>스트레스 인지</b> (평상시 스트레스를 대단히 많이 또는 많이 느끼는 경우) <b>우울감 경험</b> (최근 12개월 동안 2주 내내 일상 생활을 중단할 정도로 슬프거나 절망감을 느낀 적이 있는 경우) <b>자살 생각</b> (최근 1년 동안 죽고 싶다는 생각을 해본 경우)	<b>일반적 특성</b> (거주 지역, 유형, 성별, 학교 유형, 경제적 수 준, 학업 수준), <b>알레르기 질 한</b> 유무(진식, 알레르기 비 염, 아토피 피부염)

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	증속변수	독립변수
3	한국 의료패널	만 19-39세 성인의 성별에 따른 자살생각 영향 요인	임은정 정계현	2022	상담심리 교육복지	성인 (만 19~ 39세)	<b>자살 생각</b> (최근 1년 동안, 죽고 싶다는 생각을 한 적이 있습니까?)	<b>일반적 특성</b> (성별, 나이, 결 혼상태, 동거상태, 경제활동 참여 상태, 만성질환 상태), <b>건강행태 특성</b> (규칙적인 신체 활동 유무, 현재 흡연 유무, 현재 음주 유무, 체질량지수, 병의원 미충족 의료 경험 유 무), <b>정신건강 특성</b> (자살 생 각 유무, 스트레스 인지 정 도, 주관적 건강상태, 우울 감, 불안감, 정신건강 문제로 전문가 상담 유무, 약물복용 유무)
4	한국 의료패널	건강행위와 건 강상태가 베이 비람세대 여성 의 자살생각에 미치는 영향: 2017년 한국 의료패널 자료 이용	박민정	2021	한국융합 학회논문지	여성 (1955년~ 1963년생)	<b>자살 생각</b> (최근 1년간 죽고 싶다는 생 각을 해본 적이 있습니까?)	<b>인구사회학적 특성</b> (연령, 배 우자 유무, 교육 수준, 경제 활동 여부, 연간 총 가구소 득), <b>건강행위</b> (흡연, 음주, 수 면시간, 신체활동), <b>신체건강</b> (만성질환 유무, 시력문제 유 무, 청력문제 유무, 섭식문제 유무, 활동제한 유무), <b>정신 건강</b> (스트레스, 좌절감, 불안, 우울)

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	증속번호	독립번호
5	한국 의료패널	노인의 자살생각에 영향을 미치는 요인에 관한 연구: 객관적 및 주관적 건강상태에 따른 집단별 비교를 중심으로	성혜연 이성규 나지훈	2021	한국사회 복지조사 연구	노인 (55세 이상)	<b>자살 생각</b> (최근 1년간 죽고 싶다는 생각을 해 본 적이 있는가?)	<b>인구사회학적 특성</b> (연령, 성별, 최종 학력, 결혼상태, 정보조금 수급 여부), <b>신체건강 특성</b> (장애 여부, 미충족 의료 경험 여부), <b>정신건강 특성</b> (불안, 우울, 스트레스, 문제음주 경험 여부)
6	한국 의료패널	한국 성인의 생애주기별 자살생각 영향 요인	방소연	2021	한국산학기술학회 논문지	성인 (만 19세 이상)	<b>자살 생각</b> (최근 1년간 죽고 싶다는 생각을 해 본 적이 있는가?)	<b>일반적 특성</b> (연령, 성별, 교육 정도, 배우자 유무, 경제 활동 유무, 가구당 수입), <b>신체적 요인</b> (졸연상태, 음주상태, 규칙적인 신체활동, 체질량지수, 만성질환, 주관적 건강상태), <b>정신적 요인</b> (스트레스, 불안, 우울, 약물복용 여부, 주말 수면시간, 삶의 질)
7	한국 의료패널	중년여성의 신체활동과 우울 및 자살생각에 대한 연구	김영희	2020	차세대 융합기술 학회논문지	여성 (40~60세)	<b>자살 사고</b> (최근 1년간 죽고 싶다는 생각을 해 본 적이 있는가?) <b>우울</b> (최근 1년간 2주 이상 연속으로 일상생활에 지장이 있을 정도로 많이 슬퍼거나 불행하다고 느낀 적이 있는가?)	<b>일반적 특성</b> (나이, 소득, 결혼상태, 교육, 음주, 흡연, 경제활동, BMI, 정신-신체적 스트레스), <b>신체활동</b> (격렬한 신체활동, 중등도 신체활동, 걷기), <b>우울, 자살 사고</b>

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	증속번호	독립번호
8	한국 의료패널	노인의 신체기능 상태에 따른 정신건강에 미치는 영향에 서 성별 차이: 2016년 한국 의료패널 자료를 바탕으로	방소연 임은정	2020	상담심리 교육복지	노인 (만 65세 이상)	<p><b>우울</b> (최근 1년 동안 2주 이상 연속으로 일상생활에 지장이 있을 정도로 많이 슬펐거나 불행하다고 느낀 경험)</p> <p><b>불안</b> (지난 한 달 동안 미래에 대해 불확실하다고 느끼거나 불안해 한 경험)</p> <p><b>스트레스</b> (지난 한 달 동안 살아가는 데 신체적·정신적으로 감당하기 힘들다고 느낀 경험을 각각 항상 그렇다(5점)에서 전혀 없었다(1점)의 5점 척도로 측정)</p> <p><b>자살충동</b> (최근 1년 동안 죽고 싶다고 생각해 본 경험 유무)</p>	<p><b>일반적 특성</b>(연령, 교육 수준, 배우자 동거, 경제활동, 가구 총소득), <b>건강 관련 특성</b>(주관적 건강상태, 현재 흡연, 음주, 규칙적인 신체활동, 체질량지수, 만성질환), <b>신체기능 상태</b>(신체적 일상생활 활동(ADL), 도구적 일상생활 활동(IADL)), <b>정신건강</b>(우울, 불안, 스트레스, 자살충동)</p>
9	한국 복지패널	한국 노인의 자살생각에 관한 응복합에 중요한 : 패널 자료분석	김현주	2019	한국융합 학회논문지	노인 (만 65세 이상)	<p><b>자살 생각</b> (최근 지난 1년간 죽고 싶은 생각을 해 본 적이 있는지 여부)</p>	<p><b>일반적 특성</b>(성별, 연령), <b>공적 복지서비스</b> 수급(국민기초생활보장 수급 여부와 의료급여 수급 여부 및 1, 2종 여부), 노인장기요양보험 수급 여부), 주관적 건강상태, 우울, 자아존중감</p>



구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	종속변수	독립변수
10	한국 복지패널	물질적 박탈과 음주가 자살행위에 미치는 영향 - 한국복지패널데이터 2012-2017년 종단연구를 중심으로	김미숙 김윤영	2018	한국케어매니지먼트연구	저소득층 (중위소득 60% 이하)	<b>자살 행위</b> 1) <b>자살 생각</b> (지난 한 해 동안 진지하게 자살 생각을 했는지 여부) 2) <b>자살 계획</b> (지난 한 해 동안 구체적으로 자살 계획을 했는지 여부) 3) <b>자살 시도</b> (지난 한 해 동안 자살 생각을 했는지 여부)	<b>인구학적 특성</b> (연령, 성별, 교육 수준, 종교, 결혼 상태), <b>물질적 박탈</b> (음식/주거/생활 필수제/의료/교육/신용 박탈), <b>음주</b> (음주 빈도, 평균 음주량, 과음빈도, 문제음주 경험)
11	한국 의료패널	성인의 체질량 지수(BMI)와 자살생각의 관계 = 2009~2013년 한국 의료패널자료를 활용한 연구	이중익	2018	한국콘텐츠학회논문지	성인	<b>자살 생각</b> (최근 1년간 죽고 싶다는 생각을 해 본 적이 있습니까?)	<b>성별, 연령, 교육 수준, 소득 수준, 개인지출 의뢰비, 체질량 지수, 만성질환 수, 우울감</b>
12	한국 복지패널	사회경제적 박탈 경험에 노인의 자살생각에 미치는 영향: 6가지 박탈 유형을 중심으로	강동훈 김윤태	2018	한국노년학	노인 (65세 이상)	<b>자살 생각</b> (귀하께서는 조사일 현재를 기준으로 지난 한 해 동안 자살하는 것에 대해 진지하게 생각한 적이 한 번이라도 있습니까?)	<b>인구학적 요인</b> (성별, 연령, 결혼 상태, 교육 수준, 소득, 거주 지역), <b>박탈 경험</b> (식생활/주거/교육/사회보장/직업/사회화적/건강)

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	증속번호	독립번호
13	한국 의료패널	자녀의 아토피 피부염 유병이 주양육자의 정신건강에 미치는 영향	최세문 이해영	2017	대한 보건연구	여성 (아토피 피부염 환자의 주양육자, 어머니)	<p><b>심한 스트레스</b> (본인 또는 가족의 질병, 자녀 교육, 가족 갈등 등 원인 별 스트레스 수준 질문)</p> <p><b>우울감</b> (최근 1년 동안 2주 이상 연속으로 일상생활에 지장이 있을 정도로 많이 슬퍼거나 불행하다고 느낀 적이 있었습니까?)</p> <p><b>자살 생각 경험</b> (지난 1년 동안 죽고 싶다는 생각을 해 본 적이 있습니까?)</p>	<p><b>사회경제적 요인</b>(연령, 학력 수준, 직업, 경제활동 참여 유무, 대가족 여부, 2명 이상의 자녀 수 여부, 혼인관계 유지 여부, 가구 소득수준),</p> <p><b>아토피 피부염 환자</b>(아토피 피부염 유병 소아청소년)의</p> <p><b>주양육자(어머니) 여부</b></p>
14	정신질환 실태 역학조사	불안 장애와 주요 우울장애에서 나타나는 신체 증상과 증상군에 따른 자살 사고, 계획, 행동과의 관계 고찰	안준석 외	2016	정신신체의학	성인 (만 18세 이상 74세 이하)	<p><b>자살 사고</b> (자살하는 것에 진지하게 생각한 적이 한 번이라도 있습니까?)</p> <p><b>자살 계획</b> (자살하려고 구체적으로 계획을 세운 적이 있습니까?)</p> <p><b>자살 시도</b> (자살을 시도한 적이 있습니까?)</p>	<p><b>인구학적 특성</b>(성별, 나이, 학력, 결혼 상태, 취업 상태, 월수입, 거주지역), <b>진단군</b>(우울장애 진단, 불안장애 진단, 중보진단군), <b>신체 증상 특성</b>(통증 증상군, 소화기 증상군, 가성신경증상군)</p>

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	증속변수	독립변수
15	한국 의료패널	노인의 계층인식이 자살충동영향에 미치는 영향과 스트레스와 주관적 건강인식의 이중매개효과	남석인 이희정 김성준 김보경	2016	보건사회 연구	노인 (65세 이상)	<b>자살 충동</b> (최근 1년간 죽고 싶다는 생각을 해 본 적이 있습니까?)	<b>인구사회학적 특성</b> (성별, 연령, 교육 수준, 혼인상태, 만성질환 여부, 개인별 연간 근로소득, 의료보장형태), <b>주관적 계층인식, 스트레스, 주관적 건강인식</b>
16	한국 의료패널	노인의 성별에 따른 자살생각과 영향요인 - 2010년도 한 국가의료패널자료를 이용하여	박민정	2015	한국사료 분석학회	노인 (65세 이상)	<b>자살 생각</b> (최근 1년간 죽고 싶다는 생각을 해 본 적이 있습니까?)	<b>인구사회학적 요인</b> (성별, 연령, 교육 수준, 배우자 유무, 경제활동 여부), <b>정신건강 요인</b> (우울감 유무, 정신과 약 투약 여부, 경제적 스트레스, 질병 스트레스, 가족 간 갈등 스트레스, 사고 스트레스), <b>신체건강 요인</b> (주관적 건강상태, 활동 제한 여부, 만성질환 유무, 흡연 여부, 음주 여부, 운동 여부 및 규칙적인 식사 여부)

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	중속변수	독립변수
17	한국 의료패널	신체활동 특성이 노인 만성 질환자의 정신 건강에 미치는 영향	박재우 황병덕	2015	대한 보건연구	노인 (1956년생 이하(55세 이상))	<p><b>정신건강 특성</b></p> <p>1) <b>스트레스 정도</b> (지난 한 달 동안, 살아감에 정신적, 신체적으로 감당하기 힘들다고 느낀 적이 있습니까?)</p> <p>2) <b>우울감 경험</b> (최근 1년 동안의 사건으로 일상생활에 지장이 있을 정도로 많이 슬펐거나 불행하다고 느낀 경험 유무)</p> <p>3) <b>자살충동 경험</b> (최근 1년 동안의 사건으로 일상생활에 지장이 있을 정도로 죽고 싶다고 생각할 때 본 경험 유무)</p>	<p><b>노인 분류</b>(55~65세 연소노인, 65~75세 중간노인, 75세 이상 고령노인), <b>만성질환 유무</b>, <b>인구사회학적 특성</b>(성별, 교육 수준(중졸 이하고졸/대졸 이상), 경제활동 유무, 혼인상태(혼인/중혼인/중 이혼)), <b>신체활동 특성</b>(주관적 건강상태, 신체활동량 유무, 흡연경험 유무, 과음 유무, 건강신체 활동 도량 유무), <b>동반상병 지수</b>(CCI 지수, Charlson comorbidity Index)</p>
18	한국 아동-청소년 정신건강 실태조사	인터넷 게임 과몰입이 청소년의 자살생각에 미치는 영향: 우울감의 매개효과 분석	민다경 이한나 송인한	2014	디지털 융복합연구	청소년 (초(4~6학 년)·중·고 등학생)	<p><b>자살 생각</b> (최근 1년간 자살 생각을 해 본 적이 있는지)</p>	<p><b>인터넷 게임 과몰입, 우울감, 인구사회학적 특성</b>(성별, 경제수준, 학년, 학업 성적), <b>스트레스 정도</b>(부모/형제자매/외모/신체건강/심리 건강/경제적 문제/친구/이성/선·후배/선생님/진로/학업 등으로 인한 스트레스), <b>부모지원, 친구관계, 교사관계</b></p>

구분	데이터명	논문/보고서명	저자	출판년도	학술지명	정책 대상	증속번호	독립번호
19	국민건강 영양조사	우리나라 노인 의 자살 생각 에 영향을 미 치는 요인에 관한 연구	손신영	2014	보건의로 산업학회지	노인 (65세 이상)	<b>자살 생각</b> (최근 1년 동안 죽고 싶다는 생각을 해 본 적이 있습니 까?)	<b>일반적 특성</b> (성별, 연령, 학 력, 소득, 경제활동 여부), <b>긴 강행태 특성</b> (음주, 흡연, 중 등도 신체활동, 체중조절 노 력, 지각된 스트레스), <b>건강 관련 특성</b> (주관적 건강상태, 질한 및 사고중독으로 인해 몸이 아프거나 불편했던 경 험, 활동제한 여부, 우울)
20	국민건강 영양조사	여성자살 현황 및 정책방안	김영택 김인순 주영수 이지연	2009	한국여성 정책연구원	여성 (19세 미만, 20-59세, 60세 이상)	<b>자살 생각</b> (최근 1년 동안 죽고 싶다는 생각을 해 본 적이 있습니 까?)  <b>자살 시도</b> (최근 1년 동안 실제로 자살 시도를 해 본 적이 있습니 까?)	<b>사회인구학적 변수</b> (소득, 성 별, 연령, 결혼 상태, 직업 형 태), <b>건강 변수</b> (주관적 건강, 질병이환), <b>정신건강 변수</b> (스 트레스 정도, 2주 연속 우울 감 여부)

## 2. 자살 고위험군 관련 선행연구 실증 분석 결과

자살 고위험군 관련 선행연구 실증 분석 결과를 살펴보면, 동일한 정책 대상이라 할지라도 어떠한 데이터를 활용하느냐에 따라 통계적으로 유의한 변수가 다양하게 나타난다. 동일한 데이터를 활용하는 경우에도 모형 설정에 따라 통계적으로 유의한 변수가 다르게 나타나고 회귀 계수의 부호도 다르게 나타남을 알 수 있다.

〈표 3-27〉 선행연구 실험 분석 결과: 독립변수의 통계적 유의성

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
1	청년	설문 조사 전문 업체의 온라인 마스터 샘플 데이터	강지원 외 (2023)	같지만 다른 것들, 청년: 정별 자살생각과 자살 시도 영향 요인의 탐색 연구	자살 생각	<p>[남성] 연령&lt;(1), 배우자 유무(유배우자, 무배우자)&lt;(1), <b>교육 수준</b>(대졸 이상, 대학 재학)*1, 고졸 이하&lt;(1), 1인 가구 여부(다인가구, 1인 가구)&lt;(1), 고용 형태(정규직, 비정규직)&lt;(1), 개인사업자/프리랜서)&lt;(1), 무직)&lt;(1), 실직 경험(없음, 있음)&lt;(1), 월평균 소득(소득수준)&lt;(1), 주관적 경제 상태(좋은 편, 보통&gt;1, 나쁜 편)&lt;(1), <b>주관적 건강상태</b>(좋은 편, 보통&gt;1, 나쁜 편**)&lt;(1), <b>수면의 질</b>(좋은 편, 보통*1, 나쁜 편&lt;(1), 우울감&lt;(1), 스트레스&lt;(1), 사회 안전 인식(좋은 편, 보통&lt;(1), 나쁜 편&lt;(1)과 사회 신뢰도(높은 편, 보통&lt;(1), 나쁜 편&lt;(1), 차별 경험(없음, 있음)&lt;(1), <b>좌절된 대인관계 욕구</b>(집이 되는 느낌***)&lt;(1), 좌절된 소속감***)&lt;(1), <b>가족탄력성</b>(가족탄력성)&lt;(1, 상호 존중과 사회 참여)&lt;(1)과 어린 시절 가족관계(좋은 편, 보통&gt;1, 나쁜 편)&lt;(1), 어린 시절 학대 경험(없음, 있음)&lt;(1)</p> <p>[여성] 연령&lt;(1), 배우자 유무(유배우자, 무배우자)&lt;(1), 교육 수준(대졸 이상, 대학 재학&lt;(1), 고졸 이하&lt;(1), 1인 가구 여부(다인가구, 1인 가구)&lt;(1), 고용 형태(정규직, 비정규직)&lt;(1), 개인사업자/프리랜서)&lt;(1), 무직)&lt;(1), 실직 경험(없음, 있음)&lt;(1), 월평균 소득(소득수준)&lt;(1), 주관적 경제 상태(좋은 편, 보통&gt;1, 나쁜 편)&lt;(1), 주관적 건강 상태(좋은 편, 보통=1, 나쁜 편&lt;(1), 수면의 질(좋은 편, 보통&lt;(1), 나쁜 편)&lt;(1), <b>우울감</b>**&lt;(1), 스트레스&lt;(1), 사회 안전 인식(좋은 편, 보통&gt;1, 나쁜 편)&lt;(1)과 사회 신뢰도(높은 편, 보통&lt;(1), 나쁜 편&lt;(1), 차별 경험(없음, 있음)&lt;(1), 좌절된 대인관계 욕구(집이 되는 느낌)&lt;(1), 좌절된 소속감)&lt;(1), 가족탄력성(가족탄력성)&lt;(1, 상호 존중과 수용)&lt;(1), 질적 우대감&lt;(1), 경제적 안정과 협력)&lt;(1, 가족문화와 사회 참여)&lt;(1)과 어린 시절 가족관계(좋은 편, 보통&gt;1, 나쁜 편)&lt;(1), 어린 시절 학대 경험(없음, 있음)&lt;(1)</p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
1	청년	심문 조사 전문 업체의 온라인 모니터링 샘플 데이터	강지원 외 (2023)	다들, 성별, 자살생각과 자살 시도 영향 요인의 탐색 연구	자살 시도 경험	<p>[남성]</p> <p>연령&lt;&lt;1&gt;, 배우자 유무&lt;유배우자, 무배우자&gt;(1), 교육 수준&lt;대졸 이상, 대학 재학&gt;, 고졸 이하&gt;, 1인 가구 여부&lt;다인가구, 1인 가구&gt;(1), 고용 형태&lt;정규직, 비정규직&gt;(1), 개인사업자&lt;프리랜서&gt;(1, 무직)&gt;(1), 실직 경험&lt;없음, 있음&gt;(1), <b>월평균 소득&lt;소득수준&gt;</b>(1), 주관적 경제 상태&lt;좋은 편, 보통&gt;(1, 나쁜 편)&lt;(1), 주관적 건강 상태&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1), 수면의 질&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1), <b>우울감</b>(1), 스트레스&lt;(1), 사회 안전 인식&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1)과 사회 신뢰도&lt;높은 편, 보통&gt;(1, 나쁜 편)&gt;(1), 차별 경험&lt;없음, 있음&gt;(1), 좌절된 대인관계 욕구&lt;심어&gt;(1)가 되는 느낌&gt;(1, 좌절된 소속감&lt;(1), <b>가족탄력성&lt;(1, 상호 존중과 수용&lt;(1, 질적 우대감&gt;(1, 경제적 안정과 협력&gt;(1, 가족문화와 사회 참여&lt;(1)과 어린 시절 가족관계&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1), 어린 시절 학대 경험&lt;없음, 있음&gt;(1)</b></p> <p>[여성]</p> <p>연령&lt;&lt;1&gt;, 배우자 유무&lt;유배우자, 무배우자&gt;(1), 교육 수준&lt;대졸 이상, 대학 재학&gt;, 고졸 이하&gt;, 1인 가구 여부&lt;다인가구, 1인 가구&gt;(1), 고용 형태&lt;정규직, 비정규직&gt;(1), 개인사업자&lt;프리랜서&gt;(1, 무직)&gt;(1), <b>실직 경험&lt;없음, 있음&gt;(1), 월평균 소득&lt;소득수준&gt;(1), 주관적 경제 상태&lt;좋은 편, 보통&gt;(1, 나쁜 편)&lt;(1), 주관적 건강 상태&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1), 수면의 질&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1), <b>우울감</b>(1), 스트레스&lt;(1), 사회 안전 인식&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1)과 <b>사회 신뢰도&lt;높은 편, 보통&gt;(1, 나쁜 편)&lt;(1), 차별 경험&lt;없음, 있음&gt;(1), 좌절된 대인관계 욕구&lt;심어&gt;(1)가 되는 느낌&gt;(1, 좌절된 소속감&lt;(1), 가족간강성&lt;가족탄력성&gt;(1, 상호 존중과 수용&lt;(1, 질적 우대감&gt;(1, 경제적 안정과 협력&gt;(1, 가족문화와 사회 참여)&gt;(1)과 어린 시절 가족관계&lt;좋은 편, 보통&gt;(1, 나쁜 편)&gt;(1), 어린 시절 학대 경험&lt;없음, 있음&gt;(1)</b></b></p>



구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
2	청소년 2020년 청소년 건강행태 조사	알레르기 질환 청소년 년의 정신 건강 : 2020년 청소년 건강행태 조사 인 조사 활동	박초희 (2022)	스트레스 인지	<p><b>알레르기 질환이 있는 청소년의 정신건강에 영향을 미치는 요인</b>  <b>학교 유형(중학생, 고등학생)</b>, <b>성별(남학생, 여학생)</b>,  <b>경제적 수준(상, 중상, 중, 하)</b>, <b>학업성적(상, 중상, 중, 하)</b></p> <p><b>알레르기 질환이 있는 청소년의 정신건강에 영향을 미치는 요인</b>  <b>학교 유형(중학생, 고등학생)</b>, <b>성별(남학생, 여학생)</b>,  <b>경제적 수준(상, 중상, 중, 하)</b>, <b>학업성적(상, 중상, 중, 하)</b></p>	
3	성인 2019년 한국 의료패널 연간 데이터	만 19-39 세 성인의 성별에 따 른 자살생 각 영향 요인	임은정, 정계현 (2022)	자살 생각	<p><b>알레르기 질환이 있는 청소년의 정신건강에 영향을 미치는 요인</b>  <b>학교 유형(중학생, 고등학생)</b>, <b>성별(남학생, 여학생)</b>,  <b>경제적 수준(상, 중상, 중, 하)</b>, <b>학업성적(상, 중상, 중, 하)</b></p>	<p><b>알레르기 질환이 있는 청소년의 정신건강에 영향을 미치는 요인</b>  <b>학교 유형(중학생, 고등학생)</b>, <b>성별(남학생, 여학생)</b>,  <b>경제적 수준(상, 중상, 중, 하)</b>, <b>학업성적(상, 중상, 중, 하)</b></p> <p>[전체]          나이(19-29세&lt;1, 30-30세), 결혼상태(유&lt;1, 무), 경제활동 참여 상태(유&lt;1, 무), 만성질환 상태(유&lt;1, 무), 규칙적인 신체활동 유무(유&lt;1, 무), 현재 흡연 유무(유&lt;1, 무), 현재 음주 유무(유&lt;1, 무), 체질량지수(저체중&lt;1, 정상&lt;1, 비만 전단계&lt;1, 비만), 병의원 미충족 의뢰 경험 유무(유&lt;1, 무), 스트레스 인지 정도(낮음, 보통&lt;1, 높음), <b>주관적 건강상태(나쁨*<sup>1</sup>&lt;1, 보통&lt;1, 좋음)</b>, 우울감(유&lt;1, 무), 불안감(유&lt;1, 무), <b>정신건강 문제로 전문가 상담 유무(유*<sup>1</sup>&lt;1, 약물복용 유무(유)1, 무)</b></p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
3	성인	2019년 한국 의료패널 연간 데이터	임은정, 정계현 (2022)	만 19-39 세 성인의 성별에 따른 자살생각 요인	자살 생각	<p>[남성]                      나이(19-29세&lt;1, 30-30세), 결혼상태(유&lt;1, 무), 경제활동 참여 상태(유&lt;1, 무), 만성질환 상태(유&lt;1, 무), 규칙적인 신체활동 유무(유&lt;1, 무), 현재 흡연 유무(유&lt;1, 무), 현재 음주 유무(유&lt;1, 무), 체질양지수(저체중&lt;1, 정상&lt;1, 비만 진단계&lt;1, 비만), 병의원 미충족 의료 경험 유무(유&lt;1, 무), 스트레스 인지 정도(낮음, 보통&lt;1, 높음&lt;1), 주관적 건강상태(나쁨&lt;1, 보통&lt;1, 좋음), 우울감(유&lt;1, 무), 불안감(유*1, 무), 정신건강 문제로 전문가 상담 유무(유, 무=0), 약물복용 유무(유=0, 무)</p> <p>[여성]                      나이(19-29세&lt;1, 30-30세), 결혼상태(유&lt;1, 무), 경제활동 참여 상태(유&lt;1, 무), 만성질환 상태(유&lt;1, 무), 규칙적인 신체활동 유무(유&lt;1, 무), 현재 흡연 유무(유&lt;1, 무), 현재 음주 유무(유&lt;1, 무), 체질양지수(저체중&lt;1, 정상&lt;1, 비만 진단계=0, 비만), 병의원 미충족 의료 경험 유무(유&lt;1, 무), 스트레스 인지 정도(낮음, 보통&lt;1, 높음&lt;1), 주관적 건강상태(나쁨&lt;1, 보통&lt;1, 좋음), 우울감(유&lt;1, 무), 불안감(유&lt;1, 무), 정신건강 문제로 전문가 상담 유무(유, 무=0), 약물복용 유무(유=0, 무)</p>
4	여성	2017년 한국 의료패널	박민정 (2021)	건강행위와 건강상태가 베이비부머대 여성의 자살생각에 미치는 영향: 2017년 한국의 자료 이용	자살 생각	<p><b>배우자 유무</b>(유, 무*), <b>연간 가구 총소득</b>(높음, 중간&lt;1, 낮음&lt;1), <b>흡연</b>(예&lt;1, 아니오), <b>수면시간</b>(정상, 비정상), <b>신체활동</b>(격렬한 활동, 중등도 활동&lt;1, 걷기&lt;1), <b>사력문제</b>(유&lt;1, 무), <b>심각문제</b>(유*1, 무), <b>활동제한</b>(유*1, 무), <b>스트레스</b>, <b>좌절감</b>, <b>불안</b>, <b>우울</b>(유*1, 무)</p>

구분	정책 대상	데이티명	저자	연구명	종속변수	독립변수
5	노인	2018년 한국 의료패널	성혜연, 이성규, 나지훈 (2021)	노인의 자살생각에 영향을 미치는 요인들에 관한 연구: 객관적 및 주관적 건강상태에 따른 진단별 비교를 중심으로	자살 생각	<p>[모델 1] <b>객관적 건강상태</b> = 나쁨 and <b>주관적 건강상태</b> = 나쁨  <b>연령</b>(&lt;1), <b>성별</b>(남자*&lt;1, 여자), <b>학력</b>(&lt;1), 결혼상태(기타 결혼상태, 혼인 중&lt;1), 정부보조금 수급 여부(비수급, 수급&lt;1), 미충족 의료 경험(없음, 있음&lt;1), 장애 유무(없음, 있음&lt;1), <b>불안-우울</b>(없음, 있음**&lt;1), 문제음주 경험(없음, 있음&lt;1), <b>스트레스</b>**(&lt;1)</p> <p>[모델 2] <b>객관적 건강상태</b> = 좋음 and <b>주관적 건강상태</b> = 나쁨  <b>연령</b>(&lt;1), <b>성별</b>(남자&lt;1, 여자), <b>학력</b>(&lt;1), 결혼상태(기타 결혼상태, 혼인 중&lt;1), 정부보조금 수급 여부(비수급, 수급&lt;1), 미충족 의료 경험(없음, 있음&lt;1), 장애 유무(없음, 있음&lt;1), <b>불안-우울</b>(없음, 있음**&lt;1), 문제음주 경험(없음, 있음&lt;1), <b>스트레스</b>(&lt;1)</p> <p>[모델 3] <b>객관적 건강상태</b> = 나쁨 and <b>주관적 건강상태</b> = 좋음  <b>연령</b>(&lt;1), <b>성별</b>(남자&lt;1, 여자), <b>학력</b>(&lt;1), 결혼상태(기타 결혼상태, 혼인 중&lt;1), 정부보조금 수급 여부(비수급, 수급&lt;1), <b>미충족 의료 경험</b>(없음, 있음**&lt;1), 장애 유무(없음, 있음&lt;1), <b>불안-우울</b>(없음, 있음**&lt;1), 문제음주 경험(없음, 있음&lt;1), <b>스트레스</b>**(&lt;1)</p> <p>[모델 4] <b>객관적 건강상태</b> = 좋음 and <b>주관적 건강상태</b> = 좋음  <b>연령</b>(&lt;1), <b>성별</b>(남자&lt;1, 여자), <b>학력</b>(&lt;1), 결혼상태(기타 결혼상태, 혼인 중&lt;1), 정부보조금 수급 여부(비수급, 수급&lt;1), 미충족 의료 경험(없음, 있음&lt;1), 장애 유무(없음, 있음&lt;1), <b>불안-우울</b>(없음, 있음**&lt;1), 문제음주 경험(없음, 있음**&lt;1), <b>스트레스</b>**(&lt;1)</p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
6	성인	2016년 한국 의료패널	방소연 (2021)	한국 성인의 생애주기별 자살 생각 영향 요인	자살 생각	<p>[청·장년기] 배우자 유무(유&gt;1, 무), 주관적 건강 상태(좋은, 보통&gt;1, 나쁨&gt;1), 스트레스(&lt;1), 불안***(&lt;1), 우울(예, 아니오***&gt;1), 약물 복용(예***&gt;1, 아니오), 삶의 질**(&lt;1)</p> <p>[중년기] 주관적 건강 상태(좋은, 보통&gt;1, 나쁨&gt;1), 스트레스***(&lt;1), 불안***(&lt;1), 우울(예, 아니오***&gt;1), 약물 복용***(&lt;1), 수면 시간(&lt;1), 삶의 질**(&lt;1)</p> <p>[노년기] 가구 소득(상, 중상&gt;1, 중&gt;1, 중하&gt;1, 하&gt;1), 체질량지수(비만, 과체중&lt;1, 정상**&lt;1, 저체중*(&lt;1), 주관적 건강 상태(좋은, 보통&gt;1, 나쁨&gt;1), 스트레스***(&lt;1), 불안***(&lt;1), 우울(예, 아니오***&gt;1), 약물 복용***(&lt;1), 수면 시간(&lt;1), 삶의 질**(&lt;1)</p>
7	여성	2016년 한국 의료패널	김영희 (2020)	중년 여성들의 신체활동과 우울 및 자살생각에 대한 융합적 연구	우울  자살 사고	<p>나이(40~50, 51~60**&gt;1), 소득(낮음, 높음*(&lt;1), 결혼상태(배우자 있음, 배우자 없음**&gt;1), 정신·신체적 스트레스(없음, 있음**&gt;1), 자살 사고(아니오, 예*(&gt;1), 걷기(아니오, 예)(&lt;1), 중등도 신체활동(아니오, 예&lt;1), 격렬한 신체활동(아니오, 예&gt;1)</p> <p>소득(낮음, 높음*(&lt;1), 흡연(아니오, 예*(&gt;1), 정신·신체적 스트레스(없음, 있음**&gt;1), 우울(아니오, 예**(&gt;1), 걷기(아니오, 예)(&lt;1), 중등도 신체활동(아니오, 예&lt;1), 격렬한 신체활동(아니오, 예&gt;1)</p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
8	노인	2016년 한국 의료패널	방소연 임은정 (2020)	한국 노인 체기능 상 태에 따른 정신건강 에 미치는 영향에서 성별 차 이: 2016 년 한국의 노년 자 료패널 자 료를 바탕 으로	우울	[남성] 신체기능 상태(ADL, IADL) [여성] 신체기능 상태(ADL, IADL) * 독립성 검정(카이제곱 검정 또는 T 검정) 실시
					불안	[남성] 신체기능 상태(ADL, IADL) [여성] 신체기능 상태(ADL, IADL) * 독립성 검정(카이제곱 검정 또는 T 검정) 실시
					스트레스	[남성] 신체기능 상태(ADL, IADL) [여성] 신체기능 상태(ADL, IADL) * 독립성 검정(카이제곱 검정 또는 T 검정) 실시
					자살충동	[남성] 신체기능 상태(ADL, IADL) [여성] 신체기능 상태(ADL, IADL) * 독립성 검정(카이제곱 검정 또는 T 검정) 실시
9	노인	2016년 한국 복지패널	김현주 (2019)	한국 노인 의 자살생 각에 관한 응답률 에 측 요 인 : 패 널 자 료 분석	자살 생각	성별, 공적 복지서비스 수급(국민기초생활보장 수급 여부와 의료급여 수급 여부 및 1, 2종 여부, 노인장기요양보험 수급 여부), <b>주관적 건강상태</b> *(=0.767<1), <b>우울</b> ****(=1.113>1), 자아존중감

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
10	저소득층	2012년 ~2017년 한국 복지패널	김미숙 김운영 (2018)	물질적 박탈과 은둔형 고립의 위험에 미치는 영향- 한국 복지패널 데이터	자살 생각	연령***<0, 성별***>, 교육 수준***<0, 종교>, 결혼 상태***<0, 음식박탈***<0, 주거박탈***<0, 의료박탈***<0, 교육박탈***<0, 신용박탈***<0, 음주빈도>, 평균 음주량<0, 과음빈도>, 문제음주 경험1<0, 문제음주 경험2>, 문제음주 경험3***<0, 문제음주 경험4***>, 문제음주 경험5***>
					자살 계획	연령>, 성별***>, 교육 수준<0, 종교* >, 결혼 상태>, 음식박탈***<0, 주거박탈>, 의료박탈***<0, 교육박탈>, 신용박탈***<0, 음주빈도>, 평균 음주량<0, 과음빈도>, 문제음주 경험1>, 문제음주 경험2>, 문제음주 경험3<0, 문제음주 경험4***>, 문제음주 경험5***>
					자살 시도	연령* >, 성별***>, 교육 수준<0, 종교>, 결혼 상태>, 음식박탈***<0, 주거박탈* <0, 의료박탈***<0, 교육박탈>, 신용박탈***<0, 음주빈도>, 평균 음주량>, 과음빈도>, 문제음주 경험1>, 문제음주 경험2>, 문제음주 경험3* >, 문제음주 경험4>, 문제음주 경험5>
11	성인	2009년 ~2013년 한국 의료패널	이종익 (2018)	성인의 체질량지수(BMI)와 자살생각의 관계 = 2009 ~ 2013년 한국의료패널 자료를 활용한 연구	자살 생각	<모델1> 성별>, 연령***>, 교육 수준***<0, 소득수준***<0, 의료비<0, 만성질환 수><0, 우울감***>>  <모델2> 성별***>, 연령***>, 교육 수준***<0, 소득수준***<0, 의료비<0, 체질량지수***>, 만성질환 수>
						<모델3> 성별>, 연령***>, 교육 수준***<0, 소득수준***<0, 의료비<0, 체질량지수***>, 만성질환 수<0, 우울감***>>

\* 3개 모델 모두 로지스틱 회귀 분석 실시

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
12	노인	2014년 한국 복지패널	강동훈 김윤태 (2018)	사회경제적 박탈이 노인의 자살 생각에 미치는 영향 연구명: 6가지 박탈 유형 을 중심으로	자살 생각	<p>1) 사회경제적 박탈이 노인의 자살 생각에 미치는 영향 성별&lt;1, 연령&gt;1, 결혼 상태&lt;1, 교육 수준&lt;1, 소득&lt;1, 거주 지역&lt;1, 사회경제적 박탈*&gt;1</p> <p>2) 식생활 박탈이 노인의 자살 생각에 미치는 영향 성별&lt;1, 연령&lt;1, 결혼 상태&lt;1, 교육 수준&lt;1, 소득&lt;1, 거주 지역&lt;1, 식생활 박탈1***&gt;1, 식생활 박탈2***&gt;1</p> <p>3) 주거 박탈이 노인의 자살 생각에 미치는 영향 성별&lt;1, 연령&lt;1, 결혼 상태&lt;1, 교육 수준&lt;1, 소득&lt;1, 거주 지역&lt;1, 주거 박탈1*&gt;1, 주거 박탈2&lt;1</p> <p>4) 사회보장 박탈이 노인의 자살 생각에 미치는 영향 성별&lt;1, 연령&lt;1, 결혼 상태&lt;1, 교육 수준&lt;1, 소득&lt;1, 거주 지역&lt;1, 사회보장 박탈1&gt;1, 사회보장 박탈2&gt;1</p> <p>5) 사회적 박탈이 노인의 자살 생각에 미치는 영향 성별&lt;1, 연령&lt;1, 결혼 상태&lt;1, 교육 수준&lt;1, 소득&lt;1, 거주 지역&lt;1, 사회적 박탈1&gt;1, 사회적 박탈2**&gt;1</p> <p>6) 건강 박탈이 노인의 자살 생각에 미치는 영향 성별&lt;1, 연령&lt;1, 결혼 상태&lt;1, 교육 수준&lt;1, 소득&lt;1, 거주 지역&lt;1, 건강 박탈***&gt;1</p> <p>7) 직업 박탈이 노인의 자살 생각에 미치는 영향 성별&lt;1, 연령&lt;1, 결혼 상태&lt;1, 교육 수준&lt;1, 소득&lt;1, 거주 지역&lt;1, 직업 박탈***&gt;1</p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
13	여성	2010년 ~2012년 한국 의료패널	최세문 이혜영 (2017)	자녀의 아 토피 피부 염 유병이 주양육자 의 정신건강에 미치는 영향	우울감	<p>[2010년] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)* &gt; 1, 0~7세** &lt; 1, 8~12세*** &gt; 1, 13~18세*** &lt; 1)</b></p> <p>[2011년] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)* &lt; 1, 0~7세** &lt; 1, 8~12세*** &gt; 1, 13~18세*** &lt; 1)</b></p> <p>[2012년] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)* &gt; 1, 0~7세** &lt; 1, 8~12세*** &gt; 1, 13~18세*** &gt; 1)</b></p> <p>[패널 2010~2012] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)* &gt; 1, 0~7세** &lt; 1, 8~12세*** &gt; 1, 13~18세*** &gt; 1)</b></p> <p>[패널 2010~2012, 가구소득 하위 40%] 연령, 혼인관계 유지 여부, 학력수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)* &gt; 1, 0~7세** &lt; 1, 8~12세*** &gt; 1, 13~18세*** &lt; 1)</b></p> <p>[패널 2010~2012, 가구소득 상위 60%] 연령, 혼인관계 유지 여부, 학력수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)* &gt; 1, 0~7세** &lt; 1, 8~12세*** &gt; 1, 13~18세*** &gt; 1)</b></p>



구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
13	여성	2010년 ~2012년 한국 의료패널	최세문 이혜영 (2017)	자녀의 아 토피 피부 염 유병이 주 양육 자 의 정신진 강에 미치 는 영향	자살 생각 경험	<p>[2010년] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)**&gt;1, 0~7세**&gt;1, 8~12세**&gt;1, 13~18세**&gt;1)</b></p> <p>[2011년] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)**&gt;1, 0~7세**&gt;1, 8~12세**&gt;1, 13~18세**&gt;1)</b></p> <p>[2012년] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)**&gt;1, 0~7세**&gt;1, 8~12세**&gt;1, 13~18세**&gt;1)</b></p> <p>[패널 2010~2012] 연령, 혼인관계 유지 여부, 학력수준, 가구 소득수준, 직업, 아토피 피부염 유병 소아청소년(0~18세)&gt;1, 0~7세&gt;1, 8~12세&gt;1, 13~18세&gt;1)</p> <p>[패널 2010~2012, 가구소득 하위 40%] 연령, 혼인관계 유지 여부, 학력수준, 직업, 아토피 피부염 유병 소아청소년(0~18세)&gt;1, 0~7세&gt;1, 8~12세&lt;1, 13~18세&gt;1)</p> <p>[패널 2010~2012, 가구소득 상위 60%] 연령, 혼인관계 유지 여부, 학력수준, 직업, <b>아토피 피부염 유병 소아청소년(0~18세)&gt;1, 0~7세&gt;1, 8~12세&gt;1, 13~18세**&gt;1)</b></p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
14	성인	2011년 정신질환 실태 역학조사	안준석 외 (2016)	불안 장애와 주요 우울증에 대한 연구	자살 사고	전체 증상군, 통증 증상군, 소화가 증상군, 가성신경증상군 * 독립성 검정(카이제곱 검정 또는 T 검정) 실시
				신체 증상군에 대한 연구	자살 계획	전체 증상군, <b>통증 증상군*</b> , 소화기 증상군, 가성신경증상군 * 독립성 검정(카이제곱 검정 또는 T 검정) 실시
				행동 계획에 대한 연구	자살 시도	전체 증상군, 통증 증상군, <b>소화가 증상군*</b> , <b>가성신경증상군*</b> * 독립성 검정(카이제곱 검정 또는 T 검정) 실시
15	노인	2012년 한국 의료패널	남석인, 이희정, 김성준, 김보경 (2016)	노인의 계층별 인지 기능 저하에 미치는 영향과 스트레스와 주관적 건강 인식의 연관성	자살 충동	성별>0, 연령>0, 교육 수준>0, 혼인상태>0, 만성질환 여부<0, 개인별 연간 근로소득<0, 의료보장형태<0, <b>주관적 계층인식***</b> <0

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
16	노인	2010년 한국 의료패널	박민정 (2015)	노인의 성별에 따른 자살생각 과 영향요 인 - 2010년도한국 의료패널 자료를 이 용하여	자살 생각	<p>[남성] 교육 수준&lt;초졸 이하, 중졸&lt;1, 고졸&lt;1, 대졸 이상&lt;1&gt;, 경제활동 여부&lt;예, 아니오&gt;1, 우울감 유무&lt;없음, 있음***&gt;1, 정신과 약 투약 여부&lt;아니오, 예&gt;1, 경제적 스트레스&lt;없음, 있음&gt;1, 질병 스트레스&lt;없음, 있음&gt;1, 가족 간 갈등 스트레스&lt;없음, 있음&gt;1, 사고 스트레스&lt;없음, 있음&gt;1, 주관적 건강상태&lt;좋다, 보통이다&lt;1, 나쁘다*&gt;1, 활동 제한 여부&lt;없음, 있음&gt;1, 흡연 여부&lt;아니오, 예&gt;1, 음주 여부&lt;아니오, 예&lt;1&gt;, 운동 여부&lt;예, 아니오&gt;1, 규칙적인 식사 여부&lt;예, 아니오*&gt;1)</p> <p>[여성] 배우자 유무&lt;있음, 없음&gt;1, 우울감 유무&lt;없음, 있음***&gt;1, 정신과 약 투약 여부&lt;아니오, 예&gt;1, 경제적 스트레스&lt;없음, 있음&gt;1, 질병 스트레스&lt;없음, 있음&gt;1, 가족 간 갈등 스트레스&lt;없음, 있음&gt;1, 사고 스트레스&lt;없음, 있음&gt;1, 주관적 건강상태&lt;좋다, 보통이다&lt;1, 나쁘다&gt;1, 활동 제한 여부&lt;없음, 있음&gt;1, 만성질환 유무&lt;없음, 있음&gt;1, 음주 여부&lt;아니오, 예&gt;1, 운동 여부&lt;예, 아니오&gt;1, 규칙적인 식사 여부&lt;예, 아니오&gt;1)</p>
17	노인	2011년 한국 의료패널	박재우 황병덕 (2015)	신체 활동 특성이 노인 만성질환자의 정신 건강에 미치는 영향	우울감 경험	<p>[모텔] 주관적 건강상태&lt;매우 좋음, 좋음***&gt;1, 좋지 않음***&gt;1, 흡연경험 유무&lt;아니오, 예&gt;1, 파음 유무&lt;아니오, 예&gt;1, 신체활동 제한 유무&lt;아니오, 예&gt;1, 건강신체 활동 도달 유무&lt;아니오, 예*&gt;1, 동반상병 지수&lt;0점, 1점&gt;1, 2점&gt;1, 3점 이상&gt;1)</p> <p>[모텔2] 성별&lt;남, 여*&gt;1, 노인 분류&lt;55~65세&lt;1, 65~75세&lt;1, 75세 이상&gt;1&gt;, 교육 수준&lt;중졸 이하, 고졸&gt;1, 대졸 이상&lt;1&gt;, 경제활동 유무&lt;아니오, 예*&gt;1, 혼인 상태&lt;앞의 중, 혼인 중 아님&gt;1, 주관적 건강상태&lt;매우 좋음, 좋음***&gt;1, 좋지 않음***&gt;1, 흡연경험 유무&lt;아니오, 예&gt;1, 파음 유무&lt;아니오, 예&gt;1, 신체 활동 제한 유무&lt;아니오, 예&gt;1, 건강신체 활동 도달 유무&lt;아니오, 예*&gt;1, 동반상병 지수&lt;0점, 1점&gt;1, 2점&gt;1, 3점 이상&gt;1)</p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
17	노인	2011년 한국 의료패널	박재우 황병덕 (2015)	신체 활동 특성이 노인 만성질환 환자의 정신 건강에 미치는 영향	자살충동 경험	<p>[모델1]  <b>주관적 건강상태</b>(매우 좋음, 좋음<sup>***</sup>&gt;1, 좋지 않음<sup>***</sup>&gt;1), 흡연경험 유무(아니오, 예&lt;1), 과음 유무(아니오, 예&gt;1), 신체활동 제한 유무(아니오, 예&gt;1), <b>건강신체 활동 도달 유무</b>(아니오, 예&lt;1), 동반상병 지수(0점, 1점&gt;1, 2점&gt;1, 3점 이상&gt;1)</p> <p>[모델2]  <b>성별</b>(남, 여*1), 노인 분류(55~65세, 65~75세&lt;1, 75세 이상&lt;1), 교육 수 준(중졸 이하, 고졸&lt;1, 대졸 이상&lt;1), <b>경제활동 유무</b>(아니오, 예*1), 혼인상태(혼인 중, 혼인 중 아님&gt;1), <b>주관적 건강상태</b>(매우 좋음, 좋음<sup>***</sup>&gt;1, 좋지 않음<sup>***</sup>&gt;1), 흡연경험 유무(아니오, 예&gt;1), <b>과음 유무</b>(아니오, 예*1), 신체활동 제한 유무(아니오, 예&gt;1), <b>건강신체 활동 도달 유무</b>(아니오, 예*1), 동반상병 지수(0점, 1점&gt;1, 2점&gt;1, 3점 이상&gt;1)</p>
18	청소년	2012년 한국 아동-청소년 건강 실태조사	민다경 이하나 송인한 (2014)	인터넷 게임 과몰입이 청소년의 자살생각에 미치는 영향: 우울감의 매개효과 분석	자살 생각	<b>성별</b> ***<0, 경제수준>0, <b>학년</b> ***<0, <b>학업 성적</b> *>0, <b>스트레스 정도</b> ***>0, <b>부모지원</b> ***<0, 친구관계>0, 교사관계<0, 인터넷 게임 과몰입>0, <b>우울감</b> ***>0

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
19	노인	2012년 국민건강영양조사	손신영 (2014)	우리나라 노인의 자살생각에 영향을 미치는 요인에 관한 연구	자살 생각	<p>[모델 1] 일반적 특성  <b>성별</b>(남성, 여성)*1), <b>학력</b>(초졸 이하)*1), <b>중졸</b>&gt;1, <b>고졸</b>&gt;1, <b>초대졸 이상</b>, <b>소득</b>(1분위)*1, 2분위)*1, 3분위)*1, 4분위), <b>경제활동 여부</b>(예, 아니오)*1)          [모델 2] 일반적 특성 + 건강행태 특성  <b>일반적 특성</b>: <b>성별</b>(남성, 여성)*1), <b>학력</b>(초졸 이하)*1), <b>중졸</b>&gt;1, <b>고졸</b>&gt;1, <b>초대졸 이상</b>, <b>소득</b>(1분위)*1, 2분위)*1, 3분위)*1, 4분위), <b>경제활동 여부</b>(예, 아니오)*1), <b>흡연</b>(예)*1, 아니오), <b>음주</b>(예)*1, 아니오), <b>중등도 신체활동</b>(안함)*1, 2일 이내)*1, 3일 이상), <b>지각된 스트레스</b>(낮음, 보통)*1, 높음)*1)          [모델 3] 일반적 특성 + 건강 관련 특성  <b>성별</b>(남성, 여성)*1), <b>학력</b>(초졸 이하)*1), <b>중졸</b>&gt;1, <b>고졸</b>&gt;1, <b>초대졸 이상</b>, <b>소득</b>(1분위)*1, 2분위)*1, 3분위)*1, 4분위), <b>경제활동 여부</b>(예, 아니오)*1), <b>주관적 건강상태</b>(낮음)*1, 보통)*1, 높음), <b>활동제한 여부</b>(예)*1, 아니오), <b>질한 및 사고중독으로 인해 몸이 아프거나 불편했던 경험</b>(예)*1, 아니오), <b>우울</b>(예)*1, 아니오)          [모델 4] 일반적 특성 + 건강행태 특성 + 건강 관련 특성  <b>일반적 특성</b>: <b>성별</b>(남성, 여성)*1), <b>학력</b>(초졸 이하)*1), <b>중졸</b>&gt;1, <b>고졸</b>&gt;1, <b>초대졸 이상</b>, <b>소득</b>(1분위)*1, 2분위)*1, 3분위)*1, 4분위), <b>경제활동 여부</b>(예, 아니오)*1), <b>건강행태 특성</b>: <b>흡연</b>(예)*1, 아니오), <b>음주</b>(예)*1, 아니오), <b>중등도 신체활동</b>(안 함)*1, 2일 이내)*1, 3일 이상), <b>지각된 스트레스</b>(낮음, 보통)*1, 높음)*1), <b>주관적 건강상태</b>(낮음)*1, 보통)*1, 높음), <b>활동제한 여부</b>(예)*1, 아니오), <b>질한 및 사고중독으로 인해 몸이 아프거나 불편했던 경험</b>(예)*1, 아니오), <b>우울</b>(예)*1, 아니오)</p>

구분	정책 대상	데이터명	저자	연구명	종속변수	독립변수
20	여성	2007년 국민건강영양조사	김영택 김인순 주영수 이지연 (2009)	여성 자살 현황 및 정책방안	자살 생각	<p>[모형1] 성별(남성, 여성) &gt; 1, 연령 &gt; 1, 결혼 상태(결혼, 별거, 이혼, 사별) &gt; 1, 미혼 &gt; 1, 소득(4분위, 3분위) &gt; 1, 2분위 &gt; 1, 1분위 &gt; 1, 직업(정규 규직, 비정규직) &gt; 1, 자영업자-무급가족종사자 &gt; 1, 주관적 건강상태(매우  좋음, 좋음 &lt; 1, 보통 &gt; 1, 나쁨 &gt; 1, 매우 나쁨 &gt; 1), 질병이환(아니오, 예  &gt; 1)</p> <p>[모형2] 성별(남성, 여성) &gt; 1, 연령 &gt; 1, 결혼 상태(결혼, 별거, 이혼, 사별) &gt; 1, 미 혼 &gt; 1, 소득(4분위, 3분위) &gt; 1, 2분위 &gt; 1, 1분위 &gt; 1, 직업(정규 규직, 비정규직) &gt; 1, 자영업자-무급가족종사자 &gt; 1, 주관적 건강상태(매우  좋음, 좋음 &lt; 1, 보통 &gt; 1, 나쁨 &gt; 1, 매우 나쁨 &gt; 1), 질병이환(아니오, 예  &gt; 1)</p> <p>[모형3] 성별(남성, 여성) &gt; 1, 연령 &gt; 1, 결혼 상태(결혼, 별거, 이혼, 사별) &gt; 1, 미 혼 &gt; 1, 소득(4분위, 3분위) &gt; 1, 2분위 &gt; 1, 1분위 &gt; 1, 직업(정규 규직, 비정규직) &gt; 1, 자영업자-무급가족종사자 &gt; 1, 주관적 건강상태(매우  좋음, 좋음 &lt; 1, 보통 &gt; 1, 나쁨 &gt; 1, 매우 나쁨 &gt; 1), 질병이환(아니오,  예 &gt; 1), 스트레스 정도(거의 느끼지 않는다, 조금 느끼는 편이다 &gt; 1, 많이  느끼는 편이다 &gt; 1, 대단히 많이 느낀다 &gt; 1), 2주 연속 우울감 여부(아니  오, 예 &gt; 1)</p>
					자살 시도	응답자가 층분치 않아 다변량 분석 생각

주1: \* p<.05, \*\* p<.01, \*\*\* p<.001

2: 독립성 검정 결과의 경우 통계적 유의수준만을 표기하였음.

3: 로지스틱 회귀 분석 결과의 경우 오즈비값이 1보다 크거나 < 1 작거나 < 1) 작거나 < 0) 값을 (=)을 표기하거나 회귀 계수가 0보다 크거나 < 0) 작음(< 0)을 표기함.

4: 페널 분석 결과의 경우 회귀 계수가 0보다 크거나 < 0) 작음(< 0)을 표기함.

### 제3절 소결

이 장에서는 예측 모델 적용을 위해 활용 가능한 보건복지분야의 공공데이터를 전반적으로 살펴보고 자살 고위험군과 관련된 실증 분석 연구 및 해당 선행연구에서의 분석 변수들을 살펴보았다.

이 장에서 검토한 공공데이터는 건강 및 정신건강 분야의 대표적인 공공데이터인 ‘정신건강실태조사’와 ‘국민건강영양조사’, 정책 대상 중 아동·청소년에 대한 특성 유형화를 위한 ‘청소년건강행태조사’, 그리고 중단적 연구가 가능한 보건복지분야의 대표적인 패널 데이터인 ‘한국복지패널’과 ‘한국의료패널’이다. 5개 공공데이터의 개요와 정신건강 관련 문항(자살 생각/계획/시도) 현황 및 연도별 현황을 검토하고 응답 빈도 수 등을 정리하여 살펴보았다.

그리고 5개 공공데이터명과 ‘자살’ 키워드가 초록에 포함되어 있는 KCI 등재 학술지를 검토하여 분석 변수와 분석 결과를 표로 정리하여 제시하였다. 선행연구 검토 결과, 통계적으로 유의한 변수가 다양하게 나타났다으며, 동일한 데이터를 활용한 연구라도 모형 설정에 따라 통계적으로 유의한 변수가 다르게 나타나고 회귀 계수의 부호도 다르게 나타났다. 선행연구를 살펴보면, 종속변수는 자살 생각·계획·시도 등의 연속형 또는 변수형 변수를 사용하였다. 그리고 독립변수의 경우 특성별 분류에 따라 그룹화할 수 있는데, 특성별 분류는 1) (인구·사회·경제적) 일반적 특성, 2) 건강행태 특성, 3) 신체건강 특성, 4) 정신건강 특성, 5) 심리·사회적 특성, 6) 가족 특성, 그리고 7) 기타 미분류로 정리하였다.

검토한 5개의 공공데이터 중에서 문항 수, 응답 빈도 수(표본 수) 및 시계열적 연속성 등을 고려하여 예측 모델 적용 데이터 및 심층 분석 데이터는 ‘정신건강실태조사’, ‘청소년건강행태조사’ 그리고 ‘한국복지패널’

로 선정하였다.

이 장에서 검토한 내용을 바탕으로 제4장과 제5장, 제6장에서 실제 예측 모델을 적용하여 그 결과를 제시하였다.





## 제4장

### 공공데이터를 활용한 정책 대상별 예측 모델

제1절 청소년건강행태조사를 활용한 예측 모델

제2절 정신건강실태조사를 활용한 예측 모델

제3절 소결



## 제 4 장 공공데이터를 활용한 정책 대상별 예측 모델

### 제1절 청소년건강행태조사를 활용한 예측 모델

#### 1. 청소년건강행태조사의 분석 개요

앞에서 살펴본 공공데이터 중 아동 대상 조사인 청소년건강행태조사(2021)를 활용하여 성별, 학교 구분별로 예측 모델을 구축하였다. 청소년 건강행태조사 2021의 전체 분석 대상자는 26,396명으로 가중치를 적용하지 않은 수치이다.

최근 12개월 동안 자살 생각을 한 적이 있느냐는 질문에 성별, 만 나이별 빈도는 다음과 같다. 자살 생각은 성별에 따라서, 연령에 따라서 빈도가 다르게 나타남을 알 수 있다.

본 데이터에서 예측 모형에 활용한 변수는 <표 4-2>에 정리하였다.

이 절에서 검토한 예측 모형은 Logistic regression, Elastic net, Decision Tree, Boosting 방법이다. 이 네 가지 모형을 적용하여 모형별로 유의미한 요인을 찾아보고자 한다.

〈표 4-1〉 청소년건강행태조사 2021의 데이터 정책 대상별 기초통계

성별			최근 12개월 동안 자살 생각		전체		
			1	2			
남자	만 나이	12	빈도	1167	127	1294	
			만 나이 중 %	90.2%	9.8%	100.0%	
		13	빈도	4668	473	5141	
			만 나이 중 %	90.8%	9.2%	100.0%	
		14	빈도	4709	528	5237	
			만 나이 중 %	89.9%	10.1%	100.0%	
		15	빈도	4445	472	4917	
			만 나이 중 %	90.4%	9.6%	100.0%	
		16	빈도	4052	376	4428	
			만 나이 중 %	91.5%	8.5%	100.0%	
		17	빈도	4029	387	4416	
			만 나이 중 %	91.2%	8.8%	100.0%	
		18	빈도	2622	261	2883	
			만 나이 중 %	90.9%	9.1%	100.0%	
		전체		빈도	25692	2624	28316
				만 나이 중 %	90.7%	9.3%	100.0%
	여자	만 나이	12	빈도	969	189	1158
				만 나이 중 %	83.7%	16.3%	100.0%
13			빈도	4096	838	4934	
			만 나이 중 %	83.0%	17.0%	100.0%	
14			빈도	4007	854	4861	
			만 나이 중 %	82.4%	17.6%	100.0%	
15			빈도	3724	742	4466	
			만 나이 중 %	83.4%	16.6%	100.0%	
16			빈도	3398	624	4022	
			만 나이 중 %	84.5%	15.5%	100.0%	
17			빈도	3482	657	4139	
			만 나이 중 %	84.1%	15.9%	100.0%	
18			빈도	2415	401	2816	
			만 나이 중 %	85.8%	14.2%	100.0%	
전체			빈도	22091	4305	26396	
			만 나이 중 %	83.7%	16.3%	100.0%	

〈표 4-2〉 청소년건강행태조사 2021의 예측 모형 적용 변수

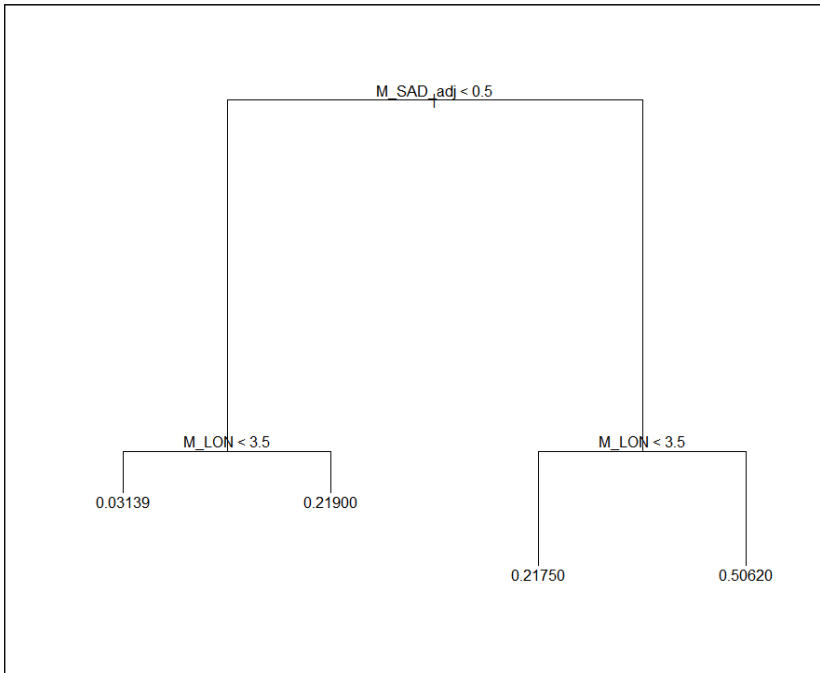
	변수명	변수 설명	
1	지역 (CTYPE)	CTYPE_adj	대도시, 중소도시 1, 군지역 0
2	주관적 건강 상태 (PR_H)	PR_HT	주관적 건강 인지 (1 매우 건강, 5 매우 건강하지 못함)
3	주관적 체형 상태 (PR_BI)	PR_BI	주관적 체형 인지 (1 마른 편, 5 매우 살찐 편)
4	고강도 신체활동 일수 (PA_VIG)	PA_VIG	고강도 신체활동 일수
5	주중 평균 학습 시간 (PA_SWD_S)	PA_SWD_S_adj	학습할 목적으로 주중에 앉아서 보낸 평균 시간(분)
6	주말 평균 학습 시간 (PA_SWK_S)	PA_SWK_S_adj	학습할 목적으로 주말에 앉아서 보낸 평균 시간(분)
7	주중 평균 학습 외 시간 (PA_SWD_N)	PA_SWD_N_adj	학습할 목적 외 주중에 앉아서 보낸 평균 시간(분)
8	주말 평균 학습 외 시간 (PA_SWK_N)	PA_SWK_N_adj	학습할 목적 외 주말에 앉아서 보낸 평균 시간(분)
9	스트레스 (M_STR)	M_STR	평상시 스트레스 인지 (1 대단히 많이 느낀다, 5 전혀 느끼지 않는다)
10	잠 피로 회복 (M_SLP_EN)	M_SLP_EN	잠으로 피로 회복을 한 정도 (1 매우 충분, 5 전혀 충분하지 않다)
11	슬픔 경험 여부 (M_SAD)	M_SAD_adj	슬픔, 절망감 경험 있음 1, 없음 0
12	외로움 경험 (M_LON)	M_LON	외로움 경험 (1 전혀 느끼지 못한다, 5 항상 외로움)
13	손상 경험 여부 (I_SCH)	I_SCH_adj	학교에서 손상 경험 있음 1, 없음 0
14	손상 병원 치료 여부 (I_SCH_TRT)	I_SCH_TRT_adj	학교에서 손상으로 병원치료 경험 있음 1, 없음 0
15	폭력 치료 경험 (V_TRT)	V_TRT	최근 1년 폭력으로 인한 치료 경험(연속형)
16	음주 경험 여부 (AC_LT)	AC_LT_adj	평생 음주 경험 있음 1, 없음 0
17	타인 음주 폐해 여부	AC_SND_adj	최근 1년 타인 음주로 인한

	변수명	변수 설명	
	(AC_SND)		폐해 있음 1, 없음 0
18	흡연 경험 여부 (TC_LT_adj)	TC_LT_adj	평생 흡연 경험 있음 1, 없음 0
19	친구 흡연 (TC_CSBF)	TC_CSBF	친한 친구 현재 흡연
20	성관계 경험 여부 (S_SI_adj)	S_SI_adj	성관계 경험 있음 1, 없음 0
21	치료 목적 외 약물 경험 여부 (DR_HAB_adj)	DR_HAB_adj	치료 목적 제외 습관적 약물 경험 있음 1, 없음 0
22	천식 여부 (AS_DG_LT_adj)	AS_DG_LT_adj	천식 평생 의사진단 있음 1, 없음 0
23	알레르기비염 여부 (RH_DG_LT_adj)	RH_DG_LT_adj	알레르기비염 평생 의사진단 있음 1, 없음 0
24	아토피피부염 여부 (ECZ_DG_LT_adj)	ECZ_DG_LT_adj	아토피피부염 평생 의사진단 있음 1, 없음 0
25	주중 스마트폰 평균 사용 시간 (INT_SPWD_TM_adj)	INT_SPWD_TM_a dj	스마트폰 주중 평균 사용시간(분)
26	주말 스마트폰 평균 사용 시간 INT_SPWK_TM_adj	INT_SPWK_TM_ad j	스마트폰 주말 평균 사용시간(분)
27	학업 성적 (E_S_RCRD)	E_S_RCRD	학업성적(1상, 5 하)
28	경제 상태 (E_SES)	E_SES	경제상태(1상 5 하)
29	코로나19 이후 경제 상태 (E_COVID19)	E_COVID19	코로나19 이후 경제적 상태 변화 (1 매우 그렇다, 5 매우 그렇지 않 다)
30	가족 동거 여부 (E_RES_adj)	E_RES_adj	거주형태, 가족과 함께 거주 1, 그 외 0
31	경제적 도움 여부 (E_AID_adj)	E_AID_adj	가정형편으로 인한 경제적 도움 경험 있음 1, 그 외 0
32	학년 (GRADE_adj)	GRADE_adj	grade 1, 2, 3학년
33	고등학교 유형 (SCHOOL_adj)	SCHOOL_adj	고등학생 활용 변수: 특성화계고 1, 일반계고 0
34	자살 생각 여부 (M_SUI_CON)	Y	최근 1년간 자살 생각 있음 1, 없음0

## 2. 남자 중학생 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 대도시, 중소도시  
에 살수록, 주관적 건강 인지가 건강하지 못하다고 생각할수록, 평상시  
스트레스 인지가 높을수록, 잠을 잔 시간이 피로 회복에 충분하지 않다고  
생각할수록, 슬픔·절망 경험이 있을수록, 외로움을 많이 느낄수록, 최근  
폭력으로 인한 치료 경험이 있을수록, 흡연 경험이 있을수록, 타인 음주  
로 인한 피해가 있을수록, 가정의 경제적 상태가 어려울수록, 현재 거주  
형태가 가족과 함께 거주하지 않을수록, 학년이 낮을수록 자살 생각을 할  
가능성이 높았다.

[그림 4-1] 남자 중학생 집단 Tree Plot



〈표 4-3〉 남자 중학생 집단 Logistic regression & Elastic net 결과

	Logistic reg Coeff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
(Intercept)	-3.851	0.385	-9.998	0.000	-3.807
지역 CTYPE_adj	0.431	0.135	3.184	0.001	0.326
주관적 건강 상태 PR_HT	0.114	0.036	3.161	0.002	0.117
주관적 체형 상태 PR_BI	0.040	0.029	1.388	0.165	0.028
고강도 신체활동 일수 PA_VIG	-0.023	0.018	-1.272	0.204	-0.013
주중 평균 학습 시간 PA_SWD_S_adj	0.000	0.000	1.887	0.059	0.000
주말 평균 학습 시간 PA_SWK_S_adj	0.000	0.000	-0.268	0.789	0.000
주중 평균 학습 외 시간 PA_SWD_N_adj	0.000	0.000	-0.225	0.822	0.000
주말 평균 학습 외 시간 PA_SWK_N_adj	0.000	0.000	0.391	0.696	0.000
스트레스 M_STR	-0.622	0.040	-15.467	0.000	-0.563
잠 피로 회복 M_SLP_EN	0.089	0.031	2.871	0.004	0.092
슬픔 경험 여부 M_SAD_adj	1.394	0.067	20.701	0.000	1.329
외로움 경험 M_LON	0.537	0.035	15.452	0.000	0.499
손상 경험 여부 I_SCH_adj	0.056	0.070	0.802	0.422	0.059
손상 병원 치료 여부 I_SCH_TRT_adj	0.085	0.080	1.057	0.291	0.073



	Logistic reg Ceoff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
폭력 치료 경험 V_TRT	0.144	0.054	2.666	0.008	0.143
음주 경험 여부 AC_LT_adj	0.200	0.070	2.858	0.004	0.183
타인 음주 폐해 여부 AC_SND_adj	0.258	0.065	3.985	0.000	0.238
흡연 경험 여부 TC_LT_adj	0.200	0.110	1.809	0.071	0.188
친구 흡연 TC_CSBF	0.109	0.061	1.793	0.073	0.099
성관계 경험 여부 S_SI_adj	0.381	0.151	2.518	0.012	0.348
치료 목적 외 약물 경험 여부 DR_HAB_adj	0.433	0.277	1.564	0.118	0.389
천식 여부 AS_DG_LT_adj	0.119	0.122	0.976	0.329	0.093
알레르기비염 여부 RH_DG_LT_adj	0.087	0.069	1.263	0.207	0.074
아토피피부염 여부 ECZ_DG_LT_adj	0.136	0.078	1.754	0.079	0.115
주중 스마트폰 평균 사용 시간 INT_SPWD_TM_adj	0.000	0.000	0.196	0.845	0.000
주말 스마트폰 평균 사용 시간 INT_SPWK_TM_adj	0.000	0.000	0.312	0.755	0.000
학업 성적 E_S_RCRD	-0.013	0.028	-0.441	0.659	0.000
경제 상태 E_SES	0.080	0.037	2.144	0.032	0.063

	Logistic reg Coeff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
코로나19 이후 경제 상태 E_COVID19	-0.015	0.036	-0.408	0.683	-0.017
가족 동거 여부 E_RES_adj	-0.448	0.192	-2.340	0.019	-0.358
경제적 도움 여부 E_AID_adj	0.152	0.108	1.401	0.161	0.140
학년 GRADE_adj	-0.137	0.041	-3.349	0.001	-0.104

의사결정나무 모형에서는 슬픔·절망감 경험 여부와 외로움 경험 정도가 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 슬픔·절망감 경험 여부, 외로움 경험 정도, 평상시 스트레스 인지 요인이 전체의 82%를 차지할 정도로 주요 요인임을 알 수 있다.

〈표 4-4〉 남자 중학생 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
슬픔 경험 여부	M_SAD_adj	33.833
외로움 경험	M_LON	30.499
스트레스	M_STR	19.344
폭력 치료 경험	V_TRT	2.308
주관적 건강 상태	PR_HT	1.410
잠 피로 회복	M_SLP_EN	1.269
친구 흡연	TC_CSBF	1.081
주중 스마트폰 평균 사용 시간	INT_SPWD_TM_adj	1.044
주중 평균 학습 시간	PA_SWD_S_adj	0.927
주말 평균 학습 외 시간	PA_SWK_N_adj	0.899
경제 상태	E_SES	0.761
타인 음주 폐해 여부	AC_SND_adj	0.737

var		rel.inf
성관계 경험 여부	S_SI_adj	0.688
주말 스마트폰 평균 사용 시간	INT_SPWK_TM_adj	0.650
흡연 경험 여부	TC_LT_adj	0.554
주중 평균 학습 외 시간	PA_SWD_N_adj	0.471
음주 경험 여부	AC_LT_adj	0.388
주말 평균 학습 시간	PA_SWK_S_adj	0.374
고강도 신체활동 일수	PA_VIG	0.344
학년	GRADE_adj	0.280
천식 여부	AS_DG_LT_adj	0.256
가족 동거 여부	E_RES_adj	0.251
손상 경험 여부	I_SCH_adj	0.247
경제적 도움 여부	E_AID_adj	0.246
코로나19 이후 경제 상태	E_COVID19	0.217
치료 목적 외 약물 경험 여부	DR_HAB_adj	0.189
아토피피부염 여부	ECZ_DG_LT_adj	0.172
주관적 체형 상태	PR_BI	0.142
지역	CTYPE_adj	0.142
학업 성적	E_S_RCRD	0.100
알레르기비염 여부	RH_DG_LT_adj	0.097
손상 병원 치료 여부	I_SCH_TRT_adj	0.079

### 3. 여자 중학생 집단의 예측 모형

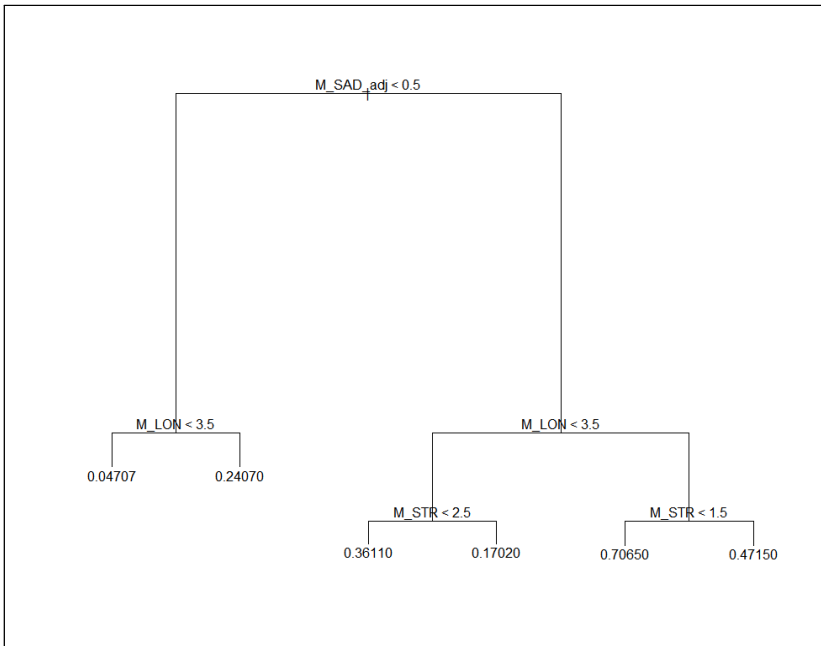
Logistic regression, Elastic net 모형 적용 결과, 주관적 건강 인지가 건강하지 못하다고 생각할수록, 학습 목적으로 주말에 앉아서 보낸 평균 시간이 많을수록, 평상시 스트레스 인지가 높을수록, 슬픔·절망 경험이 있을수록, 외로움을 많이 느낄수록, 학교에서 손상 경험이 있을수록, 최근 폭력으로 인한 치료 경험이 있을수록, 음주 경험이 있을수록, 흡연 경험이 있을수록, 치료 목적을 제외하고 습관적 약물 경험이 있을수록, 가정의 경제적 상태가 어려울수록, 현재 거주형태가 가족과 함께 거주하

지 않을수록, 학년이 낮을수록 자살 생각을 할 가능성이 높았다.

의사결정나무 모형에서는 슬픔·절망감 경험 여부와 외로움 경험 정도, 평상시 스트레스 인지 정도가 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 슬픔·절망감 경험 여부, 외로움 경험 정도, 평상시 스트레스 인지 요인이 남자 중학생과 마찬가지로 전체의 약 83%를 차지하나, 슬픔·절망감 경험 여부의 비중이 상대적으로 높게 나타났다.

[그림 4-2] 여자 중학생 집단 Tree Plot



〈표 4-5〉 여자 중학생 집단 Logistic regression &amp; Elastic net 결과

	Logistic reg Coeff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
(Intercept)	-3.486	0.367	-9.490	0.000	-3.300
지역 CTYPE_adj	0.028	0.100	0.281	0.779	0.000
주관적 건강 상태 PR_HT	0.237	0.032	7.362	0.000	0.218
주관적 체형 상태 PR_BI	0.058	0.028	2.058	0.040	0.042
고강도 신체활동 일수 PA_VIG	0.026	0.017	1.563	0.118	0.018
주중 평균 학습 시간 PA_SWD_S_adj	0.000	0.000	-0.238	0.812	0.000
주말 평균 학습 시간 PA_SWK_S_adj	0.001	0.000	3.029	0.002	0.000
주중 평균 학습 외 시간 PA_SWD_N_adj	0.000	0.000	0.638	0.524	0.000
주말 평균 학습 외 시간 PA_SWK_N_adj	0.000	0.000	1.271	0.204	0.000
스트레스 M_STR	-0.663	0.037	-17.944	0.000	-0.608
잠 피로 회복 M_SLP_EN	0.040	0.028	1.436	0.151	0.039
슬픔 경험 여부 M_SAD_adj	1.325	0.057	23.106	0.000	1.257
외로움 경험 M_LON	0.537	0.032	16.935	0.000	0.498
손상 경험 여부 I_SCH_adj	0.239	0.059	4.074	0.000	0.218
손상 병원 치료 여부 I_SCH_TRT_adj	0.048	0.080	0.603	0.546	0.048
폭력 치료 경험 V_TRT	0.324	0.080	4.075	0.000	0.281
음주 경험 여부 AC_LT_adj	0.268	0.067	4.002	0.000	0.229

	Logistic reg Ceoff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
타인 음주 피해 여부 AC_SND_adj	0.033	0.057	0.581	0.561	0.016
흡연 경험 여부 TC_LT_adj	0.524	0.123	4.269	0.000	0.453
친구 흡연 TC_CSBF	-0.003	0.061	-0.048	0.962	0.000
성관계 경험 여부 S_SI_adj	-0.073	0.170	-0.427	0.670	0.000
치료 목적 외 약물 경험 여부 DR_HAB_adj	0.733	0.287	2.555	0.011	0.606
천식 여부 AS_DG_LT_adj	-0.013	0.121	-0.107	0.915	0.000
알레르기비염 여부 RH_DG_LT_adj	0.016	0.058	0.270	0.787	0.000
아토피피부염 여부 ECZ_DG_LT_adj	0.011	0.061	0.174	0.862	0.000
주중 스마트폰 평균 사용 시간 INT_SPWD_TM_adj	0.000	0.000	-0.975	0.329	0.000
주말 스마트폰 평균 사용 시간 INT_SPWK_TM_adj	0.000	0.000	-0.427	0.669	0.000
학업 성적 E_S_RCRD	0.041	0.026	1.612	0.107	0.025
경제 상태 E_SES	0.102	0.035	2.877	0.004	0.083
코로나19 이후 경제 상태 E_COVID19	-0.034	0.033	-1.029	0.303	-0.032
가족 동거 여부 E_RES_adj	-0.492	0.182	-2.698	0.007	-0.396
경제적 도움 여부 E_AID_adj	0.086	0.092	0.936	0.349	0.055
학년 GRADE_adj	-0.132	0.034	-3.871	0.000	-0.091

(표 4-6) 여자 중학생 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
슬픔 경험 여부	M_SAD_adj	37.874
외로움 경험	M_LON	26.838
스트레스	M_STR	19.909
주관적 건강 상태	PR_HT	3.352
흡연 경험 여부	TC_LT_adj	1.522
폭력 치료 경험	V_TRT	1.324
경제 상태	E_SES	1.117
주말 스마트폰 평균 사용 시간	INT_SPWK_TM_adj	0.998
음주 경험 여부	AC_LT_adj	0.992
손상 경험 여부	I_SCH_adj	0.887
주말 평균 학습 외 시간	PA_SWK_N_adj	0.666
주말 평균 학습 시간	PA_SWK_S_adj	0.614
주중 평균 학습 시간	PA_SWD_S_adj	0.423
고강도 신체활동 일수	PA_VIG	0.423
주중 스마트폰 평균 사용 시간	INT_SPWD_TM_adj	0.402
치료 목적 외 약물 경험 여부	DR_HAB_adj	0.389
주중 평균 학습 외 시간	PA_SWD_N_adj	0.382
잠 피로 회복	M_SLP_EN	0.327
학년	GRADE_adj	0.324
학업 성적	E_S_RCRD	0.270
주관적 체형 상태	PR_BI	0.246
코로나19 이후 경제 상태	E_COVID19	0.242
가족 동거 여부	E_RES_adj	0.207
손상 병원 치료 여부	I_SCH_TRT_adj	0.120
친구 흡연	TC_CSBF	0.067
아토피피부염 여부	ECZ_DG_LT_adj	0.044
경제적 도움 여부	E_AID_adj	0.034
천식 여부	AS_DG_LT_adj	0.005
타인 음주 폐해 여부	AC_SND_adj	0.004
지역	CTYPE_adj	0.000
성관계 경험 여부	S_SI_adj	0.000
알레르기비염 여부	RH_DG_LT_adj	0.000

#### 4. 남자 고등학생 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 일반계 고등학교일 수록, 주관적 건강 인지가 건강하지 못하다고 생각할수록, 평상시 스트레스 인지가 높을수록, 슬픔·절망 경험이 있을수록, 외로움을 많이 느낄수록, 최근 폭력으로 인한 치료 경험이 있을수록, 성관계 경험이 있을수록, 천식으로 평생 의사진단이 있을수록 자살 생각을 할 가능성이 높았다.

〈표 4-7〉 남자 고등학생 집단 Logistic regression & Elastic net 결과

	Logistic reg Coeff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
(Intercept)	-3.969	0.415	-9.567	0.000	-3.977
지역 CTYPE_adj	-0.098	0.129	-0.756	0.450	0.000
고등학교 유형 SCHOOL_adj	-0.201	0.105	-1.915	0.055	-0.036
주관적 건강 상태 PR_HT	0.192	0.039	4.986	0.000	0.177
주관적 체형 상태 PR_BI	-0.009	0.033	-0.265	0.791	0.000
고강도 신체활동 일수 PA_VIG	-0.009	0.021	-0.422	0.673	0.000
주중 평균 학습 시간 PA_SWD_S_adj	0.000	0.000	-0.060	0.952	0.000
주말 평균 학습 시간 PA_SWK_S_adj	0.000	0.000	-0.859	0.390	0.000
주중 평균 학습 외 시간 PA_SWD_N_adj	0.000	0.000	-0.583	0.560	0.000
주말 평균 학습 외 시간 PA_SWK_N_adj	0.000	0.000	1.096	0.273	0.000



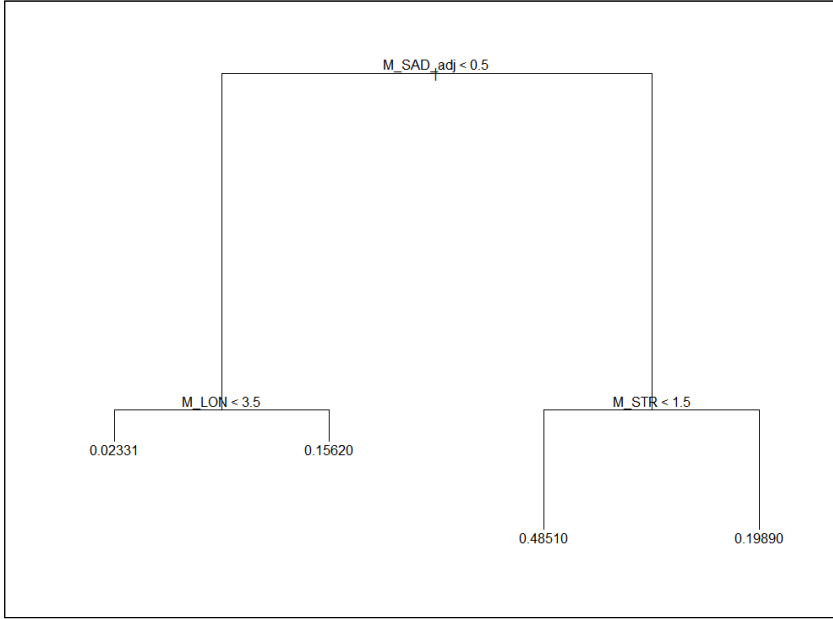
	Logistic reg Ceoff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
스트레스 M_STR	-0.608	0.046	-13.281	0.000	-0.540
잠 피로 회복 M_SLP_EN	0.037	0.035	1.058	0.290	0.021
슬픔 경험 여부 M_SAD_adj	1.462	0.080	18.319	0.000	1.385
외로움 경험 M_LON	0.469	0.039	12.180	0.000	0.437
손상 경험 여부 I_SCH_adj	0.114	0.091	1.257	0.209	0.081
손상 병원 치료 여부 I_SCH_TRT_adj	0.171	0.108	1.593	0.111	0.119
폭력 치료 경험 V_TRT	0.456	0.068	6.693	0.000	0.405
음주 경험 여부 AC_LT_adj	0.033	0.083	0.398	0.691	0.000
타인 음주 폐해 여부 AC_SND_adj	0.130	0.077	1.694	0.090	0.070
흡연 경험 여부 TC_LT_adj	0.253	0.095	2.662	0.008	0.140
친구 흡연 TC_CSBF	-0.085	0.054	-1.593	0.111	0.000
성관계 경험 여부 S_SI_adj	0.276	0.108	2.561	0.010	0.184
치료 목적 외 약물 경험 여부 DR_HAB_adj	0.546	0.261	2.094	0.036	0.459
천식 여부 AS_DG_LT_adj	0.383	0.125	3.059	0.002	0.258
알레르기비염 여부 RH_DG_LT_adj	0.009	0.077	0.114	0.909	0.000

	Logistic reg Ceoff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
아토피피부염 여부 ECZ_DG_LT_adj	0.110	0.086	1.272	0.203	0.042
주중 스마트폰 평균 사용 시간 INT_SPWD_TM_adj	0.000	0.000	0.127	0.899	0.000
주말 스마트폰 평균 사용 시간 INT_SPWK_TM_adj	0.000	0.000	-0.456	0.648	0.000
학업 성적 E_S_RCRD	0.018	0.034	0.532	0.595	0.002
경제 상태 E_SES	0.056	0.044	1.297	0.195	0.029
코로나19 이후 경제 상태 E_COVID19	0.021	0.041	0.516	0.606	0.000
가족 동거 여부 E_RES_adj	-0.135	0.123	-1.098	0.272	-0.035
경제적 도움 여부 E_AID_adj	0.061	0.121	0.503	0.615	0.000
학년 GRADE_adj	-0.019	0.046	-0.414	0.679	0.000

의사결정나무 모형에서는 슬픔·절망감 경험 여부와 외로움 경험 정도, 평상시 스트레스 인지 정도가 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 슬픔·절망감 경험 여부, 외로움 경험 정도, 평상시 스트레스 인지 요인, 최근 1년 폭력으로 인한 치료 경험 순으로 비중이 상대적으로 높게 나타났다.

[그림 4-3] 남자 고등학생 집단 Tree Plot



<표 4-8> 남자 고등학생 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
슬픔 경험 여부	M_SAD_adj	35.451
외로움 경험	M_LON	23.185
스트레스	M_STR	19.765
폭력 치료 경험	V_TRT	7.159
주관적 건강 상태	PR_HT	2.736
주말 평균 학습 외 시간	PA_SWK_N_adj	1.179
흡연 경험 여부	TC_LT_adj	0.752
치료 목적 외 약물 경험 여부	DR_HAB_adj	0.716
주중 스마트폰 평균 사용 시간	INT_SPWD_TM_adj	0.711
주중 평균 학습 외 시간	PA_SWD_N_adj	0.699
천식 여부	AS_DG_LT_adj	0.694
경제 상태	E_SES	0.648

var		rel.inf
주중 평균 학습 시간	PA_SWD_S_adj	0.644
주말 평균 학습 시간	PA_SWK_S_adj	0.626
주관적 체형 상태	PR_BI	0.598
잠 피로 회복	M_SLP_EN	0.582
주말 스마트폰 평균 사용 시간	INT_SPWK_TM_adj	0.524
성관계 경험 여부	S_SI_adj	0.477
손상 경험 여부	I_SCH_adj	0.473
친구 흡연	TC_CSBF	0.414
손상 병원 치료 여부	I_SCH_TRT_adj	0.380
아토피피부염 여부	ECZ_DG_LT_adj	0.307
고강도 신체활동 일수	PA_VIG	0.241
학업 성적	E_S_RCRD	0.212
타인 음주 폐해 여부	AC_SND_adj	0.208
코로나19 이후 경제 상태	E_COVID19	0.133
학년	GRADE_adj	0.128
알레르기비염 여부	RH_DG_LT_adj	0.091
고등학교 유형	SCHOOL_adj	0.086
가족 동거 여부	E_RES_adj	0.078
음주 경험 여부	AC_LT_adj	0.060
경제적 도움 여부	E_AID_adj	0.037
지역	CTYPE_adj	0.007

## 5. 여자 고등학생 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 대도시·중소도시에 살수록, 일반계 고등학교일수록, 주관적 건강 인지가 건강하지 못하다고 생각할수록, 평상시 스트레스 인지가 높을수록, 슬픔·절망 경험이 있을수록, 외로움을 많이 느낄수록, 최근 폭력으로 인한 치료 경험이 있는 경우, 음주 경험이 있는 경우, 흡연 경험이 있는 경우, 성관계 경험이 있는 경우, 치료 목적을 제외하고 습관적 약물 경험이 있는 경우, 스마트폰

주말 평균 사용 시간이 많을수록, 가족과 함께 거주하지 않는 경우, 가정형편으로 인한 경제적 도움을 받았을 때 자살 생각을 할 가능성이 높았다.

〈표 4-9〉 여자 고등학생 집단 Logistic regression & Elastic net 결과

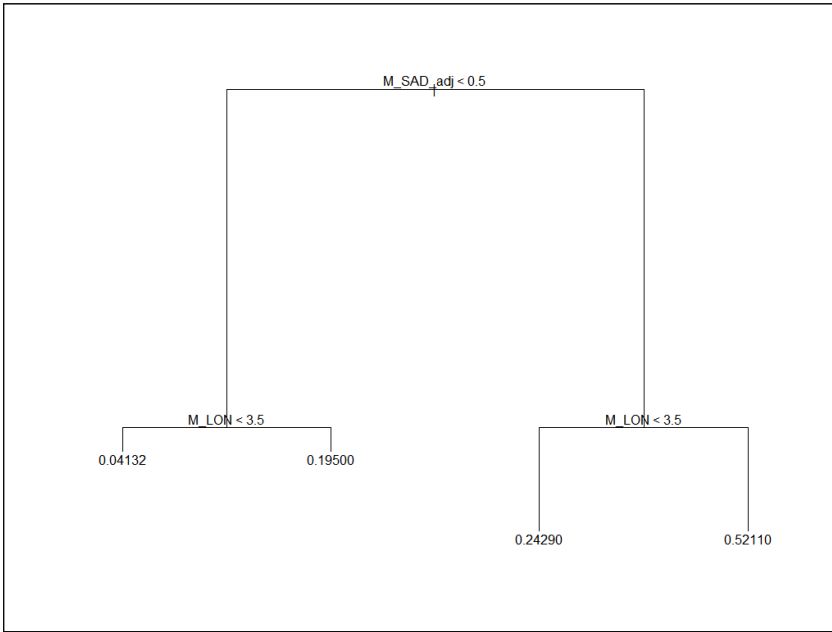
	Logistic reg Coeff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
(Intercept)	-4.496	0.395	-11.394	0.000	-4.222
지역 CTYPE_adj	0.279	0.118	2.366	0.018	0.148
고등학교 유형 SCHOOL_adj	-0.149	0.084	-1.760	0.078	-0.071
주관적 건강 상태 PR_HT	0.245	0.034	7.142	0.000	0.234
주관적 체형 상태 PR_BI	0.022	0.032	0.701	0.483	0.009
고강도 신체활동 일수 PA_VIG	0.034	0.021	1.579	0.114	0.020
주중 평균 학습 시간 PA_SWD_S_adj	0.000	0.000	1.115	0.265	0.000
주말 평균 학습 시간 PA_SWK_S_adj	0.000	0.000	0.420	0.674	0.000
주중 평균 학습 외 시간 PA_SWD_N_adj	0.000	0.000	-0.030	0.976	0.000
주말 평균 학습 외 시간 PA_SWK_N_adj	0.000	0.000	1.131	0.258	0.000
스트레스 M_STR	-0.558	0.041	-13.580	0.000	-0.517
잠 피로 회복 M_SLP_EN	0.008	0.030	0.263	0.792	0.007
슬픔 경험 여부 M_SAD_adj	1.362	0.066	20.595	0.000	1.297
외로움 경험 M_LON	0.503	0.035	14.455	0.000	0.479
손상 경험 여부 I_SCH_adj	0.095	0.075	1.267	0.205	0.088
손상 병원 치료 여부 I_SCH_TRT_adj	0.089	0.107	0.832	0.406	0.073

136 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

	Logistic reg Ceoff	Std. Error	z value	Pr(> z )	Elastic Net Coeff
폭력 치료 경험 V_TRT	0.400	0.093	4.307	0.000	0.345
음주 경험 여부 AC_LT_adj	0.192	0.067	2.854	0.004	0.151
타인 음주 피해 여부 AC_SND_adj	0.139	0.068	2.061	0.039	0.114
흡연 경험 여부 TC_LT_adj	0.534	0.096	5.549	0.000	0.449
친구 흡연 TC_CSFB	-0.062	0.053	-1.181	0.238	0.000
성관계 경험 여부 S_SI_adj	0.317	0.108	2.940	0.003	0.257
치료 목적 외 약물 경험 여부 DR_HAB_adj	0.706	0.276	2.555	0.011	0.627
천식 여부 AS_DG_LT_adj	0.110	0.122	0.908	0.364	0.056
알레르기비염 여부 RH_DG_LT_adj	-0.017	0.063	-0.276	0.783	0.000
아토피피부염 여부 ECZ_DG_LT_adj	0.125	0.066	1.894	0.058	0.091
주중 스마트폰 평균 사용 시간 INT_SPWD_TM_adj	0.000	0.000	-1.496	0.135	0.000
주말 스마트폰 평균 사용 시간 INT_SPWK_TM_adj	0.001	0.000	2.836	0.005	0.000
학업 성적 E_S_RCRD	-0.019	0.030	-0.627	0.530	0.000
경제 상태 E_SES	0.074	0.041	1.785	0.074	0.053
코로나19 이후 경제 상태 E_COVID19	-0.009	0.037	-0.241	0.809	-0.006
가족 동거 여부 E_RES_adj	-0.310	0.109	-2.832	0.005	-0.252
경제적 도움 여부 E_AID_adj	0.220	0.093	2.370	0.018	0.194
학년 GRADE_adj	-0.051	0.038	-1.332	0.183	-0.026

의사결정나무 모형에서는 슬픔·절망감 경험 여부와 외로움 경험 정도가 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

[그림 4-4] 여자 고등학생 집단 Tree Plot



부스팅 방법의 상대적 영향도를 살펴보면 슬픔·절망감 경험 여부의 비중이 전체의 40%를 차지하며, 외로움 경험 정도, 평상시 스트레스 인지 요인, 주관적 건강 인지, 흡연 경험 순으로 비중이 상대적으로 높게 나타났다.

〈표 4-10〉 여자 고등학생 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
슬픔 경험 여부	M_SAD_adj	40.240
외로움 경험	M_LON	24.836
스트레스	M_STR	14.457
주관적 건강 상태	PR_HT	4.485
흡연 경험 여부	TC_LT_adj	3.348
폭력 치료 경험	V_TRT	2.001
주말 스마트폰 평균 사용 시간	INT_SPWK_TM_adj	1.460
경제 상태	E_SES	1.232
주중 평균 학습 시간	PA_SWD_S_adj	0.794
주중 평균 학습 외 시간	PA_SWD_N_adj	0.776
성관계 경험 여부	S_SI_adj	0.719
주말 평균 학습 시간	PA_SWK_S_adj	0.648
치료 목적 외 약물 경험 여부	DR_HAB_adj	0.580
주중 스마트폰 평균 사용 시간	INT_SPWD_TM_adj	0.561
주말 평균 학습 외 시간	PA_SWK_N_adj	0.464
음주 경험 여부	AC_LT_adj	0.365
경제적 도움 여부	E_AID_adj	0.325
손상 경험 여부	I_SCH_adj	0.322
친구 흡연	TC_CSBF	0.309
타인 음주 폐해 여부	AC_SND_adj	0.305
잠 피로 회복	M_SLP_EN	0.300
고강도 신체활동 일수	PA_VIG	0.256
가족 동거 여부	E_RES_adj	0.247
손상 병원 치료 여부	I_SCH_TRT_adj	0.188
주관적 체형 상태	PR_BI	0.176
아토피피부염 여부	ECZ_DG_LT_adj	0.139
코로나19 이후 경제 상태	E_COVID19	0.129
학년	GRADE_adj	0.124
지역	CTYPE_adj	0.093
천식 여부	AS_DG_LT_adj	0.039
고등학교 유형	SCHOOL_adj	0.032
학업 성적	E_S_RCRD	0.026
알레르기비염 여부	RH_DG_LT_adj	0.025



## 제2절 정신건강실태조사를 활용한 예측 모델

### 1. 정신건강실태조사의 분석 개요

이 절에서는 정신건강실태조사(2021)를 활용하여 성별, 연령 구분별로 예측 모델을 구축하였다. 정신건강실태조사 2021의 전체 분석 대상자는 5,511명으로 가중치를 적용하지 않은 수치이다.

평생 동안 자살 생각을 한 적이 있냐(자살 사고)는 질문에 성별, 연령 구분별 빈도는 다음과 같다. 자살 생각은 성별에 따라서, 연령 구분에 따라서 빈도가 다르게 나타남을 알 수 있다.

〈표 4-11〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 빈도표

	빈도	퍼센트
남자(18_39)	851	15.4
남자(40_64)	1397	25.3
남자(65_79)	509	9.2
여자(18_39)	802	14.6
여자(40_64)	1424	25.8
여자(65_79)	528	9.6
전체	5511	100.0

〈표 4-12〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 평생 자살 사고 기초 통계

		Y_평생 자살 사고		전체
		.00	1.00	
남자(18_39)	빈도	782	69	851
	분석 집단 중 %	91.9%	8.1%	100.0%
남자(40_64)	빈도	1260	137	1397
	분석 집단 중 %	90.2%	9.8%	100.0%

		Y_평생 자살 사고		전체
		.00	1.00	
남자(65_79)	빈도	463	46	509
	분석 집단 중 %	91.0%	9.0%	100.0%
여자(18_39)	빈도	716	86	802
	분석 집단 중 %	89.3%	10.7%	100.0%
여자(40_64)	빈도	1255	169	1424
	분석 집단 중 %	88.1%	11.9%	100.0%
여자(65_79)	빈도	461	67	528
	분석 집단 중 %	87.3%	12.7%	100.0%
	빈도	4937	574	5511
	분석 집단 중 %	89.6%	10.4%	100.0%

평생 자살 계획 여부에 응답한 대상자는 전체의 2.4%에 해당하며, 정책집단별 자살 계획(평생) 여부 기초통계는 다음과 같다.

〈표 4-13〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 평생 자살 계획 기초 통계

		(생성)평생 자살 계획		전체
		Negative	Positive	
남자(18_39)	빈도	832	19	851
	분석 집단 중 %	97.8%	2.2%	100.0%
남자(40_64)	빈도	1368	29	1397
	분석 집단 중 %	97.9%	2.1%	100.0%
남자(65_79)	빈도	499	10	509
	분석 집단 중 %	98.0%	2.0%	100.0%
여자(18_39)	빈도	779	23	802
	분석 집단 중 %	97.1%	2.9%	100.0%
여자(40_64)	빈도	1390	34	1424
	분석 집단 중 %	97.6%	2.4%	100.0%
여자(65_79)	빈도	513	15	528
	분석 집단 중 %	97.2%	2.8%	100.0%
	빈도	5381	130	5511
	분석 집단 중 %	97.6%	2.4%	100.0%

평생 자살 시도 경험 여부에 응답한 대상자는 전체의 1.8%이고, 정책 집단별 자살 시도(평생) 여부 기초통계는 다음과 같다.

〈표 4-14〉 정신건강실태조사 2021의 데이터 정책 대상 집단별 평생 자살 시도 기초 통계

		(생성)평생 자살 시도		전체
		Negative	Positive	
남자(18_39)	빈도	837	14	851
	분석 집단 중 %	98.4%	1.6%	100.0%
남자(40_64)	빈도	1376	21	1397
	분석 집단 중 %	98.5%	1.5%	100.0%
남자(65_79)	빈도	502	7	509
	분석 집단 중 %	98.6%	1.4%	100.0%
여자(18_39)	빈도	786	16	802
	분석 집단 중 %	98.0%	2.0%	100.0%
여자(40_64)	빈도	1395	29	1424
	분석 집단 중 %	98.0%	2.0%	100.0%
여자(65_79)	빈도	517	11	528
	분석 집단 중 %	97.9%	2.1%	100.0%
	빈도	5413	98	5511
	분석 집단 중 %	98.2%	1.8%	100.0%

평생 자살 계획 또는 자살 시도 여부에서 둘 다 응답한 대상자는 전체의 1.3%에 해당하며 기초통계는 다음과 같다.

〈표 4-15〉 정신건강실태조사 2021의 데이터 자살 계획 또는 자살 시도 여부 기초통계

	빈도	퍼센트
둘 다 negative	5355	97.2
둘 중 하나만 positive	84	1.5
둘 다 positive	72	1.3
전체	5511	100.0

평생 자살 계획 또는 자살 시도 여부에서 둘 중 하나라도 응답한 대상자는 전체의 2.8%에 해당하며, 정책집단별 기초통계는 다음과 같다.

〈표 4-16〉 정신건강실태조사 2021의 데이터 정책집단별 자살 계획 또는 자살 시도 여부 기초통계

		자살 시도 또는 계획		전체
		negative	positive	
남자(18_39)	빈도	830	21	851
	분석 집단 중 %	97.5%	2.5%	100.0%
남자(40_64)	빈도	1361	36	1397
	분석 집단 중 %	97.4%	2.6%	100.0%
남자(65_79)	빈도	497	12	509
	분석 집단 중 %	97.6%	2.4%	100.0%
여자(18_39)	빈도	777	25	802
	분석 집단 중 %	96.9%	3.1%	100.0%
여자(40_64)	빈도	1380	44	1424
	분석 집단 중 %	96.9%	3.1%	100.0%
여자(65_79)	빈도	510	18	528
	분석 집단 중 %	96.6%	3.4%	100.0%
	빈도	5355	156	5511
	분석 집단 중 %	97.2%	2.8%	100.0%

본 데이터에서 예측 모형에 활용한 변수는 다음 표에 정리하였다.

〈표 4-17〉 정신건강실태조사 2021의 예측 모형 적용 변수

	변수명	변수 설명	
1	읍면동 (dong)	dong_adj	동이면 1, 읍/면 0
2	만 나이 (A2)	A2	만으로 몇 살입니까?
3	결혼 여부 (A4)	A4_adj	결혼한 적 있음 1, 그 외 0

	변수명	변수 설명	
4	자녀 수 (A5)	A5	자녀는 몇 명이나 두었습니까?
5	직업 여부 (A7)	A7_adj	직업 없음 1, 직업 있음 0
6	정규직 여부 (A7B)	A7B_adj	정규직 1, 비정규직 0
7	교육 수준 (A9)	A9	교육 수준
8	의료급여 여부 (DM2)	DM2_adj	의료급여 1, 그 외 0
9	월평균 가구 소득 (DM4_R)	DM4_R_adj	한 달 평균 가구의 총소득
10	전문가 상의 여부 (SR1)	SR1_adj	정신건강 문제로 전문가 상의한 적 있음 1, 그 외 0
11	약물 치료 여부 (SR10)	SR10_adj	정신건강 문제로 전문가로부터 약물처방이 나 약물치료 받은 적 있음 1, 그 외 0
12	신체 질환 여부 (Q2_1_1)	Q2_1_1_adj	신체질환 의사진단 여부 하나라도 1이면 예, 그 외 아니오
13	만성 통증 여부 (Q2_3)	Q2_3_adj	만성통증 예 1, 아니오 0
14	삶의 질 (Q3_2_1)	Q3_2_1	삶의 질 만족도
15	주관적 건강상태 (Q3_2_2)	Q3_2_2	주관적 건강상태
16	현재 삶 만족도 (Q3_3)	Q3_3	현재 삶에 대한 만족도
17	알코올 사용 장애 여부 (AUD)	AUD_adj	평생 유병률 알코올 사용장애 1, 아니오 0
18	니코틴 사용 장애 여부 (TUD)	TUD_adj	평생 유병률 니코틴 사용장애 1, 아니오 0
19	우울장애 여부 (DepD)	DepD_adj	평생 유병률 우울장애 1, 아니오 0
20	불안장애 여부 (AnxD)	AnxD_adj	평생 유병률 불안장애 1, 아니오 0
21	외상 후 스트레스 장애 여부 (PTSD)	PTSD_adj	평생 유병률 외상 후 스트레스 장애 1, 아니 오 0

	변수명	변수 설명	
22	공황장애 여부 (PanD)	PanD_adj	평생 유병률 공황장애 1, 아니오 0
23	사회공포증 여부 (SoPho)	SoPho_adj	평생 유병률 사회공포증 1, 아니오 0
24	범불안장애 여부 (GAD)	GAD_adj	평생 유병률 범불안장애 1, 아니오 0
25	특정공포증 여부 (SpPho)	SpPho_adj	평생 유병률 특정공포증 1, 아니오 0
26	게임중독 고위험군 여부 (GameRisk)	GameRisk_adj	게임중독 고위험군 1, 아니오 0
27	평생 자살 사고 여부 (Y)	Y	평생 자살 사고 여부

이 절에서 검토한 예측 모형은 Logistic regression, Elastic net, Random Forest, Decision Tree, Boosting 방법이다. 이 다섯 가지 모형을 적용하여 각 모형별 유의미한 요인을 찾아보고자 한다.

## 2. 남자 18~39세 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 만 나이가 많을수록, 정신건강 문제로 전문가와 상의한 경험이 있는 경우, 삶의 질 만족도가 낮을수록, 현재 삶에 대한 만족도가 낮을수록, 우울장애가 있는 경우에 자살 생각을 할 가능성이 높았다.

〈표 4-18〉 남자 18~39세 집단 Logistic regression &amp; Elastic net 결과

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
(Intercept)	-3.037	1.814	-1.674	0.094	-0.022
읍면동 dong_adj	0.035	0.416	0.084	0.933	0.000
만 나이 A2	0.084	0.034	2.486	0.013	0.009
결혼 여부 A4_adj	-0.262	0.470	-0.558	0.577	0.000
자녀 수 A5	-0.459	0.354	-1.296	0.195	0.000
직업 여부 A7_adj	0.469	0.751	0.624	0.532	0.000
정규직 여부 A7B_adj	0.226	0.409	0.552	0.581	0.000
교육 수준 A9	0.087	0.203	0.426	0.670	0.000
의료급여 여부 DM2_adj	0.726	1.153	0.629	0.529	0.000
월평균 가구 소득 DM4_R_adj	0.001	0.001	0.780	0.435	0.000
전문가 상의 여부 SR1_adj	2.119	1.169	1.812	0.070	1.158
약물 치료 여부 SR10_adj	-0.339	1.286	-0.264	0.792	0.000
신체 질환 여부 Q2_1_1_adj	-0.054	0.783	-0.068	0.945	0.000
만성 통증 여부 Q2_3_adj	0.321	0.823	0.390	0.697	0.000
삶의 질 Q3_2_1	-0.584	0.343	-1.700	0.089	-0.491
주관적 건강상태 Q3_2_2	0.271	0.276	0.984	0.325	0.000

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
현재 삶 만족도 Q3_3	-0.315	0.171	-1.847	0.065	-0.174
알코올 사용 장애 여부 AUD_adj	0.204	0.396	0.514	0.607	0.000
니코틴 사용 장애 여부 TUD_adj	0.602	0.397	1.516	0.130	0.166
우울장애 여부 DepD_adj	1.237	0.547	2.262	0.024	1.442
불안장애 여부 AnxD_adj	-27.660	1225.575	-0.023	0.982	0.000
외상 후 스트레스 장애 여부 PTSD_adj	29.187	1225.575	0.024	0.981	0.373
공황장애 여부 PanD_adj	27.460	1225.576	0.022	0.982	0.000
사회공포증 여부 SoPho_adj	14.736	866.612	0.017	0.986	0.000
범불안장애 여부 GAD_adj	28.620	1225.575	0.023	0.981	0.265
특정공포증 여부 SpPho_adj	27.174	1225.574	0.022	0.982	0.000
게임중독 고위험군 여부 GameRisk_adj	-0.135	0.426	-0.316	0.752	0.000

Random Forest의 특정 변수에 대한 상대적 중요도 지수는 모든 변수를 사용했을 때와 그 변수를 포함하지 않고 모형을 구축할 때, 예측오차 성능의 차이로 변수의 중요도를 나타내는데, 우울장애 유무, 정신건강 문제로 전문가와 상의한 경험, 주관적 건강상태, 외상 후 스트레스 장애 유

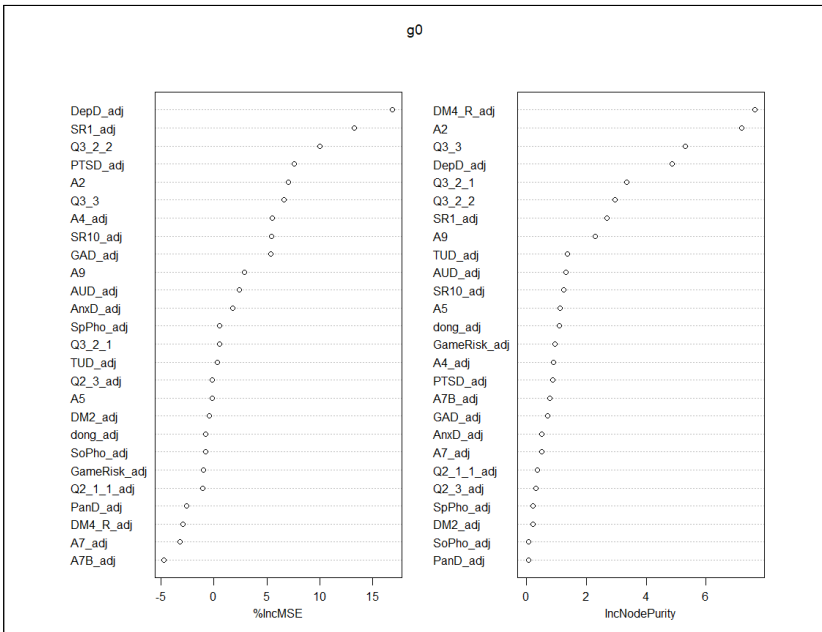


무 순으로 나타났다.

의사결정나무 모형에서는 우울장애 유무, 삶의 질 만족도, 현재 삶에 대한 만족도, 불안장애 유무, 만 나이, 교육 수준, 정신건강 문제로 전문가와 상의한 경험이 자살 생각에 영향을 미친 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 우울장애 유무, 현재 삶에 대한 만족도, 정신건강 문제로 전문가와 상의한 경험, 삶의 질 만족도, 만 연령, 가구 총소득, 주관적 건강상태 순으로 비중이 상대적으로 높게 나타났다.

[그림 4-5] 남자 18~39세 집단 Random Forest Importance Plot

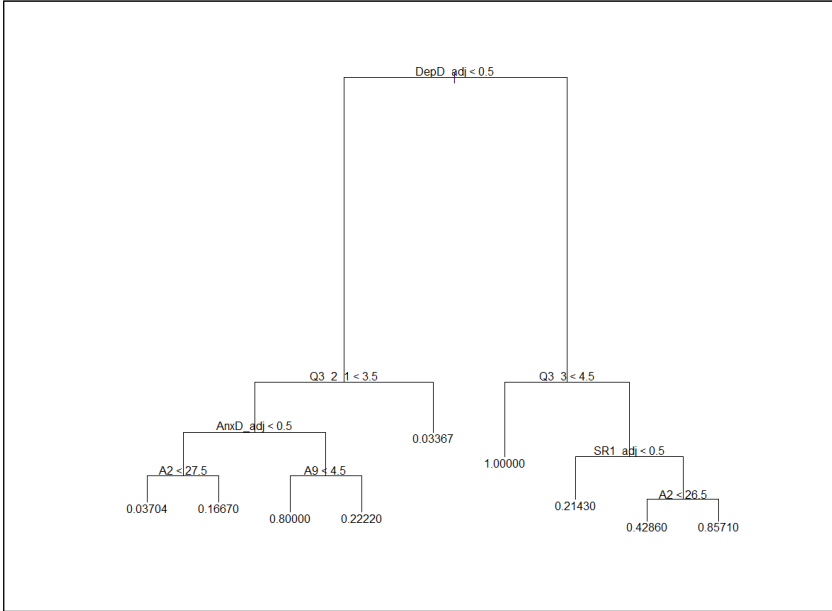


〈표 4-19〉 남자 18~39세 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
우울장애 여부	DepD_adj	32.837
현재 삶 만족도	Q3_3	16.789
전문가 상의 여부	SR1_adj	14.794
삶의 질	Q3_2_1	10.078
만 나이	A2	6.869
월평균 가구 소득	DM4_R_adj	3.755
주관적 건강상태	Q3_2_2	3.133
불안장애 여부	AnxD_adj	2.620
니코틴 사용장애 여부	TUD_adj	2.283
약물 치료 여부	SR10_adj	2.238
교육 수준	A9	1.720
알코올 사용장애 여부	AUD_adj	0.924
결혼 여부	A4_adj	0.406
자녀 수	A5	0.403
직업 여부	A7_adj	0.283
신체 질환 여부	Q2_1_1_adj	0.190
읍면동	dong_adj	0.172
만성 통증 여부	Q2_3_adj	0.172
특정공포증 여부	SpPho_adj	0.154
정규직 여부	A7B_adj	0.124
범불안장애 여부	GAD_adj	0.034
게임중독 고위험군 여부	GameRisk_adj	0.023
의료급여 여부	DM2_adj	0.000
외상 후 스트레스 장애 여부	PTSD_adj	0.000
공황장애 여부	PanD_adj	0.000
사회공포증 여부	SoPho_adj	0.000

[그림 4-6] 남자 18~39세 집단 Tree Plot



### 3. 여자 18~39세 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 등에 살수록, 정신 건강 문제로 전문가와 상의한 경험이 있는 경우, 현재 삶에 대한 만족도가 낮을수록, 우울장애가 있는 경우, 공황장애가 있는 경우, 사회공포증이 있는 경우에 자살 생각을 할 가능성이 높았다.

〈표 4-20〉 여자 18~39세 집단 Logistic regression & Elastic net 결과

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
(Intercept)	0.283	1.524	0.186	0.853	-0.093
읍면동 dong_adj	0.743	0.383	1.939	0.052	0.178
만 나이 A2	0.004	0.031	0.117	0.907	0.000
결혼 여부 A4_adj	0.137	0.411	0.334	0.739	0.000
자녀 수 A5	-0.256	0.249	-1.028	0.304	-0.042
직업 여부 A7_adj	0.422	0.614	0.687	0.492	0.102
정규직 여부 A7B_adj	0.098	0.302	0.325	0.745	0.000
교육 수준 A9	-0.129	0.185	-0.697	0.486	-0.048
의료급여 여부 DM2_adj	1.187	1.349	0.880	0.379	0.339
월평균 가구 소득 DM4_R_adj	0.000	0.001	-0.121	0.904	0.000
전문가 상의 여부 SR1_adj	2.055	0.684	3.006	0.003	0.984
약물 치료 여부 SR10_adj	0.047	0.786	0.060	0.952	0.653
신체 질환 여부 Q2_1_1_adj	-0.187	1.030	-0.182	0.856	0.000
만성 통증 여부 Q2_3_adj	-0.314	1.132	-0.277	0.782	0.000
삶의 질 Q3_2_1	-0.257	0.331	-0.777	0.437	-0.254
주관적 건강상태 Q3_2_2	0.081	0.241	0.336	0.737	-0.055
현재 삶 만족도 Q3_3	-0.341	0.147	-2.326	0.020	-0.159
알코올 사용 장애 여부 AUD_adj	0.469	0.454	1.032	0.302	0.216

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
니코틴 사용 장애 여부 TUD_adj	0.469	0.710	0.661	0.509	0.229
우울장애 여부 DepD_adj	1.269	0.417	3.043	0.002	0.786
불안장애 여부 AnxD_adj	0.445	0.935	0.476	0.634	0.071
외상 후 스트레스 장애 여부 PTSD_adj	-1.338	0.929	-1.440	0.150	-0.043
공황장애 여부 PanD_adj	-2.968	1.465	-2.026	0.043	-0.980
사회공포증 여부 SoPho_adj	-2.085	1.067	-1.954	0.051	-0.284
범불안장애 여부 GAD_adj	-1.825	1.422	-1.283	0.199	-0.414
특정공포증 여부 SpPho_adj	0.525	0.927	0.566	0.571	0.288
게임중독 고위험군 여부 GameRisk_adj	0.081	0.608	0.133	0.894	0.000

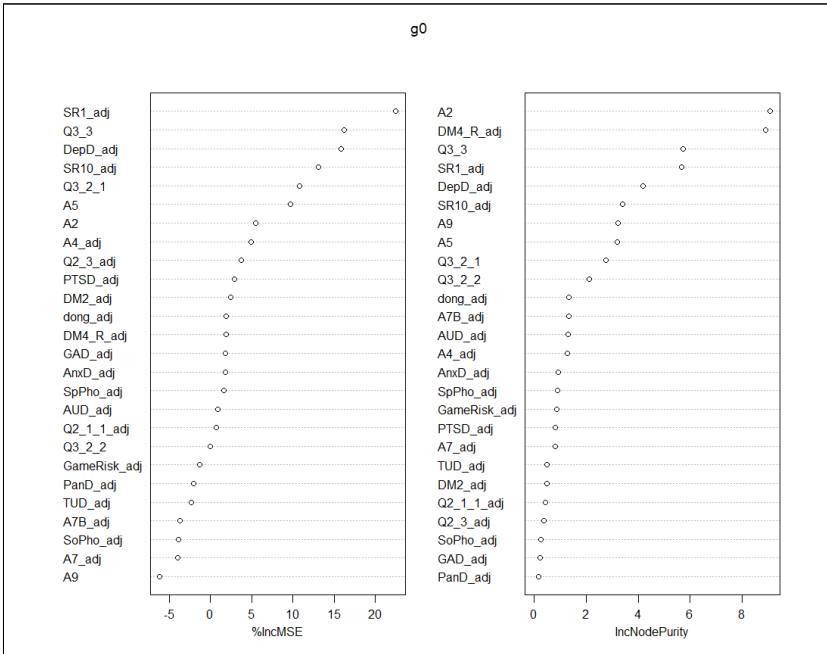
Random Forest의 상대적 중요도 분석 결과, 자살 생각에 영향을 미치는 요인은 정신건강 문제로 전문가와 상의한 경험, 현재 삶에 대한 만족도, 우울장애 유무, 정신건강 문제로 전문가로부터 약물처방이나 약물 치료 받은 적 있는 경우, 삶의 만족도 순으로 나타났다.

의사결정나무 모형에서는 정신건강 문제로 전문가와 상의한 경험, 현재 삶에 대한 만족도, 우울장애 유무, 만 나이, 가구 총소득, 교육 수준이 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 정신건강 문제로 전문가와 상의한 경험, 우울장애 유무, 현재 삶에 대한 만족도, 정신건강 문제로 전문가로부터 약물처방이나 약물 치료 받은 경험, 연령, 가구 총소득 순으로

비중이 상대적으로 높게 나타났다.

[그림 4-7] 여자 18~39세 집단 Random Forest Importance Plot



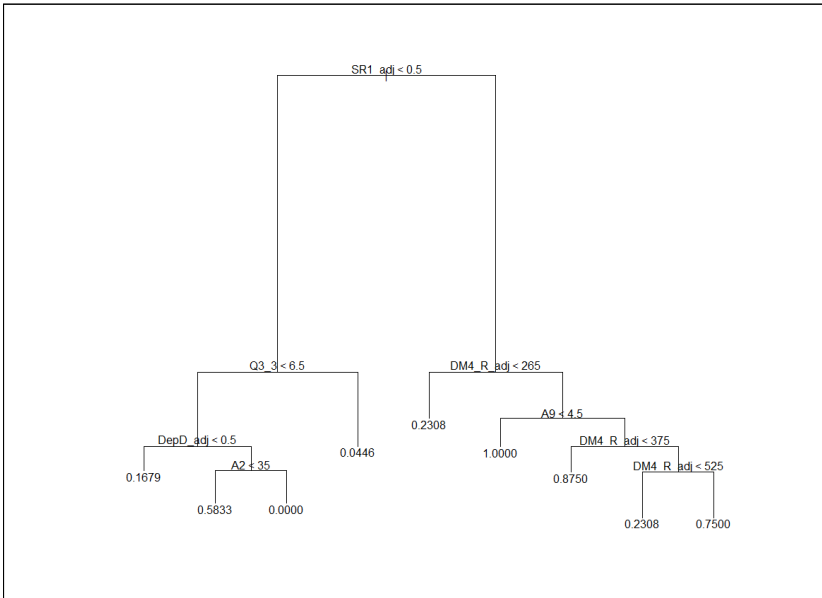
<표 4-21> 여자 18~39세 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
전문가 상의 여부	SR1_adj	28.787
우울장애 여부	DepD_adj	15.761
현재 삶 만족도	Q3_3	12.310
약물 치료 여부	SR10_adj	10.083
만 나이	A2	6.855
월평균 가구 소득	DM4_R_adj	6.709
삶의 질	Q3_2_1	5.561
특정공포증 여부	SpPho_adj	2.916
교육 수준	A9	1.849

var		rel.inf
불안장애 여부	AnxD_adj	1.791
결혼 여부	A4_adj	1.462
자녀 수	A5	1.279
알코올 사용장애 여부	AUD_adj	1.123
정규직 여부	A7B_adj	0.975
읍면동	dong_adj	0.963
주관적 건강상태	Q3_2_2	0.482
직업 여부	A7_adj	0.421
게임중독 고위험군 여부	GameRisk_adj	0.379
니코틴 사용장애 여부	TUD_adj	0.263
외상 후 스트레스 장애 여부	PTSD_adj	0.018
사회공포증 여부	SoPho_adj	0.014
의료급여 여부	DM2_adj	0.000
신체 질환 여부	Q2_1_1_adj	0.000
만성 통증 여부	Q2_3_adj	0.000
공황장애 여부	PanD_adj	0.000
범불안장애 여부	GAD_adj	0.000

[그림 4-8] 여자 18~39세 집단 Tree Plot



#### 4. 남자 40~64세 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 연령이 높을수록, 자녀가 적을수록, 직업이 없는 경우, 의료급여 수급권자인 경우, 의료급여 수급권자가 아닌 경우엔 가구 총소득이 높을수록, 현재 삶에 대한 만족도가 낮을수록, 알코올 사용장애가 있는 경우, 우울장애인 경우, 외상 후 스트레스 장애인 경우, 공황장애인 경우, 범불안장애인 경우, 특정공포증이 있는 경우 자살 생각을 할 가능성이 높았다.

〈표 4-22〉 남자 40~64세 집단 Logistic regression & Elastic net 결과

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net_coeff
(Intercept)	-3.865	1.325	-2.916	0.004	-2.515
읍면동 dong_adj	0.373	0.273	1.366	0.172	0.230
만 나이 A2	0.037	0.018	2.044	0.041	0.020
결혼 여부 A4_adj	-0.022	0.413	-0.053	0.958	0.000
자녀 수 A5	-0.272	0.156	-1.746	0.081	-0.136
직업 여부 A7_adj	-2.547	1.044	-2.440	0.015	-1.227
정규직 여부 A7B_adj	-0.084	0.260	-0.322	0.748	-0.034
교육 수준 A9	0.172	0.113	1.519	0.129	0.079
의료급여 여부 DM2_adj	1.278	0.599	2.135	0.033	0.935
월평균 가구 소득 DM4_R_adj	0.001	0.001	2.148	0.032	0.001
전문가 상의 여부 SR1_adj	-0.165	1.011	-0.163	0.871	0.000



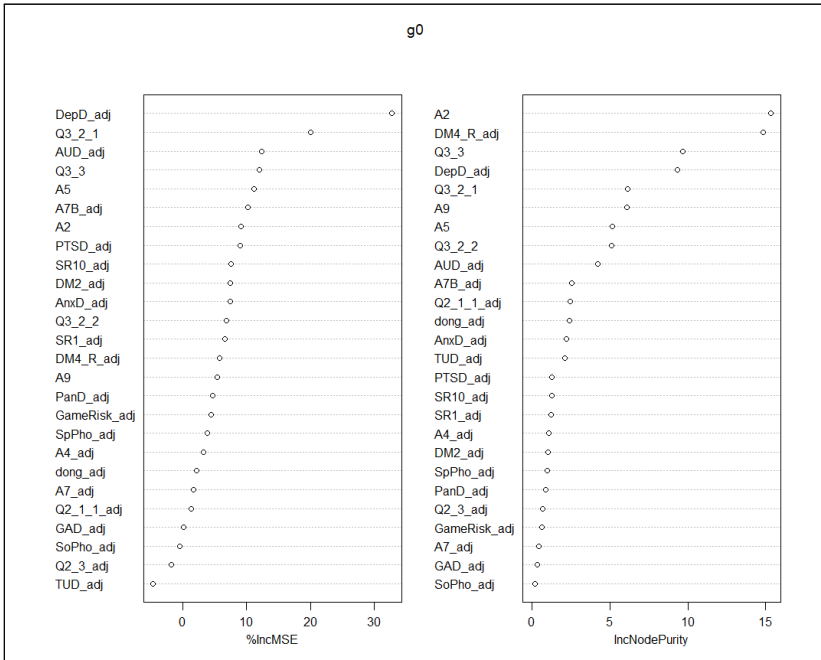
	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net_coeff
약물 치료 여부 SR10_adj	-0.085	1.081	-0.078	0.937	0.008
신체 질환 여부 Q2_1_1_adj	-0.039	0.240	-0.161	0.872	0.000
만성 통증 여부 Q2_3_adj	-0.034	0.431	-0.080	0.936	0.000
삶의 질 Q3_2_1	-0.380	0.231	-1.641	0.101	-0.292
주관적 건강상태 Q3_2_2	0.195	0.195	1.001	0.317	0.000
현재 삶 만족도 Q3_3	-0.223	0.109	-2.043	0.041	-0.135
알코올 사용 장애 여부 AUD_adj	1.047	0.219	4.782	0.000	0.821
니코틴 사용 장애 여부 TUD_adj	0.101	0.238	0.424	0.672	0.110
우울장애 여부 DepD_adj	1.727	0.370	4.674	0.000	1.380
불안장애 여부 AnxD_adj	-3.034	1.625	-1.868	0.062	0.000
외상 후 스트레스 장애 여부 PTSD_adj	3.709	1.624	2.284	0.022	0.795
공황장애 여부 PanD_adj	4.229	1.884	2.244	0.025	1.040
사회공포증 여부 SoPho_adj	2.215	1.696	1.306	0.192	0.983
범불안장애 여부 GAD_adj	2.750	1.437	1.914	0.056	0.252
특정공포증 여부 SpPho_adj	3.330	1.584	2.102	0.036	0.530
게임중독 고위험군 여부 GameRisk_adj	-0.964	0.727	-1.327	0.185	-0.356

Random Forest의 상대적 중요도 분석 결과, 자살 생각에 영향을 미치는 요인은 우울장애 유무, 삶의 질 만족도, 알코올 사용장애 유무, 현재 삶에 대한 만족도, 자녀 수, 정규직 여부 순으로 나타났다.

의사결정나무 모형에서는 우울장애 유무, 삶의 질 만족도, 현재 삶에 대한 만족도, 알코올 사용장애 유무, 만연령이 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 우울장애 유무, 현재 삶에 대한 만족도, 가구 총소득, 삶의 질 만족도, 알코올 사용장애 유무, 만연령, 불안장애, 외상 후 스트레스 장애 순으로 비중이 상대적으로 높게 나타났다.

[그림 4-9] 남자 40~64세 집단 Random Forest Importance Plot

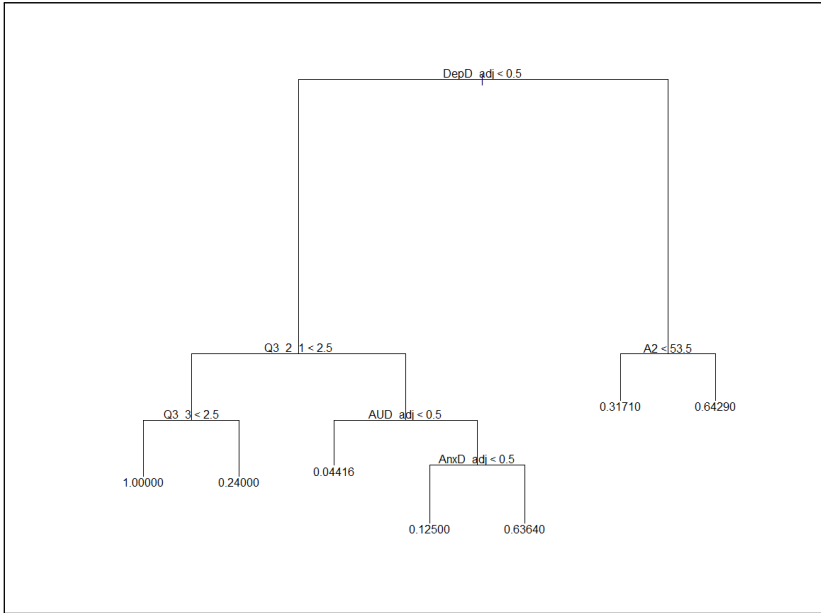


〈표 4-23〉 남자 40~64세 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
우울장애 여부	DepD_adj	22.855
현재 삶 만족도	Q3_3	11.564
월평균 가구 소득	DM4_R_adj	10.156
삶의 질	Q3_2_1	9.944
알코올 사용장애 여부	AUD_adj	9.759
만 나이	A2	8.395
불안장애 여부	AnxD_adj	3.845
외상 후 스트레스 장애 여부	PTSD_adj	3.573
교육 수준	A9	3.421
의료급여 여부	DM2_adj	3.011
주관적 건강상태	Q3_2_2	2.395
약물 치료 여부	SR10_adj	2.081
자녀 수	A5	1.748
특정공포증 여부	SpPho_adj	1.638
전문가 상의 여부	SR1_adj	1.422
직업 여부	A7_adj	0.996
읍면동	dong_adj	0.990
니코틴 사용장애 여부	TUD_adj	0.795
신체 질환 여부	Q2_1_1_adj	0.356
정규직 여부	A7B_adj	0.341
게임중독 고위험군 여부	GameRisk_adj	0.271
결혼 여부	A4_adj	0.205
만성 통증 여부	Q2_3_adj	0.204
범불안장애 여부	GAD_adj	0.032
공황장애 여부	PanD_adj	0.000
사회공포증 여부	SoPho_adj	0.000

[그림 4-10] 남자 40~64세 집단 Tree Plot



## 5. 여자 40~64세 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 알코올 사용장애가 있는 경우, 니코틴 사용장애가 있는 경우, 우울장애가 있는 경우에 자살 생각을 할 가능성이 높았다.

〈표 4-24〉 여자 40~64세 집단 Logistic regression &amp; Elastic net 결과

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net_coeff
(Intercept)	-0.423	1.392	-0.304	0.761	-1.231
읍면동 dong_adj	0.233	0.264	0.884	0.377	0.000
만 나이 A2	-0.005	0.017	-0.275	0.784	0.000
결혼 여부 A4_adj	0.047	0.576	0.082	0.935	0.000
자녀 수 A5	-0.051	0.138	-0.370	0.711	0.000
직업 여부 A7_adj	-0.755	0.719	-1.049	0.294	0.000
정규직 여부 A7B_adj	-0.029	0.200	-0.143	0.886	0.000
교육 수준 A9	-0.039	0.122	-0.320	0.749	0.000
의료급여 여부 DM2_adj	0.205	0.655	0.313	0.754	0.000
월평균 가구 소득 DM4_R_adj	0.001	0.001	1.288	0.198	0.000
전문가 상의 여부 SR1_adj	-0.875	0.763	-1.147	0.251	0.000
약물 치료 여부 SR10_adj	0.588	0.805	0.730	0.466	0.000
신체 질환 여부 Q2_1_1_adj	-0.302	0.253	-1.193	0.233	0.000
만성 통증 여부 Q2_3_adj	0.349	0.318	1.098	0.272	0.000
삶의 질 Q3_2_1	-0.339	0.217	-1.561	0.119	-0.233

160 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

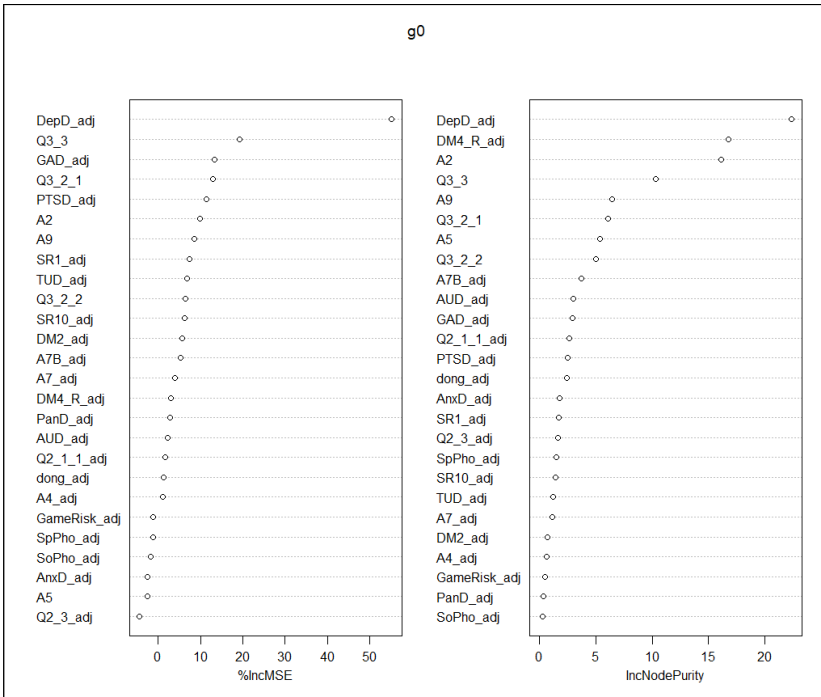
	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net_coeff
주관적 건강상태 Q3_2_2	-0.207	0.197	-1.052	0.293	-0.077
현재 삶 만족도 Q3_3	-0.040	0.094	-0.428	0.669	-0.028
알코올 사용 장애 여부 AUD_adj	0.723	0.341	2.120	0.034	0.606
니코틴 사용 장애 여부 TUD_adj	1.369	0.669	2.047	0.041	0.712
우울장애 여부 DepD_adj	2.496	0.269	9.295	0.000	2.299
불안장애 여부 AnxD_adj	0.654	0.580	1.127	0.260	0.043
외상 후 스트레스 장애 여부 PTSD_adj	0.870	0.613	1.418	0.156	0.684
공황장애 여부 PanD_adj	0.997	1.710	0.583	0.560	0.000
사회공포증 여부 SoPho_adj	-0.661	0.838	-0.789	0.430	0.000
범불안장애 여부 GAD_adj	0.053	0.576	0.091	0.927	0.040
특정공포증 여부 SpPho_adj	-0.471	0.568	-0.830	0.406	0.000
게임중독 고위험군 여부 GameRisk_adj	0.639	0.757	0.844	0.399	0.000

Random Forest의 상대적 중요도 분석 결과, 우울장애 유무, 현재 삶에 대한 만족도, 범불안장애 유무, 삶의 질 만족도, 외상 후 스트레스 장애 순으로 나타났다.

의사결정나무 모형에서는 우울장애 유무, 외상 후 스트레스 장애 유무, 알코올 사용장애 유무가 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 우울장애 유무가 전체의 60%를 차지하고, 가구 총소득, 삶의 질 만족도, 알코올 사용장애 순으로 비중이 상대적으로 높게 나타났다.

[그림 4-11] 여자 40~64세 집단 Random Forest Importance Plot



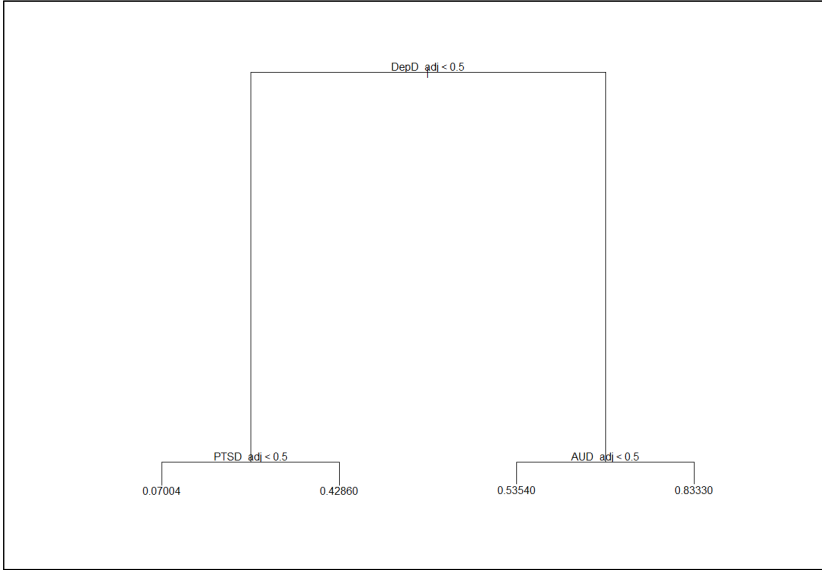
〈표 4-25〉 여자 40~64세 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
우울장애 여부	DepD_adj	60.655
월평균 가구 소득	DM4_R_adj	5.922
삶의 질	Q3_2_1	5.356
알코올 사용장애 여부	AUD_adj	4.743
외상 후 스트레스 장애 여부	PTSD_adj	4.631
현재 삶 만족도	Q3_3	4.526
범불안장애 여부	GAD_adj	3.633
만 나이	A2	3.017
주관적 건강상태	Q3_2_2	1.842
불안장애 여부	AnxD_adj	1.076
교육 수준	A9	0.878
만성 통증 여부	Q2_3_adj	0.860
니코틴 사용장애 여부	TUD_adj	0.658
특정공포증 여부	SpPho_adj	0.469
신체 질환 여부	Q2_1_1_adj	0.333
자녀 수	A5	0.322
전문가 상의 여부	SR1_adj	0.305
약물 치료 여부	SR10_adj	0.230
정규직 여부	A7B_adj	0.160
읍면동	dong_adj	0.158
직업 여부	A7_adj	0.156
의료급여 여부	DM2_adj	0.045
결혼 여부	A4_adj	0.026
공황장애 여부	PanD_adj	0.000
사회공포증 여부	SoPho_adj	0.000
게임중독 고위험군 여부	GameRisk_adj	0.000



[그림 4-12] 여자 40~64세 집단 Tree Plot



## 6. 남자 65~79세 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 가구 총소득이 높을수록, 정신건강 문제로 전문가와 상담한 경험이 있는 경우, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 적이 있는 경우, 삶의 질 만족도가 낮을수록, 우울장애가 있는 경우에 자살 생각을 할 가능성이 높았다.

〈표 4-26〉 남자 65~79세 집단 Logistic regression & Elastic net 결과

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
(Intercept)	-13.357	1038.027	-0.013	0.990	-1.398
읍면동 dong_adj	0.391	0.531	0.736	0.462	0.103
만 나이 A2	-0.034	0.054	-0.634	0.526	-0.007
결혼 여부 A4_adj	13.761	1038.020	0.013	0.989	0.000
자녀 수 A5	0.204	0.247	0.824	0.410	0.000
직업 여부 A7_adj	1.763	1.445	1.220	0.222	0.088
정규직 여부 A7B_adj	0.155	0.400	0.388	0.698	0.000
교육 수준 A9	0.120	0.207	0.581	0.561	0.037
의료급여 여부 DM2_adj	-0.178	1.258	-0.142	0.887	0.000
월평균 가구 소득 DM4_R_adj	0.002	0.001	1.889	0.059	0.001
전문가 상의 여부 SR1_adj	3.165	1.203	2.632	0.009	1.096
약물 치료 여부 SR10_adj	-2.320	1.353	-1.716	0.086	0.036
신체 질환 여부 Q2_1_1_adj	0.248	0.433	0.573	0.567	0.043
만성 통증 여부 Q2_3_adj	0.258	0.534	0.483	0.629	0.070
삶의 질 Q3_2_1	-1.024	0.443	-2.314	0.021	-0.258

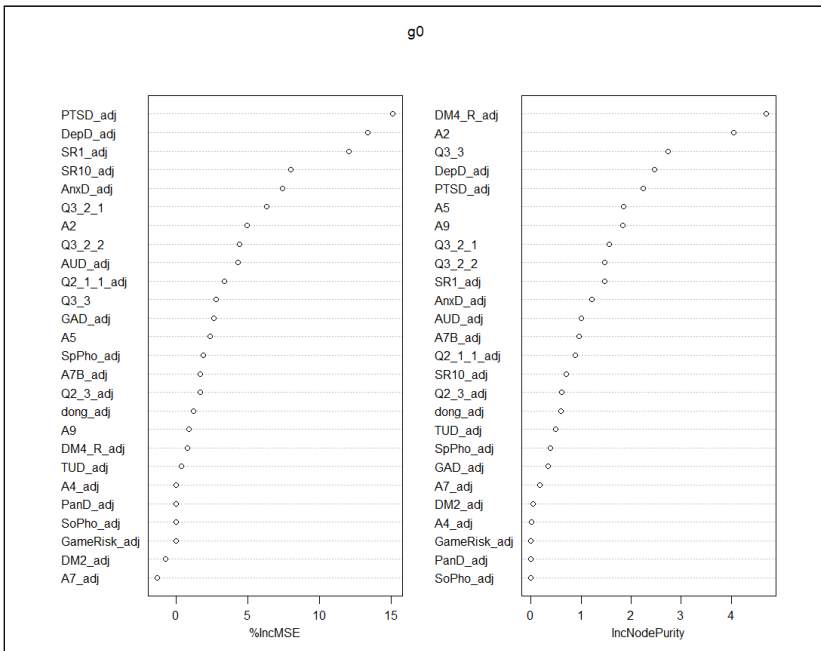
	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
주관적 건강상태 Q3_2_2	-0.092	0.415	-0.221	0.825	-0.120
현재 삶 만족도 Q3_3	0.117	0.203	0.576	0.565	-0.003
알코올 사용 장애 여부 AUD_adj	-0.035	0.468	-0.074	0.941	0.060
니코틴 사용 장애 여부 TUD_adj	-0.828	0.592	-1.398	0.162	-0.072
우울장애 여부 DepD_adj	1.817	0.611	2.973	0.003	1.124
불안장애 여부 AnxD_adj	1.544	2.261	0.683	0.494	0.801
외상 후 스트레스 장애 여부 PTSD_adj	2.320	2.100	1.105	0.269	1.617
공황장애 여부 PanD_adj	0.493	2.060	0.239	0.811	0.407
사회공포증 여부 SoPho_adj	-0.240	2.130	-0.113	0.910	0.000
범불안장애 여부 GAD_adj	-13.548	2399.545	-0.006	0.995	0.000

Random Forest의 상대적 중요도 분석 결과, 외상 후 스트레스 장애 유무, 우울장애 유무, 정신건강 문제로 전문가와 상의한 경험, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 적이 있는 경우 순으로 나타났다.

의사결정나무 모형에서는 우울장애 유무, 가구 총소득, 정신건강 문제로 전문가와 상의한 경험, 불안장애 유무가 자살 생각에 영향을 미치는 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 우울장애, 불안장애, 정신건강 문제로 전문가와 상의한 경험, 가구 총소득, 삶의 질 만족도 순으로 비중이 상대적으로 높게 나타났다.

[그림 4-13] 남자 65~79세 집단 Random Forest Importance Plot

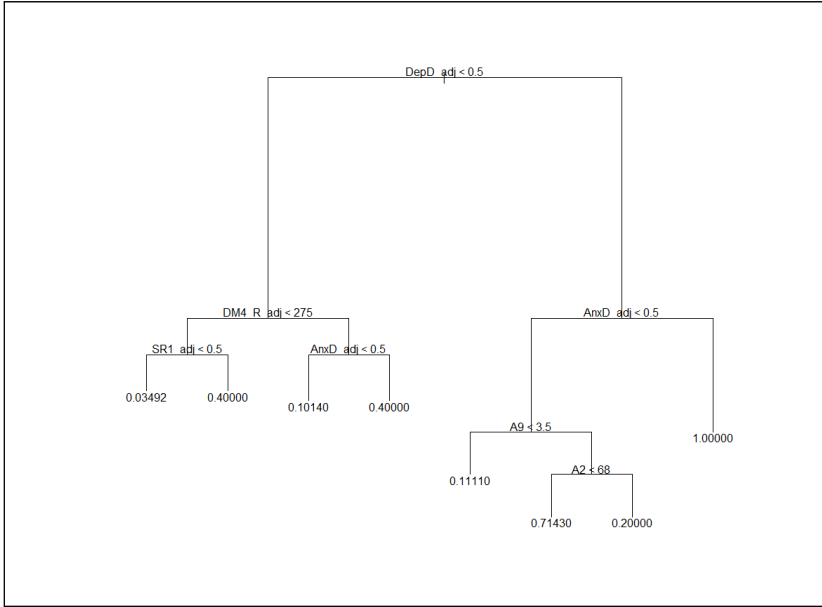


〈표 4-27〉 남자 65~79세 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
우울장애 여부	DepD_adj	27.271
불안장애 여부	AnxD_adj	18.625
전문가 상의 여부	SR1_adj	12.257
월평균 가구 소득	DM4_R_adj	10.823
삶의 질	Q3_2_1	8.978
만 나이	A2	5.534
주관적 건강상태	Q3_2_2	4.478
현재 삶 만족도	Q3_3	3.510
교육 수준	A9	2.935
알코올 사용장애 여부	AUD_adj	1.853
자녀 수	A5	1.112
만성 통증 여부	Q2_3_adj	0.718
니코틴 사용장애 여부	TUD_adj	0.632
신체 질환 여부	Q2_1_1_adj	0.514
정규직 여부	A7B_adj	0.479
읍면동	dong_adj	0.248
약물 치료 여부	SR10_adj	0.032
결혼 여부	A4_adj	0.000
직업 여부	A7_adj	0.000
의료급여 여부	DM2_adj	0.000
외상 후 스트레스 장애 여부	PTSD_adj	0.000
공황장애 여부	PanD_adj	0.000
사회공포증 여부	SoPho_adj	0.000
범불안장애 여부	GAD_adj	0.000
특정공포증 여부	SpPho_adj	0.000
게임중독 고위험군 여부	GameRisk_adj	0.000

[그림 4-14] 남자 65~79세 집단 Tree Plot



## 7. 여자 65~79세 집단의 예측 모형

Logistic regression, Elastic net 모형 적용 결과, 자녀의 수가 적을 수록, 가구 총소득이 높을수록, 정신건강 문제로 전문가와 상의한 적이 있는 경우, 우울장애가 있는 경우 자살 생각을 할 가능성이 높았다.

〈표 4-28〉 여자 65~79세 집단 Logistic regression &amp; Elastic net 결과

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
(Intercept)	-18.061	1669.256	-0.011	0.991	-0.797
읍면동 dong_adj	0.386	0.428	0.902	0.367	0.000
만 나이 A2	0.052	0.043	1.190	0.234	0.000
결혼 여부 A4_adj	15.243	1669.253	0.009	0.993	0.000
자녀 수 A5	-0.602	0.188	-3.202	0.001	-0.241
직업 여부 A7_adj	0.590	0.957	0.616	0.538	0.000
정규직 여부 A7B_adj	-0.033	0.448	-0.073	0.942	0.000
교육 수준 A9	-0.123	0.181	-0.679	0.497	0.000
의료급여 여부 DM2_adj	0.359	0.624	0.576	0.565	0.000
월평균 가구 소득 DM4_R_adj	0.002	0.001	2.315	0.021	0.000
전문가 상의 여부 SR1_adj	1.658	0.858	1.934	0.053	0.246
약물 치료 여부 SR10_adj	-1.289	0.937	-1.375	0.169	0.000
신체 질환 여부 Q2_1_1_adj	-0.111	0.339	-0.329	0.742	0.000
만성 통증 여부 Q2_3_adj	-0.374	0.401	-0.931	0.352	0.000
삶의 질 Q3_2_1	0.065	0.333	0.195	0.845	0.000

170 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

	Logistic coeff	Std. Error	z value	Pr(> z )	Elastic Net coeff
주관적 건강상태 Q3_2_2	-0.297	0.294	-1.010	0.313	0.000
현재 삶 만족도 Q3_3	-0.230	0.146	-1.567	0.117	-0.158
알코올 사용 장애 여부 AUD_adj	1.255	0.832	1.509	0.131	0.918
니코틴 사용 장애 여부 TUD_adj	15.400	2399.545	0.006	0.995	0.000
우울장애 여부 DepD_adj	1.428	0.421	3.394	0.001	1.090
불안장애 여부 AnxD_adj	-2.330	2.424	-0.961	0.337	0.000
외상 후 스트레스 장애 여부 PTSD_adj	3.577	2.652	1.349	0.177	0.401
공황장애 여부 PanD_adj	0.546	1.391	0.392	0.695	0.000
사회공포증 여부 SoPho_adj	-14.331	1559.133	-0.009	0.993	0.000
범불안장애 여부 GAD_adj	1.316	2.542	0.518	0.605	0.000
특정공포증 여부 SpPho_adj	2.000	2.428	0.824	0.410	0.000

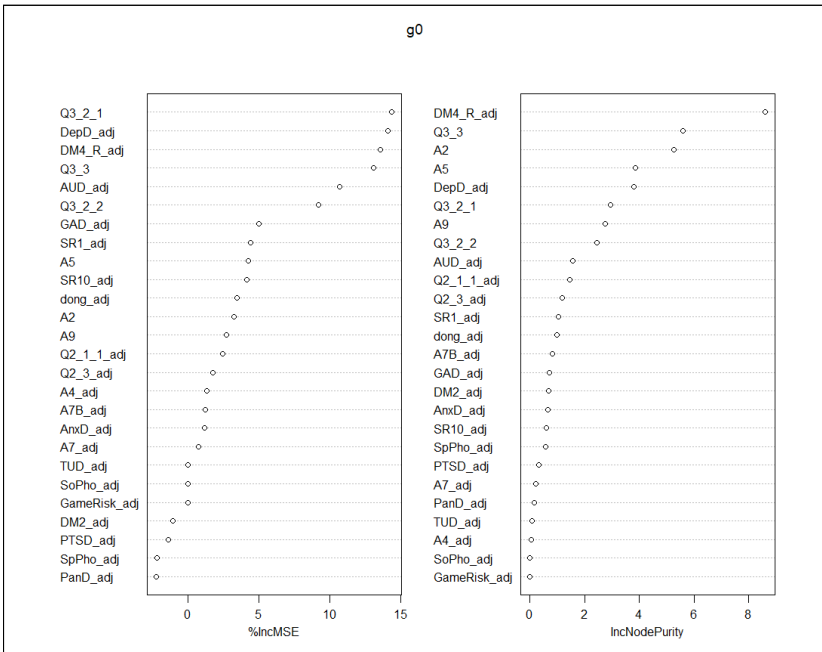


Random Forest의 상대적 중요도 분석 결과, 삶의 질 만족도, 우울장애 유무, 가구 총소득, 현재 삶에 대한 만족도, 알코올 사용장애 유무, 주관적 건강상태 순으로 나타났다.

의사결정나무 모형에서는 우울장애 유무, 알코올 사용장애 유무, 자녀 수, 현재 삶에 대한 만족도, 가구 총소득, 만연령이 자살 생각에 중요한 요인으로 도출되었다.

부스팅 방법의 상대적 영향도를 살펴보면, 우울장애, 가구 총소득, 현재 삶에 대한 만족도, 자녀 수, 만연령, 교육 수준, 정신건강 문제로 전문가와 상의한 경험 유무 순으로 비중이 상대적으로 높게 나타났다.

[그림 4-15] 여자 65~79세 집단 Random Forest Importance Plot

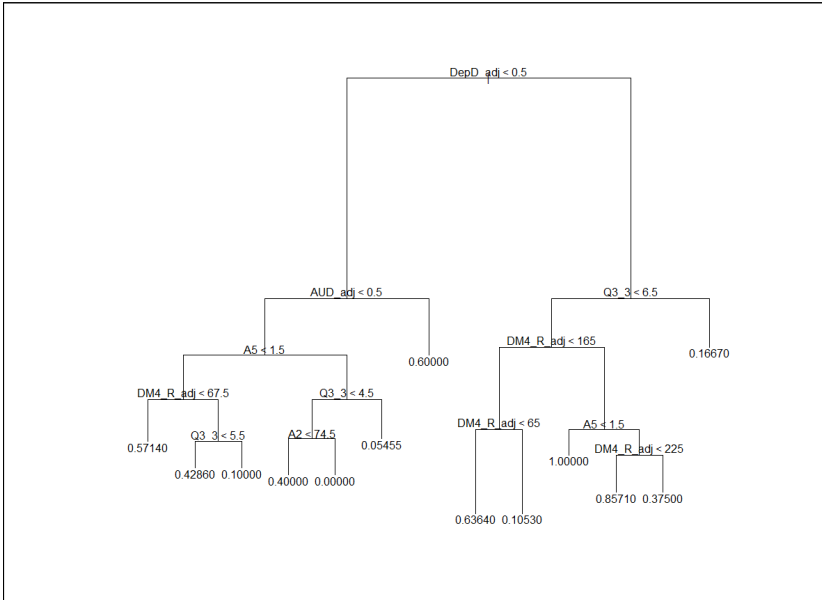


〈표 4-29〉 여자 65~79세 집단 Boosting Relative Importance Plot

(단위: %)

var		rel.inf
우울장애 여부	DepD_adj	20.845
월평균 가구 소득	DM4_R_adj	19.237
현재 삶 만족도	Q3_3	18.603
자녀 수	A5	10.700
만 나이	A2	7.782
교육 수준	A9	5.191
전문가 상의 여부	SR1_adj	5.057
주관적 건강상태	Q3_2_2	3.818
삶의 질	Q3_2_1	3.704
읍면동	dong_adj	2.052
신체 질환 여부	Q2_1_1_adj	1.046
만성 통증 여부	Q2_3_adj	0.480
약물 치료 여부	SR10_adj	0.465
알코올 사용장애 여부	AUD_adj	0.342
의료급여 여부	DM2_adj	0.298
불안장애 여부	AnxD_adj	0.169
특정공포증 여부	SpPho_adj	0.101
정규직 여부	A7B_adj	0.060
범불안장애 여부	GAD_adj	0.047
결혼 여부	A4_adj	0.000
직업 여부	A7_adj	0.000
니코틴 사용장애 여부	TUD_adj	0.000
외상 후 스트레스 장애 여부	PTSD_adj	0.000
공황장애 여부	PanD_adj	0.000
사회공포증 여부	SoPho_adj	0.000
게임중독 고위험군 여부	GameRisk_adj	0.000

[그림 4-16] 여자 65~79세 집단 Tree Plot



### 제3절 소결

이 장에서는 청소년건강행태조사와 정신건강실태조사를 활용하여 집단별로 여러 예측 모형을 적용하였다. 예측 모형은 각 모형의 장단점이 있기 때문에, 그 집단의 특성을 다각도로 이해하기 위해서는 여러 예측 모형으로 유의미한 요인을 살펴보는 것도 중요한 부분이다. 모형의 예측력을 판단하기 위해서는 모형에 대한 평가가 필수적이나, 이 장에서는 모형의 예측력을 판단하기보다는 각 모형에서 유의미한 요인을 도출하는데 집중하였다. 정신건강실태조사 남자 65~79 집단의 경우, 전체 샘플 수는 509이나, 자살 생각이 있는 대상자 수는 그중에 46명이다. 5 fold

CV와 10 fold CV를 적용하여 테스트 샘플을 정하고 각 모형의 예측력을 평가할 수는 있지만, 여기에서는 모형의 예측력보다는 각 정책집단별 유의미한 특성의 차이를 도출해내는 것이 더 의미 있는 분석이라고 판단하였다.

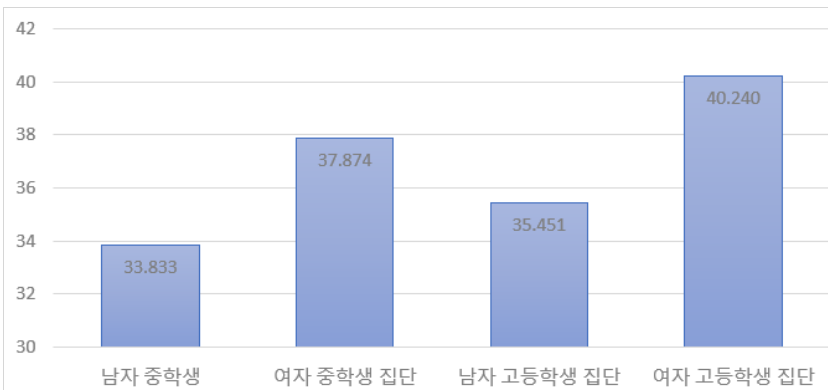
청소년건강행태조사 분석의 종속변수는 최근 1년 동안 자살 생각을 한 적이 있느냐는 문항을 사용하였다. 정신건강실태조사에서 최근 1년 동안 자살 생각을 한 대상자는 전체의 1.3%로 예측 모형을 적용하기에는 자료의 불균형이 심하기 때문에, 정신건강실태조사의 경우 종속변수로 평생 동안 자살 생각을 한 적이 있느냐는 문항을 사용하였다.

아동·청소년 집단인 청소년건강행태조사 대상자는 성별과 중고등학생을 구분하여 4개 집단으로 나누어 예측 모형을 구축하였다. 분석 결과, 전체적으로 주관적 건강 인지, 슬픔·절망감 경험 여부, 외로움 경험 정도, 평상시 스트레스 인지 정도 요인이 자살 생각에 중요한 요인으로 도출되었다. 각 집단별 특징으로 남자 중학생 집단에서는 최근 폭력으로 인한 치료 경험이 있을수록, 흡연 경험이 있을수록, 타인의 음주로 인한 폐해가 있을수록, 가정의 경제적 상태가 어려울수록, 현재 거주형태가 가족과 함께 거주하지 않을수록, 학년이 낮을수록 자살 생각을 할 가능성이 높았다. 여자 중학생 집단에서는 학습 목적으로 주말에 앉아서 보낸 평균 시간이 많을수록, 평상시 스트레스 인지가 높을수록, 슬픔·절망 경험이 있을수록, 외로움을 많이 느낄수록, 학교에서 손상 경험이 있을수록, 최근 폭력으로 인한 치료 경험이 있을수록, 음주 경험이 있을수록, 흡연 경험이 있을수록, 치료 목적을 제외하고 습관적 약물 경험이 있을수록, 가정의 경제적 상태가 어려울수록, 현재 거주형태가 가족과 함께 거주하지 않을수록, 학년이 낮을수록 자살 생각을 할 가능성이 높았다. 남자 고등학생 집단에서는 일반계 고등학교일 때, 최근 폭력으로 인한 치료 경험이

있는 경우, 성관계 경험이 있는 경우 자살 생각을 할 가능성이 높았다. 여자 고등학생 집단 모형에서는 최근 폭력으로 인한 치료 경험이 있는 경우, 음주 경험이 있는 경우, 흡연 경험이 있는 경우, 성관계 경험이 있는 경우, 치료 목적을 제외하고 습관적 약물 경험이 있는 경우, 스마트폰 주말 평균 사용 시간이 많을수록, 가족과 함께 거주하지 않는 경우, 가정형편으로 인한 경제적 도움을 받았을 때 자살 생각을 할 가능성이 높았다. 부스팅의 상대적 영향도 분석 결과, 여자 고등학생 집단은 다른 집단에 비해 슬픔·절망감 경험 여부의 비중이 상대적으로 더 높았다.

“슬픔·절망감 경험 여부”는 네 집단 모두 자살 고위험군을 예측하는 데 매우 중요한 요인이지만 집단별로 비교해보면 차이가 있다. 이는 아동·청소년 집단에서 성, 중고등학생 구분에 따라 슬픔·절망감을 경험한 수준이 다르고, 집단별로 유의미한 요인이 다름을 의미하기도 한다. Boosting 방법의 “슬픔·절망감 경험 여부”에 대한 상대적 영향도(%)를 비교한 그림은 다음과 같다.

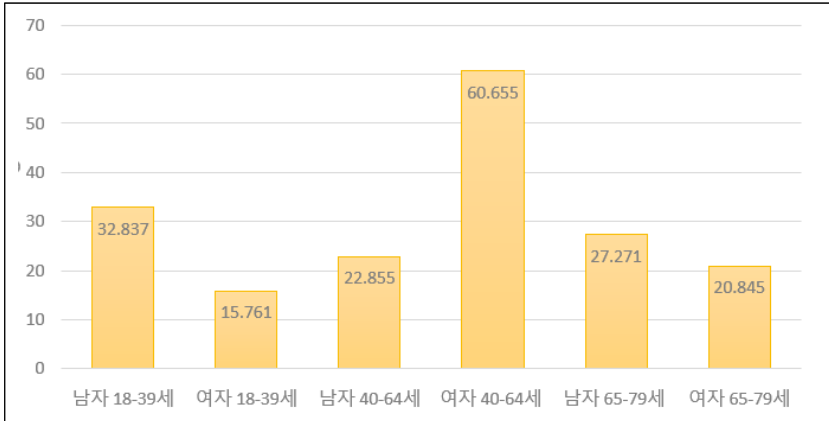
[그림 4-17] Boosting의 “슬픔·절망감 경험 여부” 상대적 영향도(%) 비교



정신건강실태조사에서는 연령과 성별로 6개의 집단을 나누어 각 정책 집단별 예측 모형을 구축하였다. 전체적으로는 우울장애가 있는 경우, 삶의 질 만족도, 현재 삶에 대한 만족도가 자살 생각에 유의미한 영향을 미치는 요인으로 도출되었다. 남자 18~39세 집단에서는 만연률이 높을수록, 정신건강 문제로 전문가와 상의한 경험이 있는 경우에 자살 생각을 할 가능성이 높았다. 여자 18~39세 집단에서는 정신건강 문제로 전문가와 상의한 경험, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 적이 있는 경우, 연령, 가구 총소득이 자살 생각에 영향을 미치는 중요한 요인이었다. 남자 40~64세 집단 예측 모형에서는 연령이 높을수록, 자녀가 적을수록, 직업이 없는 경우, 의료급여 수급권자일 경우, 의료급여 수급권자가 아닌 경우엔 가구 총소득이 높을수록, 알코올 사용장애가 있는 경우, 불안장애, 외상 후 스트레스 장애가 중요한 요인으로 도출되었다. 여자 40~64세 집단 예측 모형에서는 부스팅의 상대적 영향도에서 우울장애 유무가 전체의 60%를 차지할 정도로 중요한 요인으로 도출되었다. 알코올 사용장애, 니코틴 사용장애, 외상 후 스트레스 장애가 자살 생각에 영향을 미치는 중요한 요인이라고 볼 수 있다. 남자 60~79세 집단 예측 모형에서는 가구 총소득, 불안장애, 정신건강 문제로 전문가와 상담한 경험이 있는 경우, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 경험이 있는 경우에 자살 생각을 할 가능성이 높았다. 여자 65~79세 집단 예측 모형에서는 자녀의 수, 가구 총소득, 만연률, 정신건강 문제로 전문가와 상의한 경험이 있는 경우, 알코올 사용장애가 있는 경우에 자살 생각을 할 가능성이 높았다.

청소년건강행태조사의 “슬픔·절망감 경험 여부” 요인과 마찬가지로, 정신건강실태조사의 “우울 장애 여부”는 매우 유의미한 요인이지만, 집단별로 그 수준 차이가 아동·청소년 집단보다 컸다.

[그림 4-18] Boosting의 “우울 장애 여부” 상대적 영향도(%) 비교



이러한 집단별 특성 차이는 정책 대상별 맞춤형 서비스를 제공한다고 할 경우 대상자 우선순위를 정할 때 유용하게 활용될 수 있을 것이다.







## 제5장

### 심층 분석 I : 자살 고위험군에 대한 위험 인자 탐색을 위한 패널데이터 생존 분석 모형 적용

제1절 한국복지패널의 분석 개요

제2절 통계량 산출 및 집단 간 비교

제3절 생존 분석 결과

제4절 소결



## 제 5 장

# 심층 분석 I : 자살 고위험군에 대한 위험 인자 탐색을 위한 패널데이터 생존 분석 모형 적용

자살 고위험군을 케어할 수 있는 실효적 자살예방 정책 구축을 위해 한국복지패널 데이터에서 자살 고위험군을 정의하고, 이와 관계성이 높은 인자들을 탐색하여 정제된 코호트를 구축함으로써 정책 수립 시 근거자료로 활용하고자 한다. 패널 데이터의 정신건강 관련 문항 분석을 통해 자살 생각을 event로 고려하여 이와 관련된 요인들을 심층적으로 살펴볼 필요가 있다. 그러므로 복지패널 데이터의 자살 생각 문항 분석의 일환으로 생존 모형의 적용 가능성을 검토하고 정교한 예측 모형을 개발하고자 한다.

본 장에서는 한국복지패널의 자살 생각 문항에 대한 패턴 분석을 실시하고, 해당 데이터에서 자살 생각과 관련된 정보 수집 및 위험 인자로 평가되는 다양한 공변량들을 수집하였다. 정제된 자료를 바탕으로 자살 생각 여부에 관한 모형 수립이 가능한 코호트를 구축하고, 기초 통계량을 산출하였다. 이를 바탕으로 한국복지패널의 데이터를 생존 모형에 적용할 수 있는지 그 가능성을 검토하고 분석 결과를 정리하였다.

## 제1절 한국복지패널의 분석 개요

본 장에서는 2012년 7차부터 2021년 16차까지의 한국복지패널 데이터를 활용하여 실증 분석을 실시하였다. 독립변수(공변량, covariates)는 기저(baseline)인 7차 자료를 활용하였으며, 종속변수(결과변수, outcome)는 기저(baseline) 이후부터 발생하는 사건을 관찰한 것으로 가정하여 8~16차 자료를 활용하였다. 선행 연구인 김양숙, & 이순희(2020)의 국립정신건강센터 연구보고서를 기초로 유의미한 변수들을 선정하였고, 정제된 결과 모형 구축을 위한 N=4,463명의 대상자와 44개의 변수를 갖는 코호트를 구축하였다.

### 1. 분석 변수

개인의 자살 생각에 영향을 미치는 변수로 인구사회학적 특성, 심리적 요인, 사회문화적 요인을 고려할 수 있으며, 이러한 특성을 갖는 변수는 소득, 건강상태, 살아온 환경, 개인의 가치관, 주변 환경에 대한 만족도, 생활습관, 성별, 교육 수준이 있다. 따라서 위와 같은 공변량들을 자살 생각에 영향을 미칠 만한 변수들로 고려하고자 한다. 결측 비율을 고려하여 본 분석에서 사용한 변수는 총 44개 중 37개가 해당된다.

〈표 5-1〉 변수 설명 및 결측 비율

	변경된 변수명	변경된 변수 범주 설명	기존 변수명	기존 변수 범주 설명	결측 비율
1	경상소득	소수집 4자리 (단위: 만 원)	h07_cin (경상소득)	변경된 변수와 동일	0%
2	저소득	1. 일반가구 2. 저소득층 가구	h07_hc (저소득)	변경된 변수와 동일	0%
3	건강상태	1: 건강하다(1, 2) 2: 보통(3) 3: 건강하지 않다(4, 5)	h0702_2 (가구원1 건강상태) h0702_11 (가구원2 건강상태) h0702_20 (가구원3 건강상태) h0702_29 (가구원4 건강상태) h0702_38 (가구원5 건강상태) h0702_47 (가구원6 건강상태) h0702_56 (가구원7 건강상태) h0702_65 (가구원8 건강상태) h0702_74 (가구원9 건강상태)	1. 아주 건강하다 2. 건강한 편이다 3. 보통이다 4. 건강하지 않은 편이다 5. 건강이 아주 안좋다	6%
4	만성질환	1: 비해당(0) 2: 6개월 미만(1, 2) 3: 6개월 이상(3)	h0702_aq1 (가구원1 만성질환) h0702_aq2 (가구원2 만성질환) h0702_aq3 (가구원3)	0. 비해당 1. 3개월 미만 투병, 투약하고 있다. 2. 3~6개월 투병, 투약하고 있다. 3. 6개월 이상	6%

184 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

변경된 변수명	변경된 변수 범주 설명	기존 변수명	기존 변수 범주 설명	결측 비율	
		만성질환) h0702_aq4 (가구원4 만성질환) h0702_aq5 (가구원5 만성질환) h0702_aq6 (가구원6 만성질환) h0702_aq7 (가구원7 만성질환) h0702_aq8 (가구원8 만성질환) h0702_aq9 (가구원9 만성질환)	투병, 투약 하고 있다.		
5	근로 계약	1: 계약직(1) 2: 비계약직(2)	h0703_3aq1 (근로계약기간 설정 여부)	1. 계약기간이 정해져 있음 2. 계약기간이 정해져 있지 않음	45%
6	근로 지속 가능성	1: 지속 가능(1) 2: 지속 불가능(2)	h0703_3aq2 (근로 지속 가능성)	1. 특별한 사유 (본인의 중대 한 과실, 폐 업 등 사업체 자체의 소멸 또는 고용조 정, 천재지변 등)가 없는 한 계속 근로 가 가능함 2. 본인의 의사 와 무관하게 회사의 사정 에 따라 언제 든지 해고될 수 있음	45%
7	건강보험료	1: 없음(2)	h0705_4	1. 있다	5%

제5장 심층 분석 1: 자살 고위험군에 대한 위험 인자 탐색을 위한 패널데이터 생존 분석 모형 적용 185

	변경된 변수명	변경된 변수 범주 설명	기존 변수명	기존 변수 범주 설명	결측 비율
	미납 경험	2: 있음(1)	(건강보험료 미납 경험)	2. 없다	
8	집 점유	0: 기타(2, 3, 4, 5) 1: 자가(1)	h0706_3 (집의 등기상 점유형태)	1. 자가 2. 전세 3. 보증부월세 4. 월세(사글세) 5. 기타	0%
9	tot_debt	tot_debt: 금융기관 대출/일반사채/카드빚 중 하나라도 있으면 있음, 없으면 없음	h0709_aq1 (금융기관 대출)	단위: 만 원	0%
			h0709_aq2 (일반사채)	단위: 만 원	0%
10	tot_debt_p	tot_debt_p: 금융기관 대출, 일반사채, 카드빚의 총합(단위: 만 원)	h0709_aq3 (카드빚)	단위: 만 원	0%
11	미취업자	1: 해당(3, 4, 5) 2: 미해당(1, 2)	p0702_1 (근로유형)	1. 임금근로자 2. 자영업, 고용주 3. 무급가족 종사자 4. 미취업자 (근로능력 있음) 5. 미취업자 (근로능력 없음)	0%
12	실업	1: 없음(2) 2: 있음(1)	p0702_2 (사직 또는 폐업 경험 여부)	1. 있다 2. 없다	29%
13	근로시간	시간	p0702_8 (규칙적으로 일한 경우 주당 평균 근로시간)	변경된 변수와 동일	47%
14	건강 만족도	1: 만족(4, 5) 2: 불만족(1, 2) 3: 보통(3)	p0703_5 (건강 만족도)	1. 매우 불만족 2. 대체로 불만족 3. 그저 그렇다 4. 대체로 만족 5. 매우 만족	0%

	변경된 변수명	변경된 변수 범주 설명	기존 변수명	기존 변수 범주 설명	결측 비율
15	가족의 수입 만족도	위와 같음	p0703_6 (가족의 수입 만족도)	위와 같음	0%
16	주거환경 만족도		p0703_7 (주거환경 만족도)		
17	가족관계 만족도		p0703_8 (가족관계 만족도)		
18	직업 만족도		p0703_9 (직업 만족도)		
19	사회적 친분 관계 만족도		p0703_10 (사회적 친분 관계 만족도)		
20	여가생활 만족도	위와 같음	p0703_11 (여가생활 만족도)	위와 같음	0%
21	전반적 만족도		p0703_12 (전반적 만족도)		
22	가부장적 관념	1: 그렇지 않다(1, 2) 2: 보통(3) 3: 그렇다(4, 5)	p0704_3aq4 (남성의 임무는 밖에서 돈을 버는 것이고, 여성의 임무는 가정과 가족을 돌보는 것임)	1. 전혀 그렇지 않다 2. 그렇지 않다 3. 그저 그렇다 4. 그렇다 5. 매우 그렇다	0%
23	smoke	1: 비흡연(2) 2: 흡연(1)	p0705_3aq5 (현재 담배를 피우는지 여부)	1. 피움 2. 피우지 않음	0%
24	drink	1: 월 1회 이하(1, 5) 2: 월 2~4(2) 3: 주 2회 이상(3, 4)	p0705_2 (1년간 평균 음주량)	1. 월 1회 이하 2. 월 2~4회 3. 주 2~3회 4. 주 4회 이상 5. 전혀 마시지 않는다	0%
25	식욕	1: 주 1일 미만(1) 2: 주 1-4일(2, 3) 3: 주 5일 이상(4)	p0705_9 (식욕이 없음)	1. 극히 드물다 (일주일 1일 미만)	0%



제5장 심층 분석 1: 자살 고위험군에 대한 위험 인자 탐색을 위한 패널데이터 생존 분석 모형 적용 187

	변경된 변수명	변경된 변수 범주 설명	기존 변수명	기존 변수 범주 설명	결측 비율
				2. 가끔 있었다 (일주일에 1-2일간) 3. 종종 있었다 (일주일에 3-4일간) 4. 대부분 그랬다(일주일에 5일 이상)	
26	우울	위와 같음	p0705_11 (상당히 우울)	위와 같음	0%
27	불면		p0705_13 (잠을 설침)		
28	무기력	위와 같음	p0705_19 (무엇을 해 나갈 엄두가 나지 않음)	위와 같음	0%
29	좌절감		p0705_22 (나는 실패한 사람이라는 느낌이 든다)		
30	bad	1: 없음(0, 1) 2: 있음(2, 3, 4)	p0705_4aq9 (모욕적, 악의적인 이야기 경험) p0705_4aq10 (신체적 폭력의 위협 경험) p0705_4aq11 (신체적 폭력을 행사 당한 경험)	1. 전혀 없음 2. 1~2번 3. 3~5번 4. 6번 이상 0. 비해당	0%
31	가족생활 만족도	1: 만족(5, 6, 7) 2: 보통(4) 3: 불만족(1, 2, 3) 4: 비해당(0)	p0705_aq1 (가족생활에 대한 만족도)	1. 매우 불만족 2. 불만족 3. 약간 불만족 4. 보통 5. 약간 만족	0%

	변경된 변수명	변경된 변수 범주 설명	기존 변수명	기존 변수 범주 설명	결측 비율
				6. 만족 7. 매우 만족 0. 해당 없다	
32	배우자 관계 만족도	위와 같음	p0705_aq2 (배우자와의 관계에 대한 만족도)	위와 같음	0%
33	자녀 관계 만족도		p0705_aq3 (자녀와의 관계에 대한 만족도)		
34	자녀들의 형제자매 관계 만족도		p0705_aq4 (자녀들의 형제자매 관계에 대한 만족도)		
35	성별	1: 남 2: 여	h0701_4 (가구원1 성별) h0701_16 (가구원2 성별) h0701_28 (가구원3 성별) h0701_40 (가구원4 성별) h0701_52 (가구원5 성별) h0701_64 (가구원6 성별) h0701_76 (가구원7 성별) h0701_88 (가구원8 성별) h0701_100 (가구원9 성별)	변경된 변수와 동일	0.34%
36	year		year (나이)		0%
37	edulevel	1: 고졸 이하 2: 대졸 이상	edulevel (교육 수준)	변경된 변수와 동일	0%
38	부모님이 일찍 돌아가신 경험	1. 그렇지 않다 2. 그렇다	np0706_31 (부모님이)	변경된 변수와 동일	100% (결측)

	변경된 변수명	변경된 변수 범주 설명	기존 변수명	기존 변수 범주 설명	결측 비율
		3. 모름	일찍 돌아가신 경험)		개수: 4,463)
39	부모님이 돌아가신 당시 나이	나이	np0706_32 (부모님이 돌아가신 당시 나이)	변경된 변수와 동일	100% (결측 개수: 4,463)
40	부모님이 이혼 경험	1. 그렇지 않다 2. 그렇다 3. 모름	np0706_33 (부모님이 이혼한 경험)	변경된 변수와 동일	100% (결측 개수: 4,463)
41	최종 학력	1. 중학교 졸업 이하 2. 고등학교 중퇴, 졸업 3. 전문대학 재학, 중퇴, 졸업 4. 대학교(4년제) 재학, 중퇴, 졸업 5. 대학원 이상	p0707_3aq1 (최종 학력)	변경된 변수와 동일	100% (결측 개수: 4,463)
42	아동기 경제력	1. 매우 가난 2. 가난 3. 보통 4. 부유 5. 매우 부유	np0706_2 (아동기의 경제적 생활상태)	변경된 변수와 동일	100% (결측 개수: 4,463)
43	건강보험료 미납 이유	1. 건강보험을 별로 이용할 일이 없어서 2. 건강보험료를 납부할 돈이 없어서 3. 내는 돈에 비해 혜택이 적어서 4. 제도에 대한 불신 때문에 5. 납기기한을 잊어버려서 6. 기타	h0705_5 (건강보험료 미납 이유)	변경된 변수와 동일	98.9% (결측 개수: 4,415)
44	건강보험료 미납 기간	연간: 개월	h0705_6 (건강보험료 미납 기간)	변경된 변수와 동일	98.9% (결측 개수: 4,415)

주: 38부터 44까지의 변수들은 결측 비율로 인해 분석에서 제외하였음.  
자료: 7~16차 한국복지패널 데이터를 참고하여 저자 작성

## 2. 분석 수행 내용 및 기초 설명

### 가. 통계량 산출 및 집단 간 비교

정규성 가정( $P$ -값  $> 0.05$  with Shapiro-Wilk test)을 만족하는 연속 변수에 대해서는 평균 $\pm$ 표준편차로 표현되며, 그룹 간 비교를 위한  $P$ -값은 Student's  $t$ -test를 이용하여 계산했다. 정규성 가정이 만족되지 않는 경우 중앙값(사분위수 범위[interquartile range, IQR])을 추정하고 Mann-Whitney U 검정을 사용하여  $P$ -값을 계산했다. 범주형 변수는 개수(%)로 표시하며,  $P$ -값은 카이제곱 검정(chi-squared test)이나 피셔의 정확 검정(Fisher's exact test)을 이용하여 계산했다.

### 나. 생존 분석(survival analysis)

#### 1) 카플란 마이어 곡선(Kaplan-Meier curve)

카플란 마이어 곡선은 연구 기간 동안에 사건이 발생한 시점마다 구간 생존율을 산출하고, 마지막까지 추적된 대상자에 대한 생존율을 구한 뒤, 최종적으로 누적 생존율을 산출하는 방법이다. 사건의 관찰 기간에 따라 순서대로 자료를 정렬한 뒤, 구간별로 관찰된 대상자 수 중에서 생존자 수의 비율로 정의된 구간 생존율을 산출한다. 누적 생존율은 구간별 구간 생존율을 차례대로 곱하여 산출한다.

시점별 누적 생존율에 대해 계단식 그림으로 표현한 것이 카플란 마이어 곡선이며, 우하향하는 형태를 보인다. 본 연구에서는 자살 생각 여부에 따라 그룹을 정의했으며, 그룹 간 차이를 나타내는  $P$ -값을 산출하여

함께 표현했다. 자살 생각 여부에 따른 그룹 간 비교에는 로그-순위 검정 (log-rank test)을 사용했다.

## 2) 콕스의 비례위험 모형

콕스 비례위험 모형은 생존과 관련된 반응변수와 공변량 사이의 관계를 설명하는 데에 가장 많이 사용되는 모형 중 하나이다.

생존시간 데이터는  $(t_i, \mathbf{x}_i(t), \delta_i)$   $i = 1, 2, \dots, n$ 로 표현되며,  $t_i$ 는 연구 기간 동안  $i$ 번째 개체가 관측된 시간,  $\delta_i$ 는  $i$ 번째 개체가 사건이 발생한 경우 1, 중도절단인 경우 0인 지시변수,  $\mathbf{x}_i(t)$ 는  $i$ 번째 개체의 공변량 벡터  $(x_{i1}(t), x_{i2}(t), \dots, x_{ip}(t))^T$ 를 의미한다.

편의상  $t_i (i = 1, 2, \dots, n)$ 를 정렬한  $t_{(1)} < t_{(2)} < \dots < t_{(n)}$  을  $t_1, t_2, \dots, t_n$ 로 놓는다.

$i$ 번째 개체에 대한 콕스 비례위험 모형은 다음과 같이 표현된다.

$$\lambda(t|\mathbf{x}_i) = \lambda_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x}_i).$$

여기서  $\lambda(t|\mathbf{x}_i)$ 는  $i$ 번째 개체의 시점  $t$ 에서 위험함수(hazard),  $\lambda_0(t)$ 는 임의의 기저위험함수,  $\boldsymbol{\beta}$ 는  $(\beta_1, \beta_2, \dots, \beta_p)^T$ 인 계수벡터이다.

동점이 없는 생존시간 데이터에 대해  $\boldsymbol{\beta}$ 는 다음의 부분우도함수를 최대화함으로써 얻어진다.

$$PL(\boldsymbol{\beta}) = \prod_{i=1}^n \left\{ \frac{y_i \exp(\boldsymbol{\beta}^T \mathbf{x}_i)}{\sum_{j \in R_i} y_j \exp(\boldsymbol{\beta}^T \mathbf{x}_j)} \right\}^{\delta_i} = \prod_{i=1}^n \left\{ \frac{y_i r_i(\boldsymbol{\beta}, t)}{\sum_{j \in R_i} y_j r_j(\boldsymbol{\beta}, t)} \right\}^{\delta_i}.$$

여기서  $r_i(\boldsymbol{\beta}, t) = \exp(\boldsymbol{\beta}^T \mathbf{x}_i) \equiv r_i(t)$ 는  $i$ 번째 개체에 대한 위험점수 (risk score),  $y_i = y_i(t)$ 는  $i$ 번째 개체가  $t$ 시점에서 위험집합에 속하면

1, 아니면 0인 지시변수이다.

따라서 로그 부분우도함수  $\log PL(\beta) \equiv l(\beta)$ 를  $\beta$ 에 대해 미분하면 점수함수(score function)  $U(\beta)$ 를 얻을 수 있고, 최대부분우도추정량  $\hat{\beta}$ 은 다음의 부분우도방정식을 풀어 구할 수 있다.

$$U(\hat{\beta}) = \frac{\partial l(\hat{\beta})}{\partial \beta} = \mathbf{0}.$$

하지만 부분우도방정식으로부터 닫힌 형식의 해(closed form solution)  $\hat{\beta}$ 를 구할 수 없으므로 다음의 Newton-Raphson 알고리즘을

$\hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} + I^{-1}(\hat{\beta}^{(k)}) \cdot U(\hat{\beta}^{(k)})$ ,  $l(\hat{\beta}^{(k+1)}) \approx l(\hat{\beta}^{(k)})$  될 때까지 반복하여 근사적인 해  $\hat{\beta}$ 를 구한다. 여기서  $I(\beta) = -\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^T}$  인 정보행렬(information matrix)이다.

이렇게 구한  $\hat{\beta}$ 을 기반으로 누적위험함수를 구하기 위해 다음과 같은 Breslow 추정량을 이용한다.

$$\hat{H}(t|\mathbf{x}) = \sum_{t_{(i)} \leq t} \frac{d_i}{\sum_{j \in R_i} \exp(\beta^T \mathbf{x}_j)}, \quad d_i = \sum_{j=1}^n \delta_j I(t_j = t_i).$$

또한, 최대부분우도추정량  $\hat{\beta}$ 은 마팅계일의 중심극한정리에 의해 근사적으로 정규분포를 따른다( $\hat{\beta} - \beta \sim N(0, I^{-1}(\beta))$ ).

### 3) 모형 평가 및 다중공선성 진단

Concordance-Index(C-index): C-index란 Survival Analysis에서 가장 많이 사용되는 모형 평가 지표이다. 대상자의 정확한 생존 시간을 평가하지 않고, 대신 여러 대상의 생존 시간(또는 위험)을 상대적으로

비교한다. 즉, 사망 순서를 잘 예측하는지를 판단하는 지표로 사용된다. C-index는 0과 1 사이의 값을 갖는데, 1에 가까울수록 정확히 예측한다고 해석하고, 0.5에 가까울수록 무작위로 예측한다고 평가한다.

분산팽창인수(variance inflation factor, VIF): 분산팽창인수는 다중공선성 발생 여부를 확인하기 위한 값이다. 다중공선성이란 예측 변수들 간의 상관성을 가짐을 의미하며, 일반적으로 해당 값이 10 이상이면 높은 다중공선성을 가지고 있다고 판단하여 공변량을 모형에 포함시키는 데 있어 추가적인 조치를 취한다.

## 다. 활용된 통계 패키지 및 통계적 수준

통계 분석은 R version 4.3.1 프로그램(The R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>)을 사용하여 수행하였다. 양측검정에서 P-값이 0.05보다 작으면 통계적으로 유의한 것으로 간주하였다.

## 제2절 통계량 산출 및 집단 간 비교

### 1. 결과변수의 빈도 및 비율

아래는 8차부터 16차(총 9회)에 걸쳐 조사된 자살 생각의 횟수별 빈도 및 비율에 대한 표이다. <표 5-2>에 따르면, 자살 생각을 6회 한 비율은 0.09%, 5회, 4회 한 비율은 모두 0.29%, 3회 한 비율은 1.03%, 2회 한 비율은 2.49%, 1회 한 비율은 8.6%, 그리고 한 번도 하지 않은 비율은 87.21%로 나타났다.

〈표 5-2〉 자살 생각 결과변수의 빈도 및 비율

자살 생각	빈도	비율(%)	누적 빈도	누적 비율(%)
6회	4	0.09	4	0.09
5회	13	0.29	17	0.38
4회	13	0.29	30	0.67
3회	46	1.03	76	1.7
2회	111	2.49	187	4.19
1회	384	8.6	571	12.79
없음	3892	87.21	4463	100
전체 대상자 수	4463	100		

자료: 7~16차 한국복지패널 데이터를 참고하여 저자 작성

누적 비율로 보았을 때, 자살 생각을 6회 이상 한 비율은 0.09%, 5회 이상의 비율은 0.38%, 4회 이상의 비율은 0.67%, 3회 이상의 비율은 1.7%, 2회 이상 한 비율은 4.19%, 1회 이상 한 비율은 12.79%로, 1회 이상 및 2회 이상 자살 생각을 한 누적 비율이 모두 4% 이상의 비율을 보이고 있다. 자살 생각 1회는 대상자의 당시 상황에 따라 충동적 사고로도 간주할 수 있다고 판단하여 결과변수를 정의하였다.

본 장에서는 해당 표를 기반으로 결과변수의 정의를 1) 1회 이상의 자살 생각을 가진 자, 2) 2회 이상의 자살 생각을 가진 자로 나누어 분석하였다.

## 2. 자살 생각 여부에 따른 차이 검정

자살 생각 여부는 1회 이상의 자살 생각과 2회 이상의 자살 생각으로 각각 나누어 표를 작성하였다. 범주형 변수의 경우 개수(비율 %)로 통계량을 산출하고, 그룹 간 차이는 카이제곱 검정 또는 피셔의 정확 검정을



수행하였다. 연속형 변수의 경우 정규성 검정 결과에 따라 평균±표준편차와 중앙값[제1사분위 수, 제3사분위 수]으로 표기하였고 그룹 간 차이는 정규성 검정 결과에 따라 모수적 또는 비모수적 방식을 선택하여 P-값을 산출하였다.

〈표 5-3〉에 따르면, 자살 생각이 없었던 그룹과 자살 생각이 1회 이상 있었던 그룹 사이에 유의수준 0.05하에서 유의한 차이가 있다고 볼 수 있는 변수는 경상소득, year, 저소득, 건강상태, 만성질환, 근로계약, 근로지속 가능성, 집 점유, 미취업자, 건강 만족도, 가족의 수입 만족도, 주거환경 만족도, 가족관계 만족도, 직업 만족도, 사회적 친분 관계 만족도, 여가생활 만족도, 전반적 만족도, 가부장적 관념, drink, 식욕, 우울, 불면, 무기력, 좌절감, bad, 가족생활 만족도, 배우자 관계 만족도, 자녀 관계 만족도, 자녀들의 형제자매 관계 만족도, 성별, 그리고 edulevel이다.

따라서 경상소득과 나이에서 자살 생각을 한 번도 해본 적 없는 집단과 한 번 이상 해본 집단 사이에 유의한 차이가 존재한다. 저소득 여부, 건강의 상태, 만성질환의 정도, 근로 계약기간의 설정 여부, 근로 지속 가능성, 집의 등기상 점유형태, 근로유형, 건강 만족도, 가족의 수입 만족도, 주거환경 만족도, 가족관계 만족도, 직업 만족도, 사회적 친분 관계 만족도, 여가생활 만족도, 전반적 만족도, 가부장적 관념의 정도, 1년 평균 음주량의 정도, 식욕이 없는 정도, 우울한 정도, 잠을 설치는 정도, 무기력감을 느끼는 정도, 좌절감을 느끼는 정도, 모욕적 및 악의적인 이야기의 경험이나 신체적 폭력을 행사 및 위협당한 경험의 빈도, 가족생활에 대한 만족도, 배우자와의 관계에 대한 만족도, 자녀와의 관계에 대한 만족도, 자녀들의 형제자매 관계에 대한 만족도, 남성 대비 여성의 자살 생각이, 교육 수준이 자살 생각을 1회 이상 해보는 것과 유의한 관계가 있다고 볼 수 있다.

〈표 5-3〉 자살 생각 1회 이상을 결과변수의 사건으로 정의 시, 기초 통계량 및 집단 간 비교 결과(자살 생각 없음 vs. 자살 생각 1회 이상)

변수	전체	자살 생각 없음	자살 생각 1회 이상	P-값	검정 방법	
	4,463명	3,892명 (87.2%)	571명 (12.8%)			
경상소득 (mean (SD))	4811.50 ±3333.39	4997.46 ±3381.26	3543.91 ±2662.98	<0.001	Mann- Whitney U test	
<b>경상소득 (median[IQR])</b>	4249.00 [2612.00, 6103.00]	4395.00 [2793.00, 6254.00]	2954.00 [1606.00, 4974.00]			
tot_debt_p (mean (SD))	213.56 ±1188.62	210.36 ±1135.16	234.59 ±1495.09	0.851	Mann- Whitney U test	
tot_debt_p (median[IQR])	13.00 [4.00, 43.00]	13.00 [4.00, 43.00]	14.00 [5.00, 36.00]			
근로시간 (mean (SD))	46.93 ±14.05	46.97 ±13.76	46.59 ±16.51	0.865	Mann- Whitney U test	
근로시간 (median[IQR])	45.00 [40.00, 54.00]	45.00 [40.00, 54.00]	45.00 [40.00, 55.25]			
year (mean (SD))	59.44 ±12.22	59.00 ±12.17	62.44 ±12.14	<0.001	Mann- Whitney U test	
<b>year (median[IQR])</b>	60.00 [50.00, 70.00]	59.00 [50.00, 69.00]	64.00 [53.00, 73.00]			
저소득 (%)	일반 가구	3565(79.9)	3212(82.5)	353(61.8)	<0.001	Chi- square test
	저소득	898(20.1)	680(17.5)	218(38.2)		
건강 상태 (%)	건강	2997(71.8)	2721(74.7)	276(52.0)	<0.001	Chi- square test
	보통	744(17.8)	624(17.1)	120(22.6)		
	not 건강	432(10.4)	297(8.2)	135(25.4)		
만성 질환 (%)	비해당	2486(59.6)	2250(61.8)	236(44.4)	<0.001	Chi- square test
	6개월 미만	229(5.5)	204(5.6)	25(4.7)		
	6개월 이상	1458(34.9)	1188(32.6)	270(50.8)		
근로 계약 (%)	계약직	603(24.6)	516(23.4)	87(34.4)	<0.001	Chi- square test
	비계약직	1851(75.4)	1685(76.6)	166(65.6)		

변수		전체	자살 생각 없음	자살 생각 1회 이상	P-값	검정 방법
		4,463명	3,892명 (87.2%)	571명 (12.8%)		
근로 지속 가능성 (%)	지속 가능	1792(73.0)	1654(75.1)	138(54.5)	<0.001	Chi- square test
	지속 불가능	662(27.0)	547(24.9)	115(45.5)		
건강 보험료 미납 경험 (%)	없음	4214(98.9)	3746(99.0)	468(98.1)	0.15	Chi- square test
	있음	48(1.1)	39(1.0)	9(1.9)		
집 점유 (%)	기타	1739(39.0)	1443(37.1)	296(51.8)	<0.001	Chi- square test
	점유	2724(61.0)	2449(62.9)	275(48.2)		
tot _debt (%)	없음	2286(51.3)	2003(51.5)	283(49.6)	0.427	Chi- square test
	있음	2171(48.7)	1884(48.5)	287(50.4)		
미취업자 (%)	미해당	2910(65.2)	2595(66.7)	315(55.2)	<0.001	Chi- square test
	해당	1553(34.8)	1297(33.3)	256(44.8)		
실업 (%)	없음	2795(88.3)	2491(88.4)	304(88.1)	0.963	Chi- square test
	있음	369(11.7)	328(11.6)	41(11.9)		
건강 만족도 (%)	만족	2672(59.9)	2438(62.6)	234(41.0)	<0.001	Chi- square test
	불만족	781(17.5)	591(15.2)	190(33.3)		
	보통	1010(22.6)	863(22.2)	147(25.7)		
가족의 수입 만족도 (%)	만족	1063(23.8)	980(25.2)	83(14.5)	<0.001	Chi- square test
	불만족	1917(43.0)	1587(40.8)	330(57.8)		
	보통	1483(33.2)	1325(34.0)	158(27.7)		
주거 환경 만족도 (%)	만족	2766(62.0)	2472(63.5)	294(51.5)	<0.001	Chi- square test
	불만족	595(13.3)	492(12.6)	103(18.0)		
	보통	1102(24.7)	928(23.8)	174(30.5)		
가족 관계 만족도 (%)	만족	3728(83.5)	3329(85.5)	399(69.9)	<0.001	Chi- square test
	불만족	166(3.7)	112(2.9)	54(9.5)		
	보통	569(12.7)	451(11.6)	118(20.7)		

변수		전체	자살 생각 없음	자살 생각 1회 이상	P-값	검정 방법
		4,463명	3,892명 (87.2%)	571명 (12.8%)		
직업 만족도 (%)	만족	2347(52.6)	2137(54.9)	210(36.8)	<0.001	Chi- square test
	불만족	759(17.0)	595(15.3)	164(28.7)		
	보통	1357(30.4)	1160(29.8)	197(34.5)		
사회적 친분관계 만족도 (%)	만족	3401(76.2)	3028(77.8)	373(65.3)	<0.001	Chi- square test
	불만족	153(3.4)	114(2.9)	39(6.8)		
	보통	909(20.4)	750(19.3)	159(27.8)		
여가생활 만족도 (%)	만족	1743(39.1)	1591(40.9)	152(26.6)	<0.001	Chi- square test
	불만족	1148(25.7)	943(24.2)	205(35.9)		
	보통	1571(35.2)	1357(34.9)	214(37.5)		
전반적 만족도 (%)	만족	2627(58.9)	2395(61.5)	232(40.6)	<0.001	Chi- square test
	불만족	335(7.5)	233(6.0)	102(17.9)		
	보통	1501(33.6)	1264(32.5)	237(41.5)		
가부장적 관념 (%)	그렇지 않다	2787(62.4)	2452(63.0)	335(58.7)	0.005	Chi- square test
	보통	531(11.9)	440(11.3)	91(15.9)		
	그렇다	1145(25.7)	1000(25.7)	145(25.4)		
smoke (%)	비흡연	3571(80.0)	3132(80.5)	439(76.9)	0.052	Chi- square test
	흡연	892(20.0)	760(19.5)	132(23.1)		
drink (%)	월 1회 이하	2545(57.0)	2187(56.2)	358(62.7)	0.013	Chi- square test
	월 2~4회	1019(22.8)	904(23.2)	115(20.1)		
	주 2회 이상	899(20.1)	801(20.6)	98(17.2)		
식욕 (%)	주 1일 미만	3717(83.3)	3348(86.0)	369(64.6)	<0.001	Chi- square test
	주 1-4일	703(15.8)	517(13.3)	186(32.6)		
	주 5일 이상	43(1.0)	27(0.7)	16(2.8)		

변수		전체	자살 생각 없음	자살 생각 1회 이상	P-값	검정 방법
		4,463명	3,892명 (87.2%)	571명 (12.8%)		
우울 (%)	주 1일 미만	3409(76.4)	3126(80.3)	283(49.6)	<0.001	Chi- square test
	주 1-4일	1003(22.5)	732(18.8)	271(47.5)		
	주 5일 이상	51(1.1)	34(0.9)	17(3.0)		
불면 (%)	주 1일 미만	3144(70.4)	2878(73.9)	266(46.6)	<0.001	Chi- square test
	주 1-4일	1188(26.6)	921(23.7)	267(46.8)		
	주 5일 이상	131(2.9)	93(2.4)	38(6.7)		
무기력 (%)	주 1일 미만	3789(84.9)	3409(87.6)	380(66.5)	<0.001	Chi- square test
	주 1-4일	622(13.9)	454(11.7)	168(29.4)		
	주 5일 이상	52(1.2)	29(0.7)	23(4.0)		
좌절감 (%)	주 1일 미만	3486(78.1)	3171(81.5)	315(55.2)	<0.001	Chi- square test
	주 1-4일	937(21.0)	699(18.0)	238(41.7)		
	주 5일 이상	40(0.9)	22(0.6)	18(3.2)		
bad (%)	없음	3767(84.4)	3313(85.1)	454(79.5)	0.001	Chi- square test
	있음	696(15.6)	579(14.9)	117(20.5)		
가족 생활 만족도 (%)	만족	3463(77.6)	3133(80.5)	330(57.8)	<0.001	Chi- square test
	보통	756(16.9)	611(15.7)	145(25.4)		
	불만족	209(4.7)	125(3.2)	84(14.7)		
	비해당	35(0.8)	23(0.6)	12(2.1)		
배우자 관계 만족도 (%)	만족	2534(56.8)	2321(59.6)	213(37.3)	<0.001	Chi- square test
	보통	590(13.2)	488(12.5)	102(17.9)		
	불만족	229(5.1)	170(4.4)	59(10.3)		
	비해당	1110(24.9)	913(23.5)	197(34.5)		

변수		전체	자살 생각 없음	자살 생각 1회 이상	P-값	검정 방법
		4,463명	3,892명 (87.2%)	571명 (12.8%)		
자녀 관계 만족도 (%)	만족	3238(72.6)	2882(74.0)	356(62.3)	<0.001	Chi- square test
	보통	365(8.2)	287(7.4)	78(13.7)		
	불만족	111(2.5)	64(1.6)	47(8.2)		
	비해당	749(16.8)	659(16.9)	90(15.8)		
자녀들의 형제자매 관계 만족도 (%)	만족	2692(60.3)	2390(61.4)	302(52.9)	<0.001	Chi- square test
	보통	336(7.5)	269(6.9)	67(11.7)		
	불만족	86(1.9)	54(1.4)	32(5.6)		
	비해당	1349(30.2)	1179(30.3)	170(29.8)		
성별 (%)	남	1849(41.6)	1664(42.9)	185(32.4)	<0.001	Chi- square test
	여	2599(58.4)	2213(57.1)	386(67.6)		
edu level (%)	대졸 이상	1568(35.1)	1446(37.2)	122(21.4)	<0.001	Chi- square test
	고졸 이하	2895(64.9)	2446(62.8)	449(78.6)		

〈표 5-3〉과 〈표5-4〉를 비교하면 tot\_debt\_p, 근로시간, 건강보험료 미납 경험, tot\_debt, 그리고 실업 변수는 그룹 간 차이가 동일하게 유의성이 발생하지 않았다. 하지만 1회 이상 때 유의했던 bad 변수는 유의하지 않게 나타났다.

〈표 5-4〉 자살 생각 2회 이상을 결과변수의 사건으로 정의 시, 기초 통계량 및 집단 간 비교 결과(자살 생각 1회 이하 vs. 자살 생각 2회 이상)

변수	전체	자살 생각 1회 이하	자살 생각 2회 이상	P-값	검정 방법	
	4463명	4276명 (95.8%)	187명(4.2%)			
경상소득 (mean (SD))	4811.50 ±3333.39	4896.80 ±3351.76	2860.97 ±2087.14	<0.001	Mann- Whitney U test	
<b>경상소득 (median[IQR])</b>	4249.00 [2612.00, 6103.00]	4328.00 [2696.50, 6160.00]	2108.00 [1348.00, 4159.50]			
tot_debt_p (mean (SD))	213.56 ±1188.62	214.23 ±1202.24	197.90 ±806.02	0.451	Mann- Whitney U test	
tot_debt_p (median[IQR])	13.00 [4.00, 43.00]	13.00 [4.00, 43.00]	14.50 [5.00, 45.50]			
근로시간 (mean (SD))	46.93 ±14.05	46.92 ±13.99	47.40 ±16.22	0.988	Mann- Whitney U test	
근로시간 (median[IQR])	45.00 [40.00, 54.00]	45.00 [40.00, 54.00]	42.00 [40.00, 59.00]			
year (mean (SD))	59.44 ±12.22	59.27 ±12.21	63.44 ±11.70	<0.001	Mann- Whitney U test	
<b>year (median[IQR])</b>	60.00 [50.00, 70.00]	60.00 [50.00, 70.00]	66.00 [56.00, 73.00]			
저소득 (%)	일반 가구	3565(79.9)	3473(81.2)	92(49.2)	<0.001	Chi- square test
	저소득	898(20.1)	803(18.8)	95(50.8)		
건강 상태 (%)	건강	2997(71.8)	2926(73.2)	71(40.6)	<0.001	Chi- square test
	보통	744(17.8)	701(17.5)	43(24.6)		
	not 건강	432(10.4)	371(9.3)	61(34.9)		
만성 질환 (%)	비해당	2486(59.6)	2430(60.8)	56(32.0)	<0.001	Chi- square test
	6개월 미만	229(5.5)	222(5.6)	7(4.0)		
	6개월 이상	1458(34.9)	1346(33.7)	112(64.0)		
근로 계약 (%)	계약직	603(24.6)	573(24.1)	30(41.1)	0.001	Chi- square test
	비계약직	1851(75.4)	1808(75.9)	43(58.9)		

202 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

변수		전체	자살 생각 1회 이하	자살 생각 2회 이상	P-값	검정 방법
		4463명	4276명 (95.8%)	187명(4.2%)		
근로 지속 가능성 (%)	지속 가능	1792(73.0)	1757(73.8)	35(47.9)	<0.001	Chi- square test
	지속 불가능	662(27.0)	624(26.2)	38(52.1)		
건강 보험료 미납 경험 (%)	없음	4214(98.9)	4079(98.9)	135(97.1)	0.07	Fisher's exact test
	있음	48(1.1)	44(1.1)	4(2.9)		
집 점유 (%)	기타	1739(39.0)	1627(38.0)	112(59.9)	<0.001	Chi- square test
	점유	2724(61.0)	2649(62.0)	75(40.1)		
tot _debt (%)	없음	2286(51.3)	2187(51.2)	99(52.9)	0.699	Chi- square test
	있음	2171(48.7)	2083(48.8)	88(47.1)		
미 취업자 (%)	미해당	2910(65.2)	2816(65.9)	94(50.3)	<0.001	Chi- square test
	해당	1553(34.8)	1460(34.1)	93(49.7)		
실업 (%)	없음	2795(88.3)	2709(88.4)	86(85.1)	0.391	Chi- square test
	있음	369(11.7)	354(11.6)	15(14.9)		
건강 만족도 (%)	만족	2672(59.9)	2614(61.1)	58(31.0)	<0.001	Chi- square test
	불만족	781(17.5)	704(16.5)	77(41.2)		
	보통	1010(22.6)	958(22.4)	52(27.8)		
가족의 수입 만족도 (%)	만족	1063(23.8)	1044(24.4)	19(10.2)	<0.001	Chi- square test
	불만족	1917(43.0)	1792(41.9)	125(66.8)		
	보통	1483(33.2)	1440(33.7)	43(23.0)		
주거 환경 만족도 (%)	만족	2766(62.0)	2681(62.7)	85(45.5)	<0.001	Chi- square test
	불만족	595(13.3)	548(12.8)	47(25.1)		
	보통	1102(24.7)	1047(24.5)	55(29.4)		
가족 관계 만족도 (%)	만족	3728(83.5)	3618(84.6)	110(58.8)	<0.001	Chi- square test
	불만족	166(3.7)	144(3.4)	22(11.8)		
	보통	569(12.7)	514(12.0)	55(29.4)		



변수		전체	자살 생각 1회 이하	자살 생각 2회 이상	P-값	검정 방법
		4463명	4276명 (95.8%)	187명(4.2%)		
직업 만족도 (%)	만족	2347(52.6)	2304(53.9)	43(23.0)	<0.001	Chi- square test
	불만족	759(17.0)	691(16.2)	68(36.4)		
	보통	1357(30.4)	1281(30.0)	76(40.6)		
사회적 친분관계 만족도 (%)	만족	3401(76.2)	3292(77.0)	109(58.3)	<0.001	Chi- square test
	불만족	153(3.4)	135(3.2)	18(9.6)		
	보통	909(20.4)	849(19.9)	60(32.1)		
여가생활 만족도 (%)	만족	1743(39.1)	1709(40.0)	34(18.2)	<0.001	Chi- square test
	불만족	1148(25.7)	1067(25.0)	81(43.3)		
	보통	1571(35.2)	1499(35.1)	72(38.5)		
전반적 만족도 (%)	만족	2627(58.9)	2578(60.3)	49(26.2)	<0.001	Chi- square test
	불만족	335(7.5)	289(6.8)	46(24.6)		
	보통	1501(33.6)	1409(33.0)	92(49.2)		
가부장적 관념 (%)	그렇지 않다	2787(62.4)	2686(62.8)	101(54.0)	0.011	Chi- square test
	보통	531(11.9)	497(11.6)	34(18.2)		
	그렇다	1145(25.7)	1093(25.6)	52(27.8)		
smoke (%)	비흡연	3571(80.0)	3434(80.3)	137(73.3)	0.023	Chi- square test
	흡연	892(20.0)	842(19.7)	50(26.7)		
drink (%)	월 1회 이하	2545(57.0)	2422(56.6)	123(65.8)	0.043	Chi- square test
	월 2~4회	1019(22.8)	983(23.0)	36(19.3)		
	주 2회 이상	899(20.1)	871(20.4)	28(15.0)		
식욕 (%)	주 1일 미만	3717(83.3)	3611(84.4)	106(56.7)	<0.001	Chi- square test
	주 1-4일	703(15.8)	630(14.7)	73(39.0)		
	주 5일 이상	43(1.0)	35(0.8)	8(4.3)		

변수		전체	자살 생각 1회 이하	자살 생각 2회 이상	P-값	검정 방법
		4463명	4276명 (95.8%)	187명(4.2%)		
우울 (%)	주 1일 미만	3409(76.4)	3339(78.1)	70(37.4)	<0.001	Chi- square test
	주 1-4일	1003(22.5)	894(20.9)	109(58.3)		
	주 5일 이상	51(1.1)	43(1.0)	8(4.3)		
불면 (%)	주 1일 미만	3144(70.4)	3079(72.0)	65(34.8)	<0.001	Chi- square test
	주 1-4일	1188(26.6)	1082(25.3)	106(56.7)		
	주 5일 이상	131(2.9)	115(2.7)	16(8.6)		
무기력 (%)	주 1일 미만	3789(84.9)	3680(86.1)	109(58.3)	<0.001	Chi- square test
	주 1-4일	622(13.9)	558(13.0)	64(34.2)		
	주 5일 이상	52(1.2)	38(0.9)	14(7.5)		
취절감 (%)	주 1일 미만	3486(78.1)	3411(79.8)	75(40.1)	<0.001	Chi- square test
	주 1-4일	937(21.0)	835(19.5)	102(54.5)		
	주 5일 이상	40(0.9)	30(0.7)	10(5.3)		
bad (%)	없음	3767(84.4)	3618(84.6)	149(79.7)	0.086	Chi- square test
	있음	696(15.6)	658(15.4)	38(20.3)		
가족 생활 만족도 (%)	만족	3463(77.6)	3372(78.9)	91(48.7)	<0.001	Chi- square test
	보통	756(16.9)	699(16.3)	57(30.5)		
	불만족	209(4.7)	178(4.2)	31(16.6)		
	비해당	35(0.8)	27(0.6)	8(4.3)		
배우자 관계 만족도 (%)	만족	2534(56.8)	2483(58.1)	51(27.3)	<0.001	Chi- square test
	보통	590(13.2)	557(13.0)	33(17.6)		
	불만족	229(5.1)	212(5.0)	17(9.1)		
	비해당	1110(24.9)	1024(23.9)	86(46.0)		

변수		전체	자살 생각 1회 이하	자살 생각 2회 이상	P-값	검정 방법
		4463명	4276명 (95.8%)	187명(4.2%)		
자녀 관계 만족도 (%)	만족	3238(72.6)	3142(73.5)	96(51.3)	<0.001	Chi- square test
	보통	365(8.2)	332(7.8)	33(17.6)		
	불만족	111(2.5)	92(2.2)	19(10.2)		
	비해당	749(16.8)	710(16.6)	39(20.9)		
자녀들의 형제자매 관계 만족도 (%)	만족	2692(60.3)	2617(61.2)	75(40.1)	<0.001	Chi- square test
	보통	336(7.5)	308(7.2)	28(15.0)		
	불만족	86(1.9)	69(1.6)	17(9.1)		
	비해당	1349(30.2)	1282(30.0)	67(35.8)		
성별 (%)	남	1849(41.6)	1787(41.9)	62(33.2)	0.021	Chi- square test
	여	2599(58.4)	2474(58.1)	125(66.8)		
edu level (%)	대졸 이상	1568(35.1)	1539(36.0)	29(15.5)	<0.001	Chi- square test
	고졸 이하	2895(64.9)	2737(64.0)	158(84.5)		

### 제3절 생존 분석 결과

#### 1. 자살 생각에 관한 생존 모형 적용 결과

생존 모형은 단변수 분석과 다변수 분석으로 나누어 진행하였다. 결과 변수는 층동적 선택을 고려하여 ‘1회 이상 자살 생각’을 한 경우와 ‘2회 이상 자살 생각’을 한 경우로 사건을 나누어서 분석하였다. 분석에 사용된 공변량은 결측 비율을 고려하여 1) 결측 비율이 10% 미만인 총 33개의 공변량을 사용하였고, 2) 결측 비율이 50% 미만인 총 37개의 공변량을 각각 모형에 적합시켰다. 결측 비율이 10~50% 사이였던 4개의 변수

는 근로계약(결측 비율 45.01%), 근로 지속 가능성(45.01%), 실업(29.11%), 근로시간(46.9%)이었다. 결측 비율이 10% 미만인 공변량에 대해서만 Stepwise 변수 선택법을 사용하여 한 번 더 모형을 수립하였다. 해당 모형을 최종 모형으로 선택하고자 한다. 공변량 간의 다중공선성 여부 확인은 VIF 값을 통해 확인하였다.

생존 분석 모형을 적용한 결과를 다음의 순서에 따라 정리하였다.

**가. 결과변수를 '1회 이상의 자살 생각' vs. others로 정의**

- 1) 단변수 분석(univariable analysis)
- 2) 다변수 분석(multivariable analysis)
  - (가) 결측 비율 10% 미만 공변량 사용
    - (1) 모든 변수를 모형에 사용
    - (2) Stepwise 변수선택법 적용
  - (나) 결측 비율 50% 미만 공변량 사용(모든 변수를 모형에 사용)

**나. 결과변수를 '2회 이상의 자살 생각' vs. others로 정의**

- 1) 단변수 분석(univariable analysis)
- 2) 다변수 분석(multivariable analysis)
  - (가) 결측 비율 10% 미만 공변량 사용
    - (1) 모든 변수를 모형에 사용
    - (2) Stepwise 변수선택법 적용
  - (나) 결측 비율 50% 미만 공변량 사용(모든 변수를 모형에 사용)

## 가. 결과변수를 '1회 이상의 자살 생각' vs. others로 정의

### 1) 단변수 분석(univariable analysis)

〈표 5-5〉 1회 이상의 자살 생각에 대한 단변수 생존 분석 결과

변수		응답자 수(%)	위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
경상소득		4463(100%)	0.999	0.999	0.999	<0.001
저소득 (ref. 일반)	저소득	4463(100%)	2.696	2.277	3.192	<0.001
건강상태 (ref. 건강)	보통	4173(93.5%)	1.829	1.476	2.266	<0.001
	건강하지 않음	4173(93.5%)	3.949	3.214	4.852	<0.001
만성질환 (ref. 비해당)	6개월 미만	4173(93.5%)	1.174	0.777	1.772	0.446
	6개월 이상	4173(93.5%)	2.083	1.749	2.48	<0.001
근로계약 (ref. 계약직)	비계약직	2454(54.99%)	0.598	0.461	0.775	<0.001
근로 지속 가능성 (ref. 지속 가능)	지속 불가능	2454(54.99%)	2.387	1.863	3.057	<0.001
건강보험료 미납 경험 (ref. 없음)	있음	4262(95.5%)	1.756	0.908	3.395	0.094
집 점유 (ref. 비점유)	점유	4463(100%)	0.571	0.485	0.673	<0.001
tot_debt (ref. 없음)	있음	4457(99.87%)	1.071	0.909	1.262	0.413
tot_debt_p		4457(99.87%)	1	1	1	0.652
미취업자 (ref. 비해당)	해당	4463(100%)	1.587	1.346	1.872	<0.001
실업 (ref. 없음)	있음	3164(70.89%)	1.02	0.736	1.413	0.905
근로시간		2370(53.1%)	0.998	0.989	1.007	0.682
건강 만족도 (ref. 만족)	보통	4463(100%)	3.09	2.551	3.742	<0.001
	불만족	4463(100%)	1.729	1.407	2.125	<0.001

변수		응답자 수(%)	위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
가족의 수입 만족도 (ref. 만족)	보통	4463(100%)	2.353	1.849	2.993	<0.001
	불만족	4463(100%)	1.394	1.069	1.818	0.014
주거환경 만족도 (ref. 만족)	보통	4463(100%)	1.71	1.366	2.14	<0.001
	불만족	4463(100%)	1.532	1.27	1.848	<0.001
가족관계 만족도 (ref. 만족)	보통	4463(100%)	3.612	2.719	4.8	<0.001
	불만족	4463(100%)	2.065	1.681	2.536	<0.001
직업 만족도 (ref. 만족)	보통	4463(100%)	2.627	2.142	3.223	<0.001
	불만족	4463(100%)	1.676	1.38	2.036	<0.001
사회적 친분 관계 만족도 (ref. 만족)	보통	4463(100%)	2.612	1.878	3.632	<0.001
	불만족	4463(100%)	1.665	1.383	2.005	<0.001
여가생활 만족도 (ref. 만족)	보통	4462(99.98%)	2.176	1.765	2.684	<0.001
	불만족	4462(99.98%)	1.611	1.309	1.984	<0.001
전반적 만족도 (ref. 만족)	보통	4463(100%)	4.049	3.208	5.111	<0.001
	불만족	4463(100%)	1.871	1.561	2.242	<0.001
가부장적 관념 (ref. 그렇지 않다)	보통	4463(100%)	1.471	1.167	1.855	0.001
	그렇다	4463(100%)	1.063	0.875	1.292	0.54
smoke (ref. 비흡연)	흡연	4463(100%)	1.216	1.001	1.477	0.049
drink (ref. 월 1회 이하)	월 2~4회	4463(100%)	0.783	0.634	0.966	0.022
	주 2회 이상	4463(100%)	0.752	0.602	0.941	0.013
식욕 (ref. 주 1일 미만)	주 1-4일	4463(100%)	2.961	2.482	3.532	<0.001
	주 5일 이상	4463(100%)	4.823	2.923	7.956	<0.001
우울 (ref. 주 1일 미만)	주 1-4일	4463(100%)	3.69	3.123	4.359	<0.001
	주 5일 이상	4463(100%)	5.128	3.143	8.366	<0.001
불면 (ref. 주 1일 미만)	주 1-4일	4463(100%)	2.917	2.461	3.457	<0.001
	주 5일 이상	4463(100%)	4.019	2.861	5.647	<0.001
무기력 (ref. 주 1일 미만)	주 1-4일	4463(100%)	3.048	2.542	3.656	<0.001
	주 5일 이상	4463(100%)	6.247	4.1	9.519	<0.001

변수		응답자 수(%)	위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
좌절감 (ref. 주 1일 미만)	주 1-4일	4463(100%)	3.152	2.663	3.73	<0.001
	주 5일 이상	4463(100%)	6.874	4.274	11.054	<0.001
bad (ref. 없음)	있음	4463(100%)	1.449	1.183	1.776	<0.001
가족생활 만족도 (ref. 만족)	보통	4463(100%)	2.14	1.76	2.602	<0.001
	불만족	4463(100%)	5.288	4.161	6.721	<0.001
	비해당	4463(100%)	4.317	2.427	7.68	<0.001
배우자 관계 만족도 (ref. 만족)	보통	4463(100%)	2.181	1.723	2.762	<0.001
	불만족	4463(100%)	3.471	2.602	4.632	<0.001
	비해당	4463(100%)	2.249	1.853	2.73	<0.001
자녀 관계 만족도 (ref. 만족)	보통	4463(100%)	2.047	1.602	2.615	<0.001
	불만족	4463(100%)	4.886	3.604	6.624	<0.001
	비해당	4463(100%)	1.098	0.871	1.383	0.429
자녀들의 형제자매 관계 만족도 (ref. 만족)	보통	4463(100%)	1.885	1.446	2.456	<0.001
	불만족	4463(100%)	4.124	2.865	5.938	<0.001
	비해당	4463(100%)	1.135	0.941	1.37	0.186
성별 (ref. 남자)	여자	4448(99.66%)	1.534	1.288	1.828	<0.001
year		4463(100%)	1.023	1.016	1.03	<0.001
edulevel (ref. 대졸 이상)	고졸 이하	4463(100%)	2.09	1.711	2.553	<0.001

## 2) 다변수 분석(multivariable analysis)

### (가) 결측 비율 10% 미만 공변량 사용

#### (1) 모든 변수를 모형에 사용(N=3974, P=33, C-index=0.728)

다변수 분석은 결측 비율 10% 미만에 해당하는 33개의 모든 변수를 활용하여 결과변수에 대한 효과를 확인한 결과이다. VIF 값을 통해서 다

중공선성은 확인되지 않았다. 유의성을 보이는 변수는 건강상태, 집 점유, 가부장적 관념, 우울, 불면, 무기력, 좌절감, 그리고 성별이 해당된다.

〈표 5-6〉 1회 이상의 자살 생각에 대한 다변수 생존 분석 결과(모든 변수 활용)

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
경상소득		1.000	1.000	1.000	0.106	1.88
저소득 (ref. 일반)	저소득	1.014	0.776	1.324	0.919	1.64
건강상태 (ref. 건강)	보통	1.138	0.859	1.508	0.368	1.62
	건강하지 않음	1.591	1.140	2.222	0.006	2.01
만성질환 (ref. 비해당)	6개월 미만	1.105	0.710	1.719	0.659	1.11
	6개월 이상	1.131	0.883	1.449	0.329	1.74
건강보험료 미납 경험 (ref. 없음)	있음	0.881	0.419	1.851	0.738	1.13
집 점유 (ref. 비점유)	점유	0.775	0.630	0.953	0.016	1.21
tot_debt (ref. 없음)	있음	1.151	0.945	1.403	0.162	1.13
tot_debt_p		1.000	1.000	1.000	0.489	1.07
미취업자 (ref. 비해당)	해당	1.008	0.811	1.254	0.94	1.31
건강 만족도 (ref. 만족)	보통	0.894	0.651	1.227	0.486	2.27
	불만족	0.921	0.699	1.213	0.557	1.74
가족의 수입 만족도 (ref. 만족)	보통	0.932	0.679	1.281	0.665	2.89
	불만족	0.869	0.636	1.186	0.375	2.36
주거환경 만족도 (ref. 만족)	보통	0.860	0.635	1.165	0.33	1.46
	불만족	1.113	0.880	1.407	0.371	1.34



변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
가족관계 만족도 (ref. 만족)	보통	0.768	0.473	1.249	0.288	1.72
	불만족	1.108	0.833	1.473	0.481	1.5
직업 만족도 (ref. 만족)	보통	1.167	0.860	1.584	0.323	1.95
	불만족	1.064	0.827	1.367	0.63	1.67
사회적 친분 관계 만족도 (ref. 만족)	보통	0.782	0.472	1.295	0.339	1.44
	불만족	0.857	0.664	1.106	0.236	1.5
여가생활 만족도 (ref. 만족)	보통	1.134	0.849	1.515	0.395	2.15
	불만족	1.123	0.861	1.467	0.392	1.92
전반적 만족도 (ref. 만족)	보통	1.103	0.712	1.711	0.66	2.7
	불만족	0.959	0.728	1.263	0.764	2.11
<b>가부장적 관념</b> (ref. 그렇지 않다)	보통	1.435	1.089	1.89	<b>0.01</b>	1.18
	그렇다	0.853	0.677	1.075	0.178	1.15
smoke (ref. 비흡연)	흡연	1.323	0.983	1.781	0.065	1.68
drink (ref. 월 1회 이하)	월 2~4회	1.028	0.794	1.331	0.835	1.23
	주 2회 이상	0.950	0.706	1.278	0.735	1.47
식욕 (ref. 주 1일 미만)	주 1-4일	1.220	0.961	1.548	0.103	1.38
	주 5일 이상	1.676	0.833	3.37	0.148	1.12
<b>우울</b> (ref. 주 1일 미만)	주 1-4일	1.703	1.341	2.163	<b>&lt;0.001</b>	1.62
	주 5일 이상	1.808	0.918	3.558	0.087	1.28
<b>불면</b> (ref. 주 1일 미만)	주 1-4일	1.548	1.232	1.945	<b>&lt;0.001</b>	1.48
	주 5일 이상	1.459	0.902	2.362	0.124	1.31
<b>무기력</b> (ref. 주 1일 미만)	주 1-4일	1.119	0.870	1.44	0.38	1.41
	주 5일 이상	1.967	1.049	3.69	<b>0.035</b>	1.3

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
<b>좌절감</b> (ref. 주 1일 미만)	주 1-4일	1.546	1.232	1.94	<0.001	1.4
	주 5일 이상	2.147	1.049	4.392	0.036	1.18
bad (ref. 없음)	있음	1.263	0.981	1.627	0.071	1.26
가족생활 만족도 (ref. 만족)	보통	1.024	0.769	1.364	0.872	1.81
	불만족	1.374	0.852	2.214	0.193	2.49
	비해당	1.698	0.508	5.669	0.39	1.13
배우자 관계 만족도 (ref. 만족)	보통	1.348	0.998	1.822	0.051	1.69
	불만족	1.137	0.751	1.722	0.543	1.88
	비해당	1.188	0.877	1.609	0.267	2.05
자녀 관계 만족도 (ref. 만족)	보통	1.143	0.807	1.62	0.451	1.66
	불만족	1.058	0.574	1.949	0.856	2.12
	비해당	1.021	0.650	1.605	0.927	2.52
자녀들의 형제자매 관계 만족도 (ref. 만족)	보통	0.916	0.637	1.318	0.637	1.55
	불만족	1.300	0.709	2.385	0.396	1.65
	비해당	0.995	0.732	1.352	0.975	2.01
<b>성별</b> (ref. 남자)	여자	1.359	1.028	1.795	0.031	1.96
year		1.001	0.989	1.013	0.87	2.42
edulevel (ref. 대졸 이상)	고졸 이하	1.179	0.902	1.54	0.228	1.47

**(2) Stepwise 변수선택법 적용(N=3974, P=15, C-index=0.727)**

변수선택법으로 Stepwise 방식을 선택한 결과 총 33개의 공변량 중 15개의 변수가 선택되었으며, 33개 공변량을 사용했을 때 유의하였던 변수들은 모두 포함되었으며, 추가로 경상소득과 배우자 관계 만족도 변수들이 포함되었다.

우울 변수를 보면, 주 1일 미만보다 주 1~4일에 해당하는 HR 값이 1.705배로, 우울이 더 자주 일어날수록 자살 생각에 유의한 영향이 높아진다는 것을 의미한다. 하지만 주 5일 이상 우울한 경우에 대해선 유의성을 보이지 않았다. 이는 주 5일 이상(1.1%)에 해당하는 대상자 수가 부족하여 일반화하기 어렵기 때문으로 생각된다. 또한, 무기력 변수도 주 1일 미만보다 주 5일 이상의 HR 값이 2.006배에 해당하며, 이는 무기력이 더 자주 찾아올수록 자살 생각을 할 확률이 더 높아짐을 의미한다.

〈표 5-7〉 1회 이상의 자살 생각에 대한 다변수 생존 분석 결과(변수선택법 적용)

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
경상소득		1.000	1.000	1.000	0.04	1.29
건강상태 (ref. 건강)	보통	1.169	0.916	1.49	0.209	1.21
	건강하지 않음	1.643	1.256	2.148	<0.001	1.3
집 점유 (ref. 비점유)	점유	0.785	0.646	0.954	0.015	1.07
tot_debt (ref. 없음)	있음	1.154	0.953	1.397	0.142	1.05
가부장적 관념 (ref. 그렇지 않다)	보통	1.432	1.096	1.871	0.009	1.11
	그렇다	0.874	0.696	1.097	0.246	1.11
smoke (ref. 비흡연)	흡연	1.305	0.979	1.739	0.069	1.56
식욕 (ref. 주 1일 미만)	주 1-4일	1.216	0.960	1.541	0.105	1.36
	주 5일 이상	1.704	0.856	3.393	0.129	1.09
우울 (ref. 주 1일 미만)	주 1-4일	1.705	1.347	2.159	<0.001	1.58
	주 5일 이상	1.695	0.870	3.302	0.121	1.24

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
<b>불면</b> (ref. 주 1일 미만)	주 1-4일	1.517	1.213	1.898	<0.001	1.43
	주 5일 이상	1.398	0.867	2.254	0.169	1.29
<b>무기력</b> (ref. 주 1일 미만)	주 1-4일	1.148	0.898	1.467	0.271	1.34
	주 5일 이상	2.006	1.094	3.676	0.024	1.2
<b>좌절감</b> (ref. 주 1일 미만)	주 1-4일	1.61	1.297	1.999	<0.001	1.28
	주 5일 이상	2.382	1.198	4.738	0.013	1.09
bad (ref. 없음)	있음	1.256	0.980	1.61	0.072	1.21
<b>배우자 관계 만족도</b> (ref. 만족)	보통	1.383	1.063	1.799	0.016	1.29
	불만족	1.306	0.918	1.858	0.137	1.36
	비해당	1.227	0.948	1.588	0.12	1.48
<b>성별 (ref. 남자)</b>	여자	1.351	1.052	1.736	0.019	1.59
edulevel (ref. 대졸 이상)	고졸 이하	1.211	0.952	1.54	0.118	1.19

**(나) 결측 비율 50% 미만 공변량 사용(모든 변수를 모형에 사용)  
(N=1481, P=37, C-index=0.77)**

건강상태, 근로 지속 가능성, tot\_debt(사채 여부), 사회적 친분 관계 만족도, 여가생활 만족도, 전반적 만족도, smoke(흡연 여부), 우울, 좌절 감에 해당하는 변수들이 유의하였다. 특히 좌절감이 주 5일 이상에 해당하는 HR 값이 주 1일 미만보다 4.585배인 점을 보아 좌절감이 높을수록 자살 생각을 한 번 이상 할 확률이 높아진다고 해석할 수 있다.

결측 비율 10% 미만과 50% 미만의 분석을 비교한 결과, 결과변수를

자살 생각 1회 이상으로 정의하면 건강상태, 우울, 그리고 좌절감 변수들이 두 분석 모두 유의하게 나온다. 반면에 근로 지속 가능성, tot\_debt, 사회적 친분 관계 만족도, 여가생활 만족도, 전반적 만족도, 그리고 흡연 여부 변수들은 50% 미만에서는 유의했던 반면, 10% 미만 분석에서는 유의하지 않게 나온다. 또한 집 점유, 가부장적 개념, 불면, 무기력, 그리고 성별 변수는 결측 비율 50% 미만에서는 유의하지 않았지만, 결측 비율 10% 미만 분석에서는 유의하게 나온 변수이다.

결측 비율 50% 미만으로 확장 시, 근로계약, 근로 지속 가능성, 실업, 근로시간 4개 변수가 추가되면서 유의성을 갖는 변수들의 목록 차이가 발생하게 되었다. 다만 결측 비율 50%까지 확장하게 되면 10% 미만일 때보다 활용할 수 있는 대상자가 약 37% 수준밖에 되지 않기에 서로 동등하게 비교하기에는 무리가 있다.

〈표 5-8〉 1회 이상의 자살 생각에 대한 다변수 생존 분석 결과  
(모든 변수 활용, 결측 비율<50%)

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
경상소득		1.000	1.000	1.001	0.209	1.94
저소득 (ref. 일반)	저소득	1.031	0.565	1.882	0.921	1.54
건강상태 (ref. 건강)	보통	1.172	0.674	2.037	0.575	1.72
	건강하지 않음	2.435	1.156	5.127	0.019	1.93
만성질환 (ref. 비해당)	6개월 미만	0.848	0.335	2.145	0.727	1.29
	6개월 이상	1.396	0.868	2.247	0.169	1.83
근로계약 (ref. 계약직)	비 계약직	1.449	0.879	2.39	0.146	1.89

216 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
<b>근로 지속 가능성</b> (ref. 지속 가능)	지속 불가능	2.247	1.392	3.629	<0.001	2
건강보험료 미납 경험 (ref. 없음)	있음	2.769	0.856	8.955	0.089	1.39
집 점유 (ref. 비점유)	점유	0.838	0.573	1.226	0.363	1.29
<b>tot_debt</b> (ref. 없음)	있음	1.470	1.007	2.146	0.046	1.23
tot_debt_p		1.000	0.999	1.000	0.227	1.25
미취업자 (ref. 비해당)	해당	0.000	0.000	Inf	0.996	1
실업 (ref. 없음)	있음	0.939	0.570	1.546	0.804	1.24
근로시간		1.002	0.989	1.015	0.774	1.26
건강 만족도 (ref. 만족)	보통	0.746	0.395	1.409	0.366	1.94
	불만족	1.110	0.669	1.843	0.686	1.77
가족의 수입 만족도 (ref. 만족)	보통	1.086	0.623	1.892	0.771	2.74
	불만족	0.811	0.475	1.384	0.442	2.25
주거환경 만족도 (ref. 만족)	보통	0.647	0.356	1.174	0.152	1.44
	불만족	1.385	0.900	2.131	0.138	1.48
가족관계 만족도 (ref. 만족)	보통	0.411	0.137	1.23	0.112	1.8
	불만족	1.006	0.557	1.82	0.983	1.87
직업 만족도 (ref. 만족)	보통	0.665	0.364	1.215	0.185	1.68
	불만족	0.827	0.517	1.321	0.426	1.62
<b>사회적 친분 관계 만족도</b> (ref. 만족)	보통	0.046	0.005	0.403	0.005	1.21
	불만족	0.505	0.297	0.856	0.011	1.78
<b>여가생활 만족도</b> (ref. 만족)	보통	2.238	1.305	3.84	0.003	2.46
	불만족	1.550	0.937	2.564	0.088	2.11
<b>전반적 만족도</b> (ref. 만족)	보통	2.458	1.085	5.571	0.031	2.71
	불만족	0.793	0.473	1.33	0.379	2.19

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
가부장적 관념 (ref. 그렇지 않다)	보통	1.272	0.720	2.247	0.407	1.19
	그렇다	0.871	0.558	1.361	0.545	1.18
<b>smoke</b> (ref. 비흡연)	흡연	0.538	0.309	0.938	0.029	1.69
drink (ref. 월 1회 이하)	월 2-4회	0.725	0.463	1.133	0.157	1.3
	주 2회 이상	0.947	0.577	1.554	0.828	1.6
식욕 (ref. 주 1일 미만)	주 1-4일	0.828	0.499	1.374	0.465	1.4
	주 5일 이상	1.747	0.462	6.598	0.411	1.35
<b>우울</b> (ref. 주 1일 미만)	주 1-4일	2.316	1.453	3.693	<0.001	1.84
	주 5일 이상	4.003	0.999	16.04	0.05	1.46
불면 (ref. 주 1일 미만)	주 1-4일	1.402	0.901	2.181	0.134	1.63
	주 5일 이상	1.241	0.444	3.464	0.681	1.32
무기력 (ref. 주 1일 미만)	주 1-4일	1.271	0.746	2.168	0.378	1.85
	주 5일 이상	3.090	0.604	15.802	0.175	1.36
<b>좌절감</b> (ref. 주 1일 미만)	주 1-4일	1.848	1.199	2.849	0.005	1.53
	주 5일 이상	4.585	1.054	19.94	0.042	1.64
bad (ref. 없음)	있음	1.422	0.895	2.259	0.136	1.42
가족생활 만족도 (ref. 만족)	보통	1.410	0.812	2.449	0.222	1.94
	불만족	0.710	0.258	1.954	0.507	2.47
	비해당	0.000	0.000	Inf	0.995	1
배우자 관계 만족도 (ref. 만족)	보통	1.005	0.548	1.842	0.988	1.89
	불만족	0.897	0.410	1.96	0.785	1.74
	비해당	0.827	0.459	1.488	0.526	2.38

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
자녀 관계 만족도 (ref. 만족)	보통	1.432	0.718	2.859	0.308	1.66
	불만족	3.296	0.994	10.933	0.051	2.48
	비해당	1.649	0.731	3.719	0.228	2.82
자녀들의 형제자매 관계 만족도 (ref. 만족)	보통	0.831	0.399	1.731	0.622	1.65
	불만족	1.326	0.417	4.212	0.633	1.99
	비해당	0.912	0.527	1.581	0.743	2.19
성별 (ref. 남자)	여자	1.116	0.697	1.787	0.647	1.91
year		0.987	0.963	1.012	0.294	2.39
edulevel (ref. 대졸 이상)	고졸 이하	1.165	0.752	1.804	0.494	1.55

## 나. 결과변수를 ‘2회 이상의 자살 생각’ vs. others로 정의

### 1) 단변수 분석(univariable analysis)

〈표 5-9〉 2회 이상의 자살 생각에 대한 단변수 생존 분석 결과

변수		응답자 수(%)	위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
경상소득		4463(100%)	0.999	0.999	0.999	<0.001
저소득 (ref. 일반)	저소득	4463(100%)	4.276	3.21	5.696	<0.001
건강상태 (ref. 건강)	보통	4173(93.5%)	2.484	1.701	3.627	<0.001
	건강하지 않음	4173(93.5%)	6.341	4.504	8.929	<0.001
만성질환 (ref. 비해당)	6개월 미만	4173(93.5%)	1.358	0.619	2.98	0.445
	6개월 이상	4173(93.5%)	3.518	2.552	4.848	<0.001



변수		응답자 수(%)	위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
근로계약 (ref. 계약직)	비계약직	2454(54.99%)	0.461	0.289	0.734	0.001
근로 지속 가능성 (ref. 지속 가능)	지속 불가능	2454(54.99%)	2.995	1.892	4.74	<0.001
건강보험료 미납 경험 (ref. 없음)	있음	4262(95.5%)	2.696	0.997	7.288	0.051
집 점유 (ref. 비점유)	점유	4463(100%)	0.418	0.312	0.56	<0.001
tot_debt (ref. 없음)	있음	4457(99.87%)	0.936	0.702	1.247	0.65
tot_debt_p		4457(99.87%)	1.000	1.000	1.000	0.848
미취업자 (ref. 비해당)	해당	4463(100%)	1.885	1.415	2.51	<0.001
실업 (ref. 없음)	있음	3164(70.89%)	1.323	0.764	2.289	0.318
근로시간		2370(53.1%)	1.002	0.985	1.02	0.789
건강 만족도 (ref. 만족)	보통	4463(100%)	4.748	3.377	6.676	<0.001
	불만족	4463(100%)	2.404	1.653	3.495	<0.001
가족의 수입 만족도 (ref. 만족)	보통	4463(100%)	3.743	2.31	6.066	<0.001
	불만족	4463(100%)	1.63	0.95	2.797	0.076
주거환경 만족도 (ref. 만족)	보통	4463(100%)	2.655	1.859	3.791	<0.001
	불만족	4463(100%)	1.647	1.173	2.312	0.004
가족관계 만족도 (ref. 만족)	보통	4463(100%)	4.803	3.039	7.592	<0.001
	불만족	4463(100%)	3.395	2.456	4.692	<0.001
직업 만족도 (ref. 만족)	보통	4463(100%)	5.101	3.482	7.473	<0.001
	불만족	4463(100%)	3.123	2.149	4.54	<0.001
사회적 친분 관계 만족도 (ref. 만족)	보통	4463(100%)	3.911	2.375	6.439	<0.001
	불만족	4463(100%)	2.102	1.534	2.88	<0.001
여가생활 만족도 (ref. 만족)	보통	4462(99.98%)	3.708	2.485	5.535	<0.001
	불만족	4462(99.98%)	2.38	1.583	3.579	<0.001

220 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

변수		응답자 수(%)	위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
전반적 만족도 (ref. 만족)	보통	4463(100%)	7.907	5.287	11.824	<0.001
	불만족	4463(100%)	3.348	2.367	4.735	<0.001
가부장적 관념 (ref. 그렇지 않다)	보통	4463(100%)	1.794	1.216	2.646	0.003
	그렇다	4463(100%)	1.254	0.897	1.752	0.185
smoke (ref. 비흡연)	흡연	4463(100%)	1.48	1.07	2.046	0.018
drink (ref. 월 1회 이하)	월 2-4회	4463(100%)	0.724	0.499	1.05	0.089
	주 2회 이상	4463(100%)	0.638	0.423	0.961	0.032
식욕 (ref. 주 1일 미만)	주 1-4일	4463(100%)	3.787	2.811	5.102	<0.001
	주 5일 이상	4463(100%)	7.182	3.5	14.735	<0.001
우울 (ref. 주 1일 미만)	주 1-4일	4463(100%)	5.553	4.113	7.498	<0.001
	주 5일 이상	4463(100%)	8.274	3.981	17.194	<0.001
불면 (ref. 주 1일 미만)	주 1-4일	4463(100%)	4.477	3.287	6.096	<0.001
	주 5일 이상	4463(100%)	6.347	3.673	10.967	<0.001
무기력 (ref. 주 1일 미만)	주 1-4일	4463(100%)	3.745	2.75	5.099	<0.001
	주 5일 이상	4463(100%)	11.004	6.307	19.199	<0.001
좌절감 (ref. 주 1일 미만)	주 1-4일	4463(100%)	5.315	3.945	7.161	<0.001
	주 5일 이상	4463(100%)	13.276	6.863	25.684	<0.001
bad (ref. 없음)	있음	4463(100%)	1.391	0.974	1.986	0.07
가족생활 만족도 (ref. 만족)	보통	4463(100%)	2.948	2.117	4.105	<0.001
	불만족	4463(100%)	6.049	4.024	9.093	<0.001
	비해당	4463(100%)	9.978	4.843	20.558	<0.001
배우자 관계 만족도 (ref. 만족)	보통	4463(100%)	2.827	1.824	4.38	<0.001
	불만족	4463(100%)	3.803	2.196	6.584	<0.001
	비해당	4463(100%)	3.974	2.811	5.619	<0.001

변수		응답자 수(%)	위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
자녀 관계 만족도 (ref. 만족)	보통	4463(100%)	3.133	2.11	4.653	<0.001
	불만족	4463(100%)	6.378	3.899	10.433	<0.001
	비해당	4463(100%)	1.774	1.223	2.574	0.003
자녀들의 형제자매 관계 만족도 (ref. 만족)	보통	4463(100%)	3.086	2.000	4.764	<0.001
	불만족	4463(100%)	7.969	4.707	13.493	<0.001
	비해당	4463(100%)	1.8	1.295	2.502	<0.001
성별 (ref. 남자)	여자	4448(99.66%)	1.446	1.067	1.961	0.018
year		4463(100%)	1.029	1.017	1.042	<0.001
edulevel (ref. 대졸 이상)	고졸 이하	4463(100%)	3.01	2.026	4.472	<0.001

## 2) 다변수 분석(multivariable analysis)

### (가) 결측 비율 10% 미만 공변량 사용

#### (1) 모든 변수를 모형에 사용(N=3974, P=33, C-index=0.811)

자살 생각이 2회 이상 있었던 것으로 정의한 경우, 만성질환, 가부장적 관념, 우울, 불면, 무기력, 좌절감, 그리고 자녀들의 형제자매 관계 만족도가 유의하였다.

〈표 5-10〉 2회 이상의 자살 생각에 대한 다변수 생존 분석 결과(모든 변수 활용)

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
경상소득		1.000	1.000	1.000	0.162	2.18
저소득 (ref. 일반)	저소득	1.056	0.653	1.708	0.823	1.8
건강상태 (ref. 건강)	보통	1.101	0.667	1.819	0.706	1.64
	건강 하지 않음	1.672	0.944	2.963	0.078	2.1
<b>만성질환</b> (ref. 비해당)	6개월 미만	0.990	0.411	2.384	0.982	1.15
	6개월 이상	1.580	1.010	2.471	0.045	1.66
건강보험료 미납 경험 (ref. 없음)	있음	0.899	0.261	3.090	0.866	1.16
집 점유 (ref. 비점유)	점유	0.713	0.485	1.049	0.086	1.26
tot_debt (ref. 없음)	있음	1.073	0.737	1.563	0.712	1.19
tot_debt_p		1.000	1.000	1.000	0.572	1.08
미취업자 (ref. 비해당)	해당	1.026	0.684	1.538	0.902	1.36
건강 만족도 (ref. 만족)	보통	0.843	0.468	1.517	0.569	2.56
	불만족	1.001	0.602	1.665	0.996	1.91
가족의 수입 만족도 (ref. 만족)	보통	0.789	0.417	1.493	0.466	3.28
	불만족	0.698	0.368	1.324	0.272	2.73
주거환경 만족도 (ref. 만족)	보통	0.948	0.565	1.590	0.84	1.57
	불만족	0.940	0.606	1.460	0.784	1.39
가족관계 만족도 (ref. 만족)	보통	0.702	0.284	1.740	0.445	1.61
	불만족	1.540	0.969	2.448	0.068	1.52
직업 만족도 (ref. 만족)	보통	1.479	0.828	2.642	0.186	2.28
	불만족	1.334	0.815	2.182	0.251	2.01

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
사회적 친분 관계 만족도 (ref. 만족)	보통	0.796	0.337	1.880	0.603	1.61
	불만족	0.760	0.485	1.191	0.232	1.53
여가생활 만족도 (ref. 만족)	보통	1.093	0.610	1.959	0.765	2.73
	불만족	1.150	0.672	1.967	0.611	2.34
전반적 만족도 (ref. 만족)	보통	1.403	0.634	3.107	0.404	3.11
	불만족	1.421	0.834	2.420	0.196	2.4
<b>가부장적 관념</b> (ref. 그렇지 않다)	보통	1.735	1.063	2.833	0.027	1.22
	그렇다	0.833	0.540	1.284	0.408	1.18
smoke (ref. 비흡연)	흡연	1.627	0.945	2.802	0.079	1.82
drink (ref. 월 1회 이하)	월 2~4회	1.071	0.662	1.732	0.779	1.29
	주 2회 이상	0.818	0.458	1.463	0.498	1.49
식욕 (ref. 주 1일 미만)	주 1-4일	1.221	0.806	1.851	0.346	1.35
	주 5일 이상	1.189	0.344	4.109	0.785	1.17
<b>우울</b> (ref. 주 1일 미만)	주 1-4일	2.064	1.334	3.192	0.001	1.59
	주 5일 이상	1.823	0.588	5.651	0.299	1.29
<b>불면</b> (ref. 주 1일 미만)	주 1-4일	2.055	1.337	3.158	0.001	1.54
	주 5일 이상	1.277	0.526	3.101	0.589	1.36
<b>무기력</b> (ref. 주 1일 미만)	주 1-4일	1.008	0.646	1.572	0.972	1.4
	주 5일 이상	4.522	1.921	10.648	<0.001	1.43
<b>좌절감</b> (ref. 주 1일 미만)	주 1-4일	1.856	1.239	2.782	0.003	1.38
	주 5일 이상	2.449	0.778	7.714	0.126	1.33
bad (ref. 없음)	있음	1.574	0.969	2.555	0.067	1.41

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창 요인(VIF, Variance Inflation Factor)
가족생활 만족도 (ref. 만족)	보통	0.682	0.412	1.128	0.136	1.8
	불만족	0.563	0.232	1.367	0.204	2.23
	비해당	3.418	0.893	13.074	0.073	1.37
배우자 관계 만족도 (ref. 만족)	보통	1.433	0.821	2.502	0.206	1.78
	불만족	1.052	0.470	2.355	0.901	1.84
	비해당	1.473	0.850	2.551	0.167	2.38
자녀 관계 만족도 (ref. 만족)	보통	1.161	0.633	2.129	0.63	1.81
	불만족	0.469	0.134	1.639	0.236	1.96
	비해당	1.133	0.523	2.456	0.752	2.74
자녀들의 형제자매 관계 만족도 (ref. 만족)	보통	1.452	0.786	2.682	0.234	1.73
	불만족	3.108	1.087	8.891	0.034	1.65
	비해당	1.231	0.704	2.155	0.466	2.22
성별 (ref. 남자)	여자	1.274	0.755	2.149	0.365	2.03
year		0.996	0.974	1.019	0.742	2.41
edulevel (ref. 대졸 이상)	고졸 이하	1.321	0.787	2.220	0.292	1.41

## (2) Stepwise 변수선택법 적용(N=3974, P=13, C-index=0.802)

변수선택법으로 Stepwise 방식을 선택한 결과, 총 33개의 공변량 중 13개의 변수가 선택되었으며, 자살 생각 1회 이상을 결과변수로 정의했을 때와 비교하여 가부장적 관념, 우울, 불면, 무기력, 좌절감 등의 변수가 동일하게 유의성을 보임을 확인하였다.

자살 생각을 2회 이상으로 정의할 경우도 역시 우울과 무기력에 해당하는 변수를 해석해보면, 우울이 주 1일 미만보다 주 1~4일에 해당하는 HR 값이 2.166배로, 우울이 더 자주 일어날수록 자살 생각과 유의한 영향이 높아진다는 것을 의미한다. 이는 1차보다 더 큰 HR 값을 나타내기

때문에 우울이 2차에 더 큰 영향이 있다는 것도 알 수 있다. 또한 좌절감 변수도 주 1일 미만보다 주 1~4일의 HR이 2.1배, 그리고 주 5일 이상의 HR 값이 3.1배를 나타내며 이 역시 무기력이 자주 찾아올수록 자살 생각을 할 확률이 높아짐을 의미한다.

〈표 5-11〉 2회 이상의 자살 생각에 대한 다변수 생존 분석 결과(변수선택법 적용)

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창요 인(VIF, Variance Inflation Factor)
경상소득		1.000	1.000	1.000	0.01	1.27
건강상태 (ref. 건강)	보통	1.32	0.850	2.049	0.216	1.27
	건강하지 않음	2.029	1.267	3.247	0.003	1.42
집 점유 (ref. 비점유)	점유	0.703	0.493	1.003	0.052	1.07
가족관계 만족도 (ref. 만족)	보통	0.775	0.325	1.846	0.565	1.47
	불만족	1.69	1.090	2.620	0.019	1.37
가부장적 관념 (ref. 그렇지 않다)	보통	1.674	1.043	2.686	0.033	1.14
	그렇다	0.833	0.547	1.270	0.396	1.11
smoke (ref. 비흡연)	흡연	1.65	0.982	2.770	0.058	1.65
우울 (ref. 주 1일 미만)	주 1-4일	2.166	1.414	3.319	<0.001	1.53
	주 5일 이상	1.625	0.534	4.942	0.393	1.25
불면 (ref. 주 1일 미만)	주 1-4일	2.086	1.383	3.148	<0.001	1.41
	주 5일 이상	1.395	0.585	3.330	0.453	1.3
무기력 (ref. 주 1일 미만)	주 1-4일	1.032	0.672	1.584	0.887	1.31
	주 5일 이상	4.376	1.933	9.909	<0.001	1.3
좌절감 (ref. 주 1일 미만)	주 1-4일	2.1	1.418	3.111	<0.001	1.3
	주 5일 이상	3.1	1.060	9.064	0.039	1.16

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값	분산팽창요 인(VIF, Variance Inflation Factor)
bad (ref. 없음)	있음	1.424	0.936	2.169	0.099	1.06
가족생활 만족도 (ref. 만족)	보통	0.882	0.563	1.385	0.586	1.44
	불만족	0.657	0.309	1.400	0.277	1.62
	비해당	5.043	1.495	17.008	0.009	1.13
성별 (ref. 남자)	여자	1.419	0.878	2.293	0.153	1.7

자살 생각은 충동적 사고를 동반할 수 있는 점을 고려하여 1회만 한 경우를 배제한 2회 이상의 사건을 결과변수로 정의한 모형을 최종 모형으로 고려하였다. 또한, 전체 변수들을 모형에 포함시킨 경우와 변수선택법을 적용한 경우를 함께 정리하였다. 생존 모형에서 특정 시점  $t$ 에서 발생할 수 있는 생존확률은 다음과 같이 산출된다.

$$\hat{S}(t|X) = \hat{S}_0(t)^{\exp(\hat{\beta}X)}$$

여기서,  $\hat{S}_0(t)$ 은  $t$  시점에서의 baseline hazard 함수를 의미하며,  $X$ 는 공변량을  $\hat{\beta}$ 은 모형을 통해 추정된 회귀계수를 의미한다. 생존 모형에서는 HR에 log를 취한 값을 의미한다.

자살 생각을 2회 이상 한 경우를 결과변수로 정의하고 공변량의 결측 비율이 10% 미만인 변수들로만 활용하여 생존 모형을 수립한 예측 결과는 다음과 같았다.



〈표 5-12〉 자살 생각 2회 이상에 관한 생존 모형의 예측된 baseline hazard 함수

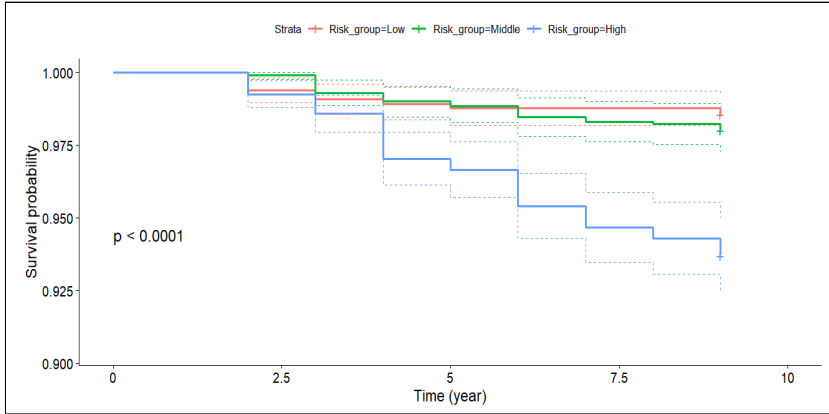
자살 생각 2회 이상 & 전체 변수 활용		자살 생각 2회 이상 & Stepwise 적용	
시간(Time)	기저 위험 (Baseline hazard)	시간(Time)	기저 위험 (Baseline hazard)
2	0.000564	2	0.00105
3	0.001213	3	0.00225
4	0.002074	4	0.00384
5	0.002369	5	0.00437
6	0.003108	6	0.00571
7	0.003522	7	0.00646
8	0.003731	8	0.00684
9	0.004223	9	0.00772

Baseline hazard 함수를 활용하여 Time=5 시점에서의 예측된 생존 확률을 산출할 수 있으며, 이 값들을 정렬하여 삼등분하여 Risk group (Low, Middle, High)인 새로운 변수를 생성한 후, 아래와 같은 표와 그림을 구할 수 있었으며, High level을 갖는 그룹들이 2회 이상의 자살 생각 비율 및 횟수가 다른 그룹에 비해 많다는 것을 확인할 수 있다.

〈표 5-13〉 자살 생각 2회 이상에 관한 생존 모형의 예측된 baseline hazard 함수

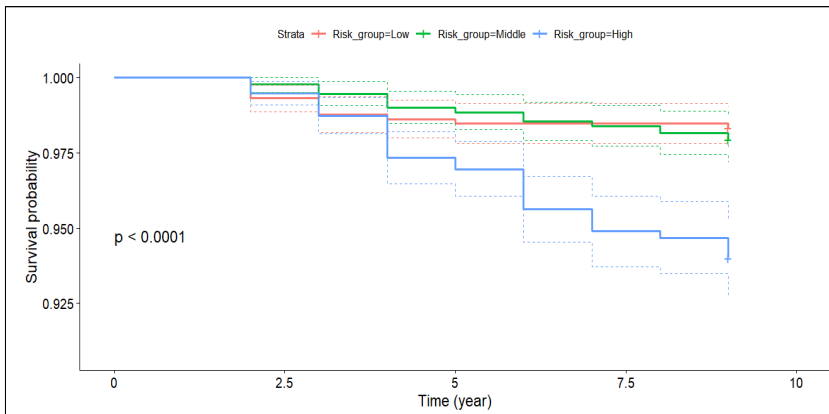
	자살 생각 2회 이상 & 전체 변수 활용			자살 생각 2회 이상 & Stepwise 적용			
	Low	Middle	High	Low	Middle	High	
2회 이상 자살 생각 여부 비율 (total =4.19%)	1.45%	1.98%	6.29%	1.68%	2.06%	6.00%	
자살 생각 횟수 (회)	2	8명	1명	10명	9명	3명	7명
	3	5명	8명	9명	7명	4명	10명
	4	2명	4명	21명	2명	6명	19명
	5	2명	2명	5명	2명	2명	5명
	6		5명	17명		4명	18명
	7		2명	10명		2명	10명
	8		1명	5명		3명	3명
	9	3명	3명	8명	2명	3명	9명

[그림 5-1] 자살 생각 2회 이상의 생존 모형을 통해서 얻은 예측된 생존확률로 구분된 위험 그룹 변수에 대한 KM curve



위험 그룹을 Time=5 시점에서 예측된 확률 값을 통해 나눈 후 KM curve를 작성한 결과, 위험도가 High인 상위 33% 해당하는 그룹의 자살 생각에 관한 확률이 다른 두 그룹에 비해 상대적으로 높음을 알 수 있다.

[그림 5-2] 자살 생각 2회 이상의 생존 모형 with stepwise을 통해서 얻은 예측된 생존확률로 구분된 위험 그룹 변수에 대한 KM curve



**(나) 결측 비율 50% 미만 공변량 사용(모든 변수를 모형에 사용)  
(N=1481, P=37, C-index=0.88)**

자살 생각을 2회 이상 한 것으로 정의한 경우, 경상소득, 건강상태, 만성질환, 근로계약, 근로 지속 가능성, 건강 만족도, 가족관계 만족도, 직업 만족도, 사회적 친분 관계 만족도, 여가생활 만족도, 우울, 불면, 무기력, 좌절감, 가족생활 만족도, 자녀 관계 만족도, 자녀들의 형제자매 관계 만족도, 그리고 edulevel에 해당하는 변수들이 유의하였다. 결측 비율 50% 미만하에서 자살 생각이 1회 이상일 때와 2회 이상일 때 모두 유의한 변수는 건강상태, 근로 지속 가능성, 사회적 친분 관계 만족도, 여가생활 만족도, 우울, 좌절감 변수였다.

특히 근로계약(여부) 변수가 비계약직에 해당하는 HR 값이 계약직의 0.287배인 점을 보아 계약직일수록 자살 생각을 두 번 이상 할 확률이 높아진다고 해석할 수 있다. 결측 비율 10% 미만과 50% 미만의 분석을 비교한 결과, 결과변수를 자살 생각 2회 이상으로 정의할 시, 만성질환, 우울, 불면, 무기력, 좌절감, 자녀들의 형제자매 만족도 변수들이 두 분석 모두 유의하게 나온다.

〈표 5-14〉 2회 이상의 자살 생각에 대한 다변수 생존 분석 결과  
(모든 변수 활용, 결측 비율<50%)

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
경상소득		1.000	1.000	1.000	0.035
저소득 (ref. 일반)	저소득	1.949	0.936	4.056	0.074
건강상태 (ref. 건 강)	보통	1.214	0.586	2.519	0.602
	건강하지 않음	3.336	1.297	8.581	0.012

230 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
만성질환 (ref. 비해당)	6개월 미만	0.000	0.000	Inf	0.997
	6개월 이상	1.965	1.018	3.793	0.044
근로계약 (ref. 계약직)	비계약직	0.287	0.149	0.554	<0.001
근로 지속 가능성 (ref. 지속 가능)	지속 불가능	0.456	0.233	0.891	0.022
건강보험료 미납 경험 (ref. 없음)	있음	5.506	0.748	40.543	0.094
집 점유 (ref. 비점유)	점유	1.459	0.758	2.808	0.258
tot_debt (ref. 없음)	있음	0.825	0.427	1.592	0.566
tot_debt_p		1.000	0.998	1.001	0.588
실업 (ref. 없음)	있음	1.395	0.636	3.062	0.406
근로시간		0.995	0.973	1.017	0.666
건강 만족도 (ref. 만족)	보통	1.478	0.615	3.551	0.383
	불만족	3.978	2.060	7.684	<0.001
가족의 수입 만족도 (ref. 만족)	보통	0.852	0.443	1.639	0.632
	불만족	0.909	0.460	1.796	0.784
주거환경 만족도 (ref. 만족)	보통	0.912	0.400	2.082	0.827
	불만족	1.081	0.547	2.137	0.822
가족관계 만족도 (ref. 만족)	보통	0.595	0.143	2.483	0.477
	불만족	2.546	1.250	5.183	0.01
직업 만족도 (ref. 만족)	보통	2.527	1.187	5.377	0.016
	불만족	1.448	0.740	2.832	0.279
사회적 친분 관계 만족도 (ref. 만족)	보통	0.000	0.000	Inf	0.997
	불만족	0.347	0.167	0.720	0.004
여가생활 만족도 (ref. 만족)	보통	0.919	0.476	1.774	0.801
	불만족	0.306	0.150	0.623	0.001
전반적 만족도 (ref. 만족)	보통	2.163	0.765	6.117	0.146
	불만족	1.378	0.717	2.650	0.336

변수		위험비 (hazard ratio)	95% 신뢰 하한	95% 신뢰 상한	P-값
가부장적 관념 (ref. 그렇지 않다)	보통	2.478	0.963	6.374	0.06
	그렇다	0.903	0.396	2.063	0.81
smoke (ref. 비흡연)	흡연	0.641	0.301	1.363	0.248
drink (ref. 월 1회 이하)	월 2-4회	0.492	0.224	1.082	0.078
	주 2회 이상	1.025	0.494	2.125	0.948
식욕 (ref. 주 1일 미만)	주 1-4일	0.668	0.278	1.605	0.367
	주 5일 이상	0.322	0.044	2.377	0.267
우울 (ref. 주 1일 미만)	주 1-4일	2.198	1.139	4.240	0.019
	주 5일 이상	0.000	0.000	Inf	0.999
불면 (ref. 주 1일 미만)	주 1-4일	2.003	1.041	3.857	0.038
	주 5일 이상	0.000	0.000	Inf	0.998
무기력 (ref. 주 1일 미만)	주 1-4일	1.806	0.870	3.749	0.113
	주 5일 이상	20.253	4.764	86.095	<0.001
좌절감 (ref. 주 1일 미만)	주 1-4일	1.927	0.992	3.742	0.053
	주 5일 이상	13.769	1.886	100.525	0.01
bad (ref. 없음)	있음	2.093	0.954	4.594	0.065
가족생활 만족도 (ref. 만족)	보통	0.392	0.184	0.834	0.015
	불만족	0.040	0.009	0.167	<0.001
	비해당	0.000	0.000	Inf	>0.999
배우자 관계 만족도 (ref. 만족)	보통	0.904	0.376	2.174	0.822
	불만족	1.244	0.299	5.179	0.764
	비해당	0.592	0.300	1.172	0.132
자녀 관계 만족도 (ref. 만족)	보통	2.394	1.047	5.475	0.039
	불만족	1.862	0.255	13.595	0.54
	비해당	4.448	2.026	9.767	<0.001
자녀들의 형제자매 관계 만족도 (ref. 만족)	보통	2.112	0.821	5.432	0.121
	불만족	4.821	1.158	20.073	0.031
	비해당	1.333	0.686	2.589	0.396
성별 (ref. 남자)	여자	0.832	0.432	1.602	0.582
year		0.972	0.943	1.000	0.054
edulevel (ref. 대졸 이상)	고졸 이하	2.079	1.001	4.318	0.049

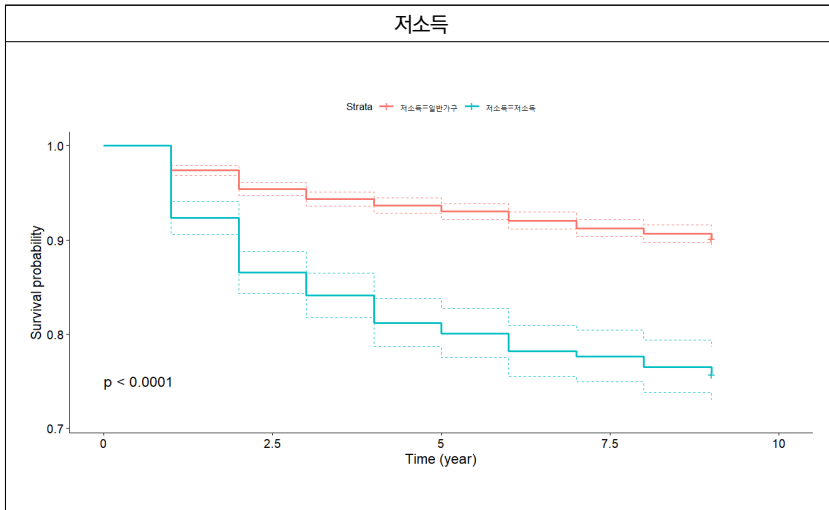
## 2. Kaplan-Meier(KM) curve: 1회 이상의 자살 생각을 결과변수로 정의

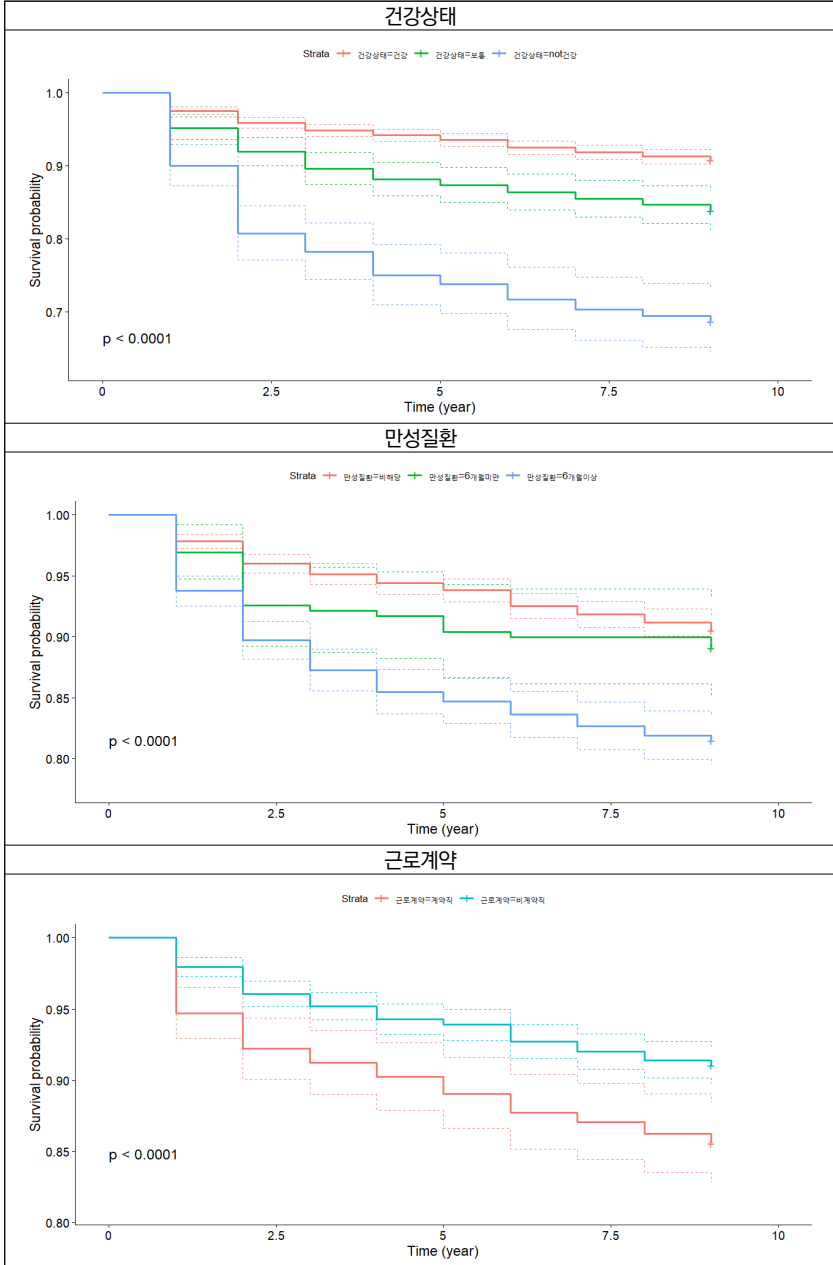
공변량 중 범주형으로 구분된 변수에 대해서만 분석하였다. KM curve는 시간(연)과 생존확률의 관계가 실선으로 표현되며, 95% 신뢰구간은 점선으로 표시되었다. KM curve에서 그룹 간 차이는 log-rank test를 통해 p-value를 산출하였고, 그림 내에 해당 값을 포함시켰다.

### 가. 1회 이상의 자살 생각을 결과변수로 정의 시 KM curve

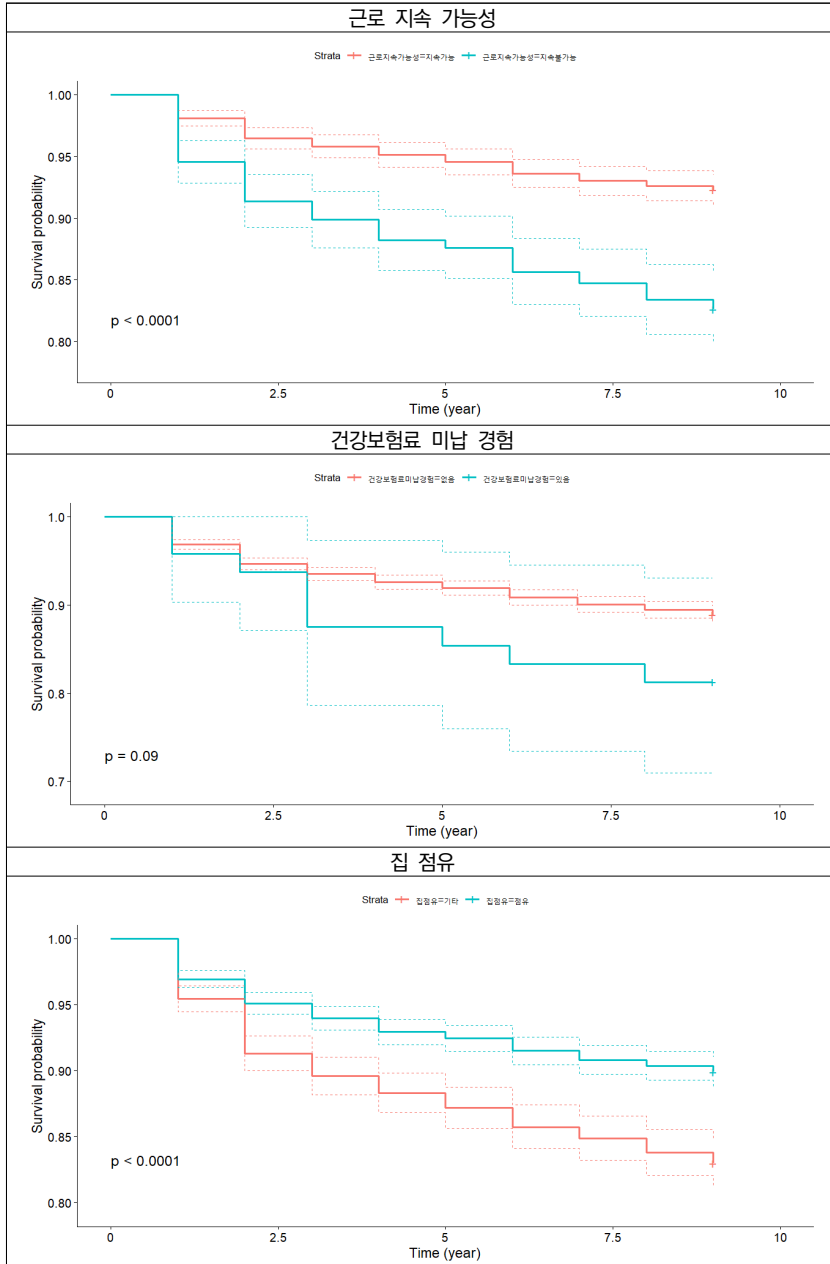
KM curve의 log-rank test를 확인한 결과, 건강보험료 미납 경험, tot\_debt, 그리고 실업 변수들을 제외한 나머지 모든 범주형 변수들이 유의한 결과를 나타냈다.

[그림 5-3] 범주형 변수에 대한 자살 생각 1회 이상에 대한 KM curve

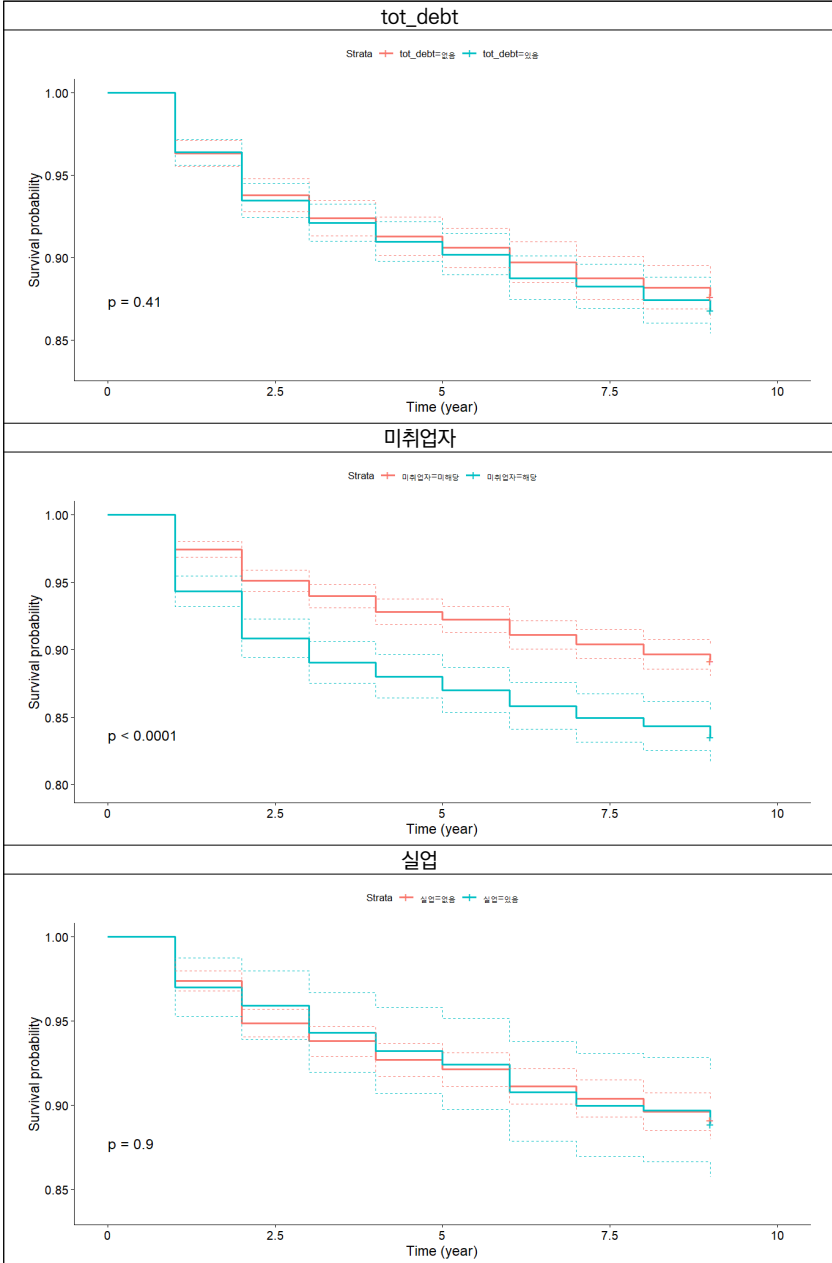


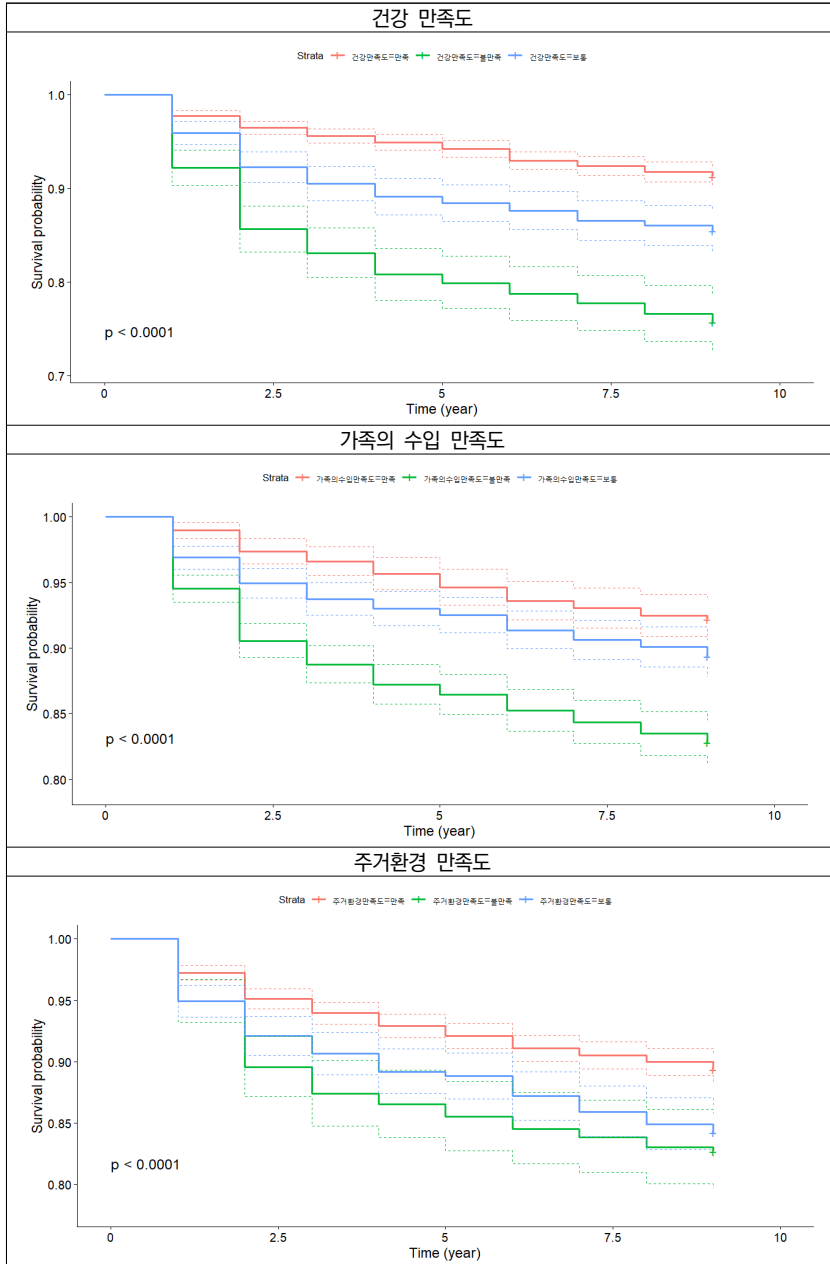


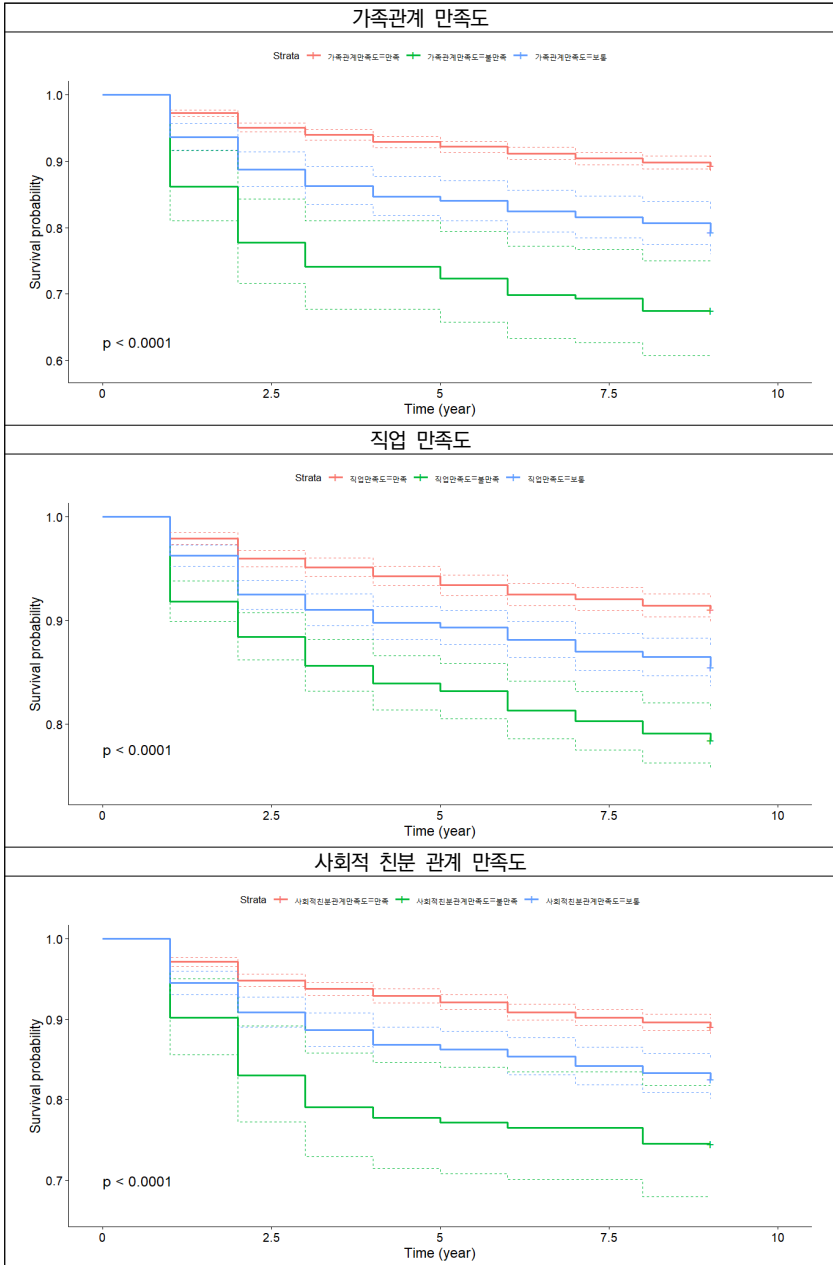
234 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로



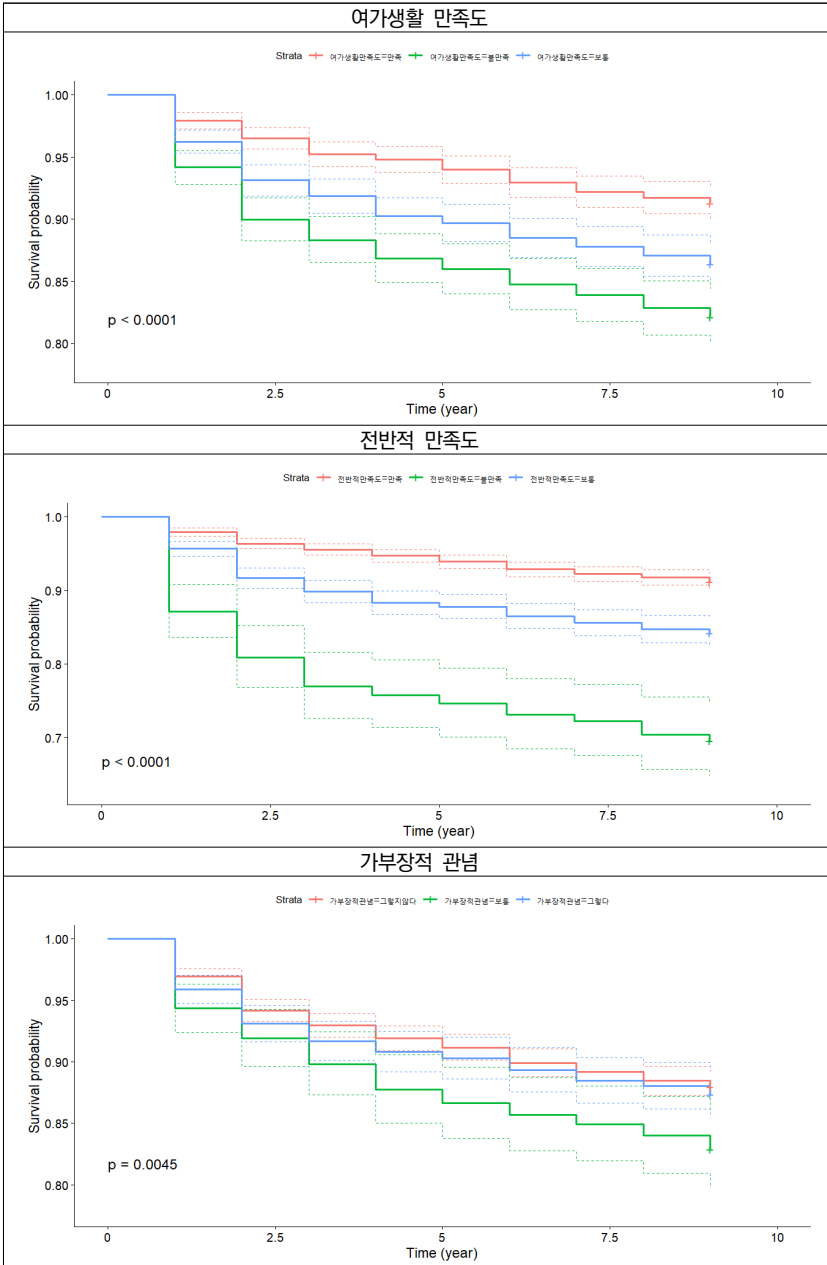


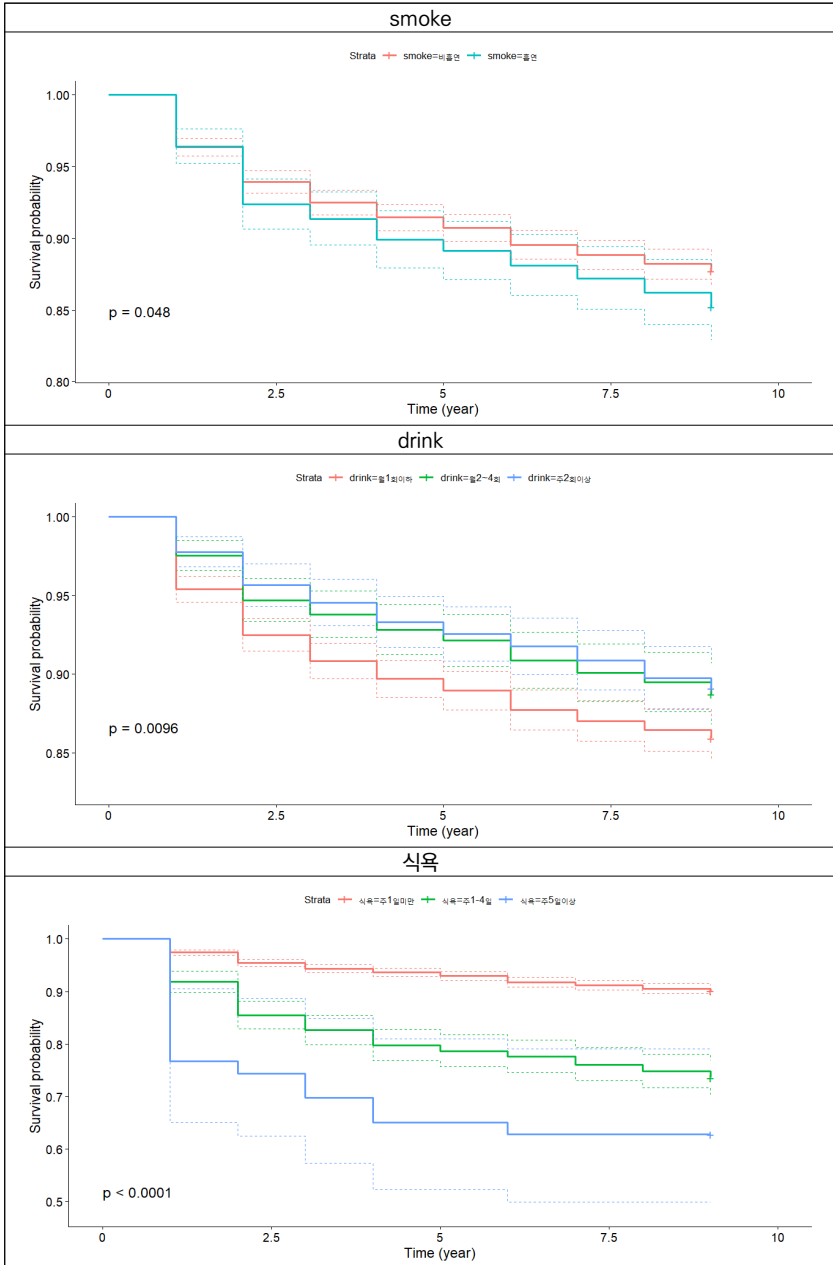




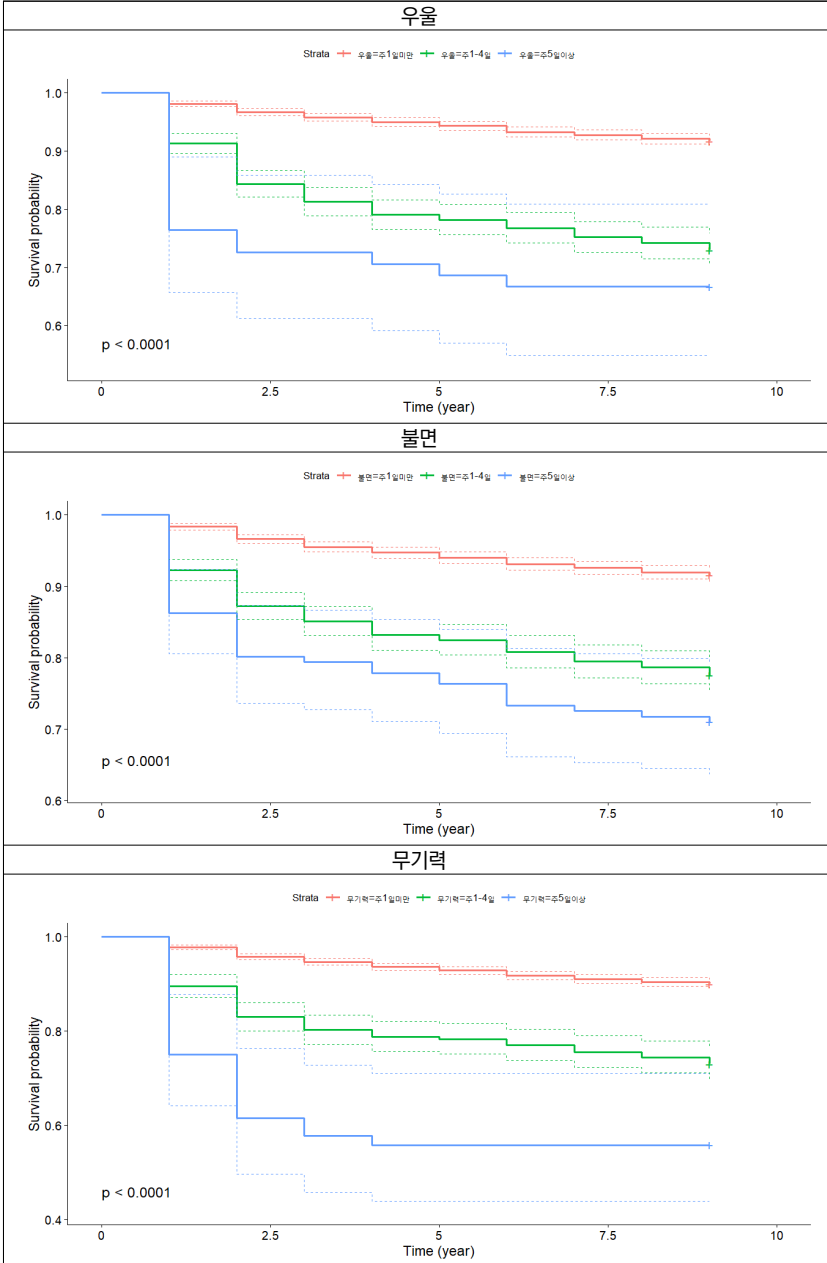


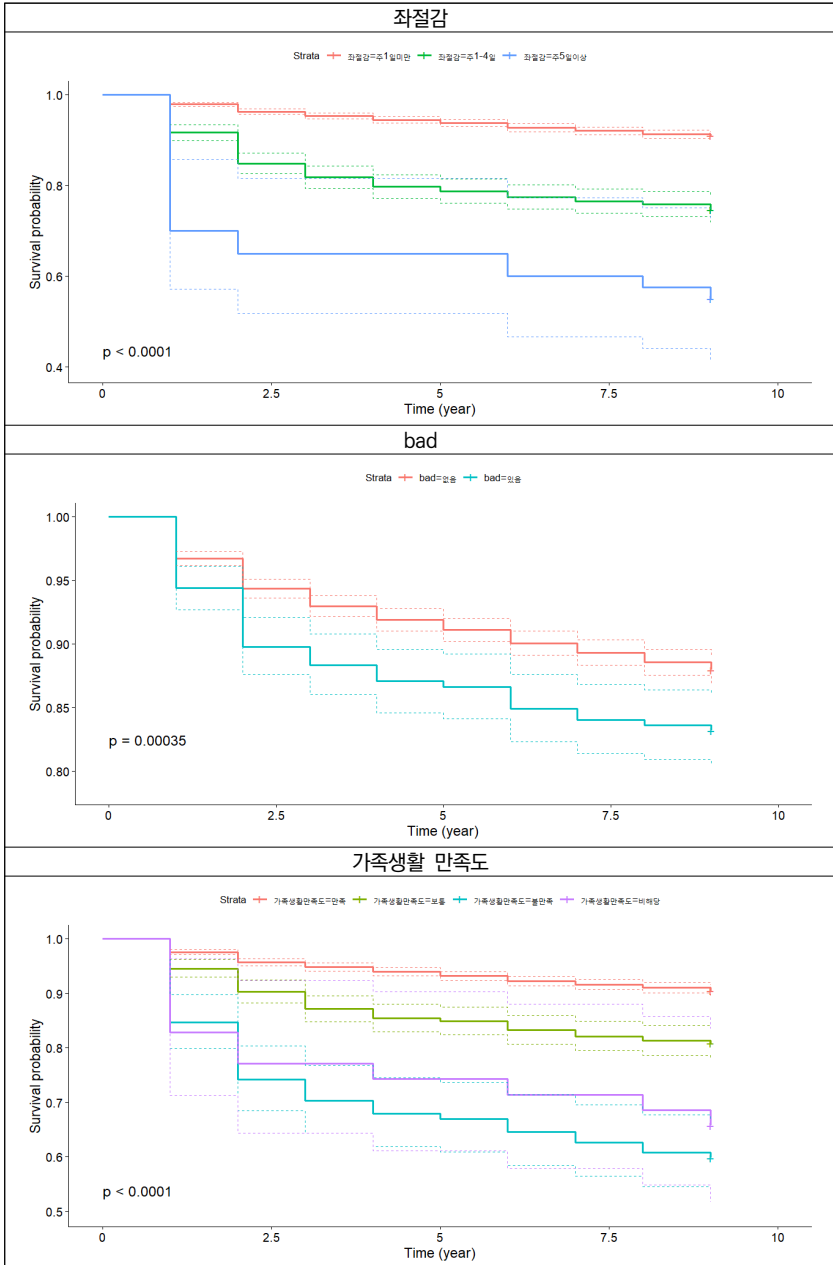
238 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로



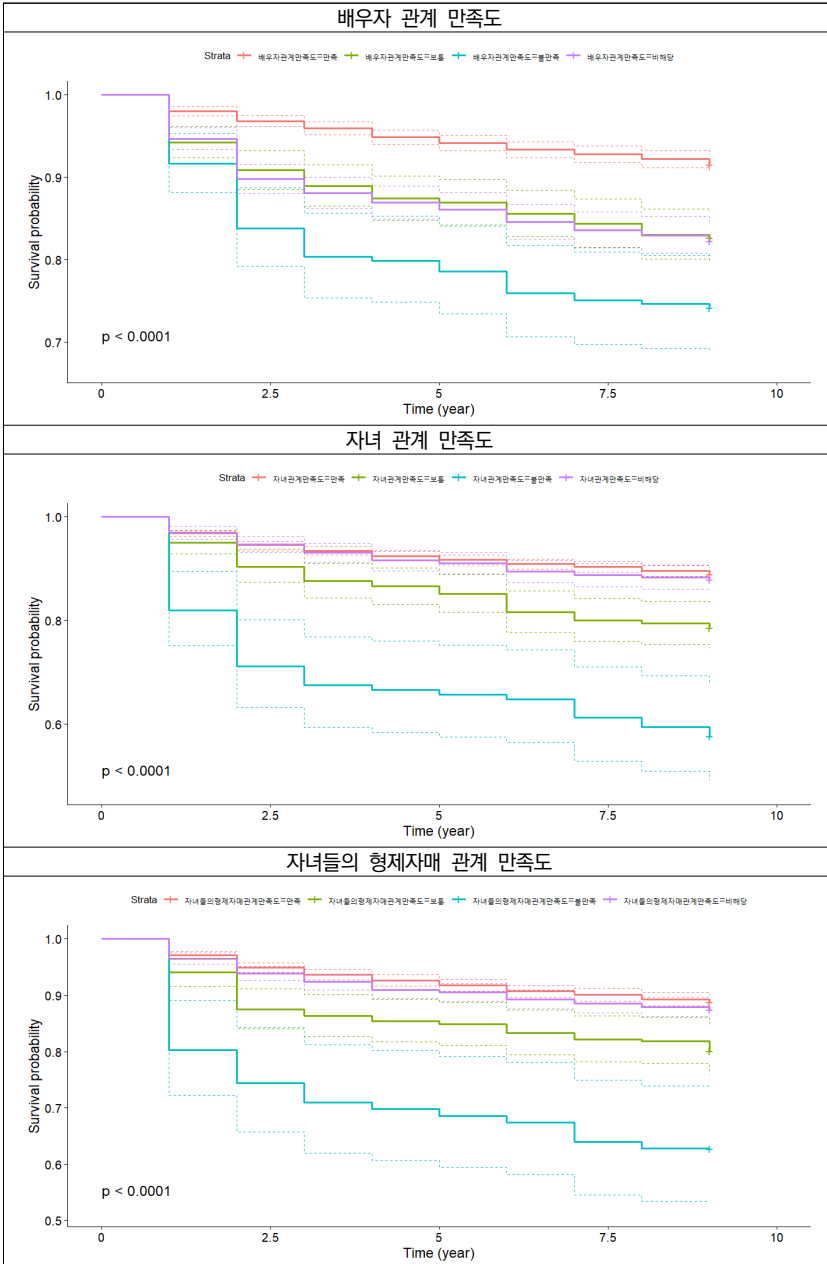


240 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

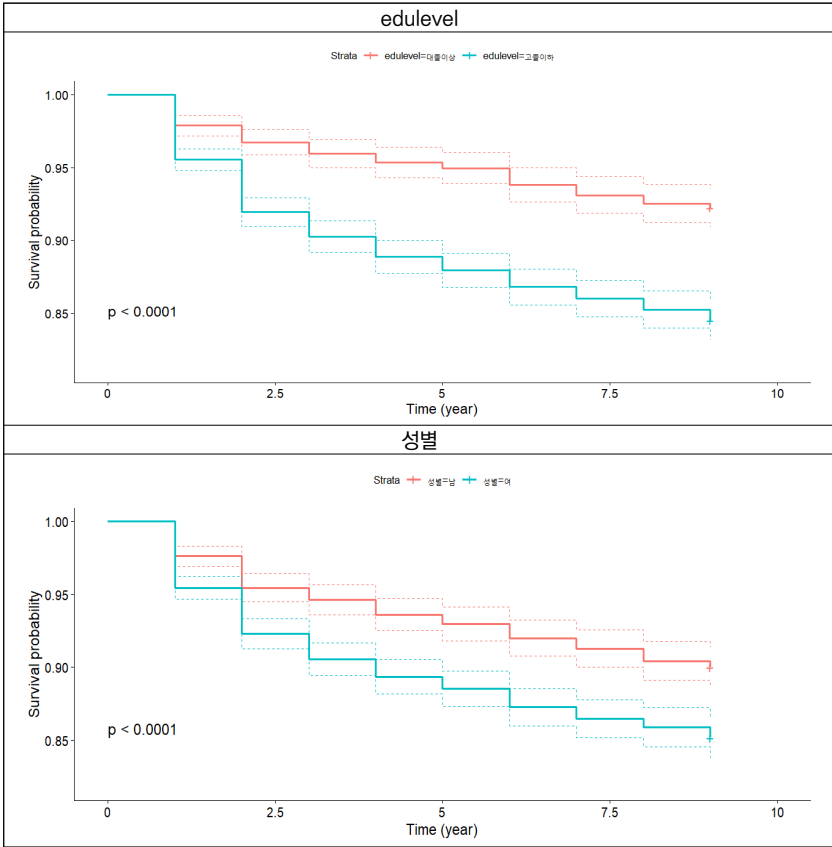




242 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

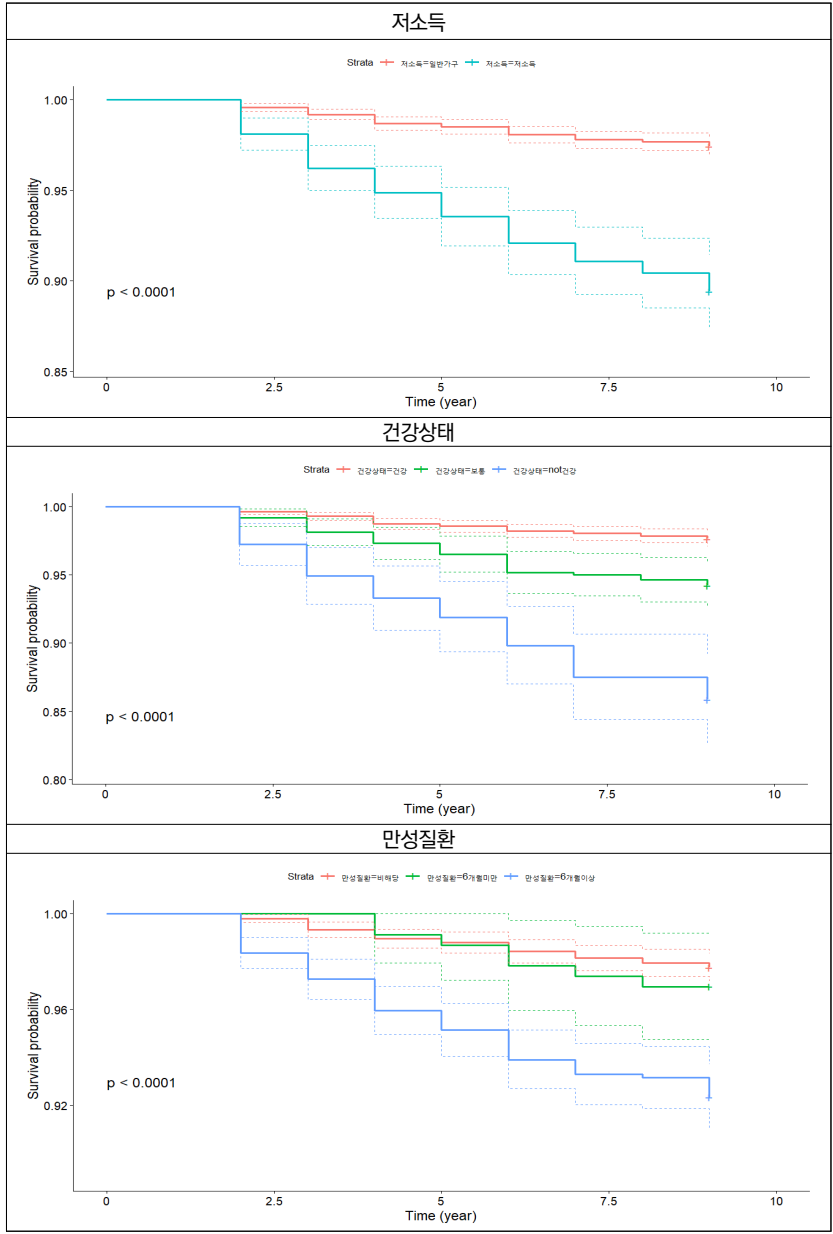


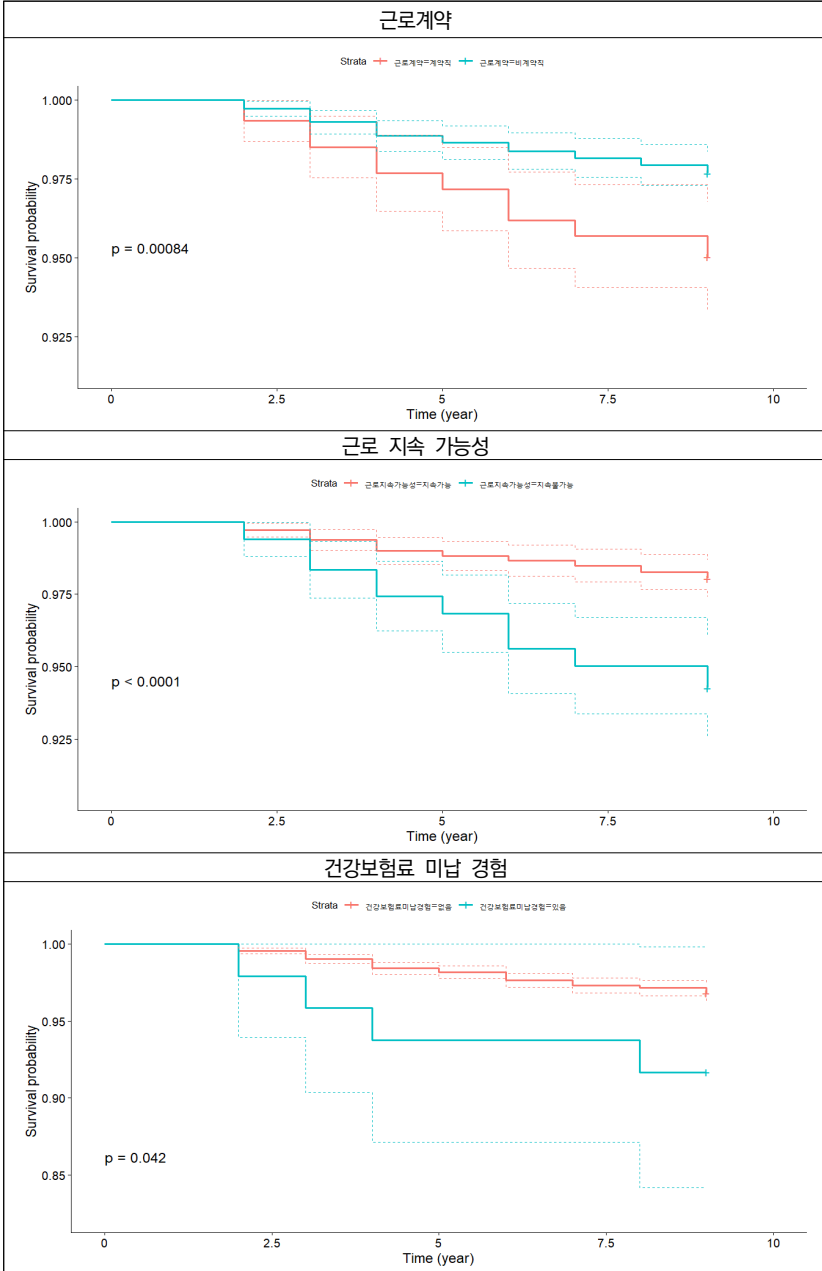


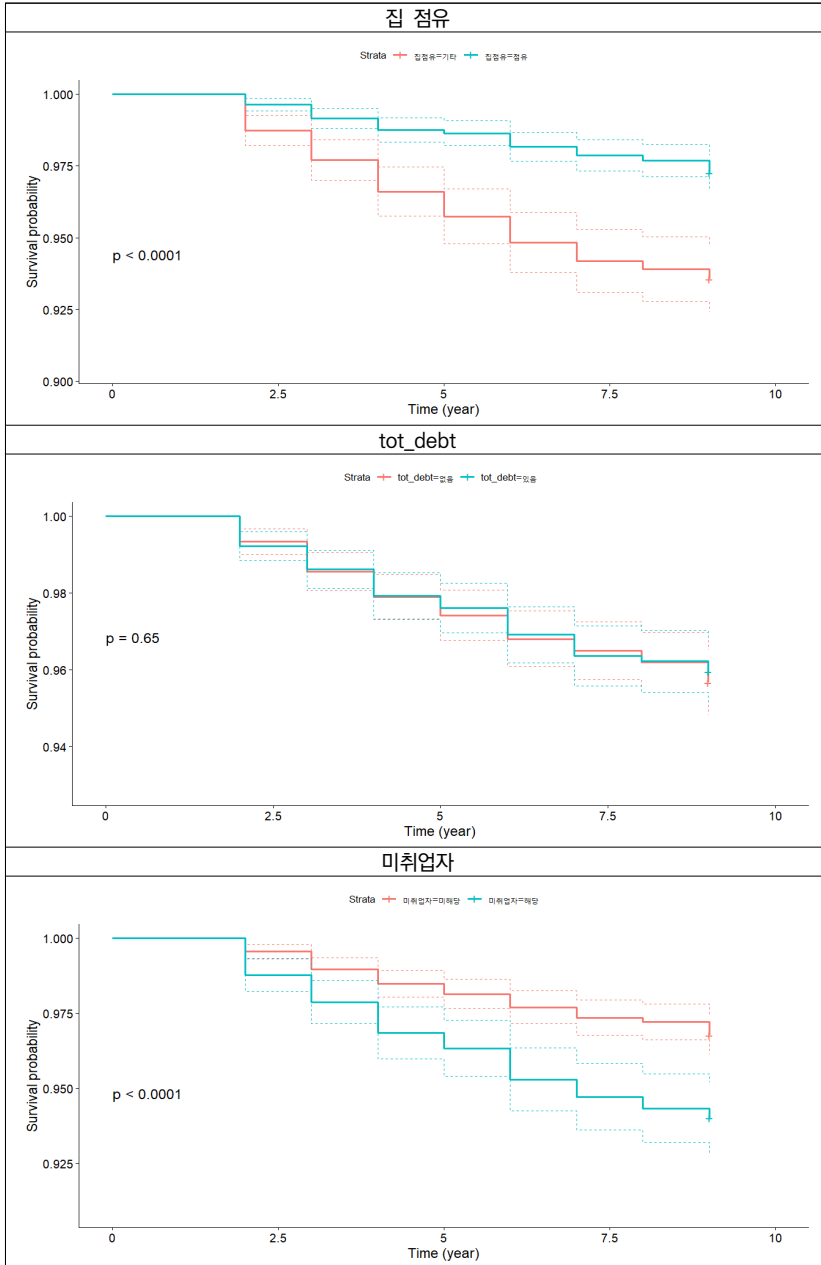


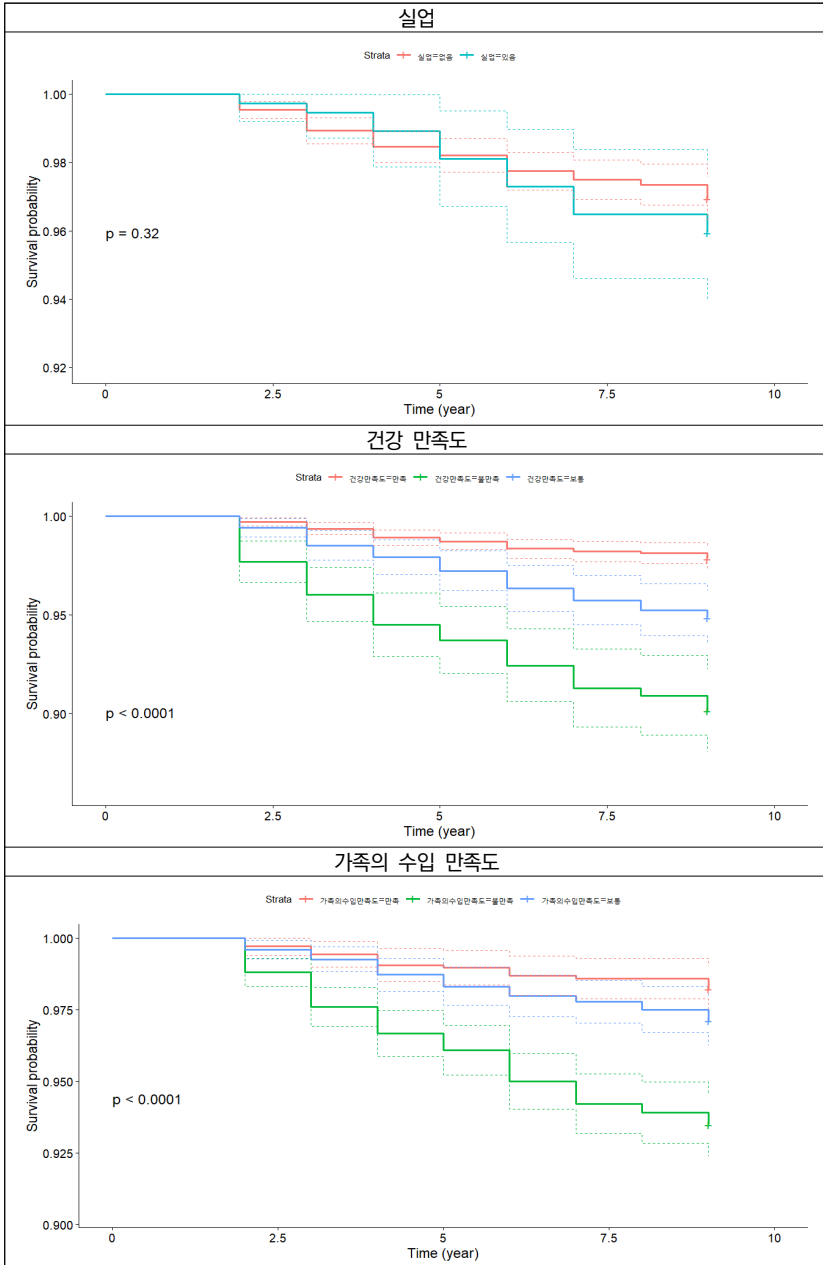
2회 이상의 자살 생각으로 결과변수를 정의하면, 1회 이상에서는 유의하지 않았던 건강보험료 미납 경험 변수는 유의하였고, 유의하였던 bad 변수는 유의하지 않았다. tot\_debt와 실업 변수들은 1회 및 2회에서 모두 유의하지 않았다.

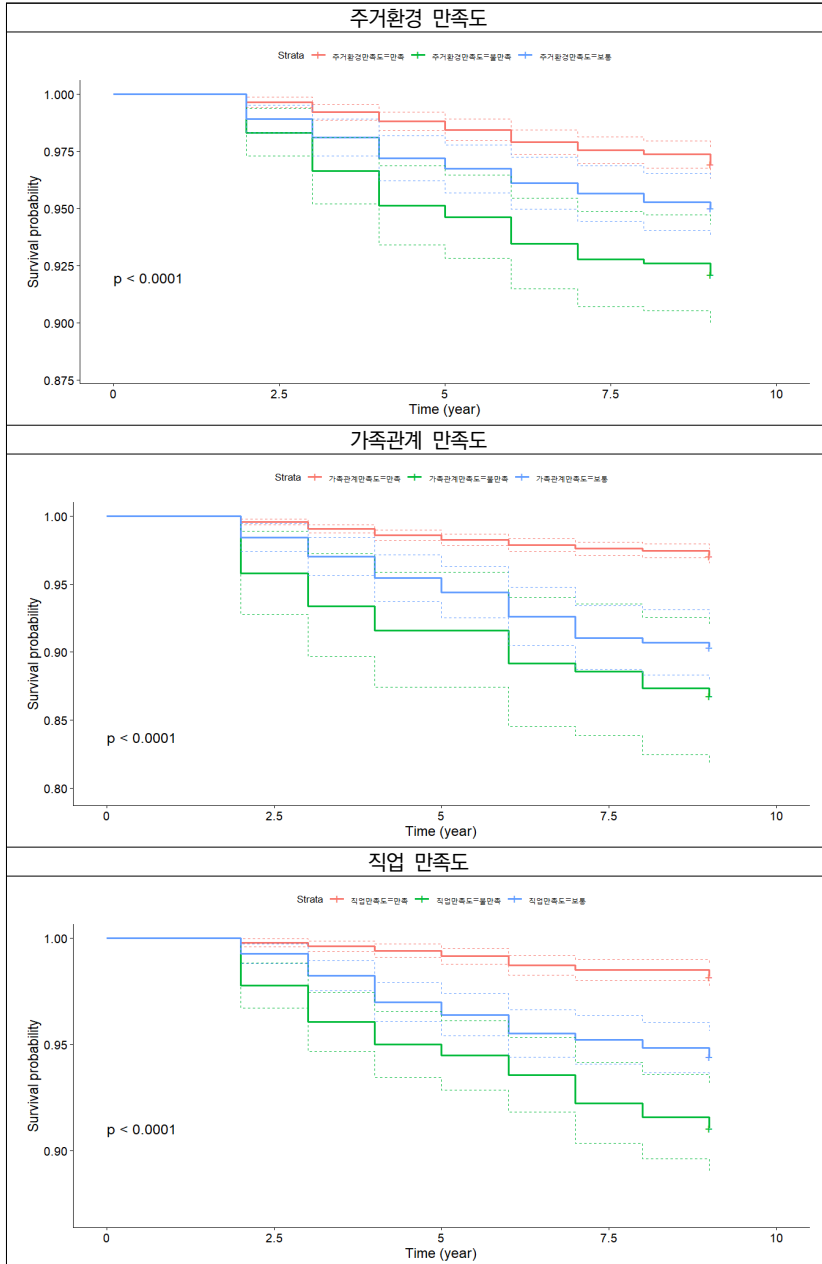
[그림 5-4] 범주형 변수에 대한 자살 생각 2회 이상에 대한 KM curve

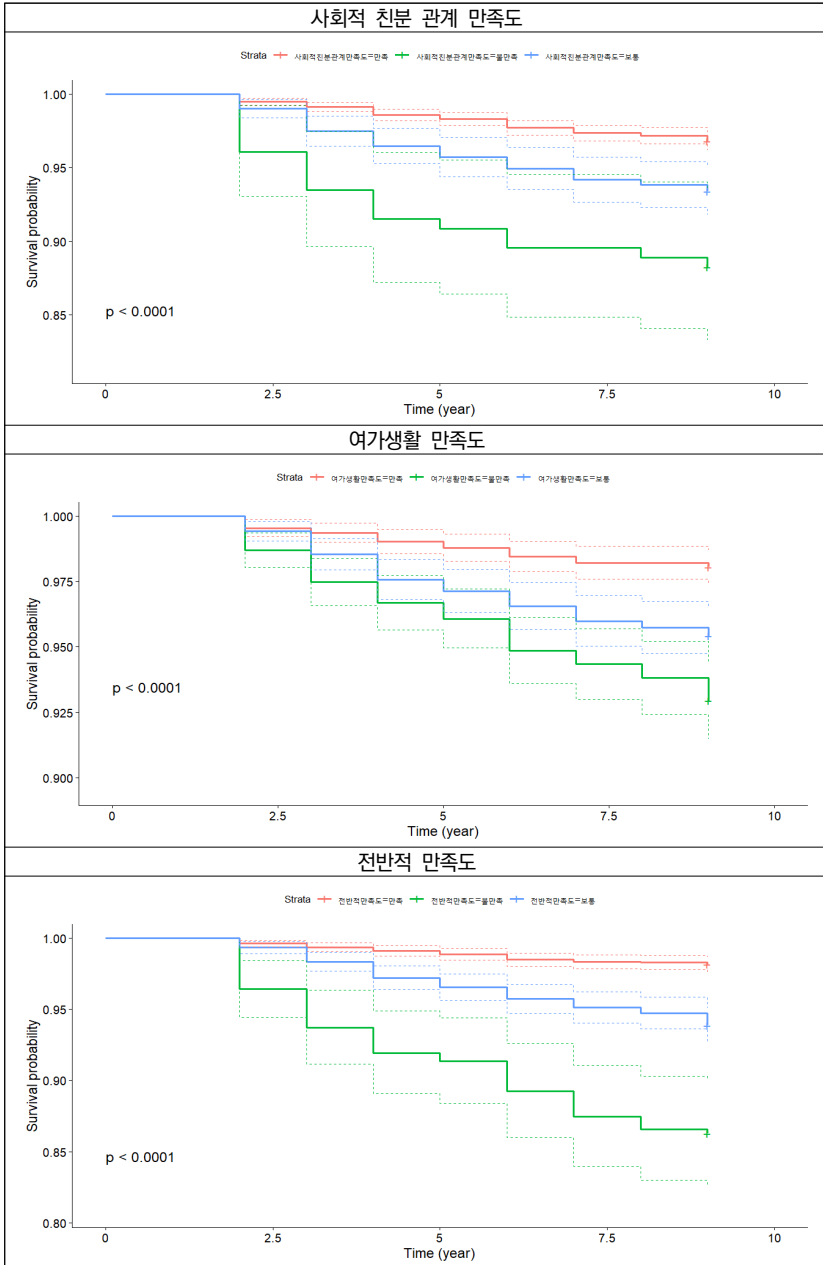




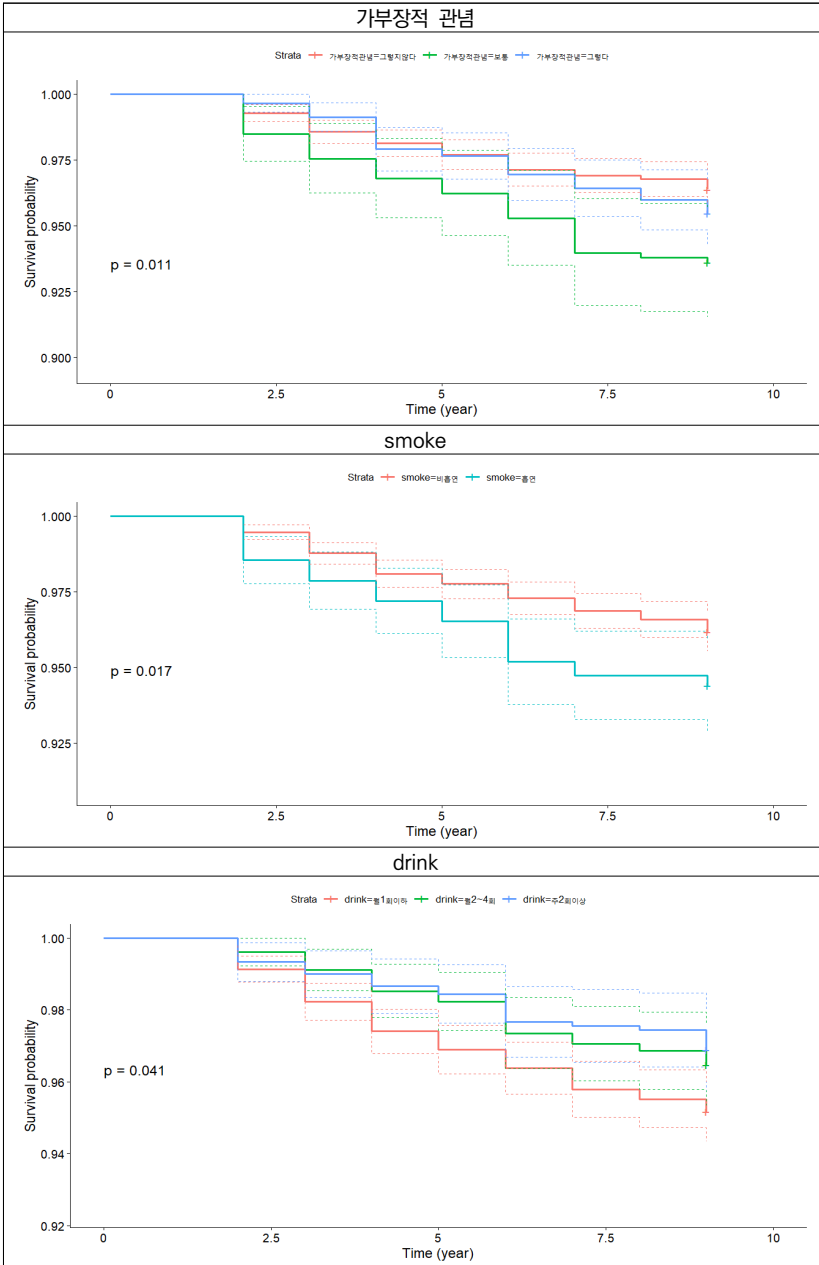






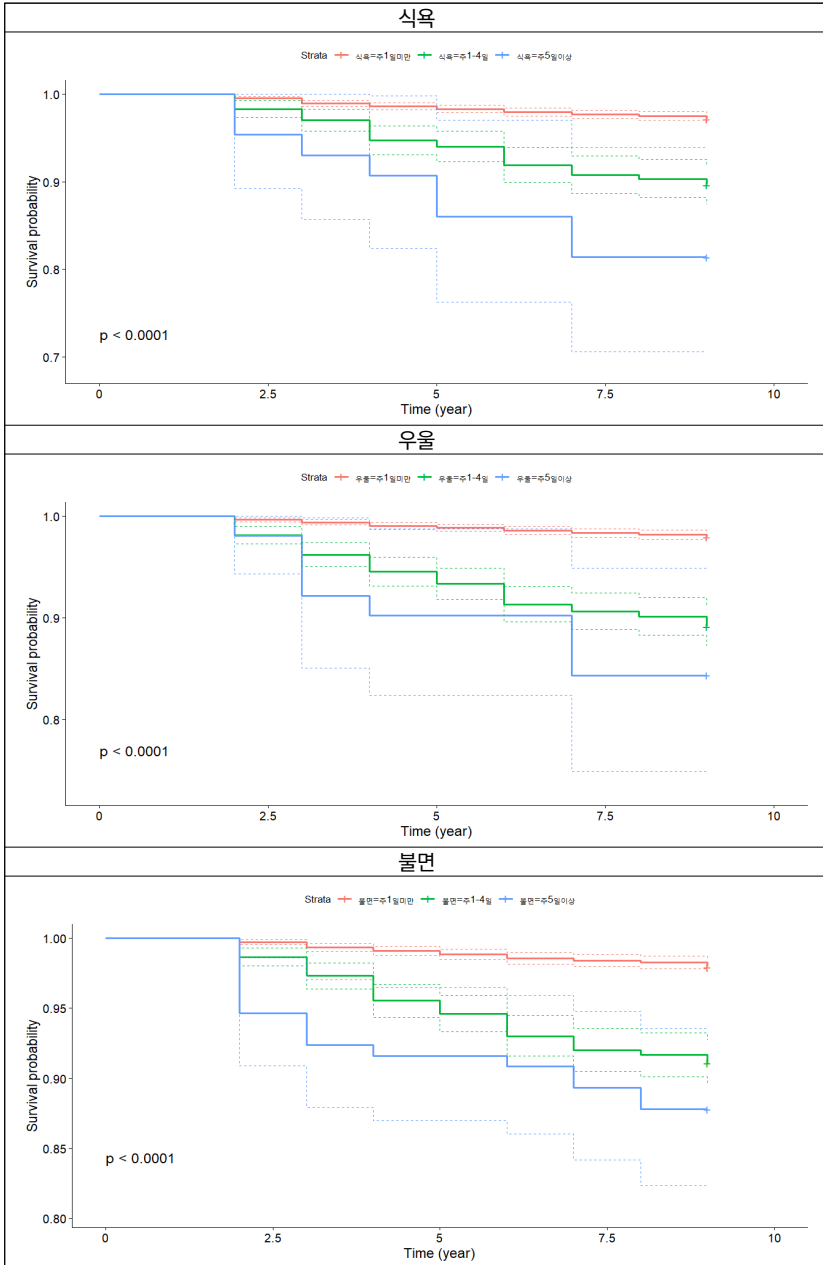


250 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로

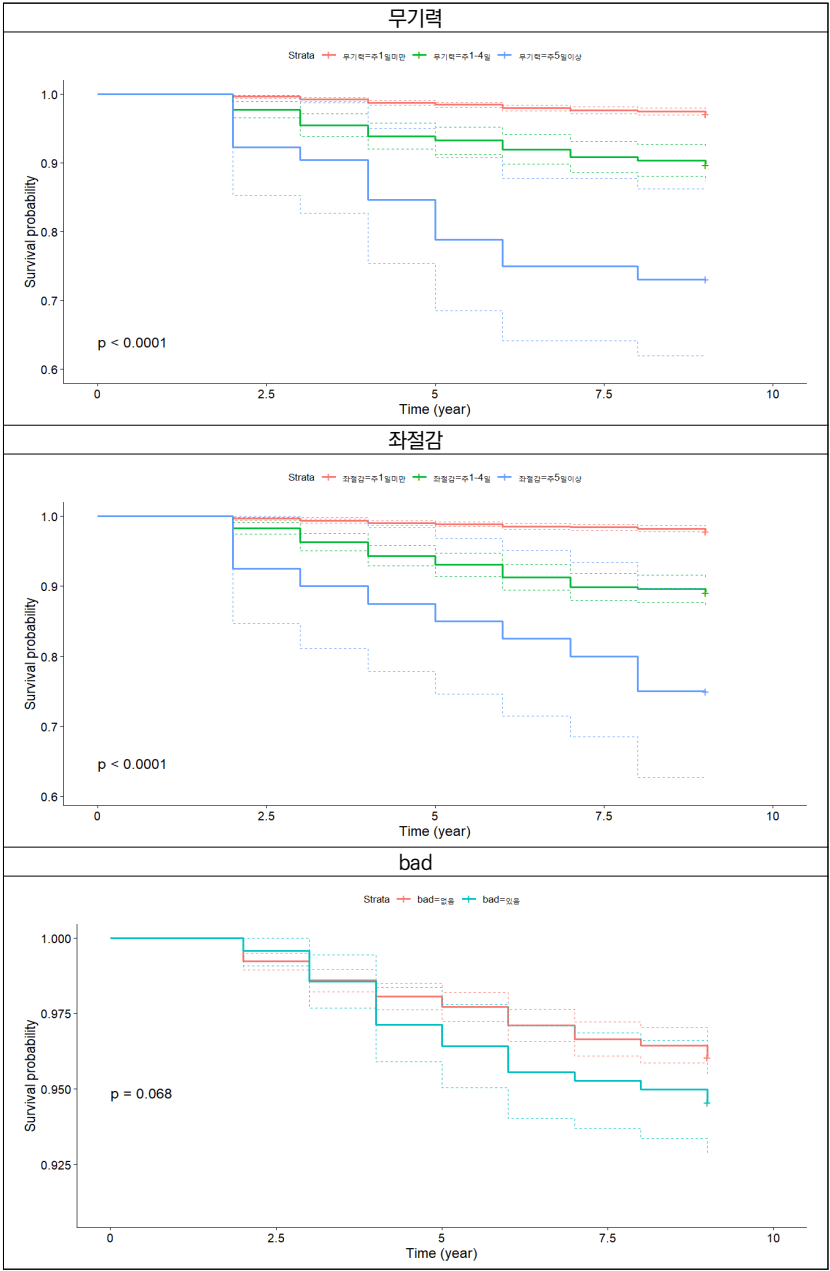


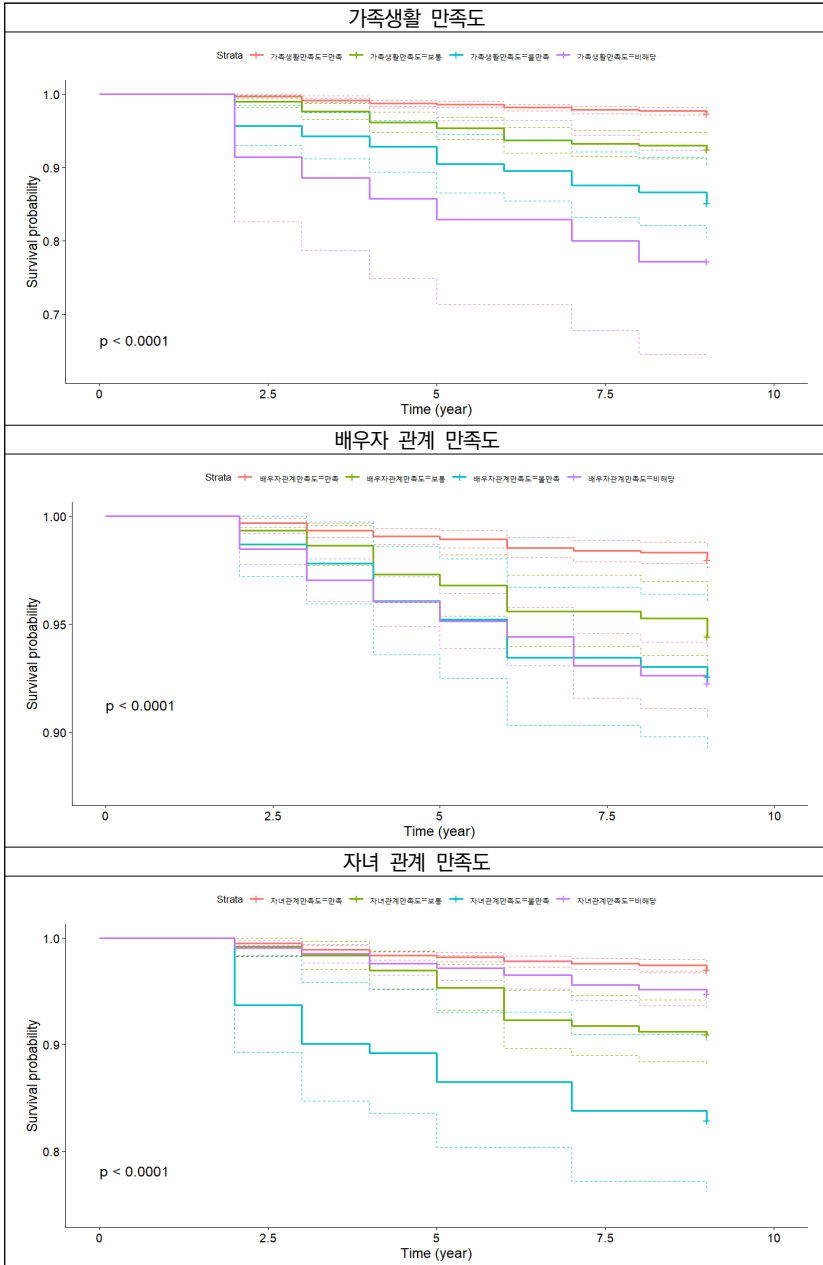


제5장 심층 분석 I : 자살 고위험군에 대한 위험 인자 탐색을 위한 패널데이터 생존 분석 모형 적용 251

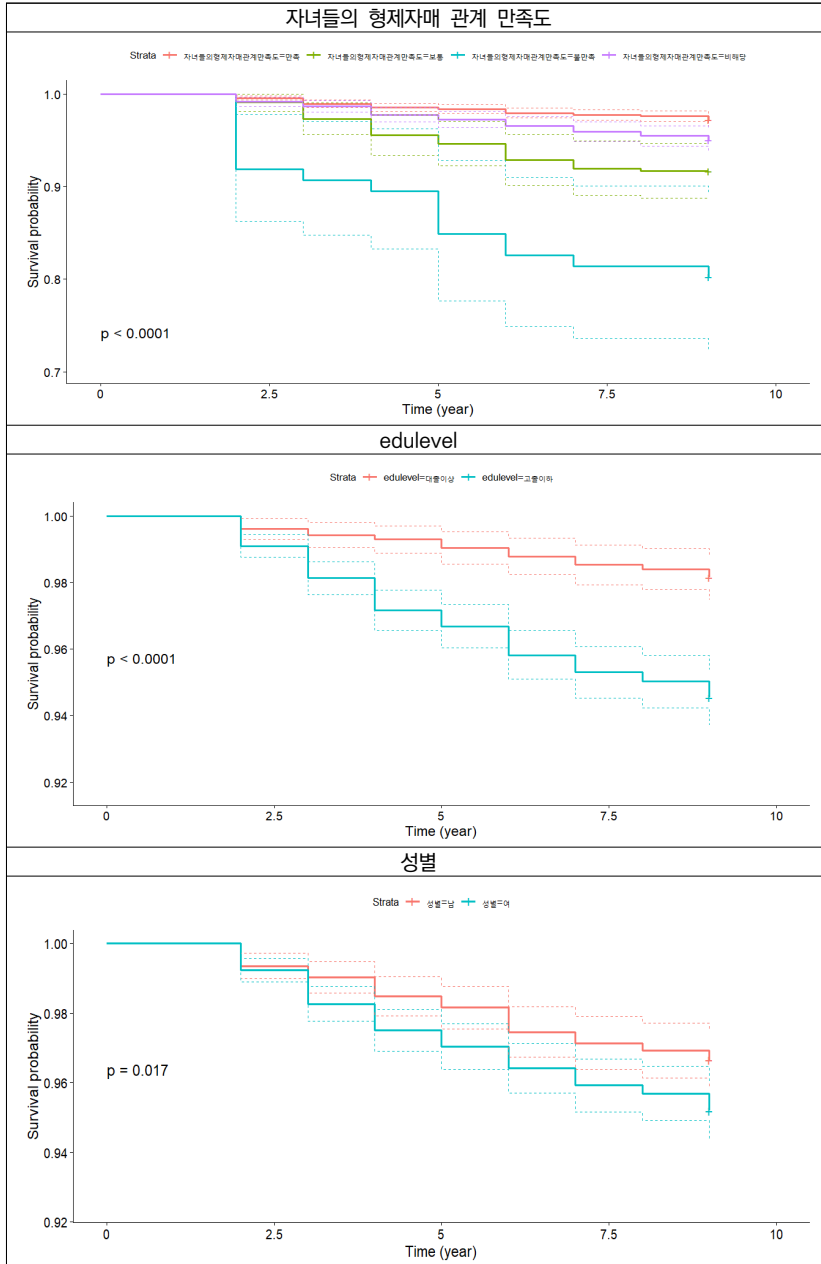


252 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로





254 보건복지분야 위기대응 예측모델링 연구-정신건강 사례를 중심으로



## 제4절 소결

활용된 모든 다변수 분석에서 유의했던 변수는 우울과 좌절감이었다. 또한, 불면, 무기력, 건강상태, 가부장적 관념 등의 변수들이 그 뒤를 이었다.

〈표 5-15〉 생존 모형 간 유의성을 지닌 변수들의 비교

결과변수 정의	자살 생각 1회 이상			자살 생각 2회 이상		
	<10%		<50%	<10%		<50%
결측 비율	전체	Stepwise	전체	전체	Stepwise	전체
변수 사용	전체	Stepwise	전체	전체	Stepwise	전체
대상자 수	3974	3974	1481	3974	3974	1481
변수 개수	33	15	37	33	13	37
경상소득		○			○	○
저소득						
건강상태	○	○	○		○	○
만성질환				○		○
근로계약						○
근로 지속 가능성			○			○
건강보험료 미납 경험						
집 점유	○	○			-	
tot_debt		-	○			
tot_debt_p						
미취업자						
실업						
근로시간						
건강 만족도						○
가족의 수입 만족도						
주거환경 만족도						
가족관계 만족도					○	○
직업 만족도						○
사회적 친분 관계			○			○

결과변수 정의	자살 생각 1회 이상			자살 생각 2회 이상		
	<10%		<50%	<10%		<50%
변수 사용	전체	Stepwise	전체	전체	Stepwise	전체
대상자 수	3974	3974	1481	3974	3974	1481
변수 개수	33	15	37	33	13	37
만족도						
여가생활 만족도			○			○
전반적 만족도			○			
가부장적 관념	○	○		○	○	
smoke		-	○		-	
drink						
식욕		-				
우울	○	○	○	○	○	○
불면	○	○		○	○	○
무기력	○	○		○	○	○
좌절감	○	○	○	○	○	○
bad		-			-	
가족생활 만족도					○	○
배우자 관계 만족도		○				
자녀 관계 만족도						○
자녀들의 형제자매 관계 만족도				○		○
성별	○	○			-	
year						
edulevel		-				○
C-index	0.728	0.727	0.77	0.811	0.802	0.88

주: Stepwise 변수선택법 적용 시 유의하지는 않았으나 모형에 포함된 변수는 '-'로 표기함.

자살 생각을 1회 이상과 2회 이상으로 각각 정의했을 시 두 모형 간의 차이는 크게 없었으나 2회 이상으로 정의할 때 가족관계 만족도, 가족생활 만족도, 자녀들의 형제자매 관계 만족도 등의 변수들이 새롭게 유의성을 갖게 되었다. 자살 생각을 1회 이상으로 정의할 때에는 집 점유, 성별,

배우자 관계 만족도 등의 변수들이 유의성을 가졌고, 이는 성별의 차이나 경제적 여유 등의 이유로 인해 발생하는 변수들이 함께 모형에서 유의하였다.

결측 비율을 50% 미만으로 확장할 때 ‘근로’와 관련된 4가지 변수들이 새롭게 추가됨에 따라 이와 관련된 근로계약, 근로 지속 가능성 등의 변수들이 유의성을 갖게 됨을 확인할 수 있었으며, 4가지 변수 추가 시 직업 만족도, 사회적 친분 관계 만족도, 여가생활 만족도 등의 사회생활과 밀접한 변수들이 새롭게 유의성을 가짐을 확인하였다.

본 장에서 실시한 실증 분석의 제한점은 다음과 같다. 수집된 자료 중 변수의 수는 많으나 의미가 중복되어 있거나 반복되어 수집된 정보까지 활용하지 못하였다. 그리고 결측 비율이 일정 범위 이상인 자료가 많았기에 자료의 손실이 일부 발생하였다. 범주형 변수의 재범주화 작업을 진행하였으나 남은 범주형 변수 중 특정 level들은 상대적으로 작은 값을 가졌던 관계로 일부 모형 수립 시 추정 값 도출에 불안정한 결과를 보였고, 이 부분은 결론의 시사점에서 다시 언급하고자 한다.







## 제6장

### 심층 분석II: 텍스트 데이터 기반 자살 고위험군 탐지를 위한 딥러닝 적용

제1절 텍스트마이닝 및 자연어처리 기반 분석 방법

제2절 텍스트의 그룹화 분석 결과

제3절 문장 임베딩을 활용한 분석 결과

제4절 소결



## 제 6 장

# 심층 분석II: 텍스트 데이터 기반 자살 고위험군 탐지를 위한 딥러닝 적용

기계학습/딥러닝 방법의 일반화 성능은 훈련데이터의 크기에 크게 의존한다. 이 장에서는 훈련데이터의 크기가 작은 데이터에 대해 내재하는 정보의 양을 최대한 활용하는 방법에 대해 살펴보고, 이를 정신건강실태조사 데이터에 적용한 결과를 살펴본다. 본 분석에서는 자기 주도 학습(self-supervised learning, SSL)을 활용하여 입력 자료의 의존성을 학습할 수 있는 딥러닝 방법에 기반한 방법을 적용하여 결과를 고찰한다.

### 제1절 텍스트마이닝 및 자연어처리 기반 분석 방법

2021년 정신건강실태조사는 자연어 형태의 질문과 보기로 구성된 객관식 문항과 자연어 형태의 질문으로 구성된 주관식 문항으로 이루어져 있다. 응답자는 질문을 읽고 본인의 상황 또는 생각에 가장 적절한 보기를 선택하거나 본인의 의견을 작성하여 조사에 참여하였다. 일반적으로 설문조사를 통해 수집된 데이터는 범주화된 보기를 기반으로 정량화하여 해석된다. 이를 통해 응답자들의 전체적인 분포를 파악하거나 종합적인 점수를 계산하여 응답자의 정신건강 상태를 비교할 수 있다. 그러나 타깃 응답의 개수를 합산하여 단일화된 수치로 평가하는 방식은 질문이 가지고 있는 문맥(context)과 이에 대한 응답 문맥을 고려하기 어렵다. 컴퓨터는 임베딩(embedding) 기법을 통해 자연어의 밀집 표현(dense representation)을 학습하여 텍스트를 이해한다. 최근 트랜스포머

(Transformer) 모델 아키텍처를 기반으로 한 다양한 딥러닝 모델이 등장하여 자연어의 임베딩 표현(embedding representation) 성능이 급격히 향상되었으며, 특히 어텐션(attention) 기법을 통해 단어와 단어 사이의 관계를 반영한 워드 임베딩 표현 학습이 가능해졌다. 이를 기반으로 자연어가 가지는 독립적인 의미와 종속적인 문맥을 고려하면서 응답자가 반응하는 풍부한 특성(feature)을 포착할 수 있다.

여기에서는 자연어 형태의 설문조사 ‘질문’과 ‘응답’ 데이터에 내포된 문맥을 반영할 수 있는 텍스트마이닝 및 자연어처리 기반 분석 방법을 적용하여 데이터를 해석해 본다.

## 1. 데이터 세트와 전처리

분석 대상은 2021년 정신건강실태조사 변수 중, B(니코틴 사용에 의한 장애), D(공포 장애 및 기타 불안 장애), E(우울 장애 및 기분부전 장애), J(알코올 사용과 관련된 장애), K(강박장애 및 외상 후 스트레스 장애), S(자살) 항목으로 선정하였다. 본 분석에서는 대표적인 응답에 대한 집단적인 패턴을 해석하기 위하여 나이, 기간과 같이 특정 응답을 요구하는 질문 변수는 제외하고 보기가 ‘예’, ‘아니오’로 이루어진 질문 변수만 택하였으며, 분석에 사용한 항목별 변수의 개수는 <표 6-1>과 같다. 사용한 변수의 번호는 <표 6-2>에 작성하였다.

〈표 6-1〉 정신장애진단도구(K-CIDI) 항목 중 사용 변수의 개수

분류 항목	총 질문 변수 개수	선정 질문 변수 개수
B(니코틴 사용에 의한 장애)	46개	30개
D(공포 장애 및 기타 불안 장애)	259개	210개
E(우울 장애 및 기분부전 장애)	84개	65개
J(알코올 사용과 관련된 장애)	71개	47개
K(강박장애 및 외상 후 스트레스 장애)	98개	75개
S(자살)	54개	5개

딥러닝 모델 적용을 위한 데이터 세트 전처리 과정은 다음과 같다.

### ① 질문별 문장 단일화

데이터 세트 입력 형식 통일을 위해 질문을 구성하는 문장이 2개 이상인 경우, 연구자가 질문을 읽고 원본 질문의 내용을 유지하며 한 개의 문장으로 수정하였다. 문장을 단일화한 질문 예시를 〈표 6-2〉에 제시하였다. 또한 질문에 포함된 특수 기호 ‘(, ‘/’, ‘)’를 삭제하였다.

〈표 6-2〉 질문별 문장 단일화 예시

설문 문항	수정 전	수정 후
E1	이제 슬프거나, 공허하거나, 우울하게 지냈던 때에 대해 질문하겠습니다. 지금까지 사는 동안, 2주 이상 거의 매일 하루 종일 슬프거나, 공허하거나, 우울하게 지낸 적이 있습니까?	지금까지 사는 동안, 2주 이상 거의 매일 하루 종일 슬프거나, 공허하거나, 우울하게 지낸 적이 있습니까?
J17_11	술을 끊거나 줄이고 난 뒤, 처음 며칠 동안에 겪을 수 있는 몇 가지 문제에 대해 질문하겠습니다_경련성 발작을 했습니까?	술을 끊거나 줄이고 난 뒤, 처음 며칠 동안 경련성 발작을 했습니까?

## ② 문맥 조건 추가

〈표 6-3〉 질문별 문맥 조건 추가 예시

설문 문항	수정 전	수정 후
E10_II	그런 기간들 중에, 2주 이상, 거의 매일, 말하거나 움직이는 게 평소보다 느려진 적이 있습니까?	우울했거나, 흥미를 잃었거나, 기운이 없었던 때, 2주 이상, 거의 매일, 말하거나 움직이는 게 평소보다 느려진 적이 있습니까?
K24	그 일 이후, 계속해서 그 일에 대한 나쁜 꿈이나 악몽에 시달렸습니까?	스트레스나 충격적인 체험을 하고 난 이후, 계속해서 사건에 대한 나쁜 꿈이나 악몽에 시달렸습니까?

일부 질문에는 이전의 질문과 이어지는 경우, 지시대명사를 사용했으므로 반복되는 내용이 생략되어 있다. 모든 조건에 대한 설명 없이도 질문 간 연결성을 이해하고 답할 수 있는 사람과 달리, 컴퓨터는 각 문장을 독립적으로 처리하기 때문에 질문을 적절히 해석하기 위해서는 모든 정보에 대해 구체적으로 서술해주어야 한다. 따라서 연구자가 질문을 읽고, 질문의 내용을 유지하도록 누락된 정보를 이전의 질문에서 찾아 추가하였다. 문맥 조건을 추가한 질문 예시를 〈표 6-3〉에 제시하였다.

## 2. 문장 임베딩(sentence embedding)

트랜스포머(Vaswani et al., 2017) 모델 아키텍처 중 인코더 층(encoder layer)만으로 구현된 BERT(Bidirectional Encoder Representations from Transformers, Devlin, Chang, Lee, & Toutanova, 2018)는 마스크 언어 모델링(masked language modeling) 기법을 통해 문맥을 양방향(bidirectional)으로 읽고 좋은 표현을 학습하는 데 효과적이다. 크게 두 종류의 자기주도 지도학습(self-su-

ervised learning)의 태스크를 활용하는데, 주어진 두 문장이 선행문장과 후행 문장을 판별하는 태스크와 문장에서의 단어를 마스킹하고 이를 예측함으로써 문맥을 파악하는 태스크로 구성되어 있다.

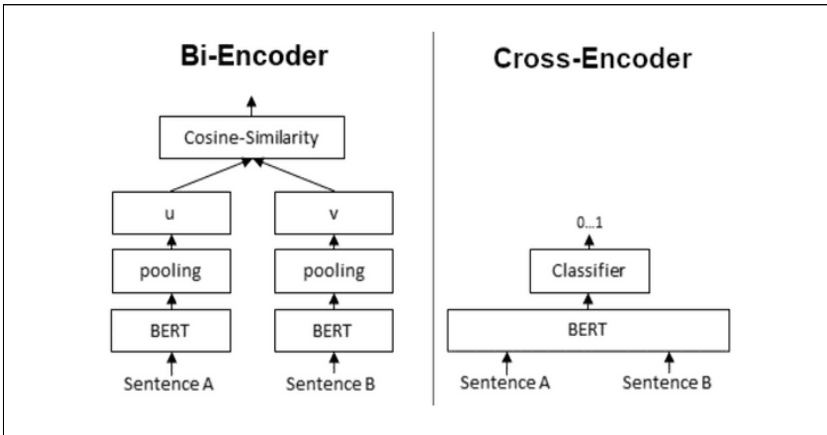
BERT에서는 총 세 가지 방법으로 문장에 대한 임베딩 벡터를 구할 수 있다. 첫 번째로, [CLS] 토큰의 임베딩 벡터를 활용하는 방법이 있다. [CLS] 토큰에는 문장을 구성하는 전체 토큰들의 종합적인 표현이 담기게 된다. 그러나 파인 튜닝(fine-tuning)을 하지 않은 사전 학습된 모델에서 [CLS] 토큰의 임베딩을 문장 전체의 임베딩으로서 사용할 경우 표현이 정확하지 않을 수 있다. 두 번째로, 모든 토큰의 임베딩 벡터를 평균 풀링하는 방법이 있다. 이는 문장을 구성하는 모든 토큰에 대한 전체적인 문맥에 관심을 둔다. 세 번째로, 모든 토큰의 임베딩 벡터를 최대 풀링하는 방법이 있다. 이는 문장을 구성하는 모든 토큰 중 가장 핵심적인 문맥에 관심을 둔다. 그러나 BERT는 cross-encoder 구조를 기반으로 두 개의 문장 쌍을 한 번에 모델에 입력받아 문장 간 연관성을 계산하는데, 문장별로 별개의 독립적인 문장 임베딩을 계산하지 못한다는 한계가 있다.

sentence-BERT(Reimers, & Gurevych, 2019)는 이러한 한계점들을 보완하여 좋은 문장 표현을 얻을 수 있도록 사전 학습된 BERT를 파인 튜닝한다. [그림 6-1]은 sentence-BERT와 BERT 구조도를 나타낸다. BERT와 달리 sentence-BERT는 Bi-Encoder 구조를 통해 두 개의 문장을 각각 모델에 통과시켜 독립적인 문장 임베딩 벡터를 구한다. 따라서 비교할 문장이 바뀌어도 다시 연산할 필요가 없어 연산 횟수가 줄어들어 계산 효율성이 증가한다.

이후 문장 쌍 분류 또는 회귀 태스크를 통해 모델을 파인 튜닝한다. 문장 쌍 분류 태스크는 삼 네트워크 아키텍처를 활용해 두 문장의 유사도 확률을 소프트맥스 함수에 입력한 후 교차 엔트로피 손실(cross-en-

trophy loss)을 최소화하는 방향으로 가중치를 업데이트한다. 문장 쌍 회귀 태스크는 트리플넷 네트워크 아키텍처를 활용해 기준(anchor) 문장과 긍정 문장(합의), 부정 문장(모순)의 세 가지 문장에 대해 기준 문장의 긍정 문장과 유사도는 높아지고, 부정 문장과 유사도는 낮아지는 방향으로 학습한다. 이러한 추가 학습을 거친 sentence-BERT는 문장의 문맥과 의미를 효과적으로 학습할 수 있다.

[그림 6-1] sentence-BERT와 BERT의 모델 아키텍처



자료: Nils Reimers. (2023). SBERT.net(SBERT의 파이썬 라이브러리). <https://www.sbert.net/examples/applications/cross-encoder/README.html>에서 2023. 12. 21. 인출.

### 3. 실험 환경

대규모의 일반적인 텍스트 데이터에 대해 사전 학습된(pre-trained) 모델을 수행하고자 하는 특정 태스크에 전이 학습(transfer learning) 또는 파인 튜닝하는 방식이 자연어처리 분야에서 우수한 성능을 보이고 있다. 대규모 데이터 세트에 대한 모델 학습은 상당한 컴퓨팅 자원과 비용, 시간이 소요되는데, Huggingface(<https://huggingface.co/>) hub 플



랫폼에서는 자연어 이해(Natural Language Understanding, NLU) 및 자연어 생성(Natural Language Generation, NLG)을 위한 다양한 범용 아키텍처를 배포하고 있다. 따라서 사용자는 모델을 직접 학습시킬 필요 없이 사전 학습된 모델의 가중치를 공유 받아 임베딩을 추출하거나 다운 스트림 태스크(downstream task)를 위한 추가 학습을 수행할 수 있다.

〈표 6-4〉 sentence-BERT 모델별 configuration

Pre-trained model	Base model	Dimension of embedding	Pooling mode
paraphrase-multilingual -MiniLM-L12-v2	Bert	384	mean tokens
xlm-r-100langs-bert-base-nli -stsb-mean-tokens	XLMLRoberta	768	mean tokens
distiluse-base-multilingual -cased-v1	DistilBert	768	mean tokens
paraphrase-multilingual -mpnet-base-v2	XLMLRoberta	768	mean tokens

자료: 허깅페이스 홈페이지의 내용을 참고하여 저자 작성  
<https://huggingface.co/sentence-transformers>에서 2023.12.27. 인출

분석에서는 한국어에 대해 사전 학습된 sentence-BERT multi-lingual 다국어 모델을 사용해 설문 조사 질문과 응답에 대한 문장 임베딩 벡터를 구하였다. 파생 모델별로 학습 데이터 세트의 종류 및 크기, 세부 아키텍처, 파라미터 개수 등 학습조건이 다르기 때문에 동일한 입력에 대해 서로 다른 임베딩을 출력한다. 본 분석에 적합한 모델을 선정하기 위하여, 입력에 대한 문장 임베딩 벡터 출력값을 네 가지 모델에 대해 비교하였다. 모델의 구체적인 구성(configuration)은 〈표 6-4〉와 같다.

비교 1에서는 응답 ‘예’와 응답 ‘아니오’를 각각 하나의 문장으로 취급하여, 두 문장의 임베딩 벡터 간 코사인 유사도를 모델별로 비교하였다.

비교 2에서는 랜덤하게 선택한 다섯 가지의 질문(항목별 1개)에 대해 응답 ‘예’와 ‘아니오’를 질문에 각각 결합한 문장을 생성하여 이들의 문장 임베딩 벡터 간 코사인 유사도를 모델별로 비교하였다. 비교 3에서는 질문에 대해 응답을 바꿔가며 코사인 유사도를 구해 모델별로 응답이 전체 문장 임베딩에 미치는 영향의 정도를 비교하였다.

〈표 6-5〉 (비교 1, 비교 2) 응답에 따른 문장 임베딩 코사인 유사도

모델명	응답[예] /응답 [아니오]	질문-응답[예] / 질문-응답[아니오]				
		B14	D65_2	E2	J7B	K22_1]
paraphrase-multilingual -MiniLM-L12-v2	0.804	0.982	0.989	0.979	0.980	0.993
xlm-r-100langs-bert-base -nli-stsb-mean-tokens	0.674	<b>0.730</b>	<b>0.565</b>	<b>0.582</b>	<b>0.801</b>	<b>0.539</b>
distiluse-base-multilingua l-cased-v1	0.728	0.990	0.991	0.994	0.981	0.991
paraphrase-multilingual -mpnet-base-v2	<b>0.619</b>	0.996	0.995	0.995	0.995	0.997

모델 paraphrase-multilingual-mpnet-base-v2는 〈표 6-5〉와 같이 ‘예’와 ‘아니오’의 두 문장의 코사인 유사도가 0.619로 가장 낮았지만, 질문과 결합하였을 때는 반대로 가장 높은 유사도를 보였다. 이는 응답의 종류가 문장의 전체적인 문맥에 미미한 영향을 미치고 있다는 것을 의미한다. 즉, ‘지금까지 사는 동안, 한 달 이상 매일 담배를 피운 적이 있습니까? 예’의 문장과 ‘지금까지 사는 동안, 한 달 이상 매일 담배를 피운 적이 있습니까? 아니오’의 응답이 주는 긍정 또는 부정의 의미보다 질문의 목적에 더욱 초점을 두고 있기 때문으로 짐작된다.

한편, 모델 xlm-r-100langs-bert-base-nli-stsb-mean-tokens는 〈표 6-6〉과 같이 ‘예’와 ‘아니오’가 질문에 결합되었을 때, 다른 모델들에

비해 상대적으로 전체적인 문맥을 이해하고 있음을 확인할 수 있다. 또한, <표 6-6>의 B1A\_A의 질문에 ‘예’와 ‘아니오’라고 응답한 경우는 0.770으로 다른 모델들보다 낮은 값을 가지고, 질문이 유사한 B1A\_A와 B1A\_E의 유사도는 ‘예’와 ‘아니오’ 각각의 경우에 높은 유사도를 보이는 특징을 보인다. 이는 ‘예’와 ‘아니오’가 문장에서 가지는 의미가 모델을 통해 적절한 가중치를 학습하여 임베딩 벡터에 반영되어 있음으로 해석할 수 있다. 설문 문항의 높은 신뢰도를 산출하는 지표인 크론바흐 alpha 지수에서는 유사 설문 문항에 대한 높은 유사도를, 비유사 문항에 대해서는 낮은 유사도를 가지는 특징을 가지는 설문 문항을 선호하는데, 본 연구에서 제안하는 방법은 이러한 점과 맥락을 같이 한다고 해석할 수 있다. 모델이 어떤 문맥에 더 큰 가중치를 두고 학습하였는지에 따라 동일한 단어, 문장에 부여되는 임베딩 값이 달라짐을 확인할 수 있다. 본 연구에서는 모델의 추가 학습 없이 사전 학습된 모델로부터 문장 임베딩을 출력받아 분석을 진행하기 위하여 모델별 임베딩 공간을 탐색하고 분석 목적과 대상에 가장 부합하는 모델로 xlm-r-100langs-bert-base-nli-stsb-mean-tokens을 선택하였다.

<표 6-6> (비교 3) 질문-응답 쌍에 따른 문장 임베딩 코사인 유사도

모델명	B1A_A-응답[예]/ B1A_A-응답[아니오]	B1A_A-응답[예]/ B1E_A-응답[예]	B1A_A-응답[아니오]/ B1E_A-응답[아니오]
paraphrase-multilingual -MiniLM-L12-v2	0.976	0.930	0.933
xlm-r-100langs-bert-bas e-nli-stsb-mean-tokens	0.770	0.977	0.998
distiluse-base-multilingu al-cased-v1	0.991	0.696	0.683
paraphrase-multilingual -mpnet-base-v2	0.995	0.958	0.9578

#### 4. t-SNE 차원 축소

t-SNE(t-distributed Stochastic Neighbor Embedding)는 비지도(unsupervised) 비선형(non-linear) 차원 축소 기법으로, 고차원의 데이터를 저차원으로 축소하여 데이터의 분포를 탐색하고 시각화하는 방법이다. 데이터 피쳐(feature)의 차원이 늘어날수록 데이터 포인트 간의 거리가 기하급수적으로 멀어지게 되고, 희소한(sparse) 형태가 되어 과적합(over fitting)이 발생하거나 성능이 저하된다. 따라서 고차원의 원본 데이터에 담긴 정보의 손실을 최소화하며 저차원의 공간으로 투영하여 데이터의 특성과 분포를 직관적으로 해석하고 핵심적인 정보를 추출할 수 있다.

비선형적 차원 축소 기법은 선형 직선으로 구분되지 않는 데이터의 분포를 고려하며, 대표적으로 t-SNE가 있다. t-SNE(Stochastic Neighbor Embedding, Roweis, & Hinton, 2002; Maaten, & Hinton, 2008)는 고차원 공간의 데이터( $x$ )와 이웃한 데이터들과의 거리를 보존하는 저차원의 표현( $y$ )을 확률적으로 학습하는 방법이다. 확률

$$p_{j|i} \text{는 } p_{j|i} = e^{-\frac{|x_i - x_j|^2}{2\sigma_i^2}} / \sum_k e^{-\frac{|x_i - x_k|^2}{2\sigma_i^2}} \text{ 로 고차원 공간에 존재하는 } x_i \text{에 대}$$

해  $x_i$ 를 중심으로 하는 가우시안 분포 밀도에 비례하여 이웃한 데이터 포인트들을 선택할 확률을 의미한다. 근사하는 확률  $q_{j|i}$ 는

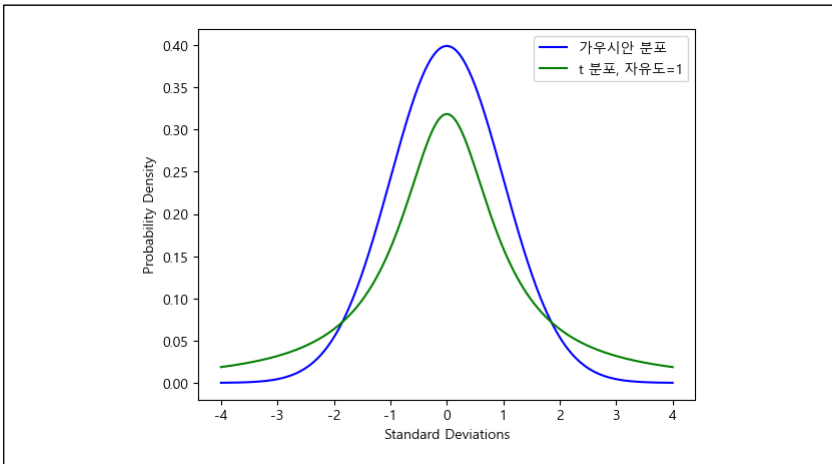
$$q_{j|i} = e^{-|y_i - y_j|^2} / \sum_k e^{-|y_i - y_k|^2} \text{ 로 저차원 공간에 축소된 표현 } y_i \text{에 대한 이}$$

웃 포인트들의 거리에 따라 선택될 확률을 의미한다. 그러므로  $p_{j|i}$ 와  $q_{j|i}$ 가 유사하다면 고차원의 데이터 간 거리가 저차원의 표현에서도 잘 보존되어 있다고 해석할 수 있다. 이들의 확률 분포는 KL divergence를 최소

화하는 방향으로  $q_{ji}$ 의  $y_i$ 들을 학습한다.

그러나 가우시안 분포를 전제하는 SNE는, 양쪽 꼬리가 상대적으로 얇다는 특징 때문에 일정 거리 이상에서는 데이터들이 선택될 확률이 비슷해지는 crowding problem이 발생하게 된다. t-SNE는 가우시안 분포(그림 6-2)와 유사하지만 양쪽 꼬리의 끝이 조금 더 두꺼운 자유도 1의 t분포를  $q_{ji}$ 계산에 사용한다. 이를 통해 t-SNE는 고차원의 데이터를 지역적인(local) 특징과 전역적인(global) 특징을 보존하며 효과적으로 저차원 공간에 축소시키고, 데이터의 군집 형태를 시각화하는 데 유용한 기법이다.

[그림 6-2] 정규분포와 자유도 1인 t분포의 비교 t-분포



## 제2절 텍스트의 그룹화 분석 결과

### 1. 질문-응답 쌍을 기준으로 그룹화

응답에 따른 질문의 임베딩 공간을 해석하기 위하여 <표 6-6>을 근거로 모델 xlm-r-100langs-bert-base-nli-stsb-mean-tokens를 사용하였다. 또한, 응답 간 차이를 극대화하기 위하여 자연어 상태에서 같은 의미를 지닌 ‘예’와 ‘네’, ‘아니오’와 ‘아니요’를 각각 조합하여 질문(변수)에 매핑한 후 코사인 유사도를 비교하였다. 랜덤하게 선택한 5개의 질문을 대상으로 계산한 결과는 <표 6-7>과 같으며, 본 분석에서는 코사인 유사도가 가장 낮은 빈도수가 높은 ‘네’와 ‘아니오’ 조합을 사용하여 분석을 진행하였다.

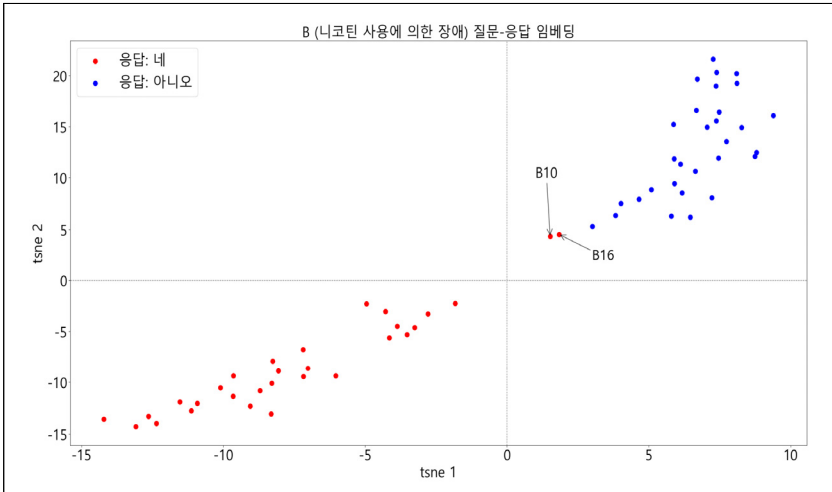
<표 6-7> 응답 조합별 동일 변수에 대한 코사인 유사도

질문 문항	예-아니오	예-아니요	네-아니오	네-아니요
B8	0.888	0.887	0.782	0.801
D65_6	0.737	0.893	0.651	0.569
E38	0.607	0.550	0.593	0.554
J19	0.790	0.756	0.691	0.700
K43_2	0.856	0.848	0.743	0.767

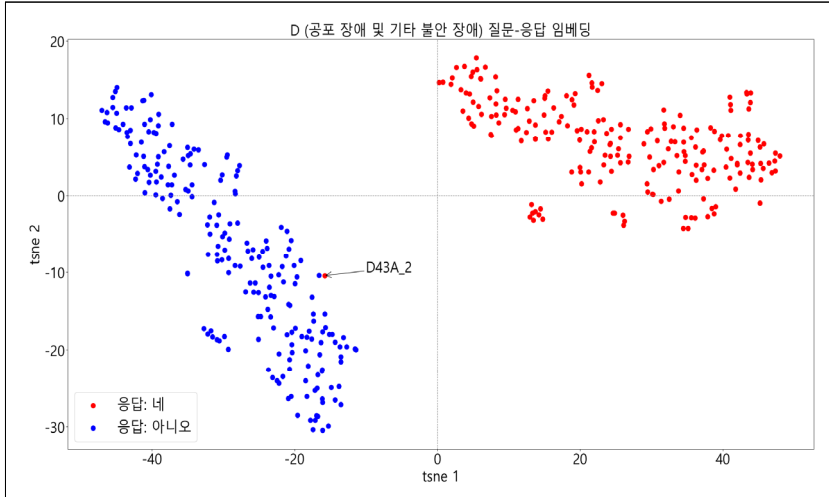
전처리를 마친 질문에 응답 ‘네’와 ‘아니오’를 각각 매핑하여 하나의 질문에 대해 두 개의 입력 데이터 쌍(예시: ‘동물, 높은 곳, 폭풍우, 갇힌 장소에 있는 것, 피를 보는 것들 중 어떤 것이라도 심하게 두려워하거나 피한 적이 있습니까? 네’, ‘동물, 높은 곳, 폭풍우, 갇힌 장소에 있는 것, 피를 보는 것들 중 어떤 것이라도 심하게 두려워하거나 피한 적이 있습니까?’)

까? 아니오’)을 생성하였다. 각 설문 문장 임베딩 벡터는 평균 풀링 (average pooling)으로 구한 768차원의 벡터를 산출하였으며, t-SNE 를 활용해 2차원으로 축소하여 항목별로 시각화하였다. B(니코틴 사용에 의한 장애), D(공포 장애 및 기타 불안 장애), E(우울 장애 및 기분부전 장애), J(알코올 사용과 관련된 장애), K(강박장애 및 외상 후 스트레스 장애) 질문-응답 임베딩의 t-SNE 차원 축소 시각화는 각각 [그림 6-3]에서 [그림 6-7]까지에 나타낸다. 시각화 결과, 응답 ‘네’와 ‘아니오’가 결합된 질문은 t-SNE 2차원 축소 임베딩 공간상에서 대부분 뚜렷한 구분을 보인다. 이는 문장 임베딩 벡터가 반대되는 응답의 의미를 반영하고 있음을 보여주는 것이다.

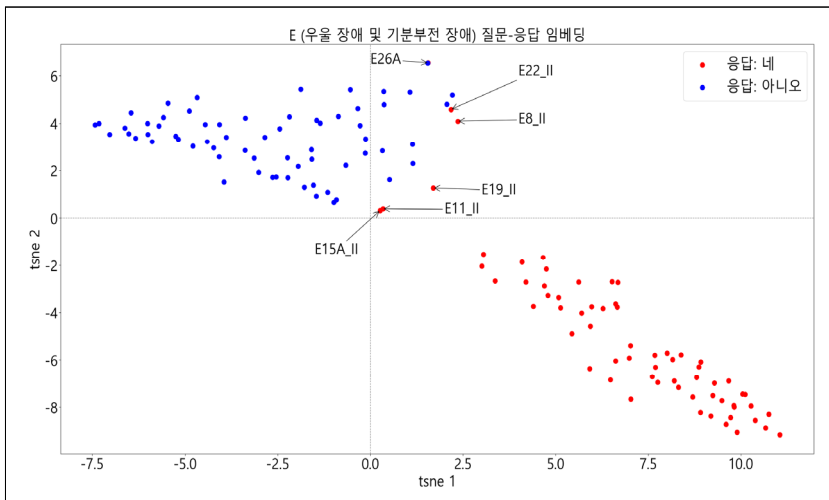
[그림 6-3] B(니코틴 사용에 의한 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화



[그림 6-4] D(공포 장애 및 기타 불안 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화

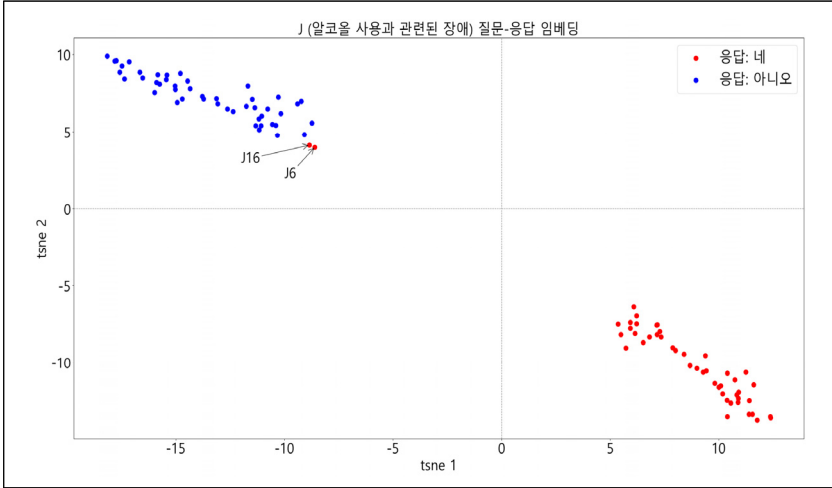


[그림 6-5] E(우울 장애 및 기분부전 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화

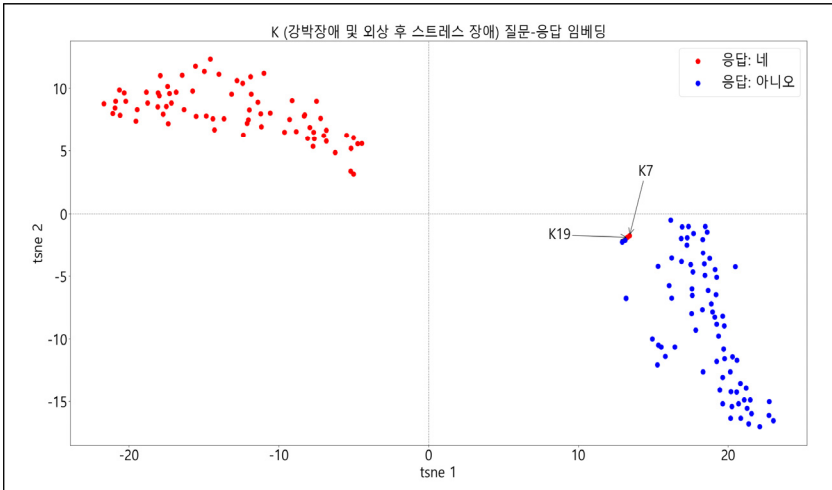




[그림 6-6] J(알코올 사용과 관련된 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화



[그림 6-7] K(강박장애 및 외상 후 스트레스 장애) 질문-응답 임베딩 t-SNE 차원 축소 시각화



t-SNE 차원 축소에 따라 군집된 변수들 중 반대 응답과 거리가 가까운 변수들을 살펴보았다. 이들 변수에 등장하는 친구, 친척, 여행, 가족, 생활, 아침, 영화 등이 내포하고 있는 긍정(positive)의 의미가 문장의 전체적인 의미를 왜곡할 가능성을 발견하였다. <표 6-8>은 반대 응답과 가까운 거리의 문항을 나타낸다. 이를 보완하기 위하여 중립적인 의미의 단어로 수정하거나 일부 단어를 삭제한 후 임베딩 공간을 비교하였다. <표 6-9>는 일부의 예시로 'D43A\_2' 문항과 'J6' 문항을 원본 질문의 표현을 수정하여 문장의 의미가 긍정과 부정의 의미의 방향을 가지게 하였다. [그림 6-8]은 [그림 6-4]와 동일한 설문 문항들에 대한 임베딩으로 'D43A\_2'의 문항의 위치가 수정된 것을 확인할 수 있다. [그림 6-9]도 [그림 6-6]과 동일한 설문 문항들에 대한 임베딩으로 'J6' 문항의 위치가 수정되었음을 볼 수 있다. 문장의 의미를 유지하며 단어들을 변경하였을 때 문장의 임베딩 공간이 달라짐을 확인할 수 있다.

<표 6-8> 반대 응답과 가까운 거리의 변수

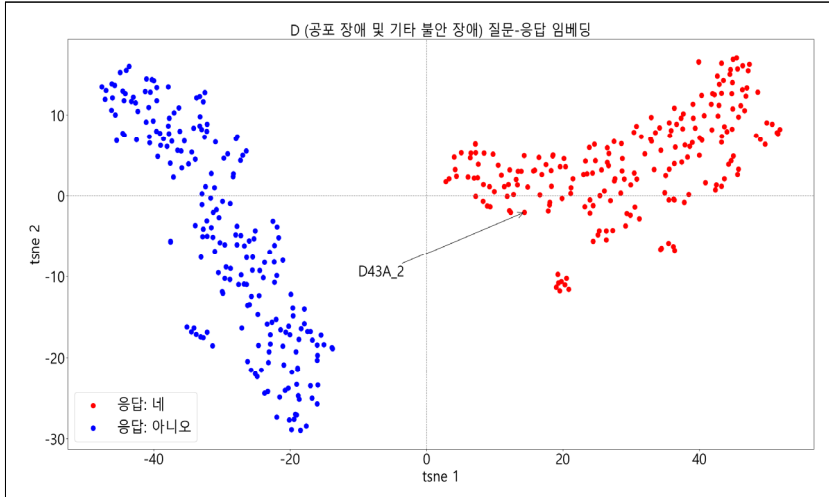
설문 문항	질문
B10	담배를 끊거나 줄인 뒤 생긴 문제 때문에, 일을 하거나 다른 사람들과 어울리는 데 지장을 받은 적이 있습니까?
B16	담배를 피우기 위해서, 운동이나 일, 친구나 친척들과의 모임 같은 중요한 활동을 포기하거나 많이 줄인 적이 있습니까?
D43A_2	버스, 기차, 또는 자동차를 타고 여행하는 상황을 심하게 두려워 한 적이 있습니까?
E8_II	(우울했거나/흥미를 잃었거나/기운이 없었던) 때, 2주 이상, 거의 매일 밤, 잠이 안 오거나, 한밤중에 깨거나, 너무 일찍 깨 잠자기가 어려웠습니까?
E11_II	(우울했거나/흥미를 잃었거나/기운이 없었던) 때, 2주 이상, 가만히 앉아 있지 못하고 왔다 갔다 하거나, 앉아 있어도 손을 가만히 두지 못하는 등 안절부절 못한 적이 있습니까?
E15A_II	(우울했거나/ 흥미를 잃었거나/기운이 없었던) 때, 집중할 수가 없어서, 평소 좋아하던 책, TV, 또는 영화를 볼 수 없었습니까?

설문 문항	질문
E19_II	(우울했거나/ 흥미를 잃었거나/기운이 없었던) 때, 너무 우울해서, 자살해야겠다는 생각을 많이 했습니까?
E22_II	(우울하거나/ 흥미를 잃거나/기운이 없이) 지냈던 2주 동안, 거의 매일, 아침에 일어날 때는 특히 기분이 나쁘다가 시간이 지날수록 기분이 나아졌습니까?
E26A	(우울하거나/ 흥미를 잃거나/기운이 없이) 지냈던 기간들 중에, 2주 이상 우울함과 관련된 증상 때문에 직장생활을 하거나, 집안, 가족, 또는 자신을 돌보는 데 지장이 많이 있었습니까?
J6	지금까지 사는 동안, 술에 취하거나 술에서 덜 깨서 학교, 직장, 또는 가정 생활에 지장을 자주 받은 적이 있습니까?
J16	술 때문에 운동, 학업, 직업, 또는 친구나 친척과의 관계 유지를 포기하거나 많이 줄이는 등 중요한 활동을 포기하거나 많이 줄인 적이 있습니까?
K7	원하지 않는 불쾌한 생각 또는 떨쳐버릴 수 없는 불쾌한 생각 때문에 생활이나 일에 지장을 받거나, 친척이나 친구들과 지내는 데 어려움이 생기거나, 또는 기분이 몹시 상했습니까?
K19	어리석은 행동을 반복하거나 일정한 순서대로 행동하거나 개수를 세는 행동 때문에 생활이나 일에 지장을 받거나, 친척이나 친구들과 지내는 데 어려움이 생기거나, 또는 기분이 몹시 상한 적이 있습니까?

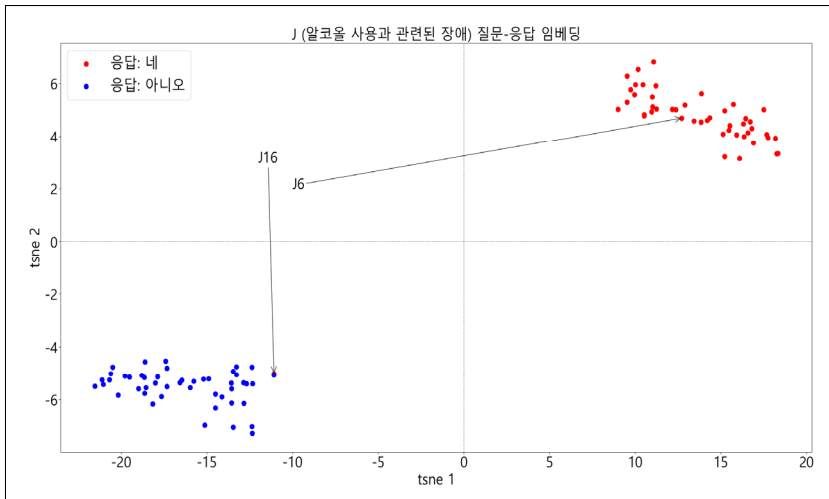
〈표 6-9〉 수정 전후 질문

설문 문항	원본 질문	변경 질문
D43A_2	버스, 기차, 또는 자동차를 타고 여행하는 상황을 심하게 두려워한 적이 있습니까?	버스, 기차, 또는 자동차를 타고 이동하는 상황을 심하게 두려워한 적이 있습니까?
J6	지금까지 사는 동안, 술에 취하거나 술에서 덜 깨서 학교, 직장, 또는 가정생활에 지장을 자주 받은 적이 있습니까?	지금까지 사는 동안, 술에 취하거나 술에서 덜 깨서 생활에 지장을 자주 받은 적이 있습니까?

[그림 6-8] D(공포 장애 및 기타 불안 장애) 수정된 질문-응답 임베딩 t-SNE 차원 축소 시각화



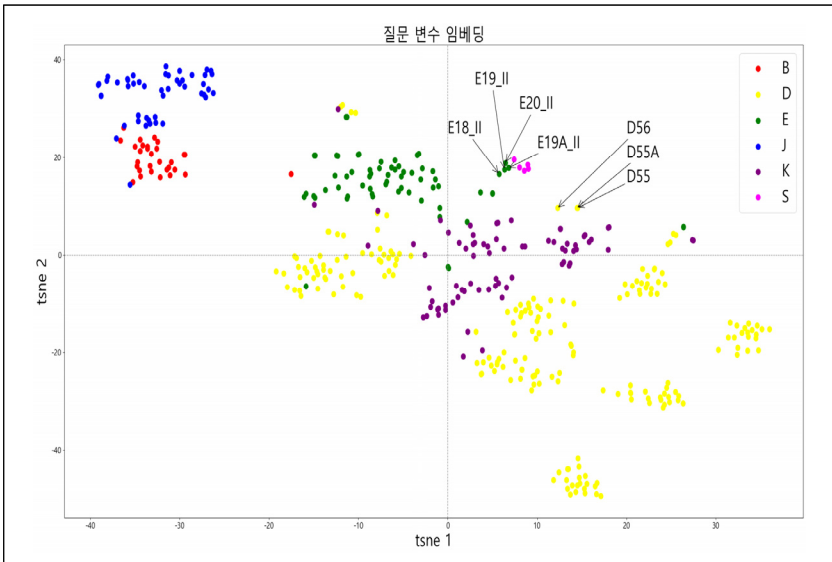
[그림 6-9] J(알코올 사용과 관련된 장애) 수정된 질문-응답 임베딩 t-SNE 차원 축소 시각화



## 2. 질문을 기준으로 그룹화

응답을 기준으로 그룹화한 것과 동일한 방법으로 설문 문항에 따른 질문의 임베딩 공간을 해석하기 위하여 질문 간 차이가 극대화될 수 있는 모델 xlm-r-100langs-bert-base-nli-stsb-mean-tokens를 사용하였다. 전처리를 마친 질문을 입력 데이터로 하여 768차원의 문장 임베딩 벡터는 t-SNE를 활용해 2차원으로 축소하여 시각화하였다.

[그림 6-10] 질문 설문군의 임베딩 t-SNE 차원 축소 시각화



[그림 6-10]은 임베딩 결과를 나타내는데 B, D, E, J, K, S의 분류 항목의 설문 문항들이 그룹을 형성하고 있음을 확인할 수 있다. 특히, B의 니코틴 사용장애와 J의 알코올 사용장애가 다른 그룹들보다 상대적으로 가깝게 분포하고 있다. D의 공포 장애 및 기타 불안 장애는 비교적 넓은

범위에 분포하고 있음을 확인하였다. 가장 밀집하게 분포하고 있는 문항들은 S(자살) 설문 문항들로 이들과 인접한 거리에 있는 다른 영역의 설문 문항들은 <표 6-10>에 정리하였다. D(공포 장애 및 기타불안 장애), E(우울장애 및 기분부전 장애)에 분류된 문항들로 이들 문항의 기술에는 직접적으로 ‘자살’, ‘죽음’, ‘목숨’ 같은 단어들이 사용되었음을 알 수 있다.

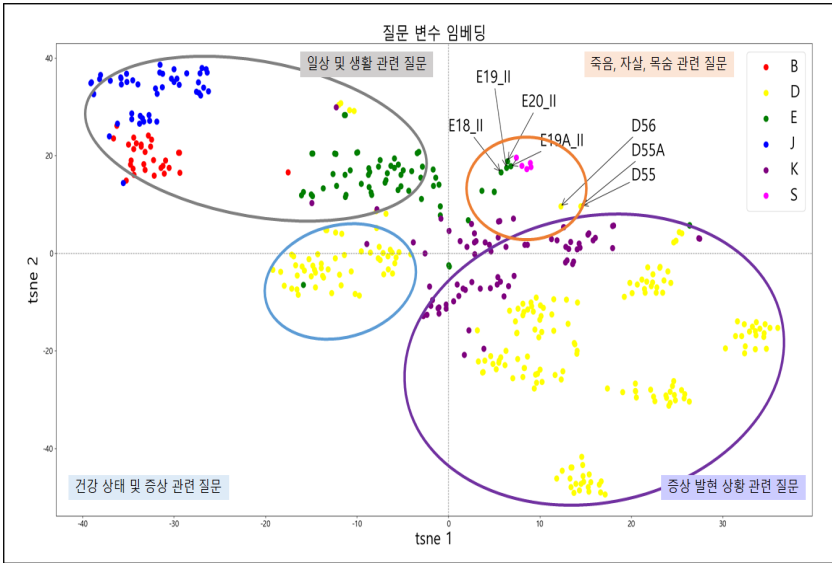
<표 6-10> S(자살) 설문 문항군과 근접한 임베딩 공간의 질문 변수

설문 문항	질문
D55	목숨이 위험한 상황에서 불안발작이 있습니까?
D55A	목숨이 위험하지 않은 상황에서도 불안발작이 있습니까?
D56	목숨이 위험하지 않은 상황에서 전혀 예상하지 못한 불안발작을 두 번 이상 겪은 적이 있습니까?
E18_II	우울했거나, 흥미를 잃었거나, 기운이 없었던 때, 2주 이상, 죽음에 대한 생각을 많이 한 적이 있습니까?
E19_II	우울했거나, 흥미를 잃었거나, 기운이 없었던 때, 너무 우울해서, 자살해야겠다는 생각을 많이 했습니까?
E19A_II	우울했거나, 흥미를 잃었거나, 기운이 없었던 때, 자살을 어떻게 할지 계획을 세웠습니까?
E20_II	우울했거나, 흥미를 잃었거나, 기운이 없었던 때, 죽음에 이를 가능성이 없어 보이는 시도도 모두 포함하여 자살을 시도했습니까?

[그림 6-11]은 설문 문항들의 분류군에 따른 t-SNE 차원 축소 공간에서의 임베딩의 분포를 나타낸다. 각 사분면을 기준으로 유사한 임베딩 공간을 보유한 질문 변수들에 대한 해석은 다음과 같다. 제1사분면에 위치한 S(자살) 항목의 질문 변수들은 일부 E(우울장애 및 기분부전 장애) 항목의 질문 변수들과 근접한데, 이에 해당하는 질문 변수 E18\_II, E19\_II, E19A\_11, E20\_II에는 공통적으로 ‘자살’과 ‘죽음’이라는 단어가 등장하였다. 또한 이들과 밀접한 거리에 이는 D(공포 장애 및 기타 불

안 장애)에는 ‘목숨이 위협’이라는 어구가 확인되었다. 이를 통해 1사분면 공간에는 죽음, 자살, 목숨과 직접적으로 관련 있는 변수들의 문맥이 포함되어 있음을 유추할 수 있다.

[그림 6-11] 질문 변수 임베딩 t-SNE 차원 축소 해석



제2사분면에 위치한 B(니코틴 사용에 의한 장애)와 J(알코올 사용과 관련된 장애) 항목은 매우 근접한 공간에 위치하고 있다. 이는 문장에 반복적으로 등장하는 ‘담배’와 ‘술’의 임베딩 공간이 유사하기 때문인 것으로 확인된다. 특히, B, J, E에는 일상과 관련된 생활 패턴 및 평상적 상황에 대한 질문이 다수 존재하였다. 이를 통해 2사분면 공간에는 일상 및 생활과 관련 있는 변수들의 문맥이 포함되어 있음을 유추할 수 있다.

제3사분면에 위치한 D(공포 장애 및 기타 불안 장애) 변수들은 건강 상태 및 증상과 관련된 질문들이 군집되어 있다. 걱정 또는 긴장한 상태에

서 발견된 구체적인 증상에 대한 질문이 다수 존재하였다. 특히 K27의 ‘땀이 나거나 심장이 빨리 뛰거나 몸이 떨리다’ 같은 특정한 신체적 현상에 대한 질문이 이와 유사한 거리에 위치하고 있음을 알 수 있다. 이를 통해 제3사분면 공간에는 건강 상태 및 증상과 관련 있는 변수들의 문맥이 포함되어 있음을 유추할 수 있다.

제4사분면에 위치한 D 항목 변수와 K(강박장애 및 외상 후 스트레스 장애) 항목 변수들은 공포심 또는 불안 증상이 발생하는 특수한 상황과 관련된 질문들이 군집되어 있다. 특정 증상의 발현을 유발하는 구체적인 상황에 대해 조사하고 있으며, 경험한 강박 증상에 관련된 특수한 상황에 대해 한 질문이 다수 존재하였다. 이를 통해 제4사분면 공간에는 증상 발현 상황과 관련 있는 변수들의 문맥이 포함되어 있음을 유추할 수 있다.

## 제3절 문장 임베딩을 활용한 분석 결과

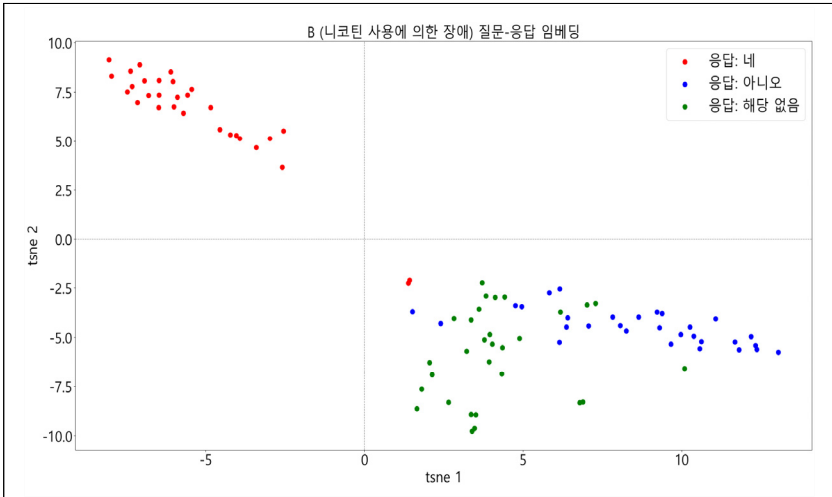
### 1. 문장 임베딩을 활용한 정신건강 집단 비교

앞 절의 분석을 통해 질문 변수가 가지는 임베딩 공간이 응답에 따라 반대되는 방향으로 나누어짐을 확인하였으며, 이를 기반으로 응답자들의 응답 데이터를 해석해 보고자 한다. 설문 조사 데이터의 변수 ID와 숫자로 범주화되어 있는 응답(1: 아니오, 5: 네)은 각각 자연어 형태로 변경하였으며, 질문 간 종속성으로 인해 NA(응답 값 없음) 처리된 문항은 ‘해당 없음’으로 치환하였다. ‘해당 없음’ 문장이 가지는 임베딩 공간을 ‘네’, ‘아니오’ 문장과 비교하기 위하여 질문-응답의 문장 임베딩을 t-SNE를 활용해 2차원으로 축소하여 항목별로 시각화하였다.

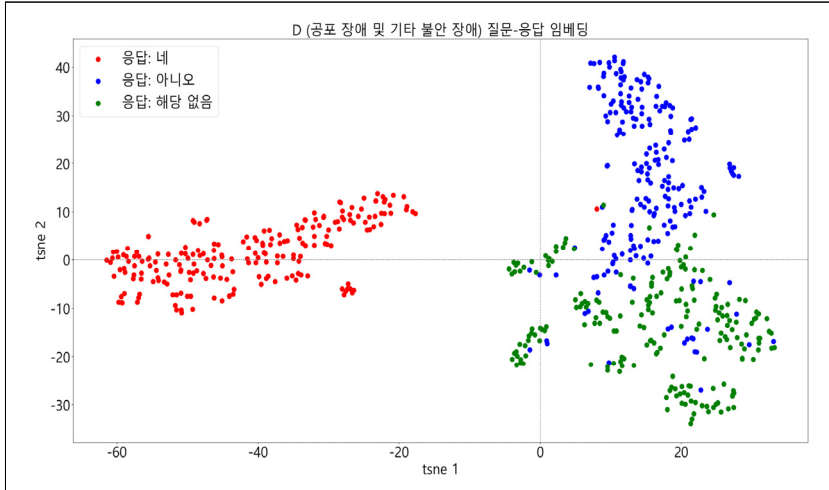


‘해당 없음’은 응답 ‘네’보다 ‘아니오’와 유사한 공간을 형성하고 있음을 알 수 있다. ‘해당 없음’이 부여되는 경우는 이전의 질문 변수에 종속되어 있는 경우에 부여된다. 예를 들어, B 항목 변수 중 첫 번째와 두 번째에 등장하는 B1A\_A(지금까지 사는 동안, 한 달 이상 매일 담배를 피운 적이 있습니까?)와 B1E\_A(지금까지 사는 동안, 한 달 이상 매일 전자담배를 피운 적이 있습니까?)에 ‘아니오’라고 응답한 응답자는 이후의 B 항목 변수에 해당하지 않기 때문에 NA(응답 값 없음) 처리된다. 따라서 해당 모델로 구한 질문-응답 쌍의 문장 임베딩은 합리적이라고 볼 수 있다. 해당없음을 포함한 질문-응답 임베딩의 t-SNE 차원 축소 시각화는 각각 [그림 6-12]부터 [그림 6-16]까지에 나타난다.

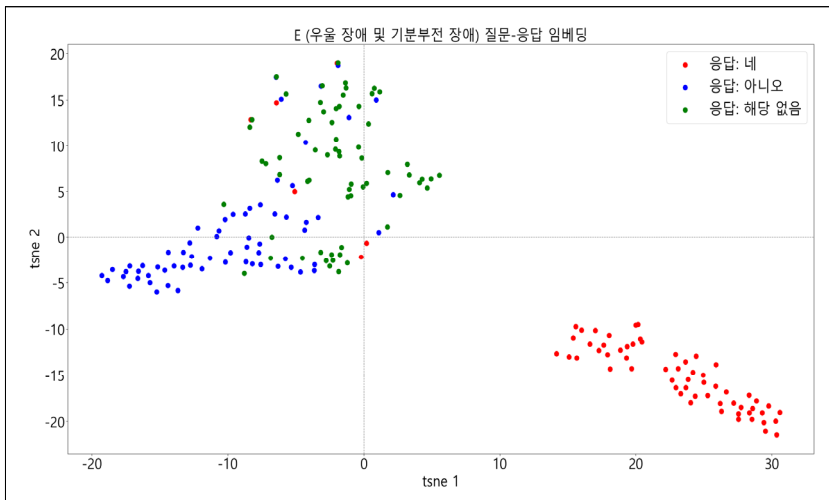
[그림 6-12] B(니코틴 사용에 의한 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화



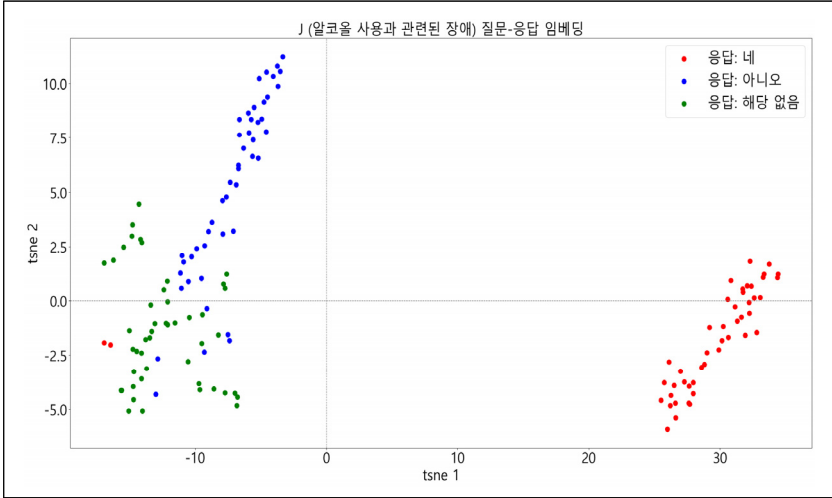
[그림 6-13] D(공포 장애 및 기타 불안 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화



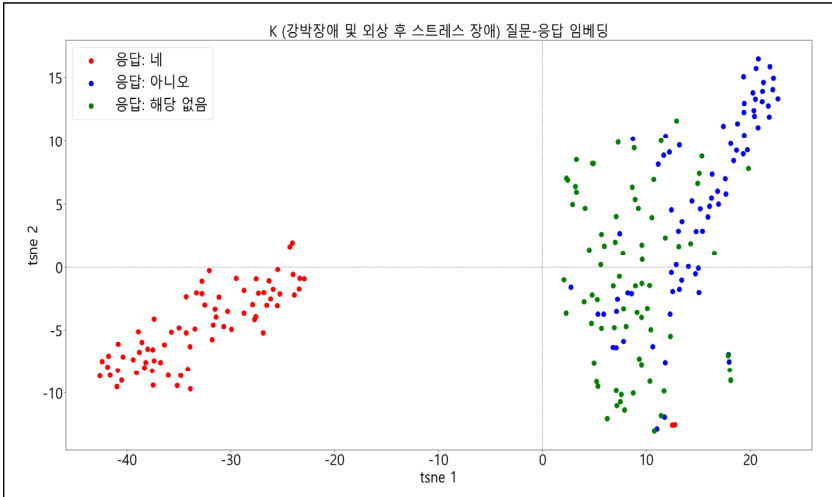
[그림 6-14] E(우울 장애 및 기분부전 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화



[그림 6-15] J(알코올 사용과 관련된 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화



[그림 6-16] K(강박장애 및 외상 후 스트레스 장애) 추가 응답 임베딩 t-SNE 차원 축소 시각화



본 분석에서는 제4장 제2절의 Y 정의(평생 자살 사고 여부)와 동일하게 S1(자살하는 것에 대해 진지하게 생각한 적이 한 번이라도 있습니까?) 변수의 응답을 기준으로 자살 고위험군을 분류하였다. S1 질문 변수에 '네'라고 응답한 고위험군(class: 1) 그룹과 '아니오'라고 응답한 저위험군(class: 0) 그룹의 이진 분류분석(binary classification)을 수행하였다. 전체 응답자 5,511명 중 574명이 고위험군에 포함되며, 불균형한(imbalanced) 데이터 분포를 고려하여 저위험군은 574명을 랜덤 샘플링하였다. 학습은 문장 임베딩 벡터로 구성된 데이터 세트와 생성변수로 구성된 데이터 세트 총 두 가지에 대해 진행하였으며, XGBOOST 모델을 사용해 성능을 비교하였다.

문장 임베딩 벡터 데이터 세트는 그룹별 574명, 총 1,148명에 대해 각 427개의 질문-응답 쌍을 매핑하였으며, 5개의 분류 항목별(B, D, E, J, K)로 t-SNE 차원 축소를 수행하였다. 총 1,148명에 대해 t-SNE 1차원 데이터 세트는 427차원의 학습 데이터를, t-SNE 2차원은 한 설문 문항당 2개의 차원을 가지므로 총 854차원을 가진 학습 데이터를 생성하였다.  $y$ (타겟 변수)는 그룹별로 라벨링하였다. 일반적으로 언어 모델에서는 perplexity이 큰 값으로 설정할수록 훈련 시 모형의 언어이해력(예측확률)이 높은 모형을 선호한다는 의미를 가지는데, t-SNE 학습 시에는 일반적인 거리 계산 시 인접한 데이터에 대한 상대적인 거리를 조절할 수 있는 인자로 활용되어 국소성(locality) 정도를 조율하는 모수로 활용된다. 이 옵션이 낮을수록 거리 계산 시 국소성이 더 강화된다고 할 수 있다. 본 연구에서는 두 값을 설정하여 t-SNE 축소 좌표를 산출하는 데 반영하였다.

생성변수 데이터 세트는 동일한 그룹에 대해 1년 유병률, 1개월 유병률, 평생 유병률 변수를 사용하였다. 단, 자살 관련 행동에 해당하는 항목

변수 중 S-Idea는 타깃인  $y$ 와 일치하기 때문에 학습 데이터에서 제외하였다. 또한, S-Idea 이외의 자살 관련 행동 변수 8개와 문장 임베딩 항목에 반영되지 않은 인터넷 게임장애 항목 9개와 게임 생활 습관 항목 1개도 제외하였다. 따라서 1,148명에 대해 60개의 생성변수를 학습 데이터로 설정하였으며,  $y$ 는 동일하게 그룹별로 라벨링하였다.

훈련 데이터와 테스트 데이터의 비율인 `split_random_state`를 실험마다 통일하며 8:2로 나누었고, min-max 정규화를 수행하였다. 모델의 파라미터는 각각 트리 모델의 개수(`n_estimators`)를 1000, 트리의 최대 깊이(`max_depth`)를 3으로 설정하였다.

〈표 6-11〉은 10번의 샘플링에 대한 결과를 나타낸다. 생성변수로 학습한 모델은 평균 0.674의 정확도를 보였으며, 설문 문항의 응답에 대한 정확도는 0.705이다. 질문-응답 쌍의 문장 임베딩 데이터로 학습한 모델은 평균값 기준으로 t-SNE 2차원 축소 데이터에서 perplexity=20일 때의 분류 성능이 0.827로, 평균적으로 가장 높은 것으로 나타났다. 설문 문항의 응답과 설문 문항의 문장의 중요도를 반영할 경우를 비교할 때, 최대 0.153의 정확도가 상승하였음을 확인하였다.

〈표 6-11〉 XGBOOST 정확도(accuracy) 결과

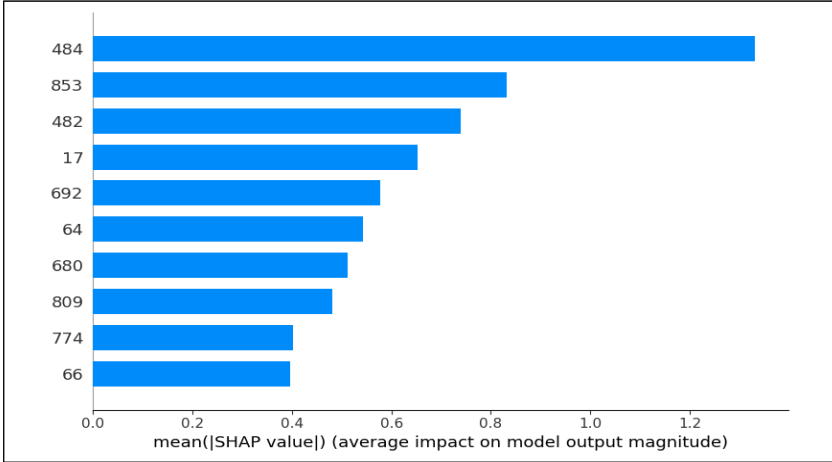
no	생성변수	설문 문항	t-SNE 1차원		t-SNE 2차원	
			perplexity 10	perplexity 20	perplexity 10	perplexity 20
1회	0.652	<b>0.752</b>	0.826	<b>0.857</b>	0.822	0.804
2회	0.661	0.678	<b>0.839</b>	0.813	0.778	0.800
3회	<b>0.696</b>	0.704	0.839	0.783	0.804	<b>0.857</b>
4회	0.687	0.696	0.826	0.835	<b>0.830</b>	0.839
5회	0.696	0.687	0.787	0.804	0.761	0.848
6회	0.657	0.713	0.787	0.800	0.778	0.835
7회	0.678	0.709	0.804	0.852	0.783	0.835
8회	0.665	0.717	0.809	0.800	0.770	0.839
9회	0.661	0.709	0.804	0.826	0.809	0.804
10회	0.683	0.687	0.783	0.809	0.761	0.804
평균 (표준 편차)	0.674 (0.016)	0.705 (0.021)	0.810 (0.021)	0.818 (0.024)	0.790 (0.025)	0.827 (0.021)

〈표 6-12〉의 혼동 행렬(confusion matrix)을 통해 생성변수만으로 학습한 모델은 저위험군을 약 45%의 오분류율(false positive error rate)을 보이고 있음을 알 수 있다. [그림 6-17]은 XGBOOST 모델의 피쳐 중요도를 살펴보기 위하여 shapley value 값을 산출하여 모델의 분류 예측 결과에 주요한 기여도를 미친 변수들을 확인해 보았다.

〈표 6-12〉 t-SNE 2차원 축소 자료(perplexity=20)의 3회차 샘플에서 혼동 행렬

생성변수			t-SNE 2차원 perplexity=20		
$y$	$\hat{y}$		$y$	$\hat{y}$	
0	0	63	0	0	100
0	1	52	0	1	15
1	0	18	1	0	18
1	1	97	1	1	97

[그림 6-17] t-SNE 2차원 perplexity=20의 3회차 샘플에서 설문 문항의 top-10 예측기여도(shapley value)



〈표 6-13〉은 해당하는 설문 문항의 기술이다. 우울함으로 인한 문제, 특수한 사건과 관련된 질문 변수는 저위험군에서는 거의 발견되지 않는다. 이러한 고위험군의 특성이 충분히 반영되는 질문 변수들에 반해, 니코틴과 알코올, 곤충, 새, 동물 같은 상황은 저위험군에서도 빈번히 발견됨을 알 수 있다. 향후 질문 변수별 자살과의 상관관계를 가중치로 하여 문장 임베딩 벡터를 학습시킨다면 고위험군 판별에 도움이 되는 질문 변수를 추출하는 데 용이할 것으로 사료된다.

〈표 6-13〉 t-SNE 2차원 자료(perplexity=20)의 3회차 샘플에서 top-10 예측기여도 (shapley value, SHAP) 상위 설문 문항

2차원의 설문 문항 번호	설문 문항	질문
484	E3_II	2주 이상 (슬프거나, 공허하거나, 우울하게 지냈던/ 또는 흥미를 잃고 지냈던) 때에, 아주 힘들게 일한 것도 아닌데, 거의 매일, 항상 기운이 없거나 피곤했습니까?
853	K45	사건으로 인해 생긴 문제 때문에, 잔치, 사회적 행사나 모임에 참석하지 못한 적이 있습니까?
482	E2	지금까지 사는 동안, 2주 이상 일, 취미, 또는 평소 좋아하던 것들 대부분에 흥미를 잃은 적이 있습니까?
17	B7_2	담배를 끊거나 줄이고 난 뒤, 처음 며칠 동안 평소보다 더 쉽게 짜증이 나거나 화를 내게 되었습니까?
692	J19	몸에 심한 병이 있어서, 술을 마시면 악화될 수 있다는 것을 알면서도 계속해서 마셨습니까?
64	D2	실제로 위험하지 않은데도, 곤충, 뱀, 새, 또는 다른 동물에게 가까이 가는 것을 피한 적이 있습니까?
680	J18_1	지금까지 사는 동안, 술 때문에 간 질환 또는 간염과 같은 건강 문제가 생긴 적이 있습니까?
809	K25	스트레스나 충격적인 체험을 하고 난 이후, 다시 사건이 일어난 것도 아닌데, 마치 그런 것처럼 갑자기 느껴지거나 행동했습니까?
774	K22_10_I	사람들에게 가끔 일어날 수 있는, 심한 스트레스나 충격적인 사건을 겪은 적이 있습니까?
66	D3	곤충, 뱀, 새, 또는 다른 동물을 (두려워 하거나/피하는) 것 때문에, 생활이나 활동에 지장을 많이 받은 적이 있습니까?

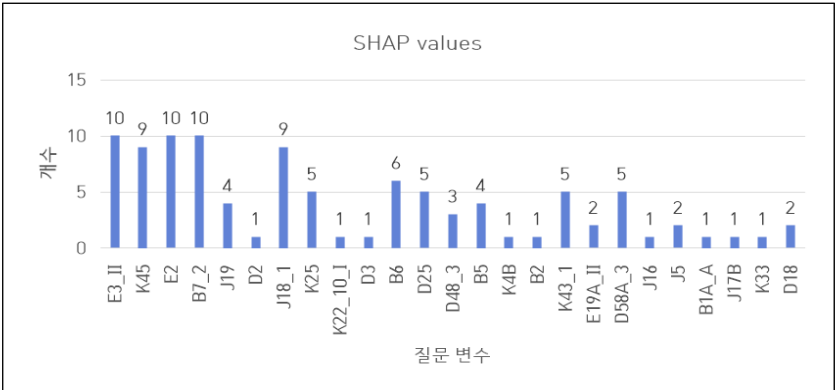
〈표 6-14〉는 t-SNE 2차원 자료(perplexity=20)의 1~10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항 번호를 나타낸다. 설문 문항 E3\_II, E2, B7\_2가 10회의 샘플에서 상위 10위 안에 모두 포함된 설문 문항임을 나타낸다. [그림 6-18]은 해당 문항의 빈도의 막대그림을 나타내며, 〈표 6-15〉는 해당하는 설문 문항을 나타낸다.



〈표 6-14〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측 기여도(shapley value, SHAP) 상위 설문 문항

sorted no	1위	2위	3위	4위	5위	6위	7위	8위	9위	10위
1회	484	17	10	149	482	853	13	680	844	398
2회	484	853	482	17	692	64	680	809	774	66
3회	484	853	482	17	680	13	188	322	10	724
4회	484	17	482	853	680	5	844	809	10	188
5회	484	853	482	17	844	680	809	10	528	398
6회	484	482	853	17	692	398	680	13	655	621
7회	484	853	17	482	680	621	188	0	13	809
8회	484	482	17	853	398	680	809	322	844	188
9회	484	482	853	188	528	17	692	398	679	13
10회	484	482	692	17	825	322	844	13	680	149

〔그림 6-18〕 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항



〈표 6-15〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측 기여도(shapley value, SHAP) 상위 설문 문항

설문 문항	횟수	질문
E3_II	10	2주 이상 (슬프거나, 공허하거나, 우울하게 지냈던/ 또는 흥미를 잃고 지냈던) 때에, 아주 힘들게 일한 것도 아닌데, 거의 매일, 항상 기운이 없거나 피곤했습니까?
E2	10	사건으로 인해 생긴 문제 때문에, 잔치, 사회적 행사나 모임에 참석하지 못한 적이 있습니까?
B7_2	10	지금까지 사는 동안, 2주 이상 일, 취미, 또는 평소 좋아하던 것들 대부분에 흥미를 잃은 적이 있습니까?
K45	9	담배를 끊거나 줄이고 난 뒤, 처음 며칠 동안 평소보다 더 쉽게 짜증이 나거나 화를 내게 되었습니까?
J18_1	9	몸에 심한 병이 있어서, 술을 마시면 악화될 수 있다는 것을 알면서도 계속해서 마셨습니까?
B6	6	실제로 위험하지 않은데도, 곤충, 뱀, 새, 또는 다른 동물에게 가까이 가는 것을 피한 적이 있습니까?
K25	5	지금까지 사는 동안, 술 때문에 간 질환 또는 간염과 같은 건강 문제가 생긴 적이 있습니까?
D25	5	스트레스나 충격적인 체험을 하고 난 이후, 다시 사건이 일어난 것도 아닌데, 마치 그런 것처럼 갑자기 느껴지거나 행동했습니까?
K43_1	5	사건으로 인해 생긴 문제 때문에, 의사가 아닌 다른 전문가와 상담을 했습니까?
D58A_3	5	불안발작 때문에 생활이나 활동에 지장을 많이 받았습니까?

고위험군과 저위험군 두 클래스 분류 모델의 예측값에 미치는 절대 기여도가 높은 상위 10개의 변수들을 확인해보았다. K 항목의 특정 사건 또는 스트레스로 인한 증상, D 항목의 불안 관련 증상이 클래스 분류에 주요한 기여를 한 것으로 생각된다. 또한, 저위험군에서도 빈번히 발견되는 B 항목과 J 항목에 대해, B7\_2, B6, J18\_1처럼 특정 증상이 발견될 경우에는 증독성 물질 의존도가 높은 상태이며, 정신적 문제의 연관성이 높을 가능성이 있다고 해석할 수 있다. E 항목의 변수가 모든 실험에서 높은

기여도를 보이는 것으로 보아 우울 관련 증상이 다른 항목들에 비해 정신 건강에 미치는 영향이 상대적으로 높다고 볼 수 있다.

## 2. 문장 임베딩을 활용한 자살 고위험군 세 집단 비교

S1 질문 변수에 ‘네’라고 응답한 고위험군 그룹은 자살 사고(class 0), 자살 계획(class 1), 자살 시도(class 2)의 세 단계로 위험 수준을 세분화할 수 있다. S1(자살하는 것에 대해 진지하게 생각한 적이 한 번이라도 있습니까?), S3(자살하려고 구체적으로 계획을 세운 적이 있습니까?), S5(자살을 시도한 적이 있습니까?) 변수의 응답을 타깃( $y$ )으로 하여 위험 수준을 분류하였다. 복수 응답이 가능한 점을 고려하여 위험 수준이 가장 높은 변수부터 ‘네’ 응답자를 선택하여 차례대로 응답자를 줄여 나가며 그룹을 라벨링을 통하여 S1(class 1)에 해당하는 응답자 574명 중 자살 사고 그룹은 421명, 자살 계획 그룹은 57명, 자살 시도 그룹은 96명으로 구분하였다. 학습은 앞의 분석과 동일한 생성변수, 선정 설문 문항, 선정 설문 문항의 문장 임베딩 벡터로 구성된 데이터 세트에 대해 진행하였으며, XGBOOST 모델을 사용해 성능을 비교하였다. 타깃 변수( $y$ )는 그룹 별로 라벨링하였다. 모델 파라미터도 앞의 분석과 동일하게 설정하였다. 단, <표 6-16>의 설문 문항은 자살 사고, 자살 계획, 자살 시도에 대한 직접적인 언급이 있기 때문에 이러한 설문 문항은 제외한 후 모델의 분류 성능도 확인해 보았다.

<표 6-17>은 XGBOOST 적합 결과를 나타낸다. 자살 관련 직접 언급 질문 변수를 제외한 경우와 제외하지 않은 경우 모두 t-SNE 2차원 축소 데이터에서 perplexity=20일 때의 분류 성능이 평균적으로 가장 높은 것으로 나타났다. <표 6-18>의 혼동 행렬을 통해 생성변수만으로 학습한

모델은 자살 시도(class 2)의 분류 성능이 t-SNE 2차원 perplexity=20으로, 학습한 모델에 비해 현저히 떨어지는 것을 알 수 있다.

〈표 6-16〉 자살 관련 직접 질문 변수

설문 문항	질문
E19_II	(우울했거나/흥미를 잃었거나/기운이 없었던) 때, 너무 우울해서, 자살해야겠다는 생각을 많이 했습니까?
E19A_II	(우울했거나/흥미를 잃었거나/기운이 없었던) 때, 자살을 어떻게 할지 계획을 세웠습니까?
E20_II	(우울했거나/흥미를 잃었거나/기운이 없었던) 때, 죽음에 이를 가능성이 없어 보이는 시도도 모두 포함하여 자살을 시도했습니까?

〈표 6-17〉 자살 고위험군 세 그룹에 대한 XGBOOST 분류 정확도

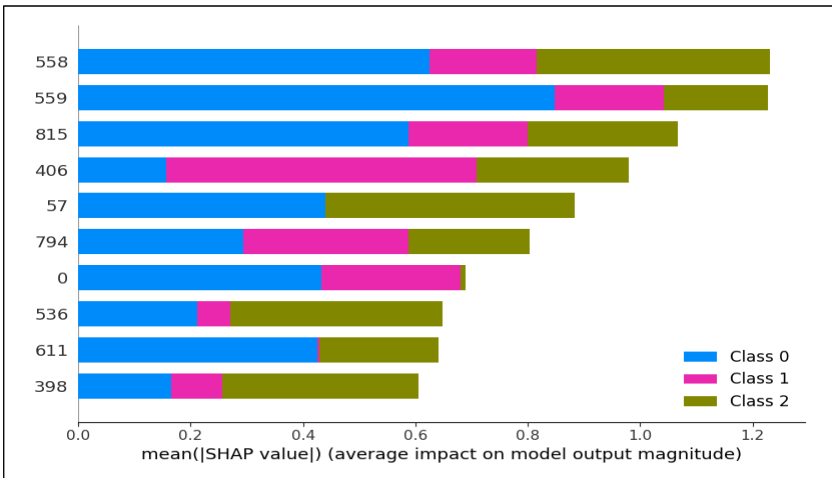
	생성변수	설문 문항	t-SNE 1차원		t-SNE 2차원	
			perplexity 10	perplexity 20	perplexity 10	perplexity 20
1회	0.713	0.643	0.783	0.852	<b>0.826</b>	<b>0.922</b>
2회	<b>0.722</b>	0.670	<b>0.791</b>	0.861	0.783	0.852
3회	0.704	0.626	0.757	0.870	0.826	0.878
4회	0.722	<b>0.687</b>	0.765	0.852	0.800	0.887
5회	0.635	0.670	0.739	0.843	0.791	0.878
6회	0.696	0.635	0.748	0.843	0.774	0.826
7회	0.696	0.626	0.748	<b>0.878</b>	0.765	0.852
8회	0.704	0.652	0.687	0.852	0.739	0.835
9회	0.687	0.643	0.757	0.826	0.783	0.878
10회	0.713	0.643	0.730	0.817	0.817	0.852
평균 (표준 편차)	0.699 (0.025)	0.650 (0.020)	0.751 (0.027)	0.849 (0.018)	0.790 (0.025)	0.866 (0.020)

〈표 6-18〉 혼동 행렬

생성변수				t-SNE 2차원 perplexity=20			
$y \backslash \hat{y}$	0	1	2	$y \backslash \hat{y}$	0	1	2
0	77	1	6	0	84	0	0
1	6	2	4	1	3	6	3
2	14	1	4	2	3	0	16

[그림 6-19]는 t-SNE 2차원 자료(perplexity=20)의 1회차 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항을 나타내는데, 57번과 611번은 class 0과 class 2를 분류하는 데 상대적인 중요도가 높은 설문 문항임을 확인할 수 있다. 반면 0번의 경우는 class 0과 class 1을 구별하는 데 상대적으로 기여도가 높은 설문 문항이라고 할 수 있다.

[그림 6-19] t-SNE 2차원 자료(perplexity=20)의 1회차 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항



〈표 6-19〉는 해당하는 설문 문항의 기술이다.

〈표 6-19〉 t-SNE 2차원 자료(perplexity=20)의 1회차 샘플에서 top-10 예측기여도 (shapley value) 상위 설문 문항

2차원상에서의 설문 문항 번호	설문 문항	질문
558	E34	지금까지 사는 동안, 가끔 괜찮은 때가 있었더라도, 2년 이상, 대부분의 날을 우울하거나 슬프게 지낸 적이 있습니까?
559	E34	지금까지 사는 동안, 가끔 괜찮은 때가 있었더라도, 2년 이상, 대부분의 날을 우울하거나 슬프게 지낸 적이 있습니까?
815	K31	스트레스나 충격적인 체험을 하고 난 이후, 위험에 대해 훨씬 더 걱정하거나 훨씬 더 조심하게 되었습니까?
406	D60	불안발작이 4주 안에 네 번 이상 생긴 적이 있습니까?
57	B16	담배를 피우기 위해서, 중요한 활동을 여러 번 포기하거나, 한 달 이상 포기한 적이 있습니까?
794	K22C	사건이 발생했을 때, 매우 무서웠습니까?
0	B1A_A	지금까지 사는 동안, 한 달 이상 매일 담배를 피운 적이 있습니까?
536	E25C	우울함과 관련된 증상 때문에 생활이나 활동에 지장을 많이 받았습니까?
611	J2A	지금까지 사는 동안, 1년에 12잔 이상을 마신 적이 있었습니까?
398	D58A_3	불안발작 때문에 생활이나 활동에 지장을 많이 받았습니까?

〈표 6-20〉은 t-SNE 2차원 자료(perplexity=20)의 1~10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항 번호를 나타낸다. 10회의 샘플에서 상위 10위 안에 모두 포함된 설문 문항으로 설문 문항 E34가 15회(2차원 각 변수별로 측정하여 최대 20회 기준임)로 가장 빈발하게 기여도가 높은 설문 문항으로 선별되었다. 다음으로 J2A, E25C, D60는 10회의 빈도로 기여도가 높은 설문 문항이다. [그림 6-20]은 해당 문항의 빈도의 막대그림을 나타내며, 〈표 6-21〉은 해당하는 설문 문항을 나타낸다.

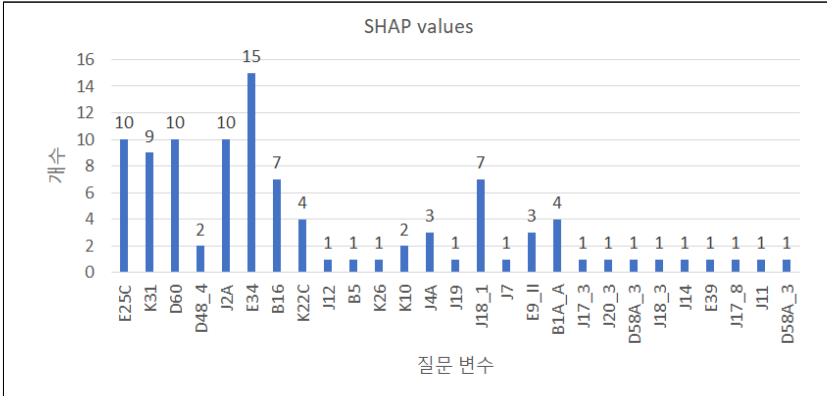
〈표 6-20〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항

sorted no	1위	2위	3위	4위	5위	6위	7위	8위	9위	10위
1회	536	324	558	815	406	794	802	611	57	570
2회	536	406	558	815	559	57	729	654	0	612
3회	536	406	558	57	559	611	612	692	674	499
4회	815	406	558	536	57	611	687	399	729	675
5회	406	815	558	536	611	635	674	794	678	665
6회	406	536	611	558	815	11	621	0	499	559
7회	324	558	406	536	611	815	57	0	674	675
8회	611	406	815	536	610	558	499	612	794	675
9회	558	406	536	611	559	815	57	675	643	630
10회	558	559	815	406	57	794	0	536	611	398

〈표 6-21〉 t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항

설문 문항	횟수	질문
E34	15	지금까지 사는 동안, 가끔 괜찮은 때가 있었더라도, 2년 이상, 대부분의 날을 우울하거나 슬프게 지낸 적이 있습니까?
J2A	10	지금까지 사는 동안, 1년에 12잔 이상을 마신 적이 있었습니까?
E25C	10	우울함과 관련된 증상 때문에 생활이나 활동에 지장을 많이 받았습니까?
D60	10	불안발작이 4주 안에 네 번 이상 생긴 적이 있습니까?
K31	9	스트레스나 충격적인 체험을 하고 난 이후, 위험에 대해 훨씬 더 걱정하거나 훨씬 더 조심하게 되었습니까?
J18_1	7	지금까지 사는 동안, 술 때문에 간 질환 또는 간염과 같은 건강 문제가 생긴 적이 있습니까?
B16	6	담배를 피우기 위해서, 운동이나 일, 친구나 친척들과의 모임 같은 중요한 활동을 포기하거나 많이 줄인 적이 있습니까?
B1A_A	4	지금까지 사는 동안, 한 달 이상 매일 담배를 피운 적이 있습니까?
K22C	4	사건이 발생했을 때, 매우 무서웠습니까?
J4A	3	지난 1년 중에, 남자는 5잔, 여자는 4잔 이상의 술을 2시간 내에 마신 날이 한달내에 5일 이상 있었나요?

[그림 6-20] t-SNE 2차원 자료(perplexity=20)의 1-10회의 각 샘플에서 top-10 예측기여도(shapley value, SHAP) 상위 설문 문항



자살 사고, 자살 계획, 자살 시도 세 클래스 분류 모델의 예측값에 미치는 절대 기여도가 높은 상위 10개의 변수들을 확인하였다. J 항목에 해당하는 질문 변수의 개수가 총 12건 발견되었다. 이는, 고위험군에 속하는 집단군이 알코올과 관련된 증상을 보유하고 있을 가능성이 높기 때문이라고 생각된다. 또한 E 항목의 우울과 관련된 질문 변수가 예측에 높은 기여를 한 것으로 확인되었다.

제6장의 자연어에 기반한 분석에서 투입한 전체 설문 문항은 <표 6-22>에 정리하였다. 이 중 집단 분류 시에 클래스의 분류 시 클래스 예측에 직접적으로 기여할 수 있는 설문 문항은 최대한 배제하였으나 내용에 따라 일부 포함되었을 가능성은 있을 수 있다.



〈표 6-22〉 정신장애진단도구(K-CIDI) 항목 중 사용 변수

분류 항목	변수	활용 비율
B (니코틴 사용에 의한 장애)	B1A_A, B1E_A, B2, B3, B4, B5, B6, B7_1, B7_2, B7_3, B7_4, B7_5, B7_6, B7_7, B7_8, B7_9, B7_10, B7_11, B7_12, B8, B10, B11, B12, B13, B13A, B14, B14A, B15, B16, B16A	30개/46개
E (우울 장애 및 기분부전 장애)	E1, E2, E3_II, E4_II, E5_II, E6_II, E7_II, E8_II, E8A_II, E9_II, E10_II, E10A_II, E11_II, E11A_II, E12_II, E12A_II, E13_II, E14_II, E15_II, E15A_II, E16_II, E17_II, E18_II, E19_II, E19A_II, E20_II, E22_II, E23_II, E24_II, E25A, E25B, E25C, E26A, E28, E29A, E29B, E29C, E30, E30A, E30B, E30C, E31, E34, E34A, E35, E36, E37, E38, E39, E40, E41, E42, E43, E44, E45, E46, E47, E48, E49, E50, E51, E52A, E52B, E52C, E54	65개/84개
J (알코올 사용과 관련된 장애)	J1, J1A, J2, J2A, J4A, J5, J5C, J6, J7, J7A, J7B, J8, J9, J11, J11A, J12, J12A, J13, J13A, J14, J14A, J15, J16, J17_1, J17_2, J17_3, J17_4, J17_5, J17_6, J17_7, J17_8, J17_9, J17_10, J17_11, J17B, J18_1, J18_2, J18_3, J18_4, J18_5, J18A, J19, J20_1, J20_2, J20_3, J20_4, J20A	48개/71개
K (강박장애 및 외상 후 스트레스 장애)	K1, K1A, K2, K3_1, K3_2, K3_3, K3_4, K3_5, K4, K4A, K4B, K5, K6, K7, K9, K10, K11, K12, K13, K14, K15, K15A, K16, K19, K20, K21, K22_1_I, K22_2_I, K22_3_I, K22_4_I, K22_5_I, K22_6_I, K22_7_I, K22_8_I, K22_9_I, K22_10_I, K22_11_I, K22_1_II, K22_2_II, K22_3_II, K22_4_II, K22_5_II, K22_6_II, K22_7_II, K22_8_II, K22_9_II, K22_10II, K22_11II, K22C, K22D, K23, K24, K25, K26, K27, K28, K29, K30, K31, K32, K33, K34, K35, K35A, K35B, K36, K37, K38, K39, K43, K43_1, K43_2, K43_3, K44, K45	75개/98개
D (공포 장애 및 기타 불안 장애)	D1, D1A, D2, D3, D4, D4A, D4B, D5, D6_1, D6_2, D6_3, D6_4, D6_5, D6_6, D6_7, D6_8, D6_9, D6_10, D6_11, D6_12, D6_13, D6_14, D9, D10, D11, D12, D12A, D12B, D13, D14_1, D14_2, D14_3, D14_4, D14_5, D14_6, D14_7, D14_8, D14_9, D14_10, D14_11, D14_12, D14_13, D14_14, D17, D18, D19, D20, D20A, D20B, D21, D22_1, D22_2, D22_3, D22_4, D22_5, D22_6, D22_7, D22_8, D22_9, D22_10, D22_11, D22_12, D22_13, D22_14, D25, D26, D27, D28, D28A, D28B, D29, D30_1, D30_2, D30_3, D30_4,	210개/259개

분류 항목	변수	활용 비율
	D30_5, D30_6, D30_7, D30_8, D30_9, D30_10, D30_11, D30_12, D30_13, D30_14, D33A_1, D33A_2, D33A_3, D33A_4, D33A_5, D33A_6, D33A_7, D34, D35A, D35B, D35C, D36, D37, D37A, D37B, D38, D39, D40_1, D40_2, D40_3, D40A_1, D40A_2, D40A_3, D40A_4, D40A_5, D40A_6, D40A_7, D40A_8, D40A_9, D40A_10, D40A_11, D40A_12, D40A_13, D43A_1, D43A_2, D43A_3, D43A_4, D44, D45, D45A, D46A, D46B, D46C, D47, D48_1, D48_2, D48_3, D48_4, D48_5, D48_6, D48_7, D48_8, D48_9, D48_10, D48_11, D48_12, D48_13, D48_14, D49, D50, D51, D51A, D51B, D54, D55, D55A, D56, D57_1, D57_2, D57_3, D57_4, D57_5, D57_6, D57_7, D57_8, D57_9, D57_10, D57_11, D57_12, D57_13, D57_14, D58, D58A_1, D58A_2, D58A_3, D59A, D59B, D59C, D60, D60A, D62, D63, D63B, D64A, D64B, D64C, D65_1, D65_2, D65_3, D65_4, D65_5, D65_6, D65_7, D65_8, D65_9, D65_10, D65_11, D65_12, D65_13, D65_14, D65_15, D65_16, D65_17, D65_18, D65_19, D65_20, D65_21, D65_22, D65_23, D65_24, D66A, D66B, D66C, D67, D68	
S (자살)	S1, S2A, S3, S4A, S5	5개/54개

## 제4절 소결

이 장에서는 정신건강실태조사 데이터를 활용하여 자살 사고, 자살 계획, 자살 시도 대상자의 응답 같은 스몰 데이터에서의 정보량을 최대화하기 위한 방안으로 기계학습과 딥러닝 방법을 수행하여 자살 고위험군의 집단적 특성을 파악해보았다. 개별 문항의 집계 점수가 가진 정보의 축약성을 극복하고자, 거대언어 모형을 활용한 설문 문항의 문맥을 파악하고 변수화하는 방법을 제안 및 활용하였다. 이로부터 산출된 점수를 활용하

여 XGBOOST의 자살 고위험군 분류분석의 정확도를 단순 설문 문항에 기초한 통상적인 방법에 비해 개선할 수 있음을 확인하였다. 질문 변수별 자살과의 상관관계를 가중치로 하여 문장 임베딩 벡터를 학습시킨다면 고위험군 식별에 도움이 되는 질문 변수를 추출하는 데 용이할 것으로 판단된다. 이 장에서 분석한 내용은 실험적으로 범주화된 응답 변수가 아닌 자연어 형태의 질문과 응답 데이터를 활용하여 자살 고위험군과 저위험군 분류 성능, 자살 고위험군 내에서 자살 사고, 자살 계획, 자살 시도 그룹의 분류 성능을 확인해 보았다는 데 의미가 있다. 자연어의 문장 임베딩에는 질문과 응답의 문맥적 관계가 반영되기 때문에, 질문의 깊이에 따라 다른 임베딩 값을 부여할 수 있고, 위험군 판별에 주요한 질문에는 더 큰 가중치를 두고 학습할 수 있을 것이다. 따라서 질문으로부터 획득할 수 있는 정보의 양이 증가할 수 있으며, 추가적으로 자연어의 문맥적 의미를 반영할 수 있다. 사전 학습된 모델들을 정신건강 자료에 대해 파인 튜닝(fine-tuning)하거나 프롬프트 튜닝(prompt-tuning)을 적용하면서 응답의 차이를 크게 하는 방향으로 학습한다면, 더 나은 성능의 임베딩 표현을 산출할 수 있을 것이라 기대한다. 그리고 충분히 학습된 모델을 보유한다면, 설문 문항으로부터 정신건강 질환 항목의 심각성 정도 또는 질환 보유 현황을 측정하는 데 미치는 영향의 크기를 정량적으로 해석할 수도 있을 것이라 기대할 수 있다.

이 연구의 한계점으로 자살 고위험군과 저위험군의 인구통계학적인 변수 정보를 모델에 반영하지 못하였는데, 이를 반영하면 결과의 정확도를 개선할 수 있을 것이다. 이 점은 추후 연구 과제로 남기고자 한다.





# 제7장

## 결론 및 시사점

제1절 결론

제2절 시사점: 정책적 활용 측면

제3절 시사점: 예측 모형 활용 측면



## 제 7 장    결론 및 시사점

### 제1절 결론

본 연구는 정신건강 사례를 중심으로 정책 대상별 특성을 파악하고 다양한 공공데이터를 활용하여 정책 대상별 예측 모델을 개발하는 것을 목적으로 하였다. 제2장에서는 정신건강 및 자살 고위험군에 대한 정의 및 범위와 관련된 위험요인에 대해 이론적 검토를 하였다. 정신건강 관련 정책 중 자살 고위험군에 대한 선별과 자살 위험에 대한 예측은 예방적 정신건강 정책의 핵심적 내용이라고 볼 수 있기 때문에 본 연구에서는 자살 고위험군에 대한 예측 모형 탐색을 중심으로 논의하였다. 제3장에서는 자살 고위험군 예측 모델을 적용하기 위해 활용 가능한 공공데이터로 ‘정신건강실태조사’와 ‘국민건강영양조사’, ‘청소년건강행태조사’, 그리고 종단적 연구가 가능한 ‘한국복지패널’과 ‘한국의료패널’을 중심으로 살펴 보았다. 이 5개 공공데이터명과 ‘자살’ 키워드가 초록에 포함되어 있는 KCI 등재 학술지를 검토하여 분석 변수와 분석 결과를 표로 정리하여 제시하였다. 검토 결과, 동일한 정책 대상이라 할지라도 어떠한 데이터를 활용하느냐에 따라 통계적으로 유의한 변수가 다양하게 나타나고, 동일한 데이터를 활용하는 경우에도 모형 설정에 따라 통계적으로 유의한 변수가 다르게 나타나고, 회귀 계수의 부호도 다르게 나타날 수 있음을 알 수 있었다.

검토한 공공데이터 중에서 문항 수, 응답 빈도 수(표본 수) 및 시계열적 연속성 등을 고려하여 예측 모델에 적용할 수 있는 데이터 및 심층 분석

에 활용할 수 있는 데이터는 ‘정신건강실태조사’, ‘청소년건강행태조사’ 그리고 ‘한국복지패널’로 선정하였다.

제4장에서는 ‘청소년건강행태조사’와 ‘정신건강실태조사’를 활용하여 정책 대상별 예측 모형을 구축하고 결과를 제시하였다. 제4장의 예측 모델링의 주요 결과를 보면, 아동·청소년의 경우 제2장, 제3장에서 언급한 아동·청소년 자살 행동의 위험요인인 개인적 특성(성별, 우울, 불안), 환경적 요인(학교 내 문제), 부정적 생활사건(스트레스 사건) 요인이 유의미한 요인으로 도출되었으며, 성인의 경우도 마찬가지로 제2장, 제3장에서 살펴본 인구사회학적 요인(실직, 주관적 스트레스), 정신증적 요인(우울 증, 외상 후 스트레스 장애, 알코올) 등이 유의미한 요인으로 도출되었다. 예측 모델링의 주요 결과를 정책 대상별로 공통적인 요인과 특징적인 요인을 나누어서 <표 7-1>에 제시하였다.

제5장에서는 한국복지패널 데이터에서의 정신건강 관련 문항을 분석했는데, 자살 생각을 event로 고려하여 이와 관련된 요인들을 심층적으로 분석하였다. 생존 분석 적용 결과, 활용된 모든 다변수 분석에서 유의했던 변수는 우울과 좌절감이었고, 불면, 무기력, 건강상태, 가부장적 관념 등의 변수들도 정신건강에 영향을 미치는 유의미한 요인으로 도출되었다. 자살 생각을 1회 이상 하는 것과 2회 이상 하는 것으로 각각 정의하였을 때, 2회 이상인 경우 가족관계 만족도, 가족생활 만족도, 자녀들의 형제자매 관계 만족도 등의 변수들이 새롭게 유의미한 변수로 도출되었다.

제6장에서는 자살 생각, 자살 계획, 자살 시도에 “예”라고 응답한 대상자들이 적은 스몰 데이터에 대해 데이터 자체에서 내재하는 정보의 양을 최대한 활용하는 방법으로 자연어처리 기반 분석 방법인 딥러닝을 정신건강실태조사 데이터에 적용해보았다. 자연어 형태의 설문 조사 질문과 응답 데이터에 내포된 문맥을 반영할 수 있기 때문에 향후 질문 변수별 자살과의 상관관계를 가중치로 하여 문장 임베딩 벡터를 학습시킨다면



자살 고위험군 판별에 도움이 되는 질문 변수를 추출하는 데 활용할 수 있다.

〈표 7-1〉 예측 모델링 주요 결과

			주요 유의미한 요인	
데이터 및 대상별 분류			요인	
청소년 건강 행태 조사	공통적		주관적 건강 인지, 슬픔·절망감 경험 여부, 외로움 경험 정도, 평상시 스트레스 인지 정도	
	특징적	남자 중학생	최근 폭력으로 인한 치료 경험, 흡연 경험, 타인 음주로 인한 피해, 가정의 경제적 상태가 어려움, 현재 거주형태가 가족과 함께 거주하지 않음, 낮은 학년	
		여자 중학생	학습할 목적으로 주말에 앉아서 보낸 평균 시간이 많음, 학교에서의 손상 경험, 최근 폭력으로 인한 치료 경험, 음주 경험, 흡연 경험, 치료 목적을 제외한 습관적 약물 경험, 가정의 경제적 상태가 어려움, 현재 거주형태가 가족과 함께 거주하지 않음, 낮은 학년	
		남자 고등학생	일반계 고등학교, 최근 폭력으로 인한 치료 경험, 성관계 경험	
		여자 고등학생	최근 폭력으로 인한 치료 경험, 음주 경험, 흡연 경험, 성관계 경험, 치료 목적을 제외한 습관적 약물 경험, 스마트폰 주말 평균 사용 시간이 많음, 가족과 함께 거주하지 않음, 가정형편으로 인한 경제적 도움을 받음	
정신 건강 실태 조사	공통적		우울장애, 삶의 질 만족도, 현재 삶에 대한 만족도	
	특징적	남자 (18_39)	만연령, 정신건강 문제로 전문가와 상의한 경험	
		여자 (18_39)	정신건강 문제로 전문가와 상의한 경험, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 경험, 만연령, 가구 총소득	
		남자 (40_64)	만연령, 자녀의 수, 직업이 없는 경우, 의료급여 수급권자, 의료급여 수급권자가 아닌 경우에는 가구 총소득이 높은 경우, 알코올 사용장애, 불안장애, 외상 후 스트레스 장애	
		여자 (40_64)	알코올 사용장애, 니코틴 사용장애, 외상 후 스트레스 장애	
		남자 (65_79)	가구 총소득, 불안장애, 정신건강 문제로 전문가와 상담한 경험, 정신건강 문제로 전문가로부터 약물처방이나 약물치료를 받은 경험	
		여자 (65_79)	자녀의 수, 가구 총소득, 만연령, 정신건강 문제로 전문가와 상의한 경험, 알코올 사용장애	

## 제2절 시사점: 정책적 활용 측면

### 1. 자살 고위험군에 대한 선별 방안

자살 고위험군에 대한 선별 및 자살 위험도를 측정하기 위한 다양한 평가 도구들이 정신건강 관련 현장에서 활용되고 있다. 이러한 도구들은 자살 고위험군에 대한 스크리닝이나 자살 위험 수준에 대한 평가를 통하여 위기 대응의 근거로 활용된다. 이때 자살 고위험군에 대한 대응 전략의 강화를 위하여 다음의 사항들이 고려되어야 한다.

첫째, 평가 도구 간의 연계를 통하여 자살 위험 수준의 우선순위를 설정할 필요가 있다. 자살 위험 선별 도구가 여러 개 활용되는 경우에 각 결과에 대한 객관적 비교가 어렵다는 한계를 가지며, 단일 평가도구를 활용할 경우에는 고위험군이 과잉 추정될 수 있다(유화정, 한유화, & 유성은, 2023). 따라서 자살 고위험군의 수준을 파악하기 위해서는 단일 측정 도구의 활용 결과를 통한 고위험군의 설정과 함께 복수의 평가 도구를 활용하여 고위험군 내에서의 개입 우선순위를 설정하는 것이 필요하다. 특히 자살 행동 경험이나 자살 시도의 심각성 등 자살 위험성에 상당한 영향을 미칠 수 있는 요인들이 측정된 문항의 결과는 개입 우선순위 설정에 핵심적인 근거로 활용될 수 있다. 연령대별, 성별로 자살 생각에 영향을 미치는 주요 요인은 제4장의 연구 결과처럼 우울 장애 등 공통적인 요인도 있지만, 연령대별 특성과 성별의 특성에 따른 자살 위험요인은 다르게 나타난다. 따라서 본 연구의 분석 결과를 바탕으로 정책 대상별로 주요 요인에 대한 가중치를 다르게 두어 평가 측도를 개발한다면 활용성이 더 높아질 것이다.

둘째, 자살 위험 정도를 신속하고 간결하게 측정할 수 있는 문항의 도

입이 요구된다. 자살 고위험군 선별을 위한 문항을 활용하는 경우에는 평가에 다소 시간이 소요되기 때문에 신속한 개입 또는 관리에 한계가 있다. 따라서 자살 행동의 예방적 측면을 강화하기 위한 단축형 척도의 개발 및 활용이 강조될 필요가 있다(윤경서 외, 2023). 이를 위해 제6장에서 기술한 심층 분석 방법을 활용할 수 있을 것이다. 단축형 척도를 개발할 때 기존 데이터의 질문 변수별 자살과의 상관관계를 가중치로 하여 문항 임베딩 벡터를 학습시킨 후, 자살 고위험군 판별에 도움이 되는 질문 변수를 추출하여 활용한다면, 자살 위험 정도를 더 빠르게 측정할 수 있을 것이다. 그리고 충분히 학습된 모델을 가지고 있다면, 설문 문항으로부터 정신건강 질환 항목의 심각성 정도 또는 질환 보유 현황을 측정하는데 미치는 영향의 크기를 정량적으로 해석할 수도 있을 것이다.

정신건강 관련 전문 기관 내에서의 전문적인 검사 이전에 지역사회나 일반적인 생활 환경 내에서 간편하고 빠르게 자살 경향을 파악하는 것이 필요하다. 이를 통하여 정신건강 문제에 대한 1차 예방을 강화하고 자살 고위험군에 대한 전문 자원의 접근 가능성을 증대시킬 수 있다.

## 2. 정신건강 문제 고위험군에 대한 대응 방안

정신건강 문제 및 자살 고위험군에 대한 대응은 그 상황에 따라 세분화될 수 있다. 자신 또는 타인에게 해를 가할 위험성의 정도와 시점에 따라 대응 시기와 방안이 달라진다. 예를 들어 자신 또는 타인에게 즉시 해를 가할 위험성을 가진 경우에는 응급 상황에 대한 대응 및 조치가 우선적으로 이루어진다. 위험 정도가 높은 경우에는 정신건강 문제 관련 자원의 즉각적인 대응이 이루어지며, 위험 정도가 낮거나 적절하게 관리될 수 있는 경우에는 정신건강 관련 기관을 통한 사례관리나 필요한 다른 자원으

로의 연계 등이 이루어질 수 있다.

〈표 7-2〉 자살 위험 척도 및 특성

척도명	특성	출처
한국판 컬럼비아 자살 심각도 척도 (Columbia Suicide Severity Rating Scale, C-SSRS)	- 자살 생각(심각도, 강도), 자살 행동(행동 및 치명도)을 구분 - 방해된, 중단된 자살 시도에 대한 문항 포함	Posner, K., et al. (2008)
한국판 자살 경향에 대한 하위 척도 (Korean Depressive Symptom Index -Suicidality Subscale, DSI-SS)	- 최근 2주 동안의 자살 생각 측정 - 자살 생각의 빈도, 생각의 강도 및 계획, 자살 생각에 대한 통제 가능성, 자살 충동 등 측정	Suh, S., Ryu, H., Chu, C., Hom, M., Rogers, M. L., & Joiner, T. (2017)
자살 행동 척도 (Suicidal Behaviors Questionnaire-Revised, SBQ-R)	- 자살 생각 또는 자살 시도 경험, 지난 1년 동안의 자살 생각 빈도, 자살에 대한 의사 전달 및 의도, 자살 시도 가능성	Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001)
한국어판 벡 자살 사고 척도 (Korean version of the Beck Scale for Suicide Ideation, K-BSS)	- 자살 사고 여부 및 심각도 측정 - 자살 욕구 수준, 자살 시도 및 자살 의도의 심각성 문항 포함	Beck, A. T., & Steer, R. A. (1993)

위기 상황에 따른 대응 모델이 적용되기 위해서는 정신건강 고위험군에 대한 선별과 위기 상황에 대한 정보 습득 및 현 상황에 대한 정보 갱신 등이 신속하게 이루어져야 한다. 이를 위해서는 정신건강 문제 또는 위기 상황과 관련된 위험/보호 요인에 대한 사전 정보 습득과 정신건강 문제의 변화 양상을 파악하기 위한 지속적인 관리가 요구되며, 정책 대상별 특성을 반영한 주요 요인에 대한 파악이 필수적이다. 특히 이러한 과정을 통하여 습득된 정보는 응급 상황에 대한 신속한 대응, 시기 및 상황에 맞는

자원 연계 등을 위하여 활용될 필요가 있다. 제5장의 한국복지패널 분석 결과, 자살 생각을 1회 이상 한 대상자와 2회 이상 한 대상자로 비교해보면, 가족관계 만족도, 가족생활 만족도, 자녀들의 형제자매 관계 만족도 등의 변수들이 자살 생각을 2회 이상 한 대상자 분석에서 유의미하게 나왔다. 이는 자살 생각을 2회 이상 했던 대상자들에게는 가족과 관련된 정책 및 서비스가 필요하다고 볼 수 있다.

이를 위해 정책 대상별로 다양한 예측 모형을 적용하여 유의미한 요인을 도출하고, 현재 분석 방법의 유의미한 변수들의 예측력이 시간이 흐르면서 추후 예측력을 follow up했을 때 안정적(stability)으로 만족할 만한 결과를 주는지에 대해 확인하고, extended/update된 모형에서도 유사한 결과가 나오는지 추적 관찰하는 것도 필요하다.

〈표 7-3〉 정신건강 상황에 따른 대응 모델

구분	상황	대응방법 및 시기	조치
A. 비상 상황	현재 자신 또는 타인에게 해를 끼칠(자살 시도, 폭력 등)	(즉시) 응급 서비스 응급, 경찰 및 구조 등	응급 서비스 도착 전까지 지속적 연락
B. 자신 또는 타인에게 즉시 해를 끼칠 수 있는 상황	자신 또는 타인에게 해를 끼칠 계획을 가지거나 관련 이력을 가진 경우	(4시간 이내) 긴급 대응 관련 위기대응팀	상황 변화 및 추가 정보에 대한 즉시 확인
C. 자신 또는 타인에게 해를 끼칠 위험이 높거나 이를 대응하기 위한 자원이 없는 경우	자신 또는 타인에게 해를 끼칠 위험성이 급격하게 증가한 상황	(24시간 이내) 긴급 대응 위기대응, 지역사회 정신건강 관련 팀	1-2일 내 연락 및 검토 상황 변화 및 추가 정보 확인
D. 증증도의 피해 또는 심각한 위험	심각한 정신건강 문제가 확인되어 우선적인 개입이 요구되는 상황	(72시간 이내) 중간 수준의 긴급 대응 지역사회 정신건강 관련 기관	전화 지원 및 상담 상황 변화에 따른 연락
E. 적절한 개입 또는 안정적인 요인들로 인한 단기적이고 낮은 수준의 위험	안정적인 상태이나 전문적인 정신건강 평가가 필요한 상황	(4주 이내) 긴급하지 않은 대응 지역사회 정신건강 관련 기관	전화 지원 및 상담 상황 변화에 따른 연락
F. 대면 응답이 요구되지 않는 정신 건강 문제	정신건강 외 적절한 지원의 필요	관련 자원을 통한 대응 상담 및 관련 자원 연계 등	자원 연계 및 관리
G. 상담 정보	요청에 따른 정보 제공	정보 제공 상담 및 관련 자원 연계 등	지속적인 사례관리

주: 원 자료 내용의 편집 및 요약

자료: Sands, N. Elsom, E. Colgate, R & Haylor, H. (2016) Development and inter-rater reliability of the UK Mental Health Triage Scale (In Press). International Journal of Mental Health Nursing, p. 333.

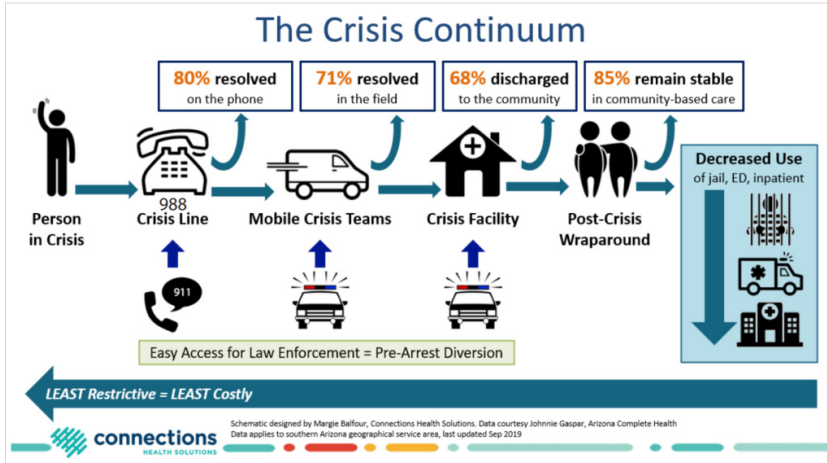
### 3. 자살 고위험군 관련 인프라

이 연구에서는 분석 대상인 개인의 연령대, 성별 등 관련 인구학적 특성에 대해 심층 분석이 이루어져 예측 모형 자체는 다양한 접근법과 요인들이 판별되었다. 반면에 예방적 측면에서는 개개인의 특성에 대해 사전에 심층적으로 모니터링할 수 있는 인프라 구축이 중요하다. 해외 사례에서는 학교나 회사에 국한하여 사전에 고위험군을 선별할 수 있는 경우가 있다.<sup>1)</sup> 그 이유는 학생이나 직원에 대한 카운슬링 등이 상대적으로 일반인 대다수보다는 교류가 이루어지는 빈도가 높고 접근 가능성에 대한 물리적 한계가 적기 때문이다. 불특정 다수 일반인에게는 이러한 접근성이 용이하지 않다는 점이 예측의 가장 어려운 점이기에 때문에 해외에서는 지역 인구밀도당 인프라 비율로 위험 예방률을 산정하는 경우가 많다[참고: <https://health.wyo.gov/publichealth/prevention/wivpp/suicide-prevention/>].

따라서 자살 예방을 위해서는 본 연구에서 이루어진 개인별 위기 분석(personal crisis) 및 예측과 함께 핫라인, 투입 배치 가능한 인력, 커뮤니티 인프라 등의 컨소시엄에 대한 실태 조사, 특히 지역적 분포나 접근 용이성이 함께 고려되어야 한다. 이 부분은 자살 고위험군에 대한 조기 개입의 근거로 지역사회 및 위기가구(저소득, 사회적 고립, 기타 사회적 위험성에 노출된 경우)의 취약성 등이 사전에 검토되고 있다는 것과 연결 지을 수 있으며, 지역사회에 기반한 정신건강 서비스가 이루어질 수 있도록 다차원적으로 검토되어야 한다.

1) <https://www.preventsuicideny.org/a-guide-for-suicide-prevention-in-new-york-schools/>에서 2023.11.29. 인출.

[그림 7-1] 위기대응 프로세스 예시



자료: Todd, C. (2022, August 18). A month into 988, call volumes up, but state behavioral health infrastructure isn't. Nevada Current. <https://www.nevadacurrent.com/2022/08/18/a-month-into-988-call-volumes-up-but-state-behavioral-health-infrastructure-isnt/>에서 2023.11.29. 인출.

### 제3절 시사점: 예측 모형 활용 측면

#### 1. 세분화 모형의 장점 및 고려사항

본 연구는 전체 데이터를 사용하지 않고 정책 대상별(성별, 연령대별)로 데이터를 나누어 세분화된 모형을 적용하였다. 특정 변수의 범주(혹은 범위)별로 데이터가 상이한 분포를 가지고 있을 경우 세분화 모형이 일반 모형보다 더 좋은 성능을 기대할 수 있다. 만약, 범주별로 데이터가 비슷한 분포를 가진다면 세분화 모형을 사용하는 것은 큰 의미를 가지지 못할 수도 있지만, 예측 모형에서 고려하는 독립변수가 성별이나 연령에 따라 달라질 수 있어서 독립변수와 반응변수와의 관계에서 성별, 연령이 교란



변수(confounding variable)가 될 수 있다. 교란변수의 고려 없이 전체적으로 독립변수와 반응변수의 관계를 분석하면 그 어떤 대상(특정 성별이며 특정 연령)에게도 해당되지 않는 평균적인 관계만을 파악하게 될 수도 있다. 이러한 경우에는 교란변수의 항목별로 독립변수와 반응변수의 관계를 구분해서 파악하는 것이 결과적으로 각 연령이나 성별에 더 적합한 정책적 대응 방안을 제시하는 데 필요한 기초 자료가 될 수 있기 때문에 더 효과적인 분석이 될 것이다.

세분화 모형의 장점은 거대 자료 분석이 용이하다는 점이다. 거대한 자료를 한 번에 분석하는 것은 시간적, 계산적 어려움이 있을 수 있기 때문에, 세분화 모형을 고려할 경우 상대적으로 크기가 작은 세부 데이터 분석을 병렬적으로 수행할 수 있으므로 효율성이 증대된다. 예를 들어,  $n$ 개의 행을 가지는 데이터에 K-nearest neighbors 분석을 수행할 경우 대략  $O(n \log(n))$ 의 계산 시간이 소요된다. 따라서 10개로 데이터를 세분화하고 각각 K-nearest neighbors 분석을 (병렬적으로) 사용하면 대략적으로 10배 이상의 시간 단축이 가능하다.

다만, 세분화하려는 대상 변수의 범주가 불균형할 경우 소수의 범주에서의 세분화 모형은 과적합이 일어날 가능성이 높다. 더 나아가, 범주마다 희소성 비율이 불균형할 경우에 모형 적합 자체가 불가능할 수 있다. 따라서 데이터 분석 시, 몇 가지의 범주를 병합해서 세분화 모형을 사용할 필요도 있다.

## 2. 희소 데이터(sparse data)

한국복지패널을 활용한 심층 분석인 생존 분석 모형 적용에서 결측 비율에 따라 공변량 활용 정도를 구분하고 각각 결과를 제시하였다. 실제

통계 분석 시, sparse data에 대한 충분한 이해가 선행되어야 한다. 예를 들어, 결측치가 실제로는 관측이 될 수 있는 값임에도 결측이 된 것인지(예: 이혼 여부의 경우 작성자가 공개를 원치 않는 경우 결측치가 될 수 있음), 혹은 결측치 그 자체로 의미를 갖는 값인지 파악해야 할 필요가 있다.

이를 위해 결측치 대체법(imputation)을 고려해볼 수 있다. 거의 모든 행에 결측값이 존재하는 경우, 행 제거를 통한 분석이 불가하기에 결측값을 대체하는 작업이 선 수행되어야 한다. 대체하는 방식에는 평균, 중앙값, 최빈값 등을 사용하는 단순한 방식부터 MICE(Multivariate Imputation by Chained Equation)로 대표되는 기계학습을 이용하는 알고리즘까지 매우 다양하게 존재한다. 분석 방향에 따라 최적의 대체 알고리즘이 다를 수 있기에 세심한 선택이 필요하다. 더 나아가, 대부분의 값이 결측치인 변수의 경우 대체하는 것이 별 의미가 없을 수 있으므로 데이터 분석 시 제거하는 것을 고려할 필요도 있다.

희소 데이터 분석을 지원하는 분석 알고리즘도 고려할 수 있다. 통계 및 기계학습 데이터 분석 방법론 구현을 지원하는 패키지(또는 모듈) 중에는 희소 데이터 및 결측치가 포함된 데이터 분석을 지원하지 않는 것들이 존재하기 때문에 방법론 및 구현 패키지를 선택할 때 희소성을 처리할 수 있는지를 먼저 고려해야 할 것이다.

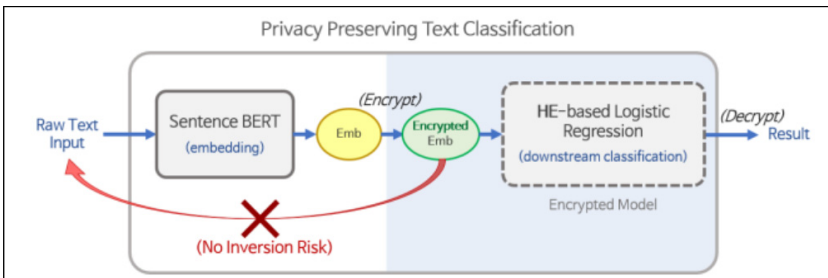
### 3. 개인정보를 보호할 수 있는 예측 모델링

향후 정신건강 데이터 분석에 기초하여 연구 방향을 다음으로 정리할 수 있다. 최근 이상동기 범죄(묻지 마 범죄)나 무차별 범죄 등이 급증하고 있는데 외래치료 명령제, 외래치료 지원제도 같은 자타해 위험성이 있는 개인을 관리하는 제도의 내실 있는 도입이 필요할 것으로 사료된다. 기존에 운영 중인 여러 프로그램들에서 획득되는 정형, 비정형의 다양한 기록

데이터는, 예를 들어 주소, 이름, 채팅 메시지, 이메일 주소, 사건기록, 약 물처방기록, 학적부 생활기록, 개인 신병 서술 등의 상담기록(텍스트)은 극민감 정보를 직접적으로 담고 있는 경우가 많으므로 프라이버시 이슈가 지속적으로 발생할 수 있다. 개인정보를 보호할 수 있는 예측 모델링 개발이나 방안이 필요하다고 할 수 있다.

본 연구에서는 사전 학습된 모델을 정신건강 자료를 이해하는 언어 모형으로 파인 튜닝하여 구축하고, 이로부터 임베딩을 추출해 특징을 추출하여 변수화하는 방안이 유효한 방법으로 활용될 수 있음을 확인하였는데, 예측 모델링을 고도화하기 위해서 극민감 정보를 활용하는 추가적인 방안이 필요하다. [그림 7-1]은 주어진 텍스트 정보로부터 sentenceBERT로 임베딩하고 이를 암호화(encryption)하여, 동형암호화(homomorphic encryption)된 상태에서 로지스틱 회귀분석을 적용하는 모식도를 나타낸다. 최종 스코어 값은 암호화되어 있으므로 암호키로 복호화(decryption)하면 실제값을 확인할 수 있다. 텍스트 원본은 노출되지 않고 로지스틱 회귀분석을 통한 위험도를 산출할 수 있음을 의미한다.

[그림 7-2] 개인정보 보호 텍스트 분류

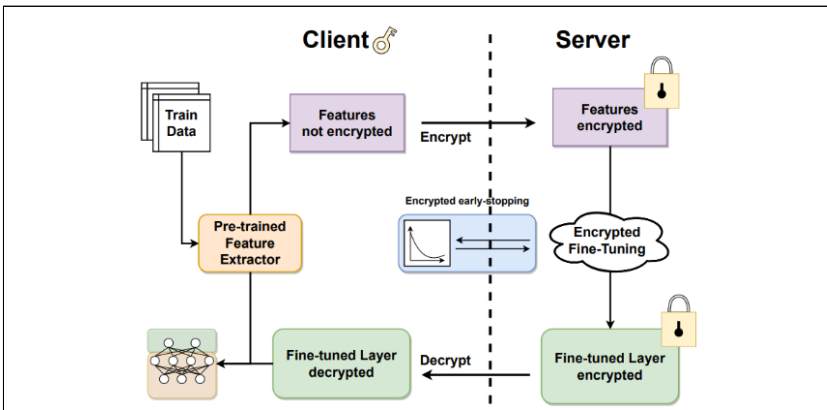


자료: Lee, G., Kim, M., Park, J. H., Hwang, S. W., & Cheon, J. H. (2022). Privacy-Preserving Text Classification on BERT Embeddings with Homomorphic Encryption. arXiv preprint arXiv:2210.02574. p.2.

좀 더 일반화하면, 개인정보를 보호하는 방안으로 이중의 장치를 고려해볼 수 있다. 첫 장치는 각 기관이 보유한 개별 모델의 임베딩 벡터로 원본 데이터(original data)를 요약하여 고차원의 성향 벡터로 변수화한다. 이를 통해, 1차적인 데이터 노출 위험도를 줄일 수 있다. 두 번째 장치는 암호화로 서로 다른 기관 간의 결합 분석 시 데이터를 보호할 수 있다.

구체적으로 살펴보면, 암호화된 앙상블 모델은 여러 기관의 원본 데이터를 합치는 것보다는 원본 데이터로부터 가공된 데이터 형태를 취합하는 구조를 가지고 있다. 가공된 데이터 형태는 논문에서의 data2vec 같은 모델의 output인 임베딩 벡터를 일컬으며, 이러한 벡터들을 여러 기관의 고유한 모델을 활용하여 산출하고 결합한다. 전술한 바와 같이 임베딩과 이를 암호화하는 이중의 데이터 보호 장치를 통해 원본 데이터를 보호할 수 있다.

[그림 7-3] Transfer learning을 활용하여 client의 암호화된 정보를 server에서 연산하는 모델 구성도



자료: Lee, S., Lee, G., Kim, J. W., Shin, J., & Lee, M. K. (2023, July). HETAL: Efficient privacy-preserving transfer learning with homomorphic encryption. In International Conference on Machine Learning (pp. 19010-19035). PMLR. p.4.

가명, 익명, 재현 등의 데이터 보호 기술들은 각 기술이 가진 장단점이 있는데, 동형암호의 경우에는 원본 자료를 암호화함으로써 야기되는, 폭증되는 데이터 크기와 연산기 등의 여러 기술적 요소의 어려움을 가지고 있다. 설령 데이터가 결합된다 하여도 이후의 데이터 분석 과업 수준에는 못 미칠 수 있다. 대안적으로 임베딩 벡터의 데이터 연계를 통해 생성형, 예측 모델링에 활용하는 방안을 고려할 수 있다. 다만, 원본 데이터가 아니므로 원본 데이터 복기 요청이 있을 경우 확인이 어렵다는 단점을 가진다.

가능한 도출할 수 있는 서비스의 방향은 크게 두 가지로 요약할 수 있다. 먼저, 정신질환자 질적 상담자 인터뷰 데이터를 BERT로 fine-tuning하여 정신건강 자료를 이해하는 언어 모형을 구축하고, 모델을 암호화한 후, 웹 등으로 인입(평가 요청)되는 암호화된 [CLS]벡터와의 연관성을 추론하여 정신질환자 텍스트 유사도를 추출(정신질환 기능 평가)하여 평가할 수 있다. 서비스 이용자는 원문을 공개하지 않고도 정신건강평가 모델을 통해 유사도를 측정할 수 있으며, 기존 대표 텍스트(barycenter)와의 유사도 평가도 가능하다. 또한, 평문에서 정신건강 관련된 서비스 기능별로 모델을 구축할 경우 다차원의 유사도 측정이 가능하다. 정신건강 실태조사의 경우에는 B/C/D/E 같은 영역이 이에 해당된다.

한편, 데이터 연계나 결합을 통해 여러 기관으로부터 원본 데이터가 아닌 임베딩 벡터를 제공받고 이 정보를 앙상블하여 의사결정(decision model)을 구축할 수 있다. 예를 들어, 정신질환과 관련된 처방이력을 기초로 LLM 등을 이용한 모델을 암호화할 경우, 이종기관 간의 데이터 결합이 가능하다.  $f_1$ 을 사회보장정보원 같은 정신건강을 평가하는 모델,  $f_2$ 를 개인건강기록시스템(personal healthcare record, PHR) 정신질환 모델,  $f_3$ 를 교육/주거데이터 등으로부터 생성된 모델이라고 하자. 이

들 데이터 연계와 결합을 통해 암호화된 모델들로부터 암호화된 정보를 추출, 활용할 수 있다. 우울증 등의 음성기록이나 영상 등의 자료는 동일한 방식으로 각각 wav2vec과 ViT 같은 Transformer를 활용한 특징추출기 같은 다양한 backend 모형을 통해 전이 학습이 가능하다. 이는 다기관 데이터를 멀티 모달리티(multimodality) 관점에서 해석할 수 있음을 시사하며, 이러한 모형 개발은 향후 연구주제로 남기고자 한다.



- 강동훈, & 김운태. (2018). 사회경제적 박탈 경험이 노인의 자살생각에 미치는 영향: 6 가지 박탈 유형을 중심으로. *한국노년학*, 38(2), 271-290.
- 강은정, 최은진, 송현중, 유근춘, 김나연, 남정자, ..., & 이연경. (2006). 국민건강 영양조사 제3기 (2005): 총괄. 보건복지부·한국보건사회연구원.
- 강지원, 장숙량, 김해송, 김민경, 김현수, 주지영, ... & 김유경. (2023). 같지만 다른 그들, 청년: 성별 자살생각과 자살 시도 영향요인의 탐색 연구. *보건사회연구*, 43(1), 69-84.
- 관계부처 합동. (2018). 자살예방 국가 행동계획.
- 국립정신건강센터, 건강보험심사평가원, & 한국보건사회연구원. (2022). 국가정신건강현황 보고서 2021. 국립정신건강센터.
- 김동배, & 안인경. (2004). 한국인의 정신건강 개념에 관한 연구. *한국사회복지학*, 56(1), 203-233.
- 김미숙, & 김윤영. (2018). 물질적 박탈과 음주가 자살행위에 미치는 영향-한국복지패널데이터 2012-2017년 종단연구를 중심으로. *한국케어매니지먼트 연구*, (29), 63-88.
- 김수진, & 김봉환. (2015). 청소년 내담자의 반복적 자해행동의 의미탐색: '고통으로 고통을 견디기'. *한국심리학회지: 상담 및 심리치료*, 27(2), 231-250.
- 김영택, 김인순, 주영수, & 이지연. (2009). 여성자살 현황 및 정책방안. 한국여성정책연구원.
- 김영희. (2020). 중년여성의 신체활동과 우울 및 자살생각에 대한 융합적 연구. *차세대융합기술학회논문지*, 4(6), 653-661.
- 김현주. (2019). 한국 노인의 자살생각에 관한 융복합 예측요인: 패널자료분석. *한국융합학회논문지*, 10(7), 391-397.
- 남석인, 이희정, 김성준, & 김보경. (2016). 노인의 계층인식이 자살충동에 미치는 영향과 스트레스와 주관적 건강인식의 이중매개효과. *보건사회연구*, 36(2), 121-150.

- 남정자, 김혜련, 이상호, 최은영, 윤강재, 박인화. (1999). 1998년도 국민건강·영양조사: 총괄보고서. 보건복지부·한국보건사회연구원.
- 뉴욕주 정신건강사무국(Office of Mental Health). (2023). <https://www.preventsuicideny.org/a-guide-for-suicide-prevention-in-new-york-schools/>에서 2023.11.29. 인출.
- 민다경, 송인한, & 이한나. (2014). 인터넷 게임 과몰입이 청소년의 자살생각에 미치는 영향: 우울감의 매개효과 분석. *Journal of Digital Convergence*, 12(9), 445-454.
- 박민정. (2015). 노인의 성별에 따른 자살생각과 영향요인-2010년도 한국의료패널자료를 이용하여. *Journal of the Korean Data Analysis Society*, 17(2), 1087-1099.
- 박민정. (2021). 건강행위와 건강상태가 베이비붐세대 여성의 자살생각에 미치는 영향: 2017년 한국의료패널 자료 이용. *한국융합학회논문지*, 12(2), 499-507.
- 박성준(2019). 청소년 자살물에 영향을 미치는 지역적 요인. *정신건강과 사회복지*, 47(1), 174-203.
- 박재우, & 황병덕. (2015). 신체활동 특성이 노인 만성질환자의 정신건강에 미치는 영향. *대한보건연구*, 41(1), 25-38.
- 박초희. (2022). 알레르기 질환 청소년의 정신건강: 2020 청소년 건강 행태 온라인 조사 활용. *산업융합연구(구 대한산업경영학회지)*, 20(1), 87-96.
- 방소연, & 임은정. (2020). 노인의 신체기능 상태에 따른 정신건강에 미치는 영향에서 성별 차이. *상담심리교육복지*, 7(4), 113-126.
- 방소연. (2021). 한국 성인의 생애주기별 자살생각 영향요인. *한국산학기술학회논문지*, 22(5), 60-70.
- 보건복지부 국립정신건강센터 정신건강연구소. *정신건강실태조사 누리집*. <https://mhs.ncmh.go.kr/front/svyAdult.do>에서 2023.6.26. 인출.
- 보건복지부 국립정신건강센터. (2021). 2021년 정신건강실태조사 보고서.
- 보건복지부 보도자료. (2021.1.14.). 대한민국 정책브리핑(온국민 마음건강 종



- 합대책 브리핑). <https://www.korea.kr/briefing/policyBriefingView.do?newsId=156431976>에서 2023.12.27. 인출
- 보건복지부(2021). 제2차 정신건강복지 기본계획.
- 보건복지부, & 한국생명존중희망재단. (2021), 2021 자살예방백서, 한국생명존중희망재단.
- 보건복지부. (2022. 8. 10). 2022년 2분기 「코로나19 국민 정신건강 실태조사」 결과 발표 보도참고자료.
- 성혜연, 이성규, & 나지훈. (2021). 노인의 자살생각에 영향을 미치는 요인에 관한 연구: 객관적 및 주관적 건강상태에 따른 집단별 비교를 중심으로. 한국사회복지조사연구, 69, 117-143.
- 손신영. (2014). 우리나라 노인의 자살생각에 영향을 미치는 요인에 관한 연구. 보건의료산업학회지, 8(2), 149-160.
- 송영달, 손지아, & 박순미. (2010). 독거노인의 자살생각에 영향을 미치는 생태체계적 요인분석. 한국노년학, 30(2), 643-660.
- 안준서, 김은영, 조명재, 홍진표, 함봉진, 정인원, ..., & 이동우. (2016). 불안 장애와 주요우울장애에서 나타나는 신체 증상과 증상군에 따른 자살 사고, 계획, 행동과의 관계 고찰. 정신신체의학, 24(2), 174-183.
- 오미애, 전진아, 김은하, & 진재현. (2020). 빅데이터를 활용한 자살 고위험군 특성 분석. 보건복지부·한국보건사회연구원.
- 유화정, 한유화, & 유성은. (2023). 자살 위험 평가도구들 간의 검사 연계화: 고위험군 선별 기준 점수 비교. stress, 31(1), 25-36.
- 윤경서, 김은경, 박용천, 김대호, 최준호, 박선철, ..., & 이유상. (2023). 한글판 단축형 자살 사고 척도의 타당화. 신경정신의학, 62(3), 111-117.
- 이종익. (2018). 성인의 체질량지수 (BMI) 와 자살생각의 관계: 2009~2013년 한국의료패널자료를 활용한 연구. 한국콘텐츠학회논문지, 18(2), 616-625.
- 이충경, 박강규, 조맹제, 정은기, 김장규, 김선욱, ..., & 배안. (2001). 정신질환 실태 역학조사 보고서. 보건복지부·국립서울정신병원.
- 임은정, & 정계현. (2022). 만 19-39세 성인의 성별에 따른 자살생각 영향 요

- 인. 상담심리교육복지, 9(6), 247-260.
- 와이오밍 보건부(Department of Health). (2023). <https://health.wyo.gov/publichealth/prevention/wivpp/suicide-prevention/>에서 2023.11.29. 인출.
- 전진아, 김동진, 고든솔, 하솔잎, 이수빈, 현유림, ..., & 김성철. (2022). 제5차 자살예방기본계획 ('23-'27) 수립 연구. 보건복지부·한국보건사회연구원.
- 정송(2019). 서울 학생 자살 고위험군 지원 방안 연구. 서울특별시교육청교육연구정보원.
- 정은희, 신재동, 한겨레, 우선희, 하은솔, 김지원, ..., & 임은정. (2021). 2021년 한국복지패널 조사·분석 보고서. 한국보건사회연구원.
- 정진욱, 채희란, 천재영, 윤시몬, 이홍석, & 김울리. (2012). 정신건강 고위험자 관리체계 정립방안에 관한 연구. 한국보건사회연구원.
- 조맹제, 박종익, 배안, 배재남, 손정우, 안준호, ..., & 전승범. (2011). 2011년도 정신질환실태 역학조사. 보건복지부·서울대학교 의과대학.
- 조맹제, 함봉잔, 홍진표, 이동우, 이준영, 정인원, ..., & 이해우. (2006). 정신질환실태 역학조사 2006. 보건복지부·서울대학교 의과대학.
- 질병관리청. (2023). 국민건강영양조사 원시자료 이용지침서 제8기(2019-2021).
- 질병관리청. (2023). 국민건강영양조사 홍보홈페이지. 국민건강영양조사 개요. [https://knhanes.kdca.go.kr/knhanes/sub01/sub01\\_02.do](https://knhanes.kdca.go.kr/knhanes/sub01/sub01_02.do)에서 2023.2.19. 인출.
- 질병관리청. 국민건강영양조사. 각 연도별 원시자료.
- 질병관리청. 국민건강영양조사. 각 연도별 원시자료 및 조사표.
- 질병관리청. 청소년건강행태조사 누리집 사업소개 페이지. <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.5. 인출.
- 질병관리청. 청소년건강행태조사 누리집 조사 내용 페이지. <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.5. 인출.
- 질병관리청. 청소년건강행태조사 누리집 통계집 페이지. <https://www.kdca.go.kr/yhs/home.jsp>에서 2023.7.10. 인출.

- 질병관리청. 청소년건강행태조사. 각 연도별 조사표 및 통계.
- 최세문, & 이해영. (2017). 자녀의 아토피 피부염 유병이 주양육자의 정신건강에 미치는 영향. *대한보건연구*, 43(2), 37-42.
- 최정수, 남정자, 김정희, 고민정, 조재국, 이연희, ..., & 서희재. (2002). 2001년도 국민건강·영양조사: 총괄편. 보건복지부·한국보건사회연구원.
- 통계청. 통계설명DB. 청소년건강행태조사 통계설명자료. <https://www.narastat.kr/metasvc/index.do?confmNo=117058>에서 2023.7.5. 인출.
- 한국보건사회연구원. 한국복지패널 홈페이지 데이터 페이지. 한국복지패널 결합 데이터(1~17차). <https://www.koweps.re.kr:442/data/data/list.do>에서 2023.7.18. 인출.
- 한국보건사회연구원. 한국복지패널 홈페이지 설문지 페이지. <https://www.koweps.re.kr:442/data/survey/list.do>에서 2023.7.17. 인출.
- 한국보건사회연구원. 한국복지패널 홈페이지 코딩북 페이지. 한국복지패널 결합 데이터(1~17차) 코드북. <https://www.koweps.re.kr:442/data/book/list.do>에서 2023.7.17. 인출.
- 한국보건사회연구원. 한국복지패널 홈페이지. <https://www.koweps.re.kr:442>에서 2023.7.10. 인출.
- 한국보건사회연구원, & 국민건강보험공단. 한국의료패널 홈페이지. 설문지 페이지. <https://www.khp.re.kr:444/web/data/board/list.do?bbsid=63>에서 2023.7.17. 인출.
- 한국보건사회연구원, & 국민건강보험공단. 한국의료패널 홈페이지. 조사개요 페이지. <https://www.khp.re.kr:444/>에서 2023.7.31. 인출.
- 한국보건사회연구원, & 국민건강보험공단. 한국의료패널조사. 각 연도별 원시자료 및 조사표.
- 한국보건사회연구원, & 서울대학교 사회복지연구소. (2022). 한국복지패널 17차년도 조사자료 유저가이드. 한국보건사회연구원.
- 한국조사연구학회. (2020). 국가건강조사 2020년 표본설계. <제1세부 과제> 청소년건강행태조사 2020년 표본설계. 질병관리본부·한국조사연구학회.

- 허깅페이스 홈페이지. (2023). <https://huggingface.co/sentence-transformers>에서 2023.12.27. 인출.
- 한국보건사회연구원, & 서울대학교 사회복지연구소. 한국복지패널조사. 각 연도별 조사표.
- 홍진표, 이동우, 함봉진, 이소희, 성수정, 윤탁, ..., & 김선웅. (2016). 2016년도 정신질환실태 조사. 보건복지부·삼성서울병원.
- Beck, A. T., & Steer, R. A. (1993). Beck Anxiety Inventory Manual. The Psychological Corporation. San Antonio, TX.
- Bunton, R., Macdonald, G., & Macdonald, G. (2003). Health promotion: disciplines and diversity. Routledge.
- CDC. (2018). Centers for Disease Control and Prevention 홈페이지 CDC's VitalSigns pdf 파일. <https://www.cdc.gov/suicide/factors>에서 2023.12.27. 인출.
- Colizzi, M., Lasalvia, A., & Ruggeri, M. (2020). Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care?. *International journal of mental health systems*, 14(1), 1-14.
- Compton, M. T., & Shim, R. S. (2015). The social determinants of mental health. *Focus*, 13(4), 419-425.
- Cramer, R. J., & Kapusta, N. D. (2017). A social-ecological framework of theory, assessment, and prevention of suicide. *Frontiers in psychology*, 1756.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Hinton, G. E., & Roweis, S. (2002). Stochastic neighbor embedding. *Advances in neural information processing systems*, 15.
- Jakobsen, S. G., Nielsen, T., Larsen, C. P., Andersen, P. T., Lauritsen,

- J., Stenager, E., & Christiansen, E. (2023). Definitions and incidence rates of self-harm and suicide attempts in Europe: A scoping review. *Journal of Psychiatric Research*.
- Kiely, K. M., Brady, B., & Byles, J. (2019). Gender, mental health and ageing. *Maturitas*, 129, 76-84.
- Klonsky, E. D., May, A. M., & Saffer, B. Y. (2016). Suicide, suicide attempts, and suicidal ideation. *Annual review of clinical psychology*, 12, 307-330.
- Lee, G., Kim, M., Park, J. H., Hwang, S. W., & Cheon, J. H. (2022). Privacy-Preserving Text Classification on BERT Embeddings with Homomorphic Encryption. arXiv preprint arXiv: 2210.02574.
- Lee, S., Lee, G., Kim, J. W., Shin, J., & Lee, M. K. (2023, July). HETAL: Efficient privacy-preserving transfer learning with homomorphic encryption. In *International Conference on Machine Learning* (pp. 19010-19035). PMLR.
- Nils Reimers. (2023). SBERT.net(SBERT의 파이썬 라이브러리). <https://www.sbert.net/examples/applications/cross-encoder/README.html>에서 2023. 12. 21. 인출.
- Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment*, 8(4), 443-454.
- Posner, K., Brent, D., Lucas, C., Gould, M., Stanley, B., Brown, G., Fisher, P., Zelazny, J., Burke, A., Oquendo, M., & Mann, J. (2008). Columbia-suicide severity rating scale (C-SSRS). New York, NY: Columbia University Medical Center, 10, 2008.
- Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint

arXiv:1908.10084.

- Sands, N. Elsom, E, Colgate, R & Haylor, H. (2016). Development and inter-rater reliability of the UK Mental Health Triage Scale (In Press). *International Journal of Mental Health Nursing*.
- Saxena, S., Izutsu, T., & Tsutsumi, A. (2008). *정신건강증진: 개념·증거·실천*(김민석 외 역). 서울: 포널스출판사.
- Suh, S., Ryu, H., Chu, C., Hom, M., Rogers, M. L., & Joiner, T. (2017). Validation of the Korean depressive symptom inventory-suicidality subscale. *Psychiatry investigation*, 14(5), 577.
- Todd, C. (2022, August 18). A month into 988, call volumes up, but state behavioral health infrastructure isn't. *Nevada Current*. <https://www.nevadacurrent.com/2022/08/18/a-month-into-988-call-volumes-up-but-state-behavioral-health-infrastructure-is> nt/에서 2023.11.29. 인출.
- Vaillant GE. (2003). *Mental Health*. *Am J Psychiatry* 160, 1373-1384.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ..., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Walker, L., Moodie, R., & Herrman, H. (2004). *Promoting mental health and wellbeing*. *Hands on Health Promotion*. Melbourne (AUST): IP Communications.
- WHO. (2020). *Basic documents*, 49rd ed. Geneva, World Health Organization.
- Yuryev, A., Leppik, L., Tooding, L. M., Sisask, M., Varnik, P., Wu, J., & Varnik, A. (2010). Social inclusion affects elderly suicide mortality. *International Psychogeriatrics*, 22(8), 1337-1343.